

線形代数学 講義ノート (2020/05/03ver.)

まえがき

これは大学1年生を対象にした線形代数学の講義ノートである。本書では、大学に入学して線形代数学を初めて学ぶ学生のうち、この分野で扱う数学的諸概念の意図が分からず、闇雲に計算することに悩んでいる人を主な読者として想定し、次の点を重んじた:

- 新しい概念を導入する理由や意図をできる限り説明した。
- 天下り的な定義の導入はなるべく控えた。
- 計算や論理の飛躍を避け、詳細まで丁寧に書いた。
- 論理的に正しければそれでよいという立場をとらずに、定理の主張や証明をどのように理解すべきかも論じた。
- 線形代数学の枠組みの外にある数学にも言及した。

この本の証明部分とそれ以外の説明部分とを繋げて理解するように心がければ、読者はあまり労力を要さずに数学特有の考え方に慣れることができる。それにより、定義と定理の証明しか記述されていないような数学書においても、著者の意図を想像できるようになるだろう。また、難攻不落だが名著とされる、志の高い線形代数の本に挑戦するための基礎力(行間を埋める能力)も身につけているはずである。

本書の構成について説明しておく。第I部と第II部では連立1次方程式の解法と行列式の計算を主に扱う。第III部から第V部までは、線形空間の抽象論の初歩を踏まえた上で、行列の対角化を理解することを目標とした。なお、第V部の後半では、対角化できない行列に関する発展的な話題も取り上げている。第VI部には内積をとともなうベクトル空間の話題を集めた。多くの大学の1年次向けカリキュラムでは、直交行列による対称行列の対角化がここでの目標となるであろう。

【本書の読み方】

各章はおおよそ講義1回分の内容に相当している。各章の冒頭には、これから何を論じ、そのために何を導入するかを概説し、読者がおおまかな議論の流れを把握できるよう配慮した。各節の題目や命題に「発展」と記してあるものはやや高度な内容のものを指し、大学のカリキュラムによっては2年次相当の部分を含んでいる。難しいと感じるようであれば、ここは目を通すだけでも構わない。「よりみち」と題した部分では、線形代数学の内容からは少々離れるものの、数学への知的好奇心が高まるような話題を取り上げた。これらを通して、数学に秘められた思想の一端に読者が触れることを望んでいる。

嶺 幸太郎

目次

第 1 章	線形代数学とは何か	8
1.1	ユークリッド空間における和とスカラー倍	8
1.2	写像とその合成	9
1.3	ユークリッド空間における線形写像	11
第 I 部	行列と連立 1 次方程式	14
第 2 章	行列の演算	15
2.1	行列の成分表示	15
2.2	行列の和とスカラー倍	16
2.3	行列の積	16
2.4	行列演算の性質	18
第 3 章	行列の表し方	21
3.1	クロネッカーのデルタ	21
3.2	\sum 記号の使い方	21
3.3	成分の空白と任意性	24
3.4	転置行列	24
3.5	行列の分割	26
第 4 章	連立 1 次方程式	28
4.1	導入	28
4.2	連立 1 次方程式と行列	28
4.3	逆行列を持つ場合	29
4.4	行列の行基本変形	30
4.5	簡約行列	31
4.6	連立 1 次方程式の解法	33
4.7	連立 1 次方程式の解の形と任意定数の個数について	35
第 5 章	可逆行列	37
5.1	逆行列の性質	37
5.2	行基本変形再考	38
5.3	逆行列の求め方	40
5.4	基本行列と列基本変形 (発展)	42
第 6 章	行列の階数	43
6.1	簡約化の一意性	43
6.2	連立 1 次方程式と階数	45
6.3	同次形の方程式	46
6.4	重ね合わせの原理	47

第 II 部	行列式	50
第 7 章	行列式とは何か	51
7.1	\mathbb{R}^n 上の線形変換の体積拡大率	51
7.2	クラメルの公式	53
7.3	微積分学における行列式	54
第 8 章	置換	56
8.1	置換の定義	56
8.2	置換の表示	56
8.3	置換の積	57
8.4	巡回置換とその表示	58
8.5	置換の符号	60
第 9 章	置換の符号について	64
9.1	符号の正当性	64
9.2	文字列の並び替え (よりみち)	66
9.3	転倒数による符号の定義 (よりみち)	68
第 10 章	行列式の定義と性質	70
10.1	定義	70
10.2	カヴァリエリの原理	71
10.3	多重線形性	72
10.4	歪対称性	74
第 11 章	行列式の計算	76
11.1	サイズの小さい行列式との関係	76
11.2	計算例	77
11.3	体積との関係	79
第 12 章	行列式の性質 (証明)	81
12.1	$ A = {}^tA $ の証明	81
12.2	歪対称性の証明	82
12.3	多重線形性と歪対称性から導かれる性質	83
12.4	行列式の特徴づけ (証明)	84
12.5	命題 11.1.1 の証明	86
12.6	$ AB = A \cdot B $ の証明	87
第 13 章	余因子展開とクラメルの公式	88
13.1	余因子展開	88
13.2	余因子行列	90
13.3	クラメルの公式の証明	93
第 III 部	抽象ベクトル空間	95
第 14 章	集合概念の基礎	96
14.1	集合の包含関係	96
14.2	集合の表し方	97
14.3	外延的か内包的か	99

14.4	和集合と共通部分	100
14.5	集合論と逆理 (よりみち)	101
第 15 章	線形空間	104
15.1	ベクトル空間の公理	104
15.2	線形空間の例	105
15.3	体 K 上の線形空間 (発展)	108
第 16 章	いろいろな線形部分空間	111
16.1	定義	111
16.2	\mathbb{R}^n の部分空間	113
16.3	部分空間の様々な例	114
第 17 章	線形結合と線形独立性	117
17.1	線形結合	117
17.2	線形独立性	118
17.3	線形独立性の判定	120
第 18 章	基底	123
18.1	ベクトルの組が生成する部分空間	123
18.2	基底の定義と例	124
18.3	基底の探し方	126
18.4	一般の基底 (発展)	129
第 IV 部	線形写像	132
第 19 章	写像概念の基礎	133
19.1	像と逆像	133
19.2	全射と単射	134
19.3	逆写像とその性質	136
19.4	写像の合成	138
19.5	無限集合 (発展)	139
第 20 章	線形写像	142
20.1	線形写像の基本的性質	142
20.2	線形写像による像	145
20.3	線形写像による逆像	146
20.4	様々な線形写像の例	148
第 21 章	線形空間の同一視	152
21.1	線形同型写像	152
21.2	同型な線形空間	153
21.3	線形写像のなす空間	155
21.4	多元環とその準同型	156
21.5	線形変換と多項式	157
第 22 章	線形空間の次元	160
22.1	次元の定義	160
22.2	連立 1 次方程式の任意定数の個数	162

22.3	線形独立な最大個数	162
22.4	次元から分かること	164
22.5	無限次元の空間も含めた一般論 (発展)	168
第 23 章	次元公式と商空間	170
23.1	空間の平行移動	170
23.2	線形写像の次元公式	172
23.3	商空間 (発展)	173
23.4	商空間の例 (発展)	175
23.5	同値関係と商集合 (発展)	176
第 24 章	第 1 同型定理と短完全列 (発展)	179
24.1	商空間の線形構造	179
24.2	第 1 同型定理	181
24.3	完全系列と短完全列	182
第 25 章	線形結合の行列表示	185
25.1	線形結合の組と行列	185
25.2	ベクトルの組と行列の演算の基本性質 (付録)	186
25.3	線形結合再考	187
25.4	線形独立性の判定 (2)	188
25.5	基底の変換行列	189
第 26 章	線形写像の表現行列	191
26.1	表現行列	191
26.2	$\text{Hom}(U, V)$ と $M_{m,n}(\mathbb{R})$	194
26.3	基底の取りかたによる表現行列の違い	195
26.4	1 対 1 の対応と可換図式	197
第 27 章	双対空間 (発展)	200
27.1	定義	200
27.2	双対基底	201
27.3	双対写像	202
27.4	双対写像の基本的性質	203
27.5	二重双対	204
27.6	内積との類似性	205
第 V 部	固有空間分解	208
第 28 章	固有値と固有ベクトル	209
28.1	関数の分解	209
28.2	固有ベクトル	210
28.3	固有ベクトルからなる基底と行列の対角化	211
28.4	特性多項式	212
28.5	固有空間	214
28.6	一般の線形写像の固有空間	216

第 29 章 固有空間分解と行列の対角化	218
29.1 固有ベクトルからなる組の線形独立性	218
29.2 対角化可能条件	220
29.3 一般の線形変換の場合	226
29.4 多項式と方程式の解に関する基本的性質 (付録)	227
第 30 章 ケーリー-ハミルトンの定理	231
30.1 まえおき	231
30.2 (i) と (ii) の同値性	232
30.3 対角化可能な行列の場合	232
30.4 一般の場合	234
第 31 章 斉次形線形漸化式	237
31.1 線形漸化式と固有値	237
31.2 複素数列のなかの実数列	239
31.3 高次の線形漸化式と表現行列 (発展)	239
第 32 章 斉次形線形常微分方程式	242
32.1 線形常微分方程式と固有値	242
32.2 特性多項式が複素解をもつ場合における実数解	244
32.3 高次の線形常微分方程式 (発展)	245
第 33 章 不変部分空間と冪零部分空間 (発展)	248
33.1 不変部分空間	248
33.2 冪零部分空間	249
33.3 微分作用素とシフト作用素の一般固有ベクトル	252
33.4 冪零部分空間と安定部分空間への分解	254
33.5 直和分解 (発展)	256
第 34 章 一般固有空間分解とその応用 (発展)	258
34.1 一般固有空間分解	258
34.2 ケーリー-ハミルトンの定理 (再論)	260
34.3 線形漸化式と線形常微分方程式 (再論)	260
34.4 定理 34.1.1 の証明	261
第 VI 部 計量ベクトル空間	264
第 35 章 内積と正規直交基底	265
35.1 標準的な内積とユークリッド・ノルム	265
35.2 余弦定理と内積	266
35.3 正規直交基底	268
35.4 シュミットの直交化法	269
35.5 直交補空間	271
35.6 内積空間 (発展)	272
第 36 章 直交行列と上三角化	275
36.1 直交行列とその性質	275
36.2 直交行列による上三角化	277
36.3 直交行列による対称行列の対角化	279

第 37 章 2 次形式の分類 (発展)	284
37.1 対称行列が定める 2 次形式	284
37.2 2 次形式の分類	285
37.3 2 次形式と固有値	287
37.4 正定値性の判別法	288
37.5 2 次曲線 (よりみち)	291
37.6 多変数関数の極値判定 (よりみち)	292
第 38 章 内積空間上の線形変換	294
38.1 ピタゴラスの定理とコーシー-シュワルツの不等式	294
38.2 直交変換	296
38.3 直交射影とその鏡映	297
38.4 対称行列の性質	299
38.5 射影変換 (発展)	300
第 39 章 エルミート内積	303
39.1 エルミート内積とノルム	303
39.2 複素内積空間における直交性	304
39.3 エルミート内積から定まるノルムの性質	306
39.4 随伴行列と随伴変換	306
39.5 ユニタリ行列	309
39.6 エルミート行列の性質	310
39.7 正規行列	311
39.8 フーリエ級数展開 (発展)	313

第1章 線形代数学とは何か

線形代数学とは、線形写像（線形性）を分析するための理論の総体である。しかしこの説明は同語反復と変わらず、まったく説明になっていない。数学用語としての線形写像を定義するには、集合、写像、ベクトル、スカラー倍、ユークリッド空間といった様々な技術用語をも説明せねばならない。このような細かい話は本章の本文にゆずることにして、まずはこの理論がもたらす影響のあらましについて述べよう。

線形写像とは、大雑把に言えば比例概念の多変数化に相当する。したがって一つの見方として、線形写像は定数関数の次に単純な関数といえる。単純なもの、つまり取るに足らぬものを分析することに何の意義があるのかと疑問に思う初学者もいるかもしれない。その疑問への答えは、複雑なものをそのまま分析できる才があるならば、そうすればいいし、それが難しそうであれば、何らかの方法で己の身の丈に合う段階まで単純化して分析するしかない、ということである。

自然科学が、自然界の現象を出来る限り単純なモデルに落として分析する学問であるとするならば、定数関数の次に単純な構造を持つ線形写像がそこで大きな役割を果たすであろうことは想像に難くない。そして実際、我々人間が感じる数理科学的知覚の多くに線形性が内在している。そこで、今後どのような専門分野に進むにせよ、将来おそらく避けては通れないであろう何らかの単純な関数（線形写像）の扱いを想定し、あらかじめ一般論として線形写像の分析法をまとめて理解しておく。これが理工系の学問を志す大多数の者が線形代数学を学ぶゆえんである。こんにちでは、理工系分野で求められる基本的な素養の一つとして線形代数学は認められている。

もちろん、上の文脈とは無縁ではあるものの知的興味本位から線形代数を学んでみたい、という人もいるだろう。本書では、そのような方も読者の中にと想定して、技術論よりも考え方に焦点を当てた解説を試みている。

それでは、「線形写像」の定義と、その分析において鍵となる概念である「行列」の説明を試みるために、やや退屈ではあるが、集合や写像、そしてベクトルの和とスカラー倍についての復習から始める。

1.1 ユークリッド空間における和とスカラー倍

いくつかのものの集まりのことを集合 (set) という。集合を構成しているもの一つ一つを要素 (element) または元という。集合の表記の仕方の一つとして、集合を構成する要素をすべて並べて中括弧でくくる方法がある。例えば、りんご、みかん、スイカの3つの要素からなる集合は、

$$\{ \text{りんご, みかん, スイカ} \}$$

と表される。こうした表し方は、本書では多くは使わないものの、稀に用いるゆえ忘れないでほしい。

数を構成要素とする集合のうちいくつかは慣例で特別なアルファベットが割り当てられており、次のような記号を用いる¹：

- \mathbb{N} : 自然数²(natural number) 全体のなす集合のこと。
- \mathbb{Z} : 整数 (integer, integral number) 全体のなす集合のこと³。
- \mathbb{Q} : 有理数 (rational number) 全体のなす集合のこと。

¹太字 $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}, \mathbf{C}$ を用いる場合もある。

²集合論を学ぶと、自然数に0を含めたほうが多くの表記において整合性が取れることが分かる。しかし、ここでは高校までの慣例に従い、自然数に0は含まれないものとする。

³記号 \mathbb{Z} はドイツ語で数を意味する Zahlen に由来する。

- \mathbb{R} : 実数 (real number) 全体のなす集合のこと.
- \mathbb{C} : 複素数 (complex number) 全体のなす集合のこと.

\mathbb{R} や \mathbb{C} においては, 加減乗除の四則演算が定まっている. これは既知のこととして, 話を進めよう⁴. ある x が集合 X の元であるとき $x \in X$ と書く. そうでないとき, $x \notin X$ と書く.

例 1.1.1. $5 \in \mathbb{N}$, $-1 \notin \mathbb{N}$, $\frac{12}{13} \in \mathbb{Q}$, $\sqrt{2} \notin \mathbb{Q}$, $\text{ぶどう} \notin \{\text{りんご}, \text{みかん}, \text{スイカ}\}$.

二つの実数 $x, y \in \mathbb{R}$ による並び順を込めた意味での組 (x, y) たち全体からなる集合を \mathbb{R}^2 と表す. ここでいう“並び順を込めた意味”とは, (x, y) と (y, x) は違うものと見なすということである⁵. 同様に, 各 $n \in \mathbb{N}$ について, n 個の実数の並び順を込めた意味での組 (x_1, x_2, \dots, x_n) たち全体からなる集合を \mathbb{R}^n と表し, これを n 次元ユークリッド空間 (n -dimensional Euclidean space) という. また, \mathbb{R}^n の元のことをベクトル (vector) と呼ぶ. \mathbb{R}^2 のベクトル (x, y) は, 直交する 2 本の座標軸が与えられた平面 (これを座標平面という) 上の点の位置 (座標) とも見なされる. すなわち, \mathbb{R}^2 の各ベクトルは座標平面上の各点と対応づけられる. 同様に \mathbb{R}^3 の各ベクトルは 3 次元の座標空間上の各点と対応づけられる.

高校で学習したように, \mathbb{R}^2 や \mathbb{R}^3 のベクトルの間には和とスカラー倍が定まっているのであった. 同様のことが \mathbb{R}^n においても定義される. すなわち, \mathbb{R}^n の二つのベクトル $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$ および実数 $r \in \mathbb{R}$ に対して, ベクトルの和 $(x_1, \dots, x_n) + (y_1, \dots, y_n)$ およびベクトルのスカラー倍 (scalar multiplication) $r(x_1, \dots, x_n)$ を次で定める:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n), \quad r(x_1, \dots, x_n) := (rx_1, \dots, rx_n).$$

【補足】 上で用いた記号 $:=$ は, 新たな概念である左辺を右辺で定めていることを表す.

ベクトルを一文字で表す場合は, $\mathbf{x} = (x_1, \dots, x_n)$ のように太字で書くことがある. 太字を用いるのは 1 変数と混同しないようにするための措置であり, これは 1 年次向け教育における慣例となっている. $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (ただし $n = 2, 3$) に対して, $\mathbf{0}, \mathbf{x}, \mathbf{x} + \mathbf{y}, \mathbf{y}$ の 4 点を頂点とする四角形が平行四辺形になることは図示によって確認できる. もちろん $n \geq 4$ の場合についても同様のことが成り立つ. 我々は 3 次元の空間に住んでいるため高次元の世界を直接に見ることはできないけれども, このように高次元の空間がもつ性質のいくつかを想像することができる.

微積分ではベクトルを横書きにするのが慣例である. 一方で, 2 章で定める行列の積との関係から線形代数ではベクトルを縦書きにしたほうが都合が良いことが多い. 横に並べたベクトルを行ベクトル (row vector) といい縦に並べたベクトルを列ベクトル (column vector) という. 以降ではどちらも \mathbb{R}^n の元とみなし, 用途に応じて使い分けることがある.

$$\text{行ベクトル} : (x_1, \dots, x_n), \quad \text{列ベクトル} : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

1.2 写像とその合成

高校までの数学で現れる写像 (関数) はほとんどが 1 変数であり, 多変数の場合についても実数値関数のみを扱った. それゆえ写像についてことさら細かい概念は必要なかったのであるが, これからは \mathbb{R}^n のベクトルを代入すると \mathbb{R}^m のベクトルが与えられるような写像を考えるため, 何を代入すると何が得られるのか明確にする必要が生じる. そこで, 写像について改めて定義を述べておこう.

⁴このようなことを書く理由には, 数とは何かという哲学的問いを深め, より厳密な立場から数を定義するという思想があるからである.

⁵より正確には, $x = y$ であったときを除いて (x, y) と (y, x) を異なるものと見なすということ.

定義 1.2.1. 集合 X の各元 x に対して集合 Y の元を一つ与える操作 (対応) を考える. このような操作を X から Y への写像 (**map, mapping**) と呼び, X をこの写像の定義域 (**domain**) という. X から Y への写像を記号 f を用いて表す場合, $x \in X$ に対応する Y の元を $f(x)$ と書く. この表記を通して, “ x を f で写像する” あるいは, “ f に x を代入する” といった表現もなされる. また, どの集合の元に対してどの集合の元を対応させる操作なのか明示するために, 次のような記号・図式が用いられる:

$$f: X \rightarrow Y, \quad f: X \ni x \mapsto f(x) \in Y, \quad \begin{array}{ccc} f: X & \longrightarrow & Y \\ \cup & & \cup \\ x & \longmapsto & f(x). \end{array}$$

なお, 写像と関数 (**function**) はほぼ同義語である. 対応される元が数となるような写像 (すなわち上の定義において集合 Y が数を要素とする集合である場合) のことを関数と呼ぶことが多い.

定義域の元がベクトル \mathbf{x} の場合, $f(\mathbf{x})$ を成分表示して正確に書けば $f((x_1, \dots, x_n))$ となり二重括弧が煩わしい. そこで, 括弧を一つ減らして $f(x_1, \dots, x_n)$ と書くのが慣例となっている.

例 1.2.2. 各実数 x に対して, $f(x) := 3x$ と定めれば, これは3倍の実数を対応させる写像 $f: \mathbb{R} \rightarrow \mathbb{R}$ である.

いくつかの写像が与えられているとき, それらを用いて新たな写像を構成する操作を数学では頻繁に行う. こうした操作の中で最も基本的なものが写像の合成である.

定義 1.2.3. 二つの写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$ が与えられているとする. このとき, 各 $x \in X$ に対して, Z の元 $g(f(x))$ を対応させる写像を f と g の合成 (**composition**) と呼び, 記号 $g \circ f: X \rightarrow Z$ で表す. すなわち, $(g \circ f)(x) := g(f(x))$ である. 誤解がなければ括弧を略して $(g \circ f)(x)$ を $g \circ f(x)$ と書く.

【注意】 合成 $g \circ f$ を定めるには, 各 $f(x)$ が g の定義域の要素でなければならない.

合成関数の簡単な例を, 比例関係にある関数を通して見てみよう.

定義 1.2.4. ある実数 a を用いて $f(x) := ax$ と定められる関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ を比例関数という⁶.

例 1.2.5. 関数 $g: \mathbb{R} \rightarrow \mathbb{R}$ および $f: \mathbb{R} \rightarrow \mathbb{R}$ を $g(x) := ax$, $f(x) := bx$ と定めれば $g \circ f$ も比例関数であり, $g \circ f(x) = (ab)x$ である. 実際, $g \circ f(x) = g(f(x)) = g(bx) = a(bx) = (ab)x$.

さて, 比例関数の例をいくつか列挙しているうちに, 次の対応に気がつくのではなからうか.



すなわち, 実数全体と比例関数全体は1対1に対応しているのである. しかも, この対応は数の間の掛け算と関数の間の合成も上手く関係づけられている. つまり, 比例関数 g および f に対応する実数をそれぞれ a, b とすると, $g \circ f$ に対応する実数は ab である (例 1.2.5). あまりにも簡単なことを述べているため拍子抜けしてしまうかもしれない. しかしながら, このように異なる二つの概念にうまく対応を与えることで, それらが本質的に同じものであると見抜くことこそが代数学の神髄といっても過言ではない⁷. いまの話では実数をただの数と思うだけでなく, 比例関数とも思えるということであり, これは実数に対する見方を広げたことを意味している. こうした考え方の多次元版として行列の概念が現れる. これを次節で詳しく述べよう.

備考 1.2.6. 二つの比例関数 $f(x) = ax$ および $g(x) = bx$ に対して, これらの和によって表される関数 $h(x) = f(x) + g(x)$ もまた比例関数である. 実際, $h(x) = (a + b)x$ と書ける. つまり, 実数の和と比例関数の和についても上手く対応づけられていることが分かる.

⁶このような関数への一般的な呼び名は与えられておらず, 「比例関数」は本書でのみ通じる用語である.

⁷代数学を「加減乗除による四則演算の技法を高めていく学問」と説明することがよくあるが, これは説明を放棄したいときの逃げ口上であって, 代数学の本質をつくものではない.

1.3 ユークリッド空間における線形写像

比例関数の多変数版に相当する線形写像は次のように定義される:

定義 1.3.1. 次の性質 (i) および (ii) を持つ写像 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ を線形写像 (linear map) という:

- (i) すべての $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ に対して, $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$,
- (ii) すべての $\mathbf{x} \in \mathbb{R}^n$ および各 $r \in \mathbb{R}$ について, $f(r\mathbf{x}) = rf(\mathbf{x})$.

すなわち, ベクトルの和を取ってから代入しても代入してから和を取っても同じ結果が得られ, また, スカラー倍をほどこしてから代入したものは代入してからスカラー倍をほどこしたものに一致するような写像のことである. 上の二つの性質を線形性 (linearity) という. 比例関数が線形写像であることを確認してみよう.

命題 1.3.2. 比例関数 $f(x) = ax$ は線形写像である.

Proof. 線形写像の性質 (i) および (ii) が成立することを確認すればよい.

- 各 $x, y \in \mathbb{R}$ に対して, $f(x + y) = a(x + y) = ax + ay = f(x) + f(y)$. ゆえに (i) は成立する.
- 各 $x \in \mathbb{R}$ および $r \in \mathbb{R}$ に対して, $f(rx) = a(rx) = r(ax) = rf(x)$. ゆえに (ii) は成立する.

以上より f は, 性質 (i) および (ii) を持つことが分かった. ゆえに f は線形写像である. □

上の証明の最後に用いた記号 \square は証明終 (q.e.d. ...quod erat demonstrandum) を意味する. 線形写像の重要な例は 20 章で与える. ここでは \mathbb{R}^2 の間の線形写像を挙げよう.

例題 1.3.3. あらかじめ実数 a, b, c, d を与えておき, 写像 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ を $f\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$ と定める. このとき次に答えよ.

- (1) $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ および $\mathbf{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ を f に代入した値 (これはベクトル値である) を求めよ.

解答例:

$$f(\mathbf{x}) = f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a1 + b0 \\ c1 + d0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}, \quad f(\mathbf{y}) = f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a0 + b1 \\ c0 + d1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}.$$

- (2) f が線形写像であることを示せ.

解答例: まず線形写像の性質 (i) を確認しよう. 各 $\mathbf{a} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{b} = \begin{pmatrix} x' \\ y' \end{pmatrix}$ に対して,

$$\begin{aligned} f(\mathbf{a} + \mathbf{b}) &= f\left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix}\right) = f\begin{pmatrix} x + x' \\ y + y' \end{pmatrix} = \begin{pmatrix} a(x + x') + b(y + y') \\ c(x + x') + d(y + y') \end{pmatrix} \\ &= \begin{pmatrix} (ax + by) + (ax' + by') \\ (cx + dy) + (cx' + dy') \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} + \begin{pmatrix} ax' + by' \\ cx' + dy' \end{pmatrix} = f(\mathbf{a}) + f(\mathbf{b}). \end{aligned}$$

ゆえに線形写像の性質 (i) は成立する.

つぎに線形写像の性質 (ii) を確認しよう. 各 $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ および $r \in \mathbb{R}$ に対して,

$$f(r\mathbf{x}) = f\left(r\begin{pmatrix} x \\ y \end{pmatrix}\right) = f\begin{pmatrix} rx \\ ry \end{pmatrix} = \begin{pmatrix} a(rx) + b(ry) \\ c(rx) + d(ry) \end{pmatrix} = \begin{pmatrix} r(ax + by) \\ r(cx + dy) \end{pmatrix} = r\begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = rf(\mathbf{x}).$$

ゆえに線形写像の性質 (ii) は成立する. □

実は、 \mathbb{R}^2 の間の線形写像は上の例題で与えた形のものですべて出つくしている。すなわち、 $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ を任意の線形写像として、 $\begin{pmatrix} a \\ c \end{pmatrix} := g \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ および $\begin{pmatrix} b \\ d \end{pmatrix} := g \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ とおくと $g \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$ が各 $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ について成立する。これは線形写像の性質 (i) および (ii) を用いて次のように確かめられる:

$$\begin{aligned} g \begin{pmatrix} x \\ y \end{pmatrix} &= g \left(x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = g \left(x \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) + g \left(y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \\ &= xg \begin{pmatrix} 1 \\ 0 \end{pmatrix} + yg \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x \begin{pmatrix} a \\ c \end{pmatrix} + y \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} ax \\ cx \end{pmatrix} + \begin{pmatrix} by \\ dy \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}. \end{aligned}$$

以上により、 \mathbb{R}^2 の間の任意の線形写像 $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ は、たった四つの数 a, b, c, d で特徴づけられることが分かった。

備考 1.3.4. 上の計算の通り、 \mathbb{R}^2 上の線形写像における各ベクトルの行き先は x 軸および y 軸と平行なベクトル $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ の行き先のみで決まる。これは線形写像の顕著な性質の一つであり、これによって物事をより単純なものに分解したり、分解したものを総合したりすることが可能になる。上の計算は、更に一般的な形で命題 20.1.11 にて論じられる。

さて、線形代数学では四つの数を次のような形にならべたものを用いる:

$$\begin{pmatrix} 1 & 0 \\ 2 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

これらを 2×2 行列と呼ぶことにしよう。すると上の議論は、線形写像と 2×2 行列が 1 対 1 に対応することを示唆している。すなわち、行列 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ に対して線形写像 $f \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$ を対応させれば、これは 1 対 1 対応である。ここで、前節で述べた比例関数と実数の間の 1 対 1 対応を思い起こせば、次の問題が想起される。

問題 1.3.5. \mathbb{R}^2 の間の線形写像における和と合成に上手く対応するように、 2×2 行列の間に和 (足し算) と積 (掛け算) を定めることができるか。

実は、次のようにして 2×2 行列に和と積を定めればよいことが分かっている。

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} F & G \\ H & I \end{pmatrix} := \begin{pmatrix} a+F & b+G \\ c+H & d+I \end{pmatrix}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} F & G \\ H & I \end{pmatrix} := \begin{pmatrix} aF+bH & aG+bI \\ cF+dH & cG+dI \end{pmatrix}.$$

行列の和の定義が自然なものと思える一方で積の定義がやや複雑になるのは、上の問題で挙げた要請に答えるためである。より詳しく述べれば、次の命題を成立させることが行列理論の前提になっている:

命題 1.3.6. \mathbb{R}^2 から \mathbb{R}^2 への線形写像全体と 2×2 行列全体の間で定まる先程の対応は 1 対 1 である。また、二つの線形写像 $g, f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ に対応する行列をそれぞれ A, B とすれば、これらの和 $g + f$ および合成 $g \circ f$ もまた線形写像であり、 $g + f$ および $g \circ f$ に対応する行列はそれぞれ $A + B, AB$ である。

この対応は 2 変数に限らず、 \mathbb{R}^n から \mathbb{R}^m への線形写像全体と $m \times n$ 行列全体の間に対応として一般に成り立つことを、ここで予告しておく。

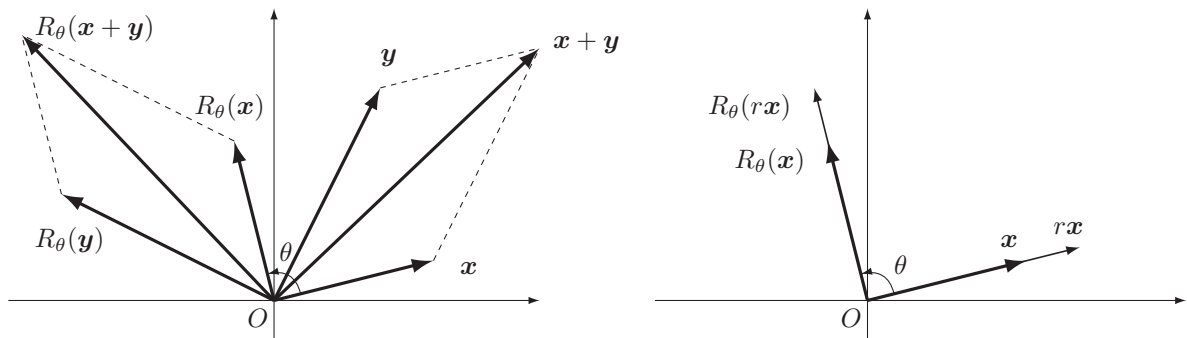
線形代数学で扱う行列は、線形写像を数値化 (データ化) するために考えだされた概念である。そして、抽象的な線形写像を行列によってデータ化することで、その写像に現れる現象を具体化・可視化して分析しやすくする、というのが行列理論のねらいである。本書の第 I 部および第 II 部では、行列に関する基本的な計算を理解することを目標とする。また第 III 部以降では、行列を駆使した線形写像の分析について論じていく。第 III 部に入るまでのしばらくのあいだ、線形写像そのものは表立って出てはこないものの、行列理論の目的が線形写像の分析にあるということ念頭において学習してもらいたい。

練習 1.3.7. 二つの線形写像 $\xi, \eta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ を $\xi \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$, $\eta \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} Fx + Gy \\ Hx + Iy \end{pmatrix}$ で定める. このとき $\xi \circ \eta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (aF + bH)x + (aG + bI)y \\ (cF + dH)x + (cG + dI)y \end{pmatrix}$ となることを確かめよ.

よりみち (加法定理)

\mathbb{R}^2 の各ベクトル \mathbf{x} に対して, 原点 O を中心に \mathbf{x} を θ 回転させたベクトルを対応させる写像 $R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ を考える. これは明らかに線形写像である. 実際, 定義 1.3.1 における線形性 (i) および (ii) が成り立つことは次のようにして理解できる.

- (i) O を頂点に持つ平行四辺形を O を中心に回転させれば, これも O を頂点に持つ平行四辺形である. ゆえに $R_\theta(\mathbf{x} + \mathbf{y})$ は二つのベクトル $R_\theta(\mathbf{x})$, $R_\theta(\mathbf{y})$ で張られる平行四辺形の頂点となる. これは $R_\theta(\mathbf{x} + \mathbf{y}) = R_\theta(\mathbf{x}) + R_\theta(\mathbf{y})$ を意味している.
- (ii) 回転によってベクトルの長さが変化することはない. したがって \mathbf{x} と $r\mathbf{x}$ における長さの比と, これらを R_θ で写像した $R_\theta(\mathbf{x})$ と $R_\theta(r\mathbf{x})$ における長さの比は共に $1 : r$ である. 更に, \mathbf{x} と $r\mathbf{x}$ は平行ゆえ θ 回転後の $R_\theta(\mathbf{x})$ と $R_\theta(r\mathbf{x})$ も平行である. 以上のことから $R_\theta(r\mathbf{x}) = rR_\theta(\mathbf{x})$ が成り立つ.



ベクトル $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ および $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ を θ 回転させたベクトルはそれぞれ $\begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$, $\begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$ であるから,

線形写像 R_θ に対応する行列は $A_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ となる.

さて, ベクトルを β 回転させた後に更に α 回転させる操作と, 一度に $\alpha + \beta$ 回転させる操作の結果は同じである. これは $R_\alpha \circ R_\beta = R_{\alpha+\beta}$ を意味する. $R_{\alpha+\beta}$ に対応する行列は $A_{\alpha+\beta} = \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}$ であり, $R_\alpha \circ R_\beta$ に対応する行列は $A_\alpha A_\beta$ ゆえ

$$\begin{aligned} A_\alpha A_\beta &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \\ &= \begin{pmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -\cos \alpha \sin \beta - \sin \alpha \cos \beta \\ \sin \alpha \cos \beta + \cos \alpha \sin \beta & -\sin \alpha \sin \beta + \cos \alpha \cos \beta \end{pmatrix}. \end{aligned}$$

$A_{\alpha+\beta} = A_\alpha A_\beta$ の成分を比較することで加法定理:

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta,$$

$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta,$$

を得る. 線形写像を知る者にとって加法定理は自明の理といえるだろう.

第I部

行列と連立1次方程式

第2章 行列の演算

この章では行列 (matrix) に関する三つの演算, すなわち和およびスカラー倍, 積を導入する. 1章で見たように, 線形写像に対して定義される演算を行列の言葉で読み替えたものになることを想定し, これらの演算の定義を与えている.

2.1 行列の成分表示

$m \times n$ 個の数を矩形に並べ括弧で囲んだものを m 行 n 列の行列あるいは (m, n) -行列, $m \times n$ 行列などという. 数学書は横書きで記述するゆえ横に並ぶ文字列が行 (row) であり, 縦に並ぶ文字列が列 (column) である. 行の数と列の数の組 (m, n) を行列の型 (type) あるいはサイズ (size) という. また, 行列を構成するために並べた数のことを, その行列の成分 (entry, element) と呼ぶ.

例 2.1.1. 4行5列の行列:

$$A = \begin{pmatrix} 1 & 0 & 4 & 2 & 1 \\ 5 & 7 & 1 & 0 & 9 \\ 4 & 3 & 2 & 9 & 8 \\ 3 & 1 & 1 & 2 & 4 \end{pmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 4 & 2 & 1 \\ 5 & 7 & 1 & 0 & 9 \\ 4 & 3 & 2 & 9 & 8 \\ 3 & 1 & 1 & 2 & 4 \end{bmatrix}.$$

行列の成分を囲む括弧は柔らかいものでも堅いものでも構わない. 板書では黒板の余白を有効に利用するため堅い括弧を用いることが多いと思うが, 気分によっては柔らかいほうを用いることもある. 前節で定めた行ベクトルとは $1 \times n$ 行列であり, 列ベクトルとは $n \times 1$ 行列のことである. 今後, 成分が n 個あるベクトルを n 次ベクトルと呼ぶことにしよう.

A を (m, n) -行列とする. 各 $i = 1, \dots, m$ および $j = 1, \dots, n$ について, A の i 行 j 列目の数を (i, j) -成分と呼ぶ. 例 2.1.1 における行列 A の 3 行目とは 5 次行ベクトル $(4, 3, 2, 9, 8)$ のことであり, 2 列目と

は 4 次列ベクトル $\begin{pmatrix} 0 \\ 7 \\ 3 \\ 1 \end{pmatrix}$ のことである. A の $(3, 2)$ -成分は 3 である.

一般の $m \times n$ 行列 A を成分表示すると次のようになる. A の (i, j) -成分を a_{ij} とするとき¹,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

毎回このような表示を用いては手間がかかるゆえ, 場合によっては, これを

$$A = [a_{ij}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}}$$

と書く. ここで, 上式では変数 i, j のどちらが行を表すのか不明である. そこで, 上式に現れる添え字 $i=1, \dots, m, j=1, \dots, n$ について, 上段に現れる変数が行に, 下段のそれが列に対応することと本書では約束する. 例え

¹ (i, j) -成分を a_{ij} と書くと, $(1, 23)$ -成分と $(12, 3)$ -成分の表記が共に a_{123} となり区別がつかない. このように誤解の恐れがある場合はカンマで区切って, それぞれ $a_{1,23}$ および $a_{12,3}$ と書く.

ば、上段と下段の変数に使う文字を入れ替えた行列 $B = [a_{ij}]_{\substack{j=1,2,3 \\ i=1,2,3}}$ について、 B の $(2, 3)$ -成分は $j = 2$, $i = 3$ とした場合の a_{ij} のこと、つまり a_{32} である (a_{23} ではない). なお、普段は行の変数に i を、列の変数に j を用いることが暗黙の前提となる場合が多い. そこで、 (i, j) の動く範囲や、どちらの変数が行を指すかに誤解の生じる恐れがないときは、上式を $A = [a_{ij}]$ と略記する.

二つの行列が等しいことを次で定める.

定義 2.1.2. $m \times n$ 行列 $A = [a_{ij}]$ および $\ell \times r$ 行列 $B = [b_{kh}]$ が等しいとは、 A, B のサイズが等しく、さらに各 (i, j) -成分が一致することである. すなわち、 $m = \ell$ かつ $n = r$ であり、更に各 i, j について $a_{ij} = b_{ij}$ が成り立つということである. A, B が等しい行列であるとき、 $A = B$ と書く.

したがって、例 2.1.1 における行列 A, B について、 $A = B$ である. 上の定義によれば、 n 次列ベクトルと n 次行ベクトルは、行列としては異なるものである. しかしながら、これらはユークリッド空間 \mathbb{R}^n の元としては同じ位置を示す場合があり、同じ位置を指すベクトルを異なるものと考えれば混乱が生じる恐れがあるだろう. そこで、 \mathbb{R}^n のベクトルについて論じる場合は、縦横どちらを用いてもよいが、その議論の最中はいずれか一方のみを用いると約束したい.

2.2 行列の和とスカラー倍

行列の和とスカラー倍の定義は、 \mathbb{R}^n のベクトルのそれとほとんどかわらない.

定義 2.2.1. サイズが等しい二つの $m \times n$ 行列 $A = [a_{ij}]$ および $B = [b_{ij}]$ に対して、 $z_{ij} := a_{ij} + b_{ij}$ を成分とする $m \times n$ 行列 $[z_{ij}]$ を A, B の和 (**sum**) といい、これを $A + B$ で表す. すなわち:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

A と B の和が定まるのは A, B のサイズが一致する場合のみである.

次で定められる、実数と行列の間の演算における実数のことをスカラー (**scalar**) と呼ぶ.

定義 2.2.2. $m \times n$ 行列 $A = [a_{ij}]$ および実数 $r \in \mathbb{R}$ について、 $w_{ij} := ra_{ij}$ を成分とする $m \times n$ 行列 $[w_{ij}]$ を A の r 倍といい、これを rA で表す. すなわち:

$$r \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ra_{11} & ra_{12} & \cdots & ra_{1n} \\ ra_{21} & ra_{22} & \cdots & ra_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ ra_{m1} & ra_{m2} & \cdots & ra_{mn} \end{bmatrix}.$$

A の実数倍たちを総称してスカラー倍 (**scalar multiplication**) という.

2.3 行列の積

何度も言うように、行列の演算とは線形写像のそれに対応するものである. \mathbb{R}^2 を定義域とする線形写像のベクトル値は、 $ax + by$ という形の数をもつことを 1 章で見た. これは 1 個 a 円のリンゴ x 個と 1 個 b 円のメロン y 個を購入するには合わせていくら必要か (答えは $ax + by$ 円) というたぐいの計算を複数回行うことに相当している. このように行列の理論とは、算数で扱うような単純計算を一般化し、昇華したものにほかならない.

さて、まずは $m \times n$ 行列 $A = [a_{ij}]$ と n 次列ベクトル $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ の間の積 $A\mathbf{x}$ を、やや天下りの的では

あるものの定義してしまおう。 $A\mathbf{x}$ は次で定義される m 次列ベクトルである：

$$A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} := \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.$$

ここで A の列の数と \mathbf{x} の成分数が等しく、 A の行の数と $A\mathbf{x}$ の成分数が等しいことに注意しておく。

例 2.3.1.

$$\begin{bmatrix} 1000 & 500 & 0 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1000 \cdot 4 + 500 \cdot 5 + 0 \cdot 1 \\ 3 \cdot 4 + 2 \cdot 5 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 6500 \\ 23 \end{bmatrix}.$$

例 2.3.2. 日帰りの団体旅行の計画があり、一人あたり次のような準備が必要であると見積もられている。また参加家族は次のように構成されているとする。

	大人	学生	幼児		斎藤	田端	嶺	...	
交通費(円)	1000	500	0		大人	4	1	2	...
おにぎり(個)	3	2	1		学生	5	2	2	...
⋮	⋮	⋮	⋮		幼児	1	1	0	...

例えば斎藤家に必要な準備を知るには、見積もり表の数値を成分とする行列と斎藤家の構成データによる列ベクトルの積を取ればよい。その計算は例 2.3.1 の通りであり、従って交通費 6500 円、おにぎり 23 個が必要となる。同様の計算が田端家(交通費 2000 円、おにぎり 8 個)や嶺家(交通費 3000 円、おにぎり 10 個)においてもなされ、これらの計算を一度に行うものとして行列の積は定義される。つまり、次のような計算を想定している。

$$\begin{bmatrix} 1000 & 500 & 0 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 5 & 2 & 2 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 6500 & 2000 & 3000 \\ 23 & 8 & 10 \end{bmatrix}.$$

行列の積の形式的な定義は次の通りである。

定義 2.3.3. $m \times n$ 行列 $A = [a_{ij}]$ と $n \times \ell$ 行列 $B = [b_{jk}]$ に対して、 $z_{ik} := a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}$ を成分とする $m \times \ell$ 行列 $[z_{ik}]$ を A と B の積 (product) といい $A \cdot B$ で表す。すなわち、

$$[a_{ij}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}} \cdot [b_{jk}]_{\substack{j=1, \dots, n, \\ k=1, \dots, \ell}} = \left[\sum_{j=1}^n a_{ij}b_{jk} \right]_{\substack{i=1, \dots, m, \\ k=1, \dots, \ell}}.$$

通常は、積の記号 \cdot (ドット) を略して AB と書く。

AB の (i, k) -成分とは、 A の i 行目 (これは n 次行ベクトル) と B の k 列目 (これは n 次列ベクトル) の積である²。行列の積 AB を定めるには、 A の列の数と B の行の数が一致せねばならないことに注意せよ。これは、線形写像 f と g の合成 $g \circ f$ を考えるとき、 f に代入して得られたベクトルが g の定義域の元でなければならないことに対応している。また A の行の数や B の列の数はいくらあってもよい。これは例 2.3.2 において、見積もり表にお茶 (ml)、お菓子代 (円)、入場料 (円) などのデータを加えたり、家族の構成表に別の家族のデータを加えたりしても上手く計算ができることに対応している。

²これはベクトル間の内積に相当する数と考えることもできる。この点については例 3.4.3(2) でもう一度述べる。

例 2.3.4. (1) 行列の積の計算に慣れないうちは次のように補助線を引いておくと見やすく計算できる。左側の行列を行について分割し、右側の行列を列について分割している。

$$\begin{aligned} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 3 & 2 & 0 & 2 \\ 1 & 1 & 2 & 0 \end{array} \right] \left[\begin{array}{c|c|c} 0 & 3 & 1 \\ 2 & 0 & 1 \\ 0 & 2 & 0 \end{array} \right] &= \begin{bmatrix} 1 \cdot 0 + 0 \cdot 2 + 1 \cdot 0 & 1 \cdot 3 + 0 \cdot 0 + 1 \cdot 2 & 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 0 \\ 3 \cdot 0 + 2 \cdot 2 + 0 \cdot 0 & 3 \cdot 3 + 2 \cdot 0 + 0 \cdot 2 & 3 \cdot 1 + 2 \cdot 1 + 0 \cdot 0 \\ 1 \cdot 0 + 1 \cdot 2 + 2 \cdot 0 & 1 \cdot 3 + 1 \cdot 0 + 2 \cdot 2 & 1 \cdot 1 + 1 \cdot 1 + 2 \cdot 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 5 & 1 \\ 4 & 9 & 5 \\ 2 & 7 & 2 \end{bmatrix}. \end{aligned}$$

上の一つ目の等号の後の細かい計算はノートに書かずに暗算できるようにしておくこと。

(2) 次の二つの計算を混同しないよう注意せよ。

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} d \\ e \\ f \end{bmatrix} = ad + be + cf, \quad \begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} d & e & f \end{bmatrix} = \begin{bmatrix} ad & ae & af \\ bd & be & bf \\ cd & ce & cf \end{bmatrix}.$$

行列の積の計算を行うには、単純ではあるものの多くの計算を繰り返さなければならない。例えば3次正方行列どうしの積を計算するには、一つの成分を求めるのに掛け算を3回、足し算を2回、合わせて5回の計算を行う。したがって、すべての成分を求めるには計45回の計算が必要になる。このうち一つでも計算を誤れば正しい結果は得られない。理論の理解と計算の正確性は別次元の話であり、計算練習によって自身の計算精度を確かめておくといよい(試験対策のためである)。

2.4 行列演算の性質

行列に関するいくつかの概念をここでまとめて定めておく。

- すべての成分がゼロになる $m \times n$ 行列を零行列 (zero matrix) とよび O_{mn} と書く。すなわち、 $O_{mn} = [0]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}}$ である。行列のサイズに誤解が生じない場合は、これを O と略記する。
- $n \times n$ 行列のことを n 次正方行列 (square matrix of order n) という。
- n 次正方行列 $A = [a_{ij}]$ において (i, i) -成分 a_{ii} ($i = 1, \dots, n$) を A の対角成分 (diagonal entry) という。
- すべての対角成分が1で、それ以外の成分がすべてゼロとなる正方行列を単位行列 (unit matrix) あるいは恒等行列 (identity matrix) という。本書では n 次単位行列を E_n で表し、行列のサイズに誤解が生じない場合は E と略記する³。

$$E_1 = 1, \quad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- $(-1)A$ を $-A$ と書く。また、 $A + (-B)$ のことを $A - B$ と書く。

命題 2.4.1. 行列の演算は次の性質を満たす。ただし、行列 A, B, C の各サイズは演算が定義されることを前提とし、 a, b を実数とする。

³単位行列を表す記号には、通常 I あるいは $E, \mathbf{1}$ などを用いる。これは identity matrix(英) あるいは Einheitsmatrix(独) の頭文字による。一般に、代数演算においてどこでも変わらない元のことを単位元 (identity element または identity) という。数の単位である1は実数の積に関する単位元であり、単位行列 E_n は n 次正方行列の積に関する単位元である。

- (1) $A + B = B + A$, (2) $A + O = A$, (3) $(A + B) + C = A + (B + C)$, (4) $A + (-A) = O$,
 (5) $AE = A$, (6) $EA = A$, (7) $AO = O$, (8) $OA = O$,
 (9) $0A = O$, (10) $1A = A$, (11) $(ab)A = a(bA)$, (12) $(aA)B = a(AB)$,
 (13) $(aA)(bB) = (ab)(AB)$, (14) $a(A + B) = aA + aB$, (15) $(a + b)A = aA + bA$,
 (16) $A(B + C) = AB + AC$, (17) $(A + B)C = AC + BC$, (18) $(AB)C = A(BC)$.

A が正方行列でない場合は (5) と (6) における E のサイズが異なることに注意せよ。また (7) における左辺の O と右辺の O もサイズが違う可能性がある。(8) についても同様である。

上の性質のうち (3) および (11),(12),(18) は結合律と呼ばれる。これは、どちらの演算を先に行っても結果が同じになることを意味し、それゆえ $A + B + C$ という表記が許されることになる。同様のことが abA , aAB , ABC についても言える。とくに、正方行列 A の k 個の積を取る演算はどの部分の積から順に計算しても性質 (18) により結果は同じであり、これを A^k と書く。すなわち、自然数 k について、

$$A^k := \underbrace{AA \cdots A}_{k \text{ 個の積}}.$$

また、 A の k 個の和は、 k による A のスカラー倍 kA に等しいことが性質 (10) および (15) から導かれる:

$$\underbrace{A + A + \cdots + A}_{k \text{ 個の和}} = 1A + 1A + \cdots + 1A = (1 + 1 + \cdots + 1)A = kA.$$

(14) から (17) までの 4 つの性質は分配律と呼ばれる。(13) をスカラー律という。(4) は、(9) および (10), (15) から導ける:

$$A + (-A) = 1A + (-1)A = (1 + (-1))A = 0A = O.$$

このように演算の性質を抽出する利点は、一つには行列の演算の定義の詳細に触れずとも議論を進められることにある(練習 2.4.3 も見よ)。

さて、本来ならば、命題 2.4.1 に挙げた性質すべてが成立することを証明しなければならない。しかし、すべてに時間を割く暇はないから、ここでは代表的なものをいくつか取り出して、(1) および (5), (16), (18) について紹介するにとどめる。これらの証明ができれば、おそらく他の性質も容易に証明できよう。

(1) $A + B = B + A$ の証明. $A = [a_{ij}]$, $B = [b_{ij}]$ とおくと、

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A.$$

□

(16) $A(B + C) = AB + AC$ の証明. A を (m, n) -行列, B, C を (n, ℓ) -行列とし, $A = [a_{ij}]$, $B = [b_{jk}]$, $C = [c_{jk}]$ とおくと、

$$\begin{aligned} A(B + C) &= A([b_{jk}] + [c_{jk}]) = [a_{ij}][b_{jk} + c_{jk}] = \left[\sum_{j=1}^n a_{ij}(b_{jk} + c_{jk}) \right] = \left[\sum_{j=1}^n (a_{ij}b_{jk} + a_{ij}c_{jk}) \right] \\ &= \left[\sum_{j=1}^n a_{ij}b_{jk} + \sum_{j=1}^n a_{ij}c_{jk} \right] = \left[\sum_{j=1}^n a_{ij}b_{jk} \right] + \left[\sum_{j=1}^n a_{ij}c_{jk} \right] = [a_{ij}][b_{jk}] + [a_{ij}][c_{jk}] = AB + AC. \end{aligned}$$

□

性質 (5) および結合律 (18) の証明は、次章においていくつかの記号の使い方を導入したうえで行おう。

例 2.4.2. サイズが同じ正方行列 A, B について、 $AB = BA$ は一般には成り立たない。例えば $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ とすれば、 $AB = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$, $BA = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$ である。

練習 2.4.3. A, B を n 次正方行列とする. 次の等式が成り立つか述べ, 正しければ証明し, 正しくなければ反例を挙げよ.

$$(a) (A + 2E)(A + E) = A^2 + 3A + 2E, \quad (b) (A + B)^2 = A^2 + 2AB + B^2.$$

(a) : 正しい. 実際, 次のように計算する. 誤解がないよう $C = (A + E)$ とおく.

$$\begin{aligned} (A + 2E)(A + E) &= (A + 2E)C = AC + 2EC = AC + 2C = A(A + E) + 2(A + E) \\ &= A^2 + A + 2A + 2E = A^2 + 3A + 2E. \end{aligned}$$

(b) : 一般には成立しない. $C = (A + B)$ とおき左辺を展開してみると,

$$(A + B)^2 = (A + B)C = AC + BC = A(A + B) + B(A + B) = A^2 + AB + BA + B^2.$$

したがって, 仮に式 (b) 左辺と右辺が等しいとすれば, 両辺から $A^2 + AB + B^2$ を引くことで $BA = AB$ を得る. しかし例 2.4.2 で見たように, これは一般には成り立たない.

発展 (代数構造)

代数構造とは集合の元に対して定義される何らかの演算のことであり, 数学では色々な代数構造を持った集合が扱われている. 一番なじみが深いものは \mathbb{R} や \mathbb{C} , \mathbb{Q} のように加減乗除の四則演算が成立する世界のことで, これを体 (**field**) という. \mathbb{Z} は体ではない. 何故なら, 整数どうしの割り算が整数にならないからである. 無理数全体も体ではない.

ある集合の各元どうしについて和, 差, 積の演算および二つの特別な元 O と E (これらをそれぞれ零元, 単位元という) が定義されており, 命題 2.4.1 の性質 (1) から (8), および (16) から (18) が成立する代数構造を環 (**ring**) という. 例えば, \mathbb{Z} は零元 O として 0 を, 単位元 E として 1 を採用することで環とみなせる. なお, 環の定義に単位元の存在を外す立場もある. この立場では, 例えば偶数全体は単位元を含まない環である.

練習 2.4.3 は, 行列に限らず一般に, 環に対して問われるべき問題である. まったく同じ議論により, 任意の環において等式 (a) は正しいことが分かる. 等式 (b) についてはどうだろうか. すべての元について $AB = BA$ を満たす環を可換環 (**commutative ring**) といい, 可換環においては (b) は成立する (ゆえに \mathbb{Z} においても成り立つ). そうでない環において等式 (b) は一般には成り立たない. その理由も練習 2.4.3 で述べた通りである. このように, 抽象的に性質を挙げておくと, 全く同じ論法で別の世界の話についても同時に議論することができる. 数学において抽象的な定義を採用する理由は, こうした汎用性を考慮したことによるのである.

他にも, まだまだ代数構造はたくさんある. 例えば, スカラー倍が環の中で定まっており, 命題 2.4.1 の性質 (9) から (15) が成立する (したがって命題 2.4.1 の性質すべてを満たす) 代数構造を多元環または代数 (**algebra**) という. 自然数 n を固定しておき, 各成分に実数を持つ n 次正方行列全体の集合を $M_n(\mathbb{R})$ とすれば, これは多元環である. とくに $n = 1$ について, \mathbb{R} 自身は, 数としての積をスカラー倍でもあると思うことで多元環ともみなせる.

いま, かなり抽象性の高い話をしており, 読者は既に食傷気味になっているかもしれない. ちなみに, 線形代数学で主として扱う代数構造は線形空間 (ベクトル空間) と呼ばれるものである. また, あまたある代数構造のうち最も重要なものは, 対称性を記述する群である. 線形空間については 15 章以降で論じる. 群については 7 章のコラムで紹介する.

第3章 行列の表し方

行列の表し方および記号の使い方について、いくつかの補足事項を説明する。ここで述べられていることは約束事であって、数学的に深い意味があるというわけではない。

3.1 クロネッカーのデルタ

定義 3.1.1. 次で定める δ_{ij} をクロネッカーのデルタ (Kronecker delta) という。

$$\delta_{ij} := \begin{cases} 1 & i = j \text{ のとき,} \\ 0 & i \neq j \text{ のとき.} \end{cases}$$

クロネッカーのデルタを用いれば、単位行列は $E = [\delta_{ij}]$ と表せる。

命題 2.4.1 (5) $AE = A$ の証明. $A = [a_{ij}]$ を (m, n) -行列, $E = [\delta_{jk}]$ を n 次単位行列とする。行列 $AE = [z_{ik}]$ の (i, k) -成分を定義にしたがって計算すると、

$$\begin{aligned} z_{ik} &= \sum_{j=1}^n a_{ij} \delta_{jk} = a_{i1} \cdot \delta_{1k} + \cdots + a_{i,(k-1)} \cdot \delta_{k-1,k} + a_{ik} \cdot \delta_{kk} + a_{i,k+1} \cdot \delta_{k+1,k} + \cdots + a_{in} \cdot \delta_{nk} \\ &= a_{i1} \cdot 0 + \cdots + a_{i,(k-1)} \cdot 0 + a_{ik} \cdot 1 + a_{i,k+1} \cdot 0 + \cdots + a_{in} \cdot 0 = a_{ik}. \end{aligned}$$

ゆえに AE と A の各 (i, k) -成分は等しく, $AE = A$. □

3.2 \sum 記号の使い方

和の記号 \sum や積の記号 \prod の使い方を詳しく説明しておこう。 n 個の数 a_1, \dots, a_n たちの和 $a_1 + a_2 + \cdots + a_n$ を

$$\sum_{i=1}^n a_i \quad \text{あるいは} \quad \sum_{i=1, \dots, n} a_i$$

と表す。積 $a_1 a_2 \cdots a_n$ については

$$\prod_{i=1}^n a_i \quad \text{あるいは} \quad \prod_{i=1, \dots, n} a_i$$

と表す。また, a_{ij} ($i = 1, \dots, m, j = 1, \dots, n$) たちの総和を $\sum_{\substack{i=1, \dots, m, \\ j=1, \dots, n}} a_{ij}$ と書く。こうした表記をより一

般的な立場から眺めると, 次のような説明ができる。

P を, ある変数に関する条件とする。ここで変数は多変数としてもよい。

例 3.2.1. P の具体例に以下のようなものがある:

- $P(i)$: “ $i \in \mathbb{N}$ ”
- $P(i)$: “ $i = 1$ または $i = 2$ または \dots または $i = n$ ”

注: この条件文 $P(i)$ を略して, 我々は “ $i = 1, \dots, n$ ” と書いている。

- $P(i)$: “ $i \neq 1$ かつ $i \neq 2$ ”

注: この条件文 $P(i)$ は “ $i = 1, 2$ ” の否定に相当するから, これを略して, “ $i \neq 1, 2$ ” と書く¹.

- $P(i, j)$: “ $i = 1, \dots, m$ かつ $j = 1, \dots, n$ ”

注: 通常は「かつ」を省略して記述することが多い. 一方, 「または」は勝手に略してはならない. 更に $m = n$ のとき, この条件文 $P(i, j)$ を略して “ $i, j = 1, \dots, n$ ” と書く.

- $P(i, j)$: “ $1 \leq i \leq j \leq 3$ ”

- $P(i, j)$: “ $1 \leq i \leq 5$ かつ $1 \leq j \leq 5$ ”

注: この条件文 $P(i, j)$ を略して, “ $1 \leq i, j \leq 5$ ” と書く.

定義 3.2.2. $P(i)$ たちが成立するような i をすべて動かして, これらの i に対応する a_i たちの総和を取るとき, これを $\sum_{P(i)} a_i$ と書く. 積 $\prod_{P(i)} a_i$ についても同様に定める.

例えば $P(i)$ として “ $i = 1, \dots, n$ ” を考えれば, $\sum_{i=1, \dots, n} a_i = a_1 + \dots + a_n$ である. ほかにも次のような使い方があある.

例 3.2.3. (1) $X = \{1, 2, 3\}$ のとき, $\sum_{i \in X} a_i = a_1 + a_2 + a_3$.

$$(2) \quad \sum_{1 \leq i \leq 5 \text{ かつ } i \text{ は整数}} a_i = a_1 + a_2 + a_3 + a_4 + a_5.$$

ただし, 通常は i が整数であることは暗黙裡に認めていることが多く, 「かつ i は整数」の部分は省略される.

$$(3) \quad \sum_{i=1,2, j=1,2,3} a_{ij} = a_{11} + a_{12} + a_{13} + a_{21} + a_{22} + a_{23}.$$

$$(4) \quad \sum_{1 \leq i < j \leq 3} a_{ij} = a_{11} + a_{12} + a_{13} + a_{22} + a_{23} + a_{33}.$$

$$(5) \quad \prod_{1 \leq i < j \leq 3} (x_i - x_j) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3).$$

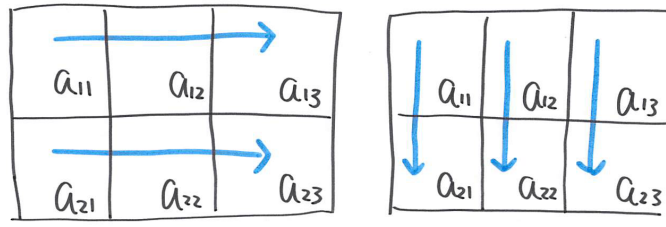
$$(6) \quad \sum_{\substack{i=1, \dots, n, \\ j=1, \dots, n}} a_{ij} = \sum_{i,j=1, \dots, n} a_{ij} = \sum_{1 \leq i, j \leq n} a_{ij}.$$

なお, \sum や \prod における添え字 i の動く範囲に誤解がない場合は, \sum_i あるいは \prod_i と略記することがある. 本書ではこのような曖昧な表記は行わない.

各 a_{ij} (ただし $i = 1, \dots, m, j = 1, \dots, n$) の総和を下図のように二通りの方法で順番に足していく. 足す順番を入れ替えても総和は等しいことから, 次の等式を得る:

例 3.2.4. $\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \right) = \sum_{j=1}^n \left(\sum_{i=1}^m a_{ij} \right)$. この等式に現れる括弧は略すことが多い.

¹この文と “ $i \neq 1$ または $i \neq 2$ ” を混同しないよう注意すること. なお余談になるが, “ $i \neq 1$ または $i \neq 2$ ” とは, i はどんな数でもよいことを意味している.



行列の積の結合律を証明しよう。

命題 2.4.1 (18) $(AB)C = A(BC)$ の証明. $A = [a_{ij}]$ を $m \times n$ 行列, $B = [b_{jk}]$ を $n \times \ell$ 行列, $C = [c_{kh}]$ を $\ell \times r$ 行列とし, $(AB)C$ および $A(BC)$ の各成分を積の定義にしたがって計算すると次のようになる.

$$\begin{aligned} (AB)C &= ([a_{ij}][b_{jk}])[c_{kh}] = \left[\sum_{j=1}^n a_{ij}b_{jk} \right] [c_{kh}] \quad \left(\text{ここで } x_{ik} := \sum_{j=1}^n a_{ij}b_{jk} \text{ とおく} \right) \\ &= [x_{ik}][c_{kh}] = \left[\sum_{k=1}^{\ell} x_{ik}c_{kh} \right] = \left[\sum_{k=1}^{\ell} \left(\sum_{j=1}^n a_{ij}b_{jk} \right) c_{kh} \right] = \left[\sum_{k=1}^{\ell} \left(\sum_{j=1}^n a_{ij}b_{jk}c_{kh} \right) \right]. \end{aligned}$$

$$\begin{aligned} A(BC) &= [a_{ij}]\left([b_{jk}][c_{kh}]\right) = [a_{ij}] \left[\sum_{k=1}^{\ell} b_{jk}c_{kh} \right] \quad \left(\text{ここで } y_{jh} := \sum_{k=1}^{\ell} b_{jk}c_{kh} \text{ とおく} \right) \\ &= [a_{ij}][y_{jh}] = \left[\sum_{j=1}^n a_{ij}y_{jh} \right] = \left[\sum_{j=1}^n \left(a_{ij} \sum_{k=1}^{\ell} b_{jk}c_{kh} \right) \right] = \left[\sum_{j=1}^n \left(\sum_{k=1}^{\ell} a_{ij}b_{jk}c_{kh} \right) \right]. \end{aligned}$$

$(AB)C$ と $A(BC)$ の各 (i, h) -成分が等しいことは例 3.2.4 より分かる. □

発展 (写像の合成の結合律)

結合律 $(AB)C = A(BC)$ を線形写像の言葉で述べれば, それは写像の合成に関する結合律 $(h \circ g) \circ f = h \circ (g \circ f)$ のことである. 合成に関する結合律は, 線形写像に限らずとも一般の写像について成立する:

命題 3.2.5. 三つの写像 $f: X \rightarrow Y, g: Y \rightarrow Z, h: Z \rightarrow W$ について, $(h \circ g) \circ f = h \circ (g \circ f)$.

Proof. 二つの写像が等しいとは, いかなる元を代入してもその結果が一致するということである. f の定義域のいかなる元 $x \in X$ についても $((h \circ g) \circ f)(x) = (h \circ (g \circ f))(x)$ となることを示そう. まず $h \circ g$ を $\phi, f(x)$ を y とおくことで

$$((h \circ g) \circ f)(x) = \phi \circ f(x) = \phi(f(x)) = (h \circ g)(y) = h(g(y)) = h(g(f(x))).$$

次に $\psi = g \circ f$ とおくことで

$$(h \circ (g \circ f))(x) = h \circ \psi(x) = h(\psi(x)) = h((g \circ f)(x)) = h(g(f(x))).$$

ゆえに $(h \circ g) \circ f = h \circ (g \circ f)$ である. □

線形写像の合成と行列の積が対応することの詳しい説明は, ユークリッド空間上の線形写像については 21.3 節において, また一般の線形空間上の線形写像については 26.2 節において行う.

3.3 成分の空白と任意性

行列 $A = [a_{ij}]$ を成分表示するとき、成分が0となる部分は何も書かずに空白で表すことがある。また、まとまった領域においてすべての成分が0のとき、これらをまとめて O で表す。例えば、単位行列は次のように書かれる。

$$\begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}, \begin{bmatrix} 1 & & O \\ & 1 & \\ O & & \ddots \\ & & & 1 \end{bmatrix}.$$

あまり重要でない成分は * と書かれる。同じ記号 * を用いるものの、各成分において異なる数が入っていてもよいと考える。例えば、 $B = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 4 \end{bmatrix}$ を略して $\begin{bmatrix} 2 & * & * \\ & 3 & * \\ & & 4 \end{bmatrix}$ と書く。このように略しても、 B^n の対角成分は計算できる:

$$B^2 = \begin{bmatrix} 2 & * & * \\ & 3 & * \\ & & 4 \end{bmatrix} \begin{bmatrix} 2 & * & * \\ & 3 & * \\ & & 4 \end{bmatrix} = \begin{bmatrix} 2^2 & * & * \\ & 3^2 & * \\ & & 4^2 \end{bmatrix}, \quad B^n = \begin{bmatrix} 2^n & * & * \\ & 3^n & * \\ & & 4^n \end{bmatrix}.$$

また、まとまった領域において成分情報が不要であるとき、これらをまとめて * と書く。

定義 3.3.1. 対角成分より下の成分がすべて0なる行列を上三角行列 (**upper triangular matrix**) という。すなわち、次のような行列のことである:

$$\begin{bmatrix} a_1 & & & \\ & a_2 & * & \\ & & \ddots & \\ O & & & a_n \end{bmatrix}.$$

線形写像を行列を用いて数値化する際には、出来る限り複雑でない行列によって表すことが望ましい。正方行列を用いて表現できる任意の線形写像は、座標 (基底) を上手く与えることでその行列を上三角行列に取れることを後に学ぶことになる。

3.4 転置行列

(m, n) -行列を成分表示し、各 (i, i) -成分を結ぶ線を軸に裏返すことで得られる (n, m) -行列のことを、もとの行列の転置行列という。形式的には次のように定義する。

定義 3.4.1. (m, n) -行列 $A = [a_{ij}]$ に対して、 $b_{kh} := a_{hk}$ (ただし $k = 1, \dots, n, h = 1, \dots, m$) を成分とする (n, m) -行列 $[b_{kh}]$ を A の転置行列 (**transposed matrix**) といい、これを tA と書く。

例 3.4.2. (1) $A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ とし、 $b_{kh} := a_{hk}$ ($k = 1, 2, h = 1, 2$) と定めれば、

$$\begin{aligned} b_{11} &= a_{11} = \alpha, & b_{12} &= a_{21} = \gamma, \\ b_{21} &= a_{12} = \beta, & b_{22} &= a_{22} = \delta, \end{aligned} \quad \text{であるから} \quad {}^tA = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}.$$

$$(2) \quad {}^t \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix} = \begin{bmatrix} 1 & 6 & 11 \\ 2 & 7 & 12 \\ 3 & 8 & 13 \\ 4 & 9 & 14 \\ 5 & 10 & 15 \end{bmatrix}.$$

- (3) 上三角行列の転置行列を下三角行列 (**lower triangular matrix**) と呼ぶ. また, 上三角行列と下三角行列をまとめて三角行列 (**triangular matrix**) と呼ぶ. 上三角かつ下三角な行列は対角行列である.

例 3.4.3. (1) 列ベクトルをそのまま書くと行数を稼いでしまうため, 次のように行ベクトルの転置で表すことがある:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = {}^t(x_1, \dots, x_n).$$

- (2) サイズが等しい二つのベクトルの内積 (inner product) を行列としての積と転置を用いて表すことができる. すなわち, 行ベクトル \mathbf{x}, \mathbf{y} に対して, これらの内積を $\mathbf{x} \cdot \mathbf{y}$ で定め, 列ベクトル \mathbf{x}, \mathbf{y} に対して, これらの内積を ${}^t\mathbf{x} \cdot \mathbf{y}$ で定める.

転置行列の真の意味, すなわち線形写像としての意味をここで述べるのは難しい². いまは, A の各行 (あるいは列) どちらの内積の全情報を得るための操作と考えておけばよいだろう. A の各行どうしの内積を成分とする (m, m) -行列は $A {}^tA$ で与えられる. 各列どうしの内積は, (n, n) -行列 ${}^tA A$ で与えられる. 転置行列の成分表示を形式的に表現する際に, 次の点に注意する必要がある.

備考 3.4.4. (1) 上の定義 3.4.1 において, tA の行と列を表す添え字 k, h を文字 i, j に置き換えれば, $[b_{ij}]_{\substack{i=1, \dots, n, \\ j=1, \dots, m}}$ となる. すなわち,

$${}^t \left([a_{ij}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}} \right) = [b_{ij}]_{\substack{i=1, \dots, n, \\ j=1, \dots, m}} = [a_{ji}]_{\substack{i=1, \dots, n, \\ j=1, \dots, m}}. \quad (3.4.1)$$

上の左辺と右辺では i, j の動く範囲が異なっていることに注意すること. 多くの文献において行列 $[a_{ij}]$ の転置行列は $[a_{ji}]$ と書かれる. しかし, これら二つの間で i, j の動く範囲が異なることから, 初学者は混乱することもあるだろう. 例えば次の (2) を見よ.

- (2) (m, n) -行列 $A = [a_{ij}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}}$ に対して, 各成分に関する添え字の文字のみを入れ替えた $B = [a_{ji}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}}$ や, これに変数の範囲を指す上下の段を入れ替えた $C = [a_{ji}]_{\substack{j=1, \dots, n \\ i=1, \dots, m}}$ は A の転置行列ではない³. この件について具体例で考えてみよう.

いま六つの数 a_{ij} ($i = 1, 2, j = 1, 2, 3$) を $a_{11} = 1, a_{12} = 2, a_{13} = 3, a_{21} = 4, a_{22} = 5, a_{23} = 6$ と定め, $(2, 3)$ -行列 A を $A = [a_{ij}]_{\substack{i=1, 2, \\ j=1, 2, 3}}$ と定める. このとき, $B = [a_{ji}]_{\substack{i=1, 2, \\ j=1, 2, 3}}$, $C = [a_{ji}]_{\substack{j=1, 2, 3 \\ i=1, 2}}$ と定めようとするれば, A, B, C の成分表示は次のようになる:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = [a_{ji}]_{\substack{i=1, 2, \\ j=1, 2, 3}} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \end{bmatrix} = \begin{bmatrix} 1 & 4 & ? \\ 2 & 5 & ? \end{bmatrix},$$

$$C = [a_{ji}]_{\substack{j=1, 2, 3 \\ i=1, 2}} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ ? & ? \end{bmatrix}.$$

つまり, a_{31} と a_{32} は未だ定めていないゆえ, B や C は定義が不適切であることが分かる. 一方, ${}^tA = [a_{ij}]_{\substack{j=1, \dots, n, \\ i=1, \dots, m}}$ であるが, 転置行列をこのように表すことは稀であり, ふだんは式 (3.4.1) の右辺のように書く.

命題 3.4.5. (1) ${}^t(A + B) = {}^tA + {}^tB$. (2) ${}^t(AB) = {}^tB {}^tA$.

²これは与えられた線形写像の双対写像と呼ばれるものに相当する.

³ただし例外として $m = n$ の場合, すなわち A が正方行列のときに限り, $B = {}^tA$ および $C = A$ が成り立つ.

Proof. (1): ${}^t(A+B) = {}^t([a_{ij}] + [b_{ij}]) = {}^t([a_{ij} + b_{ij}]) = [a_{ji} + b_{ji}] = [a_{ji}] + [b_{ji}] = {}^tA + {}^tB$.

(2): (m, n) -行列 $A = [a_{ij}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}}$, および (n, ℓ) -行列 $B = [b_{jk}]_{\substack{j=1, \dots, n, \\ k=1, \dots, \ell}}$ について, 式 (3.4.1) より

$${}^tA = [a_{ji}]_{\substack{i=1, \dots, n, \\ j=1, \dots, m}}, \quad {}^tB = [b_{kj}]_{\substack{j=1, \dots, \ell, \\ k=1, \dots, n}}$$

である. ${}^tB {}^tA$ および ${}^t(AB)$ の各成分をそれぞれ計算しよう.

$$\begin{aligned} {}^tB {}^tA &= [b_{kj}]_{\substack{j=1, \dots, \ell, \\ k=1, \dots, n}} \cdot [a_{ji}]_{\substack{i=1, \dots, n, \\ j=1, \dots, m}} = [b_{\beta\alpha}]_{\substack{\alpha=1, \dots, \ell, \\ \beta=1, \dots, n}} \cdot [a_{\gamma\beta}]_{\substack{\beta=1, \dots, n, \\ \gamma=1, \dots, m}} \quad (\text{ここで添え字の文字を書き換えた}) \\ &= [y_{\alpha\beta}]_{\substack{\alpha=1, \dots, \ell, \\ \beta=1, \dots, n}} \cdot [x_{\beta\gamma}]_{\substack{\beta=1, \dots, n, \\ \gamma=1, \dots, m}} \quad (\text{ここで } y_{\alpha\beta} := b_{\beta\alpha}, x_{\beta\gamma} := a_{\gamma\beta} \text{ と置いた}) \\ &= \left[\sum_{\beta=1}^n y_{\alpha\beta} x_{\beta\gamma} \right]_{\substack{\alpha=1, \dots, \ell, \\ \gamma=1, \dots, m}} = \left[\sum_{\beta=1}^n b_{\beta\alpha} a_{\gamma\beta} \right]_{\substack{\alpha=1, \dots, \ell, \\ \gamma=1, \dots, m}} = \left[\sum_{\beta=1}^n a_{\gamma\beta} b_{\beta\alpha} \right]_{\substack{\alpha=1, \dots, \ell, \\ \gamma=1, \dots, m}} \\ &= \left[\sum_{j=1}^n a_{kj} b_{ji} \right]_{\substack{i=1, \dots, \ell, \\ k=1, \dots, m}} \quad (\text{ここで再び添え字の文字を書き換えた}). \end{aligned}$$

一方で,

$$\begin{aligned} {}^t(AB) &= {}^t \left([a_{ij}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}} \cdot [b_{jk}]_{\substack{j=1, \dots, n, \\ k=1, \dots, \ell}} \right) = {}^t \left(\left[\sum_{j=1}^n a_{ij} b_{jk} \right]_{\substack{i=1, \dots, m, \\ k=1, \dots, \ell}} \right) \\ &= {}^t \left([z_{ik}]_{\substack{i=1, \dots, m, \\ k=1, \dots, \ell}} \right) \quad (\text{ここで } z_{ik} := \sum_{j=1}^n a_{ij} b_{jk} \text{ と置いた}) \\ &= [z_{ki}]_{\substack{i=1, \dots, \ell, \\ k=1, \dots, m}} = \left[\sum_{j=1}^n a_{kj} b_{ji} \right]_{\substack{i=1, \dots, \ell, \\ k=1, \dots, m}}. \end{aligned}$$

以上より, ${}^tB {}^tA$ と ${}^t(AB)$ の (i, k) -成分はすべて一致している. □

3.5 行列の分割

次のように行列をいくつかの小さい行列に分割し, あたかも行列を成分にもつ行列であるかのように見なして計算してもよい.

$$\begin{aligned} \left[\begin{array}{cc|cc} 1 & 0 & 3 & 4 & 5 \\ 6 & 7 & 8 & 1 & 0 \\ 11 & 12 & 13 & 0 & 1 \end{array} \right] \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 6 & 7 \\ 1 & 2 \\ 0 & 1 \end{array} \right] &= \left[\begin{array}{cc} [1 \ 2] \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] + 3 [6 \ 7] + [4 \ 5] \left[\begin{array}{cc} 1 & 2 \\ 0 & 1 \end{array} \right] \\ \left[\begin{array}{cc} 6 & 7 \\ 11 & 12 \end{array} \right] \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] + [8 \ 13] [6 \ 7] + \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] \left[\begin{array}{cc} 1 & 2 \\ 0 & 1 \end{array} \right] \end{array} \right] \\ &= \left[\begin{array}{cc} [1 \ 2] + [18 \ 21] + [4 \ 13] \\ \left[\begin{array}{cc} 6 & 7 \\ 11 & 12 \end{array} \right] + \left[\begin{array}{cc} 48 & 56 \\ 78 & 91 \end{array} \right] + \left[\begin{array}{cc} 1 & 2 \\ 0 & 1 \end{array} \right] \end{array} \right] = \left[\begin{array}{cc} 23 & 36 \\ 55 & 65 \\ 89 & 104 \end{array} \right]. \end{aligned}$$

こうした分割による積の計算と元々の定義による積の計算が一致することは, 各成分がどのような成分の積たちの和になっているか調べれば分かる⁴. 分割計算を行う際の留意点は, 積がきちんと定まるような分割を行う必要があることである.

⁴具体的な計算を通して理解できることから, ここでは形式的な証明は省略する. 例えば巻末の文献 [1] に証明が載っている.

例 3.5.1. (m, n) -行列 A と (n, ℓ) -行列 B において, A, B をそれぞれ行ベクトル表示および列ベクトル表示しよう. つまり, $\mathbf{a}_1, \dots, \mathbf{a}_m$ を n 次行ベクトル, $\mathbf{b}_1, \dots, \mathbf{b}_\ell$ を n 次列ベクトルとして

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}, \quad B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_\ell \end{bmatrix},$$

と分割する. このとき, 積 $AB = [z_{ik}]$ の (i, k) -成分とは, A の i 行目と B の k 列目の積のことであった (すなわち $z_{ik} = \mathbf{a}_i \mathbf{b}_k$). AB の分割として, 次のような表し方がある:

$$A \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_\ell \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_\ell \end{bmatrix}, \quad \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} B = \begin{bmatrix} \mathbf{a}_1 B \\ \mathbf{a}_2 B \\ \vdots \\ \mathbf{a}_m B \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_\ell \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \mathbf{b}_1 & \mathbf{a}_1 \mathbf{b}_2 & \cdots & \mathbf{a}_1 \mathbf{b}_\ell \\ \mathbf{a}_2 \mathbf{b}_1 & \mathbf{a}_2 \mathbf{b}_2 & \cdots & \mathbf{a}_2 \mathbf{b}_\ell \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_m \mathbf{b}_1 & \mathbf{a}_m \mathbf{b}_2 & \cdots & \mathbf{a}_m \mathbf{b}_\ell \end{bmatrix}.$$

なお, A の行ベクトル表示および B の列ベクトル表示をカンマで区切って, それぞれ $A = {}^t[t\mathbf{a}_1, \dots, t\mathbf{a}_m]$, $B = [\mathbf{b}_1, \dots, \mathbf{b}_\ell]$ と書くこともある.

行列の分割による計算は, 行列の性質を帰納法を用いて示す際に威力を発揮する. 行列のサイズに関する帰納法において, 分割によって小さくなった行列に帰納法の仮定を適用するという論法がよく使われている.

を解くことと、行列に関する次の関係式

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (4.2.2)$$

を満たす n 次列ベクトル $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ を探すことは同じことである。式 (4.2.1) あるいは式 (4.2.2) の形で与えられる方程式のことを連立 1 次方程式 (system of linear equations) という。

【補足】 連立 1 次方程式 (4.2.1) における変数に現れる各係数を並べた行列 $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ のことを、方程式 (4.2.1) の係数行列 (coefficient matrix) と呼ぶ。

以後、連立 1 次方程式を解くことは、既知の行列 A および列ベクトル \mathbf{b} が与えられたとき、 $A\mathbf{x} = \mathbf{b}$ を満たす未知のベクトル \mathbf{x} を求めることと解釈し、話を進めよう。

4.3 逆行列を持つ場合

まずは最も簡単な場合の解法を紹介する。それは A が正方行列であり、更に $AB = BA = E$ を満たす正方行列 B が存在する場合である。このような B のことを A の逆行列 (inverse) という。 A が逆行列を持つ場合は、次が成り立つ:

命題 4.3.1. 正方行列 A の逆行列 B が存在するとき、連立 1 次方程式 $A\mathbf{x} = \mathbf{b}$ は唯一解 $\mathbf{x} = B\mathbf{b}$ を持つ。

ここで老婆心ながら、大学レベルの数学を学ぶ際に必要な論理力を養うことを念頭に、上の命題の間違った証明を紹介しておく。次の証明のどこがまずいのか各自で考え、まずい理由が分かった者は正しい証明を自分でつけてみよ。

命題 4.3.1 の偽証明. $A\mathbf{x} = \mathbf{b}$ の両辺に左から B をかけると $BA\mathbf{x} = B\mathbf{b}$ 。この左辺を計算すると $BA\mathbf{x} = E\mathbf{x} = \mathbf{x}$ ゆえ $\mathbf{x} = B\mathbf{b}$ である。以上より、方程式 $A\mathbf{x} = \mathbf{b}$ は唯一解 $\mathbf{x} = B\mathbf{b}$ を持つ。 \square

上の証明の何がまずいのか分からなかった者は、これと類似する議論によって、次の間違った命題も正しいと判断してしまうであろう。

偽命題 4.3.2. 連立 1 次方程式

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \quad (4.3.1)$$

は唯一解 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ を持つ。

【備考】 方程式 (4.3.1) に解はない。なぜなら 3 行目に関する等式は $0x_1 + 0x_2 = 1$ であり、この式を満たす実数 x_1, x_2 は存在しないからである。

偽命題 4.3.2 の偽証明. 式 (4.3.1) の両辺に左から $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ をかけると

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 3 \\ 2 \end{bmatrix} \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 3 \\ 2 \end{bmatrix}. \end{aligned}$$

以上より, 方程式 (4.3.1) は唯一解 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ を持つ. □

上に挙げた二つの偽証明は, いずれも方程式の両辺に左から特別な行列 B をかけて解を導いた点で一致している. これらの証明のどこが悪かったかといえば, それは暗黙のうちに解が存在することを仮定したうえで議論をし, そうして導いた解を唯一解と述べてしまった点にある. 実際, 偽命題 4.3.2 の偽証明では, そもそも方程式に解が存在しないにも関わらず, ベクトル $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ が方程式の解であると仮定して, 式変形を続けていった. 本来ならば, これらの議論で分かることは「解が存在するとすれば, その形はこれこれである」という事実のみである. したがって解の唯一性を示したいのであれば, これとは別に「解が存在するかどうか」の証明も提示せねばならない.

以上の反省を踏まえたうえで, 改めて命題 4.3.1 を示そう.

命題 4.3.1 の証明. 列ベクトル \mathbf{a} が方程式 $A\mathbf{x} = \mathbf{b}$ の解であると仮定すれば, $A\mathbf{a} = \mathbf{b}$ が成り立つ. この両辺に左から A の逆行列 B をかけると $BA\mathbf{a} = B\mathbf{b}$. この左辺を計算すると $BA\mathbf{a} = E\mathbf{a} = \mathbf{a}$ ゆえ $\mathbf{a} = B\mathbf{b}$ である. したがって方程式に解が存在するとすれば, その解は唯一解 $\mathbf{x} = B\mathbf{b}$ のみであることが分かった. また, 実際に $\mathbf{x} = B\mathbf{b}$ はこの方程式の解である. 何故なら, $A\mathbf{x} = AB\mathbf{b} = E\mathbf{b} = \mathbf{b}$ ゆえ $A\mathbf{x} = \mathbf{b}$ を満たすからである. □

さて, ここで命題 4.3.1 について注釈しておこう. この命題は理論的にはすっきりとしているものの, これを直ちに応用しようとすれば, いくつかの欠点に気づく. 一つは, そもそも逆行列が存在するかどうかをどうやって判断するのかという点である. また, 逆行列 B を求める方法も今はまだ分からない. 更に, 逆行列が存在しない場合に, そうとは知らずにこの命題に準じた方向で方程式の求解に挑んでしまったら, 計算が徒労に終わってしまうのだろうか. 実は, これらの事柄はすべて繋がっていることが後で分かる (詳しくは 5 章を見よ).

4.4 行列の行基本変形

A を (m, n) -行列とする. このとき $A\mathbf{x} = \mathbf{b}$ は, 変数の数が n 個, 式の数が m 個の連立 1 次方程式である. 方程式を解くにあたり, 各係数に現れる数が単純 (例えば係数に 0 が多く現れる場合) なほど解は求めやすい. そこで, 係数になるべく単純になるよう式を変形する方法を考えたい. そのためには, 方程式を変形してもその式の本質的な意味が変わらない操作が何かをまず考える必要がある. 例えば, 次の 3 つの操作 (式の書き換え) は方程式の本質を変えない. すなわち, 変形する前と後で解の形が変わらない.

- (i) 実数 $r \neq 0$ について, 一つの式を両辺を r 倍した式に書き換える,
- (ii) m 個の式の順番を並べ替える,
- (iii) ある式の何倍かを別の式に加えることで, 新しい式に書き換える.

上の3つの操作で解の形が変化しない理由は、再び3つの操作のいずれかを行うことで元の方程式の形に戻ることができるからである。この3つの操作に対応する行列の変形操作を行基本変形 (elementary row operations) という。すなわち、次の3つの変形のことである:

- (1) 実数 $r \neq 0$ について、一つの行を r 倍する,
- (2) 二つの行を入れ替える,
- (3) ある行の何倍かを別の行に加える。

(m, n) -行列 A と m 次列ベクトル \mathbf{b} を並べた、 $(m, n+1)$ -行列 $[A|\mathbf{b}]$ を考えよう。

$$[A|\mathbf{b}] := \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right].$$

これを連立1次方程式 $A\mathbf{x} = \mathbf{b}$ に関する拡大係数行列あるいは単に拡大行列 (augmented matrix) という。行列 $[A|\mathbf{b}]$ に行基本変形の操作 (1) から (3) のいずれかを一回ほどこした行列を $[A'|\mathbf{b}']$ としよう。このとき、 n 次列ベクトル \mathbf{a} について、 \mathbf{a} が方程式 $A\mathbf{x} = \mathbf{b}$ の解であることと、 \mathbf{a} が方程式 $A'\mathbf{x} = \mathbf{b}'$ の解であることは同値である。これは、連立1次方程式に対して (i) から (iii) のいずれの式変形を行っても解が変化しないことから分かる。また、行基本変形を $[A|\mathbf{b}]$ に何度ほどこしてもこの状況は変わらない。したがって次を得る:

命題 4.4.1. 行列 $[A|\mathbf{b}]$ に有限回の行基本変形をほどこしたものを $[A'|\mathbf{b}']$ とすれば、方程式 $A\mathbf{x} = \mathbf{b}$ の解全体と方程式 $A'\mathbf{x} = \mathbf{b}'$ の解全体は一致する。

さて、連立方程式 $A\mathbf{x} = \mathbf{b}$ を解くには、行基本変形を用いてより簡単な連立方程式 $A'\mathbf{x} = \mathbf{b}'$ に帰着させればよいことが分かった。では、ここでいう「簡単」とはどういう意味だろうか。例えば A が正方行列であり、 $[A|\mathbf{b}]$ を運よく $[E|\mathbf{b}']$ に変形できたとすれば、これは簡単だと言えるだろう (実際、 $\mathbf{x} = \mathbf{b}'$ が唯一の解である)。一般の場合には単位行列になることは望めないものの (そもそも正方行列でなければ絶対に不可能である)、方程式の解が直ちに分かるような行列の形があり、それは簡約形と呼ばれている。

4.5 簡約行列

簡約行列を定義するために、まず主成分なる概念を導入する。行列 A の零ベクトルでない各行ベクトルに対して、その行のゼロでない最初の成分のことを、その行の主成分と呼ぶ。次の行列において下線が引かれた成分が各行の主成分である。

$$\left[\begin{array}{ccccc} \underline{2} & 0 & 1 & 5 & 4 \\ 0 & 0 & \underline{1} & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \underline{1} \end{array} \right].$$

定義 4.5.1. 次の (I) から (IV) すべてを満たす行列 $[a_{ij}]$ は簡約形であるといい、このような行列を簡約行列と呼ぶ。

- (I) 零ベクトルとなる行は零ベクトルでない行よりも下段にある。
- (II) 零ベクトルでない行の主成分は1である。
- (III) 第 i 行の主成分が第 j_i 列にあるとすれば、 $j_1 < j_2 < j_3 < \dots$ が成り立つ。つまり、各行の主成分は下にある行ほど右側にある。

(IV) ある行の主成分を含む列について、その列ベクトルの成分は主成分を除いてすべて0である。すなわち、(III)における記号および(II)の条件を合わせれば、第*i*行が零ベクトルでないとき、第*j_i*列は*i*行成分が1でそれ以外の成分は0となる。

【補足】上の性質を満たす行列の名称は参考書によってまちまちであり、例えば上の条件は行を基準にして定めた概念であることから行簡約形と呼ぶ場合もある。また、「簡約」という語句は「既約 (reduced)」とも言い換えられる。更には、定義4.5.1の条件(I)と(III)を満たす行列を行階段形 (row echelon form) であるといい¹、簡約形のことを行簡約階段形 (reduced row echelon form) と呼ぶこともある。主成分についても同様に様々な言い回しがあり、行に関するかなめ (pivot) あるいは先頭の成分 (leading entry)、先頭の係数 (leading coefficient) などとも呼ばれている。

練習 4.5.2. 次の行列のうち簡約行列はどれか。

$$(1) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2) \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad (3) \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

$$(4) \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (5) \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (6) \begin{bmatrix} 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

答え. (1)は(I)が成立しない。(2)は2列目について(IV)が不成立。(3)は2行目において(II)が不成立で、更に3列目において(IV)も成り立たない。(4)は(III)が不成立、(5)は(II)が不成立である。(6)は(I)から(IV)すべてを満たすゆえ簡約行列である。

行列*A*に行基本変形を何回かほどこして簡約行列*B*に変形させる操作を*A*を簡約化するといい、このとき*B*を*A*の簡約化と呼ぶ。簡約化の方法は次の手順で得られる。ただし、これはあくまでも形式的なアルゴリズムであり、実際に簡約化を行う際は、各自の判断で手順を入れ替えたほうが効率が良い場合もある。

- (1) 零ベクトルとなる行たちが下段に並ぶよう行の入れ替えを行う。
- (2) 主成分を含む行について、第*i*行の主成分が*j_i*列目にあるとする。このとき $j_1 \leq j_2 \leq j_3 \leq \dots$ が成り立つように行の入れ替えを行う。
- (3) $j_i = j_{i+1}$ となる場合は*i*行の何倍かを*i+1*行に加え、*i+1*行目の主成分であった(*i+1, j_i*)-成分を0にする。
- (4) $j_1 < j_2 < j_3 < \dots$ となるまで上の(1)から(3)の操作を繰り返す。
- (5) 行のスカラー倍を行い、各行の主成分を1にする。
- (6) ある行(これを第*i*行とする)の主成分を含む列において、その主成分以外の成分が0になるよう*i*行の何倍かを他の行に加える。

例 4.5.3. 次は簡約行列への行基本変形の一例である。実際の変形においては、2つ目と3つ目の変形を、4つ目と5つ目の変形を同時に行うなどして、計算過程をできるだけ省略する。

$$\begin{bmatrix} 2 & 0 & 1 & 5 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{1 \text{ 行} \times 2 \text{ 行}} \begin{bmatrix} 2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{1 \text{ 行} \times 4 \text{ 行}} \begin{bmatrix} 2 & 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{2 \text{ 行} \times 4 \text{ 行} \times 3} \begin{bmatrix} 2 & 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{入れ替え}} \begin{bmatrix} 2 & 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{1 \text{ 行} \times \frac{1}{2}} \begin{bmatrix} 1 & 0 & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

¹参考書によっては、行階段形に条件(II)も仮定する。

なお、行列の簡約化は唯一通りに定まる。これは、どのような行基本変形の手順を踏もうとも、最終的に得られる簡約化は必ず一致するということである。この証明は6.1節で述べよう(定理6.1.1)。

4.6 連立1次方程式の解法

簡約行列 $[B|\mathbf{b}]$ における連立1次方程式 $B\mathbf{x} = \mathbf{b}$ の解法を具体例を通して説明する。 $[B|\mathbf{b}]$ の主成分に関するある条件について、二つのケースに分けて考える。

(ケース1)：まずは $[B|\mathbf{b}]$ の最後の列 \mathbf{b} に主成分が含まれる場合である。例えば、

$$[B|\mathbf{b}] = \left[\begin{array}{cccc|c} 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right] \text{ のとき, } B\mathbf{x} = \mathbf{b} \text{ は } \begin{bmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

このとき、最後の行に関する等式を方程式の言葉で書き直せば、 $0x_1 + 0x_2 + 0x_3 + 0x_4 = 1$ となる。このような式を満たす \mathbf{x} は存在しないゆえ、したがって方程式 $B\mathbf{x} = \mathbf{b}$ の解は存在しない。

(ケース2)：次に $[B|\mathbf{b}]$ の最後の列 \mathbf{b} に主成分が含まれない場合を考える。例えば、

$$[B|\mathbf{b}] = \left[\begin{array}{ccccc|c} 0 & 1 & 3 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \text{ のとき, } B\mathbf{x} = \mathbf{b} \text{ は, } \begin{bmatrix} 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}.$$

このとき、主成分のない B の列に対応する変数 x_i を任意定数とおく と簡単に解が得られることをこれから見ていこう。

いま、ベクトル \mathbf{x} が方程式 $B\mathbf{x} = \mathbf{b}$ を満たしているとする。この例では主成分を含む B の列は第2列と第4列であるから、そうでない1,3,5列に対応する変数 x_1, x_3, x_5 について $c_1 := x_1, c_2 := x_3, c_3 := x_5$ とおこう。すると、残りの x_2 および x_4 は1行目と2行目に対応する方程式

$$(\#1) \begin{cases} 0x_1 + 1x_2 + 3x_3 + 0x_4 + 2x_5 = 4, \\ 0x_1 + 0x_2 + 0x_3 + 1x_4 + 1x_5 = 5, \end{cases}$$

を移項することで次を満たさねばならないことが分かる。

$$(\#2) \begin{cases} x_2 = 4 - 3x_3 - 2x_5 & = 4 - 3c_2 - 2c_3, \\ x_4 = 5 - x_5 & = 5 - c_3. \end{cases}$$

このように簡単に整理できる理由は簡約行列が特殊な形をしているためである。以上により、方程式 $B\mathbf{x} = \mathbf{b}$ の解 \mathbf{x} が存在するとすれば、ベクトル \mathbf{x} は次の形になっていることが分かった：

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} c_1 \\ 4 - 3c_2 - 2c_3 \\ c_2 \\ 5 - c_3 \\ c_3 \end{bmatrix}.$$

逆に、今度は実数 c_1, c_2, c_3 を勝手に選び、上の式によって \mathbf{x} を定めてみよう。このとき \mathbf{x} が方程式 $B\mathbf{x} = \mathbf{b}$ を満たすことは明らかである。何故なら、上でもって \mathbf{x} を定めるということは、式 $(\#2)$ が満たされるように定めたことを意味し、式 $(\#2)$ を移項すれば元々の連立方程式 $(\#1)$ が得られるからである。こ

これまでの議論により, 方程式 $B\mathbf{x} = \mathbf{b}$ の解は存在し, しかも解の形は次のもので出してくされていることが分かった:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} c_1 \\ 4 - 3c_2 - 2c_3 \\ c_2 \\ 5 - c_3 \\ c_3 \end{bmatrix}, \quad \text{ただし } c_1, c_2, c_3 \text{ は任意の定数である.}$$

【補足】 c_1, c_2, c_3 がいかなる実数であろうとも, 上の列ベクトルは方程式の解となる. このように, どのような数でもよいが定数であり変数ではない数のことを任意定数という.

また, 解の形を次のように整理しておくとも後々の議論において都合が良い.

$$\begin{bmatrix} c_1 \\ 4 - 3c_2 - 2c_3 \\ c_2 \\ 5 - c_3 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 0 \\ 5 \\ 0 \end{bmatrix} + c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

上のような解の表記が何を意味するのか説明するために, 次元を落として \mathbb{R}^3 のベクトルについて考えてみよう. まず二つの平行でないベクトル $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^3$ を取って固定しよう. 次に $\mathbf{y} = c\mathbf{a}_1 + d\mathbf{a}_2$ なる点 (ただし c, d は任意の数) のなす集合 H を考えると, H は \mathbb{R}^3 の中で原点を通る平面になる. 更に新たなベクトル $\mathbf{a}_0 \in \mathbb{R}^3$ を取り, $\mathbf{x} = \mathbf{a}_0 + c\mathbf{a}_1 + d\mathbf{a}_2$ なる点のなす集合 W を考えれば, これは平面 H をベクトル \mathbf{a}_0 方向に平行移動した平面である.

連立 1 次方程式の解全体の集合は常にこのような形をしている. それは, いまのような考察を \mathbb{R}^n において行えば理解できる. (m, n) -行列 A による方程式 $A\mathbf{x} = \mathbf{b}$ の解全体を W としよう. 本節で与えた連立 1 次方程式の解法によれば, 固定された n 次元列ベクトル $\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_k$ および任意定数 c_1, \dots, c_k を用いて, W の元は次のような形で表示される:

$$\mathbf{x} = \mathbf{a}_0 + c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_k\mathbf{a}_k. \quad (4.6.1)$$

W は, 次で表される点 \mathbf{y} のなす空間 H (これは \mathbb{R}^n の原点を含む k 次元の空間²とみなせる) を \mathbf{a}_0 方向へ平行移動した集合となる:

$$\mathbf{y} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \dots + c_k\mathbf{a}_k.$$

例 4.6.1. 上で解いた連立 1 次方程式 $\begin{bmatrix} 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$ の解全体の集合は, 次で表すことので

きるベクトル

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

たち全体のなす集合 H (これは \mathbb{R}^5 の原点を通る 3 次元の空間である) を,

$$\begin{bmatrix} 0 \\ 4 \\ 0 \\ 5 \\ 0 \end{bmatrix}$$

W に一致する.

² 「 k 次元の空間」の数学的な定義は第 IV 部で与える. ここでは \mathbb{R}^k と同様の広がりをもった \mathbb{R}^n の部分集合のこと, と説明を濁すしかない.

本章で与えた連立1次方程式の解法は、はきだし法 (row reduction) あるいはガウス・ジョルダンの消去法 (Gauss-Jordan elimination) と呼ばれる。ここで、この解法についてまとめておこう。

まとめ (はきだし法による連立1次方程式の解法)

連立1次方程式 $Ax = b$ の解法は次のとおりである。まず拡大係数行列 $[A|b]$ の簡約化 $[B|b']$ を求める。このとき、 $Ax = b$ の解全体と $Bx = b'$ の解全体は一致する。 $[B|b']$ の最後の列 b' に主成分がある場合は方程式の解は存在しない。そうでない場合は解は存在し、方程式 $Bx = b'$ の解は、 B における主成分を含まない列に対応する変数を任意定数とおくことで簡単に表せる。なお、 B のすべての列に主成分がある場合、任意定数の数は0で、方程式の解は一つである。また、方程式の解をベクトルの和に分解しておくことで、解全体の集合が見やすくなる。

連立1次方程式の解法を知ることの重要性は、応用の立場からは論ずるまでもない。一方で、線形代数の枠組みにおいては、簡約化が線形写像の像の次元を調べるための手段となる。これに同次形連立1次方程式³の解における任意定数の個数を合わせることで、線形写像の次元公式を得る。また、次元公式は、線形写像の第1同型定理を次元の立場から理解する上でかかせない。そして、第1同型定理は線形空間の枠組のみならず、群論や環論など様々な代数理論の上で認められる定理であり、抽象代数学における基本的な考え方となる。

さて、いま述べたことと重複するが、第III部以降において、行基本変形や連立1次方程式の解の表示を用いた議論が度々登場する。そのたびに方程式の解法についていちいち復習する余裕はないゆえ、本章で述べたことを自ら説明できるくらい咀嚼していることが、今後の学習において望まれる。

4.7 連立1次方程式の解の形と任意定数の個数について

本章において、連立1次方程式 $Ax = b$ の解法の一つ (はきだし法) を与えた。ここで、式 (4.6.1) の形で与えられる $Ax = b$ の解の表し方にどれくらいの種類があるか検討しよう。例えば、次の方程式

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & | & 1 \\ 2 & 2 & | & 2 \end{bmatrix} \xrightarrow{\text{簡約化}} \begin{bmatrix} 1 & 1 & | & 1 \\ 0 & 0 & | & 0 \end{bmatrix}$$

の解として、次の四つの形が挙げられる:

$$\begin{aligned} (1) \quad \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a \begin{bmatrix} -1 \\ 1 \end{bmatrix}, & (2) \quad \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \\ (3) \quad \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ -1 \end{bmatrix}, & (4) \quad \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} + d \begin{bmatrix} 1 \\ -1 \end{bmatrix} + e \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \end{aligned}$$

ただし、 a, b, c, d, e はそれぞれ任意定数である。(1) は本章で与えた解法による解、(2) は (1) において $a = 2b$ とした解、(3) は第1変数 x を任意定数とした解、(4) は任意定数を水増しした解である。これらの例から次のことが示唆されるであろう:

- (1) と (2) の関係から、解の形は無数にあることが分かる。
- (3) のように、本章の解法とは異なる変数を任意定数とする解もある。また、(2) のように、いずれの変数も任意定数としない解もあり得る。
- (4) は、解の表示として適当ではない。解を記述する際は、最低限必要な数だけの任意定数を用いることが望まれる。

³同次形の方程式は 6.3 節で扱う。

これらの考察から、次のような疑問を持つ読者もいることと思う。

- (i) 任意定数の個数に水増しがないかどうかをどうやって判定すればよいのか。特に、本章で与えた解法において任意定数の水増しはないか。
- (ii) 任意定数の水増しがない異なる二つの解の表現を与えたときに、二つの解の任意定数の個数は必ず一致するか。
- (iii) 任意定数の定め方には、どれくらいの種類が考えられるのか。

これらの疑問への解答は、後に学ぶ線形空間論を踏まえた上でなされる(22.2節)。ここで、未定義語の羅列になることを承知で上の解答への筋道を簡単に予告しておこう。線形独立性と呼ばれる概念を通して(i)の判定がなされ、とくに本章で与えた解法において任意定数の水増しはない。(ii)はその通りであり、これは線形空間の次元を通して説明される。また、(iii)は線形空間の基底の与え方と関係している。

ところで、(ii)を認めれば「解の任意定数の個数」なる概念が定義できる。しかし、その定義の妥当性の議論(すなわち(i)と(ii)への解答)は22.2節まで待たねばならない。それまでの間の当座の約束として、「連立1次方程式の解における任意定数の個数」とは本章で与えた解法(はきだし法)における任意定数の個数のことであると定めておく。

第5章 可逆行列

ここでは逆行列を持つ行列の性質について詳しく扱う。また、前章で学んだ行基本変形を行列の理論として再考する。これによって、行列の簡約化を用いた逆行列の導出法が理解される。

5.1 逆行列の性質

逆行列の定義を改めて書いておこう。

定義 5.1.1. n 次正方行列 A が可逆 (invertible) である (あるいは正則 (non-singular, regular) であるともいう) とは, $AB = E = BA$ を満たす n 次正方行列 B が存在することである。このとき, B を A の逆行列 (inverse) という。

線形写像の言葉に戻せば逆行列は, 逆写像 (逆関数) に対応する行列になる。逆写像の詳しい定義は 19.3 節にて, 線形写像の逆写像と逆行列との関係については命題 21.3.3(4) および系 26.2.4 にて述べる。

命題 5.1.2. A の逆行列が存在すれば, それは唯一つである。

Proof. B, C が共に A の逆行列であるとする, $B = C$ が簡単に確かめられる。実際, $B = BE = B(AC) = (BA)C = EC = C$ である。□

このように, 条件を満たすものが唯一つしかないことを示す場合, 条件を満たすものを二つ挙げ, それが一致することを言えばよい。ほかに, 条件を満たし互いに異なるものがあると仮定し, 矛盾を導くという手もある。

さて, 可逆行列 A の唯一つの逆行列を A^{-1} と表すことにしよう。今後の議論の中で A^{-1} という記号が出てきたときは, それは A が可逆行列であることを前提とした議論であることに注意せよ。

例 5.1.3. 逆行列に関する簡単な性質をここでまとめておく。

- (1) E の逆行列は E 自身に等しい。
- (2) A の逆行列を B とすれば, B の逆行列は A である。これは逆行列の定義より直ちに分かる。すなわち B も可逆行列であり, $(A^{-1})^{-1} = B^{-1} = A$ が成り立つ。
- (3) 各 X_1, X_2, \dots, X_k が共に可逆ならば, それらの積 $P = X_k X_{k-1} \cdots X_2 X_1$ も可逆であり, $P^{-1} = X_1^{-1} X_2^{-1} \cdots X_{k-1}^{-1} X_k^{-1}$ である。実際,

$$\begin{aligned}(X_1^{-1} X_2^{-1} \cdots X_{k-1}^{-1} X_k^{-1})P &= (X_1^{-1} X_2^{-1} \cdots X_{k-1}^{-1} X_k^{-1})(X_k X_{k-1} \cdots X_2 X_1) \\ &= X_1^{-1} X_2^{-1} \cdots X_{k-1}^{-1} (X_k^{-1} X_k) X_{k-1} \cdots X_2 X_1 \\ &= X_1^{-1} X_2^{-1} \cdots X_{k-1}^{-1} E X_{k-1} \cdots X_2 X_1 \\ &= X_1^{-1} X_2^{-1} \cdots (X_{k-1}^{-1} X_{k-1}) \cdots X_2 X_1 \\ &= \cdots = X_1^{-1} X_1 = E.\end{aligned}$$

$P(X_1^{-1} X_2^{-1} \cdots X_{k-1}^{-1} X_k^{-1}) = E$ も同様の計算で確かめられる。

- (4) $AB = O$ なる行列 $B \neq O$ が存在すれば、 A は可逆でない。何故なら、仮に可逆であるとする、 $AB = O$ の両辺に左から A^{-1} を掛けることで $B = O$ となり、これは $B \neq O$ に矛盾するからである。ここで、 B は正方行列でなくてもよいことに注意せよ。この議論は逆行列を持たない正方行列の存在も述べている。例えば、 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ とすれば $A^2 = O$ ゆえ A は可逆でない。
- (5) 可逆行列 A について、 $(A^{-1})^k$ のことを A^{-k} と書く。これは A^k の逆行列に等しい。また $A^0 := E$ と定めれば、整数 p, q について指数法則 $A^{p+q} = A^p A^q$ および $(A^p)^q = A^{pq}$ が成り立つ。これらは、実数の整数冪に関する指数法則の証明と同じようにして示される (証明略)¹。

よりみち (行列の割り算)

行列の和、差、積の定義を 2 章で与えた。一方で、行列演算に商 (割り算) の概念はない。しかしながら、かけ算の逆演算を割り算とみなす立場においては、行列の割り算は逆行列をかけることに相当する (これは、実数における割り算が逆数をかけることと同値であることからの類推による)。ここで注意しなければならないことは、任意の実数 $a \neq 0$ について a による割り算が定まるのに対して、任意の行列 $A \neq O$ について割り算が定まるわけではない点である。行列に関しては、 A が可逆行列でなければ A による割り算は定まらない。

さて、実数 $a \neq 0$ の逆数を $\frac{1}{a}$ あるいは a^{-1} と書くのに対して、可逆行列 A の逆行列を表す記号は A^{-1} のみであり、これを $\frac{1}{A}$ (あるいは $\frac{E}{A}$) と書く慣習はない。その理由の一つは次の通りである。仮に $\frac{1}{A}$ と表記した場合に、これと別の正方行列 B との積を考えてみよう：

$$\frac{1}{A} \cdot B, \quad B \cdot \frac{1}{A}.$$

このとき、上の二つの行列を $\frac{B}{A}$ と書く誘惑にかられる人が一定数いるであろうことが想像される。しかし、一般には上の二つの行列は異なり、これら異なる行列に同じ記号 $\frac{B}{A}$ を与えてしまえば、これ以降の計算は破綻してしまう。このような理由から、行列の分数表記は行わない。なお、実数においては $\frac{1}{a} \cdot b = b \cdot \frac{1}{a}$ が成り立つゆえ、分数表記は上手く機能する。

5.2 行基本変形再考

実は、行基本変形は、ある特別な行列を左からかけることにほかならない。これについて考えよう。

定義 5.2.1. 次の三種類の m 次正方行列 $S_m(i; r)$, $W_m(i, j)$, $K_m(i, j; r)$ を基本行列 (elementary matrix) という。なお、この記号は、本書でのみ通じる記号である。

- (1) $S_m(i; r)$: E_m の i 行目を r 倍した行列。ただし $r \neq 0$ とする。

$$S_m(i; r) := \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & r & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \leftarrow i \text{ 行目}$$

¹実数の冪に関する指数法則の証明は、例えば巻末の文献 [6] を見よ。

(2) $W_m(i, j)$: E_m の i 行と j 行を入れ替えた行列.

$$W_m(i, j) := \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & 1 & & & \\ & & & & \ddots & & \\ & & 1 & 0 & & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{array}{l} \leftarrow 1 \text{ 行目} \\ \vdots \\ \leftarrow i \text{ 行目} \\ \vdots \\ \leftarrow j \text{ 行目} \\ \vdots \\ \leftarrow n \text{ 行目} \end{array}$$

(3) $K_m(i, j; r)$: E_m において, i 行目の r 倍を j 行目に加えた行列.

$$K_m(i, j; r) := \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & & \ddots & & \\ & & r & 1 & & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{array}{l} \leftarrow 1 \text{ 行目} \\ \vdots \\ \leftarrow i \text{ 行目} \\ \vdots \\ \leftarrow j \text{ 行目} \\ \vdots \\ \leftarrow n \text{ 行目} \end{array}$$

基本行列を左から掛けることは, 行基本変形を行っていることに他ならない. すなわち次が成り立つ. 各自, 実際に計算して確かめてみる.

命題 5.2.2. (m, n) -行列 A と基本行列の積について, 次が成り立つ.

(1) $S_m(i; r)A$ は A の i 行目を r 倍した行列である:

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & r & & & & \\ & & & & \ddots & & \\ & & & & & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ ra_{i1} & ra_{i2} & \cdots & ra_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

(2) $W_m(i, j)A$ は A の i 行と j 行を入れ替えた行列である:

$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & 1 & & & \\ & & & & \ddots & & \\ & & 1 & 0 & & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

(3) $K_m(i, j; r)A$ は A の i 行目の r 倍を j 行目に加えた行列である:

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ \hline & & & & r & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \\ \hline ra_{i1} + a_{j1} & \cdots & ra_{in} + a_{jn} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

したがって、行基本変形によりある行列を別の行列に変形させることは、いくつかの基本行列を左から何度もかけることに他ならない。また、基本行列自身に基本変形をほどこすことを考えると、上の命題から次が直ちに分かる。

命題 5.2.3. 基本行列の逆行列は基本行列であり、

$$(1) S_m(i; r)^{-1} = S_m(i; r^{-1}), \quad (2) W_m(i, j)^{-1} = W_m(i, j), \quad (3) K_m(i, j; r)^{-1} = K_m(i, j; -r).$$

Proof. (1) のみ示そう。 $S_m(i; r)$ に左から $S_m(i; r^{-1})$ を掛けるということは、前命題 (1) より $S_m(i; r)$ の i 行目を r^{-1} 倍することに他ならない。ゆえに $S_m(i; r^{-1})S_m(i; r) = E_m$ である。また、 $S_m(i; r^{-1})$ に左から $S_m(i; r)$ を掛けることは $S_m(i; r^{-1})$ の i 行目を r 倍することに他ならず、 $S_m(i; r)S_m(i; r^{-1}) = E_m$ である。以上より、 $S_m(i; r)^{-1} = S_m(i; r^{-1})$ 。

(2) および (3) も同様の考察から分かる。 □

さて、行列 A が k 回の行基本変形によって B に変形できるとしよう。行基本変形は左から基本行列を掛けることに他ならないから、 A を行基本変形する際に実際に行った操作に対応する基本行列を X_1, \dots, X_k とすると、次が成り立つ:

$$X_k X_{k-1} \cdots X_2 X_1 A = B.$$

このとき、可逆行列たちの積 $P = X_k X_{k-1} \cdots X_2 X_1$ は可逆であり、また $PA = B$ と表せる。一方、 B に $X_k^{-1}, X_{k-1}^{-1}, \dots, X_1^{-1}$ に対応する基本変形を順次ほどこせば A を得る。実際、

$$X_1^{-1} X_2^{-1} \cdots X_{k-1}^{-1} X_k^{-1} B = P^{-1} PA = EA = A.$$

以上を整理すると次のような主張になる。

命題 5.2.4. (m, n) -行列 A が行基本変形により B に変形するならば、次が成り立つ。

(1) $PA = B$ を満たす m 次可逆行列 P が存在する。

【補足】 実は、逆に $PA = B$ を満たす可逆行列 P があるならば、 A を B に行基本変形できる (系 5.3.2)。

(2) B を行基本変形することにより A に戻すこともできる。

行基本変形と基本行列の関係から、可逆行列の逆行列を求めることができる。これを次節で見よう。

5.3 逆行列の求め方

n 次正方行列 A の逆行列を求めるために、まずはその候補として $BA = E$ を満たす正方行列 B を探そう。ここで、行列 A が k 回の行基本変形によって E に変形できたと仮定しよう。すなわち、 A の簡約化は E であり、また前節での考察により次のように書ける:

$$X_k X_{k-1} \cdots X_2 X_1 A = E,$$

ここで、各 X_i は A を E に行基本変形する際に実際に行った操作に対応する基本行列である。したがって、 $B = X_k X_{k-1} \cdots X_2 X_1$ とおけば $BA = E$ を満たすことが分かる。この B が A の逆行列であることは次のように示される:

Proof. B は基本行列の積で表されていること、および基本行列は可逆であること、可逆行列の積は可逆であることから B は可逆である。ゆえに B の逆行列 C が存在する。このとき、 $A = EA = (CB)A = C(BA) = CE = C$ より A は C に等しい。つまり $A = B^{-1}$ であり、この両辺の逆行列を取って $A^{-1} = (B^{-1})^{-1} = B$ を得る。□

さて、上の B を少ない労力で求めるには次の式を考えればよい。

$$B = X_k X_{k-1} \cdots X_2 X_1 E.$$

この式は、 A を行基本変形によって E に変形した操作と全く同じ手順で E を変形すると B が求まることを述べている。なお、 A の E への変形を確認した後で、同様の手順で E を変形するという二度手間は不要である。なぜなら A と E を横に並べた $(n, 2n)$ -行列 $[A|E]$ についての行基本変形を行い、左半分が単位行列となる $[E|X]$ が得られれば、

$$[E|X] = B[A|E] = [BA|BE] = [E|B].$$

つまり X は我々が求める B に他ならない。また、 $[E|X]$ は明らかに簡約行列であり、したがって $[A|E]$ の簡約化である。

逆行列の求め方

A を n 次正方行列とする。 A の逆行列を求めるには、 $(n, 2n)$ -行列 $[A|E_n]$ を簡約化すればよい。 $[A|E_n]$ の簡約化が $[E_n|B]$ なる形をしているならば、 B が A の逆行列となる。

ちなみに、 $[A|E_n]$ の簡約化が $[E_n|B]$ という形にならない場合、すなわち A の簡約化が E_n でない場合は A は可逆ではない。その理由は次章で述べる定理 6.2.2 による。これを認めれば、任意の可逆行列の簡約化は単位行列になることが分かり、したがって次を得る。

定理 5.3.1. 可逆行列は基本行列の積で表せる。

Proof. B を可逆行列とし、 $A := B^{-1}$ とする。可逆行列 A を行基本変形で単位行列に変形する手順に対応する基本行列たちの積は、これまでの議論により A の逆行列、すなわち B に一致することが分かっている。以上より B は基本行列の積で表すことができる。□

系 5.3.2. A を (m, n) -行列とすれば次は同値である：

- (1) A を B に行基本変形できる。
- (2) $PA = B$ を満たす m 次可逆行列 P が存在する。

Proof. (1)⇒(2) は命題 5.2.4(1) による。逆に (2) を仮定すれば前定理により P は基本行列の積で表せる。このことは (1) を意味している。□

本節では、 $BA = E$ を満たす正方行列 B の探し方の一例を挙げて、更に B が A の逆行列となることを見た。では、本節とは別の方法で $DA = E$ を満たす正方行列 D が得られたとき、この D は必ず A の逆行列になるのだろうか。次の定理を認めれば D も A の逆行列であり、 $B = D$ となることが分かる。この定理は行列式の項目に入ってから証明する (詳細は 13 章を見よ)。

定理 5.3.3. 二つの n 次正方行列 A, D について $DA = E$ が成り立てば D は A の逆行列である。すなわち、 $AD = E$ も成り立つ。

5.4 基本行列と列基本変形 (発展)

本章では, 基本行列を左から掛けることと行基本変形の各操作が対応することを見た. では, 基本行列を右から掛けると何が起こるのであろうか. それは列に関する変形に対応するというのが答えである. 次の命題は, 命題 5.2.2 と同様に, 計算によって確かめることができる.

命題 5.4.1. (m, n) -行列 A と基本行列の積について, 次が成り立つ.

- (1) $AS_n(i; r)$ は A の i 列目を r 倍した行列である.
- (2) $AW_n(i, j)$ は A の i 列と j 列を入れ替えた行列である.
- (3) $AK_n(i, j; r)$ は A の i 列目の r 倍を j 列目に加えた行列である.

第6章 行列の階数

連立1次方程式の解における任意定数の個数と関係する量として、行列の階数と呼ばれる概念がある。実は、階数とは線形写像の像の次元として本来は定義されるものである(例 22.1.3(2))。しかしながら、ここでは行列の言葉に翻訳したうえでの定義を述べなければならず、そのためには簡約化の一意性について言及する必要がある。

階数を用いると連立1次方程式の解に関する言明を簡潔に述べることができる(命題 6.2.1)。しかし、だからといって、階数を用いて述べられた命題を丸暗記しても理解が深まることはない。階数という便利な言葉に頼らずに、連立1次方程式の解法がどんな手順であったか常に頭の中で意識しつつ命題を解釈してもらいたい。

6.1 簡約化の一意性

定理 6.1.1. 行列の簡約化は唯一通りに定まる。

Proof. 列の数に関する帰納法で示す。まず、列の数が1の場合、すなわち列ベクトル \boldsymbol{x} の簡約化について考える。列ベクトルのうちで簡約なもの、その定義から1行成分が1で他の成分がすべて0の列ベクトル \boldsymbol{e}_1 か $\mathbf{0}$ に限る。 $\boldsymbol{x} = \mathbf{0}$ の場合、 $\mathbf{0}$ に基本変形をいくらほどこしても変化せず、 $\mathbf{0}$ の簡約化は $\mathbf{0}$ 自身以外にありえない。また、 $\boldsymbol{x} \neq \mathbf{0}$ の場合、 \boldsymbol{x} を $\mathbf{0}$ に行基本変形することはできない。何故なら、もし $\mathbf{0}$ が \boldsymbol{x} の簡約化であるならば、命題 5.2.4 により $\mathbf{0}$ を行基本変形して $\boldsymbol{x} \neq \mathbf{0}$ が得られるが、先の議論によりこれは不可能だからである。ゆえに \boldsymbol{x} の簡約化は \boldsymbol{e}_1 のみである。以上より、列ベクトルの簡約化は唯一つであることが分かった。

列の数が n なる行列について簡約化が一意的であると仮定し、列の数が $n+1$ の行列についてもそうであることを示そう。 A を (m, n) -行列、 \boldsymbol{a} を m 次列ベクトルとし、 $(m, n+1)$ -行列 $[A|\boldsymbol{a}]$ の簡約化について考える。 $[B|\boldsymbol{b}]$ および $[C|\boldsymbol{c}]$ を $[A|\boldsymbol{a}]$ の簡約化としよう。このとき、 (m, n) -行列 B, C は共に簡約行列であり、とくに A の簡約化であるから、帰納法の仮定により $B = C$ である。 B の中にどの行の主成分も含まない列がある場合とそうでない場合に分けて考えよう。

まずは、 B の第 j 列が主成分を含まない場合である。各 A, B, C から第 j 列を除いた $(m, n-1)$ -行列をそれぞれ A', B', C' とする。このとき、 (m, n) -行列 $[B'|\boldsymbol{b}']$ は簡約行列であり、これはとくに行列 $[A'|\boldsymbol{a}']$ の簡約化である。 $B = C$ ゆえ同様のことが C についても成り立ち、 $[C'|\boldsymbol{c}']$ も $[A'|\boldsymbol{a}']$ の簡約化となる。したがって帰納法の仮定より、 $[B'|\boldsymbol{b}'] = [C'|\boldsymbol{c}']$ である。すなわち、 $\boldsymbol{b} = \boldsymbol{c}$ ゆえ $[B|\boldsymbol{b}] = [C|\boldsymbol{c}]$ 。

次に、 B のどの列にも主成分が含まれている場合を考える。すなわち、 $[B|\boldsymbol{b}]$ および $[C|\boldsymbol{c}]$ が次の形をしている場合である：

$$\left[\begin{array}{ccc|c} 1 & & & * \\ & \ddots & & * \\ & & 1 & * \\ & & & * \end{array} \right] = \left[\begin{array}{cc|c} E_n & & * \\ O_{m-n, n} & & * \end{array} \right].$$

更に次の二つに場合分けをして考える。

- (i) $\boldsymbol{b}, \boldsymbol{c}$ がともに、ある主成分を含む場合。

この場合は簡約行列の定義から、 $\boldsymbol{b}, \boldsymbol{c}$ は $n+1$ 行目が1でそれ以外の成分が0の列ベクトルであり、 $\boldsymbol{b} = \boldsymbol{c}$ を得る。

(ii) (i) でない場合. すなわち, \mathbf{b}, \mathbf{c} のうち少なくともいずれか一方が主成分を含まない場合.

仮に \mathbf{b} が主成分を含まないとして話を進める. このとき, $[B|\mathbf{b}]$ および $[C|\mathbf{c}]$ は次の形をしている.

$$[B|\mathbf{b}] = \left[\begin{array}{c|c} E_n & \mathbf{b}' \\ \hline O_{m-n,n} & O_{m-n,1} \end{array} \right], \quad [C|\mathbf{c}] = \left[\begin{array}{c|c} E_n & \mathbf{c}'_1 \\ \hline O_{m-n,n} & \mathbf{c}'_2 \end{array} \right].$$

$[B|\mathbf{b}], [C|\mathbf{c}]$ はともに $[A|\mathbf{a}]$ に行基本変形を繰り返して得られる行列である. ゆえに $[B|\mathbf{b}]$ に行基本変形を繰り返すことで $[C|\mathbf{c}]$ を得ることができる. したがって, ある m 次可逆行列 X によって $X[B|\mathbf{b}] = [C|\mathbf{c}]$ となる (命題 5.2.4). X を次のように分割する:

$$X = \left[\begin{array}{cc} P_{n,n} & Q_{n,m-n} \\ R_{m-n,n} & S_{m-n,m-n} \end{array} \right] = \left[\begin{array}{cc} P & Q \\ R & S \end{array} \right],$$

ここで, 上式の真ん中の行列に現れる P, Q, R, S の添え字は行列のサイズを意味し, 煩雑ゆえ以降は省略する. $X[B|\mathbf{b}]$ を計算すると次のようになる.

$$X[B|\mathbf{b}] = \left[\begin{array}{cc} P & Q \\ R & S \end{array} \right] \left[\begin{array}{c|c} E_n & \mathbf{b}' \\ \hline O_{m-n,n} & O_{m-n,1} \end{array} \right] = \left[\begin{array}{cc} P & P\mathbf{b}' \\ R & R\mathbf{b}' \end{array} \right].$$

$[C|\mathbf{c}] = X[B|\mathbf{b}]$ について成分を比較しよう.

$$\left[\begin{array}{c|c} E_n & \mathbf{c}'_1 \\ \hline O_{m-n,n} & \mathbf{c}'_2 \end{array} \right] = [C|\mathbf{c}] = X[B|\mathbf{b}] = \left[\begin{array}{cc} P & P\mathbf{b}' \\ R & R\mathbf{b}' \end{array} \right].$$

上式より $P = E_n, R = O_{m-n,n}$ が得られる. ゆえに $\mathbf{c}'_1 = P\mathbf{b}' = E_n\mathbf{b}' = \mathbf{b}', \mathbf{c}'_2 = R\mathbf{b}' = O_{m-n,n}\mathbf{b}' = O_{m-n,1}$ である. つまり, $\mathbf{b} = \mathbf{c}$ が示された.

\mathbf{c} が主成分を含まない場合は, \mathbf{b} と \mathbf{c} の役割を入れ替えていまと同様の議論を行えばよい.

以上により, いずれの場合においても $[B|\mathbf{b}] = [C|\mathbf{c}]$ となる. すなわち, $(m, n+1)$ -行列 A の簡約化は唯一通りに定まる. □

定義 6.1.2. A の簡約化を B とする. 次で定める三つの数はすべて同じ値となり, これを A の階数 (**rank**) とよび $\text{rank } A$ と書く.

- (1) B の零ベクトルでない行の数,
- (2) B の主成分の個数,
- (3) B の主成分を含む列の数.

上で定める数のうち, 一般の行列においても (1) と (2) は等しい. これらが (3) と等しいのは, 簡約行列 B において主成分を二つ以上含む列は存在しないからである.

仮に A の簡約化が二通りあるとし, それらの主成分の個数が異なっていたとすれば A の階数を定めようがない. また, これは連立 1 次方程式において任意定数の個数が異なる二通りの解の表示が存在することも意味する. このようなことが起こり得ないという主張が簡約化の一意性にほかならない.

(m, n) -行列 A について次は明らかである.

$$\text{rank } A \leq \min\{m, n\}. \tag{6.1.1}$$

ここで \min は最小値を表す記号である. すなわち, 実数を要素とする集合 X に対して, X の中で最も小さい数が存在するとき, これを X の最小値 (**minimum**) とよび $\min X$ と書く. 同様に X の最大値 (**maximum**) も定められ, これを $\max X$ と書く.

命題 6.1.3. A を行基本変形することで C が得られるならば, A と C の簡約化は一致する. したがって, $\text{rank } A = \text{rank } C$ である.

Proof. A の簡約化を B とすれば, 行基本変形により A を B に変形できる. 一方, 仮定より A を C に行基本変形できることから, 命題 5.2.4 により C を A に行基本変形できる. したがって次のような基本変形ができる:

$$C \xrightarrow{\text{行基本変形}} A \xrightarrow{\text{行基本変形}} B.$$

ゆえに簡約行列 B は C の簡約化である. □

6.2 連立 1 次方程式と階数

連立 1 次方程式 $Ax = b$ の解の形は, 拡大係数行列 $[A|b]$ の簡約化の形によって, 解が存在する場合とそうでない場合, および解が存在する場合における任意定数の個数が決まるのであった. これを階数の言葉を用いて言いなおすと次の命題になる.

命題 6.2.1. A を (m, n) -行列とする. 連立 1 次方程式 $Ax = b$ において次が成り立つ.

- (1) $Ax = b$ の解が存在するための必要十分条件は $\text{rank}[A|b] = \text{rank } A$ である.
- (2) $Ax = b$ の解が存在するとき, 解の任意定数の個数は $n - \text{rank } A$ である.
- (3) $Ax = b$ の解が唯一であるための必要十分条件は, $\text{rank}[A|b] = \text{rank } A = n$ である.

とくに, A が正方行列の場合は逆行列との関係を含めて次の主張を得る.

定理 6.2.2. n 次正方行列 A において次は同値である.

- (1) $\text{rank } A = n$, (2) A の簡約化は E である,
- (3) 任意の n 次列ベクトル b について方程式 $Ax = b$ は唯一の解を持つ,
- (4) 方程式 $Ax = 0$ は唯一の解 0 を持つ, (5) A は可逆である.

Proof. (1) から (4) までの同値性は既に述べたことのまとめにほかならない. しかしながら確認のために復習を兼ねて証明しておこう.

(1) \Rightarrow (2): $\text{rank } A = n$ とすれば, A の簡約化 B の主成分の個数は n である. よって, $n \times n$ 簡約行列 B は各行に主成分を持ち, ゆえに $B = E$ である.

(2) \Rightarrow (3): $[A|b]$ の簡約化を $[E|b']$ とする. 命題 4.4.1 より方程式 $Ax = b$ の解全体と $Ex = b'$ の解全体は等しい. 方程式 $Ex = b'$ が唯一の解 $x = b'$ を持つことは明らかであり, したがって $Ax = b$ の解も $x = b'$ 唯一である. (別証明: $\text{rank } A = \text{rank } E = n$ に注意する. $\text{rank}[E|b'] = n$ ゆえ $\text{rank}[A|b] = \text{rank}[E|b'] = n = \text{rank } A$. よって, 命題 6.2.1(3) より方程式は唯一の解を持つ.)

(3) \Rightarrow (4): (4) は (3) における $b = 0$ という特別の場合ゆえ明らか.

(4) \Rightarrow (1): 仮定より, 解の任意定数の個数は 0 である. したがって, 命題 6.2.1(2) より $0 = n - \text{rank } A$. ゆえに $\text{rank } A = n$.

(2) \Rightarrow (5): 5.3 節で論じた逆行列の求め方より得る.

(5) \Rightarrow (4): これは命題 4.3.1 の特別な場合に相当する. すなわち, $x = 0$ が方程式 $Ax = 0$ の解であることは明らかであり, 逆に x が $Ax = 0$ を満たすならば, この両辺に左から A^{-1} を掛けることで $x = 0$ を得る.

以上によって, すべての条件の同値性が示された. □

6.3 同次形の方程式

(m, n) -行列 A において, $A\mathbf{x} = \mathbf{0}$ なる方程式を同次形あるいは斉次形の方程式という¹. 同次形の方程式は $\mathbf{x} = \mathbf{0}$ を解に持つ. つまり, 任意の同次形の方程式に解が存在する.

練習 6.3.1. 同次形の連立 1 次方程式について命題 6.2.1(1) における階数の条件が成立していることを確かめよ.

解答例: A の簡約化を B とする. 方程式 $A\mathbf{x} = \mathbf{0}$ の拡大係数行列 $[A|\mathbf{0}]$ の簡約化は $[B|\mathbf{0}]$ であるから $\text{rank}[A|\mathbf{0}] = \text{rank}[B|\mathbf{0}] = \text{rank } B = \text{rank } A$. 確かに $\text{rank}[A|\mathbf{0}] = \text{rank } A$ は成り立っている.

4.6 節で述べた連立 1 次方程式の解法を同次形の場合に適用してみよう. 命題 6.2.1(2) より, 解における任意定数の個数は $k = n - \text{rank } A$ である. 上の練習にもあるように拡大係数行列 $[A|\mathbf{0}]$ の簡約化は $[B|\mathbf{0}]$ の形になり, ゆえに式 (4.6.1) において $\mathbf{a}_0 = \mathbf{0}$ となる. したがって解の表示は

$$\mathbf{x} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_k\mathbf{a}_k \quad (\text{ただし } c_1, \dots, c_k \text{ は任意定数})$$

である. すなわち, 解全体の集合は, \mathbb{R}^n の原点を含む k 次元の空間となる.

斉次形連立 1 次方程式に関する拡大係数行列の最後の列は $\mathbf{0}$ であり, したがって行基本変形を行ってもこの部分は常に $\mathbf{0}$ となる. ゆえに, この行列を簡約化する際は係数行列に限りて簡約化してもよい.

例題 6.3.2. 次の連立 1 次方程式を解け (係数行列の簡約化は例 4.5.3 に記した).

$$\begin{bmatrix} 2 & 0 & 1 & 5 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

解答例: 係数行列を簡約化すると

$$\begin{bmatrix} 2 & 0 & 1 & 5 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{簡約化}} \begin{bmatrix} 1 & 0 & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

したがって, 主成分のない 2 列と 4 列に対応する変数を任意定数として, それぞれ α, β とすれば, 求める解は

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2}\beta \\ \alpha \\ -2\beta \\ \beta \\ 0 \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -\frac{3}{2} \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}.$$

なお, 解の表示に分数が現れるのを嫌う場合は, $\frac{1}{2}\beta$ を改めて γ とおくことにより (つまり $\beta = 2\gamma$),

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} -3 \\ 0 \\ -4 \\ 2 \\ 0 \end{bmatrix} \quad \text{としてもよい.}$$

¹同次および斉次はともに homogeneous の訳語である.

6.4 重ね合わせの原理

同次形の方程式とそうでない方程式の解の間には次のような関係がある。

命題 6.4.1. (m, n) -行列 A および m 次列ベクトル \mathbf{b} が与えられているとし、次の二つの連立 1 次方程式を考える

$$(I) \quad A\mathbf{x} = \mathbf{b}, \quad (II) \quad A\mathbf{x} = \mathbf{0}.$$

さらに方程式 (II) の解 \mathbf{a} を一つ取って固定しよう。このとき次が成り立つ。

- (1) 方程式 (II) の任意の解 \mathbf{z} に対し、 $\mathbf{a} + \mathbf{z}$ は方程式 (I) の解である。
- (2) 方程式 (I) の任意の解 \mathbf{y} に応じて、方程式 (II) の解の一つである \mathbf{z} を上手く取ると、 $\mathbf{y} = \mathbf{a} + \mathbf{z}$ と表せる。

Proof. (1) : \mathbf{z} を (II) の解とすると、 $A(\mathbf{a} + \mathbf{z}) = A\mathbf{a} + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}$. ゆえに、 $\mathbf{a} + \mathbf{z}$ は方程式 (I) の解である。

(2) : \mathbf{y} を (I) の解とする。ここで $\mathbf{z} := \mathbf{y} - \mathbf{a}$ とおこう。このとき \mathbf{z} は (II) の解である。何故なら、 $A\mathbf{z} = A(\mathbf{y} - \mathbf{a}) = A\mathbf{y} - A\mathbf{a} = \mathbf{b} - \mathbf{b} = \mathbf{0}$. また、 \mathbf{z} の定め方から $\mathbf{y} = \mathbf{a} + \mathbf{z}$ であり、我々は主張を得た。□

上の命題は、同次形の方程式の解全体の集合を H とすれば、 $A\mathbf{x} = \mathbf{b}$ の解全体の集合 W が H を \mathbf{a} 方向に平行移動した集合に一致することを言っている。この事実の特別な場合については 4.6 節にて説明していた。また、方程式を解く労力の観点からは次のように捉えることもできる：

(※) 方程式 (II) の解法は、方程式 (I) の解法より幾分か易しい。そこで、あらかじめ易しい方程式 (II) の解を求めておき、更に、何らかの方法で (I) の解 \mathbf{a} を一つでよいから見つけてくる (適当な数を代入して山勘で見つけたとしてもよい)。すると、(I) の解は、(II) の解全体を \mathbf{a} 方向へ平行移動することですべて得られる。

しかし、既に連立 1 次方程式の解法を知っている我々にとっては、このような考え方は不要にも思える。そこで、微分積分学の文脈で現れる上と類似する現象を紹介しよう：

命題 6.4.2 (重ね合わせの原理). α, β, γ を既知の実数、 $b(x)$ を既知の関数とし、2 階微分可能な未知の関数 $f(x)$ に関する次の二つの微分方程式²を考える。

$$(I) \quad \alpha f''(x) + \beta f'(x) + \gamma f(x) = b(x), \quad (II) \quad \alpha f''(x) + \beta f'(x) + \gamma f(x) = 0.$$

さらに方程式 (I) の解 $a(x)$ を一つ取って固定しよう。このとき次が成り立つ。

- (1) 微分方程式 (II) の任意の解 $z(x)$ に対し、 $a(x) + z(x)$ は微分方程式 (I) の解である。
- (2) 微分方程式 (I) の任意の解 $y(x)$ に応じて、微分方程式 (II) の解の一つである $z(x)$ を上手く取ると、 $y(x) = a(x) + z(x)$ と表せる。

Proof. (1) : $z(x)$ を (II) の解とすると、

$$\begin{aligned} & \alpha(a(x) + z(x))'' + \beta(a(x) + z(x))' + \gamma(a(x) + z(x)) \\ &= \left(\alpha a''(x) + \beta a'(x) + \gamma a(x) \right) + \left(\alpha z''(x) + \beta z'(x) + \gamma z(x) \right) \\ &= b(x) + 0 = b(x). \end{aligned}$$

ゆえに、 $a(x) + z(x)$ は微分方程式 (I) の解である。

²この形の微分方程式は、線形常微分方程式と呼ばれている。

(2) : $y(x)$ を (I) の解とする. ここで $z(x) := y(x) - a(x)$ とおこう. このとき $z(x)$ は (II) の解である. 何故なら,

$$\begin{aligned} & \alpha z''(x) + \beta z'(x) + \gamma z(x) \\ &= \alpha(y(x) - a(x))'' + \beta(y(x) - a(x))' + \gamma(y(x) - a(x)) \\ &= \left(\alpha y''(x) + \beta y'(x) + \gamma y(x) \right) - \left(\alpha a''(x) + \beta a'(x) + \gamma a(x) \right) \\ &= b(x) - b(x) = 0. \end{aligned}$$

また, $z(x)$ の定め方から $y(x) = a(x) + z(x)$ であり, 我々は主張を得た. \square

(※) で述べた戦略は, 上の微分方程式 (I) を解く際にも有効である. このように, 線形代数と微分積分において類似する数理現象が観察されることが分かった. 実は高校数学においても, これらに類似する現象を学んでいることに読者はお気づきであろうか:

命題 6.4.3. r, b を既知の実数とし, 次の二つの漸化式を考える.

$$(I) \quad x_{n+1} = rx_n + b, \quad (II) \quad x_{n+1} = rx_n.$$

さらに漸化式 (I) を満たす数列 a_n を一つ取って固定しよう. このとき次が成り立つ.

- (1) 漸化式 (II) を満たす任意の数列 z_n に対し, 数列 $a_n + z_n$ は漸化式 (I) を満たす.
- (2) 漸化式 (I) を満たす任意の数列 y_n に応じて, 漸化式 (II) を満たす数列 z_n を上手く取ると, $y_n = a_n + z_n$ と表せる.

Proof. (1) : z_n が漸化式 (II) を満たす数列であるとき,

$$a_{n+1} + z_{n+1} = (ra_n + b) + rz_n = r(a_n + z_n) + b.$$

ゆえに, $a_n + z_n$ は漸化式 (I) を満たす.

(2) : y_n が漸化式 (I) を満たす数列であるとする. ここで $z_n := y_n - a_n$ とおこう. このとき数列 z_n は漸化式 (II) を満たす. 何故なら,

$$z_{n+1} = y_{n+1} - a_{n+1} = (ry_n + b) - (ra_n + b) = r(y_n - a_n) = rz_n.$$

また, z_n の定め方から $y_n = a_n + z_n$ であり, 我々は主張を得た. \square

上の三つの命題は, これらの主張自体のみならず証明までもが並行している点に着目せよ. これは次に挙げる操作たち:

- (1) ベクトル \mathbf{x} に対して $A\mathbf{x}$ を対応させる操作,
- (2) 関数 $f(x)$ に対して関数 $\alpha f''(x) + \beta f'(x) + \gamma f(x)$ を対応させる操作,
- (3) (i) 数列 (x_1, x_2, x_3, \dots) に対して数列 (x_2, x_3, x_4, \dots) を対応させる操作, および
(ii) 数列 (x_1, x_2, x_3, \dots) に対して数列 $(rx_1, rx_2, rx_3, \dots)$ を対応させる操作,

が共に線形性を満たしていることに起因する. 本節のように, 類似する証明を連立1次方程式, 微分方程式, 漸化式の各対象ごとに繰り返す行うことについて, 聡明な読者は二度あるいは三度手間のよう感じたことであろう. そこでこのような手間を省き, 上に挙げた三つの命題を同時に議論できるような枠組みとして, 我々は線形空間と呼ばれる代数構造を提案することになる. そして上に挙げた三つの命題は, 命題 23.1.4 によって統一的に証明されることを予告しておこう.

命題 6.4.3 にもとづいた一般項の導出について解説する.

例題 6.4.4. $y_n = 3y_n + 8$, $y_1 = 1$ を満たす数列 y_n の一般項を求めよ.

解答例. 次の二つの漸化式を考える:

$$(I) \ x_{n+1} = 3x_n + 8, \quad (II) \ x_{n+1} = 3x_n.$$

まず漸化式 (I) を満たす数列の一つ, どうにかして見つけたい. そこで (I) を満たす定数列がないか探してみよう. ここで定数列 (x, x, x, \dots) が (I) を満たすならば, $x = 3x + 8$ である. これを解くと $x = -4$ であり, 実際に数列 $(-4, -4, -4, \dots)$ は (I) を満たす. この数列を $a_n = -4$ とおこう. さて, y_n が漸化式 (I) を満たすことから, 命題 6.4.3(2) より, 漸化式 (II) を満たす数列 z_n を用いて

$$y_n = z_n + a_n$$

と書くことができる.

【補足】 ここで, 上式が成り立つように上手く z_n を取る具体的な方法が分からないとしても, それは問わなくてよい (気になる読者は, どうして問わなくてもよいのか考えよ). ちなみに, 今回の場合は命題 6.4.3(2) の証明を読めば, $z_n := y_n - a_n$ とすればよいことが分かる.

ここで z_n は等比数列ゆえ, その初項を c とすれば, $z_n = c3^{n-1}$ である. したがって,

$$y_n = c3^{n-1} + (-4)$$

であり, この式において $n = 1$ の場合を考えれば, $1 = c - 4$ ゆえ $c = 5$. 以上より, 一般項は $y_n = 5 \cdot 3^{n-1} - 4$ である.

上の議論のポイントは, (I) を満たす数列の一つが方程式 $x = 3x + 8$ を解くことにより簡単に見つかること, そして (I) よりも簡単な (II) を満たす数列は, 等比数列ゆえ一般項が容易に決定できることの二点である.

第II部
行列式

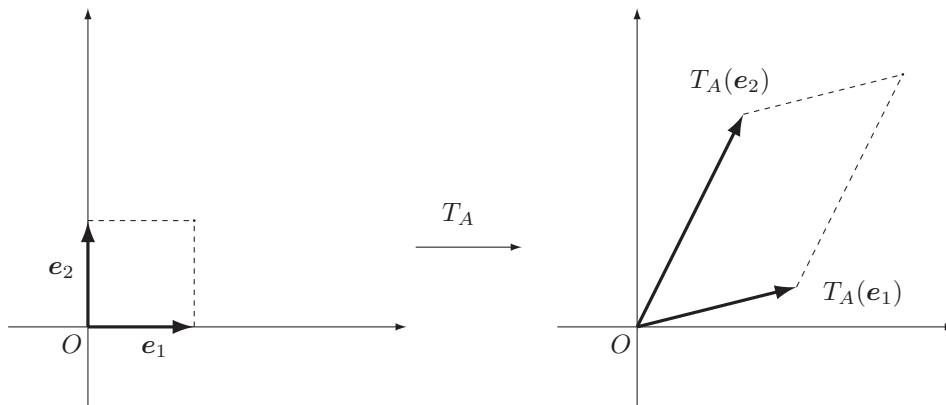
第7章 行列式とは何か

行列式とは、正方行列 A に対して定められるある量のことであり、 $\det A$ もしくは $|A|$ と書く。その定義は一言で述べるには難しく、厳密な (あるいは形式的な) 定義は後で改めて論じるとして、本章では行列式の持つ意味、および微分積分学における扱われ方について導入的な紹介をする。

行列式への数学的意味の与え方は大きく分けて二通りある。したがって、行列式の定義の仕方にも二通りの立場があると考えてよい。一つは歴史的な経緯である方程式論からの見方で、もう一つは幾何学的な観点によるものである。後者のほうがイメージが描きやすいゆえ、まずそちらから解説しよう。

7.1 \mathbb{R}^n 上の線形変換の体積拡大率

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ とし、写像 $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ を $T_A(\mathbf{x}) := A\mathbf{x}$ で定める (21.3 節以降において、この写像を線形変換と呼ぶ)。このとき、 $\det A$ とは、 \mathbb{R}^2 における 1 辺の長さが 1 の単位正方形の面積が T_A で写像されると何倍になるかを表す量である。すなわち、行列式とは \mathbb{R}^2 上の線形変換の面積拡大率である。



$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad T_A(\mathbf{e}_1) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}, \quad T_A(\mathbf{e}_2) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}.$$

単位正方形とは $\mathbf{e}_1, \mathbf{e}_2$ で張られる平行四辺形のことであり、これを T_A で写像すると $T_A(\mathbf{e}_1)$ と $T_A(\mathbf{e}_2)$ で張られる平行四辺形になる。ここで、積分において負の面積を考えたように、平行四辺形についても向きを込めた面積を導入する。面積の符号は、半周未満の回転によってベクトル $T_A(\mathbf{e}_1)$ を $T_A(\mathbf{e}_2)$ に重ねるとき、その回転が反時計回りになるか時計回りになるかで判断すればよい。 $\mathbf{e}_1, \mathbf{e}_2$ について同じことを考えればこの回転は反時計回りであるから、 $T_A(\mathbf{e}_1), T_A(\mathbf{e}_2)$ についても反時計回りならば正、そうでなければ負の面積を対応させる。図 7.1 の平行四辺形の面積は、原点と $T_A(\mathbf{e}_1) + T_A(\mathbf{e}_2)$ を頂点に持つ長方形の面積からいくつかの台形や三角形の面積を引くことで得られるから、

$$\begin{aligned} \begin{vmatrix} a & b \\ c & d \end{vmatrix} &= \text{“}T_A(\mathbf{e}_1) \text{ と } T_A(\mathbf{e}_2) \text{ で張られる平行四辺形の面積”} \\ &= (a+b)(c+d) - \left(\frac{ac}{2} + \frac{(c+c+d)b}{2} + \frac{bd}{2} + \frac{(b+a+b)c}{2} \right) \\ &= ac + ad + bc + bd - \frac{ac + 2cb + db + bd + 2bc + ac}{2} = ad - bc. \end{aligned}$$

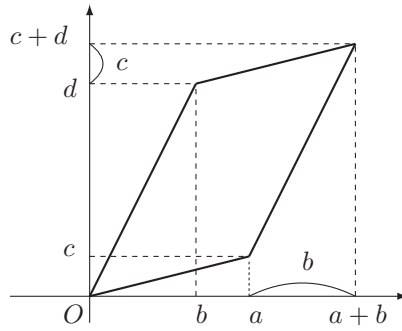


図 7.1: 平行四辺形の各頂点の座標表示

一般の n 次正方行列 A の行列式も同様に与えられる。すなわち、線形変換 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ を $T_A(\mathbf{x}) := A\mathbf{x}$ で定め、 T_A の体積拡大率を $\det A$ と定める。ここで、1次元において長さ、2次元において面積、3次元において体積とそれぞれ異なる呼ばれ方をしていた各次元に関する量は、4次元以上においてはすべて体積と呼ぶことにする。また n 次元空間において、次元が n 未満の図形を面と呼ぶことにしよう。3次元空間における2次元の図形が通常の意味での面であった。 $n = 3$ の場合 $\det A$ は、 A の成分表示に現れる3つの列ベクトルで張られる平行六面体の体積に相当し、図形の体積計算が得意な者ならば次式を得ることができるかもしれない¹:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31}. \quad (7.1.1)$$

体積拡大率を直接求める方法で4次以上の行列式を定めるのは難しい。 \mathbb{R}^n において、 n 個のベクトルで張られる図形 (これは $n - 1$ 次元の $2n$ 個の面で囲まれる) の体積の計算式はかなり複雑であり、また、普段3次元の空間に住む我々が高次元の図形を想像すること自体に困難な部分がある。そこで、 n 次元の図形の体積が持つべき性質をいくつか列挙することにより、その性質をもとに低い次元の図形の体積計算に帰着させることで、求める n 次元体積を得るという方法が考えられる。この具体的方法については11章で述べることにしよう。

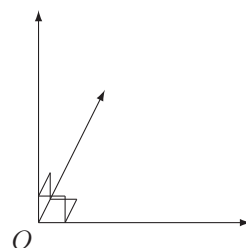
さて、行列式が体積拡大率を意味するのであれば、 T_E は元を動かさない恒等写像ゆえ $|E| = 1$ である。また、 $|AB|$ は合成写像 $T_A \circ T_B$ の体積拡大率であるから、これは $|A| \cdot |B|$ に等しい。つまり $|AB| = |A||B|$ となる。更に A が可逆であるとき、 $|A||A^{-1}| = |AA^{-1}| = |E| = 1$ より $|A^{-1}| = |A|^{-1}$ 、とくに $|A| \neq 0$ である。逆に、 $|A| \neq 0$ ならば A が可逆であることも後に示され、したがって $|A|$ は A が可逆であるかどうかを知るための指標となる。

¹式 (7.1.1) の右辺が平行六面体の体積に相当することは11章で説明する。

よりみち(多次元空間を見る)

n 本のベクトルで張られる n 次元の図形の体積計算を考えるうえで最も基本的な指針となるのはカヴァリエリの原理(切り口の面積が常に等しい2つの立体の体積は等しい…詳しくは10.2節で述べる)である。これにより、角柱や円柱など柱状の図形の体積公式を得る。すなわち、 n 次元柱の n 次元体積は、底面となる $n-1$ 次元図形の面積($n-1$ 次元体積)に高さを掛けたものに等しい。これは、特別な形の n 次行列式は $n-1$ 次行列式の計算に帰着できることを意味している。更に、カヴァリエリの原理を順次適用することで計算できる n 次行列式の種類が増えていき、最終的にはすべての行列式の値が定められる。実は、行列式の実際の計算演習においても、これと同等のことを繰り返すことになる。

ところで、3次元の空間に住む我々が n 次元の世界を想像する何か良い方法はないだろうか。右は、原点 O において3本の直線が互いに垂直に交わる様子を図示したものである。ただし、これは2次元の平面に描かれた模式的なものであり、正確な図ではない。しかしながらこのことは、3次元空間において模式的に4次元を見る方法があることを示唆している。4次元空間は座標軸が4つあるから、4本の直線が O で互いに垂直に交わることができる。この図においてそのような直線を1本加えるにはどうすればよいか考えてみよう。



物理学における第4の次元とは時間のことであった。時間の流れを記録したものの例として、我々は音楽や動画などを知っている。とくに動画には空間と時間の両方が記録されており、4次元を理解するには最良の例である。映画のフィルムを1枚ずつ重ねて束にしてみよう。これはフィルムと並行な方向に空間(2次元のフィルムに射影された3次元空間)が広がっており、束の重なる方向が時間を表している。このことから4次元の世界とは、フィルムのように薄っぺらくなった3次元空間たちの束を重ねた空間であると想像できる。したがって、先の図における第4の方向とは、束が重なって高くなる方向、すなわち、この紙面に垂直に鉛筆を立てた方向ということになる。

それでは、5本の直線が互いに垂直に交わる5次元の図を描くにはどうすればよいだろうか。それは先程の鉛筆を立てた状態、すなわち4本の直線が垂直に交わった状態を写真に撮って印刷し、4次元空間を薄っぺらい空間とみなし、写真が印刷された紙に垂直な方向に新たな鉛筆を立てればよい。これを順次繰り返し、我々は多次元空間の模式図を得る。

7.2 クラメル公式

歴史的には、行列式とは連立1次方程式の解の公式を与えるための道具として生み出されたのであった。そこで、今度はこの方針をたどってみよう。方程式の解が唯一であることは A が可逆であることと同値であり(定理6.2.2)、更に $|A| \neq 0$ と同値になることは先に述べた通りである。

簡単のため、次の2変数連立1次方程式が唯一解を持つ場合を考える:

$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

いま、 A は可逆としているから $|A| = ad - bc \neq 0$, つまり $a \neq 0$ または $c \neq 0$ である。 $a \neq 0$ の場合について拡大係数行列を簡約化してみよう。 $a \neq 0$ ゆえ a で割り算ができることに注意すると、

$$\begin{aligned} \left[\begin{array}{cc|c} a & b & y_1 \\ c & d & y_2 \end{array} \right] &\rightarrow \left[\begin{array}{cc|c} 1 & \frac{b}{a} & \frac{y_1}{a} \\ c & d & y_2 \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{b}{a} & \frac{y_1}{a} \\ 0 & d - \frac{bc}{a} & y_2 - \frac{y_1c}{a} \end{array} \right] = \left[\begin{array}{cc|c} 1 & \frac{b}{a} & \frac{y_1}{a} \\ 0 & \frac{ad-bc}{a} & \frac{ay_2 - y_1c}{a} \end{array} \right] \\ &= \left[\begin{array}{cc|c} 1 & \frac{b}{a} & \frac{y_1}{a} \\ 0 & \frac{|A|}{a} & \frac{ay_2 - y_1c}{a} \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{b}{a} & \frac{y_1}{a} \\ 0 & 1 & \frac{ay_2 - y_1c}{|A|} \end{array} \right] \rightarrow \left[\begin{array}{cc|c} 1 & 0 & \frac{y_1}{a} - \frac{ay_2 - y_1c}{|A|} \frac{b}{a} \\ 0 & 1 & \frac{ay_2 - y_1c}{|A|} \end{array} \right]. \end{aligned}$$

ここで、上式の最後の (1, 3)-成分は

$$\begin{aligned} \frac{y_1}{a} - \frac{ay_2 - y_1c}{|A|} \cdot \frac{b}{a} &= \frac{y_1|A| - (ay_2 - y_1c)b}{|A|a} = \frac{y_1(ad - bc) - (ay_2 - y_1c)b}{|A|a} \\ &= \frac{y_1ad - y_1bc - ay_2b + y_1cb}{|A|a} = \frac{y_1ad - ay_2b}{|A|a} = \frac{y_1d - y_2b}{|A|}. \end{aligned}$$

以上より、行列式の言葉で方程式の解を表すと次のようになる:

$$x_1 = \frac{\begin{vmatrix} y_1 & b \\ y_2 & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a & y_1 \\ c & y_2 \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

また、 $c \neq 0$ の場合も同様の行基本変形により上の解を得る (各自確かめよ). 上の解の公式をクラメルの公式という. 3変数の場合は次で与えられる:

定理 7.2.1 (クラメルの公式). 3次正方行列を $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ とおく. $|A| \neq 0$ ならば連立1次

方程式 $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ の解は唯一であり、次で表される:

$$x_1 = \frac{\begin{vmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_3 & a_{32} & a_{33} \end{vmatrix}}{|A|}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & y_1 & a_{13} \\ a_{21} & y_2 & a_{23} \\ a_{31} & y_3 & a_{33} \end{vmatrix}}{|A|}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & y_1 \\ a_{21} & a_{22} & y_2 \\ a_{31} & a_{32} & y_3 \end{vmatrix}}{|A|}.$$

もちろん4次以降についても同様の主張が成り立つ (証明は13章で述べる). このように、解の唯一性の判別式として、そして解の公式を一言で述べるための関数として行列式は導入された. ただし、その定義の複雑さから、手計算で解を求める際の公式としてはあまり適さない.

7.3 微積分学における行列式

本章の最後に、多変数の微積分学において比較的早い段階で学習する行列式の使用例について紹介する. 簡単のため2変数の場合に限って述べるが、いずれも一般の n 変数関数の議論にまで拡張されるものである. 証明を含めたそれらの詳細は、微積分の教科書に譲ろう.

- ヘッセ行列式

C^2 級²関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ の極値判定において行列式が現れる. 点 (a, b) における1階の偏導関数がすべて0だからといって $f(a, b)$ が極値を取るとは限らず、極値かどうかを導くには2階微分の情報を調べる必要がある (これは1変数関数のときもそうであった). 例えば2変数関数の場合には、次で定めるヘッセ行列 (**Hessian matrix**) の行列式 (これをヘッセ行列式 (**Hessian**) という):

$$H_f(a, b) := \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(a, b) & \frac{\partial^2 f}{\partial x \partial y}(a, b) \\ \frac{\partial^2 f}{\partial y \partial x}(a, b) & \frac{\partial^2 f}{\partial y^2}(a, b) \end{bmatrix}, \quad |H_f(a, b)| = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2$$

の値を用いて判定する. C^2 級ゆえ $f_{xy} = f_{yx}$ に注意せよ. 3変数以上の場合は行列式の値のみでは判定できず、ヘッセ行列が定める2次形式の挙動を詳しく見る必要があり、そこでは行列の固有値³を用いて判断する (37.6節を見よ). ヘッセ行列式は、その性質から判別式 (discriminant) とも

²すべての2階偏導関数が連続であるような関数を C^2 級または2階連続微分可能であるという.

³固有値の定義は28章で与える.

呼ばれている。

- ヤコビ行列式

面積拡大率の意味において最も重要な例は積分の変数変換である。 D および K を \mathbb{R}^2 における長方形で囲まれた集合とし、 C^1 級関数による 1 対 1 写像

$$T : K \rightarrow D, \quad T(u, v) = (\phi(u, v), \psi(u, v))$$

が与えられているとする。このとき、 D 上の 2 変数関数 $f : D \rightarrow \mathbb{R}$ の重積分、すなわち xy -平面 ($z = 0$ 平面) 上の長方形 D と曲面 $z = f(x, y)$ で挟まれた図形の体積

$$\iint_D f(x, y) \, dx dy$$

を変数変換によって K 上の関数 $f(T(u, v)) = f(\phi(u, v), \psi(u, v))$ に関する積分として表したい。 K が D に写像される際の局所的な面積拡大率は各座標に関する偏微分によって次のように計算される:

$$J_T(u, v) := \begin{bmatrix} \frac{d\phi}{du} & \frac{d\phi}{dv} \\ \frac{d\psi}{du} & \frac{d\psi}{dv} \end{bmatrix}, \quad |J_T(u, v)| = \frac{d\phi}{du} \frac{d\psi}{dv} - \frac{d\phi}{dv} \frac{d\psi}{du}.$$

この $J_T(u, v)$ をヤコビ行列 (**Jacobian matrix**) と呼び、 $|J_T(u, v)|$ をヤコビ行列式 (**Jacobian**) という。変数変換による面積の拡大分を掛けることで次の置換積分公式を得る:

$$\iint_D f(x, y) \, dx dy = \iint_K f(\phi(u, v), \psi(u, v)) |J(u, v)| \, du dv.$$

なお、上の公式は D, K がもっと一般の集合の場合 (例えば縦線形集合など) においても成り立つ。

よりみち (行列と行列式)

「行列」は matrix の訳語であり「行列式」は determinant のそれである。このように外国語とそれらの和訳でニュアンスが異なるのには歴史的な事情がある。もともと行列と行列式は別の目的をもって定められたものであった。歴史的には行列式のほうが先に生まれた。それは連立 1 次方程式の解の公式 (クラメル公式) を得るために考え出されたものである。一方、matrix なる語句は行列式の理論の中で使われ始めたのちに、数を矩形に並べた概念の総称として用いられるようになった。その後、線形写像の数値化に相当するものとして演算が定められ、今日の行列の定義に至っている。これら二語の和訳も紆余曲折があったが、これらの関係性が十分に理解されたことにより、行列・行列式なる訳が定着した。

第8章 置換

行列式の定義への道は長く険しい。ここでは行列式の定義に必要な置換について述べる。置換とはその名の通り置き換え、あるいは入れ替え方を意味する。線形代数においては行列式の定義以外に置換が現れることは稀であるものの、置換は数学を記述する言葉として重要である。とくに、群を説明する道具として初等的役割を担う。

置換に関する話題は組み合わせ論的な色彩が強く、苦手意識を持つ読者も多いように思う。しかしながら本書において置換は、行列式の定義とその性質を調べる際に用いられるのみであり、それ以降の行列式の実際の計算においては、組み合わせ論的な素養の多くを必要とするわけではない。苦手意識を克服できずとも、今後の線形代数の学習にあまり支障はないだろう。

8.1 置換の定義

置換とは、有限集合上の写像のうち、ある条件を満たすもののことをいう。ここでいう有限集合の構成要素は何でも構わない。本書では記述の簡素化のために、 n を自然数とし、集合

$$X_n := \{1, 2, \dots, n\}$$

における置換について考察する。

定義 8.1.1. X_n から X_n 自身への写像 $\sigma: X_n \rightarrow X_n$ において、 $\sigma(1), \dots, \sigma(n)$ の中に重複がないとき σ を X_n 上の置換 (permutation) という。

X_n 上の置換 σ において、文字列 $\sigma(1), \sigma(2), \dots, \sigma(n)$ は 1 から n までの文字が重複なくすべて並んでいる。したがって、 σ は 1 対 1 写像である。

例 8.1.2. 次で定める写像 $\sigma: X_3 \rightarrow X_3$ のうち (1), (2) は置換であり、(3) は置換でない。

$$(1) \begin{cases} 1 \mapsto \sigma(1) = 2 \\ 2 \mapsto \sigma(2) = 3 \\ 3 \mapsto \sigma(3) = 1 \end{cases}, \quad (2) \begin{cases} 1 \mapsto \sigma(1) = 1 \\ 2 \mapsto \sigma(2) = 3 \\ 3 \mapsto \sigma(3) = 2 \end{cases}, \quad (3) \begin{cases} 1 \mapsto \sigma(1) = 3 \\ 2 \mapsto \sigma(2) = 3 \\ 3 \mapsto \sigma(3) = 2 \end{cases}.$$

(3) が置換でないのは、 $\sigma(1)$ と $\sigma(2)$ が等しく、したがって $\sigma(1), \sigma(2), \sigma(3)$ の中に重複が見られるからである。

8.2 置換の表示

$\sigma(1) = k_1, \sigma(2) = k_2, \dots, \sigma(n) = k_n$ なる置換 $\sigma: X_n \rightarrow X_n$ を

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$$

と表す。ここで、右辺の 2 行目 k_1, k_2, \dots, k_n は文字列 $1, 2, \dots, n$ を並び変えたものである。置換の表記は行列と区別がつかないため、文脈でどちらを考えているのか留意すること。

例 8.2.1. $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$ とする. σ は写像 $\sigma: X_4 \rightarrow X_4$ であるから, $X_4 = \{1, 2, 3, 4\}$ の各元を代入できる. その値は, $\sigma(1) = 3, \sigma(2) = 1, \sigma(3) = 4, \sigma(4) = 2$ である. 紙面に余裕があるならば, これらを次のように書いてもよい:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} (1) = 3, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} (2) = 1, \\ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} (3) = 4, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} (4) = 2.$$

置換の表示を扱いやすくするために, 次の表示の仕方を許すとする.

- 置換の表し方は上下の対応のみが本質的であり, 列の並び順はあまり重要ではない. そこで, 列を入れ替えた書き方を許すとする. これによって, 後で定める逆置換の定義が簡明になる.
- σ で動かない元 ($\sigma(i) = i$ なる i のこと) に対応する列は省略してもよいとする. これは後で定める巡回置換の表記を見やすくするための措置である.

以上二つの表示の仕方を認めると, 次で表示された置換はすべて同じ写像を表す:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 1 & 3 \\ 2 & 1 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 4 \\ 3 & 4 & 1 \end{pmatrix}.$$

一方で, 上述の記法を許したことにより, 次の二つの置換 $\sigma: X_4 \rightarrow X_4$ および $\tau: X_3 \rightarrow X_3$ の区別がつかなくなったことに注意せよ:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

この不都合を逆に利用して, $m > n$ のとき, 置換 $\sigma: X_n \rightarrow X_n$ は X_m 上の置換でもあると考えることにしよう. すなわち各 $i = n+1, n+2, \dots, m$ について $\sigma(i) = i$ で定められる置換 $\sigma: X_m \rightarrow X_m$ でもあると見なす.

8.3 置換の積

二つの置換 $\sigma, \tau: X_n \rightarrow X_n$ に対して, これらの合成写像 $\sigma \circ \tau: X_n \rightarrow X_n$ もまた置換となる. これを置換どうしの積演算とみなし, $\sigma\tau$ と書く. すなわち, $\sigma\tau := \sigma \circ \tau$.

例 8.3.1. $\sigma = \begin{pmatrix} 1 & 3 & 4 \\ 4 & 1 & 3 \end{pmatrix}, \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$ のとき,

$$\begin{aligned} \sigma\tau(1) &= \sigma(\tau(1)) = \sigma(2) = 2, & \sigma\tau(2) &= \sigma(3) = 1, \\ \sigma\tau(3) &= \sigma(4) = 3, & \sigma\tau(4) &= \sigma(1) = 4, \end{aligned}$$

となるから, $\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ である. なお, σ の上段が τ の下段と同じ並びにな

るようあらかじめ並び変えておくと, 合成の表示が直ちに得られる. つまり $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix} =$

$\begin{pmatrix} 2 & 3 & 4 & 1 \\ 2 & 1 & 3 & 4 \end{pmatrix}$ であり, τ と σ を縦に並べると

$$\begin{aligned} \tau &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \\ \sigma &= \begin{pmatrix} 2 & 3 & 4 & 1 \\ 2 & 1 & 3 & 4 \end{pmatrix}. \end{aligned}$$

ここで真ん中の二段を消して, $\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$ を得る.

例 8.3.2. $\sigma = \begin{pmatrix} 2 & 4 & 9 \\ 4 & 9 & 2 \end{pmatrix}$, $\tau = \begin{pmatrix} 3 & 6 & 7 & 8 \\ 8 & 3 & 6 & 7 \end{pmatrix}$ のとき $\sigma\tau = \tau\sigma$ である (各自, 1 から 9 までの数を $\sigma\tau$ および $\tau\sigma$ にそれぞれ代入し, 結果が等しくなることを確認せよ). このように, 二つの置換 σ, τ の表示において, 同じ文字が一つも重複しないとき $\sigma\tau = \tau\sigma$ が成り立つ.

命題 3.2.5 の特別な場合として次が成り立つ:

事実 8.3.3. 置換の積に関して結合律が成立する. すなわち, 任意の三つの置換 $\sigma, \tau, \nu : X_n \rightarrow X_n$ について $\sigma(\tau\nu) = (\sigma\tau)\nu$ が成り立つ. そこで, これらの積を略して $\sigma\tau\nu$ と書く.

次の二つの概念は, 本来は写像に関する一般論において定められるものである. ここでは置換の積演算と関連して, 次の呼称を与える (一般の写像については 19 章をみよ).

定義 8.3.4. • どの元も動かさない X_n 上の恒等写像 $\begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$ を恒等置換 (**identity**) と呼び, これを id_{X_n} で表す. また, 定義域 X_n に誤解がない場合はこれを略して id と書く.

• 置換 $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$ の逆写像 $\begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ 1 & 2 & \cdots & n \end{pmatrix}$ を σ の逆置換 (**inverse**) と呼び σ^{-1} で表す.

逆置換は, 置換の表示において上下の行を入れ替えることによって得られる. ゆえに $(\sigma^{-1})^{-1} = \sigma$ である. また, $\sigma\sigma^{-1} = \text{id} = \sigma^{-1}\sigma$ もすぐに分かる.

同じ置換を繰り返し合成するときは冪を用いる:

- 自然数 n について, 置換 σ の n 回の合成 $\sigma \circ \cdots \circ \sigma$ を σ^n と書く. また $\sigma^1 = \sigma$ とする.
- 逆置換の n 回の合成 $(\sigma^{-1})^n$ を σ^{-n} と書く. これは σ^n の逆置換に等しい.
- 置換 σ に対して $\sigma^0 := \text{id}$ と約束する. この記号を導入したことにより, 各整数 m, n について指数法則 $\sigma^m \circ \sigma^n = \sigma^{m+n}$ が成り立つ. とくに, $\sigma^{-1} \circ \sigma^m = \sigma^{m-1}$ である (この事実は命題 8.4.4 の証明で用いる).

上で述べた指数法則の証明は実数の整数冪の場合 (あるいは可逆行列の整数冪の場合) と同様であるから, ここでは略す.

練習 8.3.5. 置換 σ および自然数 m に対して $\sigma^{-1} \circ \sigma^m = \sigma^{m-1}$ を示せ.

解答例. まず $\sigma^m = \sigma \circ \sigma^{m-1}$ であることを示そう. これは $m \geq 2$ の場合は明らかであり, $m = 1$ の場合も $\sigma^m = \sigma^1 = \sigma = \sigma \circ \text{id} = \sigma \circ \sigma^0 = \sigma \circ \sigma^{m-1}$ と確認することができる. ゆえに $\sigma^{-1} \circ \sigma^m = \sigma^{-1} \circ (\sigma \circ \sigma^{m-1}) = (\sigma^{-1} \circ \sigma) \circ \sigma^{m-1} = \text{id} \circ \sigma^{m-1} = \sigma^{m-1}$ となり, 求める等式を得る. \square

8.4 巡回置換とその表示

定義 8.4.1. X_n の元のうち k_1, k_2, \dots, k_r 以外の元は動かさず, $\sigma(k_1) = k_2, \sigma(k_2) = k_3, \dots, \sigma(k_{r-1}) = k_r, \sigma(k_r) = k_1$ と順にずらす置換 $\begin{pmatrix} k_1 & k_2 & \cdots & k_{r-1} & k_r \\ k_2 & k_3 & \cdots & k_r & k_1 \end{pmatrix}$ を巡回置換 (**cyclic permutation**) といい, これを省略して (k_1, k_2, \dots, k_r) と書く. また, 二つの文字からなる巡回置換 (i, j) を互換 (**transposition**) という.

巡回置換を省略して書くのは、同じ文字を何度も書く手間を省くためである。いま、置換の表示の仕方がいくつも提示され、ここで錯綜する読者も多いように思う。様々な表記法があっても、基本は最初に述べた表示に戻って考えるようにすれば誤解は少なくなるであろう。

例 8.4.2. • 次の巡回置換はすべて同じ置換を意味する:

$$(2, 5, 3) = (5, 3, 2) = (3, 2, 5) = \begin{pmatrix} 2 & 3 & 5 \\ 5 & 2 & 3 \end{pmatrix}.$$

• 上の置換に各元を代入した値は次の通りである:

$$\begin{aligned} (2, 5, 3)(1) &= \begin{pmatrix} 2 & 3 & 5 \\ 5 & 2 & 3 \end{pmatrix}(1) = 1, & (2, 5, 3)(2) &= 5, & (2, 5, 3)(3) &= 2, \\ (2, 5, 3)(4) &= 4, & (2, 5, 3)(5) &= 3, & (2, 5, 3)(6) &= 6. \end{aligned}$$

• $(k_1, k_2, \dots, k_r)^{-1} = (k_r, k_{r-1}, \dots, k_2, k_1)$. 特に $(i, j)^{-1} = (j, i) = (i, j)$.

練習 8.4.3. $\sigma : X_5 \rightarrow X_5$ を $\sigma := (3, 2, 4)(2, 3, 5)$ で定める. このとき, σ を $\begin{pmatrix} 1 & \cdots & n \\ k_1 & \cdots & k_n \end{pmatrix}$ の形で表せ. また, この σ は巡回置換であるか.

答え. X_5 の各元を代入して確認する.

$$\begin{aligned} \sigma(1) &= (3, 2, 4)(2, 3, 5)(1) = (3, 2, 4)(1) = 1, & \sigma(2) &= (3, 2, 4)(2, 3, 5)(2) = (3, 2, 4)(3) = 2, \\ \sigma(3) &= (3, 2, 4)(2, 3, 5)(3) = (3, 2, 4)(5) = 5, & \sigma(4) &= (3, 2, 4)(2, 3, 5)(4) = (3, 2, 4)(4) = 3, \\ \sigma(5) &= (3, 2, 4)(2, 3, 5)(5) = (3, 2, 4)(2) = 4. \end{aligned}$$

したがって, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 3 & 4 \end{pmatrix} = (3, 5, 4)$. ゆえに σ は巡回置換である.

複雑な事象をより単純なものに分解して考えることは、分析における基本的手段の一つである。これを置換の場合にも適用し、任意の置換をより単純な置換に分解する方法を考える。

命題 8.4.4. 任意の置換は互いに文字を共有しない巡回置換の積に分解される。

Proof. $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 6 & 2 & 7 & 5 & 3 \end{pmatrix}$ において確かめよう¹. まず 1 に σ をほどこし続けるとどのように移り変わるか、すなわち数列 $a_m := \sigma^m(1)$ がどう動くかを見る. すると, $1 \mapsto 4 \mapsto 2 \mapsto 1 \mapsto 4 \mapsto 2 \mapsto \dots$ と繰り返される. どうして繰り返されるのか説明しよう. X_n は有限集合ゆえ数列 a_m はどこかで重複する項が現れる. つまり $\sigma^k(1) = \sigma^\ell(1) = i$ なる二つの自然数 $k > \ell$ が見つかる. このとき $\sigma^{k-\ell}(1) = 1$ であることが次のように示される. 置換 σ^ℓ に X_n の元である 1 および $p = \sigma^{k-\ell}(1)$ をそれぞれ代入すると,

$$\sigma^\ell(1) = i, \quad \sigma^\ell(p) = \sigma^\ell(\sigma^{k-\ell}(1)) = \sigma^k(1) = i.$$

つまり $\sigma^\ell(p) = \sigma^\ell(1)$ であり, σ^ℓ が置換であることから $\sigma^\ell(1), \dots, \sigma^\ell(n)$ の中に重複はないゆえ $p = 1$, すなわち $\sigma^{k-\ell}(1) = 1$ でなければならない. いま, 数列 a_m において少なくとも $k - \ell$ 項目までに 1 が現れることが分かった. そして, a_m ($m \geq 1$) に初めて 1 が現れた項より先について, a_m は巡回し続ける.

さて, 1 に σ を繰り返すほどこしたときの動きは巡回置換 $(1, 4, 2)$ の働き方と同じである. 次に 1, 4, 2 に現れなかった文字, 例えば 3 について同様に移動の仕方を見てみると $3 \mapsto 6 \mapsto 5 \mapsto 7 \mapsto 3 \mapsto \dots$ を得る.

¹本来ならば一般の置換に対して証明すべきことである. しかし, 議論があまりに抽象的過ぎて読者の理解が得られなければ意味がない. そこで, 一般の置換に対する証明が再構成できるような議論を想定しつつ, ここでは特別な置換を例に挙げて論じた.

これは巡回置換 $(3, 6, 5, 7)$ と同じ働きをしている. このとき, 二つの集合 $A = \{1, 4, 2\}$ と $B = \{3, 6, 5, 7\}$ の間には重複する元は一つもない. これは何故だろうか. その理由を詳しく考察しよう.

A の元は $a_m = \sigma^m(1)$ なる元の集まりであり, B の元は $b_m = \sigma^m(3)$ なる元の集まりである. 仮に A, B の両方に含まれる元 x があるとすれば, $x = a_i = b_j$ と書ける. このとき $x = \sigma^j(3)$ より $3 = \sigma^{-j}(x)$ である. 一方, 各 a_m ($m \in \mathbb{N}$) について $\sigma^{-1}(a_m) \in A$ が成り立つ. 実際, $\sigma^{-1}(a_m) = \sigma^{-1}(\sigma^m(1)) = \sigma^{m-1}(1) = a_{m-1} \in A$ である (ここで $a_0 := \sigma^0(1) = 1$ とする). 特に, a_i に σ^{-1} を j 回ほどこすことで $\sigma^{-j}(a_i) \in A$ を得る. ところが, $3 = \sigma^{-j}(x) = \sigma^{-j}(a_i) \in A$ となってしまう, これは A に現れない元として 3 を取ってきたことに反する. 以上より, A, B に重複する元はない.

まだ X_n の文字が出つくしていない場合は, そのような文字の移動の仕方を順次見ていく. すると, X_n の各元が σ による動き方で重複なく分類されることが分かる. いまの例では $\{1, 4, 2\}$ と $\{3, 6, 5, 7\}$ の二つに分類される. このとき, $\sigma = (1, 4, 2)(3, 6, 5, 7)$ と分解されることはすぐに分かる. \square

命題 8.4.5. 巡回置換について, $(k_1, k_2, \dots, k_r) = (k_1, k_r)(k_1, k_{r-1}) \cdots (k_1, k_2)$.

上の命題の右辺は, 文字列 $k_1 k_2 \cdots k_r$ を文字列 $k_2 \cdots k_r k_1$ に入れ替える次の操作を意味している:

$$k_1 k_2 \cdots k_r \xrightarrow{(k_1, k_2)} k_2 k_1 k_3 \cdots k_r \xrightarrow{(k_1, k_3)} \cdots \xrightarrow{(k_1, k_{r-1})} k_2 \cdots k_{r-1} k_1 k_r \xrightarrow{(k_1, k_r)} k_2 \cdots k_r k_1.$$

ここで, 記号 $\xrightarrow{(k_i, k_j)}$ は文字 k_i と文字 k_j の入れ替えを表す.

命題 8.4.4 および 8.4.5 より直ちに次を得る:

系 8.4.6. 任意の置換は互換の積に分解される.

練習 8.4.7. $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 6 & 2 & 7 & 5 & 3 \end{pmatrix}$ を互換の積に分解せよ.

答え.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 6 & 2 & 7 & 5 & 3 \end{pmatrix} = (1, 4, 2)(3, 6, 5, 7) = (1, 2)(1, 4)(3, 7)(3, 5)(3, 6).$$

8.5 置換の符号

定義 8.5.1. 置換 σ が m 個の互換の積で表されるとき, $\text{sgn}(\sigma) := (-1)^m$ と定め, これを σ の符号 (**signature** または **sign**) という. $\text{sgn}(\sigma) = 1$ なる置換 σ を **偶置換 (even permutation)** と呼び, $\text{sgn}(\sigma) = -1$ なる置換 σ を **奇置換 (odd permutation)** と呼ぶ. なお, 恒等置換 id は偶置換であると約束する.

置換の偶奇性のことを **パリティ (parity)** とも呼ぶこともある.

練習 8.5.2. $n \geq 2$ ならば id_{X_n} が偶数個の互換に分解できることを示せ.

解答例: $\text{id}_{X_n} = (1, 2)(1, 2)$.

次の例 (練習 8.4.3 と同じ置換) から分かるように, 置換の互換の積への分解は一意的ではない.

$$(3, 2, 4)(2, 3, 5) = (3, 4)(3, 2)(2, 5)(2, 3) = (3, 4)(3, 5).$$

したがって, σ が奇数個の互換の積で表せ, 更に偶数個の互換の積でも表せるとなると, σ の符号を定めることはできない. しかしながら, このようなことは起こらず, 置換を互換の積で表した際の互換の個数の偶奇は必ず一致することが知られている. この事実の証明にはいくつか道具が必要となるゆえ, 次章にまわそう.

命題 8.5.3. X_n 上の置換 σ, τ について $\text{sgn}(\sigma\tau) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau)$.

Proof. σ が k 個, τ が ℓ 個の互換の積で表されたとすれば, $\sigma\tau$ は $k + \ell$ 個の互換の積で書ける. ゆえに $\text{sgn}(\sigma\tau) = (-1)^{k+\ell} = (-1)^k (-1)^\ell = \text{sgn}(\sigma) \cdot \text{sgn}(\tau)$. \square

X_n 上の置換全体のなす集合を \mathfrak{S}_n (あるいは S_n) と書き, これを n 次対称群 (symmetric group of degree n) という. \mathfrak{S}_n の元のうち偶置換のみをすべて集めた集合を \mathfrak{A}_n (あるいは A_n) と書き, これを n 次交代群 (alternating group of degree n) という. \mathfrak{S}_n の元の個数はいくつだろうか. 置換の表示を $\begin{pmatrix} 1 & \cdots & n \\ k_1 & \cdots & k_n \end{pmatrix}$ の形のみに限定すると, 置換の総数は n 個の文字による列 $k_1 k_2 \cdots k_n$ の総数に一致する. ゆえに \mathfrak{S}_n の元の総数は $n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$ 個である. 例えば \mathfrak{S}_4 の元の個数は $4! = 24$ 個である. $n \geq 2$ について, A_n の個数は \mathfrak{S}_n の元の個数のちょうど半分だけある. この事実は後で示す (命題 12.2.2).

例 8.5.4.

$$\begin{aligned} \mathfrak{S}_3 &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\} \\ &= \{ \text{id}, (2,3), (1,2), (1,2,3) = (1,3)(1,2), (1,3,2) = (1,2)(1,3), (1,3) \}. \\ A_3 &= \{ \text{id}, (1,3)(1,2), (1,2)(1,3) \}. \end{aligned}$$

練習 8.5.5. 次を示せ.

- (1) 置換 σ, τ について, $\tau\sigma = \text{id} \iff \tau = \sigma^{-1}$.

解答例: (\Rightarrow): $\tau\sigma = \text{id}$ とする. この両辺に右から σ^{-1} を合成すると

$$\begin{aligned} \tau(\sigma\sigma^{-1}) &= \text{id}\sigma^{-1} \\ \tau\text{id} &= \sigma^{-1} \\ \tau &= \sigma^{-1}. \end{aligned}$$

(\Leftarrow): 逆置換の定義より明らかである. □

- (2) 置換 $\sigma_1, \dots, \sigma_m$ について, $(\sigma_1\sigma_2\cdots\sigma_{m-1}\sigma_m)^{-1} = \sigma_m^{-1}\sigma_{m-1}^{-1}\cdots\sigma_2^{-1}\sigma_1^{-1}$.

解答例: $\sigma := \sigma_1\sigma_2\cdots\sigma_{m-1}\sigma_m$ および $\tau :=$ (右辺) とおくと, $\tau\sigma = \text{id}$ である (これは例 5.1.3(3) と類似する議論による). よって, (1) より $\tau = \sigma^{-1}$ を得る. □

- (3) 置換 σ が ℓ 個の互換の積で書けるならば, σ^{-1} も ℓ 個の互換の積で書ける.

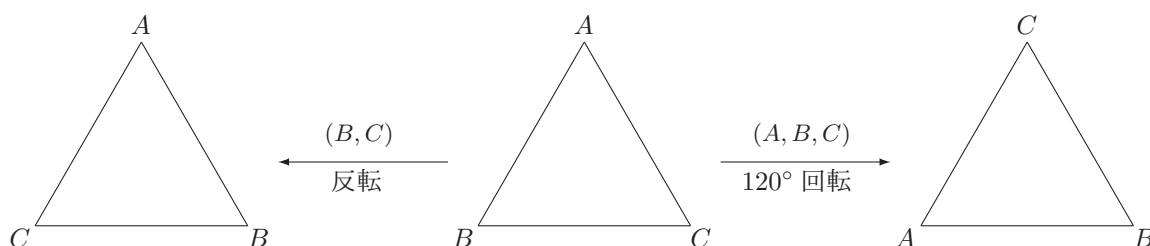
解答例: σ が ℓ 個の互換 $\sigma = (p_1, q_1)(p_2, q_2)\cdots(p_\ell, q_\ell)$ に分解されているとする. (2) より, $\sigma^{-1} = (p_\ell, q_\ell)^{-1}\cdots(p_2, q_2)^{-1}(p_1, q_1)^{-1} = (p_\ell, q_\ell)\cdots(p_2, q_2)(p_1, q_1)$ は ℓ 個の互換の積からなる. □

本章の最後に群という言葉が出てきた。群とは対称性を記述する数学用語である。群の厳密な定義はこのコラムの最後に述べるとして、その前に対称性について考えよう。

一般に、いくつかの事物において、それらが何らかの立場において対等である(つりあっている)とき、それらの関係は対称であると言われる。対称と聞いてすぐに思いつく事象は線対称や点対称など図形的・視覚的なものであろう。しかし視覚に訴えない対称性もある。例えば、式 $x^2 + xy + y^2 + zw$ において変数 x と y は対称である。何故なら、 x と y を入れ替えると $y^2 + yx + x^2 + zw$ となり、式の見た目は変化するものの、この式はもとの式と同じ意味を表すからである。同様に z と w は対称であり、 x と z は対称でない。また、“対”という字から対称とは二物の間のみの関係と思われがちであるが、必ずしもそれに限るものではない。例えばジャンケンを考えよう。ジャンケンで出す二つの手の間には勝ち負けの関係があり、これらは対等ではない。しかしながら三つの手の間の関係としてはつりあいが取れており、三者の立場は対等である。したがってジャンケンの手は対称であると考えられる。

さて、これらに共通する性質は何であろうか。それは立場を入れ替えても本質的な変化が無いということである。そこで数学や物理においては、ある構造が与えられた対象に対して、その構造が変化しない入れ替え、およびそうした入れ替えの性質のことを対称性 (symmetry) と呼んでいる。例えば原点において点対称な平面上の図形は、ベクトルの -1 倍によってお互いが入れ替わる。また、式 $x^2 + xy + y^2 + zw$ においては、式の意味が変わらないような変数の入れ替えを考えていた。ジャンケンにおいても同様である。命題「 A は B に勝ち、 B は C に勝ち、 C は A に勝つ」が成り立つような A, B, C にそれぞれジャンケンの手を対応させる。このとき、この命題が成立し続けるような A, B, C の入れ替えがジャンケンの対称性を記述する。なお、対称性とは入れ替えの総数のみを考えるのではなく、入れ替えの相互関係についても考慮する概念である。

簡単な数学的構造の対称性は置換を用いて記述することができる。そこで具体的な図形の対称性を列挙してみよう。始めに正三角形 ABC を例にとろう。 $\triangle ABC$ の頂点を別の頂点へ移動させて再び正三角形を得る操作を考える。例えば $\triangle ABC$ を 120° 回転させると頂点 A, B, C はそれぞれ B, C, A に移動する。この置き替えは巡回置換 (A, B, C) に相当している。また、頂点 A を動かさないで裏表を反転させる操作は頂点 B, C の入れ替えを意味し、これは互換 (B, C) に相当する。

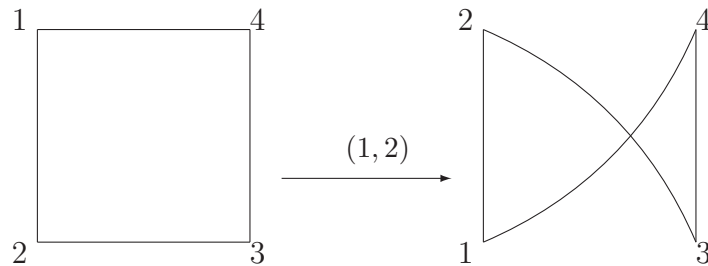


他に考えられる操作は 240° 回転に相当する巡回置換 (A, C, B) , 360° 回転 (結果的にこれは頂点を全く動かさない操作 id に等しい), B, C のいずれかを固定した反転に相当する互換 (A, C) および (A, B) であり、次の計 6 つである。

$$\{\text{id}, (B, C), (A, B), (A, B, C), (A, C, B), (A, C)\}.$$

なお、逆の操作、例えば 120° 回転の逆の操作として -120° 回転が考えられる。しかし、これは 240° 回転と結果が等しいから既に数え上げている。また、 120° 回転の後に元々 A のあった位置で頂点を固定して反転するという操作も考えられるが、この操作を全体として見ると結局、頂点 A, C の入れ替えに過ぎず、したがってこれも数え上げている。ここで、操作の組み合わせと置換の合成が対応していること $(B, C)(A, B, C) = (A, C)$ に注意したい。正三角形における頂点の移動は上の 6 種類で全てつくしている。実際、文字 A, B, C を文字 $1, 2, 3$ に置き換えれば、いま考えている置換の集合は \mathfrak{S}_3 に等しく、頂点の入れ替えをこれ以上考えることは出来ない。以上の考察により、正三角形の対称性は 3 次対称群で与えられることが分かった。

同様の考察を1, 2, 3, 4を頂点にもつ正方形に対して行ってみよう. たとえば互換(1, 2)は正方形を不変にする入れ替えではない.



何故なら, 頂点1と2のみを入れ替えると, 図形1, 2, 3, 4は正方形でなくなるからである. このような置換に注意して正方形の対称性を列挙すると, S_4 の元のうち次の計8つに数え上げられる:

$$\begin{aligned} & \{ \text{動かさない, 上下反転, 左右反転, 対角線 2-4 を軸に反転,} \\ & \quad \text{対角線 1-3 を軸に反転, } 90^\circ \text{ 回転, } 180^\circ \text{ 回転, } 270^\circ \text{ 回転} \} \\ & = \{ \text{id, } (1, 2)(3, 4), (1, 4)(2, 3), (1, 3), (2, 4), (1, 2, 3, 4), (1, 3)(2, 4), (1, 4, 3, 2) \}. \end{aligned}$$

ちなみに円や球には対称性が無限にあり, 置換を用いて記述するのは難しい. これらの対称性は, 行列式が1なる n 次直交行列全体 $SO(n)$ で記述されることを後で学ぶだろう. 図形以外の対称性はどうか. 式 $x^2 + xy + y^2 + zw$ の対称性は $\{ \text{id, } (x, y), (z, w), (x, y)(z, w) \}$ で与えられ, ジャンケンの対称性は $\{ \text{id, } (1, 2, 3), (1, 3, 2) \}$ となる. そして, 線形代数学で主に扱う構造は線形空間の構造である. \mathbb{R}^n における線形空間の構造を変えない変換の全体は, n 次可逆行列全体 $GL(n)$ と同一視され, これを一般線形群(**general linear group**)と呼ぶ.

このような“構造を変えない入れ替え全体”を抽象的に記述する言葉として, 群と呼ばれる代数構造が生みだされた. 集合 G の各元の間演算(演算記号はドット \cdot で表すか, あるいは省略することが多い)および単位元(**identity**)と呼ばれる特別な元 $e \in G$ が与えられており, 次の条件を満たすとき G を群(**group**)という:

- (1) 各 $g \in G$ について $g \cdot e = e \cdot g = g$,
- (2) 各 $g \in G$ に対して, 次の条件を満たす逆元(**inverse**) g^{-1} が存在する: $g^{-1} \cdot g = g \cdot g^{-1} = e$,
- (3) 各 $g, h, f \in G$ に対して, $(g \cdot h) \cdot f = g \cdot (h \cdot f)$ が成り立つ.

すなわち, (1) 何も入れ替えない操作 e があり, (2) 各入れ替え g に対してそれを元に戻す逆の入れ替え g^{-1} が定まっており, (3) 操作の組み合わせに関して結合律が満たされることを上の定義は述べている.

対称群や交代群, あるいはこれまで挙げてきた対称性を記述する置換の集合は, id を単位元とし, 写像の合成を演算とする群である. また, $GL(n)$ は単位行列 E_n を単位元とし, 行列の積を演算とする群である. 整数全体 \mathbb{Z} は, 0 を単位元とし, 足し算を演算とする群である. このように, 群は数学のいたるところに溢れている.

第9章 置換の符号について

本章では、まずはじめに前章で定めた置換の符号の定義に矛盾がないこと、すなわち、置換を互換の積に分解する際における互換の個数の偶奇が分解の仕方によらないことを示す。次に、置換を文字列の並び替え操作であると考え、この立場から互換の積への分解について再考する。そこでは転倒数と呼ばれる数が導入され、その議論を通してより単純な互換による分解が与えられる。

一方、転倒数を用いることで、置換の符号を別の視点から定義する方法がある。これは互換の積への分解の仕方をあらかじめ一つだけ定めておくことにより、その偶奇によって符号を定めるという方法である。そこで、こちらの方針で定義した符号が従来の性質を満たすことを再確認しておく。この方法の利点は、置換の代わりに対応する文字列を持ち出すことで、置換の詳細を語らずとも次章の冒頭で行列式が定められることにある。置換を苦手とする初学者への、ある種の教育的配慮ともいえるだろう。

なお、符号の定義に矛盾が無いと根拠なく盲信する者にとって、本章の9.1節は不要である。また、9.2節以降で述べることは置換に関する補足事項といった意味合いが強く、線形代数の議論を進める上で必ずしも必須の内容というわけではない。あえて述べたのは次の二つの理由による。一つは参考書の違いにより符号の定義が違って読者が迷わずに済むための配慮である。もう一つは、数学を応用する立場から見ても文字列の並び替えは基本的な考え方・道具となりうるからである。

9.1 符号の正当性

置換 $\sigma \in \mathfrak{S}_n$ に対して、次の数を考える：

$$s(\sigma) := \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

例 9.1.1. (1) $\sigma = (1, 2)(3, 4) \in \mathfrak{S}_4$ とすれば、

$$\begin{aligned} s(\sigma) &= \frac{\sigma(2) - \sigma(1)}{2 - 1} \cdot \frac{\sigma(3) - \sigma(1)}{3 - 1} \cdot \frac{\sigma(4) - \sigma(1)}{4 - 1} \cdot \frac{\sigma(3) - \sigma(2)}{3 - 2} \cdot \frac{\sigma(4) - \sigma(2)}{4 - 2} \cdot \frac{\sigma(4) - \sigma(3)}{4 - 3} \\ &= \frac{1 - 2}{2 - 1} \cdot \frac{4 - 2}{3 - 1} \cdot \frac{3 - 2}{4 - 1} \cdot \frac{4 - 1}{3 - 2} \cdot \frac{3 - 1}{4 - 2} \cdot \frac{3 - 4}{4 - 3} = 1. \end{aligned}$$

(2) 恒等置換 $\text{id} \in \mathfrak{S}_n$ において、 $s(\text{id}) = \prod_{1 \leq i < j \leq n} \frac{j - i}{j - i} = 1$.

$s(\sigma)$ が符号を意味することをこれから見ていく¹。さて、 $s(\sigma)$ が ± 1 の値しか取らないことはほとんど明らかであるが、念のため確認しておこう。

補題 9.1.2. $s(\sigma) = 1$ または $s(\sigma) = -1$.

Proof. $k < l$ としよう。このとき、 $s(\sigma)$ の定義式の分母において1回だけ $l - k$ が現れている。また、分子においても、 $\sigma^{-1}(l)$ と $\sigma^{-1}(k)$ のどちらが大きいかは分からないが、 $\sigma(\sigma^{-1}(l)) - \sigma(\sigma^{-1}(k)) = l - k$ あるいは $\sigma(\sigma^{-1}(k)) - \sigma(\sigma^{-1}(l)) = k - l$ のうちいずれか一方が1回だけ現れる。したがって、分子 $(l - k)$ または $(k - l)$ と分母 $(l - k)$ が相殺されて1または -1 となる。このようなことがすべての項の分母について言えるため、結局、上の式は1または -1 のいずれかのみを取り得る。□

¹したがって、 $s(\sigma)$ でもって置換の符号を定義してもよい。

次は命題 8.5.3 を $s(\sigma)$ の言葉で言い換えたものに相当している.

補題 9.1.3. 置換 $\sigma, \tau \in \mathfrak{S}_n$ について $s(\tau\sigma) = s(\tau)s(\sigma)$.

Proof. 次のように計算できる.

$$\begin{aligned} s(\tau\sigma) &= \prod_{1 \leq i < j \leq n} \frac{\tau\sigma(j) - \tau\sigma(i)}{j - i} = \left(\prod_{1 \leq i < j \leq n} \frac{\tau\sigma(j) - \tau\sigma(i)}{\sigma(j) - \sigma(i)} \right) \left(\prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} \right) \\ &= \left(\prod_{\substack{1 \leq i < j \leq n, \\ \sigma(i) < \sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{\substack{1 \leq i < j \leq n, \\ \sigma(j) < \sigma(i)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \right) \cdot s(\sigma) \\ &= \left(\prod_{\substack{1 \leq i < j \leq n, \\ \sigma(i) < \sigma(j)}} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{\substack{1 \leq i < j \leq n, \\ \sigma(j) < \sigma(i)}} \frac{\tau(\sigma(i)) - \tau(\sigma(j))}{\sigma(i) - \sigma(j)} \right) \cdot s(\sigma) \\ &\quad \uparrow \text{二つ目の } \prod \text{ において, 分母} \cdot \text{分子それぞれに } -1 \text{ を掛けた} \\ &= \left(\prod_{1 \leq k < \ell \leq n} \frac{\tau(\ell) - \tau(k)}{\ell - k} \right) \cdot s(\sigma) = s(\tau)s(\sigma). \\ &\quad \uparrow \text{この等式については後述} \end{aligned}$$

ここで, 最後から二つ目の等式は次の理由による: 1 から n の中にある二つの数の組 $i < j$ すべてを重複なく動かすとき, $k = \min\{\sigma(i), \sigma(j)\}$ および $\ell = \max\{\sigma(i), \sigma(j)\}$ とおくと, 組 $k < \ell$ も 1 から n の中を重複なくすべて動く. ゆえにこれらは同じ積を考えている. \square

上の補題を有限回適用することで, 置換 $\sigma_1, \dots, \sigma_n$ について $s(\sigma_1 \cdots \sigma_n) = s(\sigma_1) \cdots s(\sigma_n)$ を得る.

補題 9.1.4. $1 \leq k < \ell \leq n$ とする. 互換 $\sigma = (k, \ell) \in \mathfrak{S}_n$ について $s(\sigma) = -1$.

Proof. $s(\sigma)$ に現れる各項を, いくつかの場合に分けよう. まず, k, ℓ が絡まない場合と絡む場合に分けて, さらに k, ℓ と絡む場合を $i = k$ かつ $j = \ell$ の場合, そしてそれ以外の場合に細かく分ける. 次の分類は $i < j$ なる組すべてを重複なく分類している.

- $i \neq k, \ell$ かつ $j \neq k, \ell$ の場合,
- $i = k$ かつ $j = \ell$ の場合,
- $i = k$ であり $j \neq \ell$ の場合, このとき $k < j \leq n$ かつ $j \neq \ell$,
- $i = \ell$ であり $j \neq k$ の場合, このとき $\ell < j \leq n$,
- $j = k$ であり $i \neq \ell$ の場合, このとき $1 \leq i < k$,
- $j = \ell$ であり $i \neq k$ の場合, このとき $1 \leq i < \ell$ かつ $i \neq k$.

このように $s(\sigma)$ に現れる各項を分類して計算すると次のようになる:

$$\begin{aligned}
s(\sigma) &= \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} \\
&= \left(\prod_{\substack{1 \leq i < j \leq n, \\ i \neq k, \ell, j \neq k, \ell}} \frac{\sigma(j) - \sigma(i)}{j - i} \right) \left(\frac{\sigma(\ell) - \sigma(k)}{\ell - k} \right) \\
&= \left(\prod_{\substack{k < j \leq n, \\ j \neq \ell}} \frac{\sigma(j) - \sigma(k)}{j - k} \right) \left(\prod_{\ell < j \leq n} \frac{\sigma(j) - \sigma(\ell)}{j - \ell} \right) \left(\prod_{1 \leq i < k} \frac{\sigma(k) - \sigma(i)}{k - i} \right) \left(\prod_{\substack{1 \leq i < \ell, \\ i \neq k}} \frac{\sigma(\ell) - \sigma(i)}{\ell - i} \right) \\
&= \left(\prod_{\substack{1 \leq i < j \leq n, \\ i \neq k, \ell, j \neq k, \ell}} \frac{j - i}{j - i} \right) \left(\frac{k - \ell}{\ell - k} \right) \\
&\quad \left(\prod_{\substack{k < j \leq n, \\ j \neq \ell}} \frac{j - \ell}{j - k} \right) \left(\prod_{\ell < j \leq n} \frac{j - k}{j - \ell} \right) \left(\prod_{1 \leq i < k} \frac{\ell - i}{k - i} \right) \left(\prod_{\substack{1 \leq i < \ell, \\ i \neq k}} \frac{k - i}{\ell - i} \right) \\
&= 1 \cdot (-1) \cdot \left(\prod_{k < j < \ell} \frac{j - \ell}{j - k} \right) \left(\prod_{k < i < \ell} \frac{k - i}{\ell - i} \right) = -1 \cdot \left(\prod_{k < j < \ell} \frac{j - \ell}{j - k} \right) \left(\prod_{k < i < \ell} \frac{i - k}{i - \ell} \right) \\
&= -1 \cdot \left(\prod_{k < p < \ell} \frac{p - \ell}{p - k} \right) \left(\prod_{k < p < \ell} \frac{p - k}{p - \ell} \right) = -1.
\end{aligned}$$

↑ 変数を置き換えた

□

定理 9.1.5. 置換を互換の積に分解する際における互換の個数の偶奇は分解の仕方によらない.

Proof. 置換 σ が k 個の互換 $\sigma_1, \dots, \sigma_k$ の積に分解できるとすれば, 先の二つの補題から

$$s(\sigma) = s(\sigma_1 \cdots \sigma_k) = s(\sigma_1) \cdots s(\sigma_k) = (-1)^k$$

を得る. ゆえに, σ が偶数個の互換の積にも奇数個の互換の積にも分解できると仮定すると, $s(\sigma) = 1$ かつ $s(\sigma) = -1$, つまり $1 = -1$ が示され, しかしこれは起こり得ない. したがって互換の個数の偶奇は一意的である. □

上の定理の証明から直ちに次を得る.

系 9.1.6. 置換 σ について $\text{sgn}(\sigma) = s(\sigma)$.

9.2 文字列の並び替え (よりみち)

さて, 置換とは文字列を並べ替える操作とも考えられる. これを少し詳しく見てみよう. X_n 上の置換 σ を一つ固定する. n 個の文字を重複なく並べた任意の文字列 $l_1 \cdots l_n$ を別の文字列に並べ替える操作を考える. ここで, σ に対応する並び替えとは何かを考えよう. その対応のさせ方は幾つかあり, とくに自然と思われるのは次の 2 種類の方法である. この二つの方法を混同しないよう注意を促すことがこの項の目的である. ここでは具体的な $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 6 & 2 & 7 & 5 & 3 \end{pmatrix}$ を通して, 置換 σ がどういう並び替えを与えるか説明しよう.

(1) 文字 i のあった場所に文字 $\sigma(i)$ を並べる操作とみなす.

上の σ を例にとれば, 文字 1 のあった場所に文字 4 を置き, 文字 2 のあった場所に文字 1 を置き, \dots , 文字 7 のあった場所に文字 3 を置く並べ替えになる. したがって, これは文字列 1234567 を文字列 4162753 に並べ替える操作である. 抽象的に言えば, 文字列 $l_1 \dots l_n$ を文字列 $\sigma(l_1) \dots \sigma(l_n)$ に並べ替える操作である.

(2) 第 i 列目にあった文字を $\sigma(i)$ 列目に移動させる並べ替え操作とみなす.

上の σ を例にとれば, 第 1 列目にあった文字を 4 列目に置き, 第 2 列目にあった文字を 1 列目に置き, \dots , 第 7 列目にあった文字を 3 列目に置くことを意味する. したがって, 文字列 $l_4 l_1 l_6 l_2 l_7 l_5 l_3$ を文字列 $l_1 l_2 l_3 l_4 l_5 l_6 l_7$ に並べ替える操作, とくに文字列 4162753 を文字列 1234567 に並べ替える操作である. また, この操作について文字列 1234567 を並び変えると 2471635 となる. これを抽象的に述べるのはやや難しく, しいて言えば文字列 $l_{\sigma(1)} \dots l_{\sigma(n)}$ を文字列 $l_1 \dots l_n$ に並べ替える操作, あるいは文字列 $l_1 \dots l_n$ を文字列 $l_{\sigma^{-1}(1)} \dots l_{\sigma^{-1}(n)}$ に並べ替える操作となる.

上の二つの並べ替え方を念頭におきながら, 置換を互換の積に分解する方法について再考してみよう. 文字列 $123 \dots n$ において, 隣り合う二つの文字の入れ替えのみを繰り返して任意の文字列 $k_1 k_2 k_3 \dots k_n$ に変形する操作を考える. 例えば, はじめに文字 k_1 を先頭まで移動させ, 次に文字 k_2 が 2 列目になるよう移動させ \dots という操作を繰り返していけばよい.

まずは (1) の立場での入れ替えを検討する. 文字列 1234567 を 4162753 に変えるには次のような入れ替えを行えばよい.

$$(i) \quad 1234567 \xleftrightarrow{(3,4)} 1243567 \xleftrightarrow{(2,4)} 1423567 \xleftrightarrow{(1,4)} 4123567 \xleftrightarrow{(5,6)} 4123657 \\ \xleftrightarrow{(3,6)} 4126357 \xleftrightarrow{(2,6)} 4162357 \xleftrightarrow{(5,7)} 4162375 \xleftrightarrow{(3,7)} 4162735 \xleftrightarrow{(3,5)} 4162753.$$

ここで, 記号 $\xleftrightarrow{(i,j)}$ は文字 i と文字 j を入れ替える変形, すなわち (1) の意味で互換 (i, j) に対応する文字列の入れ替えを意味する. したがって, σ は上の変形を合成した互換の積に分解される:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 6 & 2 & 7 & 5 & 3 \end{pmatrix} = (3, 5)(3, 7)(5, 7)(2, 6)(3, 6)(5, 6)(1, 4)(2, 4)(3, 4).$$

このとき隣り合う列の入れ替えは何回行われるのだろうか. いまの入れ替えにおいて文字 k_i を i 列目に移動させるために行った入れ替えの回数は, k_{i+1}, \dots, k_n 中にある k_i より小さな数の個数に一致している. ゆえに, 入れ替えの総数は次の $\text{inv}(k_1, \dots, k_n)$ で与えられる:

$$\text{inv}(k_1, \dots, k_n) := \sum_{i=1}^{n-1} (k_{i+1}, \dots, k_n \text{ のうち } k_i \text{ より小さな数の個数}) \\ = i < j \text{ かつ } k_i > k_j \text{ を満たす組 } (i, j) \text{ の総数.}$$

例えば $\text{inv}(4, 1, 6, 2, 7, 5, 3) = 3 + 0 + 3 + 0 + 2 + 1 = 9$ であり, 確かに σ は 9 つの互換の積に分解されている. $\text{inv}(k_1, \dots, k_n)$ は転倒数 (**inversion number**) と呼ばれる.

次に (2) の立場において文字列の入れ替えをしてみよう. σ による文字列の入れ替えは, 4162753 を 1234567 に入れ替える操作であった. すなわち, 先程の各変形を順次逆に辿っていくことに他ならない.

$$(ii) \quad 4162753 \xleftrightarrow{(6,7)} 4162735 \xleftrightarrow{(5,6)} 4162375 \xleftrightarrow{(6,7)} 4162357 \xleftrightarrow{(3,4)} 4126357 \\ \xleftrightarrow{(4,5)} 4123657 \xleftrightarrow{(5,6)} 4123567 \xleftrightarrow{(1,2)} 1423567 \xleftrightarrow{(2,3)} 1243567 \xleftrightarrow{(3,4)} 1234567.$$

ここで, いま (2) に対応する文字の入れ替えを考えているから, 変形 (ii) における記号 $\xleftrightarrow{(i,j)}$ は i 列と j 列を入れ替える互換を指していることに注意する. これら各々の列の入れ替えの合成が σ であり, ゆえに次を得る:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 6 & 2 & 7 & 5 & 3 \end{pmatrix} = (3, 4)(2, 3)(1, 2)(5, 6)(4, 5)(3, 4)(6, 7)(5, 6)(6, 7).$$

以上は一般の置換においても成立し、とくに (2) の立場から眺めることで次を得る:

定理 9.2.1. 任意の置換 $\begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$ は $(i, i+1)$ なる形の $\text{inv}(k_1, \dots, k_n)$ 個の互換の積に分解される.

9.3 転倒数による符号の定義 (よりみち)

転倒数を用いても置換の符号が定まることを見ておこう. すなわち, 置換 $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$ に対して \tilde{s} を次で定める:

$$\tilde{s}(\sigma) := (-1)^{\text{inv}(k_1, \dots, k_n)}.$$

既に我々は定理 9.1.5 および 9.2.1 を得ているから, 上の $\tilde{s}(\sigma)$ が $\text{sgn}(\sigma)$ に一致することは分かっている. しかしながら, これらの知識を仮定せずとも符号における最も重要な性質である命題 8.5.3 が直接得られることを示そう.

例 9.3.1. (1) 恒等置換 $\text{id} \in \mathfrak{S}_n$ において, $\tilde{s}(\text{id}) = (-1)^{\text{inv}(1, \dots, n)} = (-1)^0 = 1$.

(2) 互換 $\sigma = (i, j)$ (ただし $i < j$) について $\tilde{s}(\sigma)$ を計算しよう. σ は

$$\sigma = \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & \cdots & i-1 & j & i+1 & \cdots & j-1 & i & j+1 & \cdots & n \end{pmatrix}.$$

と書ける. また文字列 $Z = 1, \dots, i-1, j, i+1, \dots, j-1, i, j+1, \dots, n$ の転倒数を数えると, j より右側にある j 未満の数は $(j-i)$ 個. $i < \ell < j$ なる ℓ については, ℓ の右側にある ℓ 未満の数は i のみであり, このような ℓ は全部で $(j-i-1)$ 個ある. また, これら以外の文字については数えなくてよいゆえ $\text{inv}(Z) = (j-i) + (j-i-1) = 2(j-i) - 1$. これは奇数である. ゆえに $\tilde{s}(\sigma) = -1$.

補題 9.3.2. 置換 $\sigma \in \mathfrak{S}_n$ および互換 τ に対して, $\tilde{s}(\sigma\tau) = -\tilde{s}(\sigma)$.

Proof. $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$, $\tau = (i, j)$ と置く (ただし $i < j$). このとき, $\sigma\tau$ の表示は次のようになる:

$$\sigma\tau = \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ k_1 & \cdots & k_{i-1} & k_j & k_{i+1} & \cdots & k_{j-1} & k_i & k_{j+1} & \cdots & k_n \end{pmatrix}.$$

二つの文字列 $X = k_1, \dots, k_n$ と $Y = k_1, \dots, k_{i-1}, k_j, k_{i+1}, \dots, k_{j-1}, k_i, k_{j+1}, \dots, k_n$ における転倒数の差 $\text{inv}(Y) - \text{inv}(X)$ を計算する. 転倒数とは, 各文字の右側にある自分より小さな数の個数の総数であったから, $\text{inv}(X)$ および $\text{inv}(Y)$ におけるそれぞれの和に関する各項の違いは k_i と k_j の間にある文字 k_ℓ ($i \leq \ell \leq j$) においてしか現れない. 文字列 X を Y に変えたときに, 文字 k_ℓ の右側にある k_ℓ より小さい数の個数がどれだけ変化するか見積もると次のようになる.

(i) k_i について.

$k_i > k_p$ を満たす p (ただし $i < p \leq j$) の個数ぶんだけ減る. 今の議論において k_j は特別な文字であるから個別に考えることにして, $k_i > k_p$ を満たす p ($i < p < j$) の個数を a 個とし, $k_i > k_j$ ならば $x = 1$, $k_i < k_j$ ならば $x = 0$ とおく. すると $k_i > k_p$ を満たす p ($i < p \leq j$) の個数は $a + x$ 個であり, 個数の変化としては $-(a + x)$ となる.

(ii) k_j について.

$k_p < k_j$ を満たす p (ただし $i \leq p < j$) の個数ぶんだけ増える. 上と同様の考えで, $k_p < k_j$ を満たす p ($i < p < j$) の個数を b 個とし, $k_i < k_j$ ならば $y = 1$, $k_i > k_j$ ならば $y = 0$ とおく. すると $k_p < k_j$ を満たす p ($i \leq p < j$) の個数は $b + y$ 個であり, ゆえに個数の変化も $b + y$ である.

(iii) k_ℓ (ただし ℓ は $i < \ell < j$ を満たす $j - i - 1$ 個の文字のいずれか) について.

$k_i > k_\ell$ ならば k_i の移動による変化はない. $k_i < k_\ell$ ならば k_i が k_ℓ よりも右側に移動したことにより 1 つ増える. なお, $k_i < k_\ell$ なる ℓ の個数は, (i) における議論から $j - i - 1 - a$ 個である. また, $k_\ell < k_j$ ならば k_j の移動による変化はなく, $k_\ell > k_j$ ならば k_j が移動して k_ℓ の右側から消えることにより 1 つ減る. $k_\ell > k_j$ なる ℓ の個数は (ii) における議論から $j - i - 1 - b$ 個である. 以上の考察から, 文字 k_ℓ たちにおける変化は, $i < \ell < j$ について総和を取り, $(j - i - 1 - a) - (j - i - 1 - b) = b - a$.

以上の変化の総和が $\text{inv}(Y) - \text{inv}(X)$ にあたるから

$$\text{inv}(Y) - \text{inv}(X) = -(a + x) + (b + y) + (b - a) = 2(b - a) + (y - x).$$

x, y の定め方より, $k_i < k_j$ のとき $(y - x) = (1 - 0) = 1$, $k_i > k_j$ のとき $(y - x) = 0 - 1 = -1$ ゆえいずれの場合においても $\text{inv}(Y) - \text{inv}(X)$ は奇数となる. したがって,

$$\tilde{s}(\sigma\tau) = (-1)^{\text{inv}(Y)} = (-1)^{\text{inv}(Y) - \text{inv}(X)} \cdot (-1)^{\text{inv}(X)} = (-1) \cdot \tilde{s}(\sigma).$$

□

上の定理は $\text{inv}(X)$ と $\text{inv}(Y)$ の値を直接提示せずに, したがって $\tilde{s}(\sigma)$ および $\tilde{s}(\sigma\tau)$ の値を求めずに証明がなされている. 証明において鍵となるのは $\text{inv}(Y) - \text{inv}(X)$ という量であった. このように, 何が本質的に重要かを見極めることが我々には求められている. いまの証明では, $\text{inv}(X)$ や $\text{inv}(Y)$ 自身, すなわち絶対的な量よりも, これらの間の関係 $\text{inv}(Y) - \text{inv}(X)$, つまり相対的な量が本質的なのであった.

最後に, 命題 8.5.3 に対応する性質は次のように示される.

命題 9.3.3. 置換 σ, τ について $\tilde{s}(\sigma\tau) = \tilde{s}(\sigma) \cdot \tilde{s}(\tau)$.

Proof. $\tau = \begin{pmatrix} 1 & \cdots & n \\ k_1 & \cdots & k_n \end{pmatrix}$ と置くと, 定理 9.2.1 により τ は $\text{inv}(k_1, \dots, k_n)$ 個の互換の積で書ける. $\ell = \text{inv}(k_1, \dots, k_n)$, $\tau = (p_1, q_1) \cdots (p_\ell, q_\ell)$ と置き, 補題 9.3.2 を ℓ 回適用すると,

$$\begin{aligned} \tilde{s}(\sigma\tau) &= \tilde{s}(\sigma(p_1, q_1) \cdots (p_{\ell-1}, q_{\ell-1})(p_\ell, q_\ell)) \\ &= \tilde{s}\left(\left(\sigma(p_1, q_1) \cdots (p_{\ell-1}, q_{\ell-1})\right) \cdot (p_\ell, q_\ell)\right) \\ &= -\tilde{s}(\sigma(p_1, q_1) \cdots (p_{\ell-1}, q_{\ell-1})) && \text{(ここで補題 9.3.2 を用いた)} \\ &= \cdots = (-1)^\ell \tilde{s}(\sigma) = \tilde{s}(\tau) \cdot \tilde{s}(\sigma) = \tilde{s}(\sigma) \cdot \tilde{s}(\tau). \end{aligned}$$

□

第10章 行列式の定義と性質

いよいよ行列式の定義に入ろう。本章では、行列式を特徴づける性質である多重線形性と歪対称性について述べる。行列式の定義を形式的に与えることもあり、これらの性質に実感が湧かない読者もいるかもしれない。そこで、行列式の列に関する性質のいくつかと、ベクトルの組で張られる図形の体積との関係についてもある程度説明を設けた。これらの幾何的な意味を知っておくと、より深い理解が得られることと思う。

10.1 定義

行列式の形式的な定義を次で与える。恐らく初学者にとって、定義を見ただけでその意味を理解するのは困難であると思う。にもかかわらず、なにゆえこの定義を採用するのかといえば、行列式に関する数々の命題を証明する際に、明示的に式が与えられていると議論を進めやすいからである。

定義 10.1.1. n 次正方行列 $A = [a_{ij}]$ に対して、 A の行列式 (**determinant**) を次の式で定め、これを $\det A$ あるいは $|A|$ と書く：

$$\det A := \sum_{\sigma \in \mathfrak{S}_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)} \right) = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

上で与えられた式が、ベクトルの列で張られる図形の符号付き体積に本当に一致するのだろうか、疑わしく感じている読者も多いのではないだろうか。しかしながら、この問題への解答は次章まで待つてほしい。11.3 節まで読めば、このような疑念は払拭されるであろう。

さて、対称群 \mathfrak{S}_n の元の総数は $n!$ であった。ゆえに n 次行列式は $n!$ 個の項の和として定義される。例えば 2 次の行列式は $2! = 2$ 項の和であり、3 次行列式は $3! = 6$ 項の和、4 次行列式は $4! = 24$ 項の和となる。行列のサイズが小さい場合について、行列式を実際書き下すと次のようになる。

例 10.1.2. 混乱を避ける必要がある場合に限り、行列の (i, j) -成分 a_{ij} を $a_{i,j}$ と表記する。

(1) $n = 1$ の場合. $\mathfrak{S}_1 = \{\text{id}\}$ および $\operatorname{sgn}(\text{id}) = 1$ より、

$$\det(a_{11}) = \operatorname{sgn}(\text{id}) a_{1,\text{id}(1)} = 1 \cdot a_{1,1} = a_{11}.$$

(2) $n = 2$ の場合. $\mathfrak{S}_2 = \{\text{id}, (1, 2)\}$, $\operatorname{sgn}(\text{id}) = 1$, $\operatorname{sgn}(1, 2) = -1$ であるから、

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \operatorname{sgn}(\text{id}) a_{1,\text{id}(1)} a_{2,\text{id}(2)} + \operatorname{sgn}(1, 2) a_{1,(1,2)(1)} a_{2,(1,2)(2)} = a_{11} a_{22} - a_{12} a_{21}.$$

(3) $n = 3$ の場合. 例 8.5.4 より

$$\mathfrak{S}_3 = \{\text{id}, (2, 3), (1, 2), (1, 2, 3) = (1, 3)(1, 2), (1, 3, 2) = (1, 2)(1, 3), (1, 3)\}$$

である。したがって、

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \sum_{\sigma \in \mathfrak{S}_3} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)} \\ &= \operatorname{sgn}(\operatorname{id}) a_{1,\operatorname{id}(1)} a_{2,\operatorname{id}(2)} a_{3,\operatorname{id}(3)} + \operatorname{sgn}(1, 2, 3) a_{1,(1,2,3)(1)} a_{2,(1,2,3)(2)} a_{3,(1,2,3)(3)} \\ &\quad + \operatorname{sgn}(1, 3, 2) a_{1,(1,3,2)(1)} a_{2,(1,3,2)(2)} a_{3,(1,3,2)(3)} + \operatorname{sgn}(1, 2) a_{1,(1,2)(1)} a_{2,(1,2)(2)} a_{3,(1,2)(3)} \\ &\quad + \operatorname{sgn}(2, 3) a_{1,(2,3)(1)} a_{2,(2,3)(2)} a_{3,(2,3)(3)} + \operatorname{sgn}(1, 3) a_{1,(1,3)(1)} a_{2,(1,3)(2)} a_{3,(1,3)(3)} \\ &= a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{12} a_{21} a_{33} - a_{11} a_{23} a_{32} - a_{13} a_{22} a_{31}. \end{aligned}$$

本章および次章で述べる行列式の諸性質は、行ベクトルに関するものと列ベクトルに関するものに分けられる。次の定理は、そのいずれか一方が示されれば、他方も直ちに得られることを意味している。

定理 10.1.3. 任意の正方行列とその転置行列の行列式は等しい。すなわち、 $\det {}^t A = \det A$ 。

行列式を、線形写像の体積拡大率と意味づける立場においては、これを、列ベクトルの組で張られる図形と関連づけたのであった。ところが列ベクトルに限定する必要はなくて、行ベクトルの組で張られる図形の体積として行列式を意味づけしても構わないことを上の定理は主張している。

定理 10.1.3 を示すには対称群の間の 1 対 1 対応を考察する補助的な議論が必要となるゆえ、証明は 12 章で行おう。

たすきがけ

2 次および 3 次の行列式の展開式を効率よく覚える手段として、たすきがけ (またはサラスの方法) と呼ばれる手法がある。右下がりの組の符号を正、左下がりの組の符号を負と考えることで展開式を得る。

$$\begin{array}{ccc} \begin{array}{|c|} \hline a_{11} \\ \hline \end{array} & \begin{array}{|c|} \hline a_{12} \\ \hline \end{array} \\ \begin{array}{|c|} \hline a_{21} \\ \hline \end{array} & \begin{array}{|c|} \hline a_{22} \\ \hline \end{array} \\ \hline & \hline \end{array} = a_{11} a_{22} - a_{12} a_{21}. \\ \text{(-)} \qquad \qquad \text{(+)} \end{array}$$

3 次の場合はやや複雑になるが、次のように分けよう。

$$\begin{array}{ccc} \text{(+)} & \begin{array}{|c|} \hline a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ \hline \end{array} & \text{(-)} & \begin{array}{|c|} \hline a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ \hline \end{array} \end{array}$$

上の図式をもとに次の展開式を得る:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{12} a_{21} a_{33} - a_{11} a_{23} a_{32} - a_{13} a_{22} a_{31}.$$

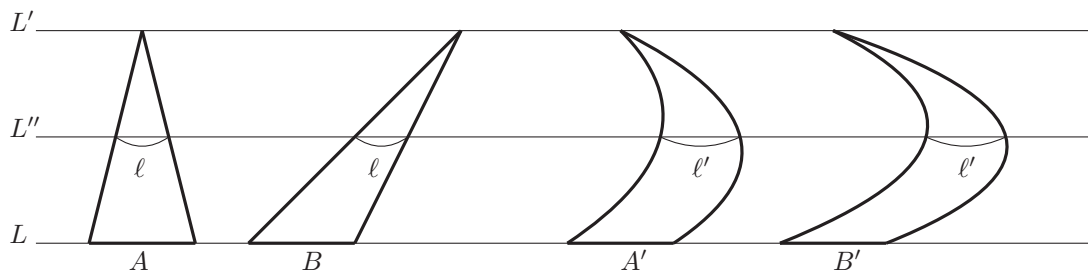
なお、4 次以上の行列式には、この様な方法は使えない。

10.2 カヴァリエリの原理

行列式の性質と図形の体積との関係を考察するにあたり、本書では何度かカヴァリエリの原理について言及する。やや寄り道となってしまふものの、この主張を正確に述べておこう。

カヴァリエリの原理.(Cavalieri's principle)

2次元の場合: 平面上の図形 A, B が, 平行な2直線 L, L' の間に挟まれているとする. このとき, L と平行な任意の直線 L'' に対して, L'' と A, B それぞれとの交わりである線分の長さが共に一致するとき, A と B の面積は等しい.



3次元の場合: \mathbb{R}^3 上の図形 A, B が, 平行な2つの平面 L, L' の間に挟まれているとする. このとき, L と平行な任意の平面 L'' に対して, L'' と A, B との交わりからなる図形の面積が共に一致するとき, A と B の体積は等しい.

高次元の場合については次のように拡張される. \mathbb{R}^n において次元がちょうど1だけ小さい $n-1$ 次元の空間のことを超平面 (hyperplane) という.

n 次元の場合: \mathbb{R}^n 上の図形 A, B が, 平行な2つの超平面 L, L' の間に挟まれているとする. このとき, L と平行な任意の超平面 L'' に対して, L'' と A, B との交わりからなる図形の $n-1$ 次元体積が共に一致するとき, A と B の n 次元体積は等しい.

よりみち (体積とは何か)

いかなる図形に対してカヴァリエリの原理が成立するのだろうか. 読者の中にはこのような素朴な疑問を持つ者もいるかと思う. しかし, これに答えるのは容易ではない. 何故なら, そもそも図形とは何か, そしてその体積とは何か, という図形や体積の定義について論じる必要があるからである. 体積の議論がやや面倒なものであることは, 微積分学において定積分 (リーマン積分) の定義に苦労したことを思い返せば想像がつくことと思う.

こうお茶を濁してばかりでは不興を買うだろうから, カヴァリエリの原理が成立している状況の一つ挙げておこう. 例えば, 二つの連続関数のグラフ $f, g: [a, b] \rightarrow \mathbb{R}$ で挟まれた図形 A の面積は, 積分 $\int_a^b (f(x) - g(x)) dx$ によって得られる. すなわち, 図形 A の面積は, $h(x) := f(x) - g(x)$ のグラフと x 軸で挟まれた図形 B の面積に等しい. ここで, y 軸と平行な直線 $x = t$ (ただし $t \in [a, b]$) で A, B を切ってみよう. すると, それら線分の長さは共に $h(t)$ である. すなわち, 図形 A, B に対してカヴァリエリの原理が成立している. 更に高次元の例としては, 重積分が累次積分 (逐次積分) によって計算されることから, 重積分により体積が計算できる図形に対してカヴァリエリの原理が成立することが分かる.

ところで, 上の例は図形の体積が積分で与えられることを仮定した上での話であった. それでは, 図形の体積が積分で与えられる根拠とは一体何であろうか. ここで話は飛ぶが, 積分論をつきつめると, 積分が一致する (体積が等しい) とはどういうことかを再考する必要に迫られ, リーマン積分とは別の方法で定積分 (重積分) を再定義することになる. この一連の理論は測度論と呼ばれ, 図形の体積の厳密化に相当するルベーグ測度から定められる積分をルベーグ積分という. そして, 測度論における重積分と累次積分の一致を主張する定理はフビニの定理と呼ばれる.

10.3 多重線形性

次の二つの命題で述べている行列式の性質は, 多重線形性 (multilinearity) と呼ばれる.

命題 10.3.1. 一つの行 (あるいは列) を r 倍すると, 行列式も r 倍になる. すなわち,

$$(1) \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ ra_{i1} & \cdots & ra_{in} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = r \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix},$$

$$(2) \begin{vmatrix} a_{11} & \cdots & ra_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & ra_{nj} & \cdots & a_{nn} \end{vmatrix} = r \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}.$$

Proof. (1) は次の式変形による:

$$(\text{左辺}) = \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots (ra_{i\sigma(i)}) \cdots a_{n\sigma(n)} = r \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = (\text{右辺}).$$

(2) は, 定理 10.1.3 を用いて (1) に帰着させることができる:

$$\begin{vmatrix} a_{11} & \cdots & ra_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & ra_{nj} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \vdots & \vdots \\ ra_{1j} & \cdots & ra_{nj} \\ \vdots & \vdots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{vmatrix} = r \begin{vmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \vdots & \vdots \\ a_{1j} & \cdots & a_{nj} \\ \vdots & \vdots & \vdots \\ a_{1n} & \cdots & a_{nn} \end{vmatrix} = r \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}.$$

□

命題 10.3.2. (1) i 行目を除くすべての行が等しい行列式どうしの和は, i 行目を互いの i 行ベクトルの和とした行列式に等しい:

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ b_{i1} & \cdots & b_{in} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ c_{i1} & \cdots & c_{in} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ b_{i1} + c_{i1} & \cdots & b_{in} + c_{in} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

(2) j 列目を除くすべての列が等しい行列式どうしの和は, j 列目を互いの j 列ベクトルの和とした行列式に等しい:

$$\begin{vmatrix} a_{11} & \cdots & b_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & b_{nj} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & c_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & c_{nj} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & b_{1j} + c_{1j} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & b_{nj} + c_{nj} & \cdots & a_{nn} \end{vmatrix}.$$

Proof. (1) は次より得る:

$$\begin{aligned} (\text{右辺}) &= \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots (b_{i\sigma(i)} + c_{i\sigma(i)}) \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathfrak{S}_n} (\text{sgn}(\sigma) a_{1\sigma(1)} \cdots b_{i\sigma(i)} \cdots a_{n\sigma(n)} + \text{sgn}(\sigma) a_{1\sigma(1)} \cdots c_{i\sigma(i)} \cdots a_{n\sigma(n)}) \\ &= \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots b_{i\sigma(i)} \cdots a_{n\sigma(n)} + \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots c_{i\sigma(i)} \cdots a_{n\sigma(n)} = (\text{左辺}). \end{aligned}$$

(2) については, 前命題 (2) の証明のように (1) に帰着することで得られる.

□

例 10.3.3.

$$\begin{aligned} \begin{vmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{vmatrix} &= \begin{vmatrix} a_1 & b_1 \\ c_1 + c_2 & d_1 + d_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c_1 + c_2 & d_1 + d_2 \end{vmatrix} \\ &= \begin{vmatrix} a_1 & b_1 \\ c_1 & d_1 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 \\ c_2 & d_2 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c_1 & d_1 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c_2 & d_2 \end{vmatrix}. \end{aligned}$$

一般には、 $\begin{vmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{vmatrix} \neq \begin{vmatrix} a_1 & b_1 \\ c_1 & d_1 \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c_2 & d_2 \end{vmatrix}$ であることに注意せよ。

正方行列 $A = [a_1, \dots, a_n]$ の各々の列ベクトルで張られる図形の符号付き体積に行列式 $|A|$ が一致することを認めたいので、多重線形性の幾何的な意味を $n = 2$ の場合について考察しよう。一つのベクトル

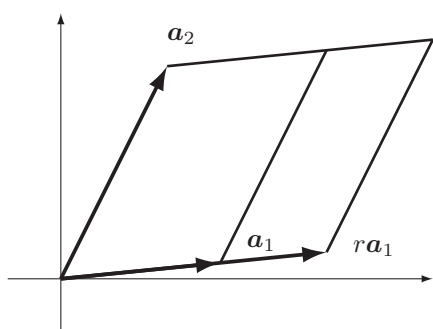


図 10.1: ベクトルのスカラー倍

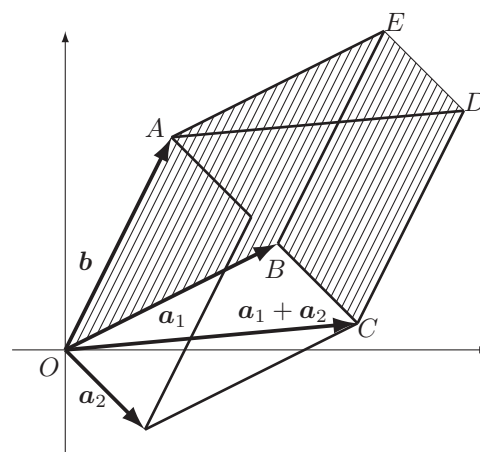


図 10.2: ベクトルの和

を r 倍すれば面積が r 倍されることは明らかである (図 10.1). 和については、図 10.2 を見よ. $a_1 + a_2, b$ で張られる平行四辺形 $A OCD$ と六角形 $A O B C D E$ の面積は等しい. これは、 b と平行な直線でこれらの図形を切ると、それぞれの長さがちょうど b の長さに一致することによる (カヴァリエリの原理). 六角形 $A O B C D E$ の面積は平行四辺形 $A O B E$ の面積 $\det(a_1, b)$ と平行四辺形 $E B C D$ の面積の和に等しい. 平行四辺形 $E B C D$ は a_2, b で張られる平行四辺形を a_1 方向に平行移動したものであるから、その面積は $\det(a_2, b)$ である. 以上より、 $\det(a_1 + a_2, b) = \det(a_1, b) + \det(a_2, b)$ となる.

10.4 歪対称性

次の性質は歪対称性 (skew-symmetry) あるいは反対称性 (antisymmetry) と呼ばれる. この命題の証明も 12 章にまわそう.

命題 10.4.1 (歪対称性). 列の入れ替え (または行の入れ替え) を行うと行列式は -1 倍される. すなわち、各 x_k ($k = 1, \dots, n$) を n 次列ベクトルとすると、 $i < j$ について、

$$(1) \det[x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_n] = -\det[x_1, \dots, x_n],$$

$$(2) \det {}^t[x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_n] = -\det {}^t[x_1, \dots, x_n].$$

歪対称性の幾何的な意味を考えよう. 列を入れ替えても、それらで張られる図形はもとの図形と同じである. したがって、列の入れ替えによる行列式の変化は、体積の符号が逆になるかどうかに限られている. $n = 2$ の場合は、列を入れ替えると第 1 列を第 2 列に重ねる際の回転の向きが逆になることから、符号も

逆になることが分かる。3次元の場合は、二つの列の入れ替えは平行六面体の底面の符号付き面積が -1 倍されることに相当し、したがって体積も -1 倍される。4次以上についても同様のことが想像されよう。

さて、 n 次列ベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_n$ の中に互いに等しい列があれば、それらで張られる図形は $n-1$ 次元以下に潰れている。よって、この図形の n 次元体積は0である。この事実を行列式の言葉で述べると次の命題になる。この性質から、行列式は代数的 (**alternating**) であると呼ばれる。一般に、交代性は歪対称性から直ちに導くことができる¹。

命題 10.4.2 (交代性). $i \neq j$ について、 i 列と j 列が等しい行列式、および i 行と j 行が等しい行列式の値はそれぞれ0である。

Proof. 正方行列 A に対して、 A の i 列と j 列を入れ替えた行列を B とすれば、命題 10.4.1 より $\det B = -\det A$ である。 A の i 列と j 列が等しいならば、 $A = B$ ゆえ $\det A = -\det A$ となり、これを移項すると $2\det A = 0$ 。したがって $\det A = 0$ である。行についても同様の考察が得られる。□

更に、多重線形性と歪対称性から次が導かれる。行列式の値を求めるための計算において、この性質は何度も利用されることになるだろう。

命題 10.4.3. ある列 (あるいは行) の何倍かを別の列 (あるいは行) に加えても行列式は変わらない。すなわち、各 \mathbf{x}_k ($k = 1, \dots, n$) を n 次列ベクトルとするとき、 $i \neq j$ について

$$(1) \det[\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j + r\mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n] = \det[\mathbf{x}_1, \dots, \mathbf{x}_n],$$

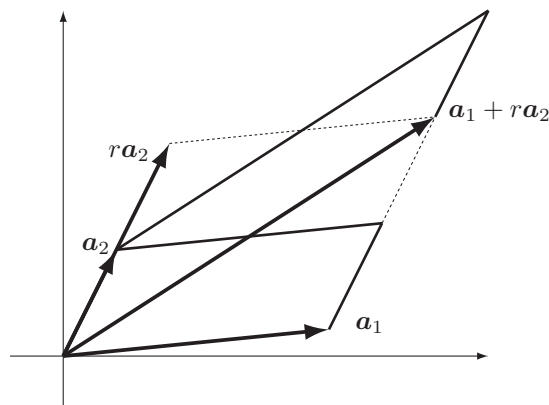
$$(2) \det {}^t[\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j + r\mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n] = \det {}^t[\mathbf{x}_1, \dots, \mathbf{x}_n].$$

Proof. 列について証明する。命題 10.3.2(2) を用いて二つの行列式に分解すると

$$\begin{aligned} \det[\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j + r\mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n] &= \det[\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n] + \det[\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, r\mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n] \\ &= \det[\mathbf{x}_1, \dots, \mathbf{x}_n] + r \det[\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n] \\ &\quad (\text{上式の第2項は、}i\text{列と}j\text{列がともに}\mathbf{x}_i\text{なる行列式ゆえ、その値は命題10.4.2より}0) \\ &= \det[\mathbf{x}_1, \dots, \mathbf{x}_n] + r \cdot 0 = \det[\mathbf{x}_1, \dots, \mathbf{x}_n]. \end{aligned}$$

行の場合についても同様の考察を行えばよい。□

右図は、命題 10.4.3 の幾何的意味を図示したものである。等式 $\det(\mathbf{a}_1 + r\mathbf{a}_2, \mathbf{a}_2) = \det(\mathbf{a}_1, \mathbf{a}_2)$ なる主張は、右図の二つの平行四辺形の面積が等しいことを意味し、これはカヴァリエリの原理からも導くことができる。



¹行列式に限らず、歪対称性を満たす写像は交代性も満たす (補題 12.3.2).

第11章 行列式の計算

行列式の値を求めるための計算例を紹介する. 多くの計算演習をこなすことで, 行列式が, ベクトルの組で張られる図形の符号付き体積を意味することを実感してもらえないだろうか. そして, この実感が妥当であるゆえんを 11.3 節において解説する.

11.1 サイズの小さい行列式との関係

行列式の値の計算においては, 次の命題を用いて, よりサイズの小さい行列式の計算に帰着させる手法を用いることが多い. この性質は, 角柱の体積が底面積と高さの積で表されることを述べたものである. 証明は次章にまわす (定理 12.5.1).

命題 11.1.1.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

上式右側の等号の幾何的な意味を説明する. n 個の n 次列ベクトル $\mathbf{a}_1, \dots, \mathbf{a}_n$ で張られる図形を D とし, その n 次元体積を $V(D)$ とする.

まずは特別な場合を考え,

$$\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, \mathbf{a}_n = \begin{bmatrix} 0 \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} \quad \text{としよう.}$$

このとき, $n-1$ 個のベクトル $\mathbf{a}_2, \dots, \mathbf{a}_n$ で張られる $n-1$ 次元以下の図形を考えよう. この図形は, \mathbb{R}^n における第 1 座標 $= 0$ を満たす点からなる超平面内に位置している. また, この図形は, $\mathbf{a}'_2 =$

$$\begin{bmatrix} a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, \mathbf{a}'_n = \begin{bmatrix} a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} \quad \text{で張られる } \mathbb{R}^{n-1} \text{ 上の図形と合同であり, ゆえにその } n-1 \text{ 次元体積は } \det[\mathbf{a}'_2, \dots, \mathbf{a}'_n]$$

である. したがって, 図形 D は底面積 $\det[\mathbf{a}'_2, \dots, \mathbf{a}'_n]$ および高さ a_{11} なる角柱であり, ゆえに体積 $V(D)$ は $a_{11} \det[\mathbf{a}'_2, \dots, \mathbf{a}'_n]$ に一致する.

\mathbf{a}_1 が一般の n 次列ベクトルの場合, 図形 D は, いま考えていた角柱を \mathbf{a}_1 方向に歪ませた形になる (図 11.1). カヴァリエリの原理によれば, D の体積はもとの角柱の体積に等しく, $V(D) = a_{11} \det[\mathbf{a}'_2, \dots, \mathbf{a}'_n]$ である.

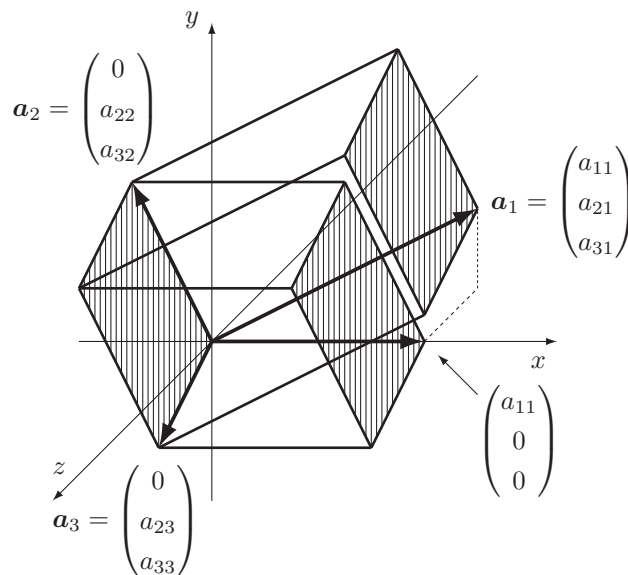


図 11.1: $n = 3$ の場合

例 11.1.2. 上三角行列の行列式は, 対角成分の積に一致する:

$$\begin{vmatrix} a_1 & & * \\ & a_2 & \\ O & \ddots & \\ & & a_n \end{vmatrix} = a_1 \begin{vmatrix} a_2 & & * \\ & a_3 & \\ O & \ddots & \\ & & a_n \end{vmatrix} = \cdots = a_1 a_2 \cdots a_n.$$

とくに, $|E| = 1$.

11.2 計算例

ここに, 行列式の計算例を挙げる. 命題 11.1.1 が適用できる形を目指して, これまで挙げてきた行列式の性質 (とくに命題 10.4.3) を用いて行列式を変形していく, というのが基本的な方針である. その変形過程は行列の行基本変形と似ているものの, スカラー倍や行の入れ替えで行列式自身に変化が生じること, および行だけではなく列についての変形も許されることに注意せよ.

例 11.2.1. (1) 次の計算では, 行に関する性質のみを用いて行列式のサイズを小さくしている. 逆に, 列の性質のみを用いてサイズを小さくすることもできる.

$$\begin{aligned} \begin{vmatrix} 12 & 16 & 32 \\ -6 & 13 & 4 \\ 15 & 10 & -20 \end{vmatrix} &= 4 \begin{vmatrix} 3 & 4 & 8 \\ -6 & 13 & 4 \\ 15 & 10 & -20 \end{vmatrix} = 4 \cdot 5 \begin{vmatrix} 3 & 4 & 8 \\ -6 & 13 & 4 \\ 3 & 2 & -4 \end{vmatrix} = 20 \begin{vmatrix} 3 & 4 & 8 \\ 0 & 21 & 20 \\ 0 & -2 & -12 \end{vmatrix} \\ &= 20 \cdot 3 \begin{vmatrix} 21 & 20 \\ -2 & -12 \end{vmatrix} = 60 \cdot (-2) \begin{vmatrix} 21 & 20 \\ 1 & 6 \end{vmatrix} \\ &= -120 \cdot (21 \cdot 6 - 20) = -120 \cdot 106 = -12720. \end{aligned}$$

(2)

$$\begin{vmatrix} 3 & 2 & 8 & 4 \\ 2 & 3 & 5 & 7 \\ 1 & 9 & 8 & 0 \\ 0 & 8 & 1 & 5 \end{vmatrix} = \begin{vmatrix} 0 & -25 & -16 & 4 \\ 0 & -15 & -11 & 7 \\ 1 & 9 & 8 & 0 \\ 0 & 8 & 1 & 5 \end{vmatrix} = (-1)^2 \begin{vmatrix} 1 & 9 & 8 & 0 \\ 0 & -25 & -16 & 4 \\ 0 & -15 & -11 & 7 \\ 0 & 8 & 1 & 5 \end{vmatrix} = \begin{vmatrix} -25 & -16 & 4 \\ -15 & -11 & 7 \\ 8 & 1 & 5 \end{vmatrix}$$

$$= (-1) \begin{vmatrix} -16 & -25 & 4 \\ -11 & -15 & 7 \\ 1 & 8 & 5 \end{vmatrix} = -(-1)^2 \begin{vmatrix} 1 & 8 & 5 \\ -16 & -25 & 4 \\ -11 & -15 & 7 \end{vmatrix} = - \begin{vmatrix} 1 & 8 & 5 \\ 0 & 103 & 84 \\ 0 & 73 & 62 \end{vmatrix}$$

↑ 1列目と2列目を入れ替えた

$$= - \begin{vmatrix} 103 & 84 \\ 73 & 62 \end{vmatrix} = -2 \begin{vmatrix} 103 & 42 \\ 73 & 31 \end{vmatrix} = -2 \begin{vmatrix} 30 & 11 \\ 73 & 31 \end{vmatrix} = -2 \begin{vmatrix} 30 & 11 \\ 13 & 9 \end{vmatrix} = -2 \begin{vmatrix} 17 & 2 \\ 13 & 9 \end{vmatrix}$$

$$= -2(17 \cdot 9 - 2 \cdot 13) = -2(153 - 26) = -2 \cdot 127 = -254.$$

(3)

$$\begin{vmatrix} 3 & 1 & 4 & 1 & 5 \\ 9 & 2 & 6 & 5 & 3 \\ 5 & 8 & 9 & 7 & 9 \\ 3 & 2 & 3 & 8 & 4 \\ 6 & 2 & 6 & 4 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 3 & 2 & -2 & 3 & -7 \\ -19 & 8 & -23 & -1 & -31 \\ -3 & 2 & -5 & 6 & -6 \\ 0 & 2 & -2 & 2 & -7 \end{vmatrix} = (-1) \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 3 & -2 & 3 & -7 \\ 8 & -19 & -23 & -1 & -31 \\ 2 & -3 & -5 & 6 & -6 \\ 2 & 0 & -2 & 2 & -7 \end{vmatrix}$$

$$= -(-1)^2 \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 3 & 2 & 3 & 7 \\ 8 & -19 & 23 & -1 & 31 \\ 2 & -3 & 5 & 6 & 6 \\ 2 & 0 & 2 & 2 & 7 \end{vmatrix} = - \begin{vmatrix} 3 & 2 & 3 & 7 \\ -19 & 23 & -1 & 31 \\ -3 & 5 & 6 & 6 \\ 0 & 2 & 2 & 7 \end{vmatrix} = - \begin{vmatrix} 3 & 2 & 3 & 7 \\ -1 & 35 & 17 & 73 \\ 0 & 7 & 9 & 13 \\ 0 & 2 & 2 & 7 \end{vmatrix}$$

$$= - \begin{vmatrix} 0 & 107 & 54 & 226 \\ -1 & 35 & 17 & 73 \\ 0 & 7 & 9 & 13 \\ 0 & 2 & 2 & 7 \end{vmatrix} = -(-1) \begin{vmatrix} -1 & 35 & 17 & 73 \\ 0 & 107 & 54 & 226 \\ 0 & 7 & 9 & 13 \\ 0 & 2 & 2 & 7 \end{vmatrix} = (-1) \begin{vmatrix} 107 & 54 & 226 \\ 7 & 9 & 13 \\ 2 & 2 & 7 \end{vmatrix}$$

$$= - \begin{vmatrix} 57 & 4 & 51 \\ 1 & 3 & -8 \\ 2 & 2 & 7 \end{vmatrix} = - \begin{vmatrix} 0 & -167 & 507 \\ 1 & 3 & -8 \\ 0 & -4 & 23 \end{vmatrix} = -(-1) \begin{vmatrix} 1 & 3 & -8 \\ 0 & -167 & 507 \\ 0 & -4 & 23 \end{vmatrix}$$

$$= \begin{vmatrix} -167 & 507 \\ -4 & 23 \end{vmatrix} = (-1) \begin{vmatrix} 167 & 507 \\ 4 & 23 \end{vmatrix} = - \begin{vmatrix} 167 & 6 \\ 4 & 11 \end{vmatrix}$$

$$= -(167 \cdot 11 - 6 \cdot 4) = -(1837 - 24) = -1813.$$

定理 11.2.2 (ヴァンデルモンドの行列式).

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Proof. $|A| = |{}^t A|$ より一つ目の等式は明らか. 二つ目の等式を行列のサイズに関する帰納法で示す. 2次の場合は $\begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = x_2 - x_1$ ゆえ等式は成立する. $n-1$ 次の場合に等式が成り立つと仮定して n 次の場合を示そう. 次の最初の等式では, n 行目, $(n-1)$ 行目, \dots , 2行目の順に変形している.

$$\begin{aligned}
& \left| \begin{array}{cccc} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{array} \right| = \left| \begin{array}{cccc} 1 & 1 & \dots & 1 \\ 0 & x_2 - x_1 & \dots & x_n - x_1 \\ 0 & x_2^2 - x_1x_2 & \dots & x_n^2 - x_1x_n \\ \vdots & \vdots & \vdots & \vdots \\ 0 & x_2^{n-1} - x_1x_2^{n-2} & \dots & x_n^{n-1} - x_1x_n^{n-2} \end{array} \right| \begin{array}{l} \text{1行} \times x_1 \text{を引く} \\ \text{2行} \times x_1 \text{を引く} \\ \vdots \\ \text{n-1行} \times x_1 \text{を引く} \end{array} \\
& = \left| \begin{array}{cccc} x_2 - x_1 & x_3 - x_1 & \dots & x_n - x_1 \\ x_2(x_2 - x_1) & x_3(x_3 - x_1) & \dots & x_n(x_n - x_1) \\ \vdots & \vdots & \vdots & \vdots \\ x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \dots & x_n^{n-2}(x_n - x_1) \end{array} \right| \\
& = (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) \left| \begin{array}{cccc} 1 & 1 & \dots & 1 \\ x_2 & x_3 & \dots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_2^{n-2} & x_3^{n-2} & \dots & x_n^{n-2} \end{array} \right| \\
& = (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) \prod_{2 \leq i < j \leq n} (x_j - x_i) \quad (\text{ここで帰納法の仮定を用いた}) \\
& = \prod_{1 \leq i < j \leq n} (x_j - x_i).
\end{aligned}$$

□

11.3 体積との関係

前章と本章を通して、行列式が図形の符号付き体積を表すと見なしても、確かに幾何的な視点と両立することを見てきた。また、両立することの根拠として、図形の n 次元体積についてカヴァリエリの原理が成り立つことを認めていたのであった。そして、行列式の計算例においても、カヴァリエリの原理と歪対称性から導かれる性質さえあれば具体的な値が原理的に得られることを見てきた。

さて、未だ我々は、図形の n 次元体積なる概念の厳密な定義は与えていないけれども、仮にそのような概念があるとすれば、 n 個の列ベクトルで張られる図形の体積について少なくともカヴァリエリの原理が認められるべきことに異論はないだろう。このことは、ベクトルの列に対してそれらの張る図形の体積を対応させる写像を考えたときに、それが多重線形性を満たすことを意味する。また、体積に符号を導入するとすれば、歪対称性が満たされると考えることは自然である。そして、当然のことながら n 次元単位立方体の体積は $1^n = 1$ であると考えよう。以上を満たす写像が行列式以外にないことを次の定理は主張している。すなわち、行列式とは、 n 個の列ベクトルで張られる図形の体積を符号付きで表す量のことである。

定理 11.3.1. 実数を成分に持つ n 次正方行列全体の集合を $M_n(\mathbb{R})$ とする¹。次の性質すべてを満たす写像 $D: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ を取れば、 D は行列式に一致する:

$\mathbf{a}_1, \dots, \mathbf{a}_n$ および $\mathbf{b}_i, \mathbf{c}_i$ をそれぞれ n 次列ベクトルとするとき、

(1) 多重線形性:

(i) $D[\mathbf{a}_1, \dots, r\mathbf{a}_i, \dots, \mathbf{a}_n] = rD[\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n]$,

(ii) $D[\mathbf{a}_1, \dots, \mathbf{b}_i + \mathbf{c}_i, \dots, \mathbf{a}_n] = D[\mathbf{a}_1, \dots, \mathbf{b}_i, \dots, \mathbf{a}_n] + D[\mathbf{a}_1, \dots, \mathbf{c}_i, \dots, \mathbf{a}_n]$,

(2) 歪対称性: $D[\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_j, \mathbf{a}_{i+1}, \dots, \mathbf{a}_{j-1}, \mathbf{a}_i, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n] = -D[\mathbf{a}_1, \dots, \mathbf{a}_n]$,

¹ $M_n(\mathbb{R})$ の括弧内にある \mathbb{R} は、行列の各成分が実数であることを意味する。複素数を成分とする n 次正方行列全体の集合は $M_n(\mathbb{C})$ と書く。

(3) 正規化 : $D(E) = 1$.

この定理の証明も次章で与えよう (定理 12.4.1).

まとめ (行列式の定義)

行列式を定義するには、大きく分けて二通りの方法がある。一つは定義 10.1.1 にあるように、置換を用いて形式的に定義する方法である。また、置換の符号は順列の転倒数を用いても定義できるゆえ、形式的な定義のみが必要な場合には、順列と転倒数を用いた次の式で定めてもよい:

$$\det A := \sum_{(k_1, \dots, k_n) \in P_n} (-1)^{\text{inv}(k_1, \dots, k_n)} a_{1k_1} a_{2k_2} \cdots a_{nk_n}.$$

ここで、 P_n は n 個の文字 $1, \dots, n$ の並べ替え (順列) 全体のなす集合とする。

もう一つの行列式の定義は、ベクトルの列で張られる図形の符号付き体積というものである。符号付き体積とは多重線形性および歪対称性を満たす $M_n(\mathbb{R})$ 上の正規化された写像 $D: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ のことであり、このような写像 D が唯一つしか存在しないことは定理 11.3.1 が保証している。更に、符号付き体積が多重線形性や歪対称性 (とくに命題 10.4.3 の性質)、および命題 11.1.1 の性質を満たすことは、カヴァリエリの原理を通して確認することができる。そして、これらの性質を通して行列式の値を求めることができる。

第12章 行列式の性質（証明）

10 および 11 章で述べた命題の証明を本章で与える。また、行列式を、線形写像の体積拡大率とみなす文脈において、7.1 節で言及した合成関数に関する性質 $|AB| = |A| \cdot |B|$ の証明も述べる。

12.1 $|A| = |{}^tA|$ の証明

補題 12.1.1. 各 $\sigma \in \mathfrak{S}_n$ に対して、 $\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$.

Proof. $\text{sgn}(\sigma)\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma\sigma^{-1}) = \text{sgn}(\text{id}) = 1$ より $\text{sgn}(\sigma^{-1})$ と $\text{sgn}(\sigma)$ の正負は一致している。□

【補足】 上の補題は練習 8.5.5(2) から導くことができる。

補題 12.1.2. 写像 $I: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ を $I(\sigma) := \sigma^{-1}$ と定めれば、これは 1 対 1 の対応である。すなわち、 σ が重複なく \mathfrak{S}_n 全体を動くとき、 $I(\sigma)$ も重複なく \mathfrak{S}_n 全体を動く。

ここで、上の結論に現れる写像 I の性質は、次の二つの性質がともに成り立つことと言い換えられる。これらの詳細については、改めて 19 章で論じる。

定義 12.1.3. σ が重複なく \mathfrak{S}_n を動けば $I(\sigma)$ も重複なく動くとき、 I を単射 (injective) であるという。また、 σ が \mathfrak{S}_n 全体を動けば $I(\sigma)$ も \mathfrak{S}_n 全体を動くとき、 I を全射 (surjective) であるという。

さらに、 I の単射性と全射性はそれぞれ次の条件に書き下すことができる：

- 単射性: 各 $\sigma_1, \sigma_2 \in \mathfrak{S}_n$ について、 $\sigma_1 \neq \sigma_2 \Rightarrow I(\sigma_1) \neq I(\sigma_2)$,
- 全射性: 各 $\tau \in \mathfrak{S}_n$ に対して、 I に代入すると τ になる置換 $\sigma \in \mathfrak{S}_n$ が存在する。

補題 12.1.2 の証明. I の単射性および全射性は次のように示される。

(単射性): 対偶である $I(\sigma_1) = I(\sigma_2) \Rightarrow \sigma_1 = \sigma_2$ を示す。 $I(\sigma_1) = I(\sigma_2)$ とすれば、 $\sigma_1^{-1} = \sigma_2^{-1}$ である。この両辺の逆置換を取れば、 $(\sigma_1^{-1})^{-1} = (\sigma_2^{-1})^{-1}$ となる。つまり $\sigma_1 = (\sigma_1^{-1})^{-1} = (\sigma_2^{-1})^{-1} = \sigma_2$ である。

(全射性): 各 $\tau \in \mathfrak{S}_n$ に対して、 $\sigma := \tau^{-1}$ とおこう。この σ を I に代入すると τ になる。実際、 $I(\sigma) = I(\tau^{-1}) = (\tau^{-1})^{-1} = \tau$ である。□

以上の補題を用いて定理 10.1.3 を証明しよう。

定理 10.1.3 (再掲). 任意の正方行列とその転置行列の行列式は等しい。すなわち、 $\det {}^tA = \det A$ 。

Proof. $A = [a_{ij}]$, ${}^tA = [b_{ij}]$ とおけば、各 i, j について $b_{ij} = a_{ji}$ である。行列式の定義および $\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$ (補題 12.1.1) から、

$$|{}^tA| = \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \quad (12.1.1)$$

$$= \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma^{-1}) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}. \quad (12.1.2)$$

右辺の総和の各項に現れる積 $a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$ の並べ替えについて考えよう。まず $\sigma(1), \dots, \sigma(n)$ を小さい順に並べ替える：

$$\sigma(k_1) = 1, \sigma(k_2) = 2, \dots, \sigma(k_n) = n.$$

つまり,

$$\sigma = \begin{pmatrix} k_1 & k_2 & \cdots & k_n \\ 1 & 2 & \cdots & n \end{pmatrix}, \quad \sigma^{-1} = \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$$

となり, ゆえに各 $i = 1, \dots, n$ について $\sigma(k_i) = i$ および $\sigma^{-1}(i) = k_i$ である. つまり $a_{\sigma(k_i)k_i} = a_{i\sigma^{-1}(i)}$ であり, 次のように並び替えができる:

$$a_{\sigma(1)1}a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{\sigma(k_1)k_1}a_{\sigma(k_2)k_2} \cdots a_{\sigma(k_n)k_n} = a_{1\sigma^{-1}(1)}a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)}. \quad (12.1.3)$$

写像 $I: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ を補題 12.1.2 で与えたものとし, 式 12.1.2 および 12.1.3 を合わせると,

$$\begin{aligned} |{}^tA| &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma^{-1})a_{1\sigma^{-1}(1)}a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(I(\sigma))a_{1,I(\sigma)(1)}a_{2,I(\sigma)(2)} \cdots a_{n,I(\sigma)(n)}. \end{aligned}$$

σ が重複なく \mathfrak{S}_n 全体を動くとき, $\tau = I(\sigma)$ も重複なく \mathfrak{S}_n 全体を動く (補題 12.1.2). ゆえに, 上式の σ による総和は次の τ による総和に書き換えても良い:

$$\sum_{\tau \in \mathfrak{S}_n} \operatorname{sgn}(\tau)a_{1\tau(1)}a_{2\tau(2)} \cdots a_{n\tau(n)} = |A|.$$

□

12.2 歪対称性の証明

補題 12.2.1. あらかじめ置換 $\tau \in \mathfrak{S}_n$ を一つ取り, 固定しておく. このとき, 写像 $F: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ を $F(\sigma) := \sigma\tau$ と定めれば, F は 1 対 1 である. すなわち, σ が重複なく \mathfrak{S}_n 全体を動くとき, $F(\sigma)$ も重複なく \mathfrak{S}_n 全体を動く.

Proof. 補題 12.1.2 の証明と同様に, 単射性および全射性を示そう.

- 単射性: 各 $\sigma_1, \sigma_2 \in \mathfrak{S}_n$ について, $\sigma_1 \neq \sigma_2 \Rightarrow F(\sigma_1) \neq F(\sigma_2)$,
- 全射性: 各 $\varphi \in \mathfrak{S}_n$ に対して, F に代入すると φ になる置換 $\sigma \in \mathfrak{S}_n$ が存在する.

(単射性): 対偶である $F(\sigma_1) = F(\sigma_2) \Rightarrow \sigma_1 = \sigma_2$ を示す. $F(\sigma_1) = F(\sigma_2)$ とすれば F の定義より $\sigma_1\tau = \sigma_2\tau$ である. この両辺に右から τ^{-1} を掛けると

$$\begin{aligned} \sigma_1\tau\tau^{-1} &= \sigma_2\tau\tau^{-1} \\ \sigma_1 \operatorname{id} &= \sigma_2 \operatorname{id} \\ \sigma_1 &= \sigma_2. \end{aligned}$$

(全射性): 各 $\varphi \in \mathfrak{S}_n$ に対して, $\sigma := \varphi\tau^{-1}$ と定める. このとき, $F(\sigma) = F(\varphi\tau^{-1}) = \varphi\tau^{-1}\tau = \varphi \operatorname{id} = \varphi$. □

上の補題において互換 $\tau = (p, q)$ (ただし $p \neq q$) を考えると, $F: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ は偶置換を奇置換に写し, 奇置換を偶置換に写す写像となる. よって F により, 偶置換と奇置換の間に 1 対 1 の対応が与えられ, これらの総数が等しいことが分かる. すなわち:

命題 12.2.2. $n \geq 2$ について交代群 A_n の元の総数は $\frac{n!}{2}$ である.

行列式の歪対称性を証明しよう:

命題 10.4.1 (再掲). 列の入れ替え (または行の入れ替え) を行うと行列式は -1 倍される.

Proof. $p < q$ とし, p 行と q 行の入れ替えについて証明する. 正方行列 $A = [a_{ij}]$ の p 行と q 行を入れ替えた行列を $B = [b_{ij}]$ とすれば, 各 $j = 1, \dots, n$ について

$$(1) i \neq p, q \implies b_{ij} = a_{ij}, \quad (2) b_{pj} = a_{qj}, \quad (3) b_{qj} = a_{pj},$$

が成り立っている. $\tau = (p, q)$ とし, $F: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ を補題 12.2.1 で与えた写像とせよ. すなわち $F(\sigma) := \sigma\tau$ である. このとき,

$$(i) k \neq p, q \implies F(\sigma)(k) = \sigma(k), \quad (ii) F(\sigma)(p) = \sigma(q), \quad (iii) F(\sigma)(q) = \sigma(p).$$

また, $\text{sgn}(F(\sigma)) = \text{sgn}(\sigma(p, q)) = \text{sgn}(\sigma) \cdot \text{sgn}((p, q)) = -\text{sgn}(\sigma)$ である. これらを用いると

$$\begin{aligned} |B| &= \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) b_{1\sigma(1)} \cdots b_{p\sigma(p)} \cdots b_{q\sigma(q)} \cdots b_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{q\sigma(p)} \cdots a_{p\sigma(q)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathfrak{S}_n} (-\text{sgn}(F(\sigma))) a_{1,F(\sigma)(1)} \cdots a_{q,F(\sigma)(q)} \cdots a_{p,F(\sigma)(p)} \cdots a_{n,F(\sigma)(n)} \\ &= - \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(F(\sigma)) a_{1,F(\sigma)(1)} \cdots a_{p,F(\sigma)(p)} \cdots a_{q,F(\sigma)(q)} \cdots a_{n,F(\sigma)(n)}. \end{aligned}$$

↑ 各項において, 積の並び順を交換した (p 番目と q 番目を入れ替えた).

σ が重複なく \mathfrak{S}_n 全体を動くとき, $\varphi = F(\sigma)$ も重複なく \mathfrak{S}_n 全体を動く (補題 12.2.1). ゆえに, 上の最後の行にある σ による総和は次の φ による総和に書き換えても良い:

$$- \sum_{\varphi \in \mathfrak{S}_n} \text{sgn}(\varphi) a_{1\varphi(1)} \cdots a_{p\varphi(p)} \cdots a_{q\varphi(q)} \cdots a_{n\varphi(n)} = -|A|.$$

列の入れ替えについては, 定理 10.1.3 を用いて行の入れ替えに帰着させれば容易に示される. \square

12.3 多重線形性と歪対称性から導かれる性質

行列式の特徴づけ (定理 11.3.1) の証明を述べる前に, 多重線形性および歪対称性から導かれるいくつかの性質について調べておこう.

n 次正方行列全体の集合を $M_n(\mathbb{R})$ と書くのであった. 本節および次節では写像 $F: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ について論じる. F の典型的な例は行列式である. 更に, 各 $X \in M_n(\mathbb{R})$ に, 列ベクトルを並べた表示 $X = [\mathbf{x}_1, \dots, \mathbf{x}_n]$ を与えておこう. このとき F は, 列ベクトルの組を代入する関数 $F(\mathbf{x}_1, \dots, \mathbf{x}_n)$ と見なすことができる. 写像 F における多重線形性および歪対称性を改めて述べておこう:

- 多重線形性: 各 $i = 1, \dots, n$ について,

$$(i) F(\mathbf{x}_1, \dots, r\mathbf{x}_i, \dots, \mathbf{x}_n) = rF(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n),$$

$$(ii) F(\mathbf{x}_1, \dots, \mathbf{y}_i + \mathbf{z}_i, \dots, \mathbf{x}_n) = F(\mathbf{x}_1, \dots, \mathbf{y}_i, \dots, \mathbf{x}_n) + F(\mathbf{x}_1, \dots, \mathbf{z}_i, \dots, \mathbf{x}_n).$$

- 歪対称性: 各 $i < j$ について,

$$F(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_j, \mathbf{x}_{i+1}, \dots, \mathbf{x}_{j-1}, \mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n) = -F(\mathbf{x}_1, \dots, \mathbf{x}_n).$$

次の主張は明らかと言えるが, ここでは形式的な証明を与えておく.

補題 12.3.1. F が多重線形ならば, 次の性質も満たす:

$$\text{多重線形性 (iii)} \quad F\left(\mathbf{a}_1, \dots, \sum_{k=1}^{\ell} r_k \mathbf{b}_k, \dots, \mathbf{a}_n\right) = \sum_{k=1}^{\ell} r_k F(\mathbf{a}_1, \dots, \mathbf{b}_k, \dots, \mathbf{a}_n).$$

Proof. 和の個数 ℓ に関する帰納法で示す. $\ell = 1$ の場合は多重線形性の性質 (i) に他ならない. 和の個数が ℓ のときに等式が成立すると仮定し, 和の個数が $\ell + 1$ の場合について示そう.

$$\begin{aligned}
& F\left(\mathbf{a}_1, \dots, \sum_{k=1}^{\ell+1} r_k \mathbf{b}_k, \dots, \mathbf{a}_n\right) \\
&= F\left(\mathbf{a}_1, \dots, \left(\sum_{k=1}^{\ell} r_k \mathbf{b}_k\right) + r_{\ell+1} \mathbf{b}_{\ell+1}, \dots, \mathbf{a}_n\right) \\
&= F\left(\mathbf{a}_1, \dots, \sum_{k=1}^{\ell} r_k \mathbf{b}_k, \dots, \mathbf{a}_n\right) + F\left(\mathbf{a}_1, \dots, r_{\ell+1} \mathbf{b}_{\ell+1}, \dots, \mathbf{a}_n\right) \quad (\text{多重線形性 (ii)}) \\
&= \sum_{k=1}^{\ell} r_k F(\mathbf{a}_1, \dots, \mathbf{b}_k, \dots, \mathbf{a}_n) + r_{\ell+1} F(\mathbf{a}_1, \dots, \mathbf{b}_{\ell+1}, \dots, \mathbf{a}_n) \quad (\text{帰納法の仮定と多重線形性 (i)}) \\
&= \sum_{k=1}^{\ell+1} r_k F(\mathbf{a}_1, \dots, \mathbf{b}_k, \dots, \mathbf{a}_n).
\end{aligned}$$

□

次の補題は, 命題 10.4.2 の証明と同様の論法により得られるゆえ, 証明は省略しよう.

補題 12.3.2 (交代性). F が歪対称性を満たすとき, $\mathbf{a}_i = \mathbf{a}_j$ (ただし $i < j$) ならば

$$F(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n) = 0.$$

歪対称性を用いて列の入れ替えを何度か行くと, それによる代入した値の変化は元の値の ± 1 倍である. このときの正負の符号は, 入れ替えに対応する置換の符号に一致する. すなわち:

補題 12.3.3. F が歪対称性を満たすとき, 任意の置換 $\sigma \in \mathfrak{S}_n$ に対して

$$F(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}) = \text{sgn}(\sigma) F(\mathbf{x}_1, \dots, \mathbf{x}_n).$$

Proof. σ を互換の積に分解した際の互換の個数 ℓ に関する帰納法により証明する. σ 自身が互換 $\sigma = (i, j)$ の場合, $F(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}) = \text{sgn}(\sigma) F(\mathbf{x}_1, \dots, \mathbf{x}_n)$ は歪対称性そのものである.

ℓ 個の互換の積で書ける任意の置換 $\tau \in \mathfrak{S}_n$ について $F(\mathbf{x}_{\tau(1)}, \dots, \mathbf{x}_{\tau(n)}) = \text{sgn}(\tau) F(\mathbf{x}_1, \dots, \mathbf{x}_n)$ が成り立つと仮定し, $\sigma := \tau \cdot (i, j)$ (ただし $i < j$ とする) について $F(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}) = \text{sgn}(\sigma) F(\mathbf{x}_1, \dots, \mathbf{x}_n)$ が成り立つことを示そう. i 列目と j 列目の添え字に注意して変形すると

$$\begin{aligned}
F(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}) &= F(\mathbf{x}_{\tau(i,j)(1)}, \dots, \mathbf{x}_{\tau(i,j)(i)}, \dots, \mathbf{x}_{\tau(i,j)(j)}, \dots, \mathbf{x}_{\tau(i,j)(n)}) \\
&= F(\mathbf{x}_{\tau(1)}, \dots, \mathbf{x}_{\tau(j)}, \dots, \mathbf{x}_{\tau(i)}, \dots, \mathbf{x}_{\tau(n)}) \\
&= -F(\mathbf{x}_{\tau(1)}, \dots, \mathbf{x}_{\tau(i)}, \dots, \mathbf{x}_{\tau(j)}, \dots, \mathbf{x}_{\tau(n)}) \quad (\text{歪対称性を用いた}) \\
&= -\text{sgn}(\tau) F(\mathbf{x}_1, \dots, \mathbf{x}_n) = \text{sgn}(\sigma) F(\mathbf{x}_1, \dots, \mathbf{x}_n). \quad (\text{帰納法の仮定を用いた})
\end{aligned}$$

□

12.4 行列式の特徴づけ (証明)

n 次正方行列 $A = [a_{ij}]$ の列ベクトル表示を $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ とする. これらの各列ベクトルを単純なベクトルの和に分解する方法を考えよう. そこで, 長さ 1 の単位列ベクトルたち

$$\mathbf{e}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^n$$

を与えておく. このとき, 各 $i = 1, \dots, n$ において

$$\mathbf{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \vdots \\ a_{ni} \end{bmatrix} = \begin{bmatrix} a_{1i} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{2i} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_{ni} \end{bmatrix} = a_{1i}\mathbf{e}_1 + \cdots + a_{ni}\mathbf{e}_n = \sum_{k_i=1}^n a_{k_i i} \mathbf{e}_{k_i} \quad (12.4.1)$$

である. 上の分解を用いて次の定理を示そう. 定理 11.3.1 は次の定理の特別な場合 ($F(E) = 1$ の場合) である.

定理 12.4.1. 写像 $F : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ が列ベクトルの組に関して多重線形性と歪対称性を満たすならば, $F(A) = \det A \cdot F(E)$ である.

Proof. 各 $A \in M_n(\mathbb{R})$ について $F(A) = |A| \cdot F(E)$ となることを示そう. そこで, $A = [a_{ij}] = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ とおく. 式 12.4.1 による分解, および補題 12.3.1 で示した多重線形性 (iii) を各列に対して適用しながら展開すると,

$$\begin{aligned} F(A) &= F(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = F\left(\sum_{k_1=1}^n a_{k_1 1} \mathbf{e}_{k_1}, \sum_{k_2=1}^n a_{k_2 2} \mathbf{e}_{k_2}, \dots, \sum_{k_n=1}^n a_{k_n n} \mathbf{e}_{k_n}\right) \\ &= \sum_{k_1=1}^n a_{k_1 1} F\left(\mathbf{e}_{k_1}, \sum_{k_2=1}^n a_{k_2 2} \mathbf{e}_{k_2}, \dots, \sum_{k_n=1}^n a_{k_n n} \mathbf{e}_{k_n}\right) \\ &= \sum_{k_1=1}^n \left(a_{k_1 1} \sum_{k_2=1}^n a_{k_2 2} F\left(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \sum_{k_n=1}^n a_{k_n n} \mathbf{e}_{k_n}\right) \right) \\ &= \cdots = \sum_{k_1=1}^n \left(a_{k_1 1} \sum_{k_2=1}^n a_{k_2 2} \left(\cdots \left(\sum_{k_{n-1}=1}^n a_{k_{n-1} n-1} \left(\sum_{k_n=1}^n a_{k_n n} F(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}) \right) \right) \cdots \right) \right) \\ &= \sum_{k_1=1}^n \left(\sum_{k_2=1}^n a_{k_1 1} a_{k_2 2} \left(\cdots \left(\sum_{k_{n-1}=1}^n a_{k_{n-1} n-1} \left(\sum_{k_n=1}^n a_{k_n n} F(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}) \right) \right) \cdots \right) \right) \\ &= \cdots = \sum_{k_1=1}^n \sum_{k_2=1}^n \cdots \sum_{k_{n-1}=1}^n \sum_{k_n=1}^n a_{k_1 1} a_{k_2 2} \cdots a_{k_{n-1} n-1} a_{k_n n} F(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}) \\ &= \sum_{1 \leq k_1, k_2, \dots, k_n \leq n} a_{k_1 1} a_{k_2 2} \cdots a_{k_n n} F(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}). \end{aligned}$$

いま得られた総和の各項において, $\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}$ の中に互いに同じものが一つでもあれば, 補題 12.3.2 より $F(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}) = 0$ であるから, そのような項は無視して構わない. 一方, 無視できない各項において, $\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}$ はいずれも異なる単位列ベクトルであり, とくに k_1, k_2, \dots, k_n に重複はない.

したがって, $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$ は置換である. このとき, 補題 12.3.3 より

$$\begin{aligned} a_{k_1 1} a_{k_2 2} \cdots a_{k_n n} F(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}) &= a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} F(\mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(2)}, \dots, \mathbf{e}_{\sigma(n)}) \\ &= a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \operatorname{sgn}(\sigma) F(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) \\ &= \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} F(E). \end{aligned}$$

以上より,

$$\begin{aligned} F(A) &= \sum_{\sigma \in \mathfrak{S}_n} (\operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} F(E)) \\ &= \left(\sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \right) F(E) = |{}^t A| \cdot F(E) = |A| \cdot F(E). \end{aligned}$$

なお、上の二行目にある括弧内の式 $\sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$ が $|A|$ に等しいことは式 (12.1.1) による。 \square

12.5 命題 11.1.1 の証明

次の定理における $r = 1$ の場合が命題 11.1.1 に相当する。

定理 12.5.1. A を r 次正方行列, X を s 次正方行列とすれば,

$$\begin{vmatrix} A & * \\ O_{s,r} & X \end{vmatrix} = |A| \cdot |X| = \begin{vmatrix} A & O_{r,s} \\ * & X \end{vmatrix}.$$

Proof. 一方の等式を示せば, それを転置することで他方の等式も得られる. ここでは右側の等式を示そう. 始めに, X が単位行列 E_s の場合について考える. そこで, $Y = \begin{bmatrix} A & O_{r,s} \\ * & E_s \end{bmatrix} = [a_{ij}]$ と置こう. Y は $n = r + s$ 次正方行列である. 行列式の定義によれば

$$|Y| = \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{r\sigma(r)} a_{r+1\sigma(r+1)} \cdots a_{n\sigma(n)} \quad (12.5.1)$$

である. ここで, 総和に現れる各項

$$\text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{r\sigma(r)} a_{r+1\sigma(r+1)} \cdots a_{n\sigma(n)} \quad (12.5.2)$$

について, この値が 0 でないための必要条件を検討しよう. まず, $O_{r,s}$ 成分に着目すると, $1 \leq i \leq r$ かつ $r+1 \leq j \leq n$ ならば $a_{ij} = 0$ である. よって $\{\sigma(1), \dots, \sigma(r)\}$ の中に r より大きな数 $\sigma(i)$ が一つでもあれば, $a_{i\sigma(i)} = 0$ ゆえこの置換 σ に関する式 (12.5.2) の値は 0 である. したがって, 式 (12.5.1) の総和における $\sigma \in \mathfrak{S}_n$ の動く範囲は,

$$\{\sigma(1), \dots, \sigma(r)\} = \{1, \dots, r\}$$

を満たす部分のみを考えればよい. また, 上の条件を満たす σ においては

$$\{\sigma(r+1), \dots, \sigma(n)\} = \{r+1, \dots, n\}$$

となっている. ここで Y の E_s 成分に着目すると, もし $\sigma(i) \neq i$ なる $i = r+1, \dots, n$ があれば Y の $(i, \sigma(i))$ -成分は E_s の対角成分ではないから $a_{i\sigma(i)} = 0$ である. つまり式 (12.5.2) の値が 0 とならないためには, 各 $i = r+1, \dots, n$ について $\sigma(i) = i$ となる必要があり, したがって σ は次の形に限られる:

$$\sigma = \begin{pmatrix} 1 & \cdots & r & r+1 & \cdots & n \\ k_1 & \cdots & k_r & r+1 & \cdots & n \end{pmatrix}, \quad \text{ただし } k_1, \dots, k_r \text{ は } r \text{ 以下の数による順列.}$$

これは σ が \mathfrak{S}_r の元とみなせることに他ならない. 以上より, 式 (12.5.1) の総和における σ が動く範囲は \mathfrak{S}_r 上のみを考えればよいことから,

$$\begin{aligned} \begin{vmatrix} A & O_{r,s} \\ * & E \end{vmatrix} &= |Y| = \sum_{\sigma \in \mathfrak{S}_r} \text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{r\sigma(r)} a_{r+1,r+1} \cdots a_{n,n} \\ &= \sum_{\sigma \in \mathfrak{S}_r} \text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{r\sigma(r)} \cdot 1 \cdots 1 = |A|. \end{aligned}$$

さて, 今度は X が一般の s 次正方行列の場合について考えよう. s 個の s 次列ベクトルの組 $X = [\mathbf{x}_1, \dots, \mathbf{x}_s]$ たち全体を定義域とする関数 $F: M_s(\mathbb{R}) \rightarrow \mathbb{R}$ を

$$F(\mathbf{x}_1, \dots, \mathbf{x}_s) := \begin{vmatrix} A & O_{r,s} \\ * & X \end{vmatrix}$$

と定めよう. すると, 行列式の列に関する多重線形性と歪対称性から, F の多重線形性と歪対称性を導くことができる. また, 既におこなった議論により $F(E) = |Y| = |A|$ である. したがって定理 12.4.1 より $F(X) = |X| \cdot F(E) = |X| \cdot |A|$. 以上により求める等式が示された. \square

例 12.5.2. A_1, \dots, A_n を正方行列とすると (各々のサイズは異なっても構わない),

$$\begin{vmatrix} A_1 & & & & \\ & A_2 & & * & \\ & & \ddots & & \\ O & & & & A_n \end{vmatrix} = |A_1| \cdot \begin{vmatrix} A_2 & & & & \\ & A_3 & & * & \\ & & \ddots & & \\ O & & & & A_n \end{vmatrix} = \dots = |A_1| \cdot |A_2| \cdot \dots \cdot |A_n|.$$

12.6 $|AB| = |A| \cdot |B|$ の証明

一般に, (m, n) -行列 A および (n, ℓ) -行列 $B = [\mathbf{b}_1, \dots, \mathbf{b}_\ell]$ に対して, $AB = [A\mathbf{b}_1, \dots, A\mathbf{b}_\ell]$ である (例 3.5.1). この事実注意到しながら次の定理を示そう.

定理 12.6.1. 積の行列式は行列式の積に等しい. すなわち, n 次正方行列 A, X について $|AX| = |A| \cdot |X|$.

Proof. n 次正方行列 $X = [\mathbf{x}_1, \dots, \mathbf{x}_n]$ たち全体を定義域とする写像 $F : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ を $F(\mathbf{x}_1, \dots, \mathbf{x}_n) := |AX|$ と定める. このとき,

$$F(\mathbf{x}_1, \dots, \mathbf{x}_n) = |A[\mathbf{x}_1, \dots, \mathbf{x}_n]| = |A\mathbf{x}_1, \dots, A\mathbf{x}_n|$$

であるから, F は多重線形性と歪対称性を満たす. 実際, 歪対称性は明らかであり, 多重線形性も次のように確認できる:

$$\begin{aligned} F(\mathbf{x}_1, \dots, \mathbf{b}_i + \mathbf{c}_i, \dots, \mathbf{x}_n) &= |A\mathbf{x}_1, \dots, A(\mathbf{b}_i + \mathbf{c}_i), \dots, A\mathbf{x}_n| = |A\mathbf{x}_1, \dots, A\mathbf{b}_i + A\mathbf{c}_i, \dots, A\mathbf{x}_n| \\ &= |A\mathbf{x}_1, \dots, A\mathbf{b}_i, \dots, A\mathbf{x}_n| + |A\mathbf{x}_1, \dots, A\mathbf{c}_i, \dots, A\mathbf{x}_n| \\ &= F(\mathbf{x}_1, \dots, \mathbf{b}_i, \dots, \mathbf{x}_n) + F(\mathbf{x}_1, \dots, \mathbf{c}_i, \dots, \mathbf{x}_n). \end{aligned}$$

$$\begin{aligned} F(\mathbf{x}_1, \dots, r\mathbf{x}_i, \dots, \mathbf{x}_n) &= |A\mathbf{x}_1, \dots, A(r\mathbf{x}_i), \dots, A\mathbf{x}_n| = |A\mathbf{x}_1, \dots, r(A\mathbf{x}_i), \dots, A\mathbf{x}_n| \\ &= r|A\mathbf{x}_1, \dots, A\mathbf{x}_i, \dots, A\mathbf{x}_n| = rF(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n). \end{aligned}$$

ゆえに定理 12.4.1 より $F(X) = |X| \cdot F(E) = |X| \cdot |AE| = |X| \cdot |A| = |A| \cdot |X|$ である. 以上により $|AX| = |A| \cdot |X|$ が示された. \square

次は 7.1 節で既に述べたことであるが念のためもう一度記しておこう.

命題 12.6.2. A が可逆行列ならば $|A| \neq 0$ であり, とくに $|A^{-1}| = |A|^{-1}$.

Proof. $1 = |E| = |AA^{-1}| = |A| \cdot |A^{-1}|$ から $|A| \neq 0$ を得る. また, この両辺を $|A|$ で割ることで $|A^{-1}| = |A|^{-1}$ を得る. \square

上の命題の逆「 $|A| \neq 0$ ならば A は可逆である」もまた正しい. その証明は次章で与えよう (定理 13.2.3).

第13章 余因子展開とクラメルの公式

11章における行列式の値の計算においては、サイズが一回り小さい行列式の計算に帰着させる手法を繰り返し用いたのであった。この手法を理論的な立場から述べたものが余因子展開である。本章では、余因子展開を通して、2次正方形行列の逆行列の公式:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{について,} \quad A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

の一般化を与える。そして、この逆行列の表示から、これまで証明を避けてきた定理 5.3.3 が導かれる。

13.1 余因子展開

定義 13.1.1. n 次正方形行列 $A = [a_{ij}]$ の第 i 行と第 j 列を取り除いた残りの成分からなる $n-1$ 次正方形行列を A_{ij} とかく。すなわち、

$$A = \begin{pmatrix} A & \begin{matrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{nj} \end{matrix} & B \\ \begin{matrix} a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \end{matrix} & & \\ C & & D \end{pmatrix} \quad \text{のとき} \quad A_{ij} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

例 13.1.2. $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ とすれば、 $A_{12} = \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix}$, $A_{22} = \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix}$ である。

列ベクトルによる成分表示を $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ とすれば、式 12.4.1 と同様にして、各 \mathbf{a}_j は次のように書ける:

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} = a_{1j}\mathbf{e}_1 + a_{2j}\mathbf{e}_2 + \cdots + a_{nj}\mathbf{e}_n.$$

したがって、補題 12.3.1 で示した行列式の多重線形性 (2) により、 $|A|$ は次の和に分解できる:

$$\begin{aligned} |A| &= |\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, a_{1j}\mathbf{e}_1 + a_{2j}\mathbf{e}_2 + \cdots + a_{nj}\mathbf{e}_n, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n| \\ &= a_{1j} |\mathbf{a}_1, \dots, \mathbf{e}_1, \dots, \mathbf{a}_n| + a_{2j} |\mathbf{a}_1, \dots, \mathbf{e}_2, \dots, \mathbf{a}_n| + \cdots + a_{nj} |\mathbf{a}_1, \dots, \mathbf{e}_n, \dots, \mathbf{a}_n|. \end{aligned} \quad (13.1.1)$$

更に, 式 13.1.1 に現れる i 番目の項を計算すると

$$\begin{aligned}
 a_{ij} \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & 1 & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} &= a_{ij}(-1)^{i-1} \begin{vmatrix} a_{i1} & \cdots & 1 & \cdots & a_{in} \\ a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} & \quad (i \text{ 行を } 1 \text{ 行へ移動させた}) \\
 &= a_{ij}(-1)^{i-1}(-1)^{j-1} \begin{vmatrix} 1 & a_{i1} & \cdots & \cdots & a_{in} \\ 0 & a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} & \quad (j \text{ 列を } 1 \text{ 列へ移動させた}) \\
 &= a_{ij}(-1)^{i+j-2} \begin{vmatrix} 1 & a_{i1} & \cdots & a_{in} \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{vmatrix} = a_{ij}(-1)^{i+j} |A_{ij}|.
 \end{aligned}$$

以上より, 式 13.1.1 は次のように書き下される:

$$|A| = a_{1j}(-1)^{1+j}|A_{1j}| + a_{2j}(-1)^{2+j}|A_{2j}| + \cdots + a_{nj}(-1)^{n+j}|A_{nj}| = \sum_{i=1}^n a_{ij}(-1)^{i+j}|A_{ij}|.$$

上の各項に現れる $(-1)^{i+j}|A_{ij}|$ を A の (i, j) -余因子 (cofactor) と呼び, この展開式のことを j 列に関する余因子展開 (cofactor expansion) という.

練習 13.1.3. j 列の代わりに i 行に関して同様の展開を行い, 次の等式を示せ.

$$|A| = a_{i1}(-1)^{i+1}|A_{i1}| + a_{i2}(-1)^{i+2}|A_{i2}| + \cdots + a_{in}(-1)^{i+n}|A_{in}| = \sum_{j=1}^n a_{ij}(-1)^{i+j}|A_{ij}|.$$

上式を i 行に関する余因子展開という.

成分に 0 が多く現れる行列式の値を求める際に余因子展開は有効である. 例えば命題 11.1.1 おける左の等号は 1 列目に関する余因子展開を, 右の等号は 1 行目に関する余因子展開を意味する.

例 13.1.4. 次の k 次正方行列について,

$$\begin{vmatrix} t & -1 & & & \\ & t & -1 & & O \\ & & \ddots & \ddots & \\ O & & & t & -1 \\ a_0 & a_1 & \cdots & a_{k-2} & t + a_{k-1} \end{vmatrix} = t^k + a_{k-1}t^{k-1} + \cdots + a_1t + a_0.$$

Proof. k に関する帰納法によって証明する. $k = 2$ の場合は次の計算で確かめられる:

$$\begin{vmatrix} t & -1 \\ a_0 & t + a_1 \end{vmatrix} = t(t + a_1) - a_0 \cdot (-1) = t^2 + a_1t + a_0.$$

サイズが $k-1$ のときに等式が成立すると仮定して、サイズが k の場合の等式を示そう。1 列目に関して左辺を余因子展開すると、第 1 項目に帰納法の仮定が適用できる:

$$\begin{aligned}
 (\text{左辺}) &= t \begin{vmatrix} t & -1 & & & \\ & \ddots & \ddots & & \\ & & t & -1 & \\ a_1 & \cdots & a_{k-2} & t+a_{k-1} & \end{vmatrix} + (-1)^{k+1} a_0 \begin{vmatrix} -1 & & & & \\ t & -1 & & & \\ & \ddots & \ddots & & \\ & & & t & -1 \end{vmatrix} \\
 &= t(t^{k-1} + a_{k-1}t^{k-2} + \cdots + a_2t + a_1) + (-1)^{k+1} a_0 \cdot (-1)^{k-1} = (\text{右辺}).
 \end{aligned}$$

なお、第 2 項目の行列式の値は、転置して例 11.1.2 を適用することで分かる。 \square

13.2 余因子行列

正方行列 A の逆行列の成分表示を与える方法を考えよう。 $C = A^{-1}$ ならば $CA = E$ を満たす。ゆえに C を行ベクトル分割し A を列ベクトル分割すれば次を満たすことになる。

$$\begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_n \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1\mathbf{a}_1 & \mathbf{c}_1\mathbf{a}_2 & \cdots & \mathbf{c}_1\mathbf{a}_n \\ \mathbf{c}_2\mathbf{a}_1 & \mathbf{c}_2\mathbf{a}_2 & \cdots & \mathbf{c}_2\mathbf{a}_n \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{c}_n\mathbf{a}_1 & \mathbf{c}_n\mathbf{a}_2 & \cdots & \mathbf{c}_n\mathbf{a}_n \end{bmatrix} = E.$$

すなわち、

$$\mathbf{c}_j\mathbf{a}_i = \begin{cases} 1 & j = i \text{ のとき,} \\ 0 & j \neq i \text{ のとき.} \end{cases}$$

上式を満たす行ベクトル \mathbf{c}_j を余因子展開を用いて導こう。

行列 $A = [a_{ij}]$ の第 j 列を別の列ベクトル \mathbf{b} に置き換えた行列 B について、 j 列に関する余因子展開を行えば、各 $i = 1, \dots, n$ について $B_{ij} = A_{ij}$ であるから、

$$\begin{vmatrix} a_{11} & \cdots & \overset{j}{\mathbf{b}_1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & \mathbf{b}_i & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & \mathbf{b}_n & \cdots & a_{nn} \end{vmatrix} = (-1)^{1+j} b_1 |A_{1j}| + (-1)^{2+j} b_2 |A_{2j}| + \cdots + (-1)^{n+j} b_n |A_{nj}|.$$

この右辺は次の二つのベクトル

$$\tilde{\mathbf{a}}_j = \begin{bmatrix} (-1)^{1+j} |A_{1j}| \\ (-1)^{2+j} |A_{2j}| \\ \vdots \\ (-1)^{n+j} |A_{nj}| \end{bmatrix} \quad \text{および} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

の内積と見ることができる。内積を行列の積として表現するには片方を転置する必要がある、ここでは

$\tilde{\mathbf{a}}_j$ を転置して

$$\begin{aligned} & \begin{array}{c} j \\ \text{列} \\ \left| \begin{array}{cccc} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & b_i & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{array} \right| \end{array} = \begin{bmatrix} (-1)^{1+j}|A_{1j}| & (-1)^{2+j}|A_{2j}| & \cdots & (-1)^{n+j}|A_{nj}| \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ & = {}^t\tilde{\mathbf{a}}_j \mathbf{b} \end{aligned} \quad (13.2.1)$$

を得る. とくに \mathbf{b} として A の各列 \mathbf{a}_i を取ると, 上式の左辺は $i = j$ の場合を除き i 列と j 列が等しい行列ゆえ 0 となり, $i = j$ の場合は $|A|$ である. すなわち,

$${}^t\tilde{\mathbf{a}}_j \mathbf{a}_i = \begin{cases} |A| & j = i \text{ のとき,} \\ 0 & j \neq i \text{ のとき.} \end{cases} \quad (13.2.2)$$

そこで, 行ベクトル ${}^t\tilde{\mathbf{a}}_1, {}^t\tilde{\mathbf{a}}_2, \dots, {}^t\tilde{\mathbf{a}}_n$ を成分とする行列を \tilde{A} としよう:

$$\tilde{A} := \begin{bmatrix} {}^t\tilde{\mathbf{a}}_1 \\ {}^t\tilde{\mathbf{a}}_2 \\ \vdots \\ {}^t\tilde{\mathbf{a}}_n \end{bmatrix} = \begin{bmatrix} (-1)^{1+1}|A_{11}| & (-1)^{2+1}|A_{21}| & \cdots & (-1)^{n+1}|A_{n1}| \\ (-1)^{1+2}|A_{12}| & (-1)^{2+2}|A_{22}| & \cdots & (-1)^{n+2}|A_{n2}| \\ \vdots & \vdots & \vdots & \vdots \\ (-1)^{1+n}|A_{1n}| & (-1)^{2+n}|A_{2n}| & \cdots & (-1)^{n+n}|A_{nn}| \end{bmatrix}. \quad (13.2.3)$$

ここで, $\tilde{A} = [a_{ij}^*]$ と成分表示すれば $a_{ij}^* = (-1)^{j+i}|A_{ji}|$ である (転置を取ったため添え字 i, j の位置が入れ替わっていることに注意せよ). すると式 (13.2.2) より $\tilde{A}A = |A|E$ を得る:

$$\tilde{A}A = \begin{bmatrix} {}^t\tilde{\mathbf{a}}_1 \\ {}^t\tilde{\mathbf{a}}_2 \\ \vdots \\ {}^t\tilde{\mathbf{a}}_n \end{bmatrix} [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n] = \begin{bmatrix} |A| & & & \\ & |A| & & \\ & & \ddots & \\ & & & |A| \end{bmatrix} = |A|E.$$

練習 13.2.1. 行に関する余因子展開に対して今と同様の議論を行うことで, $A\tilde{A} = |A|E$ を示せ.

定義 13.2.2. n 次正方行列 $A = [a_{ij}]$ に対して, 本節で定義した $\tilde{A} = [a_{ij}^*]$ (ただし $a_{ij}^* := (-1)^{j+i}|A_{ji}|$) を A の余因子行列 (**adjugate matrix**) という.

【補足】 A の (i, j) -余因子を (i, j) 成分とする行列, すなわち \tilde{A} の転置行列のことを **cofactor matrix** と言う. これを直訳すると「余因子行列」となるが, 邦語の文献では adjugate matrix のことを余因子行列と呼ぶのが慣例となっており, 本書もこれに従った.

これまでの議論により, $\tilde{A}A = A\tilde{A} = |A|E$ である. とくに $|A| \neq 0$ の場合は, 次の逆行列の公式を得る:

定理 13.2.3. 正方行列 A について, $|A| \neq 0$ ならば A は可逆であり $A^{-1} = \frac{1}{|A|}\tilde{A}$ である.

Proof. $\frac{1}{|A|}\tilde{A}$ が A の逆行列であることは直ちに確かめられる:

$$\frac{1}{|A|}\tilde{A} \cdot A = \frac{1}{|A|} \cdot |A|E = E, \quad A \cdot \frac{1}{|A|}\tilde{A} = \frac{1}{|A|}(A\tilde{A}) = \frac{1}{|A|} \cdot |A|E = E.$$

□

例 13.2.4. 2次正方行列について, 逆行列の公式 (定理 13.2.3) を適用してみよう. $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ とすれば,

$$\begin{aligned} a_{11}^* &= (-1)^{1+1}|A_{11}| = a_{22}, & a_{12}^* &= (-1)^{2+1}|A_{21}| = -a_{12}, \\ a_{21}^* &= (-1)^{1+2}|A_{12}| = -a_{21}, & a_{22}^* &= (-1)^{2+2}|A_{22}| = a_{11}, \end{aligned}$$

であるから, $\tilde{A} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$. 更に $|A| = a_{11}a_{22} - a_{12}a_{21} \neq 0$ のとき,

$$A^{-1} = \frac{1}{|A|}\tilde{A} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

命題 12.6.2 と定理 13.2.3 を合わせて次の同値性を得る. とくに, 条件 $|A| \neq 0$ は, 定理 6.2.2 の諸条件と同値である.

系 13.2.5. 正方行列 A について, $|A| \neq 0 \iff A$ は可逆である.

こうして定理 5.3.3 を示すための準備が整った.

定理 5.3.3 (再掲). 二つの n 次正方行列 A, D について $DA = E$ が成り立てば D は A の逆行列である. すなわち, $AD = E$ も成り立つ.

Proof. $DA = E$ ならば $|D| \cdot |A| = |E| = 1$ より, $|A| \neq 0$ である. ゆえに定理 13.2.3 より A は可逆であり, 逆行列 A^{-1} が存在する. このとき, $D = DE = D(AA^{-1}) = (DA)A^{-1} = EA^{-1} = A^{-1}$ より $D = A^{-1}$ である. \square

余因子行列を用いた逆行列の公式には, 次のような応用もある.

命題 13.2.6. 上三角行列 A について次が成り立つ.

- (1) A の余因子行列 \tilde{A} も上三角である.
- (2) A が可逆ならば, A^{-1} も上三角である.

Proof. (1): A のサイズを n とする. $A = [a_{ij}]$ とおけば, A は上三角ゆえ $n \geq i > j \geq 0$ ならば $a_{ij} = 0$ である. \tilde{A} が上三角であることをいうには, $n \geq i > j \geq 0$ を満たす各 i, j について, $a_{ij}^* = (-1)^{j+i}|A_{ji}| = 0$ を示せばよい. A が上三角であること, および $j < i$ より, A_{ji} も上三角である (図 13.1).

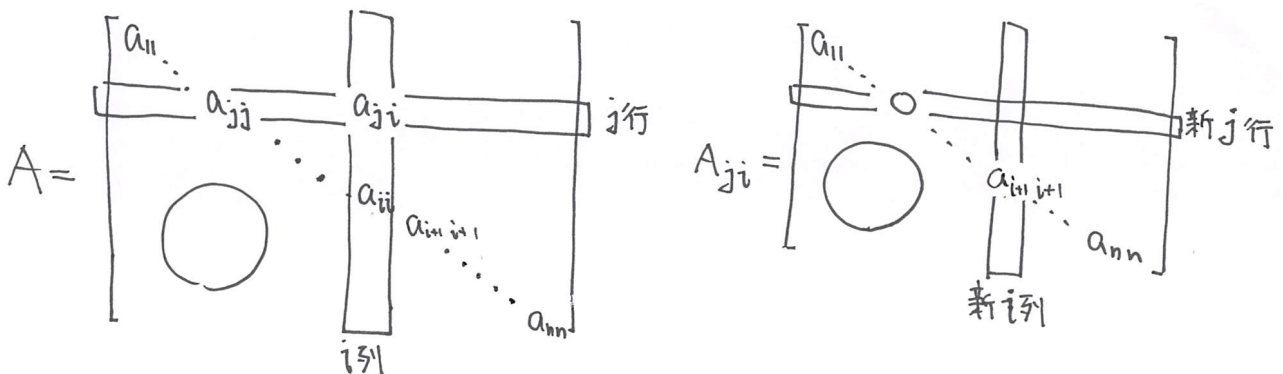


図 13.1: $1 < j < i < n$ の場合

いま, $j < n$ であるから, A_{ji} の第 j 行は, A の $j+1$ 行目から何らかの成分を一つを除いた $n-1$ 次行ベクトルである. また, A_{ji} の第 j 列は, いま $j < i$ としているゆえ, A の j 列目から何らかの成分を一つ除いた $n-1$ 次列ベクトルである. したがって, A_{ji} の (j, j) -成分は, $a_{j+1, j} = 0$. 以上より A_{ji} は上三角かつ対角成分に 0 を含むゆえ $|A_{ji}| = 0$ (例 11.1.2). つまり $a_{ij}^* = (-1)^{j+i}|A_{ji}| = 0$ であり, \tilde{A} は上三角である.

(2): $A^{-1} = \frac{1}{|A|}\tilde{A}$ ゆえ (1) より明らか. \square

発展 (無限次元の行列)

本章で証明を与えた定理 5.3.3 は有限次元の仮定の下でしか成り立たない. 次元の定義は 22 章で与えられ, ここではその詳細は控えよう. しかしながら, 成分が無限個ある行列を考えることで, 無限次元空間における定理 5.3.3 の反例を簡単に挙げるができるゆえ, ここで紹介しておく. 無限個の成分をもつ行列 A, B を次で定める:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

無限個の成分をもつ行列の間の積は, 一般には定まらない. 何故なら, 積の各成分を得るには無限個の数の和を考える必要があり, それは一般には収束せずに数が定まらないからである. しかし, 上の A, B のように, ほとんどすべての成分が 0 の場合は, 各成分の計算は有限個の和と考えられるゆえ定めることができる. BA および AB を計算すれば次のようになる:

$$BA = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = E, \quad AB = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

したがって $BA = E$ だからといって $AB = E$ とは限らない.

有限の世界で成り立つことが無限の世界で成り立たない理由の多くは, 「二つの集合において, それぞれに含まれる要素の個数が一致する」という概念を無限集合にまで拡張したとき, その様子が有限集合の場合と著しく異なる点に起因している. 詳しくは, 無限集合については 19.5 節を, 上の無限行列に対応する線形写像については例 26.1.5 を見よ.

13.3 クラメルの公式の証明

定理 13.3.1 (クラメルの公式). n 次正方行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ (ただし各 \mathbf{a}_j は n 次列ベクトル) において, A が可逆 (すなわち $|A| \neq 0$) ならば, 連立 1 次方程式 $A\mathbf{x} = \mathbf{b}$ の唯一解 $\mathbf{x} = {}^t[x_1, \dots, x_n]$ は, $\mathbf{x} = A^{-1}\mathbf{b}$ である (命題 4.3.1). このとき, 各成分 x_i は次の式で与えられる:

$$x_j = \frac{|\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n|}{|A|} \quad (j = 1, \dots, n).$$

Proof. A の余因子行列を \tilde{A} とする. 定理 13.2.3 より $A^{-1} = \frac{1}{|A|}\tilde{A}$ であり, これと式 (13.2.3) から

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A^{-1}\mathbf{b} = \frac{1}{|A|}\tilde{A}\mathbf{b} = \frac{1}{|A|} \begin{bmatrix} {}^t\tilde{\mathbf{a}}_1 \\ {}^t\tilde{\mathbf{a}}_2 \\ \vdots \\ {}^t\tilde{\mathbf{a}}_n \end{bmatrix} \mathbf{b} = \frac{1}{|A|} \begin{bmatrix} {}^t\tilde{\mathbf{a}}_1\mathbf{b} \\ {}^t\tilde{\mathbf{a}}_2\mathbf{b} \\ \vdots \\ {}^t\tilde{\mathbf{a}}_n\mathbf{b} \end{bmatrix}.$$

したがって、各成分 x_j は、

$$x_j = \frac{1}{|A|} {}^t \tilde{\mathbf{a}}_j \mathbf{b} = \frac{1}{|A|} \cdot \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & b_i & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}.$$

最後の等式は式 (13.2.1) による. □

余因子行列を持ちださずともクラメルの公式を示すことはできる. 行列式への理解を深めるために, 余因子行列を用いない別証明を与えておこう:

クラメルの公式の別証明. $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ と置けば

$$\mathbf{b} = A\mathbf{x} = [\mathbf{a}_1, \dots, \mathbf{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n$$

である. 行列式 $|\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n|$ を式 (13.1.1) と同様にして展開すると,

$$\begin{aligned} & |\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n| \\ &= |\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n| \\ &= x_1 |\mathbf{a}_1, \dots, \mathbf{a}_1, \dots, \mathbf{a}_n| + \cdots + x_j |\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n| + \cdots + x_n |\mathbf{a}_1, \dots, \mathbf{a}_n, \dots, \mathbf{a}_n|. \end{aligned}$$

ここで, 上式の第 i 項に現れる行列式は, $i = j$ の場合を除き i 列と j 列がともに \mathbf{a}_i であるから, その値は 0 である. また第 j 項は $x_j |A|$ である. 以上より,

$$|\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n| = x_j |A|.$$

これを变形して求める等式を得る. □

第III部

抽象ベクトル空間

第14章 集合概念の基礎

本章と第VI部の19章では、先に1.1および1.2節で述べた集合と写像に関する概念の発展として、これらのより繊細な使い方について解説する。ただし、線形代数学の文脈に現れる部分に限って取り上げることとした。これらの概念を扱う理念や、集合概念の基盤となるより根本的な論理の規則等については、巻末にある文献[6]を参照してもらいたい。

14.1 集合の包含関係

包含関係は、二つの集合の一致を示す際に必須となる概念である。

定義 14.1.1. 二つの集合 A, B が与えられており、 A のいかなる元も B に属するとき、 A は B の部分集合 (subset) であるといい、 $A \subset B$ あるいは $B \supset A$ と表す。

とくに、 A 自身は A の部分集合である。 $A \subset B$ であるとき「 A は B に含まれる」と述べることもある。この表現は、 A が B の元であること (つまり $A \in B$) と誤解される恐れもあるゆえ注意したい。

例 14.1.2. (1) $\{\text{りんご}, \text{スイカ}\} \subset \{\text{りんご}, \text{みかん}, \text{スイカ}\}$ である。

(2) $\mathbb{N} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ である。これらの記号の意味は1.1節を参照せよ。

(3) $A = \{1\}$ (一点のみの集合) とし、 $X = \{1, A\}$ (自然数1と集合 A の二つの元からなる集合) とすれば、 $A \in X$ と $A \subset X$ が共に成り立つ。

一方、集合 A が集合 B の部分集合でないとは、 A のすべての元が B に属するわけではないこと、つまり B に属さない A の元が存在することを意味する。このとき、 $A \not\subset B$ と書く。

例 14.1.3. 集合 $A = \{2, 9, 11, 30\}$ は集合 $B = \{2, 9, 15, 26, 30, 37\}$ の部分集合ではない。何故なら $11 \in A$ および $11 \notin B$ であり、 A のすべての元が B に属するわけではないからである。

次で定める特別な集合は全ての集合の部分集合となる:

定義 14.1.4. いかなる元も含まない集合を空集合 (empty set) とよび、これを \emptyset と書く。

命題 14.1.5. 集合 X に対して、空集合 \emptyset は X の部分集合である。

Proof. 背理法で示す。 \emptyset が X の部分集合でないと仮定しよう。このとき部分集合の定義により、 X に属さない \emptyset の元 a が存在する。とくに $a \in \emptyset$ であり、これは \emptyset が元を含まないことに矛盾する。ゆえに \emptyset は X の部分集合である。□

二つの集合 A, B が等しいとは、 A を構成する元と B を構成する元とが一致するということである。これは、 A の元は B の元でもあり、また B の元は A の元でもあることにほかならない。すなわち、次が成り立つ:

$$\text{集合 } A, B \text{ について, } \quad A = B \iff A \subset B \text{ かつ } B \subset A. \quad (14.1.1)$$

例 14.1.6. 集合 $A = \{\text{りんご}, \text{スイカ}\}$ と集合 $B = \{\text{りんご}, \text{スイカ}, \text{スイカ}\}$ は等しい。実際、包含関係の定義14.1.1によれば $A \subset B$ および $B \subset A$ が成り立つ。したがって式(14.1.1)より $A = B$ である。集合 B にスイカが二つ入っているわけではないことに注意しよう。スイカを二つ含む集合を考えたいのであれば、例えばスイカ1、スイカ2とラベルを貼り、 $\{\text{りんご}, \text{スイカ1}, \text{スイカ2}\}$ と書けばよい。

例 14.1.7. 本章以前の議論においても、式 (14.1.1) の左向き「 $A \subset B$ かつ $B \subset A \implies A = B$ 」を暗黙裡のうちに何度か用いていた。

- (1) 例 1.3.3 およびその後の議論において、 \mathbb{R}^2 から \mathbb{R}^2 への線形写像全体の集合 X と $f \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$ なる形で表せる写像 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 全体の集合 Y が一致することを確認した。例 1.3.3 において $Y \subset X$ を述べて、その後の議論において $X \subset Y$ を示している。
- (2) 命題 4.3.1 の説明では、方程式 $Ax = b$ の解全体の集合 X と一点からなる集合 $Y = \{Bb\}$ が一致することを述べている。まず、 a を方程式の解（つまり $a \in X$ ）とすれば、 a は Bb に一致すること（つまり $a \in Y$ ）を示した。これは $X \subset Y$ を示すことに相当している。次に Bb が方程式の解であることを確認した。これは $Y \subset X$ を示すことに他ならない。
- (3) 一般の連立 1 次方程式の解法 (4.6 節) における一般解の表示についても同様のことを行った。方程式 $Ax = b$ の解全体の集合 X と $a_0 + c_1 a_1 + c_2 a_2 + c_3 a_3$ で表されるベクトル全体の集合 Y が一致することを、 $X \subset Y$ および $Y \subset X$ の両方を確認することによって示している。

よりみち (集合が等しいとはどういうことか)

二つの集合が一致するとはどういうことか改めて考えてみると、雲をつかむような、とりとめもない思索しかできないことに気づく。先程、式 (14.1.1) が成立することをもっともらしく述べたが、実は「 $A \subset B$ かつ $B \subset A \implies A = B$ 」の説明はしていない。これは本当に正しい事実だろうか。例えば、赤い袋 A の中に自然数 1 と 2 のみが入っていると、青い袋 B にも 1 と 2 のみが入っているとしよう。この二つの袋は色が違っているにも関わらず一致していると言えるのだろうか。「集合と考えている場合は一致する」と言いたいところではあるけれども、その根拠に式 (14.1.1) を用いるわけにはいかない。何故なら、いま式 (14.1.1) を説明するための議論をしているからである。では、どうやって A と B が一致することを導けばよいのだろうか。

このように、集合が一致することを説明するのは意外に難しいのである。そこで集合論では、式 (14.1.1) を公理として定め、外延性公理と呼んでいる。より素朴な立場では、式 (14.1.1) が集合が等しいことの定義であると考えてもよいだろう。

14.2 集合の表し方

数学に限らず、何かしらの概念を規定しようと思うと、大きく分けて二通りの方法があることに気づく。新たな概念 A を規定するにあたり、 A であるものをすべて列挙する方法を外延的な定義といい、 A が持っている性質によって規定する方法を内包的な定義という。例えば、正多面体の定義として、次の二通りの述べ方がある：

- **外延的定義:** 正四面体および正六面体、正八面体、正十二面体、正二十面体なる図形を総称して正多面体という。
- **内包的定義:** 各頂点が同じ数の面と接し、すべての面が合同な正多角形となる多面体¹を正多面体という。

集合の定め方についてもこのことは例外ではなく、外延的な記述と内包的な記述の両方が用いられる。これは、次の二つの行為に本質的な違いがないことから読者も容易に想像がつくことと思う。

- 新しい概念 A を定める、
- 概念 A が指すものの全体の集合を与える。

¹より正確には、ここでは凸多面体のみを考えている。

ここでは、集合の外延的な記述と内包的な記述について例を挙げながら解説したい。これから挙げる例で見るように、集合を規定する場合、中括弧“{, }”を用いることが慣例となっている。また、集合を定める中括弧内で用いられる区切りの記号として、本書では“|”を用いる。文献によっては、区切りの記号にコロン“:”を用いる場合もある。

まず、外延的記法の例を挙げてみよう。

例 14.2.1. 次の例はいずれも外延的な記法である。

- (1) りんご、みかん、スイカの3つの要素からなる集合を { りんご, みかん, スイカ } と書く。
- (2) あらかじめ数列 a_n (ただし $n \in \mathbb{N}$) が与えられているとき、 a_1, a_2, a_3, \dots をすべて集めた集合のことを

$$\{a_n \mid n \in \mathbb{N}\}$$

と表す。

- (3) 正の偶数全体の集合は次のように表される:

$$\{2, 4, 6, 8, 10, 12, \dots\}, \quad \text{あるいは} \quad \{2n \mid n \in \mathbb{N}\}.$$

ただし、最初の記法は、12以降の列に偶数がもれなく現れるという保証はどこにも書かれておらず、曖昧な表記と言える。出来る限り誤解を避けたいのであれば、後の記法のほうが望ましい。

- (4) 写像 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) = x^2$ で定める。このとき、定義域 \mathbb{R} の元を f に代入した値 $f(x)$ の範囲を表す集合を次のように書く:

$$\{x^2 \mid x \in \mathbb{R}\}, \quad \text{あるいは} \quad \{f(x) \mid x \in \mathbb{R}\}.$$

この集合は0以上の実数全体の集合に一致する。

以上のように外延的記法には様々な変種があり、この記法の形式的定義および使い方を統一的に説明することは難しい。強いて定めるとすれば、上の(4)を念頭に「ある写像の像として定められる集合」となる。写像の像については定義19.1.1を見よ。

一方、内包的な記法は次のように形式的に説明することができる:

定義 14.2.2 (内包的記法). あらかじめ集合 X が与えられており、更に、 X に属する元 x たちに関する条件 $P(x)$ が与えられているとする²。このとき、 X の元のうち $P(x)$ が成立する元のみを全て集めた集合、すなわち X において $P(x)$ が成立する範囲を

$$\{x \in X \mid P(x)\}$$

と書く³。

例 14.2.3. 次の集合の表し方はいずれも内包的である。

- (1) 集合 $\{x \in \mathbb{N} \mid x = 2m \text{ を満たす } m \in \mathbb{N} \text{ が存在する}\}$ は正の偶数全体の集合である。
- (2) (m, n) -行列 A および m 次列ベクトル \mathbf{b} に対して集合 $\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{b}\}$ は連立1次方程式 $A\mathbf{x} = \mathbf{b}$ の解全体からなる集合である。
- (3) \mathbb{R}^2 の部分集合 $\{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} = (x, y) \text{ と置くと, } x > 0 \text{ かつ } y > 0\}$ を第1象限という。
- (4) 集合 $\{x \in \mathbb{R} \mid x \neq x\}$ は空集合 \emptyset に等しい。条件 $x \neq x$ を満たす $x \in \mathbb{R}$ が一つもないからである。

² P の例は3.2節を参照せよ。

³数理論理学では、条件「 $P(x)$ かつ $x \in X$ 」を改めて $Q(x)$ と置くことにより、この集合を $\{x \mid Q(x)\}$ と表す。

定義 14.2.4. a, b を実数とする. a 以上かつ b 以下の実数をすべて集めた集合を $[a, b]$ と書き, これを閉区間と呼ぶ. a より大きくかつ b 未満の実数をすべて集めた集合を (a, b) と書き, これを開区間と呼ぶ. また, a 以上 b 未満の実数全体の集合を $[a, b)$, a より大きく b 以下の実数全体の集合を $(a, b]$ と書き, これらを半開区間と呼ぶ. 更に, a 以上の実数全体を $[a, \infty)$, a より大きい実数全体を (a, ∞) と書き, a 以下の実数全体, および a 未満の実数全体をそれぞれ $(-\infty, a]$, $(-\infty, a)$ と書く.

定義 14.2.4 に現れた集合に $\mathbb{R} = (-\infty, \infty)$ を加えたものを総称して区間と呼ぶ. 一点集合 $\{a\} = [a, a]$ や空集合 $\emptyset = [2, 1]$ (2 以上かつ 1 以下の数は存在しない) も区間である.

練習 14.2.5. 定義 14.2.4 で与えたそれぞれの区間について, これを集合の内包的な表記で表せ.

解答例 (抜粋): $[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$, $(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$, $[a, \infty) = \{x \in \mathbb{R} \mid x \geq a\}$.

発展 (もう一つの区間の定義)

定義 14.2.4 は, 区間をすべて列挙した定め方であるから外延的と言える. 一方で, 条件「 $a, b \in I$ かつ $a < x < b \implies x \in I$ 」を満たす \mathbb{R} の部分集合 I のこととして区間を定義する流儀もある (内包的な区間の定義). これらの定義が一致することは明らかではなく, 証明には実数の連続性の公理を要する. 詳しくは解析の本を参照せよ.

14.3 外延的か内包的か

集合の表記が外延的なものか内包的なものかは文脈で判断すること. なかには外延的とも内包的ともとれる記法がある:

例 14.3.1. 集合 X に対して, X の元を並べた四つ組 (a, b, c, d) たち全体のなす集合を X^4 とし, X の元を成分とする 2 次正方形行列全体のなす集合を $M_2(X)$ で表す. このとき, 次の記法はいずれも $M_2(\mathbb{Z})$ を表す.

(1) 写像 $F: \mathbb{R}^4 \rightarrow M_2(\mathbb{R})$ を $F(a, b, c, d) := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ と定めておく. $M_2(\mathbb{Z}) = \{F(\mathbf{x}) \mid \mathbf{x} \in \mathbb{Z}^4\}$.

(2) 上の \mathbf{x} や $F(\mathbf{x})$ を具体的に成分表示した表記: $M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid (a, b, c, d) \in \mathbb{Z}^4 \right\}$.

(3) $M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z} \right\}$, $M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}) \mid a, b, c, d \in \mathbb{Z} \right\}$.

(4) $M_2(\mathbb{Z}) = \left\{ A \in M_2(\mathbb{R}) \mid A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ をみたす } a, b, c, d \in \mathbb{Z} \text{ が存在する} \right\}$.

(1) が外延的記法で (4) が内包的記法である. この中間の記法として (2) や (3) の記法もしばしば用いられる. これらを外延的とするか内包的とするかは記述者の立場に委ねられよう.

集合の表示が外延的か内包的かを厳密に分類したいのであれば, 上の例における (2) や (3) のような表記を認めないことと約束し, 記号 “|” の左側に現れる集合の元を表す記号が裸のまま用いられていれば内包的, 文字が添え字づけられていれば外延的とすればよい. ここで, 元を表す記号が裸であるとは, $x \in X$ というように, 元を表す記号が一つの文字 x のみからなる場合を指す. また, 添え字づけられているとは, x_n あるいは x^2 , $f(x)$, x_f といったように, 元そのものが複数の文字 (または記号) を用いて表されていることを指す. なお, 添え字づけられた元は, ある写像によって代入された値であるとみなすことができる⁴.

⁴例えば実数列 a_n の各項は, $n \in \mathbb{N}$ に対して $a_n \in \mathbb{R}$ を対応させる写像 $f: \mathbb{N} \rightarrow \mathbb{R}$ を用いることで, f に各自然数を代入した値と見なせる.

しかしながら、記述された集合が意味するところに誤解がないのであれば、例 14.3.1 における (2) あるいは (3) のような表記を認めても数学的議論に支障はなく、多くの文献でこのような記法が用いられている。本書もこれに準じる。

練習 14.3.2. 次の集合の記法は外延的か内包的か答えよ。

$$(1) \{x \in \mathbb{Z} \mid x \text{ は奇数}\}, \quad (2) \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}, \quad (3) \{(x, \sin x) \mid x \in \mathbb{R}\}.$$

解答例: (1) と (2) が内包的表記であり、(3) が外延的表記である。ただし、(2) は例 14.3.1 でいうところの中間の記法ともみなせる。なお、(1) は奇数全体の集合、(2) は \mathbb{R}^2 上の原点を中心とする半径 1 の円周上の点全体の集合、(3) は $f(x) := \sin x$ で定められる関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ のグラフ上の点全体の集合を表す。

方程式を解くとは、内包的記述を外延的記述に書き直す行為にほかならない:

練習 14.3.3. 次で与えられる集合 X の外延的表記を与えよ。

$$(1) X = \{x \in \mathbb{R} \mid x^2 - 1 = 0\}.$$

解答例: $x^2 - 1 = 0$ を満たす数は $x = 1$ および $x = -1$ である。ゆえに $X = \{1, -1\}$ 。

$$(2) X = \left\{ \mathbf{x} \in \mathbb{R}^5 \mid \begin{bmatrix} 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix} \right\}.$$

解答例: この連立 1 方程式の解法は 4.4 項で述べた通りであり、その解全体の集合は、

$$X = \left\{ \begin{bmatrix} 0 \\ 4 \\ 0 \\ 5 \\ 0 \end{bmatrix} + c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix} \mid c_1, c_2, c_3 \in \mathbb{R} \right\}.$$

14.4 和集合と共通部分

ここで紹介する記号は、本書では、後半のごく一部の命題を証明する際にのみ用いられる。

定義 14.4.1. X を全体集合とする。 X の部分集合 A, B に関して、次のような集合が新たに定義される:

- A に含まれる元と B に含まれる元をまとめた集合を A と B の和集合 (または合併集合) といい、これを $A \cup B$ と書く。すなわち、 $A \cup B := \{x \in X \mid x \in A \text{ または } x \in B\}$ 。 $A \cup B$ を「 A または B 」と読む。
- A と B の両方に含まれる元をすべて集めた集合を A と B の共通部分といい、これを $A \cap B$ と書く。すなわち、 $A \cap B := \{x \in X \mid x \in A \text{ かつ } x \in B\}$ 。 $A \cap B$ を「 A かつ B 」と読む。
- A に含まれ、かつ B に含まれない元をすべて集めた集合を A と B の差集合といい、これを $A \setminus B$ と書く。すなわち、 $A \setminus B = \{x \in X \mid x \in A \text{ かつ } x \notin B\}$ 。 $A \setminus B$ を「 A 引く B 」と読む。また、 A として全体集合を取る場合に、 $X \setminus B$ を B の補集合という。

備考 14.4.2. (1) 二つの条件「 $x \in A$ かつ $x \in B$ 」と「 $x \in B$ かつ $x \in A$ 」は同値である。また、「 $x \in A$ または $x \in B$ 」と「 $x \in B$ または $x \in A$ 」も同値である。この事実を通して、 $A \cap B = B \cap A$ 、 $A \cup B = B \cup A$ が導かれる。

(2) 差集合を $A - B$ と書く文献もあるが、本書ではこの記法は用いない。線形代数では $A, B \subset \mathbb{R}^n$ のときに $A - B := \{\mathbf{a} - \mathbf{b} \mid \mathbf{a} \in A, \mathbf{b} \in B\}$ と定義することが多く、これとの混同を避けるためである。また、のちに商空間について U/V という記法を用いる。こちらも $A \setminus B$ と似ているゆえ注意すること。

(3) $\mathbb{N} \cup \{0\}$ は、0以上の整数全体の集合に等しい。

(4) 「 x を \mathbb{R}^n の零でないベクトルとする」は「 $x \in \mathbb{R}^n \setminus \{0\}$ とする」と言い換えられる。

次に、複数個の部分集合についての和集合と共通部分を定めよう。いま、集合 A_1, A_2, A_3 が与えられているとき、これらの間の和集合として、次の二つの集合が考えられる：

$$(A_1 \cup A_2) \cup A_3, \quad A_1 \cup (A_2 \cup A_3).$$

しかしながら上の二つは集合として一致することから、以降では括弧を略して $A_1 \cup A_2 \cup A_3$ と書く。また、この集合は次のように表すこともできる：

$$A_1 \cup A_2 \cup A_3 = \{x \in X \mid i = 1, 2, 3 \text{ のいずれかにおいて } x \in A_i \text{ が成り立つ}\}.$$

共通部分については

$$A_1 \cap A_2 \cap A_3 = \{x \in X \mid \text{各 } i = 1, 2, 3 \text{ において } x \in A_i \text{ が成り立つ}\}$$

となる。これらを考慮して、複数個の集合における和集合および共通部分を次のように定める。

定義 14.4.3. n を自然数とし、全体集合 X の n 個の部分集合 A_1, A_2, \dots, A_n に対して、

- $A_1 \cup \dots \cup A_n := \{x \in X \mid i = 1, \dots, n \text{ のいずれかにおいて } x \in A_i \text{ が成り立つ}\}.$
- $A_1 \cap \dots \cap A_n := \{x \in X \mid \text{各 } i = 1, \dots, n \text{ について } x \in A_i\}.$

上で定めた和集合および共通部分をそれぞれ $\bigcup_{i=1}^n A_i, \bigcap_{i=1}^n A_i$ と表記することもある。

14.5 集合論と逆理 (よりみち)

内包的な集合の記述を導入したことにより、多様性に富んだ集合の表現が可能になった。しかし、そこには大きな落とし穴が潜んでいることが知られている。それは次の枠内における議論であり、ラッセルの逆理と呼ばれている。

ラッセルの逆理

考えられ得るすべての集合を集めた集合を U とする (つまり U は集合の集合である)。更に条件 $P(X)$ を “ $X \notin X$ ” と定め、集合 Y を次で定義する：

$$Y = \{X \in U \mid X \notin X\}.$$

このとき、集合 Y 自身は Y の元となるであろうか。 $Y \in Y$ および $Y \notin Y$ のいずれかが成立するはずである。どちらが正しいのか検討しよう。

- (1) $Y \in Y$ と仮定すると、定義によれば Y は条件 $P(X)$ を満たす集合の集まりであったから、その元である Y 自身も条件を満たす。すなわち $Y \notin Y$ が成立する。しかしこれは $Y \in Y$ と仮定したことに矛盾する。
- (2) そこで、 $Y \notin Y$ と仮定する。つまり Y は条件 $P(X)$ を満たしており、 $P(X)$ を満たす集合全体の集まりが Y であったから、 $Y \in Y$ である。しかしこれは $Y \notin Y$ という仮定に矛盾する。

こうして、いずれの場合も矛盾が生じてしまった。

数学における議論あるいは証明は、一般の諸科学と比べても厳密性が非常に高いものであると多くの人が認識していることだろう。しかし、ラッセルの逆理によると、数学の論理にも少々あやふやな部分があるということなのだろうか。また、そうではないとするのであれば、上のような矛盾を排除するかたちで数学の理論（特に集合論）を再構成することは可能なのだろうか。現在では、逆理を回避するための技術が得られており、初学者がこの点について不安がる必要はないことになっている。本節では、この点について概略的な説明をしておこう。

逆理を回避するための技術論を検討するのであれば、そもそも「証明」とは何か再考する必要があるだろう。そのための模範となった理論はユークリッド幾何学である。ユークリッド幾何学では、いくつかの公理を前提として演繹的に数多くの定理を導いていた。これと同様にして集合論においても、集合に関するいくつかの公理を認め、それらの公理と論理的に正しい命題⁵および三段論法などの推論規則を有限回だけ用いて別の命題を導くことを「証明」と定めるのである。こうした立場で展開する集合論を公理的集合論と呼ぶ⁶。

集合論の公理として何を採用すべきかという基準は、もちろん数学者各個人の価値観によって論点が分かれることかもしれない。しかしながら、現在では ZFC⁷と呼ばれる公理系が多くの数学者の同意を得て、一般的に用いられている。ZFC の詳細を書く余裕はないが、この公理系においては集合全体の集合 U は構成できず、したがって枠内の議論における Y も定義できない（詳しい理由は本章末のコラムを見よ）。かくしてラッセルのパラドックスは避けられるのである。

ところで、良い公理系を導入した理論において、もはやラッセルの逆理は生じないにしても、それではラッセルの逆理とは別の矛盾も絶対に生じないという保証はあるのだろうか。もし、新たな矛盾論法が見つかってしまったならば、その矛盾を排除するようなより頑強な公理系が作れるかどうかを検討せねばならない。こうした不安を解消するためにも、ZFC において矛盾が導かれることはない（これを無矛盾であるという）ことを証明しようという組織的な試みがなされた（この試みはヒルベルト・プログラムと呼ばれる）。ヒルベルト・プログラムにおける最終的な答えはゲーデルによって与えられており、彼によれば、自然数論を含む無矛盾ないかなる公理系においても、その公理系の内部で自身の無矛盾性を証明することは出来ない（ゲーデルの第 2 不完全性定理）というのである。無矛盾性が保証されることは決してないというゲーデルの回答は悲観すべきことだろうか。この議論は数学界の内部にとどまらず多くの人が興味を持ち、様々な論争を巻き起こすことになった。

ところで、自己言及によって矛盾を導くというラッセルの逆理と構造の似た逆理がいくつか知られている。例えば「この文は間違っている」という文は正しいか、それとも間違っているのか。正しいとすれば、その文面通り間違っているから矛盾であり、間違っているとすれば「間違っている」ことは間違いということで正しいことになり、やはり矛盾を得る。こうした日常言語の世界を我々はどうか捉えるべきかという課題もある。論理学に加えて言語学や認識論といった様々な背景を抱えたこの難問も広く論じられ、その回答のうち代表的なものとして、例えばウィトゲンシュタインの『論理哲学論考』が挙げられる。

⁵ $A \Rightarrow A$ や $(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$ といった命題は A, B にどんな論理式を代入しても真である。こういった論理式は恒真式と呼ばれ、数理論理学において厳密に定義される。

⁶ これに対して、公理化せずに感覚的に集合を扱う立場を素朴集合論と言う。

⁷ ツェルメロ＝フレンケルの公理系 (ZF) に選択公理 (Axiom of Choice) を加えた公理系のこと。

公理的集合論における集合概念の厳密な定義とは何であろうか。これは数理論理学を専攻しなくても気になることであろう。結論を先に述べてしまうと、集合自体に確固たる定義などはない。例えばユークリッド幾何学における「直線」については、公理によって直線の性質がいくつか仮定されるのみであって、直線そのものが定義されるわけではない。それと同様に、集合論においても定義が厳密なのは公理であり集合自身ではないのである。

しかしそうすると、今度は我々が数学で用いる集合らしきものが、公理的集合論の文脈における集合であるのかどうかという不安にさいなまれるかもしれない。この点において、大概是、次の二点さえおさえておけば十分である。一つは、実数全体 \mathbb{R} や複素数全体 \mathbb{C} 、ユークリッド空間 \mathbb{R}^n など日常的に現れる集合は集合論の公理を組み合わせることで構成できることが分かっており、これらが集合であること (集合として存在すること) を疑う余地はないということである。もう一つは、我々が新たに構成する集合についての注意であり、これについては置換公理と分出公理をおさえておけばよい。置換公理とは大まかに言えば、外延的な記述によって集合を与えてよいという公理に相当し、分出公理は、内包的な記述によって集合 X の部分集合を定義してよいという公理に相当する。これらの公理のおかげで、実質的な数学に現れる対象が集合であるかないかを我々が意識する必要はないのである。

分出公理についてももう少し詳しく言うと、それは次のように述べられる:

分出公理: X を集合、 P を x に関する論理式とすれば、集合 $A := \{x \in X \mid P(x)\}$ が存在する。

分出公理のかなめは文頭が「 X を集合とすれば～」となっている点である。集合全体 U を集合であると仮定して上の X に適用するとラッセルの逆理によって矛盾が生じることから、したがって U は集合ではない、集合全体をなすような集合は存在しないという結論に至る。これより詳しい事情については公理的集合論の専門書に譲ろう。

第15章 線形空間

線形空間は線形代数学において主題となる代数構造である。公理化された代数構造を論じる理由は、2.4節で述べたように演算の定義にいちいち戻らなくても議論ができるという点、そして様々な空間を同時に論じることができる点にある。例えば、 \mathbb{R}^n において示された性質が関数のなす集合においても示され、それらの証明に使われた技法もほとんど同じというのであれば、それらを同時に証明できるような枠組みを与えておくと手間が省ける(6.4節における三つの命題の類似性を思いだそう)。このように数学では、汎用性を重視して抽象的な代数構造を導入している。

15.1 ベクトル空間の公理

\mathbb{R}^n における和とスカラー倍の性質のなかで特に重要と思われる部分を抽出することで、我々は線形空間の定義を得る:

定義 15.1.1. 集合 V に対して和 $+$ とスカラー倍 \cdot の演算が定められており、さらに特別な元 $\mathbf{0} \in V$ が与えられているとする。これらの演算が次に述べるベクトル空間の公理を満たすとき、四つ組 $(V, \mathbf{0}, +, \cdot)$ を線形空間 (linear space) またはベクトル空間 (vector space) と呼ぶ。

ベクトル空間の公理

$\mathbf{a}, \mathbf{b}, \mathbf{c} \in V, r, s \in \mathbb{R}$ とする。

I. 各元 $\mathbf{a}, \mathbf{b} \in V$ に対して和 $\mathbf{a} + \mathbf{b} \in V$ が定まっており、次の性質を満たす:

$$(1) \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad (2) (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}), \quad (3) \mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}.$$

II. 各元 $\mathbf{a} \in V$ および $r \in \mathbb{R}$ に対してスカラー倍 $r \cdot \mathbf{a} \in V$ が定まっており、次の性質を満たす:

$$(4) r \cdot (s \cdot \mathbf{a}) = (rs) \cdot \mathbf{a}, \quad (5) (r + s) \cdot \mathbf{a} = (r \cdot \mathbf{a}) + (s \cdot \mathbf{a}), \quad (6) r \cdot (\mathbf{a} + \mathbf{b}) = (r \cdot \mathbf{a}) + (r \cdot \mathbf{b}), \\ (7) 1 \cdot \mathbf{a} = \mathbf{a}, \quad (8) 0 \cdot \mathbf{a} = \mathbf{0}.$$

上の性質を満たす $\mathbf{0}$ のことを V の零元 (zero element) または零ベクトル (zero vector) と呼ぶ。また、線形空間の元を総称してベクトルと呼ぶ。

零ベクトルが V の元であることを強調し、これを $\mathbf{0}_V$ と書くこともある。慣例ではスカラー倍の演算記号 \cdot は省略して $r \cdot \mathbf{a}$ を $r\mathbf{a}$ と書き、更に四つ組 $(V, \mathbf{0}, +, \cdot)$ を V と略記する。

ベクトル空間の公理 (1) から (8) を直ちに暗記しないと以後の線形代数学の理解に支障がでるかといえば、そのようなことはない。何故なら、線形空間における演算は \mathbb{R}^n における演算と同様に無意識のうちに処理されるからである。しかし、「 \mathbb{R}^n の演算と似たような演算をもつ集合」と曖昧に線形空間を定義するわけにもいかず、上のように形式的な定義を与えた。

例 15.1.2. ベクトル空間の公理における性質 (1) から (8) は、命題 2.4.1 における性質 (1), (3), (2), (11), (15), (14), (10), (9) に相当する¹。命題 2.4.1 から次が従う:

(1) \mathbb{R}^n における原点 $\mathbf{0} := (0, \dots, 0)$ を零元とみなすことにより、1.1節で定めた \mathbb{R}^n における和とスカラー倍はベクトル空間の公理を満たす。したがって \mathbb{R}^n は線形空間である。

¹したがって、多元環はベクトル空間である。多元環とは、ベクトル空間に分配法則と結合法則を満たす積を付加した代数構造のことを指す (詳しい定義は 20 ページのコラムを見よ)。

(2) (m, n) -行列全体のなす集合を $M_{m,n}(\mathbb{R})$ と書く. 2.2 章で述べたように $M_{m,n}(\mathbb{R})$ には和とスカラー倍が定義されている. 零行列 $O_{m,n}$ を零元とみなすことにより, $M_{m,n}(\mathbb{R})$ における和とスカラー倍はベクトル空間の公理を満たし, したがって $M_{m,n}(\mathbb{R})$ は線形空間である.

我々は \mathbb{R}^n と $M_{n,1}(\mathbb{R})$ (あるいは $M_{1,n}(\mathbb{R})$) を同一視していた. また, 行列式の項目において n 次正方行列全体の集合を $M_n(\mathbb{R})$ と書いていた. つまり, $M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$ である.

15.2 線形空間の例

線形空間の例をいくつか挙げよう. はじめの例は \mathbb{R}^n の部分集合として具体的に表示できるものである. この例の一般化については次章で更に詳しく述べる.

例 15.2.1. $W := \{(x, y) \in \mathbb{R}^2 \mid x - y = 0\}$ とすれば, W は直線 $y = x$ 上の点全体を表す. W における和とスカラー倍の演算を, \mathbb{R}^2 における和とスカラー倍の演算によって定めよう. すると, $\mathbf{0} \in \mathbb{R}^2$ が W に含まれること, および W の各元どうしの和や W の元のスカラー倍が再び W の元となることが次のように確認できる:

Proof. $0 - 0 = 0$ より $(x, y) = (0, 0)$ は条件「 $x - y = 0$ 」を満たす. ゆえに $\mathbf{0} = (0, 0) \in W$ である. 次に, $\mathbf{a} = (a_1, a_2) \in W, \mathbf{b} = (b_1, b_2) \in W, r \in \mathbb{R}$ と仮定し, $\mathbf{a} + \mathbf{b}, r\mathbf{a} \in W$ を示そう. $\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2)$ が W に含まれることを示すには, $(x, y) = (a_1 + b_1, a_2 + b_2)$ が条件「 $x - y = 0$ 」を満たすことをいえばよい. $\mathbf{a}, \mathbf{b} \in W$ より, $(x, y) = (a_1, a_2)$ および $(x, y) = (b_1, b_2)$ について条件「 $x - y = 0$ 」が成立する. つまり $a_1 - a_2 = 0, b_1 - b_2 = 0$ である. ゆえに $(a_1 + b_1) - (a_2 + b_2) = (a_1 - a_2) + (b_1 - b_2) = 0 + 0 = 0$. したがって, $(x, y) = (a_1 + b_1, a_2 + b_2)$ は条件「 $x - y = 0$ 」を満たし, $\mathbf{a} + \mathbf{b} \in W$ である. また $r\mathbf{a} = (ra_1, ra_2)$ が W に含まれることを示すには, $(x, y) = (ra_1, ra_2)$ が条件「 $x - y = 0$ 」を満たすことをいえばよい. これも $(ra_1) - (ra_2) = r(a_1 - a_2) = r \cdot 0 = 0$ と直ちに分かる. \square

W における和とスカラー倍がベクトル空間の公理を満たすことは, \mathbb{R}^n における演算がそうであることから明らかであり, したがって W は線形空間である.

ユークリッド空間 \mathbb{R}^n とは n 本の座標軸を持つ空間のことであった. ここから類推される空間として, 無限個の座標軸を持つ空間を考えよう.

例 15.2.2. 無限個の実数の組 (x_1, x_2, x_3, \dots) 全体のなす集合を $\mathbb{R}^{\mathbb{N}}$ と書く. これは, 無限に続く実数列全体のなす集合とも考えられる. \mathbb{R}^n の場合と同様にして, $\mathbb{R}^{\mathbb{N}}$ にも次のように和とスカラー倍が定められる:

- $(x_1, x_2, x_3, \dots) + (y_1, y_2, y_3, \dots) := (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots),$
- $r(x_1, x_2, x_3, \dots) := (rx_1, rx_2, rx_3, \dots).$

すなわち, 各座標ごとに和とスカラー倍を取っている. これらの演算はベクトル空間の公理を満たし, $\mathbb{R}^{\mathbb{N}}$ はベクトル空間となる. なお, 数列 (x_1, x_2, x_3, \dots) を $(x_n)_{n \in \mathbb{N}}$ と記すこともある. この記法を用いて上の演算を書けば次のようになる:

$$(x_n)_{n \in \mathbb{N}} + (y_n)_{n \in \mathbb{N}} = (x_n + y_n)_{n \in \mathbb{N}}, \quad r(x_n)_{n \in \mathbb{N}} = (rx_n)_{n \in \mathbb{N}}.$$

次に多項式のなす空間を考えよう.

定義 15.2.3. 文字 x および非負整数 m , 実数 a_0, a_1, \dots, a_m を用いて

$$f(x) = \sum_{i=0}^m a_i x^i \quad \left(= a_m x^m + a_{m-1} x^{m-1} + \dots + a_2 x^2 + a_1 x + a_0 \right) \quad (15.2.1)$$

と表される式のことを実数を係数とする多項式 (**polynomial**) という. ここでは形式的に $x^0 = 1$ と定める. 上の式において $a_i \neq 0$ を満たす i の中で最大のものを n とするとき, f の次数 (**degree**) を n と定め, これを記号 $\deg f$ で表す. また, このとき, f を **n 次多項式 (polynomial of degree n)** という.

次は明らかであろう。

命題 15.2.4. 多項式 $P(x), Q(x)$ について, $\deg P(x)Q(x) = \deg P(x) + \deg Q(x)$.

多項式の次数が n 以下であるとき, これを n 次以下の多項式と呼ぼう. $m > n$ のとき, n 次以下の多項式は m 次以下の多項式でもある. また, 二つの多項式 $2x^2 + 3x - 1$ および $0x^3 + 2x^2 + 3x - 1$ を同一視し, これらの多項式は等しいと考えよう². 一般的に述べれば, n 次以下の多項式はすべて $\sum_{i=0}^n a_i x^i$ なる形で表せるということである.

例 15.2.5. 実数を係数とする多項式全体の集合を $\mathbb{R}[x]$ と書き, これを実数係数多項式環 (**polynomial ring**) という. また, その中で n 次以下の多項式全体のなす部分集合を $\mathbb{R}[x]_n$ と書く. 例えば $\mathbb{R}[x]_2 = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ である. さて, $\mathbb{R}[x]$ において和とスカラー倍を次のように定めると, これらは再び多項式になる:

$$\left(\sum_{i=0}^n a_i x^i\right) + \left(\sum_{i=0}^n b_i x^i\right) := \sum_{i=0}^n (a_i + b_i) x^i, \quad r \left(\sum_{i=0}^n a_i x^i\right) := \sum_{i=0}^n (ra_i) x^i.$$

これらの演算がベクトル空間の公理を満たすことは容易に確かめられ, したがって $\mathbb{R}[x]$ は線形空間である. なお, $\mathbb{R}[x]$ における零元とは, すべての係数 a_i が 0 となる多項式のことである. これを関数と見なせば, どんな数を代入しても 0 に値をとる定数関数を意味する. また, n 次以下の多項式において和やスカラー倍を行うと, n 次以下の多項式が得られることから, $\mathbb{R}[x]_n$ も線形空間である.

多項式には, 形式的な式とみなす立場と, x を変数とする関数とみなす立場がある. 前者は, 二つの多項式 $p(x) = \sum_{i=0}^n a_i x^i$ と $q(x) = \sum_{i=0}^n b_i x^i$ が等しいことを, 各 $i = 0, \dots, n$ について $a_i = b_i$ であると定める立場である. 後者は, $p(x) = q(x)$ がいかなる定義域の元 x に対しても成り立つことを $p = q$ の定義とする立場である. なお, 後者の定義においてはあらかじめ定義域を宣言しておかねばならない. 厳密には, これら二つの立場を区別すべきであるが, 線形代数の初歩を学ぶにあたってはこだわる必要はない. 次の命題はその根拠となる.

命題 15.2.6. $n + 1$ 個以上の元を含む \mathbb{R} の部分集合 I を定義域とする二つの n 次以下の多項式関数 $p(x) = \sum_{i=0}^n a_i x^i$, $q(x) = \sum_{i=0}^n b_i x^i$ について, 次は同値である.

$$(1) \text{ 各 } x \in I \text{ について } p(x) = q(x), \quad (2) \text{ 各 } i = 0, \dots, n \text{ について } a_i = b_i.$$

Proof. (2) \Rightarrow (1) は明らかゆえ (1) \Rightarrow (2) を示す. $p(x) = q(x)$ より関数 $p(x) - q(x) = \sum_{i=0}^n (a_i - b_i) x^i$ は 0 に値を取る定数関数である. I は $n + 1$ 個以上の元を含むことから, $n + 1$ 個の異なる数 $t_1, \dots, t_{n+1} \in I$ を取ることができ, これらの数を $p(x) - q(x)$ に代入することで次の式を得る.

$$\begin{cases} t_1^n(a_n - b_n) + t_1^{n-1}(a_{n-1} - b_{n-1}) + \dots + t_1(a_1 - b_1) + 1 \cdot (a_0 - b_0) & = 0 \\ t_2^n(a_n - b_n) + t_2^{n-1}(a_{n-1} - b_{n-1}) + \dots + t_2(a_1 - b_1) + 1 \cdot (a_0 - b_0) & = 0 \\ & \vdots \\ t_n^n(a_n - b_n) + t_n^{n-1}(a_{n-1} - b_{n-1}) + \dots + t_n(a_1 - b_1) + 1 \cdot (a_0 - b_0) & = 0 \\ t_{n+1}^n(a_n - b_n) + t_{n+1}^{n-1}(a_{n-1} - b_{n-1}) + \dots + t_{n+1}(a_1 - b_1) + 1 \cdot (a_0 - b_0) & = 0. \end{cases} \quad (15.2.2)$$

上式に現れる各 $(a_i - b_i)$ にかかる係数を成分とする $n + 1$ 次正方行列 A を次で定める:

$$A := \begin{bmatrix} t_1^n & t_1^{n-1} & \dots & t_1 & 1 \\ t_2^n & t_2^{n-1} & \dots & t_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_n^n & t_n^{n-1} & \dots & t_n & 1 \\ t_{n+1}^n & t_{n+1}^{n-1} & \dots & t_{n+1} & 1 \end{bmatrix}.$$

²二つの多項式が形式的な意味で等しいことを正確に定義するのは意外に面倒である (107 ページのコラムを見よ).

すると式 15.2.2 は, $\mathbf{x} = {}^t(a_n - b_n, \dots, a_0 - b_0)$ が $A\mathbf{x} = \mathbf{0}$ の解であることを意味している. ヴァンデルモンドの公式 (定理 11.2.2) より $|A| = \pm \prod_{1 \leq i < j \leq n+1} (t_j - t_i)$ であり, 各 t_i は異なる数ゆえ $|A| \neq 0$. ゆえに A は可逆であり方程式 $A\mathbf{x} = \mathbf{0}$ の解は唯一解 $\mathbf{x} = \mathbf{0}$ のみである. したがって $a_i - b_i = 0$ ($i = 0, \dots, n$). \square

よりみち (文字列と多項式)

二つの文字列 $0x^3 + 2x^2 + 3x + 1$ と $3x^2 + 1 + 3x + (-1)x^2$ は明らかに異なる文字列であるが, くだん我々はこれらを多項式として等しいと考えている. ここで, 形式的に与えられた二つの文字列が多項式として等しいことの定義を与えてみよう. 以下では多項式とみなすべき式として, 式 (15.2.1) をさらに一般化した

$$a_0x^{n_0} + a_1x^{n_1} + \dots + a_mx^{n_m} \quad (\text{ここで } m, n_i \in \mathbb{N} \cup \{0\}, a_i \in \mathbb{R})$$

で表される形式的な文字列を考え, このように書ける文字列全体のなす集合を $\text{Pol}(\mathbb{R})$ としよう. なお, 単項式の引き算 $-ax^k$ は $+(-a)x^k$ の略式であるとし, ここでは引き算による表示は考えないとする.

さて, 各文字列 $w = a_0x^{n_0} + a_1x^{n_1} + \dots + a_mx^{n_m} \in \text{Pol}(\mathbb{R})$ に対して, 数列 $(a_0^w, a_1^w, a_2^w, \dots)$ を

$$\begin{aligned} a_n^w &:= \text{“文字列 } w \text{ において, } x \text{ の次数がちょうど } n \text{ になる項の係数たちの総和”} \\ &= \text{“各 } i = 0, \dots, m \text{ のうち, } m_i = n \text{ を満たす } i \text{ に関する } a_i \text{ たちの総和”} \end{aligned}$$

と定めれば, 写像 $Q: \text{Pol}(\mathbb{R}) \rightarrow \mathbb{R}^{\mathbb{N}}$ ($Q(w) = (a_0^w, a_1^w, a_2^w, \dots)$) が得られる. ここで, 二つの文字列 $v, w \in \text{Pol}(\mathbb{R})$ が同じ式であることを, $Q(v) = Q(w)$ が成り立つことと定める. この「二つの文字列は同じ式である」という関係は $\text{Pol}(\mathbb{R})$ 上の同値関係をなし, その商集合 $\mathbb{R}[x] := \text{Pol}(\mathbb{R})/\simeq$ の元を多項式と呼ぶ (同値関係や商集合について詳しくは 23.5 節を見よ). 例えば, 二つの文字列 $v = x + 1$ と $w = 1 + x$ は集合 $\text{Pol}(\mathbb{R})$ の元としては異なり, 多項式 $\mathbb{R}[x]$ の元としては等しい (すなわち $[v] = [w]$). くだん我々は, 同値類 $[w] \in \mathbb{R}[x]$ を単に w と略して書き, 同値類としての等式 $[v] = [w]$ を単に $v = w$ と書いて済ましている. 上でいう「多項式として等しい」 (の正確な定義) とは, ここでいう「文字列が同じ式である」ことにほかならない.

多項式を関数とみなす立場からは, 次のような一般化が考えられる.

例 15.2.7. \mathbb{R} 上の区間 I を定義域とする実数値連続関数全体のなす集合を $C(I)$ と書く. $C(I)$ において, 和とスカラー倍を定義しよう. 新たな関数 h を定義するということは, 定義域の各元を h に代入した値を定めることに他ならない. そこで, 次のように演算を定義する.

- $f, g \in C(I)$ に対して関数 $(f + g): I \rightarrow \mathbb{R}$ を次のように定める:
各 $x \in I$ について, $(f + g)(x) := f(x) + g(x)$.
- $f \in C(I)$ および $r \in \mathbb{R}$ に対して関数 $(rf): I \rightarrow \mathbb{R}$ を次のように定める:
各 $x \in I$ について, $(rf)(x) := rf(x)$.

$f + g$ や rf が再び連続関数となること (すなわち $f + g, rf \in C(I)$) の証明は微積分の教科書に譲ろう³. $C(I)$ における零元は 0 に値をとる定数関数である. 上で定義された演算がベクトル空間の公理を満たすことは容易に確かめられ, したがって, $C(I)$ は線形空間である.

例 15.2.5 における多項式を区間 I を定義域とする関数とみなす立場においては, 多項式はいずれも連続関数であるから $\mathbb{R}[x] \subset C(I)$ となる. このとき, 例 15.2.5 における和とスカラー倍の定義と, 例 15.2.7 におけるそれは一致する.

³取束列による連続性の定義を採用すれば, 高校数学の範囲で示せることである.

例 15.2.8. (1) 関数の連続性は数列の収束概念を用いて定められていた. これに類似する方法で, 収束概念が定まる空間を定義域とする関数についても連続性が定められる. 例えば \mathbb{R}^2 の点列 \mathbf{x}_n が点 $\mathbf{x} \in \mathbb{R}^2$ に収束することは, \mathbf{x}_n と \mathbf{x} の距離⁴が 0 に収束することと定められる. 収束概念が定まる図形 X (例えば X を \mathbb{R}^2 の部分集合とすればよい) を定義域とする実数値連続関数全体のなす集合を $C(X)$ と書けば, $C(X)$ も例 15.2.7 と同様にして線形空間となる.

(2) 実は, 関数のなす空間を線形空間とみなすために, 関数を連続関数のみに制限する必要はない. より一般に, 集合 X を定義域とする実数値関数全体のなす集合を \mathbb{R}^X とすれば, 例 15.2.7 と同様に加とスカラー倍を定めることで \mathbb{R}^X は線形空間となる.

なお, この例において $X = \mathbb{N}$ とする場合と例 15.2.2 における数列空間 $\mathbb{R}^{\mathbb{N}}$ の間には自然な 1 対 1 の対応があり, これらは同一の概念とみなすことができる. 実際, 数列空間の各元 (x_1, x_2, \dots) は $x(n) := x_n$ なる関数 $x : \mathbb{N} \rightarrow \mathbb{R}$ に対応し, 逆に関数 $y : \mathbb{N} \rightarrow \mathbb{R}$ は実数列 $y_n := y(n)$ (すなわち $(y(1), y(2), \dots)$) に対応する. この対応は, それぞれの和とスカラー倍の演算に関しても整合的である⁵. \mathbb{N} を定義域とする関数と数列の違いは, $x(1), x(2), \dots$ と書くか, あるいは x_1, x_2, \dots と書くかという, 僅かな記号上の違いしかないのである.

例 15.2.9 (発展). $C(X)$ には次のようにして積も定めることができる.

$$f, g \in C(X) \text{ に対して関数 } fg \text{ を次で定める: 各 } x \in X \text{ について, } (fg)(x) := f(x)g(x).$$

このとき, $C(X)$ における演算は, $1 \in \mathbb{R}$ に値を取る定数関数 $\mathbf{1}$ を単位元とすることで命題 2.4.1 における (1) から (18) すべての性質を満たす⁶. すなわち, $C(X)$ は多元環となる. また, 多項式の積は再び多項式になることから, $\mathbb{R}[x]$ も多元環となる. $\mathbb{R}[x]_n$ は多元環にはならない. 何故なら, n 次多項式どうしの積は $2n$ 次の多項式となり, これは $\mathbb{R}[x]_n$ に含まれないからである.

15.3 体 K 上の線形空間 (発展)

これまでの議論において, 行列に現れる成分およびスカラー倍の係数, 連立 1 次方程式に現れる係数はいずれも実数であるとしていた. しかし, 扱う数を実数に限るべき確たる根拠はどこにもなかった. 仮にあるとすれば, 初学者にとってイメージが描きやすいということに尽きるだろう. そこで, 行列に現れる成分を有理数に限る場合や, あるいは複素数も認める場合について考えてみよう.

仮に成分を有理数に限定した場合, 行列演算を繰り返して行っても成分に現れるのは有理数に限られ, 行列式の値も有理数である. また有理数係数の連立 1 次方程式の解も有理数である. 一方, 行列の成分に複素数を認める場合は, やはり行列演算後の成分や行列式の値は複素数となる. 複素数を係数とする連立 1 次方程式の解も複素数である. このように実数および有理数, 複素数に共通した現象が生じる背景には, これらが四則演算で閉じているという共通項がある.

したがって, 四則演算が定まる代数構造 K さえ与えられれば, 行列の成分やスカラー倍の係数, 連立 1 次方程式に現れる係数に K の元を取ることで, K に関する線形代数の世界が考えられうる. 四則演算が与えられる代数構造は体 (field) と呼ばれる. その形式的な定義は代数学の専門書を参照されたい⁷.

例 15.3.1. (1) 実数全体 \mathbb{R} および有理数全体 \mathbb{Q} , 複素数全体 \mathbb{C} はそれぞれ体である.

(2) 整数全体 \mathbb{Z} は, 整数どうしの割り算が整数になるとは限らないゆえ体ではない. 整数を成分とする行列において, 行列演算後の各成分や行列式は確かに整数となる. しかし, 整数係数の連立 1 次方程式は, 整数でない有理数を解に持つことがある. これは整数が割り算で閉じていないことに起因している.

⁴ $\mathbf{x}_n = (a_n, b_n)$ と $\mathbf{x} = (a, b)$ の距離はピタゴラスの定理により $\sqrt{(a_n - a)^2 + (b_n - b)^2}$ と計算できる.

⁵ ここでいう対応が「和とスカラー倍の演算に関して整合的である」ことの正確な定義は, 線形性を満たすことにほかならない.

⁶ 命題 2.4.1 では単位元の記号に $\mathbf{1}$ ではなく E を用いている.

⁷ K が可換な環であり, かつ K から零元を除いた集合が積演算に関して群となるとき, K を体という.

- (3) 無理数全体は体ではない. 無理数どうしの和が無理数になるとは限らないからである. 例えば $\sqrt{2}$ および $1 - \sqrt{2}$ は共に無理数であるが (練習 15.3.2), その和 $\sqrt{2} + (1 - \sqrt{2}) = 1$ は有理数である.

練習 15.3.2. $\sqrt{2}$ が無理数であることを認めたとえで $1 - \sqrt{2}$ が無理数であることを示せ.

解答例: 仮に $1 - \sqrt{2}$ が無理数でないとすればこれは有理数であり, $\sqrt{2}$ は有理数どうしの引き算 $\sqrt{2} = 1 - (1 - \sqrt{2})$ で表せる. したがって $\sqrt{2}$ は有理数となり, これは $\sqrt{2}$ が無理数であることに反する. \square

体 K に関する線形代数学では, 次で定める線形空間を対象とする.

定義 15.3.3. K を体とする. 集合 V に対して和 $+$, および K の元に関するスカラー倍 \cdot の演算が定められており, さらに特別な元 $\mathbf{0} \in V$ が与えられているとする. これらの演算が次の条件 (ベクトル空間の公理) をすべて満たすとき, 四つ組 $(V, \mathbf{0}, +, \cdot)$ を体 K 上のベクトル空間または線形空間と呼ぶ.

I. 各元 $\mathbf{a}, \mathbf{b} \in V$ に対して和 $\mathbf{a} + \mathbf{b} \in V$ が定まっており, 次の性質を満たす:

$$(1) \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad (2) (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}), \quad (3) \mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}.$$

II. 各元 $\mathbf{a} \in V$ および $r \in K$ に対してスカラー倍 $r \cdot \mathbf{a} \in V$ が定まっており, 次の性質を満たす:

$$(4) r \cdot (s \cdot \mathbf{a}) = (rs) \cdot \mathbf{a}, \quad (5) (r + s) \cdot \mathbf{a} = (r \cdot \mathbf{a}) + (s \cdot \mathbf{a}), \quad (6) r \cdot (\mathbf{a} + \mathbf{b}) = (r \cdot \mathbf{a}) + (r \cdot \mathbf{b}), \\ (7) 1 \cdot \mathbf{a} = \mathbf{a}, \quad (8) 0 \cdot \mathbf{a} = \mathbf{0}.$$

本書で主題とする線形空間は \mathbb{R} 上の線形空間である. しかし, 体として \mathbb{R} 以外のものを採用しても, 多くの場合に同等の議論が得られることを頭の片隅に留めておきたい. 実は, 体として複素数全体 \mathbb{C} を採用したほうが実数の場合よりも理論が綺麗になる部分があり, とくに, 第 V 部以降では体 \mathbb{C} 上の線形空間についても論じる.

例 15.3.4. K を体とする. K として \mathbb{R} を考えていた場合と同様に次が成り立つ:

- (1) K の元を n 個並べた組 (x_1, \dots, x_n) 全体からなる集合を K^n と書く. これは K 上の線形空間である.
- (2) K の元を成分とする (m, n) -行列全体のなす集合を $M_{m,n}(K)$ と書く. また, $M_n(K) := M_{n,n}(K)$ と定める. これらは K 上の線形空間である.

行列やベクトルの成分にどのような数を考えているかを明示する場合に, 次の語句が用いられる.

- $M_{m,n}(\mathbb{R})$ の元を実行列という. とくに m, n のいずれかが 1 の場合, これを実ベクトルと呼ぶ.
- $M_{m,n}(\mathbb{C})$ の元を複素行列という. とくに m, n のいずれかが 1 の場合, これを複素ベクトルと呼ぶ.

四則演算が成立する体 K 自身は、体の演算としての和と積を線形空間における和とスカラー倍とみなすことで線形空間になる。例えば実数直線 \mathbb{R} は座標軸が1つしかない線形空間である。次に、二つの体 K, L が与えられており、 $K \subset L$ が成り立つ場合を考えよう。このとき大きい体 L は、小さい体 K 上の線形空間とみなすこともできる。例えば $\mathbb{R} \subset \mathbb{C}$ であることから \mathbb{C} は \mathbb{R} 上の線形空間となる。このことは複素平面を通して、 \mathbb{C} は \mathbb{R}^2 と対応づけられることから分かる。一方で、 $\mathbb{Q} \subset \mathbb{R}$ について \mathbb{R} を \mathbb{Q} 上の線形空間とみたとき、この線形空間の性質を調べるのは容易でない (実際、無限次元となる)。一般に、大小関係のある二つの体の間にある対称性 (すなわち群) を調べる理論を体論 (ガロア理論) という。

複素数体 \mathbb{C} の定義を思いだそう。実数の世界に方程式 $x^2 = -1$ を満たす元を新たに加え、さらに四則演算が成り立つよう数空間を広げることで \mathbb{C} を得る。 $x^2 = -1$ を満たす数は二つ存在し (これを a, b としよう)、このうち一方を虚数単位として i と書き、このときもう一方は $-i$ と書かれる。ここで一つの疑問が現れる。ある人が虚数単位として数 a を選び参考書 A を書き、別の人が虚数単位に b を選んで参考書 B を書いたとすれば、参考書 A の i は参考書 B の $-i$ に相当する。したがって、二つの参考書で述べられている議論を比較しようと思えば面倒な翻訳作業 (変数変換) が必要なのは当然である。しかし、現実にはこのような作業は必要なく、参考書 A, B を並行して読む際に翻訳を意識する必要はない。これは不思議なことではないか。

体および群の概念が生まれた背景には、四則演算と根号のみを用いた5次方程式の解の公式の非存在証明があった。方程式の解を付け加えた体を考えて、解の対称性と方程式の関係を見極めることで、解の公式の非存在性が理解されたのである。また、上で述べた翻訳作業が必要ないことは、 $x^2 = -1$ の解の対称性を通して理解されている。

第16章 いろいろな線形部分空間

例 15.2.1 における W と \mathbb{R}^2 の関係や、例 15.2.5 における $\mathbb{R}[x]_n$ と $\mathbb{R}[x]$ の関係のように、より大きな線形空間の部分集合として実現される線形空間の例がいくつも考えられる。これらを総称する概念として線形部分空間なる概念を得る。

本章にて部分空間の数多くの例を紹介する。このことから、線形空間の枠組みで論じることのできる対象がいかに豊富であるか理解されることと思う。

16.1 定義

線形空間 V の部分集合 W が線形空間となるための条件を考えよう。 W における和とスカラー倍の演算がベクトル空間の公理を満たすことは、既に V における演算がそうであることから直ちに得られる。ゆえに、 W が線形空間となるためには、 W 内での演算結果が再び W に含まれること (このことを W が演算で閉じているという)、および零元があればよいことが分かる。こうして我々は次の定義に至る：

定義 16.1.1. 線形空間 $V = (V, \mathbf{0}_V, +, \cdot)$ の部分集合 W が次の性質 (i) から (iii) をすべて満たすとき、 $W = (W, \mathbf{0}_V, +, \cdot)$ もまた線形空間となる。

$$(i) \mathbf{0}_V \in W, \quad (ii) \mathbf{a}, \mathbf{b} \in W \implies \mathbf{a} + \mathbf{b} \in W, \quad (iii) \mathbf{a} \in W, r \in \mathbb{R} \implies r\mathbf{a} \in W.$$

このとき W を V の線形部分空間 (linear subspace) または部分ベクトル空間という。本書では、これらを部分空間と略称で述べる¹。

例 16.1.2. 線形空間 V に対して、 V 自身は V の部分空間である。また V の零元のみからなる集合 $\{\mathbf{0}\}$ も V の部分空間である。これら二つの部分空間のことを、 V の自明な部分空間という。

部分空間の例は本章の後半で述べる。その前に、部分空間になるための条件 (i)~(iii) をよく理解するために、 \mathbb{R}^2 の部分集合のなかで部分空間にならない例を挙げよう。条件 (i)~(iii) のいずれか一つでも満たさなければ部分空間にはなり得ないことから、否定的例はいくらでも簡単に列挙できる。そこで、三つの性質のうち二つは満たすものの、残りの一つを満たさないような例、つまり、あと一步で部分空間にならない例をここでは考える。

例 16.1.3. (1) 条件 (i) と (ii) を満たすが (iii) を満たさない例：

$W := \{(x, y) \in \mathbb{R}^2 \mid x \geq 0 \text{ かつ } y \geq 0\}$ とすれば W は (i) と (ii) を満たし、(iii) を満たさない。

Proof. $(x, y) = (0, 0)$ が条件「 $x \geq 0$ かつ $y \geq 0$ 」を満たすことから $\mathbf{0} = (0, 0) \in W$ である。また、 $\mathbf{w} = (x_1, y_1)$, $\mathbf{v} = (x_2, y_2)$ とし、 $\mathbf{w}, \mathbf{v} \in W$ と仮定すれば、 $x_1, y_1, x_2, y_2 \geq 0$ であり、ゆえに $x_1 + x_2 \geq 0$, $y_1 + y_2 \geq 0$ である。したがって、 $(x, y) = (x_1 + x_2, y_1 + y_2)$ は条件「 $x \geq 0$ かつ $y \geq 0$ 」を満たす。よって $\mathbf{w} + \mathbf{v} = (x_1 + x_2, y_1 + y_2) \in W$ 。つまり W は (ii) を満たす。 W が (iii) を満たさないことを示すには、(iii) を満たさない反例の一つを挙げればよい。例えば $\mathbf{a} = (1, 0)$, $r = -1$ とすれば $\mathbf{a} \in W$, $r \in \mathbb{R}$ である。しかしながら $r\mathbf{a} = -1(1, 0) = (-1, 0)$ であり、 $(x, y) = (-1, 0)$ は条件「 $x \geq 0$ かつ $y \geq 0$ 」を満たさないゆえ $r\mathbf{a} \notin W$ 。□

¹数学では様々な空間概念が与えられており、ゆえに何を対象としているかによって部分空間の意味は異なる。

(2) 条件 (i) と (iii) を満たすが (ii) を満たさない例:

$W := \{(x, y) \in \mathbb{R}^2 \mid x = 0 \text{ または } y = 0\}$ とすれば W は (i) と (iii) を満たし, (ii) を満たさない.

Proof. $(x, y) = (0, 0)$ が条件「 $x = 0$ または $y = 0$ 」を満たすことから $\mathbf{0} = (0, 0) \in W$ である. また, $\mathbf{w} = (a, b) \in W$, $r \in \mathbb{R}$ と仮定すれば, a, b のうち少なくともいずれか一方は 0 である. ゆえに ra, rb のいずれか一方は 0 であり, $(x, y) = (rx, rb)$ は条件「 $x = 0$ または $y = 0$ 」を満たす. したがって $r\mathbf{w} = (rx, ry) \in W$ である. W が (ii) を満たさないことを示そう. 例えば, $\mathbf{a} = (1, 0)$, $\mathbf{b} = (0, 1)$ とすれば $\mathbf{a}, \mathbf{b} \in W$ である. $\mathbf{a} + \mathbf{b} = (1, 0) + (0, 1) = (1, 1)$ であり, $(x, y) = (1, 1)$ は条件「 $x = 0$ または $y = 0$ 」を満たさないゆえ $\mathbf{a} + \mathbf{b} \notin W$. \square

(3) 条件 (ii) と (iii) を満たすが (i) を満たさない例:

(iii) において $r = 0$ を適用することで (i) が得られるゆえ, このような例は存在しないと考えたいところである. しかし, 実際には次の例が与えられる:

W を空集合とすれば, W は条件 (ii) と (iii) を満たすが (i) を満たさない.

Proof. 空集合は元を含まない集合ゆえ, とくに零元も含まず, したがって W は (i) を満たさない. (ii) を満たすことは次のように背理法で示される². もし仮に (ii) を満たさないとすれば, それは $\mathbf{w}, \mathbf{v} \in W$ であるにも関わらず $\mathbf{w} + \mathbf{v} \notin W$ となる例があるということである. この例において, とくに $\mathbf{w} \in W$ であり, したがって W は元を含む. これは W が元を含まない集合であったことに反する. 以上より W は (ii) を満たさねばならない. 同様の論法を用いて, W が (iii) を満たすことも示される. \square

部分空間となるための条件 (ii) と (iii) は次のようにまとめることができる.

命題 16.1.4. 定義 16.1.1 における条件 (ii) と (iii) が共に成立することと, 次の条件が成立することは同値である:

$$(iv) \quad r, s \in \mathbb{R}, \mathbf{a}, \mathbf{b} \in W \implies r\mathbf{a} + s\mathbf{b} \in W.$$

Proof. (ii) と (iii) を仮定して (iv) を示そう. $r, s \in \mathbb{R}$, $\mathbf{a}, \mathbf{b} \in W$ とすれば, (iii) より $r\mathbf{a}, s\mathbf{b} \in W$ である. これに (ii) を適用し $r\mathbf{a} + s\mathbf{b} \in W$ を得る. すなわち (iv) が成り立つ. また, (iv) において $r = s = 1$ という特別な場合が (ii) に相当し, $s = 0$ なる場合が (iii) に相当する. すなわち (iv) ならば (ii) かつ (iii) である. \square

以降, 部分空間であることを確認する際は条件 (ii) と (iii) の代わりに条件 (iv) を用いよう. これにより証明が多少は短くなるであろう. また, 条件 (iv) からは更に次の性質が導かれる.

命題 16.1.5. W を線形空間 V の部分空間とすれば, 各 $\mathbf{v}_i \in W$ および $r_i \in \mathbb{R}$ について $\sum_{i=1}^{\ell} r_i \mathbf{v}_i \in W$.

Proof. 和の個数 ℓ に関する帰納法で示す. $\ell = 1$ の場合は部分空間の性質 (iii) に他ならない. 和の個数が ℓ のとき成立すると仮定し, 和の個数が $\ell + 1$ の場合を示そう. $\mathbf{v}_i \in W$, $r_i \in \mathbb{R}$ ($i = 1, \dots, \ell + 1$) とし, $\mathbf{u} = \sum_{i=1}^{\ell+1} r_i \mathbf{v}_i$ とすれば, $\mathbf{u} = (\sum_{i=1}^{\ell} r_i \mathbf{v}_i) + r_{\ell+1} \mathbf{v}_{\ell+1}$ と書ける, 帰納法の仮定より $\sum_{i=1}^{\ell} r_i \mathbf{v}_i \in W$ であり, 性質 (iii) より $r_{\ell+1} \mathbf{v}_{\ell+1} \in W$ である. よって, 性質 (ii) より $(\sum_{i=1}^{\ell} r_i \mathbf{v}_i) + r_{\ell+1} \mathbf{v}_{\ell+1} \in W$, すなわち $\mathbf{u} \in W$ である. \square

17 章で線形結合と呼ばれる概念を導入する. これを用いて, 上の命題で述べている W の性質は「 W は線形結合で閉じている」と呼ばれる.

²より一般に, 前提が偽となる命題は真である. 論理規則の詳細については巻末の文献 [6, 付録 B] を見よ.

16.2 \mathbb{R}^n の部分空間

\mathbb{R}^n の部分空間の外延的表示と内包的表示について論じる.

例 16.2.1. \mathbb{R}^3 において互いに平行³でないベクトル $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ を取ると (ただし $\mathbf{a}, \mathbf{b} \neq \mathbf{0}$), $W = \{r_1\mathbf{a} + r_2\mathbf{b} \mid r_1, r_2 \in \mathbb{R}\}$ によって \mathbb{R}^3 上の原点 $\mathbf{0}$ および \mathbf{a}, \mathbf{b} を含む平面が定まる. この W は \mathbb{R}^3 の部分空間である.

Proof. 部分空間となるための条件 (i) および (iv) が満たされることを確認すればよい.

(i): $r_1 = r_2 = 0$ とすることで $\mathbf{0} = r_1\mathbf{a} + r_2\mathbf{b}$ と表せる. ゆえに $\mathbf{0} \in W$ である.

(iv): $\mathbf{x}, \mathbf{y} \in W$ とすれば $x_1, x_2, y_1, y_2 \in \mathbb{R}$ を用いて $\mathbf{x} = x_1\mathbf{a} + x_2\mathbf{b}$, $\mathbf{y} = y_1\mathbf{a} + y_2\mathbf{b}$ と書ける. 各 $r, s \in \mathbb{R}$ について

$$r\mathbf{x} + s\mathbf{y} = r(x_1\mathbf{a} + x_2\mathbf{b}) + s(y_1\mathbf{a} + y_2\mathbf{b}) = (rx_1 + sy_1)\mathbf{a} + (rx_2 + sy_2)\mathbf{b}$$

である. すなわち, $r\mathbf{x} + s\mathbf{y}$ は $r_1 = rx_1 + sy_1$, $r_2 = rx_2 + sy_2$ とすることで $r_1\mathbf{a} + r_2\mathbf{b}$ と表せる. ゆえに $r\mathbf{x} + s\mathbf{y} \in W$ である. \square

例 16.2.1 の証明において, \mathbf{a}, \mathbf{b} が平行でないことは用いられていない. \mathbf{a}, \mathbf{b} が平行である場合は, W は \mathbb{R}^3 の原点を通る直線になる. 一般に, \mathbb{R}^3 の部分空間とは, 原点を通る平面および原点を通る直線, 自明な部分空間 (\mathbb{R}^3 と $\{\mathbf{0}\}$) の四種に限られる. この分類の詳細は「次元」なる概念を通してなされる (例 22.4.2).

より一般に, 線形空間 V および $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ について $W = \{\sum_{i=1}^n r_i\mathbf{u}_i \mid r_1, \dots, r_n \in \mathbb{R}\}$ は V の部分空間となる (命題 18.1.2). この例が部分空間の外延的表示であるのに対して, 内包的表示は次で与えられる. これは例 15.2.1 の一般化に相当している⁴.

命題 16.2.2. (m, n) -行列 A に対して, 次で定められる W_A は \mathbb{R}^n の部分空間となる:

$$W_A := \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

Proof. 部分空間となるための条件 (i) および (iv) が満たされることを確認すればよい.

(i): $\mathbf{x} = \mathbf{0} \in \mathbb{R}^n$ とすれば, \mathbf{x} は条件 $A\mathbf{x} = \mathbf{0}$ を満たす. ゆえに $\mathbf{0} \in W_A$ である.

(iv): $\mathbf{a}, \mathbf{b} \in W_A$, $r, s \in \mathbb{R}$ とする. このとき, W_A の定義から $A\mathbf{a} = \mathbf{0}$, $A\mathbf{b} = \mathbf{0}$ が成り立つ. このとき $A(r\mathbf{a} + s\mathbf{b}) = r(A\mathbf{a}) + s(A\mathbf{b}) = r\mathbf{0} + s\mathbf{0} = \mathbf{0}$ ゆえ, $\mathbf{x} = r\mathbf{a} + s\mathbf{b}$ は条件 $A\mathbf{x} = \mathbf{0}$ を満たす. ゆえに $r\mathbf{a} + s\mathbf{b} \in W_A$ である. \square

上の W_A は斉次形連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ の解全体の空間に一致する. そこで, W_A は方程式 $A\mathbf{x} = \mathbf{0}$ の解空間と呼ばれる.

例 16.2.3. $A \neq B$ かつ $W_A = W_B$ なる例は山のようにある. 例えば, A, B が共に可逆正方行列ならば, それぞれの解空間 W_A, W_B は唯一解 $\mathbf{x} = \mathbf{0}$ のみからなる空間 $\{\mathbf{0}\}$ となる.

解空間の幾何的な意味を検討しよう. 列ベクトル $\mathbf{x} \in \mathbb{R}^n$ が $A\mathbf{x} = \mathbf{0}$ を満たすということは, A の各行と \mathbf{x} との行列としての積が 0 になるということである. これを \mathbb{R}^n 上のベクトルの内積と読み替えることにより, W_A の元であることは, A の各行と直交するベクトルであることと同値になることが分かる. この事実から, \mathbb{R}^n の任意の部分空間が W_A の形で表されることが示唆される. 例えば, 例 16.2.1 において, ベクトル \mathbf{a}, \mathbf{b} と共に直交するベクトル $\mathbf{c} \in \mathbb{R}^3$ を一つ取ると⁵, \mathbf{c} の成分を横に並べた $(1, 3)$ -行列を A とすれば, $W = W_A$ となる. 一般の場合の詳しい説明は第 VI 部で述べよう (命題 35.5.5(2)).

³線形空間における $\mathbf{0}$ でない二つのベクトル \mathbf{a}, \mathbf{b} が平行であるとは, $r\mathbf{a} = \mathbf{b}$ を満たす実数 r が存在することと定める.

⁴例 15.2.1 は, $(1, 2)$ -行列 $A = [1, -1]$ の場合に相当する. なお, 例 15.2.1 では行ベクトルで表示していたが, ここでは列ベクトルによる表示を考えている.

⁵このような \mathbf{c} が取れるかどうかという問題も解決しなければならない. \mathbb{R}^3 に限定した話では, \mathbf{a} と \mathbf{b} の外積と呼ばれるベクトル $\mathbf{a} \times \mathbf{b}$ が, この例の一つに相当する.

練習 16.2.4. 次の \mathbb{R}^2 の部分空間 W について, $W = W_A$ なる行列 A を求めよ.

(1) $W = \{\mathbf{0}\}$. 解答例: 可逆行列ならば何でもよい. 例えば $A = E_2$ とせよ.

(2) $W = \mathbb{R}^2$. 解答例: $A = O_{2,2}$, $A = O_{1,2}$ などとすればよい.

(3) $W = \{r\mathbf{y} \mid r \in \mathbb{R}\}$, ただし $\mathbf{y} = \begin{bmatrix} a \\ b \end{bmatrix} \neq \mathbf{0}$ とする (W は原点 $\mathbf{0}$ と \mathbf{y} を通る直線上の点全体を表す).

解答例: $\mathbf{z} = \begin{bmatrix} -b \\ a \end{bmatrix}$ は \mathbf{y} に直交するベクトルである. \mathbf{z} の成分を横にならべた $(1, 2)$ -行列 $A = \begin{bmatrix} -b & a \end{bmatrix}$ について, $W = W_A$ となる. $W = W_A$ を示すために, $W \subset W_A$ および $W_A \subset W$ を示そう.

($W \subset W_A$): $\mathbf{x} \in W$ を勝手に取れば, 実数 r を用いて $\mathbf{x} = r\mathbf{y}$ と書ける. このとき,

$$A\mathbf{x} = A(r\mathbf{y}) = rA\mathbf{y} = r \begin{bmatrix} -b & a \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = r(-ba + ab) = r0 = 0.$$

ゆえに $\mathbf{x} \in W_A$ である.

($W_A \subset W$): $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in W_A$ を勝手に取れば, $A\mathbf{x} = \mathbf{0}$ が成り立っている. すなわち, $-bx_1 + ax_2 = 0$ である. $\mathbf{y} \neq \mathbf{0}$ ゆえ $a \neq 0$ または $b \neq 0$ である. $a \neq 0$ の場合は実数 $r = \frac{x_1}{a}$ について $x_1 = ra$ であり, これと先の式を合わせて $ax_2 = bx_1 = b(ra) = rba$. $a \neq 0$ ゆえ両辺を a で割り, $x_2 = rb$ を得る. 以上より, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ra \\ rb \end{bmatrix} = r\mathbf{y}$. ゆえに $\mathbf{x} \in W$ である. $b \neq 0$ の場合は, $r = \frac{x_2}{b}$ とおいて同様の計算をすれば $\mathbf{x} \in W$ を得る. □

16.3 部分空間の様々な例

数学の諸分野で扱われる部分空間の例を紹介しよう.

例 16.3.1. 例 15.2.2 で与えた数列空間 $\mathbb{R}^{\mathbb{N}}$ の部分集合で, 次の漸化式を満たす数列全体を F とする:

$$x_{n+2} = x_{n+1} + x_n. \quad (n \in \mathbb{N}) \tag{16.3.1}$$

このとき, F は $\mathbb{R}^{\mathbb{N}}$ の部分空間である.

Proof. (i): すべての項が 0 なる数列 $\mathbf{0} = (0, 0, 0, \dots)$ は漸化式 (16.3.1) を満たす. ゆえに $\mathbf{0} \in F$.

(iv): $r, s \in \mathbb{R}$ とし, 数列 $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ および $\mathbf{y} = (y_n)_{n \in \mathbb{N}}$ が漸化式 (16.3.1) を満たすと仮定する. すなわち, $x_{n+2} = x_{n+1} + x_n$ および $y_{n+2} = y_{n+1} + y_n$ が成立している. この二式をそれぞれ r 倍, s 倍して和をとることで

$$rx_{n+2} + sy_{n+2} = (rx_{n+1} + sy_{n+1}) + (rx_n + sy_n)$$

を得る. これは数列 $r\mathbf{x} + s\mathbf{y} = (rx_n + sy_n)_{n \in \mathbb{N}}$ が漸化式 (16.3.1) を満たすことに他ならない. ゆえに $r\mathbf{x} + s\mathbf{y} \in F$. □

漸化式 16.3.1 を満たす数列の中で, 初項が 0 かつ第 2 項が 1 なる数列

$$a_1 = 0, a_2 = 1, a_3 = 1, a_4 = 2, a_5 = 3, a_6 = 5, a_7 = 8, a_8 = 13, a_9 = 21, \dots$$

をフィボナッチ数列 (Fibonacci sequence) という.

練習 16.3.2. 自然数 $k \geq 2$, および k 個の実数 a_0, a_1, \dots, a_{k-1} を固定する. 数列空間 $\mathbb{R}^{\mathbb{N}}$ の部分集合で, 次の漸化式を満たす数列全体を W とする:

$$x_{n+k} = a_{k-1}x_{n+k-1} + a_{k-2}x_{n+k-2} + \dots + a_1x_{n+1} + a_0x_n \quad (n \in \mathbb{N}). \quad (16.3.2)$$

このとき, W が $\mathbb{R}^{\mathbb{N}}$ の部分空間であることを示せ.

上の漸化式のことを線形漸化式 (**linear recurrence relation**) と呼び, これを満たす数列のことを線形回帰数列 (**linear recurrence sequence**) という.

次に, 微分を用いた部分空間の例を挙げる.

例 16.3.3. (1) \mathbb{R} 上の开区間 I を定義域とする C^∞ -級関数⁶全体のなす集合を $C^\infty(I)$ と書く. 定数関数 $\mathbf{0}$ は C^∞ -関数であり, また C^∞ -級関数の和やスカラー倍は再び C^∞ -関数となる. もちろん C^∞ -級関数は連続関数である. 以上より $C^\infty(I)$ は例 15.2.7 で与えた $C(I)$ の部分空間である.

(2) $C^\infty(I)$ の元の中で, 次の微分方程式を満たす関数 $y = y(x)$ 全体を W とする:

$$y^{(2)}(x) = y^{(1)}(x) + y^{(0)}(x). \quad (16.3.3)$$

ここで, $y^{(n)}(x) = \frac{d^n}{dx^n}y(x)$ は y の n 階導関数のことであり, $y^{(0)}(x) = y(x)$ とする. このとき, W は $C^\infty(I)$ の部分空間である.

Proof. (i): 定数関数 $\mathbf{0}$ の微分は再び $\mathbf{0}$ となることから, $\mathbf{0}$ が式 16.3.3 を満たすことは明らかである. ゆえに $\mathbf{0} \in W$.

(iv): $f, g \in W$ とすれば関数 f, g は式 (16.3.3) を満たす. すなわち, $f^{(2)}(x) = f^{(1)}(x) + f^{(0)}(x)$, $g^{(2)}(x) = g^{(1)}(x) + g^{(0)}(x)$ である. 各 $r, s \in \mathbb{R}$ に対して, この二式をそれぞれ r 倍, s 倍して和をとることで

$$\begin{aligned} rf^{(2)}(x) + sg^{(2)}(x) &= \left(rf^{(1)}(x) + sg^{(1)}(x) \right) + \left(rf^{(0)}(x) + sg^{(0)}(x) \right) \\ (rf(x) + sg(x))^{(2)} &= (rf(x) + sg(x))^{(1)} + (rf(x) + sg(x))^{(0)} \end{aligned}$$

を得る. これは関数 $rf(x) + sg(x)$ が微分方程式 (16.3.3) を満たすことに他ならない. ゆえに $rf + sg \in W$. \square

練習 16.3.4. 自然数 $k \geq 2$, および k 個の実数 a_0, a_1, \dots, a_{k-1} を固定する. 次の微分方程式を満たす関数 $y = y(x)$ 全体を W とする. W が $C^\infty(I)$ の部分空間であることを示せ.

$$y^{(k)}(x) = a_{k-1}y^{(k-1)}(x) + a_{k-2}y^{(k-2)}(x) + \dots + a_1y^{(1)}(x) + a_0y^{(0)}(x). \quad (16.3.4)$$

方程式 (16.3.4) は実数係数の線形常微分方程式 (**linear ordinary differential equation**) と呼ばれる.

式 (16.3.3) の両辺を n 回微分することで $y^{(n+2)}(x) = y^{(n+1)}(x) + y^{(n)}(x)$ を得る. ここに漸化式 (16.3.1) との類似性が伺えよう. 同様にして, 微分方程式 (16.3.4) から

$$y^{(n+k)}(x) = a_{k-1}y^{(n+k-1)}(x) + a_{k-2}y^{(n+k-2)}(x) + \dots + a_1y^{(n+1)}(x) + a_0y^{(n)}(x)$$

が得られ, これは漸化式 (16.3.2) と類似している. これらは式として単に似ているというだけでなく, 線形回帰数列の一般項の導出と線形常微分方程式の一般解の解法が線形代数学の枠組みにおいて並行して得られることが後に理解されるであろう (31 章および 32 章).

例 16.3.5 (発展). (1) $\mathbb{R}[x]$ の中で, 多項式 $p(x) = x - 1$ で割り切れるもの全体を集めた集合を (p) と書く. すると (p) は $\mathbb{R}[x]$ の部分空間である. この議論において p が特別な多項式 $x - 1$ である必要はなく, 任意の多項式 $p(x)$ について同様の議論が成り立つ.

⁶何回でも微分できる関数のことを C^∞ -関数という

(2) \mathbb{R} 上の区間 I の元 t を一つ固定しておく, さらに, I を定義域とする連続関数 $f: I \rightarrow \mathbb{R}$ で, $f(t) = 0$ を満たすもの全体からなる集合を \mathcal{I}_t とおこう. このとき, \mathcal{I}_t は $C(I)$ の部分空間となる.

上の例において, $\mathbb{R}[x]$ の元と (p) の元の積を取ると再び (p) の元に含まれることがわかる. また, $C(I)$ の元と \mathcal{I}_t の元の積もまた \mathcal{I}_t に含まれる. このような特別な集合はイデアル (ideal) と呼ばれ, 環論と呼ばれる代数分野において広く調べられている.

第17章 線形結合と線形独立性

線形空間の定義から直ちに導かれる基本的性質、すなわちベクトルの和とスカラー倍の性質について論じよう。本章で述べることは、形のうえではベクトルの式変形を繰り返すことに尽きる。しかし、これを単なる計算と見るのではなく、複数のベクトルの間の関係性として捉えることで、技法から理論へと考え方が昇華されるのである。これによって、より高い見地から線形空間を捉えられるようになる。なお、ベクトルたちの関係を調べるうえで技術的な部分のいくつかは連立1次方程式の掃き出し法による解法に帰着される。掃き出し法や行列の簡約化をしっかり復習したうえで本章に臨んでもらいたい。

17.1 線形結合

定義 17.1.1. 線形空間 V の元 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n \in V$ および $a_1, a_2, \dots, a_n \in \mathbb{R}$ を用いた式

$$a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_n \mathbf{u}_n = \sum_{i=1}^n a_i \mathbf{u}_i$$

を $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ による線形結合 (linear combination) あるいは1次結合という。このとき、上の線形結合における各 a_i ($i = 1, \dots, n$) を \mathbf{u}_i の係数 (coefficient) と呼ぶ。また、零ベクトルを線形結合で表す式 $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}$ が成立するとき、この等式を $\mathbf{u}_1, \dots, \mathbf{u}_n$ による線形関係あるいは1次関係という。

線形結合の標語的な解釈

\mathbf{v} が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書かれるとは、 \mathbf{v} の情報が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の情報に分解されていると考えればよい。もう少し詳しく述べれば、 \mathbf{v} の情報を得るには、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の情報と線形結合に現れる係数 a_1, \dots, a_n の値さえ分かれば十分ということである。なお、ここでいう「情報」とは、分析すべき線形写像における値のことを意味する。

例 17.1.2. \mathbb{R}^n の任意のベクトルは、次の n 個のベクトルの線形結合で書くことができる:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

実際、各 $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$ について $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_n \end{bmatrix} = a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n = \sum_{i=1}^n a_i \mathbf{e}_i$ で

ある (実は、同様の分解を式 12.4.1 において既に行っている)。上の $\mathbf{e}_1, \dots, \mathbf{e}_n$ を \mathbb{R}^n の標準ベクトルまたは基本ベクトルという。

ベクトル \mathbf{v} が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けるとしよう。 \mathbf{v} の情報を $\mathbf{u}_1, \dots, \mathbf{u}_n$ たちの情報に還元する際に、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中でその情報が不要なものがあるかもしれない。例えば次のような状況が考えられる。

(1) $\mathbf{u}_1, \dots, \mathbf{u}_n$ のいずれかの情報が、ほかのベクトルたちの情報に分解できる場合.

例えば、 \mathbf{u}_n の情報が $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ に分解されるとき、 \mathbf{v} の情報は $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の情報だけから復元できることが示唆され、 \mathbf{u}_n は不要となる. 実際、 \mathbf{v} が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けており、更に \mathbf{u}_n が $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の線形結合で書けるならば、 \mathbf{v} は $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の線形結合で書ける:

$\mathbf{v} = \sum_{i=1}^n a_i \mathbf{u}_i$, $\mathbf{u}_n = \sum_{i=1}^{n-1} b_i \mathbf{u}_i$ と書けるならば、

$$\mathbf{v} = \sum_{i=1}^n a_i \mathbf{u}_i = \left(\sum_{i=1}^{n-1} a_i \mathbf{u}_i \right) + a_n \mathbf{u}_n = \left(\sum_{i=1}^{n-1} a_i \mathbf{u}_i \right) + a_n \left(\sum_{i=1}^{n-1} b_i \mathbf{u}_i \right) = \sum_{i=1}^{n-1} (a_i + a_n b_i) \mathbf{u}_i. \quad (17.1.1)$$

(2) 線形結合に現れる係数 a_i のうちのいくつかは 0 の場合.

このとき、 $a_i = 0$ に対応する \mathbf{u}_i は不要となる. 例えば、 $\mathbf{a} = {}^t(5, 3, 0) \in \mathbb{R}^3$ を標準ベクトルに分解すると、 $\mathbf{a} = 5\mathbf{e}_1 + 3\mathbf{e}_2 + 0\mathbf{e}_3 = 5\mathbf{e}_1 + 3\mathbf{e}_2$ ゆえ \mathbf{e}_3 は不要である. しかしこれは、特別なベクトルを考えたからたまたま \mathbf{e}_3 が不要になったのであり、 \mathbb{R}^3 の別のベクトルを分解しようと思えば、 \mathbf{e}_3 が必要になることもある. とくに、 \mathbb{R}^3 のすべてのベクトルを分解しようと思えば、 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ のいずれも必須であることが分かる. この考察から、(2) の立場で \mathbf{u}_n が不要ということは、線形空間 V のいかなる元も \mathbf{u}_n を用いずに $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ のみによって分解できることと捉えるべきである. 仮にこの意味で \mathbf{u}_n が不要になる場合、 \mathbf{u}_n はとくに V の元であるから $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ たちの線形結合で書ける. すなわち、この議論の大筋については (1) に帰着される.

【注意】 逆に、(1) は (2) に帰着するとも説明できる. 実際、式 (17.1.1) の右辺は、 \mathbf{u}_n の係数が 0 となる線形結合ともみなせる.

ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に不要なものがあるかないかという状況を数学的な言葉で表すために、線形独立性なる概念を導入する.

17.2 線形独立性

以下、断りがなくとも V は線形空間であるとする.

定義 17.2.1. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ が線形独立¹(linearly independent) であるとは、

$$a_1, \dots, a_n \in \mathbb{R}, a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_n \mathbf{u}_n = \mathbf{0} \implies a_1 = a_2 = \dots = a_n = 0$$

が成り立つことをいう. 線形独立でないベクトルの組を線形従属 (linearly dependent) であるという. すなわち、少なくともいずれか一つは 0 でないような実数の組 a_1, \dots, a_n を用いて、 $a_1 \mathbf{u}_1 + \dots + a_n \mathbf{u}_n = \mathbf{0}$ と表せることである.

線形従属の定義に現れた、少なくともいずれか一つは 0 でないような実数の組 a_1, \dots, a_n のことを自明でない組という. 実数の組 a_1, \dots, a_n が自明でないことは、 $(a_1, \dots, a_n) \neq \mathbf{0}$ であることに等しい. また、自明でない実数の組 a_1, \dots, a_n を係数とする線形関係を自明でないという. 更にこれらの否定概念として、自明な実数の組、および自明な線形関係を定める. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立であるとは、それらによる線形関係が自明なものに限られることである. また、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形従属であるとは、自明でない線形関係が存在することである.

例 17.2.2. (1) \mathbb{R}^n の標準ベクトルの組 $\mathbf{e}_1, \dots, \mathbf{e}_n$ は線形独立である.

¹線形独立 (線形従属) は 1 次独立 (1 次従属) 独立とも呼ばれる.

Proof. $a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n = \mathbf{0}$ とすれば $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ ゆえ $a_1 = \cdots = a_n = 0$ である. \square

(2) m 次列ベクトル $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立であることと, m 次行ベクトル ${}^t\mathbf{u}_1, \dots, {}^t\mathbf{u}_n$ が線形独立であることは同値である. 実際, 式 $\sum_{i=1}^n a_i\mathbf{u}_i = \mathbf{0}_{m,1}$ と式 $\sum_{i=1}^n a_i {}^t\mathbf{u}_i = \mathbf{0}_{1,m}$ は互いに両辺を転置した関係にある. ゆえに, これらのベクトルの組に関する線形関係は同等である.

(3) $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に零ベクトルがあれば, これらは線形従属である. 例えば $\mathbf{u}_n = \mathbf{0}$ の場合, 自明でない実数の組 $0, \dots, 0, 1$ を係数とする線形関係 $0\mathbf{u}_1 + \cdots + 0\mathbf{u}_{n-1} + 1\mathbf{u}_n = \mathbf{0}$ が成立する. また, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に重複がある場合も線形従属である. 例えば $\mathbf{u}_1 = \mathbf{u}_2$ の場合, 自明でない実数の組 $1, -1, 0, \dots, 0$ を係数とする線形関係 $1\mathbf{u}_1 + (-1)\mathbf{u}_2 + 0\mathbf{u}_3 + \cdots + 0\mathbf{u}_n = \mathbf{0}$ が成立する.

例 17.2.3. 多項式環 $\mathbb{R}[x]$ における 3 つの元 $x^2, x, \mathbf{1}$ は線形独立である. ここで, $\mathbf{1}$ とはどんな数を代入しても $1 \in \mathbb{R}$ に値を取る定数関数とする.

Proof. $a_2, a_1, a_0 \in \mathbb{R}, a_2x^2 + a_1x + a_0\mathbf{1} = \mathbf{0}$ と仮定し, $a_2 = a_1 = a_0 = 0$ を示そう. $a_2x^2 + a_1x + a_0\mathbf{1} = \mathbf{0}$ の両辺に適当な数を三つほど代入して連立 1 次方程式を解くことで $a_2 = a_1 = a_0 = 0$ を示せるが (命題 15.2.6), ここでは微分による証明を紹介しよう. $a_2x^2 + a_1x + a_0\mathbf{1} = \mathbf{0}$ の両辺を微分すると $2a_2x + a_1 = \mathbf{0}$ であり, これを更に微分することで $2a_2 = \mathbf{0}$ を得る. つまり, いかなる数を代入しても $2a_2$ に値を取る定数関数 (左辺) と 0 に値を取る定数関数 (右辺) は等しい. ゆえに $a_2 = 0$ である. これを $2a_2x + a_1 = \mathbf{0}$ に代入して $a_1 = 0$ を得る. これらをもとの式 $a_2x^2 + a_1x + a_0\mathbf{1} = \mathbf{0}$ に代入することで $a_0 = 0$ を得る. \square

練習 17.2.4. $n + 1$ 個の組 $x^n, x^{n-1}, \dots, x, \mathbf{1} \in \mathbb{R}[x]$ が線形独立であることを帰納法を用いて示せ.

次の命題は前節の最後で考察した要不要論と線形独立性 (従属性) の関係を述べている. すなわち, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に不要なものがあることと $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形従属性は同値である. また, それらの否定を取り, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に不要なものがないことと $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形独立性は同値である.

命題 17.2.5. n 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ について次の (1) と (2) は同値である.

- (1) $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である,
- (2) $\mathbf{u}_1, \dots, \mathbf{u}_n$ のうち少なくとも一つのベクトルが他の $n - 1$ 個の線形結合で書ける.

Proof. (1) \Rightarrow (2): $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形従属であるとすれば, 自明でない実数の組 a_1, \dots, a_n を用いて $a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0}$ とできる. このとき a_1, \dots, a_n のうちいずれか一つは 0 ではない. 例えば $a_n \neq 0$ として話を進めよう. このとき, 移項により $a_n\mathbf{u}_n = -a_1\mathbf{u}_1 - \cdots - a_{n-1}\mathbf{u}_{n-1}$ を得る. $a_n \neq 0$ ゆえこの両辺を a_n で割れば $\mathbf{u}_n = \left(-\frac{a_1}{a_n}\right)\mathbf{u}_1 + \cdots + \left(-\frac{a_{n-1}}{a_n}\right)\mathbf{u}_{n-1}$. ゆえに \mathbf{u}_n は, $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の線形結合で書ける. $a_i \neq 0$ ($i = 1, \dots, n - 1$) である場合も同様にして, \mathbf{u}_i が他の $n - 1$ 個の線形結合で書けることが示される.

(2) \Rightarrow (1): \mathbf{u}_n が $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の線形結合で書ける場合を考えよう. このとき, ある実数の組 a_1, \dots, a_{n-1} を用いて $\mathbf{u}_n = a_1\mathbf{u}_1 + \cdots + a_{n-1}\mathbf{u}_{n-1}$ と書ける. これを移項して $a_1\mathbf{u}_1 + \cdots + a_{n-1}\mathbf{u}_{n-1} + (-1)\mathbf{u}_n = \mathbf{0}$ を得る. 自明でない線形関係が得られたゆえ, $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である. \mathbf{u}_n 以外のベクトルが他の $n - 1$ 個の線形結合で書ける場合についても同様の議論により $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形従属性を得る. \square

例 17.2.6. (1) \mathbb{R}^2 の $\mathbf{0}$ でない列ベクトル \mathbf{u}, \mathbf{v} が線形従属であるとは, 一方がもう一方の線形結合で書けるということであるから, これは実数 $r \neq 0$ を用いて $\mathbf{u} = r\mathbf{v}$ と書けること, すなわち \mathbf{u}, \mathbf{v} が原点を通る同一直線上にあることを意味する.

(2) 一方, \mathbb{R}^2 の列ベクトル \mathbf{u}, \mathbf{v} が線形独立であるとは, \mathbf{u}, \mathbf{v} が原点を通る同一直線上にないこと, つまり \mathbf{u}, \mathbf{v} で張られる平行四辺形が面積を持つことを意味する. これは行列 $A = [\mathbf{u}, \mathbf{v}]$ が可逆であることにほかならない. この事実を n 次に一般化した場合の証明は命題 17.3.6 で与える.

- (3) \mathbb{R}^2 の 3 つのベクトル $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ は線形従属である. これは例えば, $\mathbf{u}_3 = \frac{2}{5}\mathbf{u}_1 + \frac{1}{5}\mathbf{u}_2$ と書けることから分かる. 実は, \mathbb{R}^n の $n+1$ 個のベクトルの組は必ず線形従属になる (命題 17.3.5).

練習 17.2.7. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ について次を示せ.

- (1) n 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立ならば, そこから一つ取り除いた $n-1$ 個の組 $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ も線形独立である.

解答例: $a_1\mathbf{u}_1 + \dots + a_{n-1}\mathbf{u}_{n-1} = \mathbf{0}$ とすれば, $a_n = 0$ とおくことで $a_1\mathbf{u}_1 + \dots + a_{n-1}\mathbf{u}_{n-1} + a_n\mathbf{u}_n = \mathbf{0}$ を得る. $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形独立性より $a_1 = \dots = a_{n-1} = a_n = 0$ であり, 特に $a_1 = \dots = a_{n-1} = 0$. ゆえに $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ は線形独立である. \square

- (2) $n-1$ 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ が線形従属ならば, そこに新たなベクトル \mathbf{u}_n を加えた n 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}, \mathbf{u}_n$ も線形従属である.

解答例: (2) は (1) の対偶にほかならない. \square

17.3 線形独立性の判定

\mathbb{R}^m のベクトルの組の線形独立性の判定法を与える.

命題 17.3.1. n 個の m 次列ベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_n$ が線形独立であることと, (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ に関する斉次形連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ が唯一解を持つことは同値である.

Proof. $\mathbf{a}_1, \dots, \mathbf{a}_n$ が線形独立であるとする. n 次列ベクトル $\mathbf{x} = {}^t(r_1, \dots, r_n)$ が $A\mathbf{x} = \mathbf{0}$ を満たすならば,

$$r_1\mathbf{a}_1 + \dots + r_n\mathbf{a}_n = [\mathbf{a}_1, \dots, \mathbf{a}_n] \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} = A\mathbf{x} = \mathbf{0}.$$

$\mathbf{a}_1, \dots, \mathbf{a}_n$ の線形独立性より $r_1 = \dots = r_n = 0$ である. つまり $\mathbf{x} = \mathbf{0}$ であり, 方程式 $A\mathbf{x} = \mathbf{0}$ の解は自明なものに限る.

次に, 方程式 $A\mathbf{x} = \mathbf{0}$ が自明な解しか持たないと仮定する. このとき $r_1\mathbf{a}_1 + \dots + r_n\mathbf{a}_n = \mathbf{0}$ ならば $\mathbf{x} = {}^t(r_1, \dots, r_n)$ について $A\mathbf{x} = r_1\mathbf{a}_1 + \dots + r_n\mathbf{a}_n = \mathbf{0}$ となる. つまり $A\mathbf{x} = \mathbf{0}$ であり, この方程式は自明な解しか持たないゆえ $\mathbf{x} = \mathbf{0}$ を得る. すなわち $r_1 = \dots = r_n = 0$ である. 以上より $\mathbf{a}_1, \dots, \mathbf{a}_n$ は線形独立である. \square

いまの議論を連立 1 次方程式の解法まで戻って詳しくみると, 線形独立性の判定だけではなく, 線形従属である場合にどのベクトルが他のベクトルの線形結合で書けるかも分かる. これを次の命題を通して見てみよう.

命題 17.3.2. (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ を行基本変形により $B = [\mathbf{b}_1, \dots, \mathbf{b}_n]$ に変形できるとする. $1 \leq n_1, n_2, \dots, n_\ell \leq n$ および $i = 1, \dots, n$ に対して次が成り立つ.

$$(1) \sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k} = \mathbf{0} \iff \sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k} = \mathbf{0}.$$

$$(2) \mathbf{a}_i = \sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k} \text{ と書ける} \iff \mathbf{b}_i = \sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k} \text{ と書ける.}$$

Proof. (1): A からいくつかの列を間引いた行列 $A' = [\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_\ell}]$ を考えれば, A を B に行基本変形できることから, A' は $B' = [\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_\ell}]$ に行基本変形できる. また, $[A'|\mathbf{0}]$ を $[B'|\mathbf{0}]$ に行基本変形できる. ゆえに命題 4.4.1 より $A'\mathbf{x} = \mathbf{0}$ の解と $B'\mathbf{x} = \mathbf{0}$ の解は一致しており, $\mathbf{x} = {}^t(r_1, \dots, r_n)$ について

$$\sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k} = \mathbf{0} \iff \mathbf{x} \text{ は } A' \mathbf{x} = \mathbf{0} \text{ の解} \iff \mathbf{x} \text{ は } B' \mathbf{x} = \mathbf{0} \text{ の解} \iff \sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k} = \mathbf{0}.$$

(2): $\mathbf{a}_i = \sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k}$ とすれば, $\sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k} + (-1) \mathbf{a}_i = \mathbf{0}$ である. $\ell + 1$ 個のベクトルの組 $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_\ell}, \mathbf{a}_i$ について (1) を適用し, $\sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k} + (-1) \mathbf{b}_i = \mathbf{0}$ を得る. すなわち, $\mathbf{b}_i = \sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k}$. 逆も同様に示される. \square

上の (1) は組 $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k}$ が線形独立 (あるいは線形従属) であることと組 $\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_k}$ が線形独立 (あるいは線形従属) であることの同値性を述べている. ベクトルの組が線形従属である場合, 命題 17.2.5 により, いずれかのベクトルが他のベクトルの線形結合で書ける. 上の命題を応用して, どのベクトルが他のベクトルの線形結合で書けるか調べてみよう:

例題 17.3.3. 次の列ベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_5$ が線形独立であるかどうか判定し, 線形従属の場合はどのベクトルが他のベクトルの線形結合で書けるか答えよ.

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{a}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{a}_5 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ -3 \\ -1 \end{bmatrix}.$$

解答例: 与えられた列ベクトルを並べた行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_5]$ の簡約化を $B = [\mathbf{b}_1, \dots, \mathbf{b}_5]$ とすれば,

$$A = \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 & -3 \\ 1 & 1 & 0 & 0 & -1 \end{bmatrix} \longrightarrow B = \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$\mathbf{b}_1, \dots, \mathbf{b}_5$ が線形従属であることは成分を見れば明らかであり (線形従属性は $\text{rank } A = 3 \neq 5$ より方程式 $A\mathbf{x} = \mathbf{0}$ が自明でない解をもつことから分かる), ゆえに $\mathbf{a}_1, \dots, \mathbf{a}_5$ も線形従属である. 行列 B の成分を見れば $\mathbf{b}_3 = -\mathbf{b}_1 + \mathbf{b}_2$, $\mathbf{b}_5 = -2\mathbf{b}_1 + \mathbf{b}_2 - \mathbf{b}_4$ と書けることが分かる. したがって $\mathbf{a}_3 = -\mathbf{a}_1 + \mathbf{a}_2$, $\mathbf{a}_5 = -2\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_4$ である.

【注意】

- (1) B の成分を見れば, $\mathbf{b}_2 = 2\mathbf{b}_1 + \mathbf{b}_4 + \mathbf{b}_5$ と書けることが分かる. つまり, 他のベクトルで書けるものは $\mathbf{b}_3, \mathbf{b}_5$ に限るというわけではない. 上で $\mathbf{b}_3, \mathbf{b}_5$ を取り上げたのは, これ以外の主成分を含む列 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ が標準ベクトルであることから, $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ が線形独立であること, および $\mathbf{b}_3, \mathbf{b}_5$ が $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ の線形結合で書けることが直ちに分かるゆえである. 例えば, 組 $\mathbf{b}_1, \mathbf{b}_4, \mathbf{b}_5$ も線形独立であり, これ以外の $\mathbf{b}_2, \mathbf{b}_3$ を $\mathbf{b}_1, \mathbf{b}_4, \mathbf{b}_5$ の線形結合で表すこともできるが, それを示すのは標準ベクトルの組 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ に対して行うより骨が折れるであろう. より詳しい事情は次章の基底概念を通して説明される.
- (2) あらかじめ $\mathbf{a}_4, \mathbf{a}_5$ が線形独立であることが分かっており, これらを用いて他のベクトルを線形結合で表したい場合は, 列を並び替えて $\mathbf{a}_4, \mathbf{a}_5$ を先頭にした行列 $[\mathbf{a}_4, \mathbf{a}_5, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$ について同様の計算を行えばよい.
- (3) 行ベクトルについて同様の問題を考える場合は転置して列ベクトルの問題に変換し, 得られた答えを再び転置して行ベクトルに直せばよい.

(m, n) -行列 A による連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ の解が唯一解を持つかどうかは, 命題 6.2.1 により行列の階数を用いて判定できる. その条件は $\text{rank}[A|\mathbf{0}] = \text{rank } A = n$ である. 階数の定義から A がいかなる行列であろうと $\text{rank}[A|\mathbf{0}] = \text{rank } A$ であり, したがって条件 $\text{rank } A = n$ が $A\mathbf{x} = \mathbf{0}$ の解が唯一であるための同値条件である. 以上より次を得る:

系 17.3.4. \mathbb{R}^m の n 個の m 次列ベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_n$ が線形独立であることと, $\text{rank}[\mathbf{a}_1, \dots, \mathbf{a}_n] = n$ であることは必要十分である.

命題 17.3.5. $m < n$ について, \mathbb{R}^m の n 個の列ベクトル $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である.

Proof. (m, n) -行列 $A = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ について, 式 (6.1.1) より $\text{rank } A \leq m < n$ であり, とくに $\text{rank } A \neq n$. 系 17.3.4 より $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である. \square

正方行列に現れるベクトルの組においては次が成り立ち, これらの条件を定理 6.2.2 につけ加えることができる.

定理 17.3.6. n 次正方行列 A について次は同値である.

(1) A は可逆である, (2) A の各列は線形独立である, (3) A の各行は線形独立である.

Proof. (1) \Leftrightarrow (2): 定理 6.2.2 および命題 17.3.1 から直ちに得られる:

A は可逆 $\Leftrightarrow A\mathbf{x} = \mathbf{0}$ は唯一解をもつ $\Leftrightarrow A$ の各列は線形独立.

(1) \Leftrightarrow (3): いま示した (1) と (2) の同値性および $|A| = |{}^tA|$ より得られる:

A は可逆 $\Leftrightarrow |A| \neq 0 \Leftrightarrow |{}^tA| \neq 0 \Leftrightarrow {}^tA$ は可逆
 $\Leftrightarrow {}^tA$ の各列は線形独立 $\Leftrightarrow A$ の各行は線形独立.

\square

第18章 基底

17.1節で述べた要不要論を思い出そう。この議論の(1)に関係のある概念として線形独立性を前章で与えた。本章では要不要論の(2)で論じたことと関係する、ベクトルの組による生成について述べる。これは、 V の任意の元が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けるかどうかを定式化する概念である。これに線形独立性を合わせたものが基底であり、一般の線形空間における基底は、ユークリッド空間における座標軸のような役割を果たす。

18.1 ベクトルの組が生成する部分空間

与えられたベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に V の任意の元が分解できるかどうかはともかくとして、まず、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に分解できるベクトルの範囲を表す記号を導入しよう。

定義 18.1.1. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ において、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ たちの線形結合で書けるようなベクトルをすべて集めた V の部分集合を $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ と書く。すなわち、 $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle := \{ \sum_{i=1}^n a_i \mathbf{u}_i \mid a_1, \dots, a_n \in \mathbb{R} \}$ である。次の命題により、これは V の部分空間となる。 $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ は、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ によって生成される部分空間と呼ばれる。

命題 18.1.2. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ において、 $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ は V の部分空間である。

Proof. 部分空間となるための条件(i)および(iv)を確認すればよい。 $\mathbf{0}$ は $\mathbf{u}_1, \dots, \mathbf{u}_n$ たちの線形結合で書けるゆえ $\mathbf{0} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ である。次に、 $\mathbf{x}, \mathbf{y} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ とすれば $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$, $\mathbf{y} = \sum_{i=1}^n b_i \mathbf{u}_i$ と書ける。このとき、各 $r, s \in \mathbb{R}$ に対して $r\mathbf{x} + s\mathbf{y} = \sum_{i=1}^n (ra_i + sb_i)\mathbf{u}_i$ ゆえ $r\mathbf{x} + s\mathbf{y}$ も $\mathbf{u}_1, \dots, \mathbf{u}_n$ たちの線形結合で書ける。よって $r\mathbf{x} + s\mathbf{y} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ である。□

次は、要不要論の(1)で述べたことを、より一般的な状況に置き換えた主張である。

命題 18.1.3. V を線形空間とする。各 $\mathbf{v}_1, \dots, \mathbf{v}_\ell \in V$ が組 $\mathbf{w}_1, \dots, \mathbf{w}_n \in V$ の線形結合で書けるならば、 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で書ける元は $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で書ける。すなわち、

$$\mathbf{v}_1, \dots, \mathbf{v}_\ell \in \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle \implies \langle \mathbf{v}_1, \dots, \mathbf{v}_\ell \rangle \subset \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle.$$

Proof. $W = \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle$ とおく。 W は V の部分空間であり、命題16.1.5より W は線形結合について閉じている。仮定より $\mathbf{v}_1, \dots, \mathbf{v}_\ell \in W$ であるから、 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で書ける元は W に含まれる。ゆえに $\langle \mathbf{v}_1, \dots, \mathbf{v}_\ell \rangle \subset \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle$ 。□

上の証明では、 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で書いたときに現れる係数と $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で書いたときに現れる各係数の関係については論じなかった。これらの係数の関係は行列の積演算を通して得られる(25.1節を見よ)。

練習 18.1.4. $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ とし、 $\langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle = \mathbb{R}^n$ であるとする。また、 A を (m, n) -行列とする。このとき、各 $i = 1, \dots, k$ について $A\mathbf{v}_i = \mathbf{0}$ が成り立つならば、 $A = O$ となることを示せ。

解答例. A の各列ベクトルが零ベクトルであることを示せばよい。 A の j 列目は $A\mathbf{e}_j$ である。ここで、 $\mathbf{e}_j \in \mathbb{R}^n = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle$ より \mathbf{e}_j は $\mathbf{v}_1, \dots, \mathbf{v}_k$ の線形結合で書ける。つまり $\mathbf{e}_j = \sum_{i=1}^k r_i \mathbf{v}_i$ と表せる。ゆえに $A\mathbf{e}_j = A \left(\sum_{i=1}^k r_i \mathbf{v}_i \right) = \sum_{i=1}^k A(r_i \mathbf{v}_i) = \sum_{i=1}^k r_i A\mathbf{v}_i = \sum_{i=1}^k r_i \mathbf{0} = \mathbf{0}$ 。

18.2 基底の定義と例

次で定める基底とは、線形空間 V の各元を線形結合で表すときに過不足なく必要になるベクトルの組のことである。

定義 18.2.1. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle = V$ を満たすとき、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ は V を生成するという。また、 V を生成するような線形独立な組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を V の基底 (basis) あるいは基という。

例 18.2.2. (1) \mathbb{R}^n の標準ベクトルの組 $\mathbf{e}_1, \dots, \mathbf{e}_n$ は \mathbb{R}^n を生成し、かつ線形独立である (例 17.1.2 および例 17.2.2)。ゆえに $\mathbf{e}_1, \dots, \mathbf{e}_n$ は \mathbb{R}^n の基底であり、これを標準基底という。

(2) (i, j) -成分が 1 でそれ以外の成分がすべて 0 の (m, n) -行列を E_{ij} と書けば (ただし $1 \leq i \leq m$, $1 \leq j \leq n$)、 mn 個の行列の組 E_{ij} ($i = 1, \dots, m, j = 1, \dots, n$) は線形独立かつ $M_{m,n}(\mathbb{R})$ を生成する。ゆえにこれらは $M_{m,n}(\mathbb{R})$ の基底である。

(3) $\mathbb{R}[x]_n$ における $n+1$ 個の組 $x^n, x^{n-1}, \dots, x, \mathbf{1}$ は $\mathbb{R}[x]_n$ を生成し、かつ線形独立である (例 17.2.2(3))。ゆえに $\mathbb{R}[x]_n$ の基底である。

(4) 約束として、自明な空間 $\{\mathbf{0}\}$ は 0 個のベクトルの組からなる基底をもつとする。

ユークリッド空間 \mathbb{R}^n には座標、すなわち個々のベクトルの位置を示す情報が与えられていた。これに対して一般の線形空間においては、ベクトルの位置を定めるために基底が用いられる。 V の各元 \mathbf{v} について、基底による \mathbf{v} の線形結合表示に現れる係数を \mathbf{v} の位置情報とみなすのである。とくに \mathbb{R}^n の各元を標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ の線形結合によって書いた際に現れる各係数は、その位置を示す座標の各成分に一致している。 \mathbb{R}^n の通常の座標において、その表示が異なれば違う位置を示していたように、基底による線形結合の各係数に現れる実数の組が異なれば、線形結合が表す位置も当然異なっているべきであろう。このことは次の命題が保証している。

命題 18.2.3. $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ が線形独立であるとし、 $\mathbf{v} \in V$ がこれらの線形結合で書けるとすれば、その表し方は一通りしかない。すなわち、 $\mathbf{v} = \sum_{i=1}^n a_i \mathbf{u}_i = \sum_{i=1}^n b_i \mathbf{u}_i$ ならば $a_i = b_i$ ($i = 1, \dots, n$) である。

Proof. $\sum_{i=1}^n a_i \mathbf{u}_i = \sum_{i=1}^n b_i \mathbf{u}_i$ を移項すると線形関係 $\sum_{i=1}^n (a_i - b_i) \mathbf{u}_i = \mathbf{0}$ を得る。 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形独立性より $a_1 - b_1 = \dots = a_n - b_n = 0$ 。つまり $a_i = b_i$ ($i = 1, \dots, n$) である。□

\mathbb{R}^n の座標、つまり標準基底による位置情報と、一般の基底によるその違いを述べておこう。 \mathbb{R}^n の座標軸がそれぞれ直交するのに対して、基底を構成する各ベクトルは必ずしも直交するわけではない。そもそも、一般の線形空間においては直交なる概念が定まるとは限らないともいえる。また、 \mathbb{R}^n においては座標軸が自然な形で (先天的に) 定まるのに対して、次の例にあるように線形空間の基底の取り方は無数にある。言い換えると、線形空間においては座標軸に似た概念 (すなわち基底) を多様に定めることができる。そして、これまでとは異なる基底を与えることは、座標軸を取り換えることに相当する。基底の取り換えの前後におけるベクトルの位置情報の変化を見るための技術は本書の後半において重要な役割を担い、その一般論は 25 章で述べる。

このような事情から、基底による表示を用いた議論を行う場合、はじめにどんな基底を考えているか宣言する必要がある。以後、本書における多くの命題もこのような形で述べることになるだろう。

例 18.2.4. $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ は可逆ゆえ命題 17.3.1 より列ベクトルの組 $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ は線形独立である。 \mathbb{R}^2 の任意のベクトル \mathbf{v} は標準ベクトル $\mathbf{e}_1, \mathbf{e}_2$ の線形結合でかけ、また $\mathbf{e}_1, \mathbf{e}_2$ は $\mathbf{u}_1, \mathbf{u}_2$ の線形結合で書ける。実際、 $\mathbf{e}_1 = \mathbf{u}_1 - \mathbf{u}_2$, $\mathbf{e}_2 = \mathbf{u}_2$ である。このことは、 \mathbf{v} が $\mathbf{u}_1, \mathbf{u}_2$ の線形結合で書けることを意味する (命題 18.1.3)。すなわち $\mathbf{u}_1, \mathbf{u}_2$ は \mathbb{R}^2 を生成し、したがって \mathbb{R}^2 の基底である。

いまの例の一般化として定理 18.2.6 が得られる。つぎの補題は定理 18.2.6 の証明において必須というわけではないが、今後の抽象的議論において何度か用いる。

補題 18.2.5. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ が線形独立であるとし、さらに $\mathbf{v} \in V$ とすれば次が成り立つ.

(1) \mathbf{v} が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で表せないならば、ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}$ は線形独立である.

(2) ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}$ は線形従属ならば、 \mathbf{v} は $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で表せる.

Proof. (2) は (1) の対偶ゆえ (1) のみ示せばよい. 線形独立性を示すために線形関係 $\sum_{i=1}^n a_i \mathbf{u}_i + a_{n+1} \mathbf{v} = \mathbf{0}$ を仮定しよう. このとき $a_{n+1} = 0$ でなければならない. 何故なら、もし $a_{n+1} \neq 0$ ならば $\mathbf{v} = \sum_{i=1}^n \frac{-a_i}{a_{n+1}} \mathbf{u}_i$ となり、これは \mathbf{v} が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で表せないことに反する. ゆえに $a_{n+1} = 0$ であり、 $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}$ を得る. $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形独立性より $a_1 = \dots = a_n = 0$ である. \square

定理 18.2.6. \mathbb{R}^n における n 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ において、これらが線形独立であることと \mathbb{R}^n の基底であることは同値である.

Proof. 列ベクトルに対して示そう. 行ベクトルの場合は転置をとって列ベクトルの場合に帰着させればよい. 列ベクトル $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立であるとし、これが \mathbb{R}^n を生成することを背理法により示す. 仮に $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けないベクトル $\mathbf{v} \in \mathbb{R}^n$ があるとすれば、 $n+1$ 個の組 $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v} \in \mathbb{R}^n$ は補題 18.2.5(1) より線形独立となる. しかしこのことは命題 17.3.5 に矛盾する. ゆえに、各 $\mathbf{v} \in \mathbb{R}^n$ は $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けねばならない. \square

練習 18.2.7. 補題 18.2.5 を用いない定理 18.2.6 の別証明を与えよ.

解答例: 各 $\mathbf{b} \in \mathbb{R}^n$ が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で表せることを示そう. $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形独立ゆえ、定理 17.3.6 より正方形行列 $A = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ は可逆である. したがって連立 1 次方程式 $A\mathbf{x} = \mathbf{b}$ は唯一解 $\mathbf{x} = A^{-1}\mathbf{b}$ を持つ. $A^{-1}\mathbf{b} = {}^t[x_1, \dots, x_n]$ と成分表示すれば、 $x_1\mathbf{u}_1 + \dots + x_n\mathbf{u}_n = \mathbf{b}$ である.

座標軸の取り換えを通して関数 $y = \frac{1}{x}$ のグラフが双曲線の一種であることを確認しよう. $y = \frac{1}{x}$ のグラフ上の点は $xy = 1$ を満たす点 (x, y) と言い換えられる. ここで, 2変数関数 xy に着目する. この関数は, 変数変換により $\alpha X^2 + \beta Y^2$ なる形に変形できる. その詳細は対称行列の対角化を用いた2次形式の分類を通して理解されるのであるが (37.5節), ここでは結論だけを述べると

$$\begin{bmatrix} x \\ y \end{bmatrix} = A \begin{bmatrix} X \\ Y \end{bmatrix}, \quad \text{ただし } A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (18.2.1)$$

とすればよい. 実際, $x = \frac{1}{\sqrt{2}}X - \frac{1}{\sqrt{2}}Y, y = \frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Y$ ゆえ

$$xy = \left(\frac{1}{\sqrt{2}}X - \frac{1}{\sqrt{2}}Y \right) \left(\frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Y \right) = \frac{1}{2}X^2 - \frac{1}{2}Y^2.$$

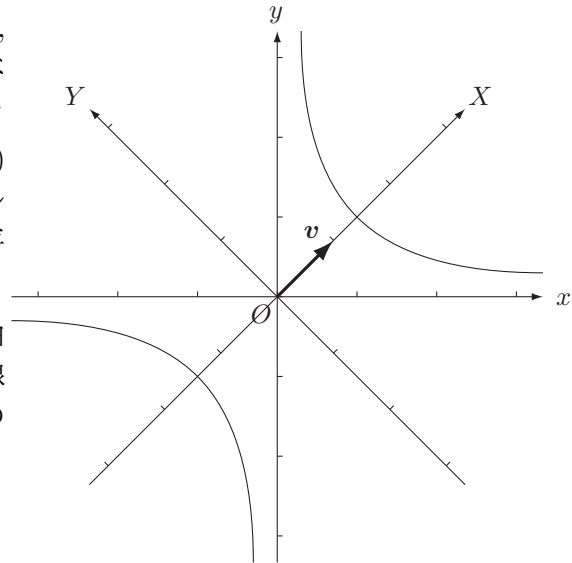
したがって, 式 (18.2.1) の関係の下で $xy = 1$ と $\frac{1}{2}X^2 - \frac{1}{2}Y^2 = 1$ は同値であり, この条件を満たす点の集合は (X, Y) -平面において双曲線を描く.

(X, Y) -平面上の各点が (x, y) -平面においてどの位置に対応するか検討しよう. これは $(X, Y) = (a, b)$ のとき, (x, y) を a, b を用いて表せばどうなるかという問いである. その答えは, A が $\frac{\pi}{4}$ 回転を表す行列であることから, 原点を中心に点 (a, b) を $\frac{\pi}{4}$ 回転させた位置を表す座標に等しい. 例えば, $(X, Y) = (1, 0), (0, 1)$ を式 (18.2.1) に代入すれば,

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ のとき, } \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ のとき, } \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

したがって, (X, Y) -平面における標準ベクトルは, (x, y) -平面における標準ベクトルを $\frac{\pi}{4}$ 回転させたベクトル $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ となる. 右図のベクトル v は (X, Y) -平面における標準ベクトル $(1, 0)$ であり, これは (x, y) -平面における標準ベクトル $(1, 0)$ を $\frac{\pi}{4}$ 回転させたベクトルに等しい. (x, y) -平面における v の座標は $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ である.

以上の考察により, 関数 $y = \frac{1}{x}$ のグラフは双曲線 $\frac{1}{2}X^2 - \frac{1}{2}Y^2 = 1$ に一致し, この曲線は, 双曲線 $\frac{1}{2}x^2 - \frac{1}{2}y^2 = 1$ を原点を中心に $\frac{\pi}{4}$ 回転させたものに等しいことが分かった.



18.3 基底の探し方

有限個のベクトルの組で生成される線形空間における基底の探し方を検討しよう. 次の命題の証明では抽象的な基底の構成法が述べられている.

命題 18.3.1. V が零でないベクトルの組 u_1, \dots, u_m によって生成されているとすれば, u_1, \dots, u_m の中からいくつかを取りだし V の基底とすることができる. とくに, 有限個のベクトルの組で生成される線形空間は基底を持つ.

Proof. V を生成するまで、線形独立性が満たされるよう元を一つずつ加えていけばよい。これは次のような手続きによってなされる。まず \mathbf{u}_{n_1} として \mathbf{u}_1 を取る。この \mathbf{u}_{n_1} が V を生成するならば、 \mathbf{u}_{n_1} は V の基底である。そうでない場合は $\mathbf{u}_2, \dots, \mathbf{u}_m$ のうち \mathbf{u}_{n_1} のスカラー倍で表せないものがある。何故なら、もし $\mathbf{u}_1, \dots, \mathbf{u}_m \in \langle \mathbf{u}_{n_1} \rangle$ とすれば命題 18.1.3 より $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_m \rangle \subset \langle \mathbf{u}_{n_1} \rangle$ であり、 V は \mathbf{u}_{n_1} によって生成されてしまう。そこで、 \mathbf{u}_{n_1} のスカラー倍で表せないベクトルを仮に \mathbf{u}_{n_2} とすれば補題 18.2.5(1) より $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ は線形独立である。 $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ が V を生成するならばこれは V の基底となる。そうでない場合は $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ を除いたベクトルのうちいずれかは $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ の線形結合で書けない。何故なら、もし $\mathbf{u}_1, \dots, \mathbf{u}_m$ すべてが $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ の線形結合で書けるとすると、命題 18.1.3 より $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_m \rangle \subset \langle \mathbf{u}_{n_1}, \mathbf{u}_{n_2} \rangle$ であり、 V は $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ によって生成されてしまう。ゆえに $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ の線形結合で表せないベクトルがあり、これを \mathbf{u}_{n_3} とする。この作業を順次繰り返していくと、いずれ線形独立な組 $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}, \mathbf{u}_{n_3}, \dots, \mathbf{u}_{n_k}$ (ただし $k \leq m$) が V を生成することになる。実際、線形結合で書けない元を新たに付け加える操作は、最大でも $\mathbf{u}_1, \dots, \mathbf{u}_m$ をすべて取りつくすことになる m 回までしか行えない。以上の手続きにより、 V の基底 $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}, \mathbf{u}_{n_3}, \dots, \mathbf{u}_{n_k}$ が得られる。□

命題 18.3.1 の証明における手順を改善すれば次のような基底の構成もできる。これは、基底の一部としたいベクトルがあらかじめ決まっているときに有効な手段となる。また、例題 17.3.3 の【注意】(2)とも関連する話題である。

命題 18.3.2. 有限個のベクトルの組で生成される線形空間 V において、線形独立な組 $\mathbf{u}_1, \dots, \mathbf{u}_i \in V$ が与えられているとき、これらに新たなベクトルを付け加えて V の基底とすることができる。

Proof. V は有限個のベクトルで生成されていることから、あるベクトル $\mathbf{u}_{i+1}, \dots, \mathbf{u}_{i+j} \in V$ を用いて $V = \langle \mathbf{u}_{i+1}, \dots, \mathbf{u}_{i+j} \rangle$ と表せる。このとき、

$$V = \langle \mathbf{u}_1, \dots, \mathbf{u}_i, \mathbf{u}_{i+1}, \dots, \mathbf{u}_{i+j} \rangle$$

でもあることに注意して、 $m = i + j$ 個の組 $\mathbf{u}_1, \dots, \mathbf{u}_m$ に対して命題 18.3.1 の証明を適用しよう。すると、 $\mathbf{u}_1, \dots, \mathbf{u}_i$ の線形独立性から、 $n_1 = 1, n_2 = 2, \dots, n_i = i$ となり、命題 18.3.1 の証明で与えた基底 $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}, \mathbf{u}_{n_3}, \dots, \mathbf{u}_{n_k}$ は $\mathbf{u}_1, \dots, \mathbf{u}_i$ をすべて含むベクトルの組となる。□

これまでに挙げてきた \mathbb{R}^n の部分空間については、行列の簡約化の理論を通して基底を見つけることができる。これを次の例題を通して説明しよう。

例題 18.3.3. 次で定める $A = [\mathbf{a}_1, \dots, \mathbf{a}_5]$ について (例題 17.3.3 と同じもの)、次の問いに答えよ。

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{a}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{a}_5 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ -3 \\ -1 \end{bmatrix}.$$

(1) $\langle \mathbf{a}_1, \dots, \mathbf{a}_5 \rangle$ の基底を求めよ。

解答例: $A = [\mathbf{a}_1, \dots, \mathbf{a}_5]$ の簡約化を $B = [\mathbf{b}_1, \dots, \mathbf{b}_5]$ とする (簡約化は例題 17.3.3 で行った)。このとき $\langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4 \rangle = \langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4, \mathbf{b}_5 \rangle$ である。よって命題 17.3.2 より $\langle \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4 \rangle = \langle \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5 \rangle$ となる。また、組 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ は \mathbb{R}^m の標準ベクトルゆえ線形独立である。ゆえに命題 17.3.2 より組 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4$ も線形独立であり、これらは $\langle \mathbf{a}_1, \dots, \mathbf{a}_5 \rangle$ の基底となる。

(2) 方程式 $A\mathbf{x} = \mathbf{0}$ の解空間 W_A の基底を求めよ。

解答例: $[A|\mathbf{0}]$ の簡約化は $[B|\mathbf{0}]$ であり, W_A の外延的表示を得るために次の方程式 $B\mathbf{x} = \mathbf{0}$ を解く:

$$\begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

掃き出し法により W_A は次のように表される:

$$W_A = \{ c_1\mathbf{u}_1 + c_2\mathbf{u}_2 \mid c_1, c_2 \in \mathbb{R} \}. \quad (\text{ただし, } \mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \\ 1 \end{bmatrix})$$

ゆえに $W_A = \langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ である. $\mathbf{u}_1, \mathbf{u}_2$ の線形独立性は, 主成分のある列に対応する行成分の情報を落とすことで理解できる. いまの例では主成分のある列 1, 2, 4 に対応する $\mathbf{u}_1, \mathbf{u}_2$ の 1, 2, 4 行を目隠しして

$$\mathbf{u}_1 = \begin{bmatrix} * \\ * \\ 1 \\ * \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} * \\ * \\ 0 \\ * \\ 1 \end{bmatrix}$$

と見ると, これらが線形独立であることは標準ベクトルがそうであることと同程度に明らかであろう. 以上より $\mathbf{u}_1, \mathbf{u}_2$ は W_A の基底である.

上の例題で行った議論を一般的に述べると次の命題になる. 特に (2) の証明は, 掃き出し法による連立 1 次方程式の解法から W_A の基底が得られることを述べている.

命題 18.3.4. (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ について次が成り立つ.

(1) \mathbb{R}^m の部分空間 $\langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ は $\text{rank } A$ 個のベクトルからなる基底を持つ.

(2) $A\mathbf{x} = \mathbf{0}$ の解空間 $W_A \subset \mathbb{R}^n$ は $n - \text{rank } A$ 個のベクトルからなる基底を持つ.

Proof. $k = \text{rank } A$ とおき, A の簡約化を $B = [\mathbf{b}_1, \dots, \mathbf{b}_n]$, B の各列のうち主成分を持つ列を $\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_k}$ とし, 主成分を持たない列を $\mathbf{b}_{r_1}, \dots, \mathbf{b}_{r_{n-k}}$ とする.

(1): ベクトルの組 $\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_k}$ は互いに異なる標準ベクトルからなるゆえ線形独立である. また, 簡約化の形から, B の各列は $\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_k}$ の線形結合で書ける. 命題 17.3.2 より組 $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k}$ は線形独立であり, A の各列は $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k}$ の線形結合で書ける. すなわち, $\mathbf{a}_1, \dots, \mathbf{a}_n \in \langle \mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k} \rangle$. これと命題 18.1.3 を合わせて $\langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle \subset \langle \mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k} \rangle$ を得る. 以上より $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k}$ は $\langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ の基底となる.

(2): 連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ を掃き出し法によって求めると, 任意定数の個数は B における主成分のない列の数 $n - k$ であるから, その一般解は $\mathbf{x} = \sum_{j=1}^{n-k} c_j \mathbf{u}_j$ と書ける. つまり $W_A = \langle \mathbf{u}_1, \dots, \mathbf{u}_{n-k} \rangle$ である. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_{n-k}$ が線形独立であることを示すために, 各 $\mathbf{u}_1, \dots, \mathbf{u}_{n-k}$ の r_j 成分 ($j = 1, \dots, n - k$) に注目しよう. B の第 r_j 列 \mathbf{b}_{r_j} は主成分を含まない列であったことから, 掃き出し法で求めた一般解において r_j 成分は任意定数としていた. このことは, \mathbf{u}_j の r_j 成分は 1 であり, $\mathbf{u}_1, \dots, \mathbf{u}_{n-k}$ のうち \mathbf{u}_j を除いた残りのベクトルの r_j 成分は 0 になっていることを意味する. (上の例題では $\mathbf{b}_{r_1} = \mathbf{b}_3$, $\mathbf{b}_{r_2} = \mathbf{b}_5$ となる. 確かに \mathbf{u}_1 の $r_1 = 3$ 成分は 1, $r_2 = 5$ 成分は 0 であり, \mathbf{u}_2 の $r_1 = 3$ 成分は 0, $r_2 = 5$ 成分は 1 となっている). ゆえに線形関係 $\sum_{j=1}^{n-k} c_j \mathbf{u}_j = \mathbf{0}$ を与えると, 各 $j = 1, \dots, n - k$ について左辺の第 r_j 成分は c_j となる. これが右辺の r_j 成分である 0 に等しいことから $c_j = 0$ ($j = 1, \dots, n - k$) を得る. すなわち, $\mathbf{u}_1, \dots, \mathbf{u}_{n-k}$ における線形関係は自明なものに限り, これらは線形独立である. \square

本節や 17.3.1 節では、ユークリッド空間 \mathbb{R}^n の部分空間における線形関係について例題を通して学んだ。一般の線形空間 V におけるベクトルの組の独立性の判定や部分空間 $W \subset V$ の基底選びは、 V に関する命題を \mathbb{R}^n に関する命題に翻訳したうえで行うことになる。実は、独立性の判定や基底選びに限らず V の分析は \mathbb{R}^n の分析を通してなされる。この翻訳の基本理念は、線形空間 V の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する条件を \mathbb{R}^n の標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ の条件で述べることにある。また、翻訳作業においては、 V のどの元が \mathbb{R}^n のどの元に対応しているのかを表示するために写像概念が用いられる。そこで次章からしばらくの間、写像に関する概念を整理することにしよう。

18.4 一般の基底 (発展)

有限個のベクトルでは生成されない線形空間もある。そのような空間における基底概念について少しだけ補足しておこう。定義 18.4.2 で与える記号 $\langle A \rangle$ は本書の以降で何度か用いる。

例 18.4.1. 多項式環 $\mathbb{R}[x]$ において、有限個の多項式の組 f_1, \dots, f_n が $\mathbb{R}[x]$ を生成することはない。なぜなら、各多項式 f_1, \dots, f_n の中で最も高い次数を m とすれば、 $m+1$ 次多項式を f_1, \dots, f_n の線形結合で表すことはできないからである。

ベクトルの組による生成や基底概念は、次のようにして無限集合の場合にも拡張される。

定義 18.4.2. 線形空間 V の部分集合 $A \subset V$ において、 A の元による線形結合で書けるベクトルをすべて集めた集合を $\langle A \rangle$ と書く。 $\langle A \rangle = V$ となるとき、 A は V を生成するという。また、 A が線形独立であるとは、 A の中から有限個取りだした相異なるベクトルの組が必ず線形独立になることをいう。さらに、 A が線形独立かつ V を生成するとき、これを V の基底あるいはハメル基底 (Hamel basis) という。

命題 18.4.3. 線形空間 V および空でない部分集合 $A \subset V$ に対して、 $\langle A \rangle$ は V の部分空間である。

Proof. A は空集合でないゆえ $\mathbf{a} \in A$ が取れる。このとき $\mathbf{0}_V = 0\mathbf{a}$ は A の元の線形結合で書けている。ゆえに $\mathbf{0}_V \in \langle A \rangle$ 。一方、 $\mathbf{x}, \mathbf{y} \in \langle A \rangle$ とすれば、 \mathbf{x}, \mathbf{y} は A の元の線形結合で書ける。ここで、 \mathbf{x}, \mathbf{y} それぞれの線形結合に現れる A の元の組は異なるかもしれないが、一部の係数を 0 とすることで共通の組 $\mathbf{a}_1, \dots, \mathbf{a}_n$ を用いて $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i$, $\mathbf{y} = \sum_{i=1}^n y_i \mathbf{a}_i$ と書くことができる。このとき、各 $r, s \in \mathbb{R}$ について $r\mathbf{x} + s\mathbf{y} = \sum_{i=1}^n (rx_i + sy_i) \mathbf{a}_i$ であるから $r\mathbf{x} + s\mathbf{y}$ は A の元の線形結合で書ける。ゆえに $r\mathbf{x} + s\mathbf{y} \in \langle A \rangle$ 。以上より $\langle A \rangle$ は部分空間である。 \square

ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および集合 $A = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ について、次が成り立つ：

- $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle = \langle A \rangle$.
- ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が V を生成することと A が V を生成することは同値である。
- 相異なるベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立であることと A が線形独立であることは同値である。
 【補足】 ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に一つでも重複があれば、これらは線形独立ではない。しかし、この場合において A が線形独立になることがある。例えば、 $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \in \mathbb{R}^2$ を $\mathbf{u}_1 := \mathbf{e}_1, \mathbf{u}_2 := \mathbf{e}_2, \mathbf{u}_3 := \mathbf{e}_2$ と定めれば、ベクトルの組 $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ は線形独立ではない。しかしながら、 $A = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2\} = \{\mathbf{e}_1, \mathbf{e}_2\}$ であり、2点集合 A は線形独立である。
- 相異なるベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が V の基底であることと A が V の基底であることは同値である。

例 18.4.4. 多項式環 $\mathbb{R}[x]$ の部分集合 $A = \{x^n \mid n \in \mathbb{Z}_{\geq 0}\}$ は線形独立である。ここで、集合 $\mathbb{Z}_{\geq 0}$ は 0 以上の整数全体を表し、 $x^0 = \mathbf{1}$ (定数関数) とする。また、 $\mathbb{R}[x]$ の各元は A の元の線形結合で書ける。ゆえに A は $\mathbb{R}[x]$ の基底である。

いま、線形空間の任意の部分集合について、それが生成する部分空間を定めた。とくに部分集合として部分空間を取れば次を得る。

命題 18.4.5. (1) $W \subset V$ が V の部分空間であるとき, $W = \langle W \rangle$.

(2) $A \subset V$ が V の部分集合であるとき, $\langle \langle A \rangle \rangle = \langle A \rangle$.

Proof. (1): $W \subset \langle W \rangle$ は明らかゆえ $\langle W \rangle \subset W$ を示す. 各 $v \in \langle W \rangle$ は W の元による線形結合で書ける. また, W は線形結合で閉じている (命題 16.1.5) ゆえ $v \in W$.

(2): $W := \langle A \rangle$ について (1) を適用すればよい. □

$\mathbb{R}^{\mathbb{N}}$ のように基底を書き下すことが難しい空間もある.

例 18.4.6. 数列空間 $\mathbb{R}^{\mathbb{N}}$ において, 第 n 項が 1 でそれ以外の項がすべて 0 となる数列を e_n と書けば, $A = \{e_n \mid n \in \mathbb{N}\}$ は線形独立である. しかし, A は $\mathbb{R}^{\mathbb{N}}$ を生成しない. 実際, すべての項が 1 となる数列 $x = (1, 1, \dots)$ を A の元から有限個を取りだした組による線形結合で書くことはできない.

そもそも線形空間に必ず基底が存在するかどうかということ自体が明らかではない. ツォルンの補題¹と呼ばれる集合論における原理を適用することで, 次の定理が証明できることが知られている. 詳しい証明は集合論の入門的参考書を参照せよ.

定理 18.4.7. いかなるベクトル空間 V も基底 $A \subset V$ を持つ. 更に, あらかじめ線形独立な部分集合 $B \subset V$ が与えられている場合は, $B \subset A$ を満たすように基底 A を取ることができる.

上の定理は基底の存在を超越的に示すものであり, 基底の形が明示的に書けることは意味しない. 基底の表示が与えられなければ, 本書全体を通しての主題でもある基底を用いた分析は行えない. このような線形空間を調べる際は, 線形空間に位相²と呼ばれる構造を導入し, 極限操作を手掛かりに分析することになる.

¹多くの理工系学部の数学科では, 2 年次の集合論の講義で学ぶことになっている.

²点列の収束発散や写像の連続性を議論できるようにするための枠組み (数学的構造) を位相 (topology) という.

部分集合 $A \subset V$ が生成する部分空間 $\langle A \rangle$ を次のように定義する流儀もある.

定義 18.4.8. 線形空間 V の部分集合 $A \subset V$ に対して, A を含む最小の V の部分空間を $\langle A \rangle$ とする.

ここでいう最小とは, 包含関係 \subset に関して最も小さいということである. 上の定義の利点は少ない言葉で済むこと, そして空集合 \emptyset は自明な部分空間 $\{0\}$ を生成することになり, \emptyset を $\{0\}$ の基底であると約束する手間が省けることにある. 一方, 欠点は, A を含む最小の部分空間はそもそも存在するかという疑問にあらかじめ答えておかねばならないことである. これは集合の共通部分をとる演算 \cap を用いて正当化される. この点について解説しよう.

V の部分集合たちを集めた集合 \mathcal{W} が与えられているとする (すなわち \mathcal{W} は集合の集合であり, このような集合は集合族と呼ばれる). このとき, 各 $W \in \mathcal{W}$ のいずれにも含まれている元をすべて集めた V の部分集合を $\bigcap \mathcal{W}$ と書く. すなわち,

$$\bigcap \mathcal{W} := \{v \in V \mid \text{各 } W \in \mathcal{W} \text{ について } v \in W\}$$

$$(\text{つまり, } v \in \bigcap \mathcal{W} \iff \text{各 } W \in \mathcal{W} \text{ について } v \in W).$$

ここで, \mathcal{W} は無数に多くの集合たちを元として含む無限集合でもよい. このとき, $\bigcap \mathcal{W}$ はそれら無限個の集合たちの共通部分に相当する集合である.

さて, A を含む V の部分空間たち全体からなる集合族を \mathcal{W} としよう. \mathcal{W} は空集合ではない. 故なら, V 自身は A を含む V の部分空間であるから $V \in \mathcal{W}$ である. このとき, $U := \bigcap \mathcal{W}$ と定めれば, U は A を含む最小の部分空間である.

Proof. 示すべきことは (1) $A \subset U$, および (2) U が部分空間であること, (3) U が A を含む部分空間の中で最小であることの三つである.

(1): $A \subset U$ を示すために任意に $a \in A$ を取る. $a \in U$ をいうには, 各 $W \in \mathcal{W}$ について $a \in W$ を示せばよい. 各 $W \in \mathcal{W}$ について W は A を含む部分空間 (つまり $A \subset W$) であった. $a \in A$ および $A \subset W$ ゆえ $a \in W$ である. 以上より「 $a \in A \implies a \in U$ 」が示された. つまり $A \subset U$.

(2): 各 $W \in \mathcal{W}$ は 0_V を含むゆえ $0_V \in U$ である. 次に $x, y \in U$ とすれば $rx + sy \in U$ となることを示そう. そのためには各 $W \in \mathcal{W}$ について $rx + sy \in W$ を示せばよい. $x, y \in U$ より各 $W \in \mathcal{W}$ において $x, y \in W$ であり, W が部分空間であることから $rx + sy \in W$ を得る. したがって $rx + sy \in U$. 部分空間となるための条件 (i) と (iv) が示されたゆえ, U は V の部分空間である.

(3): A を含む任意の V の部分空間 H について, $U \subset H$ となることを示せばよい. A を含む V の部分空間 H を勝手に取れば $H \in \mathcal{W}$ である. $U = \bigcap \mathcal{W}$ の定義により, 各 $W \in \mathcal{W}$ について $U \subset W$ であるから, とくに $U \subset H$. □

この U が定義 18.4.2 における $\langle A \rangle$ と一致することは次のように示される.

Proof. ($\langle A \rangle \subset U$): $A \subset U$ より $\langle A \rangle \subset \langle U \rangle = U$ (命題 18.4.5(1)).

($U \subset \langle A \rangle$): $\langle A \rangle$ は A を含む V の部分空間である (つまり $\langle A \rangle \in \mathcal{W}$). U の最小性より $U \subset \langle A \rangle$. □

第IV部

線形写像

第19章 写像概念の基礎

写像に関する概念のいくつかを述べる。本章では概念をひたすら提示することに終始するゆえ、読者はやや退屈を感じるかもしれない。そこで予告の意味を込めて、これらの概念が線形写像の性質とどう結び付くかを各節末で述べた。

A を (m, n) -行列 A とする。次で定められる写像 T_A は本書全体を通して何度も論じられる:

$$T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad T_A(\mathbf{x}) := A\mathbf{x}.$$

19.1 像と逆像

集合 X から集合 Y への写像 $f: X \rightarrow Y$ における X のことを定義域 (domain) と呼ぶのであった。定義域は始域 (source) と呼ばれる。また、この f における Y のことを終域 (target) と呼ぶ。定義域と始域の概念を更に細かく区別して用いる文献もある。他方で、定義域に対応する語句として値域 (range) を使う文献もあるが、これを終域の意味で使うのであれば高校数学における値域とは意味が異なる。こうした混乱を避けるため、本書では始域および値域という呼称を控えよう。高校数学において f の値域と呼んでいた概念を、本書では像と呼ぶ:

定義 19.1.1. 写像 $f: X \rightarrow Y$ および定義域の部分集合 $A \subset X$ に対して、 A の元を f に代入した値をすべて集めた Y の部分集合を f による A の像 (image) と呼び、これを $f(A)$ と書く。すなわち、外延的記述をすれば

$$f(A) := \{f(x) \mid x \in A\}.$$

とくに $f(X)$ のこと、つまり $f(x)$ の動く範囲を単に f の像と呼ぶ。

練習 19.1.2. 次で与えられる関数 f の像を求めよ。

(1) $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$. 解答: $f(\mathbb{R}) = [0, \infty)$.

(2) $f: [0, 2] \rightarrow \mathbb{R}, f(x) = x^2$. 解答: $f([0, 2]) = [0, 4]$.

(3) $f: [-2, 2] \rightarrow [0, 4], f(x) = x^2$. 解答: $f([-2, 2]) = [0, 4]$.

(4) $f: [0, 2] \rightarrow [0, 4], f(x) = x^2$. 解答: $f([0, 2]) = [0, 4]$.

定義 19.1.3. 写像 $f: X \rightarrow Y$ および終域の部分集合 $B \subset Y$ に対して、 f に代入すると B の元になるような元をすべて集めた X の部分集合を f による B の逆像 (inverse image) と呼び、これを $f^{-1}(B)$ と表す。すなわち、内包的記述をすれば

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\}.$$

また、点 $b \in Y$ に対して、一点集合 $\{b\}$ の逆像 $f^{-1}(\{b\})$ のことを中括弧を略して $f^{-1}(b)$ と書く。

一点の逆像の記号 $f^{-1}(b)$ は、 f^{-1} なる写像に b を代入した値のことではない。 $f^{-1}(b)$ は X の元ではなく、 X の部分集合である。また、一点集合になるとは限らず、複数の点を含むこともあれば空集合になる場合もある。

例 19.1.4. 次で定められる関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ および $b \in \mathbb{R}$ について $f^{-1}(b)$ を求めよ。

(1) $f(x) = x^2$, $b = 2, -2$. 解答: $f^{-1}(2) = \{\sqrt{2}, -\sqrt{2}\}$, $f^{-1}(-2) = \emptyset$.

(2) $f(x) = \sin x$, $b = 0$. 解答: $f^{-1}(0) = \{n\pi \mid n \in \mathbb{Z}\}$.

(3) $f(x) = 2^x$, $b = 1$. 解答: $f^{-1}(1) = \{0\}$.

写像 $f, g: X \rightarrow Y$ が等しい ($f = g$) とは、定義域のいかなる元を代入しても一致すること、すなわち

$$\text{各 } x \in X \text{ について } f(x) = g(x) \quad (19.1.1)$$

が成り立つことに他ならない. 上の条件 (19.1.1) が成立するとき、恒等的に f と g は等しいと言い、これを記号で $f(x) \equiv g(x)$ と書く. なお、誤解の恐れがない多くの場合において、 $f(x) \equiv g(x)$ のことを $f(x) = g(x)$ と書く. また、 f が X の各元をある一点 $b \in Y$ に対応させる定置写像 (定数関数) であるとき (つまり $f(x) \equiv b$ であるとき)、これを $f(x) = b$ あるいは $f = b$ と書くことがある. 後者の表記を認めれば、定数関数 $\mathbf{1}$ (0 次多項式) を $\mathbf{1}$ と書いてよいことになる. ただし、集合の元と写像を等号で結ぶこのような使い方には誤解が生じる恐れもあり、気をつける必要がある. 線形代数の文脈においては、すべての元を零元に対応させる写像 $f: U \rightarrow V$ のことを $f = \mathbf{0}_V$ と書く.

備考 19.1.5. 定値写像 $f: X \rightarrow Y$ ($f = b$) において、 $f(X) = \{b\}$ および $f^{-1}(b) = X$ である.

以降で学ぶこと

写像の像と逆像は、線形代数の枠組みにおいても詳しく調べられる. その理由は線形写像による部分空間の像や逆像が再び部分空間となることにある. 特に、線形写像 f の像には特別な記号が割り当てられ、 $\text{Im } f$ と書かれる. また、 f の原点による逆像 $f^{-1}(\mathbf{0})$ にも特別な記号 $\text{Ker } f$ が用いられる. 行列 A の階数が $\text{Im } T_A$ の次元に一致すること、および同次形連立 1 次方程式 $Ax = \mathbf{0}$ の解空間が $\text{Ker } T_A$ で表されること (例 22.1.3(2)) を通して次元公式 (例 23.2.5) が説明される.

19.2 全射と単射

この項で述べる全単射性 (1 対 1 の対応) なる概念は、既に、行列式の性質の証明の際にも現れていた. そこで用いられていたように、証明の細部における技術面でこの概念が有用であることは理解されよう. しかし、数学のもっと根本的な部分において全単射性の概念は現れる. それは、異なる数学的対象¹ を対応付けて同等とみなす立場を記述する際に用いられる. 例えば、1.3 節では \mathbb{R}^2 上の線形写像全体と 2 次正方行列全体 $M_2(\mathbb{R})$ が同一視できること、したがって線形写像の分析と行列計算の分析が同等であることを見た. 何をもって同等とみなすべきか、それは考えている立場や価値観によって変わってくるだろう. しかしながらいずれにせよ、何かを同等とみなすとき、そこには全単射なる概念が自然に現れることになる.

定義 19.2.1. $f: X \rightarrow Y$ を写像とする. x が重複なく X を動けば $f(x)$ も重複なく Y 上を動くとき、 f を単射 (injection) という. x が X 全体を動けば $f(x)$ も Y 全体を動くとき、 f を全射 (surjection) あるいは上への写像 (onto map) という. 単射性と全射性はそれぞれ次の条件に書き下すことができる:

- 単射性: 各 $x_1, x_2 \in X$ について、 $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$.

対偶をとれば次のようになる: 各 $x_1, x_2 \in X$ について、 $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$.

これは、各 $y \in f(X)$ について $f^{-1}(y)$ がちょうど 1 点からなる集合になることである.

つまり、各 $y \in Y$ について、 $f(x) = y$ をみたす $x \in X$ は高々一つしかない².

¹ここでいう数学的対象とは、図形や空間であったり、あるいは何らかの代数構造を持つ数空間であったりと多岐にわたる. いずれにせよ、それらの多くは集合を用いて記述されるものである.

²高々一つということは、一つもない可能性、つまり 0 個の場合もあり得る.

- **全射性:** 各 $y \in Y$ に対して, f に代入すると y になる元 $x \in X$ が存在する.

像を用いて次のよう表現してもよい: $f(X) = Y$ となること.

逆像を用いた次のような表現もできる: 各 $y \in Y$ について $f^{-1}(y) \neq \emptyset$ が成り立つ.

更に, 全射かつ単射な写像を**全単射 (bijection)** あるいは**1対1の対応**であるという.

射という字を用いることから単射性・全射性を弓矢に例えて説明すれば次のようになる. X を矢の集合, Y を的的位置 (標的) を表す集合とし, f を弓であると考え. いま, 弓 f を一つ固定し, 写像 $f: X \rightarrow Y$ とは, 弓 f を用いて矢 $x \in X$ を放つと $f(x) \in Y$ なる場所に矢が刺さると考える. このとき, 単射とは単発で当たるということである. 各々の的の位置に矢が刺さるとしても, 刺さる矢の数はせいぜい 1 本であり (矢が当たらないこともあり得る), 二本以上の矢が同じ場所に刺さることはない. 言い換えれば, もし矢 a, b がともに同じ位置に刺さった (つまり $f(a) = f(b)$) ならば, その位置に当たる矢の数は 1 本以下であるから, a と b は同一の矢ということになる. 全射とは的の全ての位置に矢が当たること, すなわち, どのような的の位置 $y \in Y$ においても, y に刺さる矢 $x \in X$ があること (つまり $f(x) = y$) を意味する. このとき, y に刺さる矢の数は 1 本以上であれば何本でも構わない.

例 19.2.2. 集合 X において X の元をまったく動かさない写像, すなわち $f(x) := x$ で定める写像 $f: X \rightarrow X$ を**恒等写像 (identity map)** といい, これを id_X と書く. id_X は全単射である.

写像が単射かどうか, あるいは全射かどうかは, f の定義式だけではなく, 定義域や終域に依存して決まるものである:

例 19.2.3. 練習 19.1.2 における写像の単射性および全射性は次のようになる.

- (1) $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) = x^2$ と定めれば, これは単射でも全射でもない.

Proof. $x_1 = 1, x_2 = -1$ と置けば, これらは共に定義域の元であり, $x_1 \neq x_2$ である. ところが $f(x_1) = f(x_2) = 1$ ゆえ, 異なる元が f で同じ元に写されている. ゆえに f は単射ではない. また, f の像は $[0, \infty)$ であり, これは終域 \mathbb{R} に一致しない. ゆえに f は全射ではない. 実際, $y = -1$ は終域の元であるが, $f(x) = y$ を満たす定義域 \mathbb{R} の元 x は存在しない. \square

- (2) $f: [0, 2] \rightarrow \mathbb{R}$ を $f(x) = x^2$ と定めれば, これは単射であり, かつ全射でない.

Proof. 単射性を示すために $a, b \in [0, 2]$ とし, $f(a) = f(b)$ と仮定しよう. このとき $a^2 = b^2$ である. これを移項して因数分解し $(a-b)(a+b) = 0$ を得る. ゆえに $a-b=0$ または $a+b=0$. $a-b=0$ ならば $a=b$ である. $a+b=0$ の場合は, $a=-b$ を得る. このとき, もし $a > 0$ とすれば $b < 0$ となり, これは $b \in [0, 2]$ (とくに $b \geq 0$) であることに矛盾する. ゆえに $a \leq 0$ であり, これと $a \geq 0$ を合わせて $a=0$ を得る. よって $b=-a=0$ であり, $a=b=0$. いずれの場合においても $a=b$ が示され, 以上より f は単射である. f が全射でないことは (1) と同様にして示される. \square

- (3) $f: [-2, 2] \rightarrow [0, 4]$ を $f(x) = x^2$ と定めれば, これは単射ではないが, 全射である.

Proof. f が単射でないことは (1) と同様にして示される. また f の像は $f([-2, 2]) = [0, 4]$ であり, これは終域に一致する. ゆえに f は全射である. \square

- (4) $f: [0, 2] \rightarrow [0, 4]$ を $f(x) = x^2$ と定めれば, これは全単射である.

Proof. 単射性は (2) と同様にして示され, 全射性は (3) と同様にして示される. \square

上の例から, 写像の定義域や終域を変更することに意味があることが分かる. とくに, 定義域を変更する際は, つぎの用語が用いられる.

定義 19.2.4. 写像 $f: X \rightarrow Y$ および部分集合 $A \subset X$ が与えられているとき、 A の各元 $a \in A$ に対して Y の元 $f(a)$ を対応させる写像を f の A における制限 (**restriction**) と呼び、これを $f|_A: A \rightarrow Y$ と書く。また、部分集合を定義域とする写像 $g: A \rightarrow Y$ に対して、新たに定めた $\tilde{g}: X \rightarrow Y$ が $\tilde{g}|_A = g$ を満たすとき、 \tilde{g} は g の拡張 (**extension**) であるという。

練習 19.2.5. 写像 $f: X \rightarrow Y$ について次を示せ。

(1) f が単射ならば、部分集合 $A \subset X$ について $f^{-1}(f(A)) = A$.

解答例: 集合が一致することを示すには、両方の包含関係を確認すればよい。各 $a \in A$ に対して $f(a) \in f(A)$ であるから $a \in f^{-1}(f(A))$ 。つまり $A \subset f^{-1}(f(A))$ である。次に、任意に $\alpha \in f^{-1}(f(A))$ を取ろう。すると $f(\alpha) \in f(A)$ であるから、ある $a_0 \in A$ を用いて $f(a_0) = f(\alpha)$ と書ける。 f の単射性より $\alpha = a_0 \in A$ 。したがって $f^{-1}(f(A)) \subset A$ 。□

(2) f が全射ならば、部分集合 $B \subset Y$ について $f(f^{-1}(B)) = B$.

解答例: $\beta \in f(f^{-1}(B))$ を任意にとれば、ある $x \in f^{-1}(B)$ を用いて $\beta = f(x)$ と書ける。 $x \in f^{-1}(B)$ より $f(x) \in B$ 、つまり $\beta \in B$ である。ゆえに $f(f^{-1}(B)) \subset B$ 。次に $b \in B$ を任意に取れば、 f の全射性より、ある $x' \in X$ を用いて $f(x') = b$ と書ける。 $f(x') \in B$ より $x' \in f^{-1}(B)$ であり、ゆえに $b = f(x') \in f(f^{-1}(B))$ 。したがって $B \subset f(f^{-1}(B))$ 。□

一般の写像においては、上の (1) および (2) が成り立つとは限らない。例えば、 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) := x^2$ と定めれば、 $A = \{2\}$ および $B = [-5, 3]$ について

$$f^{-1}(f(A)) = f^{-1}(f(\{2\})) = f^{-1}(4) = \{2, -2\} \neq A,$$

$$f(f^{-1}(B)) = f(f^{-1}([-5, 3])) = f([0, \sqrt{3}]) = [0, 3] \neq B.$$

以降で学ぶこと

線形写像 T_A の単射性は連立 1 次方程式 $Ax = \mathbf{0}$ が唯一解を持つこと、つまり A の各列ベクトルの線形独立性で特徴づけられる (例 20.3.2 および定理 20.3.4)。全射性は A の列ベクトルで生成される部分空間が終域に一致すること (例 20.2.4)、あるいは A の行の数と階数が一致することによって特徴づけられる (練習 22.4.7 (2))。

19.3 逆写像とその性質

写像 $f: X \rightarrow Y$ が全単射であるとき、次が成り立つ:

(※) 各 $y \in Y$ に応じて、 $f(x_y) = y$ を満たす $x_y \in X$ が唯一つ存在する。

この性質をもとに、 f の逆写像を次のように定める:

定義 19.3.1. 全単射 $f: X \rightarrow Y$ が与えられているとする。このとき、各 $y \in Y$ に対して (※) で与えられる x_y (つまり f に代入すると y になる定義域の元) を対応させる写像 $g: Y \rightarrow X$ ($g(y) = x_y$) を f の逆写像 (**inverse map**) とよび、この g を記号 f^{-1} で表す。

【注意】 f が全単射であり $f(x) = y$ ならば、上の x_y は x に相当し、 $f^{-1}(y) = x$ である。

f の逆写像が定義できるのは、 f が全単射のときに限る。以下、 f の逆写像について論じる際は、 f が全単射であることを暗黙のうちに前提として話を進めていると考えよ。

いま、全単射 $f: X \rightarrow Y$ の逆写像 f^{-1} が与えられているとし、 $x \in X, y = f(x)$ であるとしよう。このとき、記号 $f^{-1}(y)$ には二つの異なる意味が与えられている。一つは逆像のことであり、 $f^{-1}(\{y\})$ を略した表記のことであり、 f は全単射であるから、これは一点からなる X の部分集合 $f^{-1}(y) = \{x\}$ になる。もう一つの意味は、逆写像 $f^{-1}: Y \rightarrow X$ に y を代入した値のことであり、この立場では $f^{-1}(y)$ は X の元 x であり、部分集合ではない。記号 $f^{-1}(y)$ がどちらを意味しているかは文脈で判断しなければならないが、一点集合かあるいは一点集合の元かの違いしかなく、実質的な数学を理解するうえでは支障がないことが多い。これらの違いを厳密に区別する必要が生じるのは集合論においてのみである。一方、部分集合 $B \subset Y$ においても同様に記号 $f^{-1}(B)$ に二つの意味が与えられる。しかし、こちらは結果として同じ集合を表すことになり、どちらの意味で解釈しても構わない (練習 19.3.6)。

例 19.3.2. (1) 写像 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) := ax + b$ (ただし $a \neq 0$) と定めれば f は全単射であり、 $f^{-1}(y) = \frac{1}{a}y - \frac{b}{a}$ である。

(2) 写像 $f: [0, \infty) \rightarrow [0, \infty)$ を $f(x) = x^2$ と定めれば f は全単射であり、 $f^{-1}(y) = \sqrt{y}$ 。

(3) 写像 $f: (-\infty, 0] \rightarrow [0, \infty)$ を $f(x) = x^2$ と定めれば f は全単射であり、 $f^{-1}(y) = -\sqrt{y}$ 。

初等解析学では、関数に代入する変数に文字 x を用いることが多い。この慣習を踏襲すると逆写像に代入する変数も文字 x を用いることになる。例えば $f(x) = ax + b$ の逆写像は $f^{-1}(x) = \frac{1}{a}x - \frac{b}{a}$ と書かれる。この場合、 $f: X \rightarrow Y$ の逆写像 $f^{-1}: Y \rightarrow X$ に代入する x は Y の元である。つまり、集合 Y の元を表す文字に x を用いることになる。変数 x は必ず X の元であると勘違いしてはいけない。

次は定義 19.3.1 の下にある注意をより詳しく述べたものであり、この事実を前提として、下の命題文に現れる条件 (1) と (2) を満たす g のことを f の逆写像と定義する流儀もある。

命題 19.3.3. $f: X \rightarrow Y$ の全単射性と、次の性質 (1) と (2) を共に満たす $g: Y \rightarrow X$ が存在することは同値である。更に、この g は f の逆写像に一致する。

(1) 各 $x \in X$ について、 $g \circ f(x) = x$ が成り立つ (すなわち $g \circ f = \text{id}_X$)。

(2) 各 $y \in f(X)$ について、 $f \circ g(y) = y$ が成り立つ (すなわち $f \circ g = \text{id}_Y$)。

Proof. まず f が全単射であると仮定し、 $g = f^{-1}$ が性質 (1) と (2) を満たすことを示そう。(1) を示すために $x \in X$ を勝手に取り、 $y := f(x) \in Y$ とおく。逆写像の定義によれば、 $g(y)$ とは f に代入すると y になる X の唯一の元、すなわち x のことであり、ゆえに $g(y) = x$ 。つまり、 $g \circ f(x) = g(f(x)) = g(y) = x$ 。次に (2) を示すために $y \in Y$ を勝手取る。逆写像の定義によれば、 $g(y)$ は f に代入すると y になる元である。つまり $f(g(y)) = y$ 、すなわち $f \circ g(y) = y$ 。

次に性質 (1) と (2) を満たす g の存在を仮定し、 f の全単射性を示そう。単射性: $f(x_1) = f(x_2)$ とすれば、これらを g に代入し、 $g(f(x_1)) = g(f(x_2))$ を得る。これと (1) を合わせれば $x_1 = g \circ f(x_1) = g(f(x_1)) = g(f(x_2)) = g \circ f(x_2) = x_2$ 。つまり $x_1 = x_2$ 。全射性: 各 $y \in Y$ に対して、 $x := g(y)$ とおくと、(2) より $y = f \circ g(y) = f(g(y)) = f(x)$ 。

最後に、(1) と (2) を満たす写像 g が f^{-1} に一致することを示そう。 f^{-1} も (1) と (2) を満たすことは既に示している。とくに (2) より各 $y \in Y$ に対して、 $f \circ g(y) = y = f \circ f^{-1}(y)$ 。すなわち $f(g(y)) = f(f^{-1}(y))$ であり、 f の単射性より $g(y) = f^{-1}(y)$ 。□

備考 19.3.4. $f: X \rightarrow X$ が全単射であることの必要十分条件は、 $f \circ g = \text{id}_X = g \circ f$ を満たす $g: X \rightarrow X$ が存在することと同値である。これは、 A が可逆であることの定義 (すなわち $AB = E = BA$ を満たす B が存在すること) の写像の言葉による言い換えに相当する。

練習 19.3.5. 全単射 h について次を示せ。

(i) h^{-1} も全単射である。 (ii) $(h^{-1})^{-1} = h$ 。

解答例: $f = h^{-1}, g = h$ として命題 19.3.3 を適用すればよい。 g は (1) および (2) を満たすゆえ $f = h^{-1}$ は全単射である。また、 g は f の逆写像であるから $g = f^{-1} = (h^{-1})^{-1}$ 。すなわち $h = (h^{-1})^{-1}$ 。□

練習 19.3.6. $f: X \rightarrow Y$ を全単射とし, $B \subset Y$ とする. 逆写像 f^{-1} の B による像 I と, f の B による逆像 P が一致することを示せ (既に述べたように, I, P はいずれも記号 $f^{-1}(B)$ で表される).

解答例: ($I \subset P$): $i \in I$ とすれば, ある $b \in B$ を用いて $i = f^{-1}(b)$ と書ける. 逆写像の定義により $f(i) = b$ である. i を f に代入すると B の元になるゆえ, B による f の逆像 P に i は含まれる.

($P \subset I$): $p \in P$ とすれば, $f(p) \in B$ である. このとき $p = f^{-1}(f(p))$ ゆえ, B による f^{-1} の像 I に p は含まれる. □

以降で学ぶこと

T_A が全単射であることと A が可逆行列になることは同値になる. このとき, その逆写像は $T_{A^{-1}}$ で与えられる (命題 21.3.3). とくに, A が正方行列でなければ T_A は全単射でない. この事実を道具立てを何もせず示すことは意外に難しい:

練習 19.3.7. $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ が全単射ならば (m, n) -行列 A は正方行列であることを示せ.

解答例: $n \leq m$ および $n \geq m$ を別々に示そう.

T_A の単射性より $T_A(\mathbf{x}) = \mathbf{0}$ を満たす $\mathbf{x} \in \mathbb{R}^n$ は $\mathbf{x} = \mathbf{0}$ に限る. すなわち, 方程式 $A\mathbf{x} = \mathbf{0}$ は唯一解 $\mathbf{x} = \mathbf{0}$ を持つ. したがって命題 6.2.1(3) より $\text{rank } A = n$ であり, これと式 (6.1.1) $\text{rank } A \leq m$ を合わせて $n \leq m$ を得る.

各 $\mathbf{e}_i \in \mathbb{R}^m$ ($i = 1, \dots, m$) に対して, T_A の全射性より $A\mathbf{x}_i = \mathbf{e}_i$ を満たす $\mathbf{x}_i \in \mathbb{R}^n$ が存在する. ここで $\mathbf{x}_1, \dots, \mathbf{x}_m$ の線形独立性が示されれば, $m \leq n$ であることが導かれる. 何故なら, 仮に $m > n$ とすれば命題 17.3.5 より $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n$ は線形従属でなければならない. $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n$ の線形独立性については各自で確かめよ (命題 20.1.7(3)). □

19.4 写像の合成

全射性および単射性と写像の合成との関係について補足しておこう.

命題 19.4.1. 写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$ が与えられているとする.

- (1) f, g が共に単射ならば $g \circ f: X \rightarrow Z$ も単射である.
- (2) f, g が共に全射ならば $g \circ f: X \rightarrow Z$ も全射である.
- (3) f, g が共に全単射ならば $g \circ f: X \rightarrow Z$ も全単射である.

Proof. (1): $g \circ f(x_1) = g \circ f(x_2)$ とすれば $g(f(x_1)) = g(f(x_2))$ であり, g の単射性より $f(x_1) = f(x_2)$ である. これに f の単射性を適用し $x_1 = x_2$ を得る. (2): 各 $z \in Z$ に対して, g の全射性より $z = g(y)$ を満たす $y \in Y$ が存在する. また, この y に対して f の全射性を適用すると, $y = f(x)$ を満たす $x \in X$ が存在する. このとき $z = g \circ f(x)$. (3): (1) および (2) より明らか. □

命題 19.3.3 における全射性と単射性の証明で行った議論は次のように一般化できる.

命題 19.4.2 (発展). 写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$ が与えられているとする.

- (1) $g \circ f: X \rightarrow Z$ が単射ならば f も単射である.
- (2) $g \circ f: X \rightarrow Z$ が全射ならば g も全射である.

Proof. (1): $f(x_1) = f(x_2)$ とすれば, これらを g に代入し $g \circ f(x_1) = g \circ f(x_2)$ を得る. これに $g \circ f$ の単射性を適用し $x_1 = x_2$ を得る. すなわち f は単射である.

(2): 各 $z \in Z$ に対して, $g \circ f$ の全射性より $z = g \circ f(x)$ を満たす $x \in X$ が存在する. このとき, $y := f(x)$ とおけば $y \in Y$ であり, $z = g \circ f(x) = g(f(x)) = g(y)$. ゆえに g は全射である. □

練習 19.4.3. 写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$ を全単射とする. $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ を示せ.

解答例: 写像の合成に関する結合律 (命題 3.2.5) から $(f^{-1} \circ g^{-1}) \circ (g \circ f) = \text{id}_X$ が容易に確かめられる. 実際,

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ (g \circ f)) = f^{-1} \circ ((g^{-1} \circ g) \circ f) = f^{-1} \circ (\text{id}_Y \circ f) = f^{-1} \circ f = \text{id}_X.$$

同様にしておよび $(g \circ f) \circ (f^{-1} \circ g^{-1}) = \text{id}_Z$ も示され, ゆえに命題 19.3.3 より $f^{-1} \circ g^{-1}$ は $g \circ f$ の逆写像である. \square

備考 19.4.4. 可逆行列 B, A に対して BA の逆行列は $A^{-1}B^{-1}$ であった. この事実を写像の言葉で言い換えたものが上の練習に他ならない.

逆像についても, 練習 19.4.3 と類似の性質が成り立つ.

命題 19.4.5. 写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$, 部分集合 $C \subset Z$ に対して, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$.

Proof. $x \in (g \circ f)^{-1}(C) \iff g \circ f(x) \in C \iff g(f(x)) \in C \iff f(x) \in g^{-1}(C) \iff x \in f^{-1}(g^{-1}(C))$. \square

上の性質を有限回適用することで, n 個の写像の合成についても同様の性質がなりたつ. 例えば三つの写像の合成については次のようになる:

$$(h \circ g \circ f)^{-1}(D) = (h \circ (g \circ f))^{-1}(D) = (g \circ f)^{-1}(h^{-1}(D)) = f^{-1}(g^{-1}(h^{-1}(D))).$$

以降で学ぶこと

19.2 節末で予告したことと命題 19.4.2 を合わせると直ちに次が導かれる:

命題 19.4.6. (m, n) -行列 A および (n, r) -行列 B について次が成り立つ.

- (1) 積 AB の各列が線形独立ならば B の各列も線形独立である.
- (2) 積 AB の各列が \mathbb{R}^m を生成するならば A の各列も \mathbb{R}^m を生成する.

練習 19.4.7. これまでに学習した知識から上の (1) を示せ.

解答例: 方程式 $Bx = 0$ が唯一解をもつことを示そう. ベクトル x を方程式 $Bx = 0$ の解とすれば, $ABx = A(Bx) = A0 = 0$ ゆえ x は $ABx = 0$ の解でもある. 仮定より AB の各列は線形独立であり, 命題 17.3.1 より $ABx = 0$ の解は唯一である. ゆえに $x = 0$. 以上より, $Bx = 0$ の解は自明なものに限る. 再び命題 17.3.1 より, B の各列は線形独立である. \square

19.5 無限集合 (発展)

定義 8.1.1 をよく読むと, 置換を単射 $f: X_n \rightarrow X_n$ のことと定めている. しかし, 置換は逆写像 (逆置換) を持つゆえ全単射である. この点について補足しておこう.

命題 19.5.1. X, Y をともに n 点からなる集合とする. 写像 $f: X \rightarrow Y$ において次は同値である:

- (1) f は単射である,
- (2) f は全射である.

Proof. $X = \{x_1, \dots, x_n\}, Y = \{y_1, \dots, y_n\}$ とおいて証明しよう.

(1) \Rightarrow (2): f は単射ゆえ $f(x_1), \dots, f(x_n)$ の中に重複はない. もし f が全射でないと仮定すれば, $f(x_1), \dots, f(x_n)$ のいずれでもない $y \in Y$ が存在する. このとき Y は $n+1$ 個の元 $f(x_1), \dots, f(x_n), y$ を含むことになり, これは Y が n 点集合であることに反する. ゆえに f は全射である.

(2) \Rightarrow (1): f を全射とする. もし f が単射でないと仮定すれば, $x_i \neq x_j$ かつ $f(x_i) = f(x_j)$ なる $i \neq j$ が取れる. このとき $f(x_1), \dots, f(x_n)$ のうち $f(x_i)$ と $f(x_j)$ は等しいゆえ, 集合 $f(X) = \{f(x_1), \dots, f(x_n)\}$ の点の数は $n-1$ 以下となる. f の全射性より $Y = f(X)$ であり, つまり Y の元の数は $n-1$ 以下である. これは Y が n 点集合であることに反する. ゆえに f は単射である. \square

無限集合においては上の命題の類似は成り立たない。すなわち、 X から X への写像で、単射だが全射でない例、および全射だが単射でない例がある：

例 19.5.2. (1) $f : \mathbb{N} \rightarrow \mathbb{N}$ を $f(x) := x + 1$ と定めれば、 f は単射であるが全射でない。

(2) $g : \mathbb{N} \rightarrow \mathbb{N}$ を $f(x) := \max\{1, x - 1\}$ と定めれば、 g は全射であるが単射でない。

備考 19.5.3. 命題 19.5.1 および例 19.5.2 で見た有限集合と無限集合における様子の違いと類似する現象が、線形代数においても有限次元空間と無限次元空間における違いとして表れる。すなわち、線形写像 $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ について、 f の単射性と全射性は同値である (命題 22.4.8)。一方で、一般の無限次元線形空間において、このような主張は成立しない (例 26.1.5)。

無限集合においても、「元の総数が一致する」という性質に相当する概念を次のように定める：

定義 19.5.4. 集合 X, Y の間に全単射が存在するとき、 X と Y は対等である (あるいは濃度が等しい) という。

命題 19.5.5. X と Y が対等であり、 Y と Z が対等ならば X と Z も対等である。

Proof. 仮定より二つの全単射 $f : X \rightarrow Y$ および $g : Y \rightarrow Z$ が存在する。このとき $g \circ f : X \rightarrow Z$ は命題 19.4.1(3) より全単射であり、したがって X と Z は対等である。 \square

有限集合の場合と異なり、例 19.5.2 で見たように無限集合は自身の真部分集合³と対等になり得る。このことは少なくともガリレオの時代には既に気づかれていた。

例 19.5.6. 自然数全体 \mathbb{N} と正の偶数全体 $2\mathbb{N} = \{2n \mid n \in \mathbb{N}\}$ は対等である。実際、 $f : \mathbb{N} \rightarrow 2\mathbb{N}$ を $f(x) := 2x$ と定めればこれは全単射である。

すべての無限集合が互いに対等なわけではない。例えば次の事実は、大学1年次の範囲で十分理解できるものである。証明は、例えば巻末の文献 [6] を参照せよ。

命題 19.5.7. (1) \mathbb{N} から \mathbb{R} への全射は存在しない。したがって \mathbb{N} と \mathbb{R} は対等ではない。

(2) \mathbb{Z} や \mathbb{Q} と \mathbb{N} は対等である。

上の事実から、外延的な記法を用いて $\mathbb{R} = \{x_1, x_2, x_3, \dots\}$ と表せないことが分かる。実際、もし仮に $\mathbb{R} = \{x_1, x_2, x_3, \dots\}$ と表せるとすれば、 $f : \mathbb{N} \rightarrow \mathbb{R}$ を $f(n) := x_n$ と定めれば、これは全射であり、上の (1) に反する。

読者になじみのある無限集合について次が成り立つことが知られている。証明は、集合論の入門的な本を参照されたい。

例 19.5.8. (1) 無理数全体 $\mathbb{R} \setminus \mathbb{Q}$ や複素数全体 \mathbb{C} 、ユークリッド空間 \mathbb{R}^n は \mathbb{R} と対等である。

(2) 集合 X の部分集合をすべて集めた集合を X の冪集合とよび、これを $\mathfrak{P}(X)$ あるいは 2^X とかく。 X と $\mathfrak{P}(X)$ は対等ではない。とくに、 $X, \mathfrak{P}(X), \mathfrak{P}(\mathfrak{P}(X)), \mathfrak{P}(\mathfrak{P}(\mathfrak{P}(X))), \dots$ は互いに対等ではない。

無限集合が \mathbb{N} と対等であるとき可算であるといい、そうでないとき非可算であるという。 \mathbb{Z} や \mathbb{Q} は可算であり、 \mathbb{R} や \mathbb{C} は非可算である。

³自分自身以外の部分集合のことを真部分集合 (proper subset) と呼ぶ。

\mathbb{R} と \mathbb{R}^n は対等である。次元の異なる空間の元の個数が等しいことを読者は不思議に感じるかもしれない。しかし、ここでいう全単射 $f: \mathbb{R} \rightarrow \mathbb{R}^n$ の存在性は、代数構造や数列の極限など多くの数学的構造を無視したうえでの 1 対 1 対応があると述べているに過ぎないのである。例えば線形空間としての演算を保つ写像に限れば \mathbb{R} と \mathbb{R}^n ($n \geq 2$) の間の 1 対 1 対応を作ることはできない (命題 21.2.7)。また、数列の発散・収束性を互いに保つ 1 対 1 対応 (これを同相写像という) も \mathbb{R} と \mathbb{R}^n の間には作れないことが知られている (22.1 節のコラムを参照)。このように、二つの対象を同じとみなす (つまり 1 対 1 の対応を与える) といっても、様々な立場があり得る。

一方、何の化粧もない単なる集合に限った場合、無限集合の間の 1 対 1 対応はどこまで理解されているのだろうか。実は、実数の無限部分集合の大きさにどれくらいの種類があるかという基本的な問題ですら容易に理解されるものではなく、これは集合論の創始者であるカントールを生涯悩ませ続けた問題でもあった。すなわち、 \mathbb{N} と \mathbb{R} と対等でない実数の部分集合 $X \subset \mathbb{R}$ は存在するかという問いである。このような X は存在しないという立場を連続体仮説といい、連続体仮説 (あるいはその否定) の証明に彼は長い年月を費やしたが、いずれも証明することはできなかった。現在では、連続体仮説およびその否定のいずれも集合論の公理系からは導けないことが分かっており、これ以上この問題について論じるならば、我々は連続体仮説とその否定のどちらか一方を公理として選択する必要に迫られることになる。

連続体仮説は微積分学とも無縁ではない。逐次積分 (累次積分) の順序の入れ替え:

$$\int_a^b \int_c^d f(x, y) dx dy = \int_c^d \int_a^b f(x, y) dy dx.$$

について、上の等式が成立するための f の条件を解析学では与えている。一方、上の等式が成立しない関数の例が存在するかという問題は授業ではあまり扱わないことが多い。実は、この存在・非存在性も集合論の公理系からは導かれないこと、そして連続体仮説からは等号不成立の例が導けることが知られている。

集合論の公理系から導かれる不思議な事実についても述べておこう。それは、 \mathbb{R}^3 の原点を中心とする半径 1 の球体を有限個の集合に分割し、これらを回転と並行移動によって上手く配置しなおすと半径 2 の球体になるという定理である (バナッハ・タルスキーの逆理)。一見するとこれは体積概念と矛盾するように思えるが、体積が定義できないような複雑な集合に分割させることで、このような構成を実現させている。

以上、無限集合の奥深さを示す例の一端をかけ足ながら取り上げた。

第20章 線形写像

線形代数で扱う線形写像には二つの性格がある。一つは、分析すべき対象であり、そこには諸科学分野において現れる個々の具体的な線形写像をいかに理解するかが念頭にある。これこそが線形代数の主題であるといってもよい。そしてもう一つは、一般のベクトル空間 V の言葉をユークリッド空間の言葉に翻訳するために与える対応 (線形同型) のことである。後者は前者を分析するための道具といえる。本章ではこれら二つの区別をせず、線形写像の定義から共通して得られる一般論を展開し、これらを区別した各論は次章以降に論じる。

ところで、線形写像の多くは行列によって表現され、以降では行列と標準ベクトル e_i の間の積に関する次の性質

$$Ae_i = A \text{ の第 } i \text{ 列目}$$

を断りなく用いる (本章では例 20.2.4 で用いた)。上式を頭の片隅に留めておいてもらいたい。

20.1 線形写像の基本的性質

1 節で述べた通り、線形写像とは比例関数の一般化に相当する概念である。

定義 20.1.1. 線形空間 U から線形空間 V への写像 $f : U \rightarrow V$ が次の性質 (i) および (ii) を満たすとき、 f を線形写像 (linear map) あるいは線形作用素 (linear operator) という：

(i) すべての $x, y \in U$ に対して、 $f(x + y) = f(x) + f(y)$,

(ii) すべての $x \in U$ および $r \in \mathbb{R}$ について、 $f(rx) = rf(x)$.

【補足】 作用素と写像は同義語である。一般に、関数や集合、数列などを代入する写像のことを作用素と呼ぶことが多い。

上の性質 (i), (ii) は線形性と呼ばれる。部分空間になるための条件がまとめられたように、線形性は次の性質 (iii) にまとめられる：

$$\text{線形性 (iii) すべての } x, y \in V \text{ および } a, b \in \mathbb{R} \text{ について、} f(ax + by) = af(x) + bf(y).$$

実際、(iii) において $a = b = 1$ とした場合が (i) であり、 $a = r, b = 0$ とした場合が (ii) である。また、(i) と (ii) を用いて (iii) は次のように導かれる： $f(ax + by) = f(ax) + f(by) = af(x) + bf(y)$ 。以上より、条件「(i) かつ (ii)」と条件 (iii) は同値である。以降、線形性の確認を (iii) によって判定することとしよう。

例 20.1.2. (m, n) -行列 A に対して写像 $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ を $T_A(x) := Ax$ と定めれば、これは線形写像である。実際、線形性 (iii) は次のようにして確かめられる。

$$T_A(ax + by) = A(ax + by) = A(ax) + A(by) = a(Ax) + b(Ay) = aT_A(x) + bT_A(y).$$

上で定めた写像 T_A は今後頻繁に現れるゆえ忘れないこと。なお、ユークリッド空間の間の線形写像は必ず T_A の形で書ける (命題 21.3.4)。とくに写像 $f : \mathbb{R} \rightarrow \mathbb{R}$ のうちで線形写像となるものは、 $(1, 1)$ -行列 $A = [a]$ を用いて表される比例関数 $T_A(x) = ax$ のみである。

例 20.1.3. すべてのベクトルを零ベクトル $0_V \in V$ にうつす定置写像 $f = 0_V : U \rightarrow V$ は線形写像である。実際、 $f(ax + by) = 0_V = a0_V + b0_V = af(x) + bf(y)$ 。このような線形写像は自明な線形写像と呼ばれる。

練習 20.1.4. 次の写像は線形写像ではない. 具体的な元を代入することにより, 線形性 (i), (ii) のいずれも満たされないことを確認せよ.

- (1) 写像 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) = x^2$ と定めれば, これは線形写像ではない.
- (2) 行列式を与える写像 $\det: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ は線形写像でない.

ユークリッド空間以外の線形空間における線形写像の例は本章の最後に述べるとして, しばらくは線形性から導かれる一般論を展開しよう.

命題 20.1.5. 線形写像 $f: U \rightarrow V$ において $\mathbf{0}_U$ の行き先は $\mathbf{0}_V$ である. すなわち $f(\mathbf{0}_U) = \mathbf{0}_V$.

Proof. 各ベクトルを 0 倍すると零ベクトルになることを用いると, $f(\mathbf{0}_U) = f(0 \cdot \mathbf{0}_U) = 0 \cdot f(\mathbf{0}_U) = \mathbf{0}_V$. \square

線形性 (iii) は更に次のような有限和の形に一般化できる¹.

命題 20.1.6. $f: U \rightarrow V$ が線形性を満たすならば次も満たす:

$$\text{線形性 (iii)'} \quad \text{各 } \mathbf{u}_k \in U, r_k \in \mathbb{R} \ (k = 1, \dots, \ell) \text{ について, } f\left(\sum_{k=1}^{\ell} r_k \mathbf{u}_k\right) = \sum_{k=1}^{\ell} r_k f(\mathbf{u}_k).$$

Proof. 和の個数 ℓ に関する帰納法で示す. $\ell = 1$ の場合は線形性の性質 (ii) に他ならない. 和の個数が ℓ のときに等式が成立すると仮定し, 和の個数が $\ell + 1$ の場合について示そう.

$$\begin{aligned} f\left(\sum_{k=1}^{\ell+1} r_k \mathbf{u}_k\right) &= f\left(\left(\sum_{k=1}^{\ell} r_k \mathbf{u}_k\right) + r_{\ell+1} \mathbf{u}_{\ell+1}\right) \\ &= f\left(\sum_{k=1}^{\ell} r_k \mathbf{u}_k\right) + f(r_{\ell+1} \mathbf{u}_{\ell+1}) && \text{(線形性 (i))} \\ &= \sum_{k=1}^{\ell} r_k f(\mathbf{u}_k) + r_{\ell+1} f(\mathbf{u}_{\ell+1}) && \text{(帰納法の仮定と線形性 (ii))} \\ &= \sum_{k=1}^{\ell+1} r_k f(\mathbf{u}_k). \end{aligned}$$

\square

線形性 (iii)' は, 線形写像が線形関係を保存することを述べている:

命題 20.1.7. $f: U \rightarrow V$ を線形写像とし, $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ とする.

- (1) $\mathbf{x} \in U$ が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けるならば, $f(\mathbf{x})$ は $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ の線形結合で書ける.
(すなわち, $\mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \implies f(\mathbf{x}) \in \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle$.)
- (2) $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}_U \implies \sum_{i=1}^n a_i f(\mathbf{u}_i) = \mathbf{0}_V$.
- (3) $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ が線形独立ならば $\mathbf{u}_1, \dots, \mathbf{u}_n$ も線形独立である.
- (4) $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形従属ならば $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ も線形従属である.

Proof. (1): この主張は線形性 (iii)' の言い換えにすぎない. 実際, $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ と書けるならば $f(\mathbf{x}) = f(\sum_{i=1}^n a_i \mathbf{u}_i) = \sum_{i=1}^n a_i f(\mathbf{u}_i)$ であり, $f(\mathbf{x})$ は $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ の線形結合で書ける.

(2): $\mathbf{x} = \mathbf{0}_U$ に対して (1) の証明と同等の計算をすればよい.

(3): 線形関係 $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}_U$ を仮定すれば (2) より $\sum_{i=1}^n a_i f(\mathbf{u}_i) = \mathbf{0}_V$ である. 組 $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ の線形独立性より $a_1 = \dots = a_n = 0$. つまり $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形独立である.

(4): これは (3) の対偶にほかならない. \square

¹命題 12.3.1 と同様の議論を行っている. 行列式を先に扱う都合上, 我々は線形性よりも複雑な多重線形性を先に論じていたのである.

次の例で見るように上の命題における (1) から (4) の逆はいずれも成り立たない。逆が成り立つのは f が単射の場合に限る (命題 20.3.6)。

例 20.1.8. $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ が定める線形写像 $T_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ について次が成り立つ:

- (1) $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ とする。このとき, $f(\mathbf{x}) = \mathbf{e}_1$, $f(\mathbf{e}_1) = \mathbf{e}_1$ である。よって, $f(\mathbf{x})$ は $f(\mathbf{e}_1)$ の線形結合で書ける。しかし \mathbf{x} を \mathbf{e}_1 の線形結合で書くことはできない。
- (2) $0f(\mathbf{e}_1) + 1f(\mathbf{e}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$ である。一方で $0\mathbf{e}_1 + 1\mathbf{e}_2 = \mathbf{e}_2 \neq \mathbf{0}$ である。
- (3) $\mathbf{e}_1, \mathbf{e}_2$ は線形独立であるが, $f(\mathbf{e}_1), f(\mathbf{e}_2)$ は線形独立でない。

【補足】 線形写像 $f: U \rightarrow V$ が単射でない場合, 命題 20.3.6(1) から (4) の逆は常に成り立たない。実際, f が単射でないとするならば, 20.3 節における議論から $f^{-1}(\mathbf{0}_V)$ は 1 点集合ではない。したがって, $f(\mathbf{u}) = \mathbf{0}_V$ を満たす零でないベクトル $\mathbf{u} \in U$ が存在する。このとき, (1) $f(\mathbf{u}) = \mathbf{0}_V = f(\mathbf{0}_U)$ ゆえ $f(\mathbf{u})$ は $f(\mathbf{0}_U)$ の線形結合で書けるが, $\mathbf{u} \neq \mathbf{0}_U$ を $\mathbf{0}_U$ の線形結合で書くことはできない。(2) $1f(\mathbf{u}) + 1f(\mathbf{0}_U) = \mathbf{0}_V$ であるが, $1\mathbf{u} + 1\mathbf{0}_U = \mathbf{u} \neq \mathbf{0}_U$ である。(3) 1 つのベクトルからなる組 \mathbf{u} は線形独立であるが, 組 $f(\mathbf{u})$ は線形独立でない。

以降の内容とは関連しない些細なことになるが, 命題 20.1.7(3) および (4) を集合を用いて述べ直した次の主張は成立しない:

- (a) $f(A)$ が線形独立ならば A も線形独立である。
- (b) A が線形従属ならば $f(A)$ も線形従属である。

上の主張の反例を挙げておこう:

例 20.1.9 (発展). $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ と定めれば, $A = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ は線形従属である。ところが, 例 20.1.8 で与えた $T_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ において, $f(\mathbf{a}_1) = f(\mathbf{a}_2) = f(\mathbf{a}_3) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ より $f(A) = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$ であり, $f(A)$ は線形独立である。

【補足】 自明な線形写像は上の主張 (a) および (b) を満たす。一方で, 自明でない線形写像 $f: U \rightarrow V$ が単射でない場合, (a) および (b) の反例は必ず存在する。これを示すには後に学ぶいくつかの概念が必要となる。ここでは $U = \mathbb{R}^n$ の場合における証明を紹介しよう²。

Proof. f が単射でない場合, 前例の補足で述べたように $f^{-1}(\mathbf{0}_V)$ は 1 点集合ではない。また, $f^{-1}(\mathbf{0}_V)$ は \mathbb{R}^n の線形部分空間となることから (命題 20.3.3), これは無限集合となる。 f は自明でないゆえ零でないベクトル $\mathbf{v} \in f(\mathbb{R}^n)$ が存在し, 系 23.1.6 によれば $f^{-1}(\mathbf{v})$ の元の個数は $f^{-1}(\mathbf{0}_V)$ の元の個数と等しく, したがって $f^{-1}(\mathbf{v})$ も無限集合となる。そこで, 相異なる $n+1$ 個の元 $\mathbf{u}_1, \dots, \mathbf{u}_{n+1} \in f^{-1}(\mathbf{v}) \subset \mathbb{R}^n$ をとり, $A = \{\mathbf{u}_1, \dots, \mathbf{u}_{n+1}\}$ とすれば $f(A) = \{\mathbf{v}\}$ は線形独立である。しかしながら, A は線形従属である (命題 17.3.5)。□

練習 18.1.4 の一般化として次が成り立つ。

命題 20.1.10. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が U を生成するとし, $f, g: U \rightarrow V$ を線形写像とする。

- (1) 各 $i = 1, \dots, n$ について $f(\mathbf{u}_i) = \mathbf{0}_V$ ならば, $f = \mathbf{0}_V$ (つまり f は $\mathbf{0}_V$ に値を取る定置写像)。
- (2) 各 $i = 1, \dots, n$ について $f(\mathbf{u}_i) = g(\mathbf{u}_i)$ ならば, $f = g$ 。

Proof. (1): 各 $\mathbf{x} \in U$ について, 仮定より $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ と書ける。ゆえに $f(\mathbf{x}) = f\left(\sum_{i=1}^n a_i \mathbf{u}_i\right) = \sum_{i=1}^n a_i f(\mathbf{u}_i) = \sum_{i=1}^n a_i \mathbf{0}_V = \mathbf{0}_V$ 。

(2): $f = g$ を導くためには, 各 $\mathbf{x} \in U$ について $f(\mathbf{x}) = g(\mathbf{x})$ を示せばよい。さて, 各 $\mathbf{x} \in U$ に対して $h(\mathbf{x}) := f(\mathbf{x}) - g(\mathbf{x})$ と定めると, $h: U \rightarrow V$ も線形写像となる (各自確かめよ)。仮定より, 各 $i = 1, \dots, n$ について $h(\mathbf{u}_i) = \mathbf{0}_V$ であるから, (1) より $h = \mathbf{0}_V$ 。つまり, 各 $\mathbf{x} \in U$ について $\mathbf{0}_V = h(\mathbf{x}) = f(\mathbf{x}) - g(\mathbf{x})$ 。すなわち $f(\mathbf{x}) = g(\mathbf{x})$ であり, 以上より $f = g$ 。□

²一般の U について考える場合は n として U の次元を取ればよい。

上の (2) より, 特に, 線形写像の値は基底の行き先のみによって決定される. 逆に, 基底の行き先を指定すると, それを満たす線形写像を構成できる:

命題 20.1.11. $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とし, $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ とする ($\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形独立である必要はなく, とくに重複があっても良い). このとき, $f(\mathbf{u}_i) = \mathbf{v}_i$ を満たす線形写像 $f: U \rightarrow V$ が唯一つ存在する.

Proof. U の各元 \mathbf{x} は実数の組 a_1, \dots, a_n を用いて $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ と一意的に書ける (命題 18.2.3). f が線形性を満たすには $f(\mathbf{x}) := \sum_{i=1}^n a_i \mathbf{v}_i$ と定めるしかないが, 実際にこのように定めた f は線形性を満たす. 何故なら, 各 $\mathbf{x}, \mathbf{y} \in U$ を $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$, $\mathbf{y} = \sum_{i=1}^n b_i \mathbf{u}_i$ と書けば, $r\mathbf{x} + s\mathbf{y}$ は $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合として $r\mathbf{x} + s\mathbf{y} = \sum_{i=1}^n (ra_i + sb_i) \mathbf{u}_i$ と一意的に書けており, f の定義から $f(\mathbf{x}) = \sum_{i=1}^n a_i \mathbf{v}_i$, $f(\mathbf{y}) = \sum_{i=1}^n b_i \mathbf{v}_i$, $f(r\mathbf{x} + s\mathbf{y}) = \sum_{i=1}^n (ra_i + sb_i) \mathbf{v}_i$ である. したがって,

$$f(r\mathbf{x} + s\mathbf{y}) = \sum_{i=1}^n (ra_i + sb_i) \mathbf{v}_i = \sum_{i=1}^n ra_i \mathbf{v}_i + \sum_{i=1}^n sb_i \mathbf{v}_i = r \left(\sum_{i=1}^n a_i \mathbf{v}_i \right) + s \left(\sum_{i=1}^n b_i \mathbf{v}_i \right) = rf(\mathbf{x}) + sf(\mathbf{y}).$$

すなわち, 線形性 (iii) を満たす f は線形写像である.

f の一意性を示すには, 線形写像 $g: U \rightarrow V$ も $g(\mathbf{u}_i) = \mathbf{v}_i$ を満たすと仮定し $f = g$ を導けばよい. これは命題 20.1.10(2) より直ちに得られる. \square

【注意】 上の証明において, 各 $\mathbf{x} \in U$ が $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ と一意的に書けるとは限らないとすると f を簡単には定義できなくなる. 例えば, $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i = \sum_{i=1}^n b_i \mathbf{u}_i$ と二通りに書けるとすると $f(\mathbf{x})$ を $\sum_{i=1}^n a_i \mathbf{v}_i$ と $\sum_{i=1}^n b_i \mathbf{v}_i$ のいずれにすべきか, あらかじめ指定する必要があるが生じる. こうなると, 線形性 (iii) が満たされるように定めるのは絶望的である.

命題 20.1.12. 線形写像の合成はまた線形写像となる.

Proof. $f: U \rightarrow V$ および $g: V \rightarrow W$ を線形写像とし, これらの合成 $g \circ f: U \rightarrow W$ が線形性 (iii) を満たすことを示そう. $\mathbf{x}, \mathbf{y} \in U$ および $r, s \in \mathbb{R}$ に対して,

$$g \circ f(r\mathbf{x} + s\mathbf{y}) = g(f(r\mathbf{x} + s\mathbf{y})) = g(rf(\mathbf{x}) + sf(\mathbf{y})) = rg(f(\mathbf{x})) + sg(f(\mathbf{y})) = r(g \circ f)(\mathbf{x}) + s(g \circ f)(\mathbf{y}).$$

\square

20.2 線形写像による像

写像 $f: X \rightarrow Y$ に対して, X の元を f に代入することで得られる Y の元をすべて集めた集合を f の像と呼び $f(X)$ と書くのであった. 線形写像における像には特別な記号 $\text{Im } f$ が用いられる.

定義 20.2.1. 線形空間 U から線形空間 V への線形写像 $f: U \rightarrow V$ において, f による U の像 $f(U)$ のことを $\text{Im } f$ と書く. すなわち, $\text{Im } f := \{f(\mathbf{u}) \mid \mathbf{u} \in U\}$.

命題 20.2.2. 線形写像 $f: U \rightarrow V$ の像 $\text{Im } f$ は V の部分空間である.

Proof. 部分空間になるための条件 (i) および (iv) を確認する.

(i): $f(\mathbf{0}_U) = \mathbf{0}_V$ ゆえ $\mathbf{0}_V \in \text{Im } f$ である.

(iv): $\mathbf{v}_1, \mathbf{v}_2 \in \text{Im } f$ とすれば, これらは f に代入して得られる元である. すなわち, $f(\mathbf{u}_1) = \mathbf{v}_1$, $f(\mathbf{u}_2) = \mathbf{v}_2$ を満たす $\mathbf{u}_1, \mathbf{u}_2 \in U$ が存在する. このとき, 各 $r, s \in \mathbb{R}$ について $r\mathbf{v}_1 + s\mathbf{v}_2 = rf(\mathbf{u}_1) + sf(\mathbf{u}_2) = f(r\mathbf{u}_1 + s\mathbf{u}_2)$ である. つまり, ベクトル $r\mathbf{v}_1 + s\mathbf{v}_2 \in V$ はベクトル $r\mathbf{u}_1 + s\mathbf{u}_2 \in U$ を f に代入することで得られる. ゆえに $r\mathbf{v}_1 + s\mathbf{v}_2 \in \text{Im } f$. \square

$f: U \rightarrow V$ を線形写像とし, $W \subset U$ を U の部分空間とする. f の W による像 $f(W)$ は, f の W への制限 $f|_W: W \rightarrow V$ の像とも見なせる. また, $f|_W$ も線形写像であることから, $f(W) = \text{Im } f|_W$ と書いてもよい. 上の命題から $f(W)$ も V の部分空間となることが分かる.

命題 20.2.3. 線形写像 $f: U \rightarrow V$ および組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ において,

$$f(\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle) = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle.$$

とくに $\mathbf{u}_1, \dots, \mathbf{u}_n$ が U を生成するとき, $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は $\text{Im } f$ を生成する.

Proof.

$$\begin{aligned} f(\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle) &= \{ f(\mathbf{x}) \mid \mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \} \\ &= \left\{ f\left(\sum_{i=1}^n a_i \mathbf{u}_i\right) \mid a_1, \dots, a_n \in \mathbb{R} \right\} \quad (\text{この等号は後述する}) \\ &= \left\{ \sum_{i=1}^n a_i f(\mathbf{u}_i) \mid a_1, \dots, a_n \in \mathbb{R} \right\} \quad (f \text{ の線形性 (iii)' を用いた}) \\ &= \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle. \end{aligned}$$

上の変形で明らかでないのは $\{ f(\mathbf{x}) \mid \mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \} = \{ f(\sum_{i=1}^n a_i \mathbf{u}_i) \mid a_1, \dots, a_n \in \mathbb{R} \}$ のみゆえ, これを示そう. $A = \{ f(\mathbf{x}) \mid \mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \}$, $B = \{ f(\sum_{i=1}^n a_i \mathbf{u}_i) \mid a_1, \dots, a_n \in \mathbb{R} \}$ とする. 集合の一致を示すには両方の包含関係を確認すればよい.

($A \subset B$): 各 $f(\mathbf{x}) \in A$ において, $\mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ より a_1, \dots, a_n を用いて $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ とかける. つまり $f(\mathbf{x}) = f(\sum_{i=1}^n a_i \mathbf{u}_i) \in B$.

($B \subset A$): 各 $f(\sum_{i=1}^n a_i \mathbf{u}_i) \in B$ において, $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ とおけば $\mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ である. つまり $f(\sum_{i=1}^n a_i \mathbf{u}_i)$ は, $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ の元 \mathbf{x} を用いて $f(\mathbf{x})$ と書ける. したがって $f(\sum_{i=1}^n a_i \mathbf{u}_i) \in A$.

最後に $\mathbf{u}_1, \dots, \mathbf{u}_n$ が U を生成する場合を考えよう. このとき, $U = \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ ゆえ, $\text{Im } f = f(U) = f(\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle) = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle$. すなわち, $\text{Im } f$ の各元は, $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ の線形結合として書ける. \square

上の命題より, 線形全射は「空間を生成する」という状況を保存する写像である.

例 20.2.4. (m, n) -行列 $A = [a_1, \dots, a_n]$ による線形写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($T_A(\mathbf{x}) := A\mathbf{x}$) において, 定義域 \mathbb{R}^n の標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ を T_A に代入した値 $A\mathbf{e}_1, \dots, A\mathbf{e}_n$ は A の各列ベクトル $\mathbf{a}_1, \dots, \mathbf{a}_n$ に一致する. したがって T_A の像は, A の各列ベクトルによって生成される \mathbb{R}^m の部分空間に等しい. すなわち, $\text{Im } T_A = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$.

20.3 線形写像による逆像

写像 $f: X \rightarrow Y$ および $B \subset Y$ に対して, f に代入すると B の元となるような X の元をすべて集めた集合を f の W による逆像と呼び, $f^{-1}(W)$ と書くのであった. 線形写像においては, 特別な逆像である核が調べられる.

定義 20.3.1. 線形空間 U から線形空間 V への線形写像 $f: U \rightarrow V$ において, f による $\mathbf{0}_V$ の逆像 $f^{-1}(\{\mathbf{0}_V\})$ のことを $\text{Ker } f$ と書き, これを f の核 (kernel) という. すなわち, $\text{Ker } f := \{ \mathbf{u} \in U \mid f(\mathbf{u}) = \mathbf{0}_V \}$.

例 20.3.2. (m, n) -行列 A による線形写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($T_A(\mathbf{x}) := A\mathbf{x}$) の核とは, $A\mathbf{x} = \mathbf{0}$ を満たす点 $\mathbf{x} \in \mathbb{R}^n$ 全体からなる集合である. ゆえに $\text{Ker } T_A$ は連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ の解空間 W_A に等しい.

命題 16.2.2 において斉次連立 1 次方程式の解空間が部分空間になることを見た. 一般の線形写像においても, その核は定義域の部分空間になる. より一般に, 次が成り立つ.

命題 20.3.3. 線形写像 $f: U \rightarrow V$ および終域の部分空間 $W \subset V$ に対して, f の W による逆像 $f^{-1}(W)$ は U の部分空間である.

Proof. 部分空間になるための条件 (i) および (iv) を確認する.

(i): $f(\mathbf{0}_U) = \mathbf{0}_V \in W$ ゆえ $\mathbf{0}_U$ を f に代入すると W の元となる. すなわち, $\mathbf{0}_U \in f^{-1}(W)$ である.

(iv): $\mathbf{u}_1, \mathbf{u}_2 \in f^{-1}(W)$ とすれば, $f(\mathbf{u}_1), f(\mathbf{u}_2) \in W$ である. このとき W が部分空間であることから, 各 $r, s \in \mathbb{R}$ に対して, $f(r\mathbf{u}_1 + s\mathbf{u}_2) = rf(\mathbf{u}_1) + sf(\mathbf{u}_2) \in W$ である. すなわち, $r\mathbf{u}_1 + s\mathbf{u}_2 \in f^{-1}(W)$. \square

次の事実は, 線形代数学に限らず代数学に現れる様々な準同型写像において, 断りなく用いられる事実である. ゆえに忘れてはならない.

定理 20.3.4. 線形写像 $f: U \rightarrow V$ において次は同値である:

- (1) f は単射である. すなわち, いかなる $\mathbf{v} \in \text{Im } f$ においても, その逆像 $f^{-1}(\mathbf{v})$ が 1 点からなる,
- (2) $\text{Ker } f = \{\mathbf{0}_U\}$.

Proof. (1) \Rightarrow (2): (2) は, (1) における \mathbf{v} として $\mathbf{0}_V = f(\mathbf{0}_U) \in \text{Im } f$ を考えた特別な場合である.

(2) \Rightarrow (1): $f(\mathbf{x}) = f(\mathbf{y})$ を仮定し, $\mathbf{x} = \mathbf{y}$ を示そう. ベクトル $\mathbf{x} - \mathbf{y}$ を f に代入すると $f(\mathbf{x} - \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) = \mathbf{0}_V$. つまり $\mathbf{x} - \mathbf{y} \in \text{Ker } f$ であり, 仮定より $\text{Ker } f$ は零ベクトル $\mathbf{0}_U$ のみからなるとしていたから $\mathbf{x} - \mathbf{y} = \mathbf{0}_U$ である. これを移項して $\mathbf{x} = \mathbf{y}$ を得る. \square

例 20.3.5. (m, n) -行列 A において, T_A が単射であることと $A\mathbf{x} = \mathbf{0}$ が唯一解を持つことは同値である.

単射線形写像においては命題 20.1.7 の逆も成り立つ. すなわち, 線形単射は線形独立性を保存する写像である:

命題 20.3.6. $f: U \rightarrow V$ を線形写像とし, $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ とする. f が単射であるとき次が成り立つ:

- (1) $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i \iff f(\mathbf{x}) = \sum_{i=1}^n a_i f(\mathbf{u}_i)$.
- (2) $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}_U \iff \sum_{i=1}^n a_i f(\mathbf{u}_i) = \mathbf{0}_V$.
- (3) $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形独立である $\iff f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は線形独立である.
- (4) $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である $\iff f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は線形従属である.
- (5) $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は V の基底である $\implies \mathbf{u}_1, \dots, \mathbf{u}_n$ は U の基底である.

Proof. まず (1) を示そう. (\Rightarrow) は命題 20.1.7 で得られているゆえ (\Leftarrow) のみ証明する. $f(\mathbf{x}) = \sum_{i=1}^n a_i f(\mathbf{u}_i)$ とする. このとき, $\mathbf{y} = \sum_{i=1}^n a_i \mathbf{u}_i$ とおき, $\mathbf{x} - \mathbf{y}$ を f に代入すると,

$$f(\mathbf{x} - \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) = f(\mathbf{x}) - f\left(\sum_{i=1}^n a_i \mathbf{u}_i\right) = \sum_{i=1}^n a_i f(\mathbf{u}_i) - \sum_{i=1}^n a_i f(\mathbf{u}_i) = \mathbf{0}.$$

ゆえに $\mathbf{x} - \mathbf{y} \in \text{Ker } f$ である. f は単射ゆえ $\text{Ker } f$ は $\mathbf{0}_U$ のみからなる. ゆえに $\mathbf{x} - \mathbf{y} = \mathbf{0}_U$. すなわち $\mathbf{x} = \mathbf{y} = \sum_{i=1}^n a_i \mathbf{u}_i$.

(2) は, (1) において $\mathbf{x} = \mathbf{0}_U$ とした特別な場合である. (3) は (2) より直ちに得られる. (4) は, (3) における両条件の否定をとった条件ゆえ, これらも同値である.

最後に (5) を示す. $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形独立性は (3) より分かっている. 各 $\mathbf{x} \in U$ に対して $f(\mathbf{x}) \in V$ であり, $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ が V を生成することから $f(\mathbf{x}) = \sum_{i=1}^n a_i f(\mathbf{u}_i)$ と書ける. (1) より $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$. すなわち $\mathbf{u}_1, \dots, \mathbf{u}_n$ は U を生成する. \square

命題 20.3.6(5) の逆が成り立つためには, $\text{Im } f = V$ となる必要がある. つまり, f が全単射でなければならない (命題 21.1.5).

練習 20.3.7. 命題 20.1.11 の設定において v_1, \dots, v_n が線形独立であるとき, $f(u_i) = v_i$ を満たす線形写像 $f: U \rightarrow V$ は単射である. これを示せ.

解答例: $\text{Ker } f = \{0_U\}$ を示せばよい. そこで $x \in \text{Ker } f$ とする. u_1, \dots, u_n が U を生成することから $x = \sum_{i=1}^n a_i u_i$ と書けば,

$$0 = f(x) = f\left(\sum_{i=1}^n a_i u_i\right) = \sum_{i=1}^n a_i f(u_i) = \sum_{i=1}^n a_i v_i.$$

v_1, \dots, v_n は線形独立ゆえ $a_1 = \dots = a_n = 0$ であり, $x = \sum_{i=1}^n 0 u_i = 0_U$ を得る. すなわち $\text{Ker } f$ の元は 0_U に限る. \square

20.4 様々な線形写像の例

例 20.4.1. \mathbb{R}^n のベクトルに対して第 i 座標を対応させる写像 $\text{pr}_i: \mathbb{R}^n \rightarrow \mathbb{R}$ ($\text{pr}_i(x_1, \dots, x_n) := x_i$) は第 i 座標への射影 (**projection**) と呼ばれる. pr_i は線形写像である.

例 20.4.2. n 次正方形行列 $A = [a_{ij}]$ に対して A の対角成分の和 $\text{tr } A := \sum_{i=1}^n a_{ii}$ を A の跡 (**trace**) あるいはトレース, 対角和などと呼ぶ. 写像 $\text{tr}: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ は線形写像である.

命題 20.4.3. n 次正方形行列 A, B について $\text{tr}(AB) = \text{tr}(BA)$.

Proof. $A = [a_{ij}]_{i,j}, B = [b_{ij}]_{i,j}$ とすれば, $AB = [\sum_{k=1}^n a_{ik} b_{kj}]_{i,j}$ ゆえ $\text{tr}(AB) = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki}$. また, $BA = [\sum_{\ell=1}^n b_{p\ell} a_{\ell q}]_{p,q}$ ゆえ $\text{tr}(BA) = \sum_{p=1}^n \sum_{\ell=1}^n b_{p\ell} a_{\ell p} = \sum_{\ell=1}^n \sum_{p=1}^n a_{\ell p} b_{p\ell}$. この式に現れる添え字 ℓ, p を i, j に置き換えれば, これは $\text{tr}(AB)$ と同じ式である. \square

例 20.4.4. (1) 数列空間 $\mathbb{R}^{\mathbb{N}}$ において列を左にずらす写像, すなわち $S(x_1, x_2, x_3, \dots) := (x_2, x_3, \dots)$ で定められる写像 $S: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ はシフト作用素と呼ばれる. S は線形写像である.

(2) S は全射である. 何故なら, 各数列 $\mathbf{y} = (y_1, y_2, y_3, \dots)$ に対して, $\mathbf{x} = (0, y_1, y_2, \dots)$ と定めれば $S(\mathbf{x}) = \mathbf{y}$ となる. また, $S(\mathbf{x}) = \mathbf{0}$ なる数列 \mathbf{x} は, 初項を除いてすべて 0 なる数列であり, $\text{Ker } S = \{(x, 0, 0, \dots) \mid x \in \mathbb{R}\}$.

(3) 数列 $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ が線形漸化式 16.3.2:

$$x_{n+k} = a_{k-1}x_{n+k-1} + a_{k-2}x_{n+k-2} + \dots + a_1x_{n+1} + a_0x_n$$

を満たすならば, 数列 $S(\mathbf{x}) = (x_{n+1})_{n \in \mathbb{N}}$ もまた上の漸化式を満たす.

(4) 漸化式 $x_{n+2} = x_{n+1} + x_n$ を満たす数列全体 F は $\mathbb{R}^{\mathbb{N}}$ の部分空間である (例 16.3.1). S を F に制限した $S: F \rightarrow F$ は全単射である.

Proof. 全射性: 各 $\mathbf{y} = (y_1, y_2, y_3, \dots)$ に対して, $\mathbf{x} = (y_2 - y_1, y_1, y_2, y_3, \dots)$ と定めれば $\mathbf{x} \in F$ であり $S(\mathbf{x}) = \mathbf{y}$ となる. なお, $\mathbf{x} = S(\mathbf{y}) - \mathbf{y}$ である. 実際, これらの列の初項は等しい. また, 第 2 項以降について, 左辺の第 n 項は y_{n-1} , 右辺は $y_{n+1} - y_n$ であり, 数列 \mathbf{y} が漸化式 $y_{n+1} = y_n + y_{n-1}$ ($n \geq 2$) を満たすことからこれらも等しい³.

単射性: $\text{Ker } S = \{0\}$ を示せばよい. $\mathbf{x} = (x_1, x_2, x_3, \dots) \in F, S(\mathbf{x}) = \mathbf{0}$ とすれば (2) より \mathbf{x} の第 2 項以降はすべて 0 であり, これに漸化式の条件 $x_3 = x_2 + x_1$ を合わせて $x_1 = 0$ を得る. つまり $\mathbf{x} = \mathbf{0}$ である. \square

³実は, 初項と第 2 項が等しいことさえ分かれば, F の元は漸化式 $x_{n+2} = x_{n+1} + x_n$ を満たすゆえ第 3 項以降もすべて等しいことが分かる.

例 20.4.5. (1) I を開区間とし, I 上の実数値 C^∞ -級関数全体を $C^\infty(I)$ とする. $C^\infty(I)$ の元 f は何回でも微分できるゆえ, f の微分 f' もまた何回でも微分できる, すなわち $f' \in C^\infty(I)$. そこで, $D: C^\infty(I) \rightarrow C^\infty(I)$ を $D(f) := f'$ と定めれば D は線形写像である.

(2) D は全射である. 実際, 各 $f \in C^\infty(I)$ に対して f の原始関数を F とすれば, F を 1 回微分すると f になり, また f は何回でも微分できる. すなわち F も何回でも微分可能であり $F \in C^\infty(I)$. $D(F) = F' = f$ より D の全射性を得る. 一方, 微分すると $\mathbf{0} \in C^\infty(I)$ になる関数は定数関数しかないことから⁴, $\text{Ker } D$ は定数関数全体のなす集合になる.

(3) 関数 $y \in C^\infty(I)$ が線形微分方程式 16.3.4:

$$y^{(k)}(x) = a_{k-1}y^{(k-1)} + a_{k-2}y^{(k-2)} + \cdots + a_1y^{(1)}(x) + a_0y^{(0)}(x)$$

を満たすならば, 上の両辺を微分することで関数 $D(y)$ もまた上の微分方程式を満たすことが分かる.

(4) 微分方程式 $y^{(2)} = y^{(1)} + y^{(0)}$ を満たす C^∞ -級関数全体 W は $C^\infty(I)$ の部分空間である (例 16.3.3). D を W に制限した $D: W \rightarrow W$ は全単射である.

Proof. 全射性: 各 $y \in W$ に対して $F := D(y) - y$ (つまり $F(x) = y'(x) - y(x)$) と定めれば W は部分空間であったから $F \in W$ である. このとき $D(F) = y$, すなわち $F'(x) = y(x)$ が成り立つ. 実際, y が微分方程式 $y^{(2)}(x) = y^{(1)}(x) + y(x)$ を満たすことから, $F'(x) = (y'(x) - y(x))' = y^{(2)}(x) - y^{(1)}(x) = y(x)$.

単射性: $\text{Ker } D = \{\mathbf{0}\}$ を示せばよい. $y \in W$, $D(y) = \mathbf{0}$ とすれば (2) より y は定数関数である. つまり $y' = \mathbf{0}$, $y'' = \mathbf{0}$ であり, これに微分方程式の条件 $y''(x) = y'(x) + y(x)$, すなわち $\mathbf{0} = \mathbf{0} + y$ を合わせて $y = \mathbf{0}$ を得る. \square

微分を用いた線形写像は微分作用素と呼ばれ, とくに上の写像 D は通常 $\frac{d}{dx}$ で表す.

例 20.4.6 (発展). $a \in \mathbb{R}$ を固定し, 写像 $T: C^\infty(\mathbb{R}) \rightarrow \mathbb{R}^\mathbb{N}$ を

$$T(f) := (f(a), f^{(1)}(a), f^{(2)}(a), \dots)$$

と定めれば, これは線形写像である. また, 例 20.4.4 におけるシフト作用素 S および例 20.4.5 における微分作用素 D について $S \circ T = T \circ D$ が成り立つ. これを図式で表すと次のようになる.

$$\begin{array}{ccccc} C^\infty(\mathbb{R}) & \ni & f & \xrightarrow{T} & (f(a), f^{(1)}(a), f^{(2)}(a), f^{(3)}(a), \dots) & \in & \mathbb{R}^\mathbb{N} \\ D \downarrow & & \downarrow & & \downarrow & & \downarrow S \\ C^\infty(\mathbb{R}) & \ni & f^{(1)} & \xrightarrow{T} & (f^{(1)}(a), f^{(2)}(a), f^{(3)}(a), f^{(4)}(a), \dots) & \in & \mathbb{R}^\mathbb{N} \end{array}$$

【補足】 点 a の周りで冪級数展開 (テーラー展開) 可能であり, かつ収束半径が無限大となる関数に限れば T は単射になり, 上の図式は微分作用素とシフト作用素がほぼ同等であることを示唆している.

例 20.4.7 (発展). I を \mathbb{R} 上の開区間とし, I 上の実数値連続関数全体のなす線形空間を $C(I)$ とする. また $a, b \in I$ を固定しておく.

⁴実際, 対偶「 f が定数関数でないならば $f' \neq \mathbf{0}$ 」が次のように示される: f が定数関数でないならば, $f(x) \neq f(y)$ をみたす異なる二点 $x, y \in I$ がある. $a = \frac{f(x)-f(y)}{x-y}$ とおけば ($x \neq y$ より分母は 0 でない), $f(x) \neq f(y)$ より $a \neq 0$ である. 平均値の定理より $f'(t) = a$ を満たす $t \in I$ が存在し, とくに f' はゼロ値定数関数 $\mathbf{0}$ ではない.

(1) $S_{a,b} : C(I) \rightarrow \mathbb{R}$ を $S_{a,b}(f) := \int_a^b f(x)dx$ と定めれば $S_{a,b}$ は線形写像である. $\text{Ker } S_{a,b}$ の元は区間 $[a, b]$ における定積分が 0 となる関数である. $\text{Im } S_{a,b}$ は区間 $[a, b]$ が幅を持つかどうかで異なる. $a = b$ の場合, 区間 $[a, b]$ における積分はゼロゆえ, $\text{Im } S_{a,b} = \{0\}$. 言い換えると $\text{Ker } S_{a,b} = C(I)$ である. 一方 $a \neq b$ ならば積分は様々な値を取り得る. 実際, 各 $r \in \mathbb{R}$ に対して, 定数関数 $f(x) = \frac{r}{b-a}$ を取れば $S_{a,b}(f) = r$ である. すなわち, $S_{a,b}$ は全射であり, $\text{Im } S_{a,b} = \mathbb{R}$.

(2) $C(I)$ の元 f に対して次で定義される $C(I)$ の元 $I_a(f) : I \rightarrow \mathbb{R}$ を対応させる写像 $I_a : C(I) \rightarrow C(I)$ を考える.

$$I_a(f)(x) := \int_a^x f(t)dt.$$

I_a は線形写像である. 微分積分学の基本定理より, $I_a(f)$ は f の原始関数のうちの一つである. $\text{Im } I_a$ は $F(a) = 0$ を満たす C^1 -級関数⁵ F たち全体に一致する.

Proof. $V = \{F \in C(I) \mid F(a) = 0 \text{ かつ } F \text{ は } C^1\text{-級}\}$ とし, $\text{Im } I_a = V$ を示そう.

($\text{Im } I_a \subset V$): $\text{Im } I_a$ の各元 $F = I_a(f)$ が $F(a) = 0$ を満たすことは定義から直ちにわかる (区間 $[a, a]$ 上の定積分は 0). また, F の微分が連続関数 f になることから, F は C^1 -級である. ゆえに $F \in V$.

($V \subset \text{Im } I_a$): $F \in V$ とする. $f := F'$ とすれば, F は C^1 -級ゆえ $f \in C(I)$ である. $G := I_a(f)$ とおけば $G' = f$ である. ゆえに $(F-G)' = F' - G' = f - f = \mathbf{0}$. 微分が $\mathbf{0}$ になる関数は定数関数のみであるから⁶, $F-G$ は定数関数であり, これに a を代入すると $(F-G)(a) = F(a) - G(a) = 0 - 0 = 0$. ゆえに関数 $F-G$ は 0 値定数関数であり, $F-G = \mathbf{0}$. 以上より $F = G = I_a(f) \in \text{Im } I_a$. \square

$I_a(f) = I_a(g)$ とすれば両辺を微分して $f = g$ となるゆえ I_a は単射であり $\text{Ker } I_a = \{\mathbf{0}\}$.

(3) $\text{Im } I_a$ の各元は微分可能な関数であり, ゆえにそれらは微分作用素 $D = \frac{d}{dx}$ に代入することができる. 微分積分学の基本定理より $D \circ I_a = \text{id}_{C(I)}$ である.

例 20.4.8 (発展). X, Y を \mathbb{R} の部分集合とする (あるいはより一般に $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$ としてもよい). 連続写像 $F : X \rightarrow Y$ を一つ与えると, 次のような写像 $T_F : C(Y) \rightarrow C(X)$ が定義できる (写像の向きが F と逆になっていることに注意せよ).

$$T_F(f) := f \circ F, \quad (\text{ここで } f \in C(Y), \text{ つまり } f : Y \rightarrow \mathbb{R} \text{ は連続関数}).$$

(1) T_F は線形写像である.

Proof. まず $T_F(f) \in C(X)$ であること, すなわち $f \circ F : X \rightarrow \mathbb{R}$ が連続であることを示さねばならないが, これは微積分の教科書に譲り (一般に, 連続写像による合成は連続である), 線形性 (iii) のみ確認しよう. 各 $f, g \in C(Y)$ および $r, s \in \mathbb{R}$ について

$$T_F(rf + sg) = rT_F(f) + sT_F(g)$$

を示したい. これらは共に X を定義域とする写像であるから, 左右の両写像に各 $x \in X$ を代入した値が等しいことが分かれば上の等式を得る. $C(X)$ における和とスカラー倍の定義を思いだしな

⁵導関数が連続となる関数を C^1 -級関数という.

⁶「 $g' = \mathbf{0}$ ならば g は定数関数」は先程示した「 g は定数関数でないならば $g' \neq \mathbf{0}$ 」の対偶にあたる.

がら計算すると,

$$\begin{aligned} T_F(rf + sg)(x) &= (rf + sg) \circ F(x) = (rf + sg)(F(x)) = (rf)(F(x)) + (sg)(F(x)) \\ &= r \cdot f(F(x)) + s \cdot g(F(x)). \end{aligned}$$

$$\begin{aligned} (rT_F(f) + sT_F(g))(x) &= (rT_F(f))(x) + (sT_F(g))(x) = (r(f \circ F))(x) + (s(g \circ F))(x) \\ &= r \cdot (f \circ F)(x) + s \cdot (g \circ F)(x) = r \cdot f(F(x)) + s \cdot g(F(x)). \end{aligned}$$

ゆえに $T_F(rf + sg) = rT_F(f) + sT_F(g)$ が成り立つ. □

(2) T_F は $C(X)$ の積演算とも整合的な写像である. すなわち, $T_F(fg) = T_F(f)T_F(g)$.

Proof. 各 $x \in X$ を代入した値が一致することを示せばよい.

$$T_F(fg)(x) = (fg) \circ F(x) = (fg)(F(x)) = f(F(x)) \cdot g(F(x)).$$

$$(T_F(f)T_F(g))(x) = (T_F(f))(x) \cdot (T_F(g))(x) = (f \circ F)(x) \cdot (g \circ F)(x) = f(F(x)) \cdot g(F(x)).$$

□

よりみち (二つの行為の可換性)

中学校で数学を学び始めたころ, 次のような間違っただ式変形をした経験はないだろうか:

$$(a + b)^2 = a^2 + b^2, \quad \sqrt{a + b} = \sqrt{a} + \sqrt{b}.$$

賢明な読者ならばこのような間違いは犯さなかったかもしれないが, 上述のミス指摘されたご友人は少なくともいたことだろう. このようなミスを犯す背景には, 関数は線形性 (i): $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ を満たすべきであるという考えが, 無意識の内に潜んでいるからなのかもしれない. 我々人間には, 物事を線形であると考えてしまう思考のクセがあるのだろうか. もし, そうであるとするならば, 線形性に関する基礎事項を把握しておくことは極めて重要である.

線形性 (i) を文章で述べなおせば, 「和を取った後に代入したものと代入してから和をとったものが等しい」となる. これは「和をとる」および「代入する」という二つの行為の交換可能性 (可換性), つまり順番を入れ替えても結果が変わらないことを言っている. また, 線形性 (ii) は「スカラー倍をとる」および「代入する」という操作の可換性を意味している. 一般には, 二つの行為の順番を入れ替えれば結果は異なるのが普通であり (例えば試験勉強をしてから試験を受けるのと試験を受けてから試験勉強をするのでは成績に大きな差がでることだろう), したがって, 線形性を要求することは, 特殊な状況を想定していると考えられる. それにも関わらず, 線形写像の例が豊富に挙がるのは, どういった事情によるのであろうか.

結合律や分配法則をはじめとして, 二つの行為の可換性を前提とする数学は豊富にある. 一般に, 代数学で扱う写像では主に演算と代入の可換性を仮定する (このような写像は準同型 (**homomorphism**) と呼ばれる). 例えば置換の符号を定める写像はこうした可換性 $\text{sgn}(\sigma\tau) = \text{sgn}(\sigma)\text{sgn}(\tau)$ を満たす. 他方で, 微分積分学の基礎で扱われる連続写像は, 極限操作と代入の可換性を許すものであった:

$$\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right).$$

数学における基本的な枠組みの中で可換性がよく現れるのは何故だろうか. もしかすると, 可換性を通して物事を理解しようとする傾向が人間にはあるのかもしれない.

第21章 線形空間の同一視

認識した対象を区別する、あるいは分類するという行為は、その対象を理解するための基本的な手段である。分類(類別)とは、似たものをどうしを集め、また著しく異なっているものを分けることをいう。線形空間の分類の場合、何を基準とするのが妥当であろうか。本章では、まずはじめに、線形代数的な性質の相互翻訳が可能で1対1の対応(線形同型写像)について論じる。そして、二つの線形空間の間にこの対応があるかどうかを同一視の基準と定め、その定義の妥当性について考察する。

次に、ユークリッド空間上の線形写像のなす空間と行列のなす空間が同一視できることを見る。とくに \mathbb{R}^n から \mathbb{R}^n への線形写像全体は n 次正方行列全体と1対1に対応し、これらは線形空間であるだけでなく積演算を持つ多元環となる。本章の後半では、多元環の元に関する複雑な演算を多項式を用いて簡明に記述する手法を導入する。これは、線形代数学の一般論ではケーリー-ハミルトンの定理において最初に用いられる。また、線形微分方程式の表示にも利用できる記法である。

21.1 線形同型写像

次で与える対応によって、線形代数的な性質が相互に翻訳されることを見よう。

定義 21.1.1. 線形写像 $f: U \rightarrow V$ が全単射であるとき、これを線形同型 (linear isomorphism) あるいは単に同型という。

例 21.1.2. 単射線形写像 $f: U \rightarrow V$ において、終域を置き換えた写像 $f: U \rightarrow \text{Im } f$ は線形同型である。

例 21.1.3. U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$ とする。このとき、命題 20.1.11 より $f(\mathbf{u}_i) = \mathbf{e}_i$ を満たす線形写像 $f: U \rightarrow \mathbb{R}^n$ が存在する。この f は線形同型である。

Proof. 命題 20.2.3 より $\text{Im } f = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle = \langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle = \mathbb{R}^n$ 。すなわち f は全射である。また、単射性は練習 20.3.7 による。□

線形写像 f が全単射ならば次の命題により f^{-1} も自動的に線形写像となる。例 19.3.5(i) より f^{-1} は全単射であり、したがって f^{-1} も線形同型である。

命題 21.1.4. 線形同型 $f: U \rightarrow V$ の逆写像 $f^{-1}: V \rightarrow U$ は線形写像である。

Proof. f^{-1} が線形性 (iii) を満たすことを示そう。そこで $\mathbf{x}, \mathbf{y} \in V$ および $r, s \in \mathbb{R}$ を任意に取る。 $\mathbf{a} = f^{-1}(\mathbf{x})$, $\mathbf{b} = f^{-1}(\mathbf{y})$ とおく。逆写像の定義から $f(\mathbf{a}) = \mathbf{x}$, $f(\mathbf{b}) = \mathbf{y}$ である。また f の線形性より $f(r\mathbf{a} + s\mathbf{b}) = rf(\mathbf{a}) + sf(\mathbf{b}) = r\mathbf{x} + s\mathbf{y}$ である。この両辺をそれぞれ f^{-1} に代入することで

$$\begin{aligned} f^{-1}(f(r\mathbf{a} + s\mathbf{b})) &= f^{-1}(r\mathbf{x} + s\mathbf{y}) \\ r\mathbf{a} + s\mathbf{b} &= f^{-1}(r\mathbf{x} + s\mathbf{y}) \\ rf^{-1}(\mathbf{x}) + sf^{-1}(\mathbf{y}) &= f^{-1}(r\mathbf{x} + s\mathbf{y}). \end{aligned}$$

上の最後の式は f^{-1} が線形性 (iii) を満たすことを意味している。□

命題 20.3.6 から更に踏み込んで、線形同型では次が成立する。

命題 21.1.5. $f: U \rightarrow V$ が線形同型であるとき、 $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ について次が成り立つ。

$$\mathbf{u}_1, \dots, \mathbf{u}_n \text{ は } U \text{ の基底である} \iff f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \text{ は } V \text{ の基底である.}$$

Proof. (\Leftarrow) は命題 20.3.6(5) より得られているゆえ (\Rightarrow) のみを示せばよい. $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とする. 命題 20.3.6(3) より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は線形独立である. また, f は全射ゆえ $\text{Im } f = V$ であり, 命題 20.2.3 より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は V を生成する. \square

逆に, 基底が対応し合う線形写像は同型である:

命題 21.1.6. U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$, $f: U \rightarrow V$ を線形写像とするとき, 次が成り立つ.

f は線形同型である $\iff f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は V の基底である.

Proof. (\Rightarrow) は命題 21.1.5 より明らか. (\Leftarrow) は例 21.1.3 における証明とほとんど同じ論法で得られる. \square

命題 21.1.7. 線形同型の合成は線形同型である.

Proof. $f: U \rightarrow V$ および $g: V \rightarrow W$ を共に線形同型とすれば命題 20.1.12 より $g \circ f$ は線形写像であり, また命題 19.4.1(3) より全単射である. \square

本節において, 前章で述べた一般の線形写像について片側の矢印: \implies あるいは \impliedby しか成り立たなかった諸性質が, 線形同型写像においては必要十分 (\iff) になることを見た. これはつまり, 対応する二つの線形空間の間で相互翻訳ができることを意味している.

21.2 同型な線形空間

二つの線形空間 U, V が本質的に同じであるとはどういうことか考えよう. それは, 線形空間の枠組みにおいてそれらを区別できないことと定めるのが妥当である. すなわち, 線形空間上のあらゆる命題において U と V における真偽が一致することに他ならない. 言い換えれば, U で成り立つことと同等の現象が必ず V においても成り立ち, またその逆も言えるということである. 線形空間の定義を振り返れば, それは和とスカラー倍の演算の性質のみを規定した対象であったから, 線形空間の現象として述べることができるのは, それらの演算に関する言明のみである. したがって, U と V の間に和とスカラー倍の演算について整合的な 1 対 1 対応があるとき, 線形空間の枠組みで語れる現象のみを用いて U と V を区別することはできないことになる. 以上の考察から, 次の線形同型なる概念 (本質的に同じものとみなすこと) を得る.

定義 21.2.1. 線形空間 U, V の間に線形同型写像 $f: U \rightarrow V$ が存在するとき, U と V は線形同型である (**linearly isomorphic**) あるいは単に同型であるという. このとき, $U \simeq V$ と書く.

例 21.2.2. 2 次行ベクトル全体 $U = M_{1,2}(\mathbb{R})$ と 2 次列ベクトル全体 $V = M_{2,1}(\mathbb{R})$ について考えよう. 前者は横に成分を並べたベクトルの集合であり, 後者は縦に成分を並べたそれである. ゆえにこれらは見た目の上では異なっているとも考えられる. しかし, 線形空間的な性質 (つまり演算に関する性質) における違いのみを用いてこれらを区別することはできない. 実際, 写像 $f: U \rightarrow V$ を $f(\mathbf{x}) := {}^t \mathbf{x}$ と定めればこれは線形同型であり, U において $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ に関する線形空間的な性質 $P(\mathbf{u}_1, \dots, \mathbf{u}_n)$ が成立するとき, V において $P(f(\mathbf{x}_1), \dots, f(\mathbf{x}_n))$ が成立することが示唆される¹. すなわち, U で成り立ち V で成り立たないような線形空間の枠組みにおける現象を挙げることはできない.

いま, 線形空間を同一視する基準を与えた. ここで細かい理屈をこねれば, U 自身は U と同一視できるか, あるいは U と同型な線形空間と同型な線形空間は U と同型か, といった素朴な疑問が生じよう. 学問上のあらゆる理論は, こうした些細な疑問にも答えられるよう構築されねばならない.

命題 21.2.3. 線形空間 U, V, W について次が成り立つ.

(反射律) $U \simeq U$, (対称律) $U \simeq V \implies V \simeq U$, (推移律) $U \simeq V, V \simeq W \implies U \simeq W$.

¹ここでは「示唆される」と述べるに留め, 証明は行わない. 任意の性質 P についてこの事実を示すには, 数理論理的な枠組みにおいて「線形空間的な性質 (論理式)」を再帰的に定義し, 論理式の長さに関する帰納法を用いればよい.

Proof. (反射律): 恒等写像 $\text{id}_U : U \rightarrow U$ は線形同型ゆえ $U \simeq U$ である. (対称律): $U \simeq V$ とすれば線形同型 $f : U \rightarrow V$ が存在し, その逆写像 $f^{-1} : V \rightarrow U$ も線形同型であることから $V \simeq U$ を得る. (推移律): $U \simeq V, V \simeq W$ とすれば線形同型 $f : U \rightarrow V, g : V \rightarrow W$ が存在し, これらの合成 $g \circ f : U \rightarrow W$ が線形同型であることから $U \simeq W$ を得る. \square

よりみち (同一視すること)

一般に, 二つの対象を結ぶ記号 (集合論では, これを関係と呼ばれる概念を用いて定義する) が反射律および対称律, 推移律を満たすとき, これを同値関係 (equivalence relation) と言う. 線形代数学に限らず何らかの立場で二つの対象を同一のものとみなすとき, その同一性は同値関係になることが望ましい. これを認めるならば, 考えるべき対象に同値関係を与えるということは, 同一視する基準を与えることと言い換えてもよい.

同値関係となるような概念はこれまでもいくつか学んでいる.

例 21.2.4. (1) 平面または空間上の図形 A, B が合同であるとき, $A \equiv B$ と書く. また, 図形 A, B が相似であるとき $A \sim B$ と書く (記号 \sim は国際標準ではない). 合同 \equiv および相似 \sim はそれぞれ同値関係である.

(2) 二つの集合 A, B の間に全単射が存在するとき A と B は対等であるというのであった. 対等は同値関係である. 反射律は例 19.2.2 により, 対称律は例 19.3.5(i) に相当する. また, 推移律は命題 19.5.5 で述べた.

同型な線形空間の例を挙げよう.

例 21.2.5. (1) 線形空間 U および V において同じ個数からなる基底が取れるとき, U と V は線形同型である. 実際, U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$, V の基底を $\mathbf{v}_1, \dots, \mathbf{v}_n$ とすれば, 命題 20.1.11 により $f(\mathbf{u}_i) = \mathbf{v}_i$ を満たす線形写像 $f : U \rightarrow V$ が取れる. この写像は命題 21.1.6 より線形同型である.

(2) $M_{m,n}(\mathbb{R})$ と \mathbb{R}^{mn} は共に mn 個のベクトルからなる基底を持つ. ゆえにこれらは線形同型である.

例 21.2.6. 例 15.2.2 および例 15.2.8(2) において, 異なる集合に同一の記号 $\mathbb{R}^{\mathbb{N}}$ を与えていた. 前者の数列空間を U とし, 後者の写像空間を V としよう. このとき, 例 15.2.8(2) で与えた次の対応:

$$f : U \rightarrow V, \quad f(x_1, x_2, \dots) := "x(n) := x_n \text{ で定義される写像 } x : \mathbb{N} \rightarrow \mathbb{R}"$$

は線形同型となる. したがって, 線形空間の枠組みにおいてこれらは同一のものとみなすことができ, それゆえ, これら二つの空間に同じ記号 $\mathbb{R}^{\mathbb{N}}$ を与えている.

$M_{m,n}(\mathbb{R})$ と \mathbb{R}^{mn} は線形同型であるにもかかわらず同じ記号を用いることはない. これは, 線形空間としての構造だけではなく, 積の構造も加味した代数構造として $M_{m,n}(\mathbb{R})$ を扱うためである. $M_{m,n}(\mathbb{R})$ の元と $M_{n,r}(\mathbb{R})$ の元の間に行列としての積が定義されている一方で, \mathbb{R}^{mn} の元と \mathbb{R}^{nr} の元の間には自然な積が定まらない. ゆえに, これらを同一の記号で表すのは好ましくない.

【よりみち】 ちなみに \mathbb{R}^m の元どうしの積を座標ごとの掛け算によって定めることができる. しかしながら $m = nn$ におけるこの積演算と $M_n(\mathbb{R})$ の積演算 (行列としての積) の間に自然な対応は与えられない. 前者は可換である, すなわち常に $\mathbf{xy} = \mathbf{yx}$ が成り立つのに対し, 後者はそうではないからである. なお, \mathbb{R}^m と, m 点集合 $X_m = \{1, \dots, m\}$ 上の関数 (連続関数) 全体 $C(X_m) = \mathbb{R}^{X_m}$ の間には積演算を込めた意味での (すなわち多元環としての) 自然な 1 対 1 対応がつく. 実際, 有限数列 (第 m 項までの数列) について例 21.2.6 と同様の対応を考えれば, $C(X_m)$ との間には積演算についても整合的な 1 対 1 対応が得られる.

次は線形同型でない例である.

命題 21.2.7. 自然数 $m < n$ について \mathbb{R}^m と \mathbb{R}^n は線形同型でない.

Proof. もし線形同型 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ が存在するとすれば、定義域 \mathbb{R}^n の標準ベクトル e_1, \dots, e_n について命題 21.1.5 (より根本的には命題 20.3.6(3)) より $f(e_1), \dots, f(e_n)$ は \mathbb{R}^m における線形独立な組となり、これは命題 17.3.5 に矛盾する。ゆえに \mathbb{R}^m と \mathbb{R}^n の間に線形同型写像は存在せず、 \mathbb{R}^m と \mathbb{R}^n は線形同型でない。 \square

命題 21.2.7 の証明で用いた f の性質は単射性のみである。つまり、自然数 $m < n$ について線形単射 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ は存在しない。一方、逆向きの線形全射 $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ の非存在性も示すことができる (系 22.4.6)。

21.3 線形写像のなす空間

1.3 節で予告したように、行列とは線形写像の数値化にほかならない。すなわち、線形写像のなす空間と行列のなす空間は同一視できる。1.3 節において \mathbb{R}^2 の間で定まる線形写像について紹介した事実の一般化を本節で述べよう。そのためには、 (m, n) -行列全体 $M_{m,n}(\mathbb{R})$ に対応する線形写像のなす集合に記号を与えておかねばならない。 $M_{m,n}(\mathbb{R})$ には線形空間の構造が入っていたゆえ、対応する線形写像の集合も線形空間になることが示唆される。

定義 21.3.1. 線形空間 U から線形空間 V への線形写像全体のなす集合を $\text{Hom}(U, V)$ と書く。とくに $U = V$ のとき、これを $\text{End}(U)$ と書く。 $\text{Hom}(U, V)$ には次のように和とスカラー倍が定まり、線形空間となる：

- $f, g \in \text{Hom}(U, V)$ に対して $(f + g): U \rightarrow V$ を次の写像として定める：
各 $x \in U$ について、 $(f + g)(x) := f(x) + g(x)$.
- $f \in \text{Hom}(U, V)$ および $r \in \mathbb{R}$ に対して $(rf): U \rightarrow V$ を次の写像として定める：
各 $x \in U$ について、 $(rf)(x) := rf(x)$.

$\text{Hom}(U, V)$ における零ベクトルは、 U の各元に対して $\mathbf{0}_V$ を対応させる定値写像 $f = \mathbf{0}_V$ である。

記号 Hom は準同型写像 (homomorphism) に由来する。線形空間に限らず、与えられた代数構造に関する演算と相性のよい写像、すなわち演算との合成が交換可能 (可換) になる写像のことを準同型と言う。また、ある代数構造からそれ自身への準同型のことを自己準同型 (**endomorphism**) と言う。線形写像は線形空間に関する準同型であり、とくに線形空間上の自己準同型は線形変換²と呼ばれる。

練習 21.3.2. (1) 上で定めた和とスカラー倍において、線形写像の和が線形写像になること、および線形写像のスカラー倍が線形写像になることを示せ。すなわち、 $f, g \in \text{Hom}(U, V)$, $r \in \mathbb{R}$ ならば $f + g, rf \in \text{Hom}(U, V)$ ということである。

解答例: $f + g$ が線形写像となることを示そう。線形性 (iii) は次のように確認できる：

$$\begin{aligned} (f + g)(s\mathbf{x} + t\mathbf{y}) &= f(s\mathbf{x} + t\mathbf{y}) + g(s\mathbf{x} + t\mathbf{y}) = (sf(\mathbf{x}) + tf(\mathbf{y})) + (sg(\mathbf{x}) + tg(\mathbf{y})) \\ &= s(f(\mathbf{x}) + g(\mathbf{x})) + t(f(\mathbf{y}) + g(\mathbf{y})) = s(f + g)(\mathbf{x}) + t(f + g)(\mathbf{y}). \end{aligned}$$

\square

(2) $\text{Hom}(U, V)$ における和とスカラー倍の演算がベクトル空間の公理を満たすことを確かめよ。

(m, n) -行列 A において、線形写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ を $T_A(\mathbf{x}) := A\mathbf{x}$ と定めるのであった。行列における演算と T_A との関係についてまとめておこう。

命題 21.3.3. (1) サイズの等しい行列 A, B について $T_A = T_B \iff A = B$.

²英語の文献では、一般の線形写像を linear transformation と書くことがある。

(2) サイズの等しい行列 A, B および $r, s \in \mathbb{R}$ について $T_{rA+sB} = rT_A + sT_B$.

(3) 行列 B と A の間に積が定まるとき, $T_{BA} = T_B \circ T_A$.

(4) A が可逆 $\iff T_A$ は全単射. また, このとき, $(T_A)^{-1} = T_{A^{-1}}$.

Proof. (1): (\Leftarrow) は明らかゆえ (\Rightarrow) のみ示す. A, B の各列が等しいことが示されれば $A = B$ である. $T_A(e_i)$ および $T_B(e_i)$ はそれぞれ A, B の i 列目であり, 仮定より $T_A(e_i) = T_B(e_i)$ である. ゆえに A, B の各列は等しい.

(2): $T_{rA+sB}(\mathbf{x}) = (rA + sB)(\mathbf{x}) = r(A\mathbf{x}) + s(B\mathbf{x}) = rT_A(\mathbf{x}) + sT_B(\mathbf{x})$.

(3): $T_{BA}(\mathbf{x}) = (BA)\mathbf{x} = B(A\mathbf{x}) = BT_A(\mathbf{x}) = T_B(T_A(\mathbf{x})) = T_B \circ T_A(\mathbf{x})$.

(4): A を可逆とすれば, (3) より $T_A \circ T_{A^{-1}} = T_E = \text{id}_{\mathbb{R}^n}$, $T_{A^{-1}} \circ T_A = \text{id}_{\mathbb{R}^n}$ である. したがって命題 19.3.3 より T_A は全単射であり, その逆写像は $T_{A^{-1}}$ である. 逆に T_A が全単射であると仮定し, A の可逆性を示そう. 線形写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ が同型であることから命題 21.2.7 より $m = n$ ゆえ A は正方行列である. また, 標準基底 e_1, \dots, e_n について命題 20.3.6(3) より $T_A(e_1), \dots, T_A(e_n)$ は線形独立である. すなわち, 正方行列 A の各列ベクトルによる組は線形独立であり, 定理 17.3.6 より A は可逆である. \square

定理 21.3.4. 任意の線形写像 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ は (m, n) -行列 A_f を用いて $f = T_{A_f}$ と一意的に表せる. 更に, $\mathcal{T}: M_{m,n}(\mathbb{R}) \rightarrow \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ を $\mathcal{T}(A) := T_A$ と定めれば, これは線形同型写像である.

Proof. $A_f := [f(e_1), \dots, f(e_n)]$ と定めれば $f = T_{A_f}$ である. 実際, $T_{A_f}(e_i) = \text{“}A_f \text{の} i \text{列目”} = f(e_i)$ ゆえ T_{A_f} と f における標準基底の行き先は等しい. したがって命題 20.1.10(2) より $T_{A_f} = f$ である.

いま示したことは, \mathcal{T} の全射性にほかならない. また, \mathcal{T} の単射性は命題 21.3.3(1) による. \mathcal{T} の線形性 (iii) は命題 21.3.3(2) による. \square

上の定理における \mathcal{T} の逆写像 $\mathcal{R}: \text{Hom}(\mathbb{R}^n, \mathbb{R}^m) \rightarrow M_{m,n}(\mathbb{R})$ は, $f \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ に対して $A_f = [f(e_1), \dots, f(e_n)]$ を対応させる写像である.

21.4 多元環とその準同型

命題 2.4.1 における性質をすべて満たす代数構造のことを多元環 (algebra) と呼ぶのであった. $M_n(\mathbb{R})$ はその代表例である. $M_n(\mathbb{R})$ と対応づけられる $\text{End}(\mathbb{R}^n)$ も多元環となることを確認しておこう.

【補足】多元環の定義に単位元の存在を仮定しない流儀もあるが, 本書では単位元の存在を仮定する. 命題 2.4.1 はもともと行列の性質を述べるものであったゆえ, 単位元を表す記号に E を用いている. 一般の多元環においては, 単位元の記号として I を用いることが多い.

ここでは \mathbb{R}^n に限らずに, 一般の線形空間 U 上の線形変換について論じよう.

例 21.4.1. $\text{End}(U)$ の各元 f, g に対して, これらの積 fg を合成 $f \circ g$ によって定める. すると, 既に定めていた和とスカラー倍および今の積演算は, $E = \text{id}_U$ を単位元として命題 2.4.1 における性質をすべて満たす (練習 21.4.2). すなわち, $\text{End}(U)$ は多元環となる.

以降では, 恒等写像 $\text{id}_U: U \rightarrow U$ を $\text{End}(U)$ における単位元とみなすときは, これを単位行列と区別するために I と書くことにしよう.

練習 21.4.2. $F, G, H \in \text{End}(U)$ および $r, s \in \mathbb{R}$ について次を確認せよ:

$$(13) (rF)(sG) = (rs)(FG), \quad (16) F(G+H) = FG + FH, \quad (17) (F+G)H = FH + GH.$$

【注意】積に関する結合律: (18) $(FG)H = F(GH)$ は命題 3.2.5 において, より一般的な立場から既に示している.

解答例: 積演算が写像の合成を意味していることに注意し, 各 $\mathbf{u} \in U$ を代入した値が一致することを確認すればよい. (16) のみ示そう.

$$\begin{aligned} F(G+H)(\mathbf{u}) &= F \circ (G+H)(\mathbf{u}) = F((G+H)(\mathbf{u})) = F(G(\mathbf{u}) + H(\mathbf{u})) \\ &= F(G(\mathbf{u})) + F(H(\mathbf{u})) = F \circ G(\mathbf{u}) + F \circ H(\mathbf{u}) = (F \circ G + F \circ H)(\mathbf{u}) = (FG + FH)(\mathbf{u}) \\ &\quad (\uparrow F \text{ の線形性を用いた}) \end{aligned}$$

□

定理 21.3.4 において $m = n$ とするとき, 命題 21.3.3(3) により, $\mathcal{T} : M_n(\mathbb{R}) \rightarrow \text{End}(\mathbb{R}^n)$ は積の演算についても整合的な対応となる. このような写像は多元環における準同型と呼ばれる:

定義 21.4.3. 二つの多元環 \mathcal{M} および \mathcal{L} の間の線形写像 $\tau : \mathcal{M} \rightarrow \mathcal{L}$ が次の二つの性質を満たすとき, τ を多元環準同型 (**algebra homomorphism**) あるいは単に準同型と呼ぶ³:

(i) 各 $A, B \in \mathcal{M}$ について, $\tau(AB) = \tau(A)\tau(B)$,

(ii) τ は \mathcal{M} の単位元を \mathcal{L} の単位元にうつす.

また, 多元環準同型 τ が全単射であるとき τ^{-1} も多元環準同型となり (この事実は各自確かめよ), このとき τ を同型 (**isomorphism**) と呼ぶ.

例 21.4.4. (1) 先ほどの $\mathcal{T} : M_n(\mathbb{R}) \rightarrow \text{End}(\mathbb{R}^n)$ は多元環準同型であり, かつ同型である. これは $M_m(\mathbb{R})$ と $\text{End}(\mathbb{R}^m)$ を多元環として同一視できることを意味する.

(2) 例 20.4.8 における $T_F : C(Y) \rightarrow C(X)$ は多元環準同型である (各自確かめよ).

21.5 線形変換と多項式

前節において $\text{End}(U)$ が多元環の構造を持つこと, すなわち, 線形変換たちの間で正方行列と類似の演算ができることを見た. ここで, 写像あるいは行列による演算が複雑化したときに有効な記法を導入しよう.

以下, 多元環 \mathcal{M} の元 $A \in \mathcal{M}$ について, A^n を A の n 個の積と定める. また, 計算に現れる式を簡略にするため A^0 を \mathcal{M} の単位元と定める. とくに線形変換 $F : U \rightarrow U$ について, $F^0 = I$ とする.

定義 21.5.1. 実数係数の n 次多項式 $\Phi(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$ および線形変換 $F : U \rightarrow U$ に対して, $\Phi(t)$ における変数 t を F に置き換え, また定数項 a_0 を $a_0 I$ に置き換えると次のような線形変換が得られる:

$$a_n F^n + a_{n-1} F^{n-1} + \cdots + a_1 F + a_0 I.$$

これを略して $\Phi(F)$ と書く. また正方行列 A に対しても同様に, $\Phi(A)$ を次の正方行列として定める:

$$a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 E.$$

【補足】 この記法はより一般に, 多元環の元に対して適用してよい.

上の定義はあくまで形式的に新たな写像 $\Phi(F)$ および行列 $\Phi(A)$ を与えているのであって, 多項式 $\Phi(t)$ に F や A を代入しているわけではない.

例 21.5.2. $D : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ を $D = \frac{d}{dx}$ と定めれば, $y \in C^\infty(\mathbb{R})$ に関する条件としての次の四つの式はすべて同値である.

³条件 (i) を満たす線形写像を多元環準同型と呼び, 更に (ii) も満たすものを単位的準同型 (**unital homomorphism**) と呼ぶのが一般的な定義であるが, 本書では単位的でない多元環準同型は扱わないことから上の定義を採用する.

- $y^{(k)}(x) + a_{k-1}y^{(k-1)}(x) + a_{k-2}y^{(k-2)}(x) + \cdots + a_1y^{(1)}(x) + a_0y^{(0)}(x) = 0,$
- $D^k(y) + a_{k-1}D^{k-1}(y) + a_{k-2}D^{k-2}(y) + \cdots + a_1D(y) + a_0I(y) = \mathbf{0},$
- $(D^k + a_{k-1}D^{k-1} + a_{k-2}D^{k-2} + \cdots + a_1D + a_0I)(y) = \mathbf{0},$
- $\Phi(D)(y) = \mathbf{0}$ (ただし, $\Phi(t) := t^n + a_{n-1}t^{n-1} + \cdots + a_1t + a_0$ とする).

最後の式に現れた $\Phi(t)$ は, 上の常微分方程式の特性多項式と呼ばれる. また, 式 $\Phi(D)(y) = \mathbf{0}$ を解析学では $\Phi(D)y = 0$ と略記するのが慣例となっている.

練習 21.5.3 (ケーリー-ハミルトンの定理). $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ について, $\Phi_A(t) = t^2 - (a+d)t + (ad-bc)$ とおくと $\Phi_A(A) = O$ となることを確かめよ.

【補足】 $\Phi_A(A) = O$ を移項すれば $A^2 = (a+d)A - (ad-bc)E$ である. この左辺よりも右辺のほうが計算が容易な場合は, A の冪を右辺を用いて計算するとよい. 一般の正方行列に関するケーリー-ハミルトンの定理は 30 章で論ずる.

多項式は, $\Phi(t) = a_nt^n + a_{n-1}t^{n-1} + \cdots + a_1t + a_0$ のように展開した形以外にも, 一部の項を括弧でくくったり, 因数分解したりと無数の表示を持つ. すると, 多項式を変形してから変数 t を線形変換 F に置き換えても問題がないかという疑念が浮かぶかもしれない. 例えば, $\Phi(t) = (t-\lambda)^n$ と因数分解される場合, 定義 21.5.1 による線形変換 $\Phi(F)$ (すなわち展開した式について t を F に置き換えたもの) と $(F-\lambda I)^n$ は等しいだろうか. これは $(t-\lambda)^n$ と $(F-\lambda I)^n$ をそれぞれ二項定理を用いて実際に展開することで確かめられるだろう. より一般に, 多項式がどのように括弧でくくられて表示されていても, その多項式の変数 t を F に置き換えた写像は定義 21.5.1 のそれと一致する. この事実を次の練習を通して確認しよう.

練習 21.5.4. 線形変換 $F: U \rightarrow U$ および多項式 $\Phi, \Phi_1, \Phi_2, \Psi, \Psi_k$ について次を示せ.

(1) 実数 λ について, $\Phi(t) = \lambda t^k \Psi(t) \implies \Phi(F) = \lambda F^k \Psi(F)$.

解答例: $\Psi(t) = \sum_{i=0}^n a_i t^i$ とおけば, $\Phi(t) = \sum_{i=0}^n a_i \lambda t^{i+k}$ であるから,

$$\begin{aligned} \Phi(F) &= \text{“多項式 } a_n \lambda t^{i+n} + \cdots + a_0 \lambda t^k \text{ の } t \text{ を } F \text{ に置き換えた写像”} \\ &= a_n \lambda F^{i+n} + \cdots + a_0 \lambda F^k = \sum_{i=0}^n \lambda F^k a_n F^i = \lambda F^k \sum_{i=0}^n a_n F^i = \lambda F^k \Psi(F). \end{aligned}$$

□

(2) $\Phi(t) = \sum_{k=1}^r \Psi_k(t) \implies \Phi(F) = \sum_{k=1}^r \Psi_k(F)$.

解答例: $\Psi_k(t) = \sum_{i=0}^n a_{k,i} t^i$ ($k = 1, \dots, r$) とおけば, 例 3.2.4 により

$$\Phi(t) = \sum_{k=1}^r \left(\sum_{i=0}^n a_{k,i} t^i \right) = \sum_{i=0}^n \left(\sum_{k=1}^r a_{k,i} t^i \right) = \sum_{i=0}^n \left(\sum_{k=1}^r a_{k,i} \right) t^i$$

となるから,

$$\begin{aligned} \Phi(F) &= \text{“多項式 } \left(\sum_{k=1}^r a_{k,n} \right) t^n + \cdots + \left(\sum_{k=1}^r a_{k,1} \right) t + \left(\sum_{k=1}^r a_{k,0} \right) \text{ の } t \text{ を } F \text{ に置き換えた写像”} \\ &= \left(\sum_{k=1}^r a_{k,n} \right) F^n + \cdots + \left(\sum_{k=1}^r a_{k,1} \right) F + \left(\sum_{k=1}^r a_{k,0} \right) I \\ &= \sum_{i=0}^n \left(\sum_{k=1}^r a_{k,i} \right) F^i = \sum_{i=0}^n \left(\sum_{k=1}^r a_{k,i} F^i \right) = \sum_{k=1}^r \left(\sum_{i=0}^n a_{k,i} F^i \right) = \sum_{k=1}^r \Psi_k(F). \end{aligned}$$

□

(3) $\Phi(t) = \Phi_1(t)\Phi_2(t) \implies \Phi(F) = \Phi_1(F)\Phi_2(F)$.

解答例: $\Phi_1(t) = \sum_{i=0}^n a_i t^i$ とおけば, $\Phi(t) = \sum_{i=0}^n a_i t^i \Phi_2(t)$ である. ここで $\Psi_i(t) = a_i t^i \Phi_2(t)$ ($i = 0, \dots, n$) とおけば, (1) より $\Psi_i(F) = a_i F^i \Phi_2(F)$ が成り立つ. また, $\Phi(t) = \sum_{i=0}^n \Psi_i(t)$ であるから (2) より $\Phi(F) = \sum_{i=0}^n \Psi_i(F)$ である. 以上より,

$$\Phi(F) = \sum_{i=0}^n \Psi_i(F) = \sum_{i=0}^n a_i F^i \Phi_2(F) = \left(\sum_{i=0}^n a_i F^i \right) \Phi_2(F) = \Phi_1(F)\Phi_2(F).$$

□

【補足】 上の練習は, $\text{End}(U)$ の元 F に関する等式であるが, 一般の多元環の元についても同様の事実が成り立つ. 証明も上と全く同じである. とくに $M_n(\mathbb{R})$ の元である正方行列 A についても成り立つ.

上の性質を有限回適用することで, 多項式 $\Phi(t)$ の表し方によらず, $\Phi(t)$ における t を F に置き換えた写像は一致することが分かる. とくに $\Phi(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \cdots (t - \lambda_k)^{n_k}$ と因数分解されるとき,

$$\Phi(F) = (F - \lambda_1 I)^{n_1} (F - \lambda_2 I)^{n_2} \cdots (F - \lambda_k I)^{n_k}$$

である.

行列の積演算がそうであったように, 多元環における積演算は一般には可換ではない. しかしながら, 次の二つの形の元は可換になる:

命題 21.5.5. $\Phi(t)$ および $\Psi(t)$ を多項式とする. 多元環 \mathcal{M} の各元 $F \in \mathcal{M}$ について $\Phi(F)$ と $\Psi(F)$ は可換である. すなわち $\Phi(F)\Psi(F) = \Psi(F)\Phi(F)$.

Proof. $\Theta(t) = \Phi(t)\Psi(t)$ とおく. すると練習 21.5.4(3) により $\Theta(F) = \Phi(F)\Psi(F)$ である. 一方, $\Theta(t) = \Psi(t)\Phi(t)$ であるから $\Theta(F) = \Psi(F)\Phi(F)$ でもある. 以上より, $\Phi(F)\Psi(F) = \Theta(F) = \Psi(F)\Phi(F)$. □

いまの命題を線形変換に対して適用すれば,

系 21.5.6. $F : U \rightarrow U$ を線形変換とし, $\Phi(t)$ と $\Psi(t)$ を多項式とすれば,

$$\text{各 } \mathbf{u} \in U \text{ について, } \Phi(F) \circ \Psi(F)(\mathbf{u}) = \Psi(F) \circ \Phi(F)(\mathbf{u}).$$

多項式を用いた表示と多元環準同型に関して, 次が成り立つ.

命題 21.5.7. $\mathcal{S} : \mathcal{M} \rightarrow \mathcal{L}$ を多元環準同型とすれば, 各多項式 $\Psi(t)$ および $F \in \mathcal{M}$ について

$$\mathcal{S}(\Psi(F)) = \Psi(\mathcal{S}(F)).$$

Proof. $\Psi(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$ とおく. また \mathcal{M} および \mathcal{L} の単位元をそれぞれ $I_{\mathcal{M}}, I_{\mathcal{L}}$ とおく. \mathcal{S} が多元環準同型であることから $\mathcal{S}(F^2) = \mathcal{S}(FF) = \mathcal{S}(F)\mathcal{S}(F) = \mathcal{S}(F)^2$ である. 同様に $\mathcal{S}(F^n) = \mathcal{S}(F)^n$ が得られ, したがって

$$\begin{aligned} \mathcal{S}(\Psi(A)) &= \mathcal{S}(a_n F^n + a_{n-1} F^{n-1} + \cdots + a_1 F + a_0 I_{\mathcal{M}}) \\ &= a_n \mathcal{S}(F^n) + a_{n-1} \mathcal{S}(F^{n-1}) + \cdots + a_1 \mathcal{S}(F) + a_0 \mathcal{S}(I_{\mathcal{M}}) \\ &= a_n \mathcal{S}(F)^n + a_{n-1} \mathcal{S}(F)^{n-1} + \cdots + a_1 \mathcal{S}(F) + a_0 I_{\mathcal{L}} = \Psi(\mathcal{S}(F)). \end{aligned}$$

□

例 21.5.8. $m = n$ として定理 21.3.4 で与えた $\mathcal{T} : M_n(\mathbb{R}) \rightarrow \text{End}(\mathbb{R}^n)$ について前命題を適用すれば, 多項式 $\Psi(t)$ および n 次正方行列 A について $\mathcal{T}(\Psi(A)) = \Psi(\mathcal{T}(A))$, すなわち $T_{\Psi(A)} = \Psi(T_A)$ である.

第22章 線形空間の次元

前章において、線形空間を類別する手段、すなわち線形同型なる概念を与えた。本章では線形空間が同型であるかどうかをはかる指標として次元とよばれる量を導入する。これにより、線形空間の完全な分類を得る。

22.1 次元の定義

二つの線形空間 U と V が同型でないとは、線形同型写像 $f: U \rightarrow V$ が存在しないということである。つまり、 U と V が同型でないことを示すには非存在証明を行わねばならない。一般的な経験則として、存在証明に比べて非存在証明は難しい。非存在証明が難しいのは、長時間探し続けたが見つからなかったというだけでは証明にならず、どんなに才能あふれる者が探しても見つからないことを示さねばならない点にある。この事情と関連して、多くの非存在証明は背理法によってなされている。命題 21.2.7 においてもそうであった。

さて、命題 21.2.7 における非存在証明の鍵となる概念が何であったか振り返ろう。それは \mathbb{R}^m と \mathbb{R}^n では座標軸の数、すなわち基底におけるベクトルの個数が異なることが本質をついていた。基底におけるベクトルの個数は線形空間たちが同型であるか否かを決定する重要な量であり、この量を次元と呼ぶ。次の命題を前提として次元の定義を得る：

命題 22.1.1. 線形空間 V の基底におけるベクトルの個数は、基底の取り方によらずに一定である。

Proof. 二つのベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_m$ および $\mathbf{v}_1, \dots, \mathbf{v}_n$ がそれぞれ V の基底であることを仮定し、 $m = n$ を示せばよい。基底 $\mathbf{u}_1, \dots, \mathbf{u}_m$ に関して例 21.2.5(1) を適用すれば V と \mathbb{R}^m は同型となる。また、基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ に対しても同様の考察を行い V と \mathbb{R}^n は同型となる。すなわち、 $V \simeq \mathbb{R}^m$ 、 $V \simeq \mathbb{R}^n$ であり、これに命題 21.2.3(推移律) を合わせて $\mathbb{R}^m \simeq \mathbb{R}^n$ を得る。命題 21.2.7 より $\mathbb{R}^m \simeq \mathbb{R}^n$ となるのは $m = n$ のときに限る。□

定義 22.1.2. 線形空間 V の有限個のベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n$ が V の基底となるとき、この基底の個数 n を V の次元 (**dimension**) と呼び $\dim V$ で表す。また、基底となる有限個のベクトルの組を持つ線形空間のことを有限次元 (**finite dimensional**) であるといい $\dim V < \infty$ と書く。そのような基底を持たない線形空間を無限次元 (**infinite dimensional**) であるといい $\dim V = \infty$ と書く。

仮に命題 22.1.1 が認められなければ、一つの線形空間 V が m 個の元からなる基底と n 個の元からなる基底を持つ可能性があり ($m \neq n$)、これでは V の次元を定められない。次元の定義に命題 22.1.1 が必要なのはこのためである。行列の階数や置換の符号の定義の際にも類似の議論を行ったのを覚えているだろうか。

命題 18.3.1 により、線形空間が有限個のベクトルで生成されることと有限次元であることは同値である。

例 22.1.3. (1) $\dim \mathbb{R}^n = n$, $\dim M_{m,n}(\mathbb{R}) = mn$, $\dim \mathbb{R}[x]_n = n + 1$.

(2) (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ について、例 20.2.4 および 20.3.2 において $\text{Im } T_A = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$, $\text{Ker } T_A = W_A$ となることを見た。また、これらの基底におけるベクトルの個数は命題 18.3.4 によって得られている。すなわち、 $\dim \text{Im } T_A = \dim \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle = \text{rank } A$, $\dim \text{Ker } T_A = \dim W_A = n - \text{rank } A$ である。この事実から一般の線形写像 f においても $\dim \text{Im } f$ のことを線形写像 f の階数 (**rank**) と呼び $\text{rank } f$ と書く。また、 $\dim \text{Ker } f$ にも特別な記号が与えられており、これを $\text{null } f$

と書き, f の退化次数 (nullity) と呼ぶ. あまり概念を増やして情報が錯綜しても困るゆえ, 本書ではこれらの語句の使用は控える.

(3) 例 18.4.1 より $\mathbb{R}[x]$ は有限個のベクトルで生成されることはない. したがって無限次元である.

(4) $\mathbb{R}^{\mathbb{N}}$ は無限次元である. この事実は次節で示す (例 22.3.5).

次元を比べるだけで線形空間どうしが同型かどうかを判定できる. とくに次の定理の (3) は, 線形同型写像 $f: U \rightarrow V$ の非存在性を示すための手段となる.

定理 22.1.4. 線形空間 U, V において次が成り立つ.

(1) U が有限次元であるとき, $U \simeq V \iff \dim U = \dim V$.

(2) $U \simeq V \implies \dim U = \dim V$.

(3) $\dim U \neq \dim V \implies U \not\simeq V$.

Proof. (1): (\Leftarrow) は例 21.2.5(1) より得られているゆえ (\Rightarrow) を示せばよい. $f: U \rightarrow V$ を同型とする. いま U は有限次元であると仮定していたから, 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が取れる (つまり $\dim U = n$). 命題 21.1.5 より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は V の基底であり, $\dim V = n$. 以上より $\dim U = \dim V$.

(2): U が有限次元の場合は (1) により示されている. また V が有限次元の場合は U と V の立場を入れ替えることで (1) に帰着できる. U, V いずれも無限次元の場合は $\dim U = \dim V = \infty$ である.

(3): (2) の対偶である. □

発展 (不変量)

二つの対象 X, Y が本質的に異なることを理解するには, それらに関する特別な量 $i(X), i(Y)$ が異なることを見ればよいのではないか. このような視点で与えられる量 i のことを一般に**不変量 (invariant)** という. 次元は線形空間に関する不変量である (定理 22.1.4(3)). 数学の多くの場面において, 対象としている数学的構造を保つ 1 対 1 対応の非存在性を示す際に不変量は鍵となる役割を担う. 例えば位相幾何学 (トポロジー) では, 点列の収束・発散を保つ 1 対 1 対応 (同相写像) があることを二つの空間 (図形) が同一視できることの基準としている. 線形空間における同一視を同型と呼んだのに対して, 位相幾何における空間の同一視を同相という. 同相は同値関係である. 位相幾何的な立場でも次元という量が定義され (ルベークの被覆次元), 二つの空間が同相ならばそれらの被覆次元は一致することが知られている. すなわち, 被覆次元は位相幾何における不変量であり, 言い換えれば被覆次元の異なる空間は同相ではない. とくに \mathbb{R}^n の被覆次元が n であることが分かっており, このことから $m \neq n$ ならば \mathbb{R}^m と \mathbb{R}^n は同相ではないことが示される.

ところで, 通常の不変量 i において, $i(X) = i(Y)$ だからといって, X と Y が同一視できるかどうかは分からない. 例えば曲面の被覆次元は当然すべて 2 であり, 被覆次元は曲面の分類には適さない. このことは, 状況に応じて臨機応変に様々な不変量を使い分ける必要があることを示唆している. 一方で, $i(X) = i(Y)$ から X, Y が同一視できることを導けるような不変量 i を**完全不変量**という. 例えば, 有限次元の線形空間において次元は完全不変量である (定理 22.1.4(1)). 完璧な分類を可能にすることが完全不変量の強みである. しかしながら完全不変量が一致することの意味を読み解くと, それは同一視の仕方を焼き直したものに過ぎないことが多く, このような場合は不変量によって対象への理解が深まったとは言えないだろう. 言い換えれば, いま考えている立場において不要な情報を捨てて有益な情報のみを取り出す不変量が望まれるのであり, このような不変量の一部の情報を捨てることから必然的に完全不変量にはなりえない. 結局, どんな量が必要とされるかを見極める力が求められている. こうした考え方は数学に限らず諸科学の分析において共通に認められるものである.

ちなみに, 曲面を分類する際の有効な不変量として, オイラー標数やベッチ数, ホモロジー群などが知られている. 最後に群が登場したように不変量を数に限る必要はなく, 様々な数学的対象をその候補に挙げることができる.

22.2 連立1次方程式の任意定数の個数

4.7節にて予告していた連立1次方程式の一般解の表示に現れる任意定数の個数について再考する. 方程式 $Ax = b$ の解全体の集合 W は, 解が存在するとすれば斉次方程式 $Ax = 0$ の解空間 W_A を平行移動した集合であった. したがって, 連立1次方程式の任意定数の個数について論ずるのであれば斉次方程式についてのみ考察すれば十分である. いま, 方程式 $Ax = 0$ の解空間 W_A が次のように表されているとする.

$$W_A = \{c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \cdots + c_k \mathbf{a}_k \mid c_1, \dots, c_k \in \mathbb{R}\} = \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle.$$

4.7節で問題提起した次の3点について答えよう.

- (i) 任意定数の個数に水増しがないかどうかをどうやって判定すればよいのか. 特に, 掃き出し法による一般解の表示において任意定数の水増しはないか.

回答: 水増しがないとは $\mathbf{a}_1, \dots, \mathbf{a}_k$ の中に不要なベクトルがないということであり, これは $\mathbf{a}_1, \dots, \mathbf{a}_k$ が線形独立であることに他ならない (命題 17.2.5). つまり, 水増しがないことを示すには $\mathbf{a}_1, \dots, \mathbf{a}_k$ の線形独立性を示せばよい. また, 掃き出し法による一般解の表示において $\mathbf{a}_1, \dots, \mathbf{a}_k$ が線形独立になることは, 命題 18.3.4(2) およびその証明において得られている. すなわち掃き出し法による一般解の表示において任意定数の水増しはない.

- (ii) 任意定数の水増しのない異なる二つの一般解を与えたときに, 二つ解の任意定数の個数は必ず一致するか.

回答: 命題 22.1.1 より W_A の基底におけるベクトルの個数は一定である. $\mathbf{a}_1, \dots, \mathbf{a}_k$ に水増しがないければこれは W_A の基底であり, その個数は $\dim W_A$ に一致する. すなわち, 任意定数に水増しのない一般解の表示において, 任意定数の個数は必ず $\dim W_A$ 個になる.

- (iii) 任意定数の定め方には, どれくらいの種類が考えられるのか.

回答: この質問は $\mathbf{a}_1, \dots, \mathbf{a}_k$ の選び方がどの程度あるかを問うているのであり, それは W_A の基底の選び方のぶんだけ任意性がある.

連立1次方程式の解法として掃き出し法を特別視しない立場において, 一般解の表示における任意定数の個数を定義しようと思えば次のようになる.

定義 22.2.1. (m, n) -行列 A による連立1次方程式 $Ax = b$ の解が存在するとき, その一般解の表示における任意定数の個数とは, \mathbb{R}^n の部分空間 W_A (斉次方程式の解空間) の次元のことであり, その数は $n - \text{rank } A$ に等しい.

22.3 線形独立な最大個数

次元は, 次で与える数で特徴づけられる (命題 22.3.3(2) を参照せよ). この概念を通して, 部分空間の次元や生成系との関係について考察しよう.

定義 22.3.1. 線形空間 V の部分集合 $A \subset V$ および非負整数 n が次の性質 (i) および (ii) を満たすとき, A における線形独立な最大個数を n と定める:

- (i) n 個のベクトルからなる線形独立な組 $\mathbf{a}_1, \dots, \mathbf{a}_n \in A$ が存在する.
(ii) いかなる $n+1$ 個の組 $\mathbf{a}_1, \dots, \mathbf{a}_{n+1} \in A$ も線形従属となる.

【補足】 練習 17.2.7(2) により, 上の条件 (ii) は次の条件 (iii) およびその対偶 (iv) と同値である:

- (iii) 各 $\mathbf{a}_1, \dots, \mathbf{a}_m \in A$ について, $m \geq n+1 \implies \mathbf{a}_1, \dots, \mathbf{a}_m$ は線形従属.
(iv) 各 $\mathbf{a}_1, \dots, \mathbf{a}_m \in A$ について, これらが線形独立 $\implies m \leq n$.

次の命題は、線形独立な最大個数をとる範囲を生成元のみを制限しても、あるいは全体まで広げてよいことを述べている。

命題 22.3.2. 線形空間 V および $A \subset V$ について次が成り立つ。

- (1) A における線形独立な最大個数が n であるとき、線形独立な組 $\mathbf{a}_1, \dots, \mathbf{a}_n \in A$ は $\langle A \rangle$ の基底である。
- (2) $\langle A \rangle$ における線形独立な最大個数が n であるとき、線形独立な組 $\mathbf{w}_1, \dots, \mathbf{w}_n \in \langle A \rangle$ は $\langle A \rangle$ の基底である。
- (3) A における線形独立な最大個数が n であるとき、 $\dim \langle A \rangle = n$ 。
- (4) $\langle A \rangle$ における線形独立な最大個数が n であるとき、 $\dim \langle A \rangle = n$ 。

Proof. (1): まず A の各元がいずれも $\mathbf{a}_1, \dots, \mathbf{a}_n$ の線形結合で書けることを示そう。仮に $\mathbf{a}_1, \dots, \mathbf{a}_n$ の線形結合で書けないベクトル $\mathbf{a} \in A$ があるとすれば $n+1$ 個のベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{a}$ は補題 18.2.5(1) により線形独立であり、これは線形独立な最大個数が n であることに反する。ゆえに $A \subset \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ である。 V の部分空間 $\langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ は線形結合で閉じていることから $\langle A \rangle \subset \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ であり、以上より $\mathbf{a}_1, \dots, \mathbf{a}_n$ は $\langle A \rangle$ を生成する。

(2): いま示した「 B における線形独立な最大個数が n であるとき、線形独立な組 $\mathbf{a}_1, \dots, \mathbf{a}_n \in B$ は $\langle B \rangle$ の基底である」における B として $\langle A \rangle$ を取れば、命題 18.4.5 より $\langle B \rangle = B$ であり、(2) を得る。

(3) と (4) は、それぞれ (1) と (2) および次元の定義から明らか。□

次の命題によれば、有限次元線形空間 V において、次元の数だけ集めたベクトルの組が V を生成することと線形独立であることは同値になる。

命題 22.3.3. 有限次元線形空間 V において次が成り立つ。

- (1) V において線形独立な最大個数が定まり、その数は $\dim V$ に等しい。
- (2) V における $\dim V$ 個の線形独立な組は V の基底である。
- (3) $\dim V$ 個のベクトルの組 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{\dim V}$ が V を生成するならば、これは V の基底である。

Proof. $\dim V = n$ とし、 $\mathbf{v}_1, \dots, \mathbf{v}_n$ を V の基底とする。

(1): V における $n+1$ 個のベクトルからなる組が線形従属になることを背理法によって示そう。仮に $\mathbf{u}_1, \dots, \mathbf{u}_{n+1} \in V$ が線形独立であるとすれば、命題 18.3.2 により、これらに更にいくつかのベクトルを加えることで V の基底とすることができる。すなわち、 $n+1$ 個以上の元からなる V の基底が存在することになり、これは $\dim V = n$ に反する。したがって、 $\mathbf{u}_1, \dots, \mathbf{u}_{n+1}$ は線形従属である。また、 n 個の組 $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形独立であり、 V における線形独立な最大個数は n である。

(2): $A := \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ について、いま示した (1) と命題 22.3.2(2) より得る。

(3): 集合 $A = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ における線形独立な最大個数を k とする。 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中から線形独立な組 $\mathbf{u}_{n_1}, \dots, \mathbf{u}_{n_k}$ を取れば、命題 22.3.2(1) より $\mathbf{u}_{n_1}, \dots, \mathbf{u}_{n_k}$ は $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ の基底でとなる。仮定より $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ であり、ゆえに k 個の組 $\mathbf{u}_{n_1}, \dots, \mathbf{u}_{n_k}$ は V の基底となる。つまり $\dim V = k$ であり、 $n = k$ となる。これは組 $\mathbf{u}_{n_1}, \dots, \mathbf{u}_{n_k}$ が A の元をすべて取りつくしていることを意味する。したがって $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形独立である。□

【補足】 上の (1) の証明において命題 18.3.2 を用いたが、これが証明に本質的に必要なわけではない。命題 18.3.2 を用いずとも $\mathbf{u}_1, \dots, \mathbf{u}_{n+1}$ の線形従属性は次のようにして示すことができる:

Proof. 仮にこれらが線形独立であるとしよう。命題 20.1.11 により、各 \mathbf{u}_i を \mathbf{u}_i 自身に写す線形写像 $f: \langle \mathbf{u}_1, \dots, \mathbf{u}_{n+1} \rangle \rightarrow V$ が存在し、これは練習 20.3.7 より単射である。 $\dim V = n$ より \mathbb{R}^n と V は同型であり、また $\langle \mathbf{u}_1, \dots, \mathbf{u}_{n+1} \rangle$ と \mathbb{R}^{n+1} は同型である。これらを結ぶ同型写像を $g: V \rightarrow \mathbb{R}^n$, $h: \mathbb{R}^{n+1} \rightarrow \langle \mathbf{u}_1, \dots, \mathbf{u}_{n+1} \rangle$ とすれば、単射線形写像 $g \circ f \circ h: \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ を得る。しかし、 \mathbb{R}^{n+1} から \mathbb{R}^n への線形単射が存在しないことは命題 21.2.7 の下に述べた通りである。ここに矛盾を得ることから、 $\mathbf{u}_1, \dots, \mathbf{u}_{n+1}$ は線形従属でなければならない。□

一方で、線形空間が無限次元であることは次のように言い換えることができる。

命題 22.3.4 (発展). 線形空間 V において次は同値である.

- (1) V は無限次元である.
- (2) V は線形独立な無限部分集合を含む.
- (3) V において、いくらでも多くのベクトルからなる線形独立な組を取ることができる.
(すなわち、任意の自然数 n について、線形独立な組 $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ が存在する.)

Proof. (1) \Rightarrow (2): 各 $k \in \mathbb{N}$ において組 $\mathbf{v}_1, \dots, \mathbf{v}_k$ が線形独立となるようにベクトルの無限列 \mathbf{v}_n を見つけたい. このような \mathbf{v}_n が帰納的に取れることを示そう. いま、 \mathbf{v}_k までが得られているとし、 $\mathbf{v}_1, \dots, \mathbf{v}_k$ は線形独立であるとする. V は有限次元でないゆえ、これらは V の基底ではない. したがって、 $\mathbf{v}_1, \dots, \mathbf{v}_k$ の線形結合で表せない元 $\mathbf{v}_{k+1} \in V$ が存在する. このとき、 $k+1$ 個の組 $\mathbf{v}_1, \dots, \mathbf{v}_{k+1}$ は補題 18.2.5(1) より線形独立である. こうして帰納的に選んだ \mathbf{v}_n を集めた無限集合 $\{\mathbf{v}_n \mid n \in \mathbb{N}\}$ は線形独立である.

(2) \Rightarrow (3): 明らか.

(3) \Rightarrow (1): 対偶を示そう. V が有限次元であるならば、命題 22.3.3(1) より線形独立な最大個数は $\dim V$ であり、 V の中から $\dim V + 1$ 個のベクトルからなる線形独立な組を取ることができない. とくに、 V の中からいくらでも多くのベクトルからなる線形独立な組を取ることができない. \square

例 22.3.5. 例 18.4.6 より数列空間 $\mathbb{R}^{\mathbb{N}}$ は線形独立な無限集合 $\{\mathbf{e}_n \mid n \in \mathbb{N}\}$ を持つゆえ無限次元である.

22.4 次元から分かること

次元を用いて部分空間の分類や、線形単射や線形全射の存在性について考察しよう.

命題 22.4.1. 有限次元線形空間 V の部分空間 W は有限次元であり、 $\dim W \leq \dim V$. また、等号成立は $W = V$ のときに限る.

Proof. $\dim V = n$ とおく. V における線形独立な最大個数は n であったから、 W における線形独立な最大個数は n を越えることはない. そこでこの数を m とし、線形独立な組 $\mathbf{w}_1, \dots, \mathbf{w}_m \in W$ を取ろう. $A = W$ について命題 22.3.2(2) を適用すれば $\mathbf{w}_1, \dots, \mathbf{w}_m$ は W の基底であり、 $\dim W = m \leq n = \dim V$. また、等号 $n = m$ が成立する場合、命題 22.3.3(2) より $\mathbf{w}_1, \dots, \mathbf{w}_m$ は V の基底となり、 $W = V$ を得る. \square

例 22.4.2. \mathbb{R}^3 の部分空間 W の次元は前命題により 3 以下であり、次のように分類される.

- $\dim W = 0$ なる部分空間は $\{\mathbf{0}\}$ のみである.
- $\dim W = 1$ なる部分空間の基底は一つのベクトル $\mathbf{u} \neq \mathbf{0}$ からなり、 $W = \{r\mathbf{u} \mid r \in \mathbb{R}\}$ は \mathbf{u} と $\mathbf{0}$ を通る直線に一致する.
- $\dim W = 2$ なる部分空間は、基底 $\mathbf{u}_1, \mathbf{u}_2 \in W$ を持つ. $W = \{r\mathbf{u}_1 + s\mathbf{u}_2 \mid r, s \in \mathbb{R}\}$ であり、これは $\mathbf{u}_1, \mathbf{u}_2, \mathbf{0}$ を通る平面に一致する.
- $\dim W = 3$ なる部分空間は前命題により $W = \mathbb{R}^3$ のみに限る.

練習 22.4.3 (発展). (1) t_1, \dots, t_{n+1} を相異なる実数とする. 任意の実数 r_1, \dots, r_{n+1} に対して、 $p(t_i) = r_i$ ($i = 1, \dots, n+1$) を満たす n 次多項式 $p \in \mathbb{R}[x]_n$ が存在することを示せ.

解答: $I := \{t_1, \dots, t_{n+1}\} \subset \mathbb{R}$ とおけば、命題 15.2.6 より $\mathbb{R}[x]_n \subset C(I)$ とみなせる. 写像 $p : I \rightarrow \mathbb{R}$ を $p(t_i) := r_i$ ($i = 1, \dots, n+1$) と定めれば $p \in C(I)$ である. p が $\mathbb{R}[x]_n$ の元でもあることを示せ

ば主張を得る. $C(I) = \mathbb{R}^I \simeq \mathbb{R}^{n+1}$ ゆえ $\dim C(I) = n+1$ である¹. また, $\dim \mathbb{R}[x]_n = n+1$ であるから命題 22.4.1 より $C(I) = \mathbb{R}[x]_n$ であり, $p \in \mathbb{R}[x]_n$. \square

- (2) 上の証明の脚注で触れた f_i を多項式で表せ. すなわち, $f_i(t_i) = 1, f_i(t_j) = 0 (j \neq i)$ を満たす n 次多項式を求めよ (この答えを通して, (1) で存在を示した多項式 p の具体的な式を書き下せることが分かる).

解答: $i = 1$ のみ記す.

$$f_1(x) := \frac{1}{(t_1 - t_2)(t_1 - t_3) \cdots (t_1 - t_{n+1})} (x - t_2)(x - t_3) \cdots (x - t_{n+1}).$$

線形単射や線形全射の存在は, 次元の大小によって特徴づけられる.

命題 22.4.4. 有限次元の線形空間 U, V について次が成り立つ.

- (1) $\dim U \leq \dim V$ ならば線形単射 $f: U \rightarrow V$ が存在する.
 (2) $\dim U \geq \dim V$ ならば線形全射 $f: U \rightarrow V$ が存在する.

Proof. (1) $\dim U = n$ とおき, $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とする. $\dim V \geq n$ より, n 個のベクトルからなる線形独立な組 $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ が取れる. このとき, 例 21.2.5(1) より $f(\mathbf{u}_i) = \mathbf{v}_i$ を満たす線形同型 $f: U \rightarrow \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ があり, とくに f は単射である. f の終域を V と解釈し, 線形単射 $f: U \rightarrow V$ を得る.

(2) $\dim V = m$ および $\dim U = m + k$ (ただし $k \geq 0$) とおく. また, $\mathbf{v}_1, \dots, \mathbf{v}_m$ を V の基底とし, $\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{u}_{m+1}, \dots, \mathbf{u}_{m+k}$ を U の基底とする. 命題 20.1.11 により

$$f(\mathbf{u}_i) = \mathbf{v}_i (i = 1, \dots, m), \quad f(\mathbf{u}_{m+j}) = \mathbf{0}_V (j = 1, \dots, k)$$

を満たす線形写像 $f: U \rightarrow V$ を取れば, $\text{Im } f = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_{m+k}) \rangle = \langle \mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{0}_V, \dots, \mathbf{0}_V \rangle = V$ ゆえ f は全射である. \square

命題 22.4.5. 有限次元線形空間の間の線形写像 $f: U \rightarrow V$ について次が成り立つ.

- (1) f が単射ならば $\dim U \leq \dim V$. (2) f が全射ならば $\dim U \geq \dim V$.

Proof. (1): U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$ とすれば命題 20.3.6(3) より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \in V$ は線形独立である. $\dim V$ は, V における線形独立な最大個数に等しいゆえ $\dim V \geq n = \dim U$.

(2): V の基底を $\mathbf{v}_1, \dots, \mathbf{v}_m$ とすれば, f の全射性より $f(\mathbf{u}_i) = \mathbf{v}_i$ を満たす $\mathbf{u}_i \in U (i = 1, \dots, m)$ が存在する. 命題 20.1.7(3) より $\mathbf{u}_1, \dots, \mathbf{u}_m$ は線形独立である. $\dim U$ は, U における線形独立な最大個数に等しいゆえ $\dim U \geq m = \dim V$. \square

上の命題の対偶をとると, 命題 21.2.7 の後に述べたことが得られる.

系 22.4.6. 有限次元線形空間 U および V について次が成り立つ.

- (1) $\dim U > \dim V$ ならば線形単射 $f: U \rightarrow V$ は存在しない.
 (2) $\dim U < \dim V$ ならば線形全射 $f: U \rightarrow V$ は存在しない.

単射性や全射性は次元からも判断できる.

練習 22.4.7. 有限次元線形空間 U, V の間の線形写像 $f: U \rightarrow V$ について, 次を示せ.

¹具体的に次のような基底があることから分かる: $f_i: I \rightarrow \mathbb{R}$ を $f_i(t_i) := 1, f_i(t_j) := 0$ (ただし $j \neq i$) と定めれば, f_1, \dots, f_{n+1} は $C(I)$ の基底である. 実際, $\sum_{i=1}^{n+1} a_i f_i(x) = \mathbf{0}$ とすれば両辺に t_i を代入することで $a_i = 0$ を得るゆえ線形独立である. また, 各 $g \in C(I)$ は $g(x) = \sum_{i=1}^{n+1} g(t_i) f_i(x)$ と書けるゆえ $C(I)$ を生成している.

(1) f は単射である $\iff \dim U = \dim \operatorname{Im} f$.

解答例. (\Rightarrow): $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とすれば, $\operatorname{Im} f = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle$ である. f の単射性より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は線形独立であり, したがってこれは $\operatorname{Im} f$ の基底である. ゆえに $\dim \operatorname{Im} f = n = \dim U$.

(\Leftarrow): $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ を $\operatorname{Im} f$ の基底とすれば, f の線形性より $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形独立である. 仮定より $\dim U = \dim \operatorname{Im} f = n$ である. これと命題 22.3.3(2) を合わせれば, $\mathbf{u}_1, \dots, \mathbf{u}_n$ が U の基底であることが分かる. したがって, 命題 21.1.6 より $f: U \rightarrow \operatorname{Im} f$ は全単射である. とくに f は単射である. \square

(2) f は全射である $\iff \dim V = \dim \operatorname{Im} f$.

解答例. (\Rightarrow): f が全射ならば $\operatorname{Im} f = V$ より $\dim \operatorname{Im} f = \dim V$ である.

(\Leftarrow): $\operatorname{Im} f \subset V$ ゆえ $W = \operatorname{Im} f$ について命題 22.4.1 を適用すれば $\operatorname{Im} f = V$. すなわち f は全射である. \square

有限集合 X 上の写像 $f: X \rightarrow X$ について f の単射性と全射性は同値であった (命題 19.5.1). この類似として, 有限次元線形空間について次が成り立つ:

命題 22.4.8. 有限次元線形空間 U と V の次元が等しいとき, 線形写像 $f: U \rightarrow V$ に関する次の 3 条件は同値である:

- (1) f は単射である, (2) f は全射である, (3) f は線形同型である.

Proof. (1) と (2) の同値性のみを示せば十分である.

(1) \Rightarrow (2): f を単射とすれば練習 22.4.7(1) より $\dim U = \dim \operatorname{Im} f$ である. 仮定より $\dim U = \dim V$ ゆえ $\dim V = \dim \operatorname{Im} f$ であり, したがって練習 22.4.7(2) より f は全射である.

(2) \Rightarrow (1): f を全射とすれば練習 22.4.7(2) より $\dim V = \dim \operatorname{Im} f$ である. 仮定より $\dim U = \dim V$ ゆえ $\dim U = \dim \operatorname{Im} f$ であり, したがって練習 22.4.7(1) より f は単射である. \square

U が無限次元の場合, 上の (1) と (2) の同値性は成り立たない (例 26.1.5).

命題 22.4.9 (発展). 有限次元線形空間における線形写像 $T: U \rightarrow V$ について次が成り立つ.

(1) T は単射である $\iff S \circ T = \operatorname{id}_U$ を満たす線形写像 $S: V \rightarrow U$ が存在する.

(2) T は全射である $\iff T \circ S = \operatorname{id}_V$ を満たす線形写像 $S: V \rightarrow U$ が存在する.

Proof. (1): (\Rightarrow): $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とし, $\mathbf{v}_i := T(\mathbf{u}_i)$ ($i = 1, \dots, n$) とおけば, T の単射性より $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形独立な組である (命題 20.3.6). そこで, これらにいくつかのベクトルを加えて V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ を取る. このとき,

$$S(\mathbf{v}_i) = \begin{cases} \mathbf{u}_i & i \leq n \text{ のとき,} \\ \mathbf{0} & i > n \text{ のとき,} \end{cases}$$

を満たす線形写像 $S: V \rightarrow U$ が存在し (命題 20.1.11), これが求めるものである. 実際, 各 $i = 1, \dots, n$ について $S(T(\mathbf{u}_i)) = S(\mathbf{v}_i) = \mathbf{u}_i$ であり, 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の行き先は $S \circ T$ と id_U において等しい. ゆえに $S \circ T = \operatorname{id}_U$ である (命題 20.1.10(2)).

(\Leftarrow): $S \circ T = \operatorname{id}_U$ の単射性および命題 19.4.2 による.

(2): (\Rightarrow): $\mathbf{v}_1, \dots, \mathbf{v}_m$ を V の基底とすれば, 各 $i = 1, \dots, m$ について T の全射性より $T(\mathbf{u}_i) = \mathbf{v}_i$ を満たす $\mathbf{u}_i \in U$ が存在する. このとき $S(\mathbf{v}_i) = \mathbf{u}_i$ ($i = 1, \dots, m$) を満たす線形写像 $S: V \rightarrow U$ が存在し (命題 20.1.11), これが求めるものである. $T \circ S = \operatorname{id}_V$ は上と同様にして示せるゆえ略す.

(\Leftarrow): $T \circ S = \operatorname{id}_V$ の全射性および命題 19.4.2 による. \square

系 22.4.10. (m, n) -行列 A について次が成り立つ.

(1) $\text{rank } A = n \iff BA = E_n$ を満たす (n, m) -行列 B が存在する.

(2) $\text{rank } A = m \iff AC = E_m$ を満たす (n, m) -行列 B が存在する.

Proof. $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ について $\dim \text{Im } T_A = \text{rank } A$ に注意する (例 22.1.3(2)).

(1): (\Rightarrow): 仮定および練習 22.4.7(1) より T_A は単射であり, 前命題 (1) より $S \circ T_A = \text{id}_{\mathbb{R}^n}$ を満たす線形写像 $S: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が存在する. 定理 21.3.4 で与えた同型 $\mathcal{T}: M_{n,m}(\mathbb{R}) \rightarrow \text{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ を用いて $B := \mathcal{T}^{-1}(S)$ とすれば, これが求めるものである. 実際, $T_B = \mathcal{T}(B) = \mathcal{T}(\mathcal{T}^{-1}(S)) = S$ ゆえ, $T_{BA} = T_B \circ T_A = S \circ T_A = \text{id}_{\mathbb{R}^n}$. ゆえに BA は単位行列である.

(\Leftarrow): $T = T_A, S = T_B$ として前命題 (1) を適用すれば, T_A の単射性を得る. 練習 22.4.7(1) より $\dim \text{Im } T_A = \dim \mathbb{R}^n = n$.

(2): (\Rightarrow): 仮定および練習 22.4.7(2) より T_A は全射であり, 前命題 (2) より $T_A \circ S = \text{id}_{\mathbb{R}^m}$ を満たす線形写像 $S: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が存在する. 上の同型 \mathcal{T} を用いて $C := \mathcal{T}^{-1}(S)$ とすれば, これが求めるものであることは上と同様にして分かる.

(\Leftarrow): $T = T_A, S = T_C$ として前命題 (2) を適用すれば, T_A の全射性を得る. 練習 22.4.7(2) より $\dim \text{Im } T_A = \dim \mathbb{R}^m = m$. □

上の命題における B を A の左逆行列, C を A の右逆行列と呼ぶ. A が正方行列でない場合, その右逆行列や左逆行列が存在するならば, それらは無数にある. 上の行列 B, C の具体的な求め方は次の通りである. これは標準基底について, 命題 22.4.9 の証明の通りに構成したものである.

(1) A の列の数 n と階数が等しいから A の各列は線形独立であり, また $m \geq n$ である. ここで $k := m - n \geq 0, A = [v_1, \dots, v_n]$ とおく. \mathbb{R}^m の標準基底を e_1, \dots, e_m とし, $[A|E_m] = [v_1, \dots, v_n | e_1, \dots, e_m]$ の簡約化を $[e_1, \dots, e_n | b_1, \dots, b_m]$ とするとき, (m, m) -行列 $Q := [b_1, \dots, b_m]$ の $n+1$ 行目以降を消し去った (n, m) -行列 B が求めるものである. その理由は次の通り:

上の簡約化を言い換えれば, m 次可逆行列 X を用いて $X[A|E_m] = [e_1, \dots, e_n | b_1, \dots, b_m]$ と表せる. このとき $XE_m = [b_1, \dots, b_m] = Q$, つまり $X = Q$ である. また $QA = XA = [e_1, \dots, e_n]$ より, $T_Q: \mathbb{R}^m \rightarrow \mathbb{R}^n$ は v_1, \dots, v_n を e_1, \dots, e_n に写す. さて $e_1, \dots, e_n \in \mathbb{R}^m$ を \mathbb{R}^n の標準基底に写す線形写像の一つに, (n, m) -行列 $P := [E_n | O_{n,k}]$ による $T_P: \mathbb{R}^m \rightarrow \mathbb{R}^n$ がある. このとき, $B := PQ$ とすれば,

$$BA = PQ[v_1, \dots, v_n] = [PQv_1, \dots, PQv_n] = [Pe_1, \dots, Pe_n] = E_n.$$

この B が Q の $n+1$ 行目以降を消し去った行列に一致することは, 行列の積の各成分を計算すればすぐに分かる. 実際, $Q = \begin{bmatrix} Q_{n,m} \\ Q_{k,m} \end{bmatrix}$ とおくと,

$$PQ = \left[E_n \mid O_{n,k} \right] \begin{bmatrix} Q_{n,m} \\ Q_{k,m} \end{bmatrix} = E_n Q_{n,m} + O_{n,k} Q_{k,m} = Q_{n,m}.$$

【補足】 b_1, \dots, b_m のうち, $\langle e_1, \dots, e_n \rangle \subset \mathbb{R}^m$ に属さないものを順にとり, これを b_{n_1}, \dots, b_{n_k} とすれば, $[e_1, \dots, e_n | b_1, \dots, b_m]$ が簡約行列であることから, $b_{n_j} = e_{n+j}$ ($j = 1, \dots, k$) である. このとき, $X[v_1, \dots, v_n | e_{n_1}, \dots, e_{n_k}] = [e_1, \dots, e_n | b_{n_1}, \dots, b_{n_k}] = E_m$ であるから, 線形変換 $T_X = T_Q$ は基底 $v_1, \dots, v_n, e_{n_1}, \dots, e_{n_k}$ を標準基底 e_1, \dots, e_m に写す. さらに T_P は, 標準基底を $e_1, \dots, e_n, \mathbf{0}, \dots, \mathbf{0} \in \mathbb{R}^n$ に写す. したがって, $S = T_P \circ T_Q = T_B$ は, 命題 22.4.9 の証明で与えた S と同じものである.

(2) **【方法 1】** 各 $i = 1, \dots, m$ について連立 1 次方程式 $Ax = e_i$ の解の一つを取り, これを $u_i \in \mathbb{R}^n$ とする (仮定より解は存在する). このとき, $C := [u_1, \dots, u_m]$ が求める行列である. ここで, 連立 1 次方程式を n 回解く必要があり, これは $[A|e_1|e_2|\dots|e_n]$ を簡約化することで同時に計算できる. また, 各方程式の解は一つ取ればよいので, 任意定数が 0 のものを取ると計算が楽になる.

【方法 2】 実は $\text{rank } A = \text{rank } {}^tA$ であることが知られている (系 27.4.3). そこで, (n, m) -行列 tA の左逆行列 B を (1) の方法で取れば, $C := {}^tB$ は A の右逆行列である. 実際, $B{}^tA = E_m$ の両辺の転置を取れば, $A{}^tB = E_m$ を得る.

22.5 無限次元の空間も含めた一般論 (発展)

これまで無限集合の線形独立性と線形写像との関係については控えていたが、ここでまとめておこう。有限個のベクトルからなる組の場合における証明を参考にするこゝで、次の命題群を証明することができる。これらは余力のある読者への演習問題として残すこととし、一般の読者には軽く読み流して次章に進むことを勧める。

命題 22.5.1 (命題 20.1.7(1) 参照). $f: U \rightarrow V$ を線形写像とし, $A \subset U$ とする. このとき $x \in \langle A \rangle$ ならば $f(x) \in \langle f(A) \rangle$ である.

命題 22.5.2 (命題 20.1.11 参照). $A \subset U$ を線形空間 U の基底とし, あらかじめ写像 $f: A \rightarrow V$ を与えておく. このとき, f の拡張である線形写像 $\tilde{f}: U \rightarrow V$ が唯一つ存在する.

命題 22.5.3 (命題 20.2.3 参照). 線形写像 $f: U \rightarrow V$ および $A \subset U$ において, $f(\langle A \rangle) = \langle f(A) \rangle$. とくに A が U を生成するとき, $f(A)$ は $\text{Im } f$ を生成する.

命題 22.5.4 (命題 20.3.6 参照). $f: U \rightarrow V$ を線形写像とし, $A \subset U$ とする. f が単射であるとき次が成り立つ:

- (1) A は線形独立である $\iff f(A)$ は線形独立である.
- (2) $f(A)$ は V の基底である $\implies A$ は U の基底である.

命題 22.5.5 (命題 21.1.5 参照). $f: U \rightarrow V$ が線形同型であるとき, $A \subset U$ について次が成り立つ.

$$A \text{ は } U \text{ の基底である} \iff f(A) \text{ は } V \text{ の基底である.}$$

命題 22.5.6 (命題 21.1.6 参照). U の基底を $A \subset U$, $f: U \rightarrow V$ を線形写像とすると, 次が成り立つ.

$$f \text{ は線形同型である} \iff f(A) \text{ は } V \text{ の基底である.}$$

次は命題 22.1.1 に対応する主張である. この定理の証明は本書の枠を越えており, 証明には無限集合の濃度演算について学ぶ必要がある.

定理 22.5.7. 線形空間 V について, $A, B \subset V$ を V の基底とすれば, A と B は対等である. すなわち, 基底におけるベクトルの個数 (濃度) は, 基底の取り方によらずに一定である.

上の事実から, 有限次元でない線形空間の次元を一括して ∞ と書くのではなく, 基底の元の個数 (濃度) によって細かく分類する案が考えられる. この案による次元は, 無限次元線形空間を含めた枠組みにおける完全不変量となる:

定理 22.5.8 (定理 22.1.4 参照). 線形空間 U, V において次は同値である:

- (1) $U \simeq V$.
- (2) 対等になるような U の基底と V の基底が存在する.
- (3) いかなる U の基底 A および V の基底 B についても, A と B は対等である.

Proof. (1) \implies (2): $U \simeq V$ であると仮定し, $f: U \rightarrow V$ を線形同型とせよ. A を U の基底とすれば命題 22.5.6 により $f(A)$ は V の基底であり, また, f の単射性より A と $f(A)$ は対等である.

(2) \implies (1): 逆に, U の基底 A と, V の基底 B が対等であるとすれば, 全単射 $f: A \rightarrow B$ が存在する. このとき命題 22.5.2 より, f の拡張である線形写像 $\tilde{f}: U \rightarrow V$ が唯一つ存在する. \tilde{f} が線形同型であることは命題 22.5.6 より得る.

(2) \implies (3): 定理 22.5.7 より得られる.

(3) \implies (2): 明らか. □

次は 22.4 節で述べた命題に対応する主張である。これらの証明も余力ある読者への演習として残そう。

命題 22.5.9 (命題 22.4.4 参照). 線形空間 U, V の基底をそれぞれ A, B とすれば次が成り立つ。

- (1) $f: A \rightarrow B$ が単射ならば, その拡張である線形単射 $\tilde{f}: U \rightarrow V$ が存在する。
- (2) $f: A \rightarrow B$ が全射ならば, その拡張である線形全射 $\tilde{f}: U \rightarrow V$ が存在する。

命題 22.5.10 (命題 22.4.5 参照). 線形空間 U, V の基底をそれぞれ A, B とする。

- (1) 線形単射 $f: U \rightarrow V$ が存在するならば, 単射 $p: A \rightarrow B$ が存在する。
- (2) 線形全射 $f: U \rightarrow V$ が存在するならば, 全射 $p: A \rightarrow B$ が存在する。

発展 ($\mathbb{R}[x]$ と $\mathbb{R}^{\mathbb{N}}$ は線形同型か)

本章において, 無限次元空間の例として $\mathbb{R}[x]$ と $\mathbb{R}^{\mathbb{N}}$ を挙げたが, これらが線形同型か否かについては言及しなかった。結論を先に述べればこれらは同型ではない。この事実を示すには, 定理 22.5.8 を念頭に, $\mathbb{R}[x]$ の基底と $\mathbb{R}^{\mathbb{N}}$ の基底が対等でないことを示す戦略が考えられよう。例 18.4.4 で述べたように $\mathbb{R}[x]$ の基底 $\{x^n \mid n \in \mathbb{Z}_{\geq 0}\}$ は非負整数と 1 対 1 の対応がつき, とくに \mathbb{N} と対等である。したがって, $\mathbb{R}^{\mathbb{N}}$ の基底が \mathbb{N} と対等でないことをいかに導くかがこの問題の鍵となる。 $\mathbb{R}^{\mathbb{N}}$ に基底が存在することは定理 18.4.7 によって保証されているものの, その具体像が不明であることが問題を難しくしており, 解決の糸口は極限操作を導入することにある。

$\mathbb{R}^{\mathbb{N}}$ のベクトルの間に距離を定めれば, ベクトルの列の収束発散が定まり, これを用いて $\mathbb{R}^{\mathbb{N}}$ 上の写像の連続性を定義できる。距離には様々な定め方があるものの, 線形空間の構造と相性のよい距離ということであれば和やスカラー倍の演算を写像とみなしたときにこれらが連続写像となるような距離を導入することが望ましい。このような距離が入った線形空間を線形距離空間という。線形距離空間の理論を進めると, \mathbb{N} と対等な基底をもつ線形距離空間は完備ではないことが示される。ここで, コーシー点列が必ず収束するような空間を完備であるという (微積分学でも \mathbb{R}^n の完備性を学ぼう)。一方, $\mathbb{R}^{\mathbb{N}}$ を完備な線形距離空間にできることが知られており, 以上の事実を総合すると, $\mathbb{R}^{\mathbb{N}}$ の基底は \mathbb{N} と対等でないことが分かる。

このように, 無限次元の線形空間の理論においては, 極限概念を導入することで対象を解析する手段を広げている。線形空間への極限概念の導入の仕方も一通りではなく, その一端として大学初年級の線形代数学においても内積空間 (計量ベクトル空間) やノルム空間を学ぶことになる。

第23章 次元公式と商空間

連立1次方程式と斉次形方程式の解の関係、および線形常微分方程式とその斉次方程式の解の関係には類似性が認められていた(6.4節)。これら類似する性質が線形写像に関する命題として一般的な立場から証明できることを本章で述べる。また、この性質を通して線形写像の次元公式を示す。次元公式とは、線形写像による空間の分解を次元の視点から述べた式のことである。

一方、線形空間の分解それ自体を記述するための概念として、商空間を定める。これは、次のような要請に応じて導入されるものである：ある線形空間において、いくつかのベクトル方向があまり重要ではないと判断されたとしよう。このとき、これらの方向を捨象した空間概念をいかに与えればよいか。この要請に満足する空間が、不要と思われる方向と平行な直線を集めた集合(つまり直線を要素とする集合)、あるいは、与えられた平面と平行な平面を集めた集合(つまり平面を要素とする集合)のような形で実現されることを本章の後半で見る。

ところで、本章および次章において、 $u-v$ と表記すればよいところをわざわざ読みにくい表記で $-v+u$ と書いた箇所がいくつかある。これは、可換性を満たさない代数構造においても同様の主張が成り立ち、その際の証明を見越したことによるものである。後に群論を学ぶ際の助けになるだろう。

23.1 空間の平行移動

連立1次方程式の解の集合の表示において集合の平行移動について言及していた。ここで改めて、平行移動の正確な定義を与える。

定義 23.1.1. 線形空間 U の部分集合 $A \subset U$ およびベクトル $u \in U$ に対して、 A を u 方向に平行移動した集合を $A+u$ あるいは $u+A$ とかく(図 23.1)。すなわち、

$$A+u := \{a+u \mid a \in A\}, \quad u+A := \{u+a \mid a \in A\}.$$

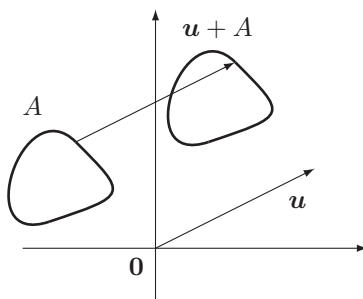


図 23.1: 図形 A の u 方向への平行移動

明らかに $A+u = u+A$ であり、この事実は和の演算の可換性 ($a+u = u+a$) に由来するものである。以降では $u+A$ という表記を主に用いることにする¹。

命題 23.1.2. A と $u+A$ は対等である。すなわち、これらの元の個数は等しい。

¹実は、 $A+u$ のほうを採用していれば、冒頭で述べた $-v+u$ という表記は必要でなくなる。しかしながら、ここでは慣例に従い $u+A$ と書くことにしたい。

Proof. 写像 $f : A \rightarrow U$ を $f(\mathbf{a}) := \mathbf{u} + \mathbf{a}$ で定めれば, $f(A) = \mathbf{u} + A$ が成り立つ. f の単射性は明らかであり, したがって A と $f(A)$ は対等である. \square

本論では, 上の集合 A として U の部分空間を主に考える.

練習 23.1.3. 部分空間 $W \subset U$ および $\mathbf{u} \in W$ において, $\mathbf{u} + W = W$ を示せ.

解答例: W が和の演算について閉じていることから, $\mathbf{u} + W \subset W$ を得る. また, 各 $\mathbf{v} \in W$ に対して $\mathbf{w} := -\mathbf{u} + \mathbf{v}$ とおけば $\mathbf{w} \in W$ であり, $\mathbf{v} = \mathbf{u} + \mathbf{w} \in \mathbf{u} + W$. つまり $W \subset \mathbf{u} + W$ である. \square

次に述べる命題は, 6.4 節で論じたことを線形写像の言葉で統一的に述べなおした主張に相当している (例 23.1.5). したがって, 読者も直ちに証明方針を予想できることと思う.

命題 23.1.4. $f : U \rightarrow V$ を線形写像とし, $\mathbf{b} \in \text{Im } f$ とする. 更に $f(\mathbf{a}) = \mathbf{b}$ を満たす $\mathbf{a} \in U$ を一つ取って固定しよう (すなわち $\mathbf{a} \in f^{-1}(\mathbf{b})$). このとき次が成り立つ:

- (1) 任意の $\mathbf{z} \in \text{Ker } f$ に対し, $\mathbf{a} + \mathbf{z} \in f^{-1}(\mathbf{b})$ である.
- (2) 任意の $\mathbf{y} \in f^{-1}(\mathbf{b})$ は, ある $\mathbf{z} \in \text{Ker } f$ を用いて $\mathbf{y} = \mathbf{a} + \mathbf{z}$ と表せる.
- (3) $f^{-1}(\mathbf{b}) = \mathbf{a} + \text{Ker } f$.

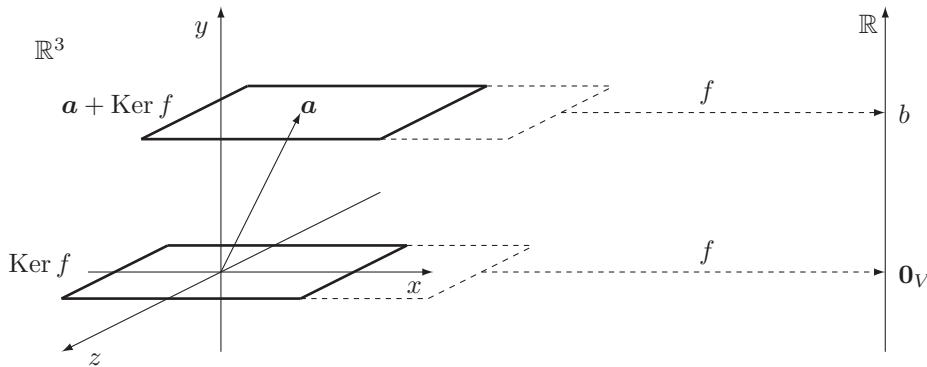


図 23.2: 線形写像の核とその平行移動

図 23.2 は線形写像 $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ ($f(x, y, z) := y$) について命題 23.1.4(3) を模式的に説明したものである. $\text{Ker } f$ は x - z 平面に一致する. また, $\mathbf{a} := (x_1, b, z_1)$ とすれば $f(\mathbf{a}) = b$ である (つまり $\mathbf{a} \in f^{-1}(b)$). $\text{Ker } f$ を \mathbf{a} 方向にずらしたものが $\mathbf{a} + \text{Ker } f$ であり, 図では平行四辺形が右斜め上にずれたように見えているが, 実際には x - z 平面と平行な方向に無限に広がる平面である. つまり, $\text{Ker } f$ を真上に持ち上げたもの ($(0, b, 0)$ 方向に平行移動させた平面) と一致する. 命題 23.1.4(3) によれば $f^{-1}(b) = \mathbf{a} + \text{Ker } f$ であり, $\mathbf{a} + \text{Ker } f$ に属する各元を f に代入すれば, その値はすべて b となる.

命題 23.1.4 の証明. (1) : $\mathbf{z} \in \text{Ker } f$ とすると, $f(\mathbf{a} + \mathbf{z}) = f(\mathbf{a}) + f(\mathbf{z}) = b + \mathbf{0} = b$. ゆえに $\mathbf{a} + \mathbf{z} \in f^{-1}(b)$ である.

(2) : $\mathbf{y} \in f^{-1}(b)$ とする. $\mathbf{z} := -\mathbf{a} + \mathbf{y}$ とおこう. このとき $f(\mathbf{z}) = f(-\mathbf{a} + \mathbf{y}) = -f(\mathbf{a}) + f(\mathbf{y}) = -b + b = \mathbf{0}$ より $\mathbf{z} \in \text{Ker } f$ である. また, \mathbf{z} の定め方から $\mathbf{y} = \mathbf{a} + \mathbf{z}$ であり, 我々は主張を得た.

(3) : 両方の包含関係 $f^{-1}(b) \subset \mathbf{a} + \text{Ker } f$ および $f^{-1}(b) \supset \mathbf{a} + \text{Ker } f$ を示せばよい. しかし, これらは (2) および (1) の主張をそれぞれ言い換えたものに過ぎない. \square

例 23.1.5. 次の線形写像および固定されたベクトル \mathbf{b} について上の命題を適用することで, 6.4 節における三つの命題がそれぞれ得られる.

- 命題 6.4.1: $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{b} \in \mathbb{R}^m$.
- 命題 6.4.2: $\alpha \frac{d^2}{dx^2} + \beta \frac{d}{dx} + \gamma I : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R}), \mathbf{b} = b(x)$.

- 命題 6.4.3: $S - rI : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ (ただし S はシフト作用素), $\mathbf{b} = (b, b, b, \dots)$.

【備考】 $(S - rI)(x_1, x_2, x_3, \dots) = (x_2 - rx_1, x_3 - rx_2, x_4 - rx_3, \dots)$ である.

命題 23.1.2 および 23.1.4(3) より, $f^{-1}(\mathbf{b}) \neq \emptyset$ であるとき $f^{-1}(\mathbf{b})$ と $\text{Ker } f$ は対等である. とくに, 線形写像の空でない逆像は互いに元の個数が等しい²:

系 23.1.6. 線形写像 $f : U \rightarrow V$ および各 $\mathbf{b}_1, \mathbf{b}_2 \in \text{Im } f$ について, $f^{-1}(\mathbf{b}_1)$ と $f^{-1}(\mathbf{b}_2)$ は対等である.

Proof. $f^{-1}(\mathbf{b}_1)$ と $f^{-1}(\mathbf{b}_2)$ はそれぞれ $\text{Ker } f$ と対等であり, したがって $f^{-1}(\mathbf{b}_1)$ と $f^{-1}(\mathbf{b}_2)$ も対等である. \square

23.2 線形写像の次元公式

有限次元の線形空間を定義域とする線形写像の次元公式を述べる. 像と核がともに有限次元であることを一応ながら確認しておく.

命題 23.2.1. 有限次元線形空間を定義域とする線形写像の像と核は有限次元である.

Proof. この写像の像は命題 20.2.3 より有限個のベクトルによって生成される. ゆえに有限次元である. また, この写像の核は有限次元線形空間の部分空間であるから, 命題 22.4.1 より有限次元である. \square

定理 23.2.2. 有限次元線形空間 U を定義域とする線形写像 $f : U \rightarrow V$ において, $\mathbf{z}_1, \dots, \mathbf{z}_n$ が $\text{Ker } f$ の基底であり, $\mathbf{u}_1, \dots, \mathbf{u}_m \in U$ について $f(\mathbf{u}_1), \dots, f(\mathbf{u}_m)$ が $\text{Im } f$ の基底であるならば, $m + n$ 個の組 $\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{z}_1, \dots, \mathbf{z}_n$ は U の基底である.

Proof. まず線形独立性を示そう. そこで線形関係 $\sum_{i=1}^m r_i \mathbf{u}_i + \sum_{j=1}^n s_j \mathbf{z}_j = \mathbf{0}_U$ を仮定する. これらに f をほどこすと

$$\begin{aligned} \sum_{i=1}^m r_i f(\mathbf{u}_i) + \sum_{j=1}^n s_j f(\mathbf{z}_j) &= f(\mathbf{0}_U) \\ \sum_{i=1}^m r_i f(\mathbf{u}_i) + \sum_{j=1}^n s_j \mathbf{0}_V &= \mathbf{0}_V \\ \sum_{i=1}^m r_i f(\mathbf{u}_i) &= \mathbf{0}_V. \end{aligned}$$

$f(\mathbf{u}_1), \dots, f(\mathbf{u}_m)$ の線形独立性より $r_1 = \dots = r_m = 0$ を得る. つまり $\sum_{j=1}^n s_j \mathbf{z}_j = \mathbf{0}_U$ であり, $\mathbf{z}_1, \dots, \mathbf{z}_n$ の線形独立性より $s_1 = \dots = s_n = 0$ を得る. 以上より $\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{z}_1, \dots, \mathbf{z}_n$ は線形独立である.

次に各 $\mathbf{y} \in U$ が $\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{z}_1, \dots, \mathbf{z}_n$ の線形結合で書けることを示そう. $\mathbf{b} := f(\mathbf{y})$ とおくと $\mathbf{b} \in \text{Im } f$ ゆえ $\mathbf{b} = \sum_{i=1}^m r_i f(\mathbf{u}_i)$ と書ける. このとき $\mathbf{a} = \sum_{i=1}^m r_i \mathbf{u}_i$ とおくと, $f(\mathbf{a}) = f(\sum_{i=1}^m r_i \mathbf{u}_i) = \sum_{i=1}^m r_i f(\mathbf{u}_i) = \mathbf{b}$ より $\mathbf{a} \in f^{-1}(\mathbf{b})$ である. この \mathbf{a}, \mathbf{b} に対して命題 23.1.4(2) を適用すれば, $\mathbf{y} \in f^{-1}(\mathbf{b})$ は $\mathbf{z} \in \text{Ker } f$ をもちいて $\mathbf{y} = \mathbf{a} + \mathbf{z}$ と書ける. $\mathbf{z}_1, \dots, \mathbf{z}_n$ は $\text{Ker } f$ の基底であったから $\mathbf{z} = \sum_{j=1}^n s_j \mathbf{z}_j$ と書け, $\mathbf{y} = \mathbf{a} + \mathbf{z} = \sum_{i=1}^m r_i \mathbf{u}_i + \sum_{j=1}^n s_j \mathbf{z}_j$ を得る. \square

いまの定理より直ちに次の公式を得る.

定理 23.2.3 (次元公式). 有限次元線形空間 U を定義域とする線形写像 $f : U \rightarrow V$ において,

$$\dim U = \dim \text{Ker } f + \dim \text{Im } f.$$

² 「元の個数」といっても $\text{Ker } f$ は線形空間ゆえ, これらは一点集合でなければ無限集合である. しかし, より一般の代数構造 (群の準同型) においても同様の主張が成り立ち, その場合は n 点集合として元の個数が等しくなることがある.

上の公式によれば、 U および $\text{Im } f$, $\text{Ker } f$ の三つの空間のうち二つの次元が分かっているならば、残った空間の次元も計算できる。つまり、三つの空間のうち二つの空間がよく分かっているならば、残りの空間についてもある程度のこと分かるということである。

備考 23.2.4. 定理 23.2.2 の状況のもとで、 $H := \langle \mathbf{u}_1, \dots, \mathbf{u}_m \rangle$ とおけば、 $f|_H : H \rightarrow \text{Im } f$ は基底を基底に写す写像ゆえ線形同型写像である (命題 21.1.6)。

例 23.2.5. (m, n) -行列 A による線形写像 $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ において、例 22.1.3(2) より $\dim \text{Im } T_A = \text{rank } A$, $\dim \text{Ker } f = n - \text{rank } A$ である。ゆえに

$$\dim \text{Ker } f + \dim \text{Im } f = (n - \text{rank } A) + \text{rank } A = n = \dim \mathbb{R}^n$$

であり、確かに T_A において次元公式は成立している。

定理 23.2.2 の設定が成り立つ状況について補足しておこう。例えば、 $\text{Ker } f$ の基底を拡張して U の基底を得たならば、これらは定理 23.2.2 の設定を満たす：

練習 23.2.6. U を有限次元とし、 $f : U \rightarrow V$ を線形写像とする。いま、 $\mathbf{z}_1, \dots, \mathbf{z}_n$ が $\text{Ker } f$ の基底であるとすれば、命題 18.3.2 より、これにいくつかのベクトルを付け加えて $\mathbf{z}_1, \dots, \mathbf{z}_n, \mathbf{u}_1, \dots, \mathbf{u}_m$ を U の基底とすることができる。このとき、 $f(\mathbf{u}_1), \dots, f(\mathbf{u}_m)$ が $\text{Im } f$ の基底となることを示せ。

解答例: 次元公式により $\dim \text{Im } f = \dim U - \dim \text{Ker } f = (n + m) - n = m$ である。命題 20.2.3 より

$$\text{Im } f = \langle f(\mathbf{z}_1), \dots, f(\mathbf{z}_n), f(\mathbf{u}_1), \dots, f(\mathbf{u}_m) \rangle = \langle \mathbf{0}_V, \dots, \mathbf{0}_V, f(\mathbf{u}_1), \dots, f(\mathbf{u}_m) \rangle = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_m) \rangle.$$

ゆえに $f(\mathbf{u}_1), \dots, f(\mathbf{u}_m)$ は $\text{Im } f$ を生成する。更に、命題 22.3.3(3) より、 $\dim \text{Im } f$ 個の組 $f(\mathbf{u}_1), \dots, f(\mathbf{u}_m)$ は $\text{Im } f$ の基底である。□

一方、 $f(\mathbf{u}_1), \dots, f(\mathbf{u}_m)$ が $\text{Im } f$ の基底であるとき、線形独立な組 $\mathbf{u}_1, \dots, \mathbf{u}_m$ を拡張して U の基底を得たとしても、その際に加えたベクトルの組が $\text{Ker } f$ の基底になるとは限らない：

例 23.2.7. 行列 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ による線形写像 $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ において、 $\text{Im } T_A = \langle \mathbf{e}_1 \rangle = \langle T_A(\mathbf{e}_1) \rangle$ および $\text{Ker } T_A = \langle \mathbf{e}_2 \rangle$ である。よって、 $T_A(\mathbf{e}_1)$ は $\text{Im } T_A$ の基底であり、 \mathbf{e}_1 に $\mathbf{u} := \mathbf{e}_1 + \mathbf{e}_2$ を付け加えた \mathbf{e}_1, \mathbf{u} は \mathbb{R}^2 の基底となる。このとき、 $\mathbf{u} \notin \text{Ker } T_A$ ゆえ \mathbf{u} は $\text{Ker } T_A$ の基底ではない。

23.3 商空間 (発展)

代入元 \mathbf{u} を動かすことにより写像 $f : U \rightarrow V$ の値 $f(\mathbf{u})$ を変化させるにはどうすればよいか、あるいは逆に $f(\mathbf{u})$ を変化させないような代入元 \mathbf{u} の動かし方を考えよう。こうした問いは数学の応用とも無縁ではない。例えば写像 f が利益やリスクと関係のある量ならば、これをできる限り都合のよい量に変化させたいと思うことは当然であろう。また、ある都合のよい性質を f の値が特徴づけるのであれば、これらの性質を保ちながら (つまり f の値を変えずに) \mathbf{u} を変化させることにどの程度の自由度があるのか知っておくに越したことはない。この問題に関係して、例えば多変数の微積分学においては、次の事実を学習するであろう：

事実 23.3.1 (よりみち). C^1 級関数 $f : \mathbb{R}^n \rightarrow \mathbb{R}$ および点 $\mathbf{a} \in \mathbb{R}^n$ が与えられているとし、 $\mathbf{x} = \mathbf{a}$ をわずかに動かして $f(\mathbf{x})$ の値を変化させることを考えよう。このとき、勾配 $\text{grad } f(\mathbf{a})$ 方向に動かすとき $f(\mathbf{x})$ の変化率³は最大となる。ここで、

$$\text{grad } f(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \frac{\partial f}{\partial x_2}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right)$$

である。また、 $\text{grad } f(\mathbf{a})$ に直行する方向に動かしたときの $f(\mathbf{x})$ の変化率は 0 である。

³ここでいう変化率とは方向微分のこと。

線形代数学の話題に戻ろう。 f が線形写像である場合、この種の問題に対する答えは次の通りである。

命題 23.3.2. 線形写像 $f : U \rightarrow V$ および $\mathbf{u}, \mathbf{z} \in U$ において、 $f(\mathbf{u}) = f(\mathbf{u} + \mathbf{z})$ となるための必要十分条件は $\mathbf{z} \in \text{Ker } f$ となることである。

Proof. $\mathbf{b} := f(\mathbf{u})$ とおこう。 $\mathbf{z} \in \text{Ker } f$ を仮定すれば、命題 23.1.4(1) より $f(\mathbf{u} + \mathbf{z}) = \mathbf{b}$ (すなわち $f(\mathbf{u}) = f(\mathbf{u} + \mathbf{z})$) である。逆に、 $f(\mathbf{u}) = f(\mathbf{u} + \mathbf{z}) = \mathbf{b}$ ならば $f(\mathbf{z}) = f(-\mathbf{u} + \mathbf{u} + \mathbf{z}) = -f(\mathbf{u}) + f(\mathbf{u} + \mathbf{z}) = -\mathbf{b} + \mathbf{b} = \mathbf{0}$ より $\mathbf{z} \in \text{Ker } f$ を得る。 \square

つまり、 $\text{Ker } f$ と平行な方向にベクトル \mathbf{u} を移動させるか否かで $f(\mathbf{u})$ が変化するかどうかが決まるのである。とくに f の値の変化を望む立場からは $\text{Ker } f$ 方向への移動は無意味である。そこで、 U の中で $\text{Ker } f$ 方向の情報を捨象する概念として商空間を与えよう。ここでは $\text{Ker } f$ に限らず、一般の部分空間 $W \subset U$ に対して商空間 U/W を定める。

定義 23.3.3. U を線形空間とし、 W をその部分空間とする。 W を平行移動させた集合 (これは U の部分集合である) をすべて集めた集合 (つまり集合族) を U の商空間 (quotient space) と呼び、 U/W と書く。すなわち、

$$U/W := \{ \mathbf{u} + W \mid \mathbf{u} \in U \}.$$

集合 $\mathbf{u} + W$ を商空間 U/W の元とみなすとき、これを略して $[\mathbf{u}]$ と書く。

$\mathbf{u} = \mathbf{u} + \mathbf{0} \in \mathbf{u} + W = [\mathbf{u}]$ より $\mathbf{u} \in [\mathbf{u}]$ であることに注意せよ。また、 $[\mathbf{0}] = \mathbf{0} + W = W$ である。

例 23.3.4. \mathbb{R}^3 における x - z 平面を $W = \{ (x, y, z) \in \mathbb{R}^3 \mid y = 0 \}$ とすれば、 \mathbb{R}^3/W は W と平行な平面によって構成される集合である。図 23.3 では有限個の平面しか描かれていないが、実際は無数の平面が連続的に並んでいる。

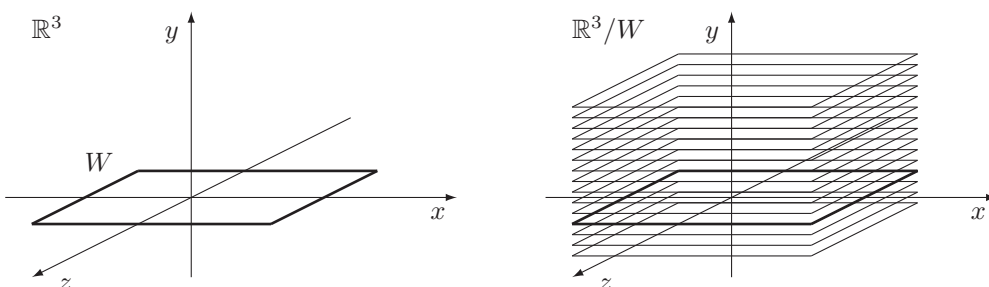


図 23.3: \mathbb{R}^3 上の平面 W による商空間 \mathbb{R}^3/W

U の元 \mathbf{u} の代わりに U/W の元 $[\mathbf{u}]$ を考えるにあたって、この対応に関する写像は商写像と呼ばれる:

定義 23.3.5. U の各元 $\mathbf{u} \in U$ に対して、 $[\mathbf{u}] \in U/W$ を対応させる写像 $q : U \rightarrow U/W$ を商写像 (quotient map) と言う。

商写像 $q : U \rightarrow U/W$ は全射である。実際、 U/W の各元 $\mathbf{u} + W$ に対して、 $q(\mathbf{u}) = \mathbf{u} + W$ である。

次の命題は、 \mathbf{u} と $\mathbf{u} + \mathbf{w}$ が商空間 U/W において同じ元を表すこと (すなわち $[\mathbf{u}] = [\mathbf{u} + \mathbf{w}]$) と $\mathbf{w} \in W$ であることの同値性を述べている (系 23.3.7)。つまり、 U の元 \mathbf{u} の代わりに $[\mathbf{u}]$ を考えるということは、 W と平行な方向を捨象することを意味している。

命題 23.3.6. 各 $\mathbf{u}, \mathbf{v} \in U$ について、次の条件はすべて同値である:

- (1) $[\mathbf{u}] = [\mathbf{v}]$,
- (2) $-\mathbf{v} + \mathbf{u} \in W$,
- (3) $-\mathbf{u} + \mathbf{v} \in W$,
- (4) $\mathbf{v} \in [\mathbf{u}]$.

Proof. W が部分空間ゆえ -1 倍の演算について閉じており、ゆえに (2) \Leftrightarrow (3) は明らかである。

(1) \Rightarrow (2): $[\mathbf{u}] = [\mathbf{v}]$ (すなわち $\mathbf{u} + W = \mathbf{v} + W$) を仮定しよう。すると $\mathbf{u} \in [\mathbf{u}] = \mathbf{v} + W$ ゆえ、ある $\mathbf{w} \in W$ を用いて $\mathbf{u} = \mathbf{v} + \mathbf{w}$ と書くことができる。この両辺に $-\mathbf{v}$ を加えて $-\mathbf{v} + \mathbf{u} = \mathbf{w} \in W$ を得る。

(2) \Rightarrow (1): $-\mathbf{v} + \mathbf{u} \in W$ を仮定しよう。 $[\mathbf{u}] = [\mathbf{v}]$ を示すには $[\mathbf{u}] \subset [\mathbf{v}]$ および $[\mathbf{v}] \subset [\mathbf{u}]$ を示せばよい。まず $[\mathbf{u}] \subset [\mathbf{v}]$ を示すために $\mathbf{a} \in [\mathbf{u}]$ を任意に取る。すると、ある $\mathbf{w} \in W$ を用いて $\mathbf{a} = \mathbf{u} + \mathbf{w}$ と書くことができる。このとき、別の $\mathbf{w}' \in W$ を用いて $\mathbf{a} = \mathbf{v} + \mathbf{w}'$ となることを示せばよいのであるが、これは次の計算により得られる。

$$\mathbf{a} = \mathbf{u} + \mathbf{w} = (\mathbf{v} - \mathbf{v}) + \mathbf{u} + \mathbf{w} = \mathbf{v} + (-\mathbf{v} + \mathbf{u}) + \mathbf{w}.$$

つまり、 $\mathbf{w}' := (-\mathbf{v} + \mathbf{u}) + \mathbf{w}$ とすれば、 $(-\mathbf{v} + \mathbf{u}), \mathbf{w} \in W$ および W が和の演算について閉じていることから $\mathbf{w}' \in W$ であり、 $\mathbf{a} = \mathbf{v} + \mathbf{w}' \in \mathbf{v} + W$ を得る。いま $[\mathbf{u}] \subset [\mathbf{v}]$ であることが分かった。 \mathbf{u} と \mathbf{v} の立場を入れ替えることで逆向きの包含関係 $[\mathbf{v}] \subset [\mathbf{u}]$ も同様にして示すことができる。その際には、(2) と (3) が同値であることから $-\mathbf{u} + \mathbf{v} \in W$ を用いればよい。

(1) \Rightarrow (4): $\mathbf{v} \in [\mathbf{v}] = [\mathbf{u}]$ より直ちに分かる。

(4) \Rightarrow (3): $\mathbf{v} \in [\mathbf{u}] = \mathbf{u} + W$ を仮定すれば、ある $\mathbf{w} \in W$ を用いて $\mathbf{v} = \mathbf{u} + \mathbf{w}$ と書ける。ゆえに $-\mathbf{u} + \mathbf{v} = \mathbf{w} \in W$. □

系 23.3.7. 各 $\mathbf{w} \in U$ について、 $[\mathbf{u} + \mathbf{w}] = [\mathbf{u}] \iff \mathbf{w} \in W$.

Proof. 前命題の \mathbf{u} および \mathbf{v} として、 $\mathbf{u} + \mathbf{w}$ および \mathbf{u} を取れば主張を得る。 □

23.4 商空間の例 (発展)

例 23.4.1. (1) \mathbb{R}^2 における x 軸上の点のなす直線 $X = \{(x, y) \in \mathbb{R}^2 \mid y = 0\}$ に対して、 \mathbb{R}^2/X は X と平行な直線からなる集合 (直線を元とする集合) である。 \mathbb{R}^2/X と y 軸上の点のなす集合 Y には自然な 1 対 1 対応がある。実際、 \mathbb{R}^2/X の元である直線 $\ell = \{(x, y) \in \mathbb{R}^2 \mid y = a\}$ に対して、 ℓ と Y との交点 $(0, a)$ を対応させる写像 $\mathbb{R}^2/X \rightarrow Y$ は全単射である。

(2) 上の例において \mathbb{R}^2/X との間には 1 対 1 対応を与える直線を Y に限る必要はなく、直線 X と平行でない任意の直線 L について、 \mathbb{R}^2/X と L の間に 1 対 1 対応がつく。実際、各直線 $\ell \in \mathbb{R}^2/X$ に対して、 ℓ と L の交点を対応させれば (ℓ と L は平行でないゆえ一点で交わる)、この対応 $\mathbb{R}^2/X \rightarrow L$ は全単射である。

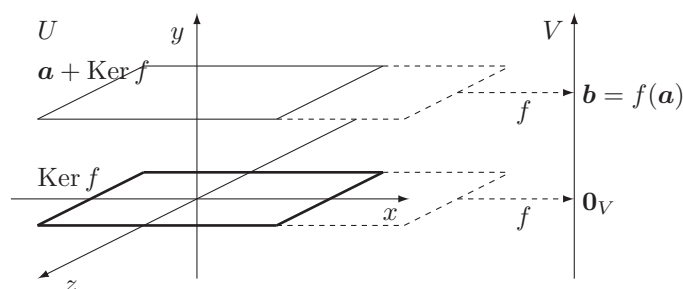
(3) \mathbb{R}^2 における直線 $y = \frac{1}{2}x$ 上の点全体を L とすれば、 \mathbb{R}^2/L と y 軸との間でも 1 対 1 対応がつく。

(4) \mathbb{R}^3 における x 軸上の点のなす直線 X に対して、 \mathbb{R}^3/X と y - z 平面の間に 1 対 1 対応がつく。

上の例から、商空間も線形空間となることが示唆される。商空間にどのように和とスカラー倍を定めるかは、次章で詳しく述べよう。

線形写像 $f: U \rightarrow V$ に対して、 f の値が動かないような代入元 \mathbf{u} の移動とは、 $\text{Ker } f$ と平行な方向に移動させることに他ならなかった。そこで、 $\text{Ker } f$ 方向を捨象した $U/\text{Ker } f$ を定義域とする f の代わりとなる写像 $F: U/\text{Ker } f \rightarrow V$ を考えることができる。

例 23.4.2. $f: U \rightarrow V$ を線形写像とする。各 $\mathbf{b} \in \text{Im } f$ に対して $f(\mathbf{a}) = \mathbf{b}$ なる $\mathbf{a} \in U$ を取れば、命題 23.1.4(3) より $f^{-1}(\mathbf{b}) = \mathbf{a} + \text{Ker } f = [\mathbf{a}] \in U/\text{Ker } f$ である。



そこで、各 $f^{-1}(b) = [a] \in U/\text{Ker } f$ に対して $b \in V$ を対応させる写像を $F: U/\text{Ker } f \rightarrow V$ と定めよう。写像の値が変化しない方向を潰したことから、 L が $U/\text{Ker } f$ 上を動けば $F(L)$ の値も必ず変化し、したがって F は単射となる。

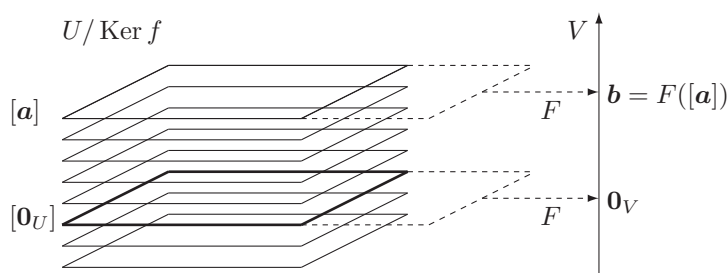


図 23.4: 線形写像 f が誘導する写像 $F: U/\text{Ker } f \rightarrow V$

実は、上の F は線形写像となる。 F の線形性および単射性の証明は定理 24.2.1(第 1 同型定理) をみよ。

23.5 同値関係と商集合(発展)

商空間 U/W の各元は W と平行な集合であるゆえ、互いに交わらないはずである。これを証明によって確認しておこう。

命題 23.5.1. 商空間 U/W において次が成り立つ。

- (1) 各 $u \in U$ に対して、商空間 U/W の要素で u を含むものは $[u]$ 唯一のみである。
- (2) 各 $L \in U/W$ および $u \in L$ について $[u] = L$ 。とくに $w \in W$ について $[w] = W = [0]$ 。
- (3) 各 $L \in U/W$ および $u \in U$ について、 $[u] = L \iff u \in L$ 。
- (4) U/W の任意の二つの元は、完全に一致するか交わらないかのいずれかである。

(すなわち、 $L, L' \in U/W$ とすれば $L \cap L' = \emptyset$ または $L = L'$ のいずれかが成り立つ⁴。)

Proof. (1): U/W の二つの元 $[a] = a + W$ と $[b] = b + W$ がともに u を含むと仮定する。すると、ある $w, w' \in W$ を用いて $u = a + w = b + w'$ と書くことができる。このとき、 $-b + a = w' - w \in W$ であり、命題 23.3.6 より $[a] = [b]$ 。すなわち、 u を含む U/W の要素は一つしかない(注意: $[u]$ は u を含む U/W の要素ゆえ $[u] = [a] = [b]$ が成り立っている)。

(2): $L \in U/W$ および $u \in L$ について、 $[u]$ と L はともに u を元として含む U/W の要素である。このような U/W の要素は (1) より一つしかないことから、 $[u] = L$ である。

(3): (\Leftarrow) は (2) で示している。また、 $[u] = L$ とすれば $u \in [u] = L$ ゆえ $u \in L$ である。

(4): $L, L' \in U/W$ および $L \cap L' \neq \emptyset$ とすれば $u \in L \cap L'$ が取れる。このとき、(1) より $L = L' = [u]$ である。 □

上の命題から、商空間 U/W は U の各元を互いに交わらないグループに分類していることが分かる。 $u \in U$ の属するグループが $[u]$ という具合である。このような交わりのない分類(グループ分け)は、線形代数の枠組みに縛られない一般論として次のように定められる:

定義 23.5.2. 集合 X の元に対して同値関係⁵ \simeq が与えられているとき、各 $x \in X$ の同値類 $[x] \subset X$ を次のように定める。

$$[x] := \{y \in X \mid y \simeq x\}.$$

⁴集合演算記号 \cap は共通部分を表す。すなわち、集合 A, B のいずれにも含まれる元をすべて集めた集合を A と B の共通部分と呼び、これを $A \cap B$ で表す。

⁵同値関係の定義は 21.2 節のコラムを見よ。

また, X の部分集合からなる次の集合族 X/\simeq を同値関係 \simeq に関する商集合という:

$$X/\simeq := \{ [x] \mid x \in X \}.$$

商集合は X の各元を互いに交わらないグループに分類しており, 命題 23.3.6 および 23.5.1 と同様の主張が一般の商集合においても成り立つ:

命題 23.5.3. X 上の同値関係 \simeq による商集合 X/\simeq において次が成り立つ.

- (1) 各 $x, y \in X$ について, $[x] = [y] \iff x \simeq y$.
- (2) 各 $x \in X$ に対して, 商集合 X/\simeq の要素で x を含むものは $[x]$ 唯一つのみである.
- (3) 各 $L \in X/\simeq$ および $x \in X$ について, $[x] = L \iff x \in L$.
(特に $y \in X, L = [y]$ とすれば, $[x] = [y] \iff x \in [y]$.)
- (4) X/\simeq の任意の二つの元は, 完全に一致するか交わらないかのいずれかである.

Proof. (1): (\Rightarrow) を示すために $[x] = [y]$ を仮定しよう. $x \simeq x$ (反射律) より $x \in [x]$ であり, $[x] = [y]$ ゆえ $x \in [y]$. つまり, $x \simeq y$ である. 次に (\Leftarrow) を示すために $x \simeq y$ を仮定しよう. $[x] = [y]$ を示すには $[x] \subset [y]$ および $[y] \subset [x]$ を示せばよい. まず $[x] \subset [y]$ を示すために $a \in [x]$ を任意に取る. このとき $a \simeq x$ であり, これと $x \simeq y$ から $a \simeq y$ を得る (推移律). つまり $a \in [y]$ である. いま $[x] \subset [y]$ であることが分かった. x と y の立場を入れ替えることで $[y] \subset [x]$ も同様に示すことができる. この証明において, 対称律から $y \simeq x$ であることに注意せよ.

(2): X/\simeq の二つの元 $[a]$ と $[b]$ がともに x を含むと仮定する. すると, $x \simeq a$ および $x \simeq b$ が成り立っている. よって対称律と推移律より $a \simeq b$ であり, (1) より $[a] = [b]$. すなわち, x を含む X/\simeq の要素は一つしかない (注意: $[x]$ は x を含む X/\simeq の要素ゆえ $[x] = [a] = [b]$ が成り立っている).

(3) および (4) は, 命題 23.5.1(2) および (3), (4) と同様に示される. □

例 23.5.4. 線形空間 U およびその部分空間 W が与えられているとする. U の各元 u, v に対して, 記号 $u \simeq v$ を「 $-v + u \in W$ を満たすこと」と定めれば, これは同値関係になる (各自確かめよ). このとき次が成り立つ:

- 各 $u \in U$ について, いま定めた \simeq に関する u の同値類と集合 $u + W$ は一致する.

Proof. 命題 23.3.6 では $u + W$ のことを $[u]$ と書いていたため, これと区別して同値関係 \simeq に関する u の同値類を記号 $]u[$ で表そう. すなわち, $]u[:= \{ v \in U \mid v \simeq u \}$ と定める. 我々が示すべきことは $]u[= [u]$ であり, これは命題 23.3.6 より得られる:

$$v \in]u[\iff u \simeq v \iff -v + u \in W \iff v \in [u].$$

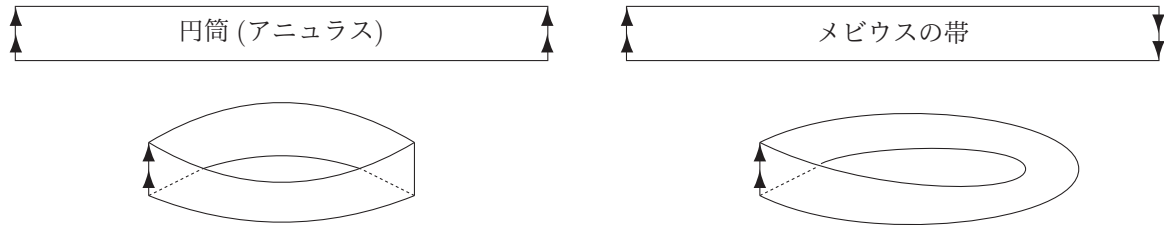
↑ 命題 23.3.6

□

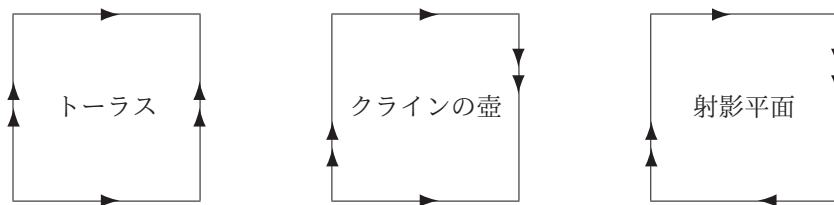
すなわち, この同値関係による商集合 U/\simeq と本章で与えた商空間 U/W は一致する.

よりみち (図形の貼り合わせ).

商集合は、対象の同一視を記述するための言葉である。代数学以外の文脈でも用いられる概念であり、例えば幾何学における商集合に図形の貼り合わせがある。最も単純な貼り合わせの例は、線分の両端点を同一視することで得る円周であろう。ここから一つ次元を挙げて、長方形 (内側も含む) の左右の両辺を貼り合わせる (すなわち両辺を同一視する) 操作を考えよう。両辺の貼り合わせ方には、向きをそろえる場合と逆にする場合の二通りが考えられ、前者による貼り合わせからは円筒 (円柱の側面) が得られ、後者からはメビウスの帯が得られる。



さらに、長方形の上下の辺どうしも貼り合わせるとすれば、次の3通りが考えられる:



左のトーラスとは、いわゆるドーナツの表面のような図形のことであり、細長い円筒の両端を自然に貼り合わせることでトーラスが得られる。この両端の貼り合わせを逆向きに行えばクラインの壺になるのであるが、折り紙などを用いて実際に貼り合わせようと思うと、どうしても面が重なりあって上手くいかないはずである。射影平面はさらに複雑な図形のように思えるだろう。実は、右の二つの図形は \mathbb{R}^3 上の図形として実現できないことが知られている。また、これらはメビウスの帯を含むことから、表裏が定まらない曲面でもある。このように、貼り合わせによって構成した新しい図形の性質を調べる際に、証明が求められる数学では貼り合わせの厳密な定義が必要となる。この要請に応える概念が同値関係および商集合なのである。

上の構成では射影平面のイメージがつかみにくいから、別の方法による構成を紹介しよう。 \mathbb{R}^3 の単位球面

$$S^2 = \{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1 \}$$

にある点に対して、 $x \simeq y$ を「 $x = y$ または $x = -y$ であること」と定めると \simeq は同値関係となる。これは、球の中心について点対称な位置にある互いの点 (対蹠点) を同一視することに相当する。先程貼り合わせによって作った射影平面と商集合 S^2/\simeq との間には自然な 1 対 1 対応がある。このことは、北半球面 $\{ (x, y, z) \in S^2 \mid z \geq 0 \}$ において今の同値関係による同一視を行うと理解し易い。北半球面に限ると、同一視すべき異なる二点は赤道上にしか現れない。そして、赤道上の点における同一視は、ちょうど上で定めた射影平面における長方形の各辺の同一視と対応付けられる。なお、球面の赤道付近の帯

$$A = \left\{ (x, y, z) \in S^2 \mid |z| \leq \frac{1}{10} \right\}$$

において、 A/\simeq はメビウスの帯となる。いまの例とは別に、射影平面には次のような構成もある。

練習 23.5.5. \mathbb{R}^3 上の原点を通る直線全体の集合を P とおくと、 P と S^2/\simeq の間に自然な 1 対 1 対応があることを確認せよ。

解答例: 各 $s \in S^2/\simeq$ に対して、同値類 s は S^2 の二点からなる部分集合であり、これらの点は互いに原点对称な位置にある。そこで、 s 上の 2 点を結ぶ直線を $F(s)$ とすれば、これは原点を通る直線である。対応 $F : S^2/\simeq \rightarrow P$ は全単射である。 \square

第24章 第1同型定理と短完全列(発展)

前章において、線形空間を分解する概念として次元公式と商空間を与えた。本章では、これらの関係について論じ、次元公式を商空間の立場から述べた主張である第1同型定理を導く。実は、線形空間に限らず多くの代数構造(群や環, 加群, 多元環など)においても第1同型定理は認められる。つまり、この定理は代数構造を分解して理解するうえで基本となる考え方であり、代数学では至る所で用いられるものである。これと関連して、第1同型定理による代数構造の分解を図式で表す短完全列についても少しだけ触れる。

24.1 商空間の線形構造

商空間における和とスカラー倍の定義を与えよう。

補題 24.1.1. 商空間 U/W において和とスカラー倍の演算を次のように定めることができる。

各 $[u], [v] \in U/W$ および $r \in \mathbb{R}$ に対して, $[u] + [v] := [u + v]$, $r[u] := [ru]$.

ここで、「定めることができる」と書いたのは次の点に考慮せねばならないからである: U/W の元 L を U の元を用いて表す方法は命題 23.5.1 に見るように何通りもあり(実際 L の元の個数ぶんだけある), 例えば $L = [u] = [u']$ としよう。このとき L のスカラー倍 rL は, $[ru]$ と $[ru']$ のいずれであると定めるべきだろうか。もし $[ru] \neq [ru']$ ならば L の代表として u と u' のいずれを選ぶのか(あるいは u, u' のいずれとも異なる L の元を代表に選ぶのか) あらかじめ決めておかなければ, 上は定義として認められない(読み手に定義が伝わらない)ことになる。しかしながら, 実際には $[ru] = [ru']$ が成り立ち, L の中の

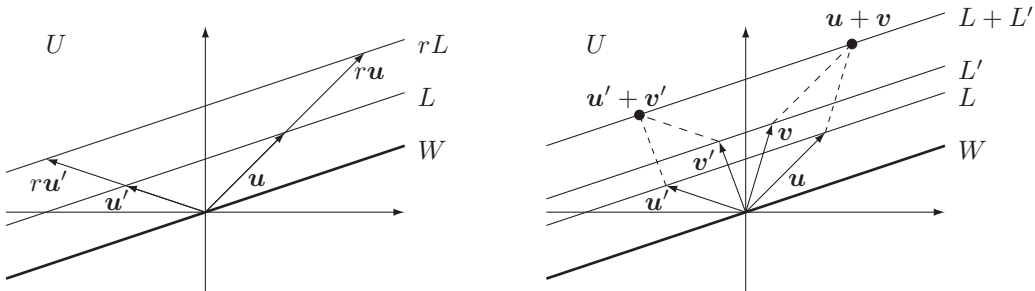


図 24.1: $L \in U/W$ のスカラー倍および $L' \in U/W$ との和

どのベクトルを代表に選んでも定義が変わることはない。つまり, L の代表元は何を選ぶか指定する必要はないのである。和の定義についても同様のことが言える(図 24.1)。

補題 24.1.1 の証明. 上で説明したように, 示すべき事は和とスカラー倍の定義が代表元の取り方によらないこと, すなわち次の二つである:

$$(1) [u] = [u'] \implies [ru] = [ru']. \quad (2) [u] = [u'] \text{ かつ } [v] = [v'] \implies [u + v] = [u' + v'].$$

(1): $[u] = [u']$ とすれば命題 23.3.6 より $-u' + u \in W$ であり, ゆえに $-ru' + ru = r(-u' + u) \in W$. 再び命題 23.3.6 より $[ru] = [ru']$ を得る。

(2): 仮定より $-u' + u, -v' + v \in W$ であり, ゆえに

$$-(u' + v') + (u + v) = (-u' + u) + (-v' + v) \in W$$

である¹. したがって $[u + v] = [u' + v']$. □

上の補題により, 任意の $u, v \in U$ および $r \in \mathbb{R}$ について次の式変形が認められることになる.

$$[u + v] = [u] + [v], \quad [ru] = r[u].$$

$[0] = W \in U/W$ を零元とみなせば, 補題 24.1.1 で定めた U/W 上の演算がベクトル空間の公理を満たすことは容易に確かめられる. したがって U/W は線形空間となる.

命題 24.1.2. 商写像 $q: U \rightarrow U/W$ は線形全射であり, $\text{Ker } q = W$.

Proof. 線形性 (iii) は直ちに確認できる:

$$q(ru + sv) = [ru + sv] = [ru] + [sv] = r[u] + s[v] = rq(u) + sq(v).$$

また, $[0] = W$ に注意すると,

$$u \in \text{Ker } q \iff q(u) = [0] \iff [u] = W \iff u \in W \quad (\text{命題 23.5.1(3)}).$$

すなわち, $\text{Ker } q = W$. □

命題 24.1.3. 有限次元線形空間 U および部分空間 $W \subset U$ について, $\dim U/W = \dim U - \dim W$.

Proof. 商写像 $q: U \rightarrow U/W$ に対して次元公式を適用すればよい. q の全射性より $U/W = \text{Im } q$ であること, また $\text{Ker } q = W$ に注意すれば,

$$\begin{aligned} \dim \text{Im } q + \dim \text{Ker } q &= \dim U \\ \dim U/W + \dim W &= \dim U \\ \dim U/W &= \dim U - \dim W. \end{aligned}$$

□

練習 24.1.4. U を有限次元線形空間とし, W をその部分空間とする. W の基底 w_1, \dots, w_n にいくつかのベクトルを付け加えて $w_1, \dots, w_n, u_1, \dots, u_m$ を U の基底とすれば $[u_1], \dots, [u_m]$ は U/W の基底となる. これを示せ.

解答例: 練習 23.2.6 と同様の論法をたどればよい. 前命題より $\dim U/W = \dim U - \dim W = (n + m) - n = m$ である. 命題 20.2.3 より

$$\text{Im } q = \langle q(w_1), \dots, q(w_n), q(u_1), \dots, q(u_m) \rangle = \langle [0], \dots, [0], [u_1], \dots, [u_m] \rangle = \langle [u_1], \dots, [u_m] \rangle.$$

ゆえに $[u_1], \dots, [u_m]$ は $\text{Im } q = U/W$ を生成する. 更に, 命題 22.3.3(3) より, $\dim U/W$ 個の組 $[u_1], \dots, [u_m]$ は U/W の基底である. □

¹いまの式変形を非可換な場合も見据えた形で行うと次のようになる:

$$\begin{aligned} -(u' + v') + (u + v) &= (-v' - u') + (u + v) = -v' + (-u' + u) + v \\ &\in -v' + (W + v) = -v' + (v + W) = (-v' + v) + W = [-v' + v] = W. \end{aligned}$$

非可換な世界では一般に $(W + v) = (v + W)$ は成立せず, 上の式変形を行うには条件 $(W + v) = (v + W)$ を仮定する必要がある. この条件と関連して, 正規部分群なる概念を群論で学ぶことになる.

24.2 第1同型定理

例 23.4.2 で述べた事実の詳細は次の通りである。この定理を応用する際は、結論そのものよりもむしろ、同型写像 F の構成 (定義) が無視できない役割を果たすことが多い。また、代数学においては、自明の定理といっても差し支えないくらい何度も用いられる事実である。

定理 24.2.1 (第1同型定理). 線形写像 $f : U \rightarrow V$ について、 $U/\text{Ker } f$ と $\text{Im } f$ は線形同型である。

Proof. 写像 $F : U/\text{Ker } f \rightarrow \text{Im } f$ を $F([\mathbf{u}]) := f(\mathbf{u})$ と定めることができる。この定義が代表元の取り方によらないこと、すなわち「 $[\mathbf{u}] = [\mathbf{v}]$ ならば $f(\mathbf{u}) = f(\mathbf{v})$ 」であることを確認しよう。実際、 $[\mathbf{u}] = [\mathbf{v}]$ ならば $-\mathbf{v} + \mathbf{u} \in \text{Ker } f$ であり、これは $f(-\mathbf{v} + \mathbf{u}) = \mathbf{0}_V$ を意味する。つまり $-f(\mathbf{v}) + f(\mathbf{u}) = f(-\mathbf{v} + \mathbf{u}) = \mathbf{0}_V$ であり、 $f(\mathbf{u}) = f(\mathbf{v})$ を得る。

いま定めた F が線形同型であることを示そう。まず線形性 (iii) は次のように確認できる：

$$F(r[\mathbf{u}] + s[\mathbf{v}]) = F([\mathbf{ru} + \mathbf{sv}]) = f(\mathbf{ru} + \mathbf{sv}) = rf(\mathbf{u}) + sf(\mathbf{v}) = rF([\mathbf{u}]) + sF([\mathbf{v}]).$$

また、各 $f(\mathbf{u}) \in \text{Im } f$ に対して、 $F([\mathbf{u}]) = f(\mathbf{u})$ であるから F は全射である。最後に単射性を示そう。 $[\mathbf{u}] \in \text{Ker } F$ とすれば $F([\mathbf{u}]) = \mathbf{0}_V$ 、すなわち $f(\mathbf{u}) = \mathbf{0}_V$ であり、したがって $\mathbf{u} \in \text{Ker } f$ である。ゆえに命題 23.5.1(2) より $[\mathbf{u}] = \text{Ker } f = [\mathbf{0}]$ であり、 $\text{Ker } F$ の元は零元のみであることが分かった。すなわち F は単射である。□

備考 24.2.2. 上の $F : U/\text{Ker } f \rightarrow \text{Im } f$ の逆写像は、各 $\mathbf{v} \in \text{Im } f$ に対して逆像 $f^{-1}(\mathbf{v})$ を対応させる写像である。ここで、集合 $f^{-1}(\mathbf{v})$ が $U/\text{Ker } f$ の元となることは命題 23.1.4(3) による。実際、 $f^{-1}(\mathbf{v}) = \mathbf{a} + \text{Ker } f = [\mathbf{a}] \in U/\text{Ker } f$ (ただし $\mathbf{a} \in f^{-1}(\mathbf{v})$) であり、これを F に代入すれば、 $F([\mathbf{a}]) = f(\mathbf{a}) = \mathbf{v}$ 。

練習 24.2.3. 次元公式を用いて、 U が有限次元の場合における線形写像 $f : U \rightarrow V$ の第1同型定理の別証明を与えよ。

解答例: 命題 24.1.3 および次元公式より $\dim U/\text{Ker } f = \dim U - \dim \text{Ker } f = \dim \text{Im } f$ 。したがって $U/\text{Ker } f$ と $\text{Im } f$ の次元は等しく、ゆえにこれらは同型である。□

有限次元の空間に限れば、上のようにして第1同型定理の主張する線形同型性は次元定理から直ちに分かってしまう。したがって、第1同型定理の真の御利益は何かと問われれば、それは無限次元空間を対象とするか、あるいは定理の証明に現れた同型写像 F を本質的に必要とするような例を挙げねばならない。しかし、何の数学理論の予備知識も仮定せずに、そのような応用例を説明するのは残念ながら困難なことである。そこで、ここでは次の例を述べるに留めておく。

例 24.2.4 (不定積分の線形性). 例 20.4.5(2) において、微分作用素 $D : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ ($D(f) := f'$) の核 $\text{Ker } D$ が定数関数全体に一致することを見た。つまり $C^\infty(\mathbb{R})/\text{Ker } D$ は、定数関数だけの差を無視した関数の空間である。すなわち、関数 $f \in C^\infty(\mathbb{R})$ および定数 $C \in \mathbb{R}$ について、

$$[f(x)] = [f(x) + C] \in C^\infty(\mathbb{R})/\text{Ker } D.$$

D は全射ゆえ $\text{Im } D = C^\infty(\mathbb{R})$ であり、これに第1同型定理を適用すれば線形同型 $\mathcal{D} : C^\infty(\mathbb{R})/\text{Ker } D \rightarrow C^\infty(\mathbb{R})$ ($\mathcal{D}([f]) := D(f) = f'$) を得る。 \mathcal{D} の逆写像 $\mathcal{I} : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})/\text{Ker } D$ は、各 $g \in C^\infty(\mathbb{R})$ に対して集合 $D^{-1}(g)$ (すなわち、 g の原始関数全体) を対応させる写像である。ここで g の原始関数の一つ G を取れば、

$$\mathcal{I}(g) = D^{-1}(g) = [G] = G + \text{Ker } D \in C^\infty(\mathbb{R})/\text{Ker } D.$$

原始関数には定数関数の個数だけ任意性があるものの、 $\mathcal{I}(g)$ は原始関数の選び方に依らずに決まる $C^\infty(\mathbb{R})/\text{Ker } D$ の元である。ふだん $\mathcal{I}(g)$ は f の不定積分とよばれ、微積分学では次の式で表される：

$$\int g(x) dx.$$

こうして我々は、不定積分が $C^\infty(\mathbb{R})/\text{Ker } D$ への線形写像として実現されること、とくに不定積分の線形性を得る。なお、例 20.4.7(2) で与えた単射 $I_a : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ は原始関数の一つを与える線形写像であったゆえ、 $q : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})/\text{Ker } D$ を商写像とすれば $\mathcal{I} = q \circ I_a$ が成り立つ。

24.3 完全系列と短完全列

定義 24.3.1. 線形写像の列

$$\cdots \xrightarrow{f_{n-2}} U_{n-1} \xrightarrow{f_{n-1}} U_n \xrightarrow{f_n} U_{n+1} \xrightarrow{f_{n+1}} U_{n+2} \xrightarrow{f_{n+2}} \cdots \quad (24.3.1)$$

が各 $n \in \mathbb{Z}$ について $\text{Im } f_n = \text{Ker } f_{n+1}$ を満たすとき、これを完全系列 (exact sequence) という。また、次のような特別な線形写像の列

$$\cdots \xrightarrow{\text{id}} \{0\} \xrightarrow{\text{id}} \{0\} \xrightarrow{f_0} W \xrightarrow{f_1} U \xrightarrow{f_2} V \xrightarrow{f_3} \{0\} \xrightarrow{\text{id}} \{0\} \xrightarrow{\text{id}} \cdots$$

を略して次のように書く²:

$$0 \xrightarrow{f_0} W \xrightarrow{f_1} U \xrightarrow{f_2} V \xrightarrow{f_3} 0 \quad (24.3.2)$$

この列が完全系列であるとき、これを短完全列 (short exact sequence) という。

列 24.3.1 が完全系列であるとき、各 $\mathbf{u} \in U_n$ について $f_n(\mathbf{u}) \in \text{Im } f_n = \text{Ker } f_{n+1}$ ゆえ $f_{n+1}(f_n(\mathbf{u})) = \mathbf{0}_{U_{n+2}}$ である。したがって合成写像 $f_{n+1} \circ f_n : U_n \rightarrow U_{n+2}$ は、定義域のすべての元を零元に写す写像である (つまり $f_{n+1} \circ f_n = \mathbf{0}_{U_{n+2}}$)。また、線形写像の列 24.3.2 において、両端の写像 f_0 および f_3 はともに自明な写像ゆえ記号を割りふらないことが多い。 f_0 は零元を零元に写す線形写像であり (つまり $f_0 = \mathbf{0}_W$)、 f_3 はすべての元を零元に写す線形写像である (つまり $f_3 = \mathbf{0}$)。

短完全列の定義から次の性質が直ちに従う。

補題 24.3.2. 線形写像の列 24.3.2 において次が成り立つ。とくに列 24.3.2 が短完全列であるとき、これらの性質が成り立っている。

(1) $\text{Im } f_0 = \text{Ker } f_1 \iff f_1 : W \rightarrow U$ は単射である (つまり W と $\text{Im } f_1$ は同型)。

(2) $\text{Im } f_2 = \text{Ker } f_3 \iff f_2 : U \rightarrow V$ は全射である。

Proof. (1): $\text{Im } f_0 = \{0_W\}$ ゆえ、 $\text{Im } f_0 = \text{Ker } f_1 \iff \text{Ker } f_1 = \{0_W\} \iff f_1$ は単射。

(2): $\text{Ker } f_3 = V$ ゆえ、 $\text{Im } f_2 = \text{Ker } f_3 \iff \text{Im } f_2 = V \iff f_2$ は全射。 □

線形写像があると、対応する短完全列を与えることができる:

例 24.3.3. (1) 任意の線形写像 $f : U \rightarrow V$ について、次は短完全列である。

$$0 \longrightarrow \text{Ker } f \xrightarrow{\text{id}} U \xrightarrow{f} \text{Im } f \longrightarrow 0$$

(2) U の部分空間 W および商写像 $q : U \rightarrow U/W$ について、次は短完全列である。

$$0 \longrightarrow W \xrightarrow{\text{id}} U \xrightarrow{q} U/W \longrightarrow 0$$

Proof. (1) で与えられた写像列が短完全列であることを示すには、前補題より次の三つの性質: $\text{id} : \text{Ker } f \rightarrow U$ が単射であること、および $\text{id}(\text{Ker } f) = \text{Ker } f$ 、 $f : U \rightarrow \text{Im } f$ が全射であることを示せばよい。しかし、これらはすべて明らかである。また、商写像 q について (1) を適用すると、命題 24.1.2 より $\text{Ker } q = W$ ゆえ (2) を得る。 □

短完全列において第 1 同型定理を適用してみよう。

命題 24.3.4. 短完全列

$$0 \longrightarrow W \xrightarrow{f_1} U \xrightarrow{f_2} V \longrightarrow 0$$

において、 f_1 の単射性より W は U の部分空間とみなすことができる ($W \simeq \text{Im } f_1 \subset U$)。そこで W と $\text{Im } f_1$ を同一視すれば、 U/W と V は同型である。

²代数学では、自明な空間 $\{0\}$ を 0 と略するのが慣例となっている。

Proof. $\text{Im } f_1 = \text{Ker } f_2$ および f_2 の全射性に注意すると, f_2 に関する第 1 同型定理により次を得る:

$$V = \text{Im } f_2 \simeq U / \text{Ker } f_2 = U / \text{Im } f_1 = U/W.$$

□

完全系列の標語的な解釈

完全系列の特別な場合である短完全列

$$0 \longrightarrow W \xrightarrow{f_1} U \xrightarrow{f_2} V \longrightarrow 0$$

を次元公式の文脈で読んでみよう. 全射線形写像 $f_2: U \rightarrow V$ に関する次元公式により,

$$\begin{aligned} \dim U &= \dim \text{Ker } f_2 + \dim \text{Im } f_2 \\ &= \dim \text{Im } f_1 + \dim \text{Im } f_2 = \dim W + \dim V. \end{aligned}$$

この式から, 短完全列の中央に位置する空間 U が左右の空間 W および V によって分解されていると見れる. 実際, w_1, \dots, w_n を W の基底, v_1, \dots, v_m を V の基底とすれば, f_2 の全射性より $f(u_i) = v_i$ ($i = 1, \dots, m$) を満たす $u_i \in U$ が取れる. このとき, 定理 23.2.2 により $u_1, \dots, u_n, f_1(w_1), \dots, f_1(w_n)$ は U の基底である. すなわち, W の基底と V の基底を通して U の基底を与えることができる.

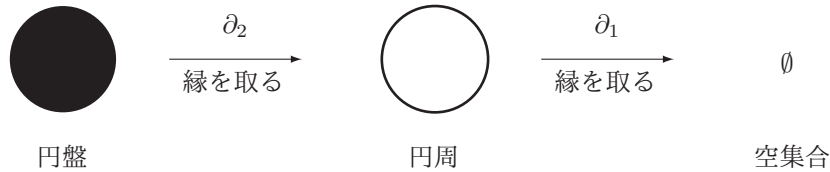
また, 例 24.3.3(2) および命題 24.3.4 によれば, 短完全列を与えることは商空間を与えることの言い換えに他ならない. 商空間を与えることは, 線形空間の各元を互いに交わらないグループに分類することを意味し, これは文字通りの空間の分解である. 短完全列から導かれる分解を荒っぽく述べれば, V の元の個数ぶんの W と合同な図形によって U は分解されている.

以上のことから, 短完全列は代数構造の分解を表す図式であることが分かる. 一般の完全系列についても, 短完全列ほどの単純明快さはないものの, ある種の分解を与えていると考えられよう. 実際, 複雑な現象をより単純な対象に分解して理解するという基本的な考え方のもとで, 代数学を援用する多くの数学分野において完全系列が扱われている.

よりみち (合成すると消える写像).

完全系列が満たすべき性質の一つに $f_{n+1} \circ f_n = \mathbf{0}$ がある. このように合成すると消えてしまう写像列の例について, ここで二つほど紹介しておこう.

ある空間上の基本的な図形に対して, その縁 (境界) を対応させる操作を考える. 一般に n 次元の図形の縁は $n-1$ 次元になる. つまり, n 次元の各図形 Δ に対して, その縁を $\partial_n(\Delta)$ と書けば, $\partial_n(\Delta)$ は $n-1$ 次元の図形である. 例えば円盤に対してこの操作を二回ほどこすと次のようになる:



他の多くの図形に対しても縁を対応させる操作を 2 回繰り返すと消えてしまうことが分かり, 空集合に対応する図形を $\mathbf{0}$ と書くとすれば, $\partial_{n-1} \circ \partial_n = \mathbf{0}$ を得る.

上の対応 ∂_n を線形代数の文脈に無理矢理持ち込むこともできる. 空間 X に配置された n 次元の各図形を基底とする線形空間を $C_n(X)$ とすると, 基底の間の写像を線形写像として拡張することにより (命題 20.1.11 あるいは 22.5.2), 線形写像 $\partial_n : C_n(X) \rightarrow C_{n-1}(X)$ を得る (これを境界作用素と呼ぶ). このとき線形写像の列

$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1}(X) \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \xrightarrow{\partial_{n-1}} \cdots$$

は $\partial_n \circ \partial_{n+1} = \mathbf{0}$ を満たす. しかしながら, 一般には上の列は完全系列にはならず (つまり $\text{Im } \partial_{n+1} \neq \text{Ker } \partial_n$), $\text{Im } \partial_{n+1} \subset \text{Ker } \partial_n$ であるに過ぎない. そこで, 上の写像の列がどれだけ完全系列から離れているかを調べる指標として商空間 $H_n(X) = \text{Ker } \partial_n / \text{Im } \partial_{n+1}$ が与えられる. $H_n(X)$ はホモロジー群と呼ばれ, 空間 X にどのくらい穴があいているかを計る量であることが知られている.

ところで, 合成すると消える写像のうち, 理工系学科の教育で最初に学ぶものといえば何であろうか. おそらくそれは電磁気学で学ぶ次の式である.

$$\text{rot} \circ \text{grad } f = \mathbf{0}, \quad \text{div} \circ \text{rot } F = \mathbf{0}. \quad (24.3.3)$$

ここで, 勾配ベクトル場 $\text{grad } f$ およびベクトル場の回転 $\text{rot } F$, ベクトル場の発散 $\text{div } g$ は次で定められるのであった:

$$\begin{aligned} \text{grad } f &:= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right), \\ \text{rot } F &:= \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right), \\ \text{div } g &:= \frac{\partial g_1}{\partial x} + \frac{\partial g_2}{\partial y} + \frac{\partial g_3}{\partial z}, \end{aligned}$$

ただし, 上に現れる関数はすべて変数 $(x, y, z) \in \mathbb{R}^3$ をパラメータとし, f は C^∞ -級関数, $F = (f_1, f_2, f_3)$ および $g = (g_1, g_2, g_3)$ は C^∞ -級ベクトル場である. 実は, これらの写像は図形の縁を対応させる写像とも無縁ではない. 勾配および回転, 発散は外微分と呼ばれる線形写像によって統一的に記述されることを後にベクトル解析を通して学ぶであろう. そして, 境界作用素と外微分の関係を双対概念と関連づけて理解することになる (ド・ラームの定理). これ以上詳しいことは, 多様体論および微分形式の専門書を参照されたい.

練習 24.3.5. 上の定義をもとに式 (24.3.3) を確認せよ.

ヒント: C^2 -級関数ゆえ偏微分の順序交換ができること (例えば $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$) を用いよ.

第25章 線形結合の行列表示

n 次元線形空間 U はユークリッド空間 \mathbb{R}^n と同型であったゆえ、 U において述べられる線形空間に関する現象はすべて線形同型 $T: U \rightarrow \mathbb{R}^n$ を通して \mathbb{R}^n の現象に書き換えられるはずである。本章では、 U の上での線形結合に関する情報を \mathbb{R}^n の情報に読みかえる技術について解説する。

25.1 線形結合の組と行列

ベクトルの間の線形結合を繰り返し計算し続けると \sum 記号が二重三重に現れ、読む側には非常に複雑に見えてしまう。これを回避する手段として、ベクトルの線形結合を行列を用いて表す記法を導入しよう。

\mathbb{R}^m の n 個のベクトルの組 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ について $[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ は (m, n) -行列ゆえ (n, ℓ) -行列 $A = [a_{ij}]$ との積を取ることができる:

$$[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1\ell} \\ a_{21} & a_{22} & \dots & a_{2\ell} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\ell} \end{bmatrix} = \left[\sum_{i=1}^n a_{i1} \mathbf{u}_i \quad \sum_{i=1}^n a_{i2} \mathbf{u}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell} \mathbf{u}_i \right].$$

上の右辺には A の第 j 列成分を係数とする $\mathbf{u}_1, \dots, \mathbf{u}_n$ による線形結合の ℓ 個の組が現れている。そこで、 \mathbb{R}^m の元とは限らない一般の線形空間 U におけるベクトルを並べた列¹ $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ と (n, ℓ) -行列 $A = [a_{ij}]$ に対して、これらの積を上と同様に定める。

$$[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1\ell} \\ a_{21} & a_{22} & \dots & a_{2\ell} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\ell} \end{bmatrix} := \left[\sum_{i=1}^n a_{i1} \mathbf{v}_i \quad \sum_{i=1}^n a_{i2} \mathbf{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell} \mathbf{v}_i \right]. \quad (25.1.1)$$

より抽象的に書けば、

$$[\mathbf{v}_h]_{h=1, \dots, n} [a_{ij}]_{\substack{i=1, \dots, n, \\ j=1, \dots, \ell}} = \left[\sum_{i=1}^n a_{ij} \mathbf{v}_i \right]_{j=1, \dots, \ell}.$$

式 (25.1.1) における $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ および右辺は U 上のベクトルを並べた組であって、 $U = \mathbb{R}^m$ でない限りこれらは行列ではないことに注意せよ。また、 A が列ベクトルである場合は、ベクトルの組と A の積は線形結合を意味する²:

$$[\mathbf{u}_1, \dots, \mathbf{u}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^n x_i \mathbf{u}_i.$$

例 25.1.1. (1) $[\mathbf{v}_1, \mathbf{v}_2] \begin{bmatrix} 3 & 2 & 1 \\ 1 & -1 & 4 \end{bmatrix} = [3\mathbf{v}_1 + \mathbf{v}_2 \quad 2\mathbf{v}_1 - \mathbf{v}_2 \quad \mathbf{v}_1 + 4\mathbf{v}_2].$

(2) $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] E = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n].$

¹ベクトルを並べる際にカンマで区切るかどうかはこだわらないことにする。

² $\sum_{i=1}^n \mathbf{u}_i x_i$ と書いてもよいが、慣例ではスカラー係数を左側に書く。このような事情から、ベクトルのスカラー係数を右側に書く流儀もある。

ベクトルの組における和とスカラー倍も定めておこう:

- $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] + [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n] := [\mathbf{v}_1 + \mathbf{w}_1, \mathbf{v}_2 + \mathbf{w}_2, \dots, \mathbf{v}_n + \mathbf{w}_n]$.
- $r[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] := [r\mathbf{v}_1, r\mathbf{v}_2, \dots, r\mathbf{v}_n]$.

25.2 ベクトルの組と行列の演算の基本性質 (付録)

前節で定めたベクトルの組と行列の積が分配法則や結合法則, スカラー律を満たすことを退屈ではあるが確認しておこう. 実際の計算では, これらの性質は意識することなく用いられるであろう.

命題 25.2.1. 線形空間 V のベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n$, および (n, ℓ) -行列 A , $r \in \mathbb{R}$ に対して

$$[\mathbf{v}_1, \dots, \mathbf{v}_n](rA) = (r[\mathbf{v}_1, \dots, \mathbf{v}_n])A = r([\mathbf{v}_1, \dots, \mathbf{v}_n]A).$$

Proof. $A = [a_{ij}]$ とおくと $rA = [ra_{ij}]$ であり,

$$\begin{aligned} (\text{左辺}) &= \left[\sum_{i=1}^n ra_{i1}\mathbf{v}_i \quad \sum_{i=1}^n ra_{i2}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n ra_{i\ell}\mathbf{v}_i \right] \\ &= \left[\sum_{i=1}^n a_{i1}(r\mathbf{v}_i) \quad \sum_{i=1}^n a_{i2}(r\mathbf{v}_i) \quad \dots \quad \sum_{i=1}^n a_{i\ell}(r\mathbf{v}_i) \right] \quad (\text{この式は上の中央の式に相当する}) \\ &= \left[r \sum_{i=1}^n a_{i1}\mathbf{v}_i \quad r \sum_{i=1}^n a_{i2}\mathbf{v}_i \quad \dots \quad r \sum_{i=1}^n a_{i\ell}\mathbf{v}_i \right] \\ &= r \left[\sum_{i=1}^n a_{i1}\mathbf{v}_i \quad \sum_{i=1}^n a_{i2}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell}\mathbf{v}_i \right] = (\text{右辺}). \end{aligned}$$

□

命題 25.2.2. 線形空間 V のベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n$, および (n, ℓ) -行列 A, B に対して

$$[\mathbf{v}_1, \dots, \mathbf{v}_n]A + [\mathbf{v}_1, \dots, \mathbf{v}_n]B = [\mathbf{v}_1, \dots, \mathbf{v}_n](A + B).$$

Proof. $A = [a_{ij}]$, $B = [b_{ij}]$ とおくと,

$$\begin{aligned} (\text{左辺}) &= \left[\sum_{i=1}^n a_{i1}\mathbf{v}_i \quad \sum_{i=1}^n a_{i2}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell}\mathbf{v}_i \right] + \left[\sum_{i=1}^n b_{i1}\mathbf{v}_i \quad \sum_{i=1}^n b_{i2}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n b_{i\ell}\mathbf{v}_i \right] \\ &= \left[\sum_{i=1}^n a_{i1}\mathbf{v}_i + \sum_{i=1}^n b_{i1}\mathbf{v}_i \quad \sum_{i=1}^n a_{i2}\mathbf{v}_i + \sum_{i=1}^n b_{i2}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell}\mathbf{v}_i + \sum_{i=1}^n b_{i\ell}\mathbf{v}_i \right] \\ &= \left[\sum_{i=1}^n (a_{i1} + b_{i1})\mathbf{v}_i \quad \sum_{i=1}^n (a_{i2} + b_{i2})\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n (a_{i\ell} + b_{i\ell})\mathbf{v}_i \right] = (\text{右辺}). \end{aligned}$$

□

命題 25.2.3. 線形空間 V のベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n$, および (n, ℓ) -行列 A , (ℓ, r) -行列 B に対して

$$([\mathbf{v}_1, \dots, \mathbf{v}_n]A)B = [\mathbf{v}_1, \dots, \mathbf{v}_n](AB).$$

Proof. $A = [a_{ij}]$, $B = [b_{jk}]$ とおけば, (n, r) -行列 $AB = [z_{ik}]$ の各成分は行列の積の定義より $z_{ik} =$

$\sum_{j=1}^{\ell} a_{ij}b_{jk}$ である。

$$\begin{aligned}
 (\text{左辺}) &= \left[\sum_{i=1}^n a_{i1} \mathbf{v}_i \quad \cdots \quad \sum_{i=1}^n a_{i\ell} \mathbf{v}_i \right] B \\
 &= \left[\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_\ell \right] B \quad (\mathbf{u}_j := \sum_{i=1}^n a_{ij} \mathbf{v}_i \ (j = 1, \dots, \ell) \text{とおいた}) \\
 &= \left[\sum_{j=1}^{\ell} b_{j1} \mathbf{u}_j \quad \cdots \quad \sum_{j=1}^{\ell} b_{jr} \mathbf{u}_j \right] = \left[\sum_{j=1}^{\ell} b_{jk} \mathbf{u}_j \right]_{k=1, \dots, r} = \left[\sum_{j=1}^{\ell} b_{jk} \left(\sum_{i=1}^n a_{ij} \mathbf{v}_i \right) \right]_{k=1, \dots, r} \\
 &= \left[\sum_{j=1}^{\ell} \left(\sum_{i=1}^n a_{ij} b_{jk} \mathbf{v}_i \right) \right]_{k=1, \dots, r} = \left[\sum_{i=1}^n \left(\sum_{j=1}^{\ell} a_{ij} b_{jk} \mathbf{v}_i \right) \right]_{k=1, \dots, r} \quad (\text{例 3.2.4 を用いた}) \\
 &= \left[\sum_{i=1}^n \left(\sum_{j=1}^{\ell} a_{ij} b_{jk} \right) \mathbf{v}_i \right]_{k=1, \dots, r} = \left[\sum_{i=1}^n z_{ik} \mathbf{v}_i \right]_{k=1, \dots, r} = [\mathbf{v}_1, \dots, \mathbf{v}_n][z_{ik}] = (\text{右辺}).
 \end{aligned}$$

□

25.3 線形結合再考

線形結合に関するいくつかの基本的な性質について、25.1 節で定めた行列による記法を用いる立場から再考する。例えば、命題 18.1.3 は次のように示される。

命題 18.1.3 (再掲). 各 $\mathbf{u}_1, \dots, \mathbf{u}_r$ が組 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で表され、また各 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ が組 $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で表されるならば、各 $\mathbf{u}_1, \dots, \mathbf{u}_r$ は $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で表される。

Proof. 各 $\mathbf{u}_1, \dots, \mathbf{u}_r$ が $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で表されることから、 (ℓ, r) -行列 B を用いて $[\mathbf{u}_1, \dots, \mathbf{u}_r] = [\mathbf{v}_1, \dots, \mathbf{v}_\ell]B$ と書ける。また、各 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ が $\mathbf{w}_1, \dots, \mathbf{w}_n$ 個の線形結合で表されることから、 (n, ℓ) -行列 A を用いて $[\mathbf{v}_1, \dots, \mathbf{v}_\ell] = [\mathbf{w}_1, \dots, \mathbf{w}_n]A$ と書ける。ゆえに、

$$[\mathbf{u}_1, \dots, \mathbf{u}_r] = [\mathbf{v}_1, \dots, \mathbf{v}_\ell]B = ([\mathbf{w}_1, \dots, \mathbf{w}_n]A)B = [\mathbf{w}_1, \dots, \mathbf{w}_n](AB).$$

この式は各 $\mathbf{u}_1, \dots, \mathbf{u}_r$ が $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で表されることを意味している。

□

いまの証明から、 \mathbf{u}_k ($k = 1, \dots, r$) を $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で書くときに現れる係数は、行列 AB の第 k 列成分に等しいことが分かる。

次は、各成分ごとに命題 18.2.3 を主張するものに他ならない。

練習 25.3.1. $\mathbf{v}_1, \dots, \mathbf{v}_m \in U$ が線形独立であるとし、 A, B を (m, n) -行列とする。このとき、 $[\mathbf{v}_1, \dots, \mathbf{v}_m]A = [\mathbf{v}_1, \dots, \mathbf{v}_m]B$ ならば $A = B$ となることを示せ。

解答例: 移項して $[\mathbf{v}_1, \dots, \mathbf{v}_m](A - B) = [\mathbf{0}_V, \dots, \mathbf{0}_V]$. $A - B = [x_{ij}]$ とおき、各成分を比較すれば $j = 1, \dots, n$ について $\sum_{i=1}^m x_{ij} \mathbf{v}_i = \mathbf{0}_V$ であり、 $\mathbf{v}_1, \dots, \mathbf{v}_m$ の線形独立性より $x_{ij} = 0$, つまり $A - B = O$ を得る。

□

次は、各成分ごとに線形性 (iii)' を並べた式である。

練習 25.3.2. A を (m, n) -行列、 $f: U \rightarrow W$ を線形写像とする。 $\mathbf{u}_j, \mathbf{v}_i \in U$ について次を示せ。

$$[\mathbf{u}_1, \dots, \mathbf{u}_n] = [\mathbf{v}_1, \dots, \mathbf{v}_m]A \implies [f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [f(\mathbf{v}_1), \dots, f(\mathbf{v}_m)]A.$$

解答例: $A = [a_{ij}]$ とすれば、仮定より、

$$[\mathbf{u}_1, \dots, \mathbf{u}_n] = \left[\sum_{i=1}^m a_{i1} \mathbf{v}_i \quad \cdots \quad \sum_{i=1}^m a_{in} \mathbf{v}_i \right] \quad \text{であるから,}$$

$$\begin{aligned}
[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] &= \left[f\left(\sum_{i=1}^m a_{i1}\mathbf{v}_i\right) \ \dots \ f\left(\sum_{i=1}^m a_{in}\mathbf{v}_i\right) \right] \\
&= \left[\sum_{i=1}^m a_{i1}f(\mathbf{v}_i) \ \dots \ \sum_{i=1}^m a_{in}f(\mathbf{v}_i) \right] = [f(\mathbf{v}_1), \dots, f(\mathbf{v}_m)]A.
\end{aligned}$$

□

次の命題は、命題 17.3.5 を一般の線形空間に拡張したものである。

命題 25.3.3. $\mathbf{v}_1, \dots, \mathbf{v}_n \in \langle \mathbf{u}_1, \dots, \mathbf{u}_m \rangle$ とする。このとき、 $m < n$ ならば $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形従属である。

Proof. 仮定より (m, n) -行列 A を用いて $[\mathbf{v}_1, \dots, \mathbf{v}_n] = [\mathbf{u}_1, \dots, \mathbf{u}_m]A$ と書ける。 $n > m$ より連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ は自明でない解 ${}^t[c_1, \dots, c_n] \neq \mathbf{0}$ を持つ。このとき、

$$\begin{aligned}
c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n &= [\mathbf{v}_1, \dots, \mathbf{v}_n] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = [\mathbf{u}_1, \dots, \mathbf{u}_m]A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \\
&= [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = 0\mathbf{u}_1 + \dots + 0\mathbf{u}_m = \mathbf{0}.
\end{aligned}$$

自明でない線形関係 $\sum_{j=1}^n c_j\mathbf{v}_j = \mathbf{0}$ があるゆえ $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形従属である。

□

練習 25.3.4. 線形空間 U の基底におけるベクトルの個数が基底の取り方によらないこと (命題 22.1.1) を前命題を用いて示せ。

解答例: $\mathbf{u}_1, \dots, \mathbf{u}_m$ および $\mathbf{v}_1, \dots, \mathbf{v}_n$ をそれぞれ U の基底とする。仮に $m < n$ とすれば前命題より $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形従属になってしまう。ゆえに $m \geq n$ でなければならない。 $\mathbf{u}_1, \dots, \mathbf{u}_m$ と $\mathbf{v}_1, \dots, \mathbf{v}_n$ の立場を入れ替えることで $n \geq m$ も示され、したがって $m = n$ 。

□

25.4 線形独立性の判定 (2)

一般のベクトル空間 U における線形結合を行列表示することで、 U における線形関係をユークリッド空間上の線形関係に読み替えられることを見よう。 $\mathbf{u}_1, \dots, \mathbf{u}_m$ を U の基底とし、 $F(\mathbf{u}_j) = \mathbf{e}_j$ なる線形同型 $F: U \rightarrow \mathbb{R}^m$ を取る。

$$\begin{array}{ccccccc}
U & : & \mathbf{u}_1, & \mathbf{u}_2, & \dots, & \mathbf{u}_m & \mathbf{v} = \sum_{i=1}^m r_i\mathbf{u}_i \\
F \downarrow & & \downarrow & \downarrow & & \downarrow & \downarrow \\
\mathbb{R}^m & : & \mathbf{e}_1, & \mathbf{e}_2, & \dots, & \mathbf{e}_m & \sum_{i=1}^m r_i\mathbf{e}_i = {}^t[r_1, \dots, r_m]
\end{array}$$

U において調べたいベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n \in U$ が与えられたとき、これに対応する $F(\mathbf{v}_1), \dots, F(\mathbf{v}_n) \in \mathbb{R}^m$ を調べるというのが基本方針である。いま $\mathbf{u}_1, \dots, \mathbf{u}_m$ は基底であるから各 $\mathbf{v}_1, \dots, \mathbf{v}_n$ は $\mathbf{u}_1, \dots, \mathbf{u}_m$ の線形結合で書ける:

$$\mathbf{v}_1 = [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{v}_2 = [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots, \quad \mathbf{v}_n = [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

これをまとめて書くために、 (m, n) -行列 $A = [a_{ij}] = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ を用いれば、

$$[\mathbf{v}_1, \dots, \mathbf{v}_n] = [\mathbf{u}_1, \dots, \mathbf{u}_m] \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right].$$

このとき、 $[F(\mathbf{v}_1), \dots, F(\mathbf{v}_n)] = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ である。実際、上式に F をほどこせば、練習 25.3.2 により

$$[F(\mathbf{v}_1), \dots, F(\mathbf{v}_n)] = [F(\mathbf{u}_1), \dots, F(\mathbf{u}_m)]A = EA = A = [\mathbf{a}_1, \dots, \mathbf{a}_n].$$

命題 20.3.6 より $\mathbf{v}_1, \dots, \mathbf{v}_n$ の線形関係と $F(\mathbf{v}_1), \dots, F(\mathbf{v}_n)$ (すなわち $\mathbf{a}_1, \dots, \mathbf{a}_n$) の線形関係は同等である。つまり、各 $\mathbf{w} \in U$ および $s_1, \dots, s_n \in \mathbb{R}$ について

$$\mathbf{w} = \sum_{j=1}^n s_j \mathbf{v}_j \iff F(\mathbf{w}) = \sum_{j=1}^n s_j \mathbf{a}_j. \quad (25.4.1)$$

が成り立つ。とくに、 $n = m$ の場合は A は正方行列であり、定理 17.3.6 により $\mathbf{a}_1, \dots, \mathbf{a}_n$ の線形独立性 (したがって $\mathbf{v}_1, \dots, \mathbf{v}_n$ の線形独立性) は A の可逆性に帰着される。以上の事実をまとめると:

命題 25.4.1. $\mathbf{u}_1, \dots, \mathbf{u}_m$ を U の基底とする。ベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n \in U$ について $[\mathbf{v}_1, \dots, \mathbf{v}_n] = [\mathbf{u}_1, \dots, \mathbf{u}_m]A$ とすれば次が成り立つ。

(1) $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ とおけば、 $\sum_{i=1}^n c_i \mathbf{v}_i = \mathbf{0}_U \iff \sum_{i=1}^n c_i \mathbf{a}_i = \mathbf{0}$.

(2) $n = m$ のとき次が成り立つ:

$$\mathbf{v}_1, \dots, \mathbf{v}_n \text{ は線形独立 (したがって } U \text{ の基底である)} \iff A \text{ は可逆.}$$

【補足】 (2) の \Leftarrow において $\mathbf{v}_1, \dots, \mathbf{v}_n$ が U を生成することは命題 22.3.3 から直ちに分かることであるが、次のように直接示すこともできる:

$[\mathbf{v}_1, \dots, \mathbf{v}_n] = [\mathbf{u}_1, \dots, \mathbf{u}_n]A$ の両辺に右から A^{-1} を掛ければ $[\mathbf{u}_1, \dots, \mathbf{u}_n] = [\mathbf{v}_1, \dots, \mathbf{v}_n]A^{-1}$. したがって、基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の各ベクトルは $\mathbf{v}_1, \dots, \mathbf{v}_n$ の線形結合で書けるゆえ、命題 18.1.3 より $\mathbf{v}_1, \dots, \mathbf{v}_n$ は U を生成する。

本節で述べた事実を用いて、一般の線形空間におけるベクトルの組の線形独立性を判定できる:

例題 25.4.2. $\mathbf{u}_1, \dots, \mathbf{u}_5$ を U の基底とする。次で与えるベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_5$ が線形独立かどうか答えよ。また、これらのベクトルで生成される部分空間 $\langle \mathbf{v}_1, \dots, \mathbf{v}_5 \rangle$ の基底を求めよ。

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1 - \mathbf{u}_3 + 2\mathbf{u}_4 + \mathbf{u}_5, & \mathbf{v}_2 &= \mathbf{u}_2 + \mathbf{u}_4 + \mathbf{u}_5, & \mathbf{v}_3 &= -\mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3 - \mathbf{u}_4, \\ \mathbf{v}_4 &= \mathbf{u}_3, & \mathbf{v}_5 &= -2\mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3 - 3\mathbf{u}_4 - \mathbf{u}_5. \end{aligned}$$

解答例: 上の線形結合を行列を用いた表示に書き直せば,

$$[\mathbf{v}_1, \dots, \mathbf{v}_5] = [\mathbf{u}_1, \dots, \mathbf{u}_5] \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 & -3 \\ 1 & 1 & 0 & 0 & -1 \end{bmatrix}.$$

上式に現れる行列は例題 17.3.3 における $A = [\mathbf{a}_1, \dots, \mathbf{a}_5]$ に一致している。 $\mathbf{a}_1, \dots, \mathbf{a}_5$ が線形独立でないことは例題 17.3.3 の通りであり、したがって、命題 25.4.1(1) より $\mathbf{v}_1, \dots, \mathbf{v}_5$ は線形独立でない。

また、例題 17.3.3 によれば $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4$ は線形独立であり、これ以外のベクトルは $\mathbf{a}_3 = -\mathbf{a}_1 + \mathbf{a}_2$, $\mathbf{a}_5 = -2\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_4$ と書ける。したがって、再び命題 25.4.1(1) より $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4$ は線形独立であり、更に式 (25.4.1) から $\mathbf{v}_3 = -\mathbf{v}_1 + \mathbf{v}_2$, $\mathbf{v}_5 = -2\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_4$ と書けることが分かる。以上により、 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4$ は $\langle \mathbf{v}_1, \dots, \mathbf{v}_5 \rangle$ の基底である。

25.5 基底の変換行列

前節にて、 U 上の各ベクトル \mathbf{v} を数値化する方法を述べた。ここで注意すべきは、数値化された情報は線形同型 F の取り方に依存すること、言い換えれば U の基底の取り方に依存することである。基底を取

りかえるたびに以前の基底による情報が役に立たなくなるようでは甚だ不便であるから、基底の取り換えにより v の数値化がどのように変化するかを調べておこう。

いま、 U は n 次元であるとし、 U において二組の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ が与えられているとする。このとき、各 $\mathbf{u}'_1, \dots, \mathbf{u}'_n \in U$ は基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書ける。これを式で書けば：

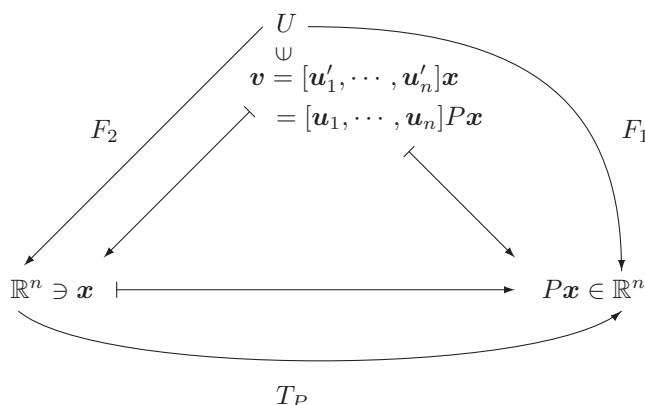
$$\begin{aligned} [\mathbf{u}'_1, \mathbf{u}'_2, \dots, \mathbf{u}'_n] &= \left[\sum_{i=1}^n a_{i1} \mathbf{u}_i, \sum_{i=1}^n a_{i2} \mathbf{u}_i, \dots, \sum_{i=1}^n a_{in} \mathbf{u}_i \right] \\ &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}. \end{aligned}$$

定義 25.5.1. 上式に現れる行列 $P = [a_{ij}]$ を基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列³という。

備考 25.5.2. 命題 25.4.1(2) により変換行列 P は可逆である。そこで $[\mathbf{u}'_1, \mathbf{u}'_2, \dots, \mathbf{u}'_n] = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]P$ の両辺に右から P^{-1} をかけることで $[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] = [\mathbf{u}'_1, \mathbf{u}'_2, \dots, \mathbf{u}'_n]P^{-1}$ を得る。すなわち、基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ による基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の変換行列は P^{-1} である。

例 25.5.3. $\mathbf{w}_1, \dots, \mathbf{w}_n \in \mathbb{R}^n$ を基底とする。標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ による基底 $\mathbf{w}_1, \dots, \mathbf{w}_n$ の変換行列は、 $[\mathbf{w}_1, \dots, \mathbf{w}_n]$ である。

備考 25.5.4. 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P とする。 $F_1 : U \rightarrow \mathbb{R}^n$ を \mathbf{u}_i を \mathbf{e}_i に写す線形同型とし、 $F_2 : U \rightarrow \mathbb{R}^n$ を \mathbf{u}'_i を \mathbf{e}_i に写す線形同型とする。



各ベクトル $\mathbf{v} = [\mathbf{u}'_1, \dots, \mathbf{u}'_n]\mathbf{x} \in U$ (ただし $\mathbf{x} \in \mathbb{R}^n$) に対応する \mathbb{R}^n の元は、同型 F_2 を用いれば $F_2(\mathbf{v}) = \mathbf{x}$ であり、同型 F_1 を用いれば、 $\mathbf{v} = [\mathbf{u}'_1, \dots, \mathbf{u}'_n]\mathbf{x} = [\mathbf{u}_1, \dots, \mathbf{u}_n]P\mathbf{x}$ より $F_1(\mathbf{v}) = P\mathbf{x}$ である。このとき、

$$F_1 = T_P \circ F_2, \quad T_P = F_1 \circ F_2^{-1}, \quad F_2 = T_{P^{-1}} \circ F_1$$

が成り立つことは上の図式から直ちに分かる。

練習 25.5.5. 上の備考において $T_P = F_1 \circ F_2^{-1}$ であることを確認せよ。

解答例: 各標準基底 $\mathbf{e}_j \in \mathbb{R}^n$ について $F_1 \circ F_2^{-1}(\mathbf{e}_j) = T_P(\mathbf{e}_j)$ であることを確かめればよい。 $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ と置けば、変換行列の定義から $\mathbf{u}'_j = [\mathbf{u}_1, \dots, \mathbf{u}_n]\mathbf{p}_j$ である。したがって、

$$\begin{aligned} F_1 \circ F_2^{-1}(\mathbf{e}_j) &= F_1(\mathbf{u}'_j) = F_1([\mathbf{u}_1, \dots, \mathbf{u}_n]\mathbf{p}_j) = [F_1(\mathbf{u}_1), \dots, F_1(\mathbf{u}_n)]\mathbf{p}_j \\ &= [\mathbf{e}_1, \dots, \mathbf{e}_n]\mathbf{p}_j = E\mathbf{p}_j = \mathbf{p}_j = P\mathbf{e}_j = T_P(\mathbf{e}_j). \end{aligned}$$

□

³「変換行列」を英語に直訳すると transformation matrix となるものの、英文献におけるこの語は、線形変換 $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ の標準基底に関する表現行列のことを指す(詳しくは例 26.1.2 および系 26.2.2(2) をみよ)。また、備考 25.5.2 における基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ による基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の変換行列 P^{-1} のことを、“transition matrix (あるいは change-of-basis matrix) from $\mathbf{u}_1, \dots, \mathbf{u}_n$ to $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ ” と言い表す。このとき、 P は “transition matrix from $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ to $\mathbf{u}_1, \dots, \mathbf{u}_n$ ” である。

第26章 線形写像の表現行列

線形写像 $f: U \rightarrow V$ をユークリッド空間の間の線形写像 (すなわち行列) に対応させる方法について述べる. 前章の対応と同様に, ここで与える対応も基底の取り方に依存することに注意しなければならない.

26.1 表現行列

U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$, V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ をあらかじめ与えておく. 線形写像 $T: U \rightarrow V$ を (m, n) -行列 A による線形写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ に翻訳する方法を考えよう. $\mathbf{v}_1, \dots, \mathbf{v}_m$ が基底であることから, 各 $T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)$ は $\mathbf{v}_1, \dots, \mathbf{v}_m$ の線形結合で書ける:

$$T(\mathbf{u}_1) = [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T(\mathbf{u}_2) = [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots, \quad T(\mathbf{u}_n) = [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

つまり, (m, n) -行列 $A = [a_{ij}]$ を用いてまとめて書けば,

$$[T(\mathbf{u}_1), T(\mathbf{u}_2), \dots, T(\mathbf{u}_n)] = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}. \quad (26.1.1)$$

定義 26.1.1. 上の設定における式 (26.1.1) を満たす行列 $A = [a_{ij}]$ を, U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ に関する $T: U \rightarrow V$ の表現行列¹ という. A の定義を次のように言い換えても良い:

$$A \text{ の第 } j \text{ 列} = \text{線形結合 } T(\mathbf{u}_j) = \sum_{i=1}^m a_{ij} \mathbf{v}_i \text{ に現れる係数を並べた列ベクトル.}$$

注意: 練習 25.3.1 により, 式 (26.1.1) を満たす行列 A は唯一つ存在する.

例 26.1.2. (m, n) -行列 B について, \mathbb{R}^n の標準基底および \mathbb{R}^m の標準基底に関する $T_B: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($T_B(\mathbf{x}) := B\mathbf{x}$) の表現行列は B に一致する.

例 26.1.3. (1) $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とする. 定義域の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および終域の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する恒等写像 $\text{id}_U: U \rightarrow U$ の表現行列は単位行列 E である.

(2) $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とし, $\mathbf{v}_1, \dots, \mathbf{v}_n$ を V の基底とする. 各 \mathbf{u}_i を \mathbf{v}_i に対応させる線形同型 $f: U \rightarrow V$ について, U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ に関する f の表現行列は単位行列 E である.

例題 26.1.4. 次の線形写像および基底について, 表現行列を求めよ.

¹英文献ではこれを representation matrix とは呼ばない. この行列を “the matrix of T with respect to the bases $\mathbf{u}_1, \dots, \mathbf{u}_n$ and $\mathbf{v}_1, \dots, \mathbf{v}_m$ ” と言います.

(1) $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ($T_A(\mathbf{x}) := A\mathbf{x}$), ただし $A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$.

基底: 定義域と終域ともに $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ とせよ.

解答例:

$$\begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} = -2\mathbf{v}_1 + 0\mathbf{v}_2, \quad \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 0\mathbf{v}_1 + 3\mathbf{v}_2.$$

ゆえに $[T_A(\mathbf{v}_1), T_A(\mathbf{v}_2)] = [\mathbf{v}_1, \mathbf{v}_2] \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$ であり, T_A の表現行列は $\begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$.

(2) $D : \mathbb{R}[x]_3 \rightarrow \mathbb{R}[x]_3$ ($D(p(x)) := \frac{d}{dx}p(x)$). 基底: 定義域と終域ともに $\mathbf{1}, x, x^2, x^3$ とせよ.

解答例:

$$D(\mathbf{1}) = 0\mathbf{1} + 0x + 0x^2 + 0x^3, \quad D(x) = 1 \cdot \mathbf{1} + 0x + 0x^2 + 0x^3, \\ D(x^2) = 0\mathbf{1} + 2x + 0x^2 + 0x^3, \quad D(x^3) = 0\mathbf{1} + 0x + 3x^2 + 0x^3.$$

ゆえに

$$[D(\mathbf{1}), D(x), D(x^2), D(x^3)] = [\mathbf{1}, x, x^2, x^3] \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

であり, 求める表現行列は $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

【備考】 上の行列を A とおけば, $A^4 = O$ である. これは 3 次多項式を 4 回微分すると必ず $\mathbf{0}$ になることに対応している.

(3) $I_{a,b} : \mathbb{R}[x]_3 \rightarrow \mathbb{R}$ ($I_{a,b}(p(x)) := \int_a^b p(x) dx$). ただし $a, b \in \mathbb{R}$ は定数とする.

$\mathbb{R}[x]_3$ の基底: $x^3, x^2, x, \mathbf{1}$. \mathbb{R} の基底: 1 .

解答例:

$$I_{a,b}(x^3) = \left[\frac{1}{4}x^4 \right]_a^b = \frac{1}{4}(b^4 - a^4), \quad I_{a,b}(x^2) = \left[\frac{1}{3}x^3 \right]_a^b = \frac{1}{3}(b^3 - a^3), \\ I_{a,b}(x) = \left[\frac{1}{2}x^2 \right]_a^b = \frac{1}{2}(b^2 - a^2), \quad I_{a,b}(\mathbf{1}) = \left[x \right]_a^b = (b - a).$$

したがって

$$[I_{a,b}(x^3), I_{a,b}(x^2), I_{a,b}(x), I_{a,b}(\mathbf{1})] = [1] \left[\frac{1}{4}(b^4 - a^4), \frac{1}{3}(b^3 - a^3), \frac{1}{2}(b^2 - a^2), b - a \right]$$

であり, 求める表現行列は (1, 4)-行列 $\left[\frac{1}{4}(b^4 - a^4), \frac{1}{3}(b^3 - a^3), \frac{1}{2}(b^2 - a^2), b - a \right]$.

【備考】 とくに $I_{a,b}(c_1x^3 + c_2x^2 + c_3x + c_4\mathbf{1}) = c_1I_{a,b}(x^3) + c_2I_{a,b}(x^2) + c_3I_{a,b}(x) + c_4I_{a,b}(\mathbf{1}) = \frac{c_1}{4}(b^4 - a^4) + \frac{c_2}{3}(b^3 - a^3) + \frac{c_3}{2}(b^2 - a^2) + c_4(b - a)$. 実際に積分計算を行う際も, このように各項ごとに積分するのであった.

例 26.1.5 (発展). 例 18.4.6 にて与えた, 数列空間 $\mathbb{R}^{\mathbb{N}}$ における線形独立な無限部分集合 $B = \{e_n | n \in \mathbb{N}\}$ について, $V = \langle B \rangle$ と定める. 次の二つのシフト作用素

$$\begin{aligned} S_+ : V &\rightarrow V & S_+(x_1, x_2, x_3, \dots) &:= (0, x_1, x_2, x_3, x_4, \dots) \\ S_- : V &\rightarrow V & S_-(x_1, x_2, x_3, \dots) &:= (x_2, x_3, x_4, x_5, x_6, \dots) \end{aligned}$$

の基底 B に関する表現行列は, 13.2 節のコラムで与えた A と B にそれぞれ等しい:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}.$$

$S_- \circ S_+ = \text{id}_V$ ゆえ BA は単位行列となる. 一方, $S_+ \circ S_-(x_1, x_2, x_3, \dots) = (0, x_2, x_3, \dots)$ ゆえ $S_+ \circ S_- \neq \text{id}_V$. つまり AB は単位行列ではない. このようなことが起きる背景には, 基底が無限集合であること (つまり線形空間が無限次元であること), そして有限集合の場合 (命題 19.5.1) とは異なり, 無限集合 B から B 自身への単射 (あるいは全射) が全単射になるとは限らないという事情が関係している. 実際, 基底の間の写像 $S_+|_B : B \rightarrow B$ は単射であり, $S_-|_B : B \cup \{\mathbf{0}\} \rightarrow B \cup \{\mathbf{0}\}$ は全射であるが, これらはいずれも全単射ではない. そして, 写像 $S_+|_B$ および $S_-|_B$ を全体に拡張した線形写像がそれぞれ S_+ および S_- である. S_+ は単射であり S_- は全射であるが, これらはいずれも全単射ではない.

表現行列の捉え方

u_1, \dots, u_n を U の基底とし, v_1, \dots, v_m を V の基底とする. また, これらの基底に関する線形写像 $T : U \rightarrow V$ の表現行列を A とする (つまり $[T(u_1), \dots, T(u_n)] = [v_1, \dots, v_m]A$). $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ とすれば, T はベクトル $[u_1, \dots, u_n]x \in U$ をベクトル $[v_1, \dots, v_m]Ax \in V$ に写す写像である. 実際, 線形性 (iii)' より

$$T([u_1, \dots, u_n]x) = [T(u_1), \dots, T(u_n)]x = [v_1, \dots, v_m]Ax.$$

つまり, 次のような図式を得る:

$$\begin{array}{ccccc} U & \ni & [u_1, \dots, u_n]x & \xrightarrow{T} & [v_1, \dots, v_m]Ax \in V \\ F \downarrow & & \downarrow & & \downarrow G \\ \mathbb{R}^n & \ni & x & \xrightarrow{T_A} & Ax \in \mathbb{R}^n \end{array}$$

ここで, $F : U \rightarrow \mathbb{R}^n$ は $F(u_j) = e_j$ を満たす線形同型写像, $G : V \rightarrow \mathbb{R}^m$ は $G(v_i) = e_i$ を満たす線形同型写像である. u_1, \dots, u_n が U の基底であることから, U の任意の元は $[u_1, \dots, u_n]x$ の形で表せることに注意すれば, 上の図式から直ちに次を得る:

$$\text{各 } u \in U \text{ について, } T_A \circ F(u) = G \circ T(u), \quad \text{すなわち, } T_A \circ F = G \circ T.$$

$T_A \circ F = G \circ T$ が成立するとき, これを次のような図式で表す.

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ F \downarrow & & \downarrow G \\ \mathbb{R}^n & \xrightarrow{T_A} & \mathbb{R}^m \end{array} \tag{26.1.2}$$

このような図式は可換図式 (commutative diagram) と呼ばれる.

26.2 Hom(U, V) と $M_{m,n}(\mathbb{R})$

Hom($\mathbb{R}^n, \mathbb{R}^m$) と $M_{m,n}(\mathbb{R})$ が同型であることを 21.3 節にて既に見たが、これらが Hom(U, V) (ただし $\dim U = n, \dim V = m$) と同型になることも容易に示唆される。 U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ を与えておき、これらの基底による表現行列を対応させる写像 $\mathcal{R} : \text{Hom}(U, V) \rightarrow M_{m,n}(\mathbb{R})$ が線形同型を与えることを見よう。

各線形写像 $f : U \rightarrow V$ の表現行列を $\mathcal{R}(f)$ とする。つまり、 $\mathcal{R}(f)$ は

$$[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m] \mathcal{R}(f)$$

を満たす (m, n) -行列である。また、これとは逆に、各 (m, n) -行列 A に対して、写像 $\mathcal{T}(A) : U \rightarrow V$ を、

$$[\mathcal{T}(A)(\mathbf{u}_1), \dots, \mathcal{T}(A)(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m] A$$

を満たす線形写像と定める (このような線形写像は命題 20.1.11 により存在する)。次は定義より明らかであろう (しかし念のため証明を記す):

- 各 $A \in M_{m,n}(\mathbb{R})$ について $\mathcal{R}(\mathcal{T}(A)) = A$ (つまり $\mathcal{R} \circ \mathcal{T} = \text{id}_{M_{m,n}(\mathbb{R})}$)。

Proof. $\mathcal{T}(A) : U \rightarrow V$ は、 $[\mathcal{T}(A)(\mathbf{u}_1), \dots, \mathcal{T}(A)(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m] A$ を満たす線形写像であり、この写像 $\mathcal{T}(A)$ の表現行列は A である。すなわち、 $\mathcal{R}(\mathcal{T}(A)) = A$ 。 \square

- 各 $f \in \text{Hom}(U, V)$ について $\mathcal{T}(\mathcal{R}(f)) = f$ (つまり $\mathcal{T} \circ \mathcal{R} = \text{id}_{\text{Hom}(U, V)}$)。

Proof. 行列 $\mathcal{R}(f)$ は $[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m] \mathcal{R}(f)$ を満たす。一方、 $\mathcal{T}(\mathcal{R}(f)) : U \rightarrow V$ は $[\mathcal{T}(\mathcal{R}(f))(\mathbf{u}_1), \dots, \mathcal{T}(\mathcal{R}(f))(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m] \mathcal{R}(f)$ を満たす線形写像であり、これらを合わせれば、 $[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathcal{T}(\mathcal{R}(f))(\mathbf{u}_1), \dots, \mathcal{T}(\mathcal{R}(f))(\mathbf{u}_n)]$ 。 $\mathcal{T}(\mathcal{R}(f))$ と f に U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を代入した値はそれぞれ等しいゆえ、 $\mathcal{T}(\mathcal{R}(f)) = f$ (命題 20.1.10(2))。 \square

したがって、命題 19.3.3 より \mathcal{R}, \mathcal{T} は全単射かつ $\mathcal{R}^{-1} = \mathcal{T}$ である。これらが線形写像であることは容易に確認できる:

練習 26.2.1. 上で定めた $\mathcal{R} : \text{Hom}(U, V) \rightarrow M_{m,n}(\mathbb{R})$ および $f, g \in \text{Hom}(U, V), r, s \in \mathbb{R}$ について次を示せ。

$$\text{線形性 (iii): } \mathcal{R}(rf + sg) = r\mathcal{R}(f) + s\mathcal{R}(g).$$

解答例: $[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m] \mathcal{R}(f)$ および $[g(\mathbf{u}_1), \dots, g(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m] \mathcal{R}(g)$ であり、このとき

$$\begin{aligned} [(rf + sg)(\mathbf{u}_1), \dots, (rf + sg)(\mathbf{u}_n)] &= r[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] + s[g(\mathbf{u}_1), \dots, g(\mathbf{u}_n)] \\ &= r[\mathbf{v}_1, \dots, \mathbf{v}_m] \mathcal{R}(f) + s[\mathbf{v}_1, \dots, \mathbf{v}_m] \mathcal{R}(g) \\ &= [\mathbf{v}_1, \dots, \mathbf{v}_m] (r\mathcal{R}(f) + s\mathcal{R}(g)). \end{aligned}$$

↑ここで命題 25.2.1 および 25.2.2 を用いた。

以上より、線形写像 $rf + sg$ の表現行列は $r\mathcal{R}(f) + s\mathcal{R}(g)$ に等しい。すなわち、 $\mathcal{R}(rf + sg) = r\mathcal{R}(f) + s\mathcal{R}(g)$ 。 \square

系 26.2.2. (1) U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ に関する $f, g : U \rightarrow V$ の表現行列が一致するならば、 $f = g$ である。

(2) 標準基底に関する $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ の表現行列が A ならば、 $f = T_A$ 。

Proof. (1): \mathcal{R} の単射性から明らか.

(2): 標準基底に関する f および T_A の表現行列は共に A である. ゆえに (1) より $f = T_A$. □

線形写像の合成と表現行列の積について, 次の関係が成立する.

命題 26.2.3. $f: U \rightarrow V, g: V \rightarrow W$ を線形写像とし, U, V, W の基底をそれぞれ $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{w}_1, \dots, \mathbf{w}_\ell$ とし, これらの基底に関する f, g の表現行列をそれぞれ A, B とする. このとき, これらの基底に関する $g \circ f$ の表現行列は BA である.

Proof. 仮定より,

$$[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]A, \quad [g(\mathbf{v}_1), \dots, g(\mathbf{v}_m)] = [\mathbf{w}_1, \dots, \mathbf{w}_\ell]B$$

である. 練習 25.3.2 に注意すれば,

$$[g(f(\mathbf{u}_1)), \dots, g(f(\mathbf{u}_n))] = [g(\mathbf{v}_1), \dots, g(\mathbf{v}_m)]A = [\mathbf{w}_1, \dots, \mathbf{w}_\ell]BA.$$

ゆえに $g \circ f$ の表現行列は BA である. □

系 26.2.4. U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ に関する線形同型 $f: U \rightarrow V$ の表現行列を A とすれば, これらの基底に関する $f^{-1}: V \rightarrow U$ の表現行列は A^{-1} である.

Proof. $W = U$ として前命題を適用する. B を $f^{-1}: V \rightarrow U$ の表現行列とすれば,

$$E = \text{“id}_U: U \rightarrow U \text{ の表現行列”} = \text{“}f^{-1} \circ f: U \rightarrow U \text{ の表現行列”} = BA.$$

同様に $AB = E$ も示される. つまり, B は A の逆行列である. □

系 26.2.4 は, 表現行列の定義から直接に示すこともできる:

系 26.2.4 の別証明. $[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_n]A$ に逆写像 f^{-1} をほどこせば, 練習 25.3.2 より $[\mathbf{u}_1, \dots, \mathbf{u}_n] = [f^{-1}(\mathbf{v}_1), \dots, f^{-1}(\mathbf{v}_n)]A$ である. この両辺に右から A^{-1} をかけて $[f^{-1}(\mathbf{v}_1), \dots, f^{-1}(\mathbf{v}_n)] = [\mathbf{u}_1, \dots, \mathbf{u}_n]A^{-1}$ を得る. すなわち, A^{-1} は f^{-1} の表現行列である. □

$U = V$ とし, U, V の基底として共に $\mathbf{u}_1, \dots, \mathbf{u}_n$ を取るとき, 命題 26.2.3 は $\mathcal{R}(gf) = \mathcal{R}(g)\mathcal{R}(f)$ を意味する. すなわち, 線形同型 $\mathcal{R}: \text{End}(U) \rightarrow M_n(\mathbb{R})$ は, さらに多元環同型でもある. 命題 21.5.7 より直ちに次が従う.

系 26.2.5. $\mathcal{R}: \text{End}(U) \rightarrow M_n(\mathbb{R})$ を上で与えた同型とすれば, 任意の多項式 $\Psi(t)$ および $f \in \text{End}(U)$ について $\mathcal{R}(\Psi(f)) = \Psi(\mathcal{R}(f))$. すなわち, U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列を A とすれば, $\Psi(f) \in \text{End}(U)$ の表現行列は $\Psi(A)$ である.

26.3 基底の取りかたによる表現行列の違い

線形写像 $T: U \rightarrow V$ の表現行列が基底の取り換えによってどう変化するか考察しよう.

U の基底: $\mathbf{u}_1, \dots, \mathbf{u}_n$ および $\mathbf{u}'_1, \dots, \mathbf{u}'_n$,

V の基底: $\mathbf{v}_1, \dots, \mathbf{v}_m$ および $\mathbf{v}'_1, \dots, \mathbf{v}'_m$,

とし, $\mathbf{u}_1, \dots, \mathbf{u}_n$ と $\mathbf{v}_1, \dots, \mathbf{v}_m$ に関する T の表現行列を A , $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ と $\mathbf{v}'_1, \dots, \mathbf{v}'_m$ に関する T の表現行列を B とする. また, 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P , 基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ による基底 $\mathbf{v}'_1, \dots, \mathbf{v}'_m$ の変換行列を Q とする. このとき,

命題 26.3.1. 上の設定のもとで $B = Q^{-1}AP$.

この命題は、それぞれの表現行列に関する可換図式をまとめた次の可換図式からほとんど明らかとも言える。

$$\begin{array}{ccccc} \mathbb{R}^n & \xleftarrow{F_2} & U & \xrightarrow{F_1} & \mathbb{R}^n \\ T_B \downarrow & & T \downarrow & & \downarrow T_A \\ \mathbb{R}^m & \xleftarrow{G_2} & V & \xrightarrow{G_1} & \mathbb{R}^m \end{array} \quad (26.3.1)$$

ここで、 F_1, F_2, G_1, G_2 は次の対応を意味する線形同型写像である：

$$\begin{aligned} F_1: F_1(\mathbf{u}_j) &= \mathbf{e}_j \in \mathbb{R}^n \text{ を満たす写像, } & G_1: G_1(\mathbf{v}_i) &= \mathbf{e}_i \in \mathbb{R}^m \text{ を満たす写像,} \\ F_2: F_2(\mathbf{u}'_j) &= \mathbf{e}_j \in \mathbb{R}^n \text{ を満たす写像, } & G_2: G_2(\mathbf{v}'_i) &= \mathbf{e}_i \in \mathbb{R}^m \text{ を満たす写像.} \end{aligned}$$

上の図式において $F_1 \circ F_2^{-1} = T_P$ および $G_1 \circ G_2^{-1} = T_Q$ である。これは変換行列の定義から直ちに得られる（詳しくは備考 25.5.4 を見よ）。可換図式 (26.3.1) における左上の \mathbb{R}^n から左下の \mathbb{R}^m への写像が近道と遠回りでも同等なことから $T_B = T_{Q^{-1}} \circ T_A \circ T_P$ であり、これは $B = Q^{-1}AP$ であることに他ならない。

命題 26.3.1 の証明. $F_1 \circ F_2^{-1} = T_P$ および $G_1 \circ G_2^{-1} = T_Q$ であり、次の可換図式が成り立つ：

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{T_P} & \mathbb{R}^n \\ T_B \downarrow & & \downarrow T_A \\ \mathbb{R}^m & \xrightarrow{T_Q} & \mathbb{R}^m \end{array}$$

可換図式 (26.3.1) から上の可換図式が導かれること（つまり $T_A \circ T_P = T_Q \circ T_B$ ）はほとんど明らかではあるが、一応確認しておこう。 $G_1 \circ T = T_A \circ F_1$ の両辺に左から G_1^{-1} を合成することで $T = G_1^{-1} \circ T_A \circ F_1$ を得る。また、同様に $T = G_2^{-1} \circ T_B \circ F_2$ でもある。 $G_1^{-1} \circ T_A \circ F_1 = G_2^{-1} \circ T_B \circ F_2$ の両辺に左から G_1 を右から F_2^{-1} それぞれ合成すると

$$\begin{aligned} G_1 \circ G_1^{-1} \circ T_A \circ F_1 \circ F_2^{-1} &= G_1 \circ G_2^{-1} \circ T_B \circ F_2 \circ F_2^{-1} \\ \text{id}_{\mathbb{R}^m} \circ T_A \circ T_P &= T_Q \circ T_B \circ \text{id}_{\mathbb{R}^n} \\ T_A \circ T_P &= T_Q \circ T_B. \end{aligned}$$

ゆえに命題 21.3.3 より $T_{AP} = T_{QB}$ および $AP = QB$ を得る。この両辺に左から Q^{-1} をかけて $B = Q^{-1}AP$ である。□

定理の主張への理解を促すため、上では図式を用いて説明した。計算結果が正しければそれでよいというのであれば、次のような証明もある。

命題 26.3.1 の別証明. $[T(\mathbf{u}'_1), \dots, T(\mathbf{u}'_n)]$ を二通りの方法で計算する。 $[\mathbf{u}'_1, \dots, \mathbf{u}'_n] = [\mathbf{u}_1, \dots, \mathbf{u}_n]P$ に T をほどこすと練習 25.3.2 より

$$[T(\mathbf{u}'_1), \dots, T(\mathbf{u}'_n)] = [T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)]P = [\mathbf{v}_1, \dots, \mathbf{v}_m]AP.$$

一方で、

$$[T(\mathbf{u}'_1), \dots, T(\mathbf{u}'_n)] = [\mathbf{v}'_1, \dots, \mathbf{v}'_m]B = [\mathbf{v}_1, \dots, \mathbf{v}_m]QB.$$

以上より $[\mathbf{v}_1, \dots, \mathbf{v}_m]AP = [\mathbf{v}_1, \dots, \mathbf{v}_m]QB$ である。練習 25.3.1 より $AP = QB$ 。したがって $Q^{-1}AP = B$ 。□

線形変換 $T: U \rightarrow U$ において、 T で写すことによりベクトルが元の位置からどう変化するかを見るのであれば、定義域と終域において同一の基底を取っておくことが望ましい。そこで、定義域の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および終域の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する T の表現行列のことを、以下では単に基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する T の表現行列と呼ぶことにしよう。前命題の特別な場合として次を得る。

系 26.3.2. $T: U \rightarrow U$ を線形変換とし、 U の二組の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ が与えられているとする。このとき、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する T の表現行列を A 、 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ に関する T の表現行列を B 、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P とすれば、 $B = P^{-1}AP$ である。

線形写像を分析する立場からは、表現行列に現れる成分があまり複雑でないことが望ましい。そのためには基底を上手く選ぶ必要がある。次章以降では、与えられた線形変換と相性のよい基底、すなわち表現行列が複雑にならないような基底の探し方を考察する。

定義 26.3.3. 同じサイズの正方行列 A, B において、 $B = P^{-1}AP$ を満たす可逆行列 P が存在するとき、 A と B は相似 (similar) であるという。

行列 A, B が相似であるとき、これらは異なる座標軸を通して同じ線形変換を表したものと捉えることができる。実際、 $B = P^{-1}AP$ とすれば、標準基底に関する T_A の表現行列は A であり、 P の列ベクトルの組からなる基底による T_A の表現行列は B である。

練習 26.3.4. 二つの行列が相似であるという関係は同値関係である。これを示せ。

命題 26.3.5. 正方行列 A, B が相似ならば、 $\text{tr } A = \text{tr } B$ 。

Proof. 仮定より、 $B = P^{-1}AP$ と書ける。命題 20.4.3 を用いれば、

$$\text{tr } B = \text{tr}(P^{-1}AP) = \text{tr}((P^{-1}A)P) = \text{tr}(P(P^{-1}A)) = \text{tr } A.$$

□

トレースは相似な行列について不変な線形写像であり、相似について不変な性質 (例えば座標軸の取り換えについて不変な性質) を調べる際の指標²となることだろう。

26.4 1対1の対応と可換図式

一般の線形写像 T を表現行列 A を用いて調べるにあたって、 T が持つ性質と T_A が持つ性質が同等であることを理解しておく必要がある。例えば、次のような性質を T が持つことと T_A が持つことは同値である:

単射性, 全射性, 像の次元が l , 核の次元が r 。

これらの同値性が可換図式 (26.1.2) における全単射 F, G を通して導かれることを確認しておこう。上の四つの性質のほかにも、 $\lambda \in \mathbb{R}$ が固有値となること、および固有値 λ に関する固有空間の次元、あるいは一般固有空間の次元について同様の対応を次章以降で論ずることになる。

線形空間の枠組みの外でも通用する基本的な事実として、次の命題が成り立つ。

命題 26.4.1. 次の可換図式が成り立っているとする (すなわち $G \circ T = S \circ F$ が成り立つ)。ここで、 U, V, R_n, R_m は集合であり (線形空間でなくてもよい)、 T, S, F, G は写像である。

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ F \downarrow & & \downarrow G \\ R_n & \xrightarrow{S} & R_m \end{array}$$

F および G が全単射であるとき次が成り立つ。

$$(1) T = G^{-1} \circ S \circ F, \quad S = G \circ T \circ F^{-1}.$$

² 「指標 (character)」は代数学の専門用語でもある。

(2) T は単射 $\iff S$ は単射, T は全射 $\iff S$ は全射, T は全単射 $\iff S$ は全単射.

(3) $X \subset U$ について, $T(X) = G^{-1}(S(F(X)))$.

(4) $Y \subset V$ について, $T^{-1}(Y) = F^{-1}(S^{-1}(G(Y)))$.

Proof. (1): $G \circ T = S \circ F$ の両辺に左から G^{-1} を合成すれば $T = G^{-1} \circ S \circ F$ を得る. また, 両辺に右から F^{-1} を合成すれば $S = G \circ T \circ F^{-1}$.

(2): 単射の合成は単射であること, および全射の合成が全射となることから, (1) より直ちに従う.

(3): 写像 $T = G^{-1} \circ S \circ F$ における X による像として, $T(X) = G^{-1} \circ S \circ F(X) = G^{-1}(S(F(X)))$ である.

(4): 逆写像 G^{-1} の Y による逆像について $(G^{-1})^{-1}(Y) = G(Y)$ であることに注意して, $T = G^{-1} \circ S \circ F$ における Y の逆像を取ると, 命題 19.4.5 より

$$T^{-1}(Y) = (G^{-1} \circ S \circ F)^{-1}(Y) = F^{-1}(S^{-1}((G^{-1})^{-1}(Y))) = F^{-1}(S^{-1}(G(Y))).$$

□

備考: (3) および (4) の証明に F の全単射性は実は不要であり, これらは G の単射性のみから導かれる:

G の単射性のみを用いた (3) と (4) の別証明.

(3): G の単射性より各 $B \subset V$ について $G^{-1}(G(B)) = B$ である (練習 19.2.5(1)). $S \circ F = G \circ T$ より $S \circ F(X) = G \circ T(X) = G(T(X))$. この両辺による G の逆像をとれば, $G^{-1}(S \circ F(X)) = G^{-1}(G(T(X))) = T(X)$.

(4): 両方の包含関係を示す. 各 $x \in T^{-1}(Y)$ に対して, $T(x) \in Y$ より $(S \circ F)(x) = G \circ T(x) = G(T(x)) \in G(Y)$. ゆえに $x \in (S \circ F)^{-1}(G(Y)) = F^{-1}(S^{-1}(G(Y)))$. 一方, 各 $w \in F^{-1}(S^{-1}(G(Y))) = (S \circ F)^{-1}(G(Y))$ に対して, $S \circ F(w) \in G(Y)$ より, ある $y \in Y$ を用いて $S \circ F(w) = G(y)$ と書ける. $G(T(w)) = G \circ T(w) = S \circ F(w) = G(y)$ より $G(T(w)) = G(y)$ であり, G の単射性から $T(w) = y \in Y$. したがって $w \in T^{-1}(Y)$ である. □

上の命題の応用として, 一般の線形写像 $T : U \rightarrow V$ の像と核が表現行列についての線形写像の像と核にそれぞれ対応することを見よう.

命題 26.4.2. U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ に関する $T : U \rightarrow V$ の表現行列を A とすれば, 次が成り立つ.

(1) $\text{Im } T \simeq \text{Im } T_A$,

(2) $\text{Ker } T \simeq \text{Ker } T_A$,

(3) $\dim \text{Im } T = \dim \text{Im } T_A = \text{rank } A$,

(4) $\dim \text{Ker } T = \dim \text{Ker } T_A = n - \text{rank } A$.

Proof. 可換図式 (26.1.2):

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ F \downarrow & & \downarrow G \\ \mathbb{R}^n & \xrightarrow{T_A} & \mathbb{R}^m \end{array}$$

について前命題を適用する.

(1): $X = U$ について前命題 (3) を適用すると

$$\text{Im } T = T(U) = G^{-1}(T_A(F(U))) = G^{-1}(T_A(\mathbb{R}^n)) = G^{-1}(\text{Im } T_A).$$

(↑ここで前命題 (3) を用いた)

すなわち, 線形同型写像 $G^{-1} : \text{Im } T_A \rightarrow \text{Im } T$ が存在するゆえ, これらは同型である.

(2): $Y = \{\mathbf{0}_V\}$ について前命題 (4) を適用すれば,

$$\text{Ker } T = T^{-1}(\{\mathbf{0}_V\}) = F^{-1}(T_A^{-1}(G(\{\mathbf{0}_V\}))) = F^{-1}(T_A^{-1}(\{\mathbf{0}\})) = F^{-1}(\text{Ker } T_A).$$

(↑ここで前命題 (4) を用いた)

線形同型写像 $F^{-1} : \text{Ker } T_A \rightarrow \text{Ker } T$ により, これらは同型である.

(3) と (4): 前半の等号は (1) と (2) より明らか. 後半の等号は, 既に例 22.1.3(2) で述べた通りである. □

$T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ における次元定理は例 23.2.5 でみた通りであり, これと上の (3) および (4) を合わせることで, 一般の線形写像 $T : U \rightarrow V$ における次元公式の別証明を得る.

第27章 双対空間(発展)

行列 A を線形写像 $f: U \rightarrow V$ の表現行列とみなすとき, tA に対応する自然な線形写像があることを解説する. その写像は V から U への写像ではなく, 特別な関数の集合からなる空間 (双対空間) における線形写像として与えられる. 本章では \mathbb{K} を一般の体 (\mathbb{Q} や \mathbb{R} , \mathbb{C} など) とし, 体 \mathbb{K} 上の線形空間の双対について論じる. これを抽象的に感じる者は $\mathbb{K} = \mathbb{R}$ として読めばよい.

双対空間がどのように応用されるかについては本書の枠を大きく超える. つまり, 本章で述べる内容は本書の以降では用いない. 本章の内容を難しいと感じる読者は, 定義のみに目を通して第 V 部に進んでよい. なお, 27.5 節まで読めば, より発展的な数学の文献においてベクトルを太字で表さなくなる理由が理解されるであろう.

27.1 定義

定義 27.1.1. 体 \mathbb{K} 上の線形空間 U に対して, $\text{Hom}(U, \mathbb{K})$ を U の双対空間 (**dual space**) と呼び, これを U^* と書く.

例えば $(\mathbb{R}^n)^*$ の元の例として, 第 i 座標への射影 $\text{pr}_i: \mathbb{R}^n \rightarrow \mathbb{R}$ が挙げられる (例 20.4.1). 他の具体例としては次を挙げるに留めよう.

例 27.1.2 (よりみち). 閉区間 $X = [a, b]$ (ただし $a < b$) 上の連続関数全体 $C(X)$ はベクトル空間をなす (例 15.2.7).

(1) $x_0 \in X$ を一つ固定しておき, $q_{x_0}: C(X) \rightarrow \mathbb{R}$ を $q_{x_0}(f) := f(x_0)$ と定めれば, q_{x_0} は連続関数を代入する関数であり, q_{x_0} は線形ゆえ $q_{x_0} \in C(X)^*$. なお, $\text{Ker } q_{x_0}$ は, 例 16.3.5(2) で与えたイデアル \mathcal{I}_{x_0} に等しい.

(2) 区間 $[a, b]$ の n 等分割 $x_0 = a < x_1 < \dots < x_n = b$ (つまり, $x_i = a + (b-a)i/n$) を与え, 分割した各区間ごとに, あらかじめ元 $\xi_i^n \in [x_{i-1}, x_i]$ ($i = 1, \dots, n$) を一つ取っておく. このとき $q_{\xi_i^n}$ の平均 $A_n = \frac{1}{n} \sum_{i=1}^n q_{\xi_i^n}$ を取れば, もちろん $A_n \in C(X)^*$ である.

(3) $S: C(I) \rightarrow \mathbb{R}$ を $S(f) := \int_a^b f(x)dx$ と定めれば $S \in C(X)^*$ である (例 20.4.7(1)).

(4) A_n は次のように書き換えられる:

$$A_n(f) = \frac{1}{n} \sum_{i=1}^n f(x_i) = \frac{1}{b-a} \sum_{i=1}^n f(\xi_i^n) \frac{b-a}{n} = \frac{1}{b-a} \sum_{i=1}^n f(\xi_i^n)(x_i - x_{i-1}).$$

したがって, 各 $f \in C(X)$ について, $\lim_{n \rightarrow \infty} A_n(f) = \frac{1}{b-a} S(f)$ である (区分求積法).

何らかの二つの対象 A, B が与えられているとしよう. このとき A, B が異なるかどうかを認識する際に, 普段我々はよく理解されている別の対象に A, B を対応づけることによって判断している. 例えば二人の人物が別人かどうかを判別する際には, 彼らの身長や体重, 生年月日, 顔の風貌などといった人物にまつわるデータ (人物に関する関数) に何らかの違いがないかを見ている. 普段の生活において意識することは少ないながらも, このように関数や写像は, 相違を認識するうえで決定的な役割を果たす.

いま, 違いを認識するには, 何らかの関数 (や写像) の値が異なる例を見つけ出せばよいことを述べた. しかしながら, 月の裏側の写真を見て, それが普段我々が見ている月の模様 (兎の餅つきと形容される柄) と異なるからといって, それを月とは異なる衛星であると結論づけてははやとちりである. この喩えからは, 単に異なるというだけではなく, 本質的な違いを見出せるような適切なデータ (関数) を用意する必要があることが分かる. この適切性は, 何を分析したいと考えているかによって千差万別であり, ここで一般論として語ることはできない. しかしながら, 日常生活において我々は絶えず「あれとこれは違う」と認識しているわけであるから, そういった色々な例を通して, それぞれの違いの本質はどこにあるのか自問自答してみると, 認識の質を深める訓練になるであろう.

さて, 分析すべき対象がベクトル空間 V であり, V の元における何らかの違いを見出す関数が線形であるならば, それは V^* の元である. そこで, 双対空間 V^* の中から, 本質的な違いを見出せるような関数を探る, という手法が数学のいくつかの分野で用いられている. 例えば形の違いを把握する量として幾何学には, コホモロジーと呼ばれる双対空間 (の商空間) の中から適切な関数を見出す理論がある.

27.2 双対基底

体 \mathbb{K} が \mathbb{R} や \mathbb{C} のとき, U^* の各元は関数であるから, 本章では U^* の元を線形関数 (あるいは単に関数) とよび, 一般の線形空間 V における $\text{Hom}(U, V)$ の元を線形写像と呼んで区別することにしよう. 線形関数 $f: U \rightarrow \mathbb{K}$ が U^* の元であることを強調するとき, f は線形汎関数 (linear functional) とも呼ばれる.

U に基底が与えられているとき, それらを用いて U^* の基底を定めることができる.

例 27.2.1. $(\mathbb{R}^n)^* = \text{Hom}(\mathbb{R}^n, \mathbb{R})$ と $M_{1,n}(\mathbb{R})$ の各元は 1 対 1 に対応するのであった (定理 21.3.4). $M_{1,n}(\mathbb{R})$ は n 次行ベクトル全体からなる集合であり, たとえば $M_{1,n}(\mathbb{R})$ の基底として t_{e_1}, \dots, t_{e_n} が取れる. したがって, これらのベクトルに対応する $T_{t_{e_1}}, \dots, T_{t_{e_n}}$ は $(\mathbb{R}^n)^*$ の基底である. なお関数 $T_{t_{e_i}}$ は第 i 座標への射影 pr_i (例 20.4.1) に等しい.

定義 27.2.2. 線形空間 U とその基底 u_1, \dots, u_n が与えられているとき, 各 $i = 1, \dots, n$ に対して線形関数 $u_i^*: U \rightarrow \mathbb{K}$ を次のように定め, これを u_i の双対 (dual) と呼ぶ.

$$\text{各 } j = 1, \dots, n \text{ について, } U \ni u_j \xrightarrow{u_i^*} \delta_{ij} \in \mathbb{K}.$$

【注意】 つまり, 関数 u_i^* は, u_i を 1 に移し, それ以外の u_j を 0 に写す. なお, 基底の行き先さえ決めれば, U の各元の行き先も定まる (命題 20.1.11).

上の u_i^* は, u_1, \dots, u_n すべてを用いて定めるものであるから, 本当は $(u_1, \dots, u_n)_i^*$ と書かなければいけない¹. しかし, 本書では与えられている基底に誤解がない限り, これを略して u_i^* と書くことにする. u_1, \dots, u_n の双対を u^1, \dots, u^n と書く流儀もある.

例 27.2.3. (1) \mathbb{R}^n において, 列ベクトルからなる標準基底 e_1, \dots, e_n を考えるとき, その双対を e_1^*, \dots, e_n^* とすれば, $e_i^* = T_{t_{e_i}}: \mathbb{R}^n \rightarrow \mathbb{R}$ ($i = 1, \dots, n$) である.

¹ U に内積が定まっており, かつ基底 u_1, \dots, u_n が正規直交基底の場合は, $u_j^*(x) = \langle u_j, x \rangle$ であり, この場合は u_j のみによって定まる概念になる.

- (2) U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$, およびその双対 $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ が与えられているとする. このとき, U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および \mathbb{K} の基底 1 に関する線形関数 $\mathbf{u}_i^* : U \rightarrow \mathbb{K}$ の表現行列は, 行ベクトル ${}^t\mathbf{e}_i \in M_{1,n}(\mathbb{K})$ に等しい.

命題 27.2.4. U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が与えられているとする. このとき,

- (1) $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ は線形独立である.
- (2) 各 $f \in U^*$ に対して, $r_i := f(\mathbf{u}_i)$ ($i = 1, \dots, n$) とおくと, $f = \sum_{i=1}^n r_i \mathbf{u}_i^*$.
- (3) $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ は U^* の基底であり, $\dim U = \dim U^*$.

【備考】 この事実を通して組 $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ を, 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の双対基底 (dual basis) という.

Proof. (1): $f = \sum_{i=1}^n r_i \mathbf{u}_i^* = \mathbf{0} \in U^*$ と仮定する. つまり $f : U \rightarrow \mathbb{K}$ は, すべてのベクトルを $0 \in \mathbb{K}$ に写す関数である. このとき各 $j = 1, \dots, n$ について, $0 = f(\mathbf{u}_j) = \sum_{i=1}^n r_i \mathbf{u}_i^*(\mathbf{u}_j) = \sum_{i=1}^n r_i \delta_{ij} = r_j$. ゆえに $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ は線形独立である.

(2): $g = \sum_{i=1}^n f(\mathbf{u}_i) \mathbf{u}_i^* : U \rightarrow \mathbb{K}$ とおく. このとき, 各 \mathbf{u}_j ($j = 1, \dots, n$) を f に代入した値と g に代入した値は共に $f(\mathbf{u}_j)$ である. つまり, 基底の行き先が一致するゆえ $f = g$ である (命題 20.1.10(2)).

(3): (1) と (2) より明らか. □

27.3 双対写像

定義 27.3.1. 体 \mathbb{K} 上の線形空間における線形写像 $T : U \rightarrow V$ に対して, 線形写像 $T^* : V^* \rightarrow U^*$ を $T^*(g) := g \circ T$ で定める. この写像を T の双対 (dual) または転置 (transpose) という.

記号 T^* は, T の随伴変換と呼ばれる似て非なる概念に対しても用いられるため注意を要する². T の双対写像を ${}^tT : V^* \rightarrow U^*$ とも書く. 本章で紹介する部分に限れば, U およびその基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$, 線形写像 $T : U \rightarrow V$ について, それぞれの双対を ${}^tU, {}^t\mathbf{u}_1, \dots, {}^t\mathbf{u}_n, {}^tT$ と書くのが自然で混乱がないように思われる. しかしながら他の文献に合わせて, 本書でも一般的に用いられている記法を用いよう.

例 27.3.2. $A = [a_{ij}] \in M_{m,n}(\mathbb{K})$ とし, $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ の双対について考えよう. 各行ベクトル $\mathbf{y} = (y_1, \dots, y_m)$ について, $T\mathbf{y} \in (\mathbb{R}^m)^*$ である. このとき, 線形関数 $T\mathbf{y}$ を双対写像 $(T_A)^*$ に代入したものは, $(T_A)^*(T\mathbf{y}) = T\mathbf{y} \circ T_A = T\mathbf{y}A$. つまりこの線形関数 $(T_A)^*(T\mathbf{y}) : \mathbb{R}^n \rightarrow \mathbb{K}$ は, 各 $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ を $T\mathbf{y}A(\mathbf{x}) = \mathbf{y}A\mathbf{x} \in \mathbb{K}$ に写す. これを成分表示して書けば,

$$\left((T_A)^*(T\mathbf{y}) \right) (\mathbf{x}) = \left([y_1, y_2, \dots, y_m] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

したがって, $(\mathbb{R}^m)^*$ および $(\mathbb{R}^n)^*$ の各元を行ベクトルと同一視するとき, $(T_A)^*$ に $T\mathbf{y}$ を代入することは, \mathbf{y} に右から A をかけることに他ならない.

さて, ふだん我々は線形写像に代入する側のベクトル, および代入して得られるベクトルをそれぞれ列ベクトルで表示することにしていた. そこで m 次行ベクトル \mathbf{y} および n 次行ベクトル $\mathbf{y}A$ を転置して列ベクトル表示すれば, ${}^t(\mathbf{y}A) = {}^tA {}^t\mathbf{y}$ となる. つまり, 標準基底の双対基底による $(T_A)^* : (\mathbb{R}^m)^* \rightarrow (\mathbb{R}^n)^*$ の表現行列は tA に等しい.

上の例のより一般的な場合として, 次が成り立つ.

命題 27.3.3. U と V を体 \mathbb{K} 上の有限次元線形空間とし, これらの基底をそれぞれ, $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ および $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ とする. さらに $T : U \rightarrow V$ を線形写像とし, いま与えた基底に関する T の表現行列を $A \in M_{m,n}(\mathbb{K})$ とすれば, 双対基底に関する $T^* : V^* \rightarrow U^*$ の表現行列は tA である.

²随伴変換は第 VI 部で論じる.

Proof. $A = [a_{ij}]$ とおく、つまり、各 $j = 1, \dots, n$ について $T(\mathbf{u}_j) = \sum_{i=1}^m a_{ij} \mathbf{v}_i$ である。各 $k = 1, \dots, m$ について、関数 $T^*(\mathbf{v}_k) = \mathbf{v}_k^* \circ T : U \rightarrow \mathbb{R}$ に $\mathbf{u}_1, \dots, \mathbf{u}_n$ を代入した値を計算すると、

$$T^*(\mathbf{v}_k)(\mathbf{u}_j) = \mathbf{v}_k^* \circ T(\mathbf{u}_j) = \mathbf{v}_k^* \left(\sum_{i=1}^m a_{ij} \mathbf{v}_i \right) = \sum_{i=1}^m a_{ij} \mathbf{v}_k^*(\mathbf{v}_i) = \sum_{i=1}^m a_{ij} \delta_{ki} = a_{kj}.$$

したがって、命題 27.2.4(2) より、 $T^*(\mathbf{v}_k) = \sum_{j=1}^n a_{kj} \mathbf{u}_j^*$ が成り立つ。つまり、

$$[T^*(\mathbf{v}_1), \dots, T^*(\mathbf{v}_m)] = [\mathbf{u}_1^*, \dots, \mathbf{u}_n^*] \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

ゆえに双対基底に関する T^* の表現行列は、 $A = [a_{kj}]_{\substack{j=1, \dots, n, \\ k=1, \dots, m}} = {}^t A$ である。 □

27.4 双対写像の基本的性質

双対写像と、もとの写像との間にある基本的な性質を抑えておこう。

補題 27.4.1. n 次元線形空間 U および $\mathbf{x} \in U$ について次が成り立つ。

- (1) $\mathbf{x} \neq \mathbf{0} \implies f(\mathbf{x}) \neq 0$ を満たす $f \in U^*$ が存在する。
- (2) 各 $f \in U^*$ について $f(\mathbf{x}) = 0 \implies \mathbf{x} = \mathbf{0}$.

Proof. (1): $\mathbf{u}_1 := \mathbf{x}$ とし、これにベクトルを加えて U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を作る。この双対基底を考え、 $f := \mathbf{u}_1^*$ とすれば、これが求めるものである。実際、 $f(\mathbf{x}) = \mathbf{u}_1^*(\mathbf{u}_1) = 1 \neq 0$.

(2): これは (1) の対偶である。 □

備考 27.4.4 にあるように、次の (1) と (2) の証明は略してもよい。

命題 27.4.2. $T : U \rightarrow V$ を有限次元線形空間の間の線形写像とする。

- (1) T は単射である $\iff T^*$ は全射である。
- (2) T は全射である $\iff T^*$ は単射である。
- (3) $\mathbf{v}_1, \dots, \mathbf{v}_k$ を $\text{Im } T$ の基底とする。このとき、これらにベクトルを加えて $\mathbf{v}_1, \dots, \mathbf{v}_m$ を V の基底を取れば、この双対基底に関して、 $T^*(\mathbf{v}_1), \dots, T^*(\mathbf{v}_k)$ は $\text{Im } T^*$ の基底である。
- (4) $\dim \text{Im } T = \dim \text{Im } T^*$.

Proof. (1): (\implies): T は単射ゆえ、 $S \circ T = \text{id}_U$ を満たす線形写像 $S : V \rightarrow U$ が存在する (命題 22.4.9(1)). 各 $f \in U^*$ に対して $g := f \circ S$ とおけば、 $T^*(g) = g \circ T = f \circ S \circ T = f$.

(\impliedby): $\mathbf{x} \in \text{Ker } T$ とする。 $\mathbf{x} = \mathbf{0}$ が示せれば T は単射である (定理 20.3.4). そのためには、各 $f \in U^*$ について $f(\mathbf{x}) = 0$ を言えばよい (補題 27.4.1(2)). $f \in U^*$ とすれば、 T^* の全射性より $f = T^*(g) = g \circ T$ を満たす $g \in V^*$ が存在する。このとき、 $f(\mathbf{x}) = g(T(\mathbf{x})) = g(\mathbf{0}) = 0$ である。

(2): (\implies): $g \in \text{Ker } T^*$ とし、 $g = \mathbf{0}_{V^*}$ を示そう。そのためには、各 $\mathbf{y} \in V$ について $g(\mathbf{y}) = 0$ を言えばよい。 $\mathbf{y} \in V$ とすれば、 T の全射性より $\mathbf{y} = T(\mathbf{x})$ を満たす $\mathbf{x} \in U$ が存在する。このとき、 $g \circ T = T^*(g) = \mathbf{0}_{U^*}$ (つまり $g \circ T(\mathbf{x}) = 0$) ゆえ、 $g(\mathbf{y}) = g(T(\mathbf{x})) = 0$.

(\impliedby): 対偶を示す。 T が全射でないとする。 $\mathbf{y} \in V \setminus \text{Im } T$ が存在する。このとき、 $\text{Im } T$ の基底を $\mathbf{v}_1, \dots, \mathbf{v}_k$ とすれば、組 $\mathbf{y}, \mathbf{v}_1, \dots, \mathbf{v}_k$ は線形独立である。よってさらにベクトルを加え $\mathbf{y}, \mathbf{v}_1, \dots, \mathbf{v}_{m-1}$ を V の基底とすることができる。この双対基底を $\mathbf{y}^*, \mathbf{v}_1^*, \dots, \mathbf{v}_{m-1}^*$ とすれば、次が成り立つ:

(*) 各 $\mathbf{x} \in U$ について, $\mathbf{y}^*(T(\mathbf{x})) = 0$.

実際, $\mathbf{x} \in U$ とすれば $T(\mathbf{x}) = \sum_{i=1}^k r_i \mathbf{v}_i$ と書けること, および双対基底の定義より $\mathbf{y}^*(T(\mathbf{x})) = \mathbf{y}^*(\sum_{i=1}^k r_i \mathbf{v}_i) = \sum_{i=1}^k r_i \mathbf{y}^*(\mathbf{v}_i) = 0$ である. 条件 (*) は $\mathbf{y}^* \circ T = \mathbf{0}_{U^*}$, つまり $\mathbf{y}^* \in \text{Ker } T^*$ を意味し, したがって T^* は単射でない.

(3): まず線形独立性を示そう. そこで $\sum_{i=1}^k r_i T^*(\mathbf{v}_i^*) = \mathbf{0}_{U^*}$ とおく. 各 $j = 1, \dots, k$ に対して, 像の定義から $T(\mathbf{u}_j) = \mathbf{v}_j$ を満たす \mathbf{u}_j が取れる. これを代入すれば,

$$\begin{aligned} 0 &= \left(\sum_{i=1}^k r_i T^*(\mathbf{v}_i^*) \right) (\mathbf{u}_j) = \left(\sum_{i=1}^k r_i (\mathbf{v}_i^* \circ T) \right) (\mathbf{u}_j) = \sum_{i=1}^k r_i (\mathbf{v}_i^* \circ T(\mathbf{u}_j)) \\ &= \sum_{i=1}^k r_i \mathbf{v}_i^*(T(\mathbf{u}_j)) = \sum_{i=1}^k r_i \mathbf{v}_i^*(\mathbf{v}_j) = \sum_{i=1}^k r_i \delta_{ij} = r_j. \end{aligned}$$

ゆえに $r_1 = \dots = r_k = 0$ であり, $T^*(\mathbf{v}_1^*), \dots, T^*(\mathbf{v}_k^*)$ は線形独立である.

$\text{Im } T^* = \langle T^*(\mathbf{v}_1^*), \dots, T^*(\mathbf{v}_m^*) \rangle$ であるから (命題 20.2.3), $T^*(\mathbf{v}_1^*), \dots, T^*(\mathbf{v}_k^*)$ が $\text{Im } T^*$ を生成することをいうには, $i = k+1, \dots, m$ について $T^*(\mathbf{v}_i^*) = \mathbf{0}_{U^*}$ を示せば十分である. $i = k+1, \dots, m$ とし, 関数 $T^*(\mathbf{v}_i^*)$ にかなる U の元を代入しても 0 になることを示そう. 任意の $\mathbf{x} \in U$ について, $\mathbf{v}_1, \dots, \mathbf{v}_k$ が $\text{Im } T$ の基底であることから, $T(\mathbf{x}) = \sum_{j=1}^k r_j \mathbf{v}_j$ と書ける. このとき

$$T^*(\mathbf{v}_i^*)(\mathbf{x}) = \mathbf{v}_i^* \circ T(\mathbf{x}) = \mathbf{v}_i^* \left(\sum_{j=1}^k r_j \mathbf{v}_j \right) = \sum_{j=1}^k r_j \mathbf{v}_i^*(\mathbf{v}_j) = \sum_{j=1}^k r_j \delta_{ij} = 0.$$

上の最後の等式において, $i \geq k+1 > k \geq j$ より $\delta_{ij} = 0$ を用いた.

(4): (3) より明らか. □

系 27.4.3. (m, n)-行列 A について, $\text{rank } {}^t A = \text{rank } A$.

Proof. 標準基底に関する $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ の表現行列は A であり, また, 標準基底の双対基底に関する $(T_A)^* : (\mathbb{R}^m)^* \rightarrow (\mathbb{R}^n)^*$ の表現行列は ${}^t A$ である (命題 27.3.3). ゆえに命題 26.4.2 (あるいは例 22.1.3(2)) より, $\text{rank } A = \dim \text{Im } T_A = \dim \text{Im } (T_A)^* = \text{rank } {}^t A$. □

備考 27.4.4. 命題 27.4.2 における (1) と (2) の同値性は, 等式 (4) から直ちに導くこともできる. 例えば, (2) の (\Leftarrow) については, T^* の単射性より $\dim \text{Im } T^* = \dim V^*$ であり (練習 22.4.7(1)), ゆえに $\dim V = \dim V^* = \dim \text{Im } T^* = \dim \text{Im } T$, すなわち T は全射である (練習 22.4.7(2)). ほかの三つの矢印についても類似する戦略で示せる. にもかかわらず, あえて定義まで戻って示したのは, 公式でブラックボックス化してしまうと理解が深まらないことがあり, これを危惧してのことである.

27.5 二重双対

行列 A について ${}^t({}^t A) = A$ であった. 一方で, 双対写像に転置記号を用いるとき, 線形写像 $T : U \rightarrow V$ について ${}^t({}^t T) = T$ ではない. ${}^t({}^t T) = (T^*)^*$ は, T の双対写像 $T^* : V^* \rightarrow U^*$ の双対写像であるから, $(T^*)^* \in \text{Hom}((U^*)^*, (V^*)^*)$ である. 実は, 自然な単射 $\mathcal{N}_U : U \rightarrow (U^*)^*$ が与えられ, さらに U が有限次元の場合これは同型となる. そして, この同型を通して, T と ${}^t({}^t T)$ を自然に同一視することができる. この事実を紹介しよう. 以下では括弧を略し, 二重双対 $(U^*)^*$ および $(T^*)^*$ のことを, それぞれ U^{**} および T^{**} と書く.

記号 $f(\mathbf{x})$ について, これを普段我々は f に \mathbf{x} を代入したものと見ている. 他方で, 記号の意味を知らぬ者がこれを見たとき, f を \mathbf{x} に代入したものの³であると予想するかもしれない. そして, その予想はあながち間違いではないのである. 実際, 次の写像 $\mathcal{N}_U : U \rightarrow U^{**}$ は, $\mathbf{x} \in U$ が U^* を定義域とする関数とみなせることを意味する:

³つまり普段の表記でいう $\mathbf{x}(f)$ のこと.

各 $\mathbf{x} \in U$ について, 関数 $\mathcal{N}_U(\mathbf{x}) : U^* \rightarrow \mathbb{K}$ を $\mathcal{N}_U(\mathbf{x})(f) := f(\mathbf{x})$ と定める.

命題 27.5.1. (1) 各 $\mathbf{x} \in U$ について, 関数 $\mathcal{N}_U(\mathbf{x}) : U^* \rightarrow \mathbb{K}$ は線形である (つまり $\mathcal{N}_U(\mathbf{x}) \in U^{**}$).

(2) 写像 $\mathcal{N}_U : U \rightarrow U^{**}$ は, 線形単射である.

(3) U が有限次元のとき, \mathcal{N}_U は線形同型である.

(4) $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とする. このとき, この双対基底 $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ の双対基底を $\mathbf{u}_1^{**}, \dots, \mathbf{u}_n^{**}$ とすれば, 各 $j = 1, \dots, n$ について $\mathcal{N}_U(\mathbf{u}_j) = \mathbf{u}_j^{**}$ が成り立つ.

(5) 線形写像 $T : U \rightarrow V$ に対して, $\mathcal{N}_V \circ T = T^{**} \circ \mathcal{N}_U$ である. すなわち次の可換図式が成り立つ.

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ \mathcal{N}_U \downarrow & & \downarrow \mathcal{N}_V \\ U^{**} & \xrightarrow{T^{**}} & V^{**} \end{array}$$

【備考】 U と V が有限次元であるとき, (3) より \mathcal{N}_U と \mathcal{N}_V は同型であり, 上の可換図式は T と T^{**} を同一視できることを意味している.

Proof. (1): $\xi = \mathcal{N}_U(\mathbf{x})$ とおく. 各 $f, g \in U^*$ および $r, s \in \mathbb{K}$ について, $\xi(rf + sg) = (rf + sg)(\mathbf{x}) = rf(\mathbf{x}) + sg(\mathbf{x}) = r\xi(f) + s\xi(g)$.

(2): 線形性: 各 $\mathbf{x}, \mathbf{y} \in U$ および $r, s \in \mathbb{K}$ について, $\mathcal{N}_U(r\mathbf{x} + s\mathbf{y}) = r\mathcal{N}_U(\mathbf{x}) + s\mathcal{N}_U(\mathbf{y})$ を示せばよい. これら二つの関数が等しいことをいうために, 両辺に各 $f \in U^*$ を代入した値が一致することを確認しよう. $\mathcal{N}_U(r\mathbf{x} + s\mathbf{y})(f) = f(r\mathbf{x} + s\mathbf{y}) = rf(\mathbf{x}) + sf(\mathbf{y})$. また $(r\mathcal{N}_U(\mathbf{x}) + s\mathcal{N}_U(\mathbf{y}))(f) = r\mathcal{N}_U(\mathbf{x})(f) + s\mathcal{N}_U(\mathbf{y})(f) = rf(\mathbf{x}) + sf(\mathbf{y})$. ゆえにこれらは等しい.

単射性: $\text{Ker } \mathcal{N}_U = \{\mathbf{0}\}$ を示す. そこで $\mathbf{x} \in \text{Ker } \mathcal{N}_U$ を仮定する. このとき, $\mathcal{N}_U(\mathbf{x}) = \mathbf{0}_{U^{**}}$ であるから, 各 $f \in U^*$ について $f(\mathbf{x}) = \mathcal{N}_U(\mathbf{x})(f) = 0$. ゆえに補題 27.4.1 より $\mathbf{x} = \mathbf{0}$. つまり $\text{Ker } \mathcal{N}_U = \{\mathbf{0}\}$ である.

(3): 単射性と $\dim U = \dim U^* = \dim U^{**}$ (命題 27.2.4) より明らか (命題 22.4.8).

(4): $j = 1, \dots, n$ とする. 各 $f \in U^*$ に対して $\mathcal{N}_U(\mathbf{u}_j)(f) = \mathbf{u}_j^{**}(f)$ を示せばよい. そのためには U^* の基底について, $\mathcal{N}_U(\mathbf{u}_j)(\mathbf{u}_i^*) = \mathbf{u}_j^{**}(\mathbf{u}_i^*)$ ($i = 1, \dots, n$) を示せば十分である. $\mathcal{N}_U(\mathbf{u}_j)(\mathbf{u}_i^*) = \mathbf{u}_i^*(\mathbf{u}_j) = \delta_{ij}$. また, $\mathbf{u}_j^{**}(\mathbf{u}_i^*) = (\mathbf{u}_j^*)^*(\mathbf{u}_i^*) = \delta_{ji} = \delta_{ij}$. ゆえにこれらは等しい.

(5): $\mathbf{x} \in U$ とし, 各 $g \in V^*$ について $(\mathcal{N}_V \circ T(\mathbf{x}))(g) = (T^{**} \circ \mathcal{N}_U(\mathbf{x}))(g)$ を示そう. まず左辺について, $(\mathcal{N}_V \circ T(\mathbf{x}))(g) = \mathcal{N}_V(T(\mathbf{x}))(g) = g(T(\mathbf{x}))$ である. 一方, 写像 $(T^*)^* : (U^*)^* \rightarrow (V^*)^*$ の定義が, 各 $F \in U^{**}$ に対して $F \circ T^*$ を対応させるものであったことに注意すれば, $T^{**} \circ \mathcal{N}_U(\mathbf{x}) = (T^*)^*(\mathcal{N}_U(\mathbf{x})) = \mathcal{N}_U(\mathbf{x}) \circ T^*$ である. ゆえに

$$(T^{**} \circ \mathcal{N}_U(\mathbf{x}))(g) = (\mathcal{N}_U(\mathbf{x}) \circ T^*)(g) = \mathcal{N}_U(\mathbf{x})(T^*(g)) = T^*(g)(\mathbf{x}) = g \circ T(\mathbf{x}) = g(T(\mathbf{x})).$$

ゆえにこれらは等しい. □

線形代数を深めていくと, 関数や写像が線形空間の元 (つまりベクトル) となり, ベクトルもまた二重双対空間の元として関数になる. このような多面的な見方を獲得すると, もはやベクトルを太字で表して誤解を避けようとする試みは無意味であり, 次第に細字で表記するようになる.

27.6 内積との類似性

U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$, およびその双対基底 $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ を取れば, 各 \mathbf{u}_i を \mathbf{u}_i^* に写す同型写像 $\iota : U \rightarrow U^*$ が得られる. この写像は基底の取り方に依存して定まるものであり, 前節で与えた $\mathcal{N}_U : U \rightarrow U^{**}$ のように, 基底の取り方によらずに定まるものではない. この意味において, 一般には U から U^* への自然な写像は存在しない.

しかしながらこの件に関して、例えば $U = \mathbb{R}^n$ の場合は、列ベクトル ${}^t(x_1, \dots, x_n) \in \mathbb{R}^n$ を行ベクトル $(x_1, \dots, x_n) \in M(1, n) \simeq (\mathbb{R}^n)^*$ に写す自然な同型写像があるではないかと思う読者もおられよう。実は内積空間においては、次の方法で基底の取り方によらない全単射 $\iota: U \rightarrow U^*$ が構成できるため、これを通して U と U^* を同一視する場合がある。以下は内積空間 (第 VI 部) について学んでから読むとよい。

備考 27.6.1. \mathbb{K} を \mathbb{R} または \mathbb{C} とし、 $(U, \langle \cdot, \cdot \rangle)$ を体 \mathbb{K} 上の内積空間とする。

- (1) 各 $\mathbf{x} \in U$ に対して、関数 $\iota(\mathbf{x}): U \rightarrow \mathbb{K}$ を $\iota(\mathbf{x})(\mathbf{y}) := \langle \mathbf{x}, \mathbf{y} \rangle$ で定めれば、 $\iota(\mathbf{x})$ は線形である (すなわち $\iota(\mathbf{x}) \in U^*$)。
 - (2) $\iota: U \rightarrow U^*$ は単射である。
 - (3) $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とすると、その双対基底 $\mathbf{u}_1^*, \dots, \mathbf{u}_n^*$ について、一般には $\iota(\mathbf{u}_i) \neq \mathbf{u}_i^*$ である。 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が正規直交基底の場合は $\iota(\mathbf{u}_i) = \mathbf{u}_i^*$ が成り立つ。
 - (4) U が体 \mathbb{R} 上の内積空間であるとき ι は線形写像である。しかしながら U が複素内積空間の場合は、 ι は線形写像ではない。
 - (5) U が有限次元ならば ι は全単射である。

Proof. (1): 各 $\mathbf{y}_1, \mathbf{y}_2 \in U$ および $r, s \in \mathbb{K}$ について、 $\iota(\mathbf{x})(r\mathbf{y}_1 + s\mathbf{y}_2) = \langle \mathbf{x}, r\mathbf{y}_1 + s\mathbf{y}_2 \rangle = r\langle \mathbf{x}, \mathbf{y}_1 \rangle + s\langle \mathbf{x}, \mathbf{y}_2 \rangle = r\iota(\mathbf{x})(\mathbf{y}_1) + s\iota(\mathbf{x})(\mathbf{y}_2)$ 。

(2): $\iota(\mathbf{x}_1) = \iota(\mathbf{x}_2)$ とすると、これらの両辺に $\mathbf{x}_1 - \mathbf{x}_2$ を代入すれば、 $\langle \mathbf{x}_1, \mathbf{x}_1 - \mathbf{x}_2 \rangle = \langle \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2 \rangle$ 。ゆえに $\langle \mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2 \rangle = 0$ であり、 $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$ 。つまり $\mathbf{x}_1 = \mathbf{x}_2$ である。

(3): 正規直交基底の定義から明らか。

(4): $\mathbb{K} = \mathbb{R}$ の場合: 各 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in U$ および $r, s \in \mathbb{R}$ について、 $\iota(r\mathbf{x}_1 + s\mathbf{x}_2)(\mathbf{y}) = \langle r\mathbf{x}_1 + s\mathbf{x}_2, \mathbf{y} \rangle = r\langle \mathbf{x}_1, \mathbf{y} \rangle + s\langle \mathbf{x}_2, \mathbf{y} \rangle = r\iota(\mathbf{x}_1)(\mathbf{y}) + s\iota(\mathbf{x}_2)(\mathbf{y}) = (r\iota(\mathbf{x}_1) + s\iota(\mathbf{x}_2))(\mathbf{y})$ 。つまり $\iota(r\mathbf{x}_1 + s\mathbf{x}_2) = r\iota(\mathbf{x}_1) + s\iota(\mathbf{x}_2)$ である。

$\mathbb{K} = \mathbb{C}$ の場合: $\iota(r\mathbf{x}_1 + s\mathbf{x}_2) = \bar{r}\iota(\mathbf{x}_1) + \bar{s}\iota(\mathbf{x}_2)$ ゆえ ι は線形ではない。

(5): $\mathbb{K} = \mathbb{R}$ の場合は、 ι の線形性と単射性から明らか (命題 22.4.8)。 $\mathbb{K} = \mathbb{C}$ の場合について全射性を示そう。 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の正規直交基底とする。このとき各 $f \in U^*$ について、 $f = \sum_{i=1}^n c_i \mathbf{u}_i^*$ と書ける (命題 27.2.4)。そこで、 $\mathbf{z} := \sum_{i=1}^n \bar{c}_i \mathbf{u}_i$ とおけば、 $\iota(\mathbf{z}) = f$ である。実際、(3) と (4) で示したことより $\iota(\mathbf{z}) = \iota(\sum_{i=1}^n \bar{c}_i \mathbf{u}_i) = \sum_{i=1}^n \bar{c}_i \iota(\mathbf{u}_i) = \sum_{i=1}^n c_i \mathbf{u}_i^* = f$ 。 \square

備考 27.6.2. $X = [a, b]$ を閉区間とし、 $C(X)$ には例 35.6.2 で与えた内積が定まっているとする。

- (1) $f \in C(X)$ とする。上で与えた $\iota(f): C(X) \rightarrow \mathbb{R}$ について $\iota(f)(g) = \int_a^b f(x)g(x) dx$ である。
- (2) $\mathbf{1} \in C(X)$ について、 $\iota(\mathbf{1})$ は例 27.1.2(3) で与えた S に等しい。
- (3) いかなる $f \in C(X)$ についても、例 27.1.2(1) で与えた q_{x_0} と $\iota(f)$ は一致しない⁴。すなわち、 $q_{x_0} \in C(X)^* \setminus \text{Im } \iota$ である (つまり $\iota: C(X) \rightarrow C(X)^*$ は全射でない)。しかしながら、もし仮に $\iota(\delta_{x_0}) = q_{x_0}$ を満たす仮想的な関数 δ_{x_0} が考えられるとすれば、それはどのような性質を持つことが望ましいかを論じることは無益ではないことが知られている。このような立場で考える仮想的な関数 δ_{x_0} をディラックのデルタ関数という。

前節で述べたように、 $\mathbf{x} \in U$ と $f \in U^*$ のどちらが代入する側であるかは、見方によっていずれにも解釈できる。そこで U の元と U^* の元を対等に扱う立場では $f(\mathbf{x})$ のことを $\langle f, \mathbf{x} \rangle$ と書く。この記号は内積と類似する性質を満たす:

命題 27.6.3. $\mathbf{x}, \mathbf{y} \in U, f, g \in U^*, k \in \mathbb{K}$ について次が成り立つ。

- $\langle f + g, \mathbf{x} \rangle = \langle f, \mathbf{x} \rangle + \langle g, \mathbf{x} \rangle, \quad \langle f, \mathbf{x} + \mathbf{y} \rangle = \langle f, \mathbf{x} \rangle + \langle f, \mathbf{y} \rangle.$

⁴証明略。仮に $\iota(f) = q_{x_0}$ とすると、 $f(x_0)$ の値はいかなる数よりも大きくなる必要があり、それは不可能である。

- $\langle kf, \mathbf{x} \rangle = k\langle f, \mathbf{x} \rangle = \langle f, k\mathbf{x} \rangle$.
- 各 $\mathbf{x} \in U$ について $\langle f, \mathbf{x} \rangle = 0 \implies f = \mathbf{0}_{U^*}$.
- 各 $f \in U^*$ について $\langle f, \mathbf{x} \rangle = 0 \implies \mathbf{x} = \mathbf{0}_U$ (補題 27.4.1).

このような内積との類似性から, $\langle f, \mathbf{x} \rangle = 0$ が成り立つとき, f と \mathbf{x} は直交するという. また, U の部分集合と U^* の部分集合のあいだにも直交性が同様に定義される.

第 V 部

固有空間分解

第28章 固有値と固有ベクトル

特別な基底によるベクトルの分解を考えよう. 分解を行う一義的な理由は, 複雑なベクトルをより単純なベクトルに分解し理解を容易にすることにある. ここでいう単純なベクトルとは, 何を基準にして考えるのだろうか. もちろんそれは, 何をどのような立場で分析したいかによって変わってくる. さて, 議論が抽象的になり過ぎないように, ここでは例として, 次の二つの単純さを考えてみよう:

- (1) 標準的と考えられるベクトルへの分解.
- (2) 与えられた線形写像と相性のよいベクトルへの分解.

これらの違いの一例を 28.1 節で与えたうえで, 以降, 第 V 部全体を通して, (2) による分解とその応用について論じる.

28.1 関数の分解

冒頭で挙げた (1) の最も典型的な例は \mathbb{R}^n のベクトルを標準基底に分解するというものである. また, 解析学で学ぶ関数の冪級数展開もそのような分解の一つと言えるだろう. 例えば, テイラーの定理における剰余項が 0 に収束することを確認することで, 次のような冪級数展開 (テイラー展開) を得る:

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \quad e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

上の三つの冪級数は \mathbb{C} 上でも収束することが確認され, $x = i\theta$ に対してこれらを適用することで次の等式が導かれる¹. ここで, i は虚数単位 (すなわち $i^2 = -1$ を満たす数) とする.

定理 28.1.1 (オイラーの公式). $e^{i\theta} = \cos \theta + i \sin \theta$.

このように, 関数の多項式による分解を通して, 例えば三角関数と指数関数の相互理解が深められるわけである. 一方で, 冪級数展開 $f(x) = \sum_{n=0}^{\infty} a_n x^n$ について次のような計算

$$\frac{d}{dx} f(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}, \quad \int f(x) dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} + C \quad (C \text{ は積分定数})$$

が期待できるものの, この計算は x の冪がずれるため少し煩雑になる. さらに n 階微分になれば煩わしさは増す一方である.

そこで, 微分作用素とより相性の良い関数による分解を提案しよう. 例えば, 閉区間 $[-\pi, \pi]$ を定義域とする関数の多くは次のように分解できることが知られている (フーリエ展開)²:

$$f(x) = \sum_{n=-\infty}^{\infty} z_n e^{inx}, \quad \text{ここで, 各 } z_n := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt \text{ は複素数.}$$

¹三角関数のテイラー展開やオイラーの公式の詳しい証明は, 例えば文献 [6] を見よ.

²ここで各係数 z_n を与える式が上のような理由が分からなかったり, あるいは三角級数展開との関係について気になったりしている読者もおられよう. これらについては 39.8 節を参照されたい.

このとき、 $\frac{d}{dx}e^{ax} = ae^{ax}$ が複素数値関数としても成り立つことをあらかじめ確認しておけば、

$$\frac{d}{dx}f(x) = \sum_{n=-\infty}^{\infty} inz_n e^{inx}, \quad \frac{d^n}{dx^n}f(x) = \sum_{n=-\infty}^{\infty} (in)^n z_n e^{inx}$$

となることが期待される。冪級数展開を微分するよりも上式のほうが計算がやさしい。

このように、与えられた線形写像と相性のよいベクトルによる分解をいかに見つけるかが第 V 部の主題である。また、この問題は線形変換の表現行列をいかに簡単に表すか、という別の基本的な問題とも密接に関連している。まずはこの二つの問題の関係から論じよう。以下では、線形空間 U から U 自身への線形変換、すなわち $\text{End}(U)$ の元のみを考えることにする³。

28.2 固有ベクトル

上で述べた微分作用素に対する指数関数のように $f(\mathbf{u}) = a\mathbf{u}$ を満たすベクトル \mathbf{u} についての議論は、そうでないベクトルについての議論よりも単純であろうことは明白である。このようなベクトルを固有ベクトルと言う。

定義 28.2.1. 線形変換 $f: U \rightarrow U$ において $f(\mathbf{u}) = \lambda\mathbf{u}$ を満たす $\lambda \in \mathbb{R}$ および零ベクトルでない $\mathbf{u} \in U$ が存在するとき、 λ を f における固有値 (eigenvalue) といい、 \mathbf{u} を固有値 λ に関する f の固有ベクトル (eigenvector) という。また、 n 次正方形行列 A による線形変換 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ の固有値および固有ベクトルをそれぞれ A における固有値および固有ベクトルと呼ぶ。

$\mathbf{u} \in U$ が f の固有値 λ_i に関する固有ベクトル \mathbf{v}_i に分解されるとき、すなわち $\mathbf{u} = \sum_{i=1}^n a_i \mathbf{v}_i$ とかけるとき $f(\mathbf{u})$ の値は次のように簡単に計算できる:

$$\begin{aligned} f(\mathbf{u}) &= f\left(\sum_{i=1}^n a_i \mathbf{v}_i\right) = \sum_{i=1}^n a_i f(\mathbf{v}_i) = \sum_{i=1}^n a_i \lambda_i \mathbf{v}_i, \\ f^2(\mathbf{u}) &= f\left(\sum_{i=1}^n a_i \lambda_i \mathbf{v}_i\right) = \sum_{i=1}^n a_i \lambda_i f(\mathbf{v}_i) = \sum_{i=1}^n a_i \lambda_i^2 \mathbf{v}_i, \\ f^3(\mathbf{u}) &= \sum_{i=1}^n a_i \lambda_i^3 \mathbf{v}_i, \quad \dots, \quad f^n(\mathbf{u}) = \sum_{i=1}^n a_i \lambda_i^n \mathbf{v}_i. \end{aligned}$$

例 28.2.2. (1) n 次対角行列 $B = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ および \mathbb{R}^n の標準ベクトル $\mathbf{e}_1, \dots, \mathbf{e}_n$ について、

$B\mathbf{e}_i = \lambda_i \mathbf{e}_i$ である。ゆえに λ_i は B の固有値であり、 \mathbf{e}_i は B の固有値 λ_i に関する固有ベクトルである。

(2) U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ について、各 \mathbf{u}_i が線形変換 $f \in \text{End}(U)$ の固有値 λ_i に関する固有ベクトルであるとすると、このとき、 $f(\mathbf{u}_i) = \lambda_i \mathbf{u}_i$ であるから、基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列は対角行列

$$\text{列} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \text{である.}$$

例 28.2.3. ある自然数 n について $f^n = \mathbf{0}$ を満たす線形変換を冪零 (nilpotent) であるという。また、 $A^n = O$ を満たす行列を冪零であるという。冪零変換の固有値は 0 のみに限る。

³定義域の基底と終域の基底を別々に取ることが許されるならば、表現行列は簡約行列に取ることができ、したがって多くを論じる必要はなくなる。例 26.1.3(2) がその典型的な場合である。

Proof. まず 0 が f の固有値となることを示そう. $\mathbf{x} \neq \mathbf{0}$ なるベクトルを一つとれば, $f^n(\mathbf{x}) = \mathbf{0}$ ゆえ, ある非負整数 m について $f^m(\mathbf{x}) \neq \mathbf{0}$ かつ $f^{m+1}(\mathbf{x}) = \mathbf{0}$ となる. このとき, $\mathbf{y} := f^m(\mathbf{x}) \neq \mathbf{0}$ とすれば, $f(\mathbf{y}) = \mathbf{0} = 0 \cdot \mathbf{y}$. つまり \mathbf{y} は固有値 0 に関する固有ベクトルである.

一方, $\mathbf{x} \neq \mathbf{0}$ を固有値 λ に関する f の固有ベクトルとすれば, $\mathbf{0} = f^n(\mathbf{x}) = \lambda^n \mathbf{x}$ である. いま $\mathbf{x} \neq \mathbf{0}$ としていたから $\lambda^n = 0$ であり⁴, ゆえに $\lambda = 0$. □

例 28.2.4. (1) 例 20.4.4 で与えたシフト作用素 $S: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ ($S((x_n)_{n \in \mathbb{N}}) := (x_{n+1})_{n \in \mathbb{N}}$) において, 初項 a 等比 λ なる等比数列 $\mathbf{x} = (a\lambda^{n-1})_{n \in \mathbb{N}}$ は S の固有値 λ に関する固有ベクトルである.

(2) $y = e^{\lambda x}$ とおくと $y' = \lambda e^{\lambda x}$ である. したがって, 微分作用素 $D: C^\infty(I) \rightarrow C^\infty(I)$ ($D(y) := y'$) において関数 $e^{\lambda x}$ は D の固有値 λ に関する固有ベクトルである.

固有値が存在しない線形変換もある.

例 28.2.5. (1) 平面 \mathbb{R}^2 における 90° 回転を表す行列 $A = \begin{bmatrix} \cos 90^\circ & -\sin 90^\circ \\ \sin 90^\circ & \cos 90^\circ \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ において, 各 $\mathbf{x} \in \mathbb{R}^2$ (ただし $\mathbf{x} \neq \mathbf{0}$) は $A\mathbf{x}$ と平行でないゆえ $A\mathbf{x} \neq \lambda\mathbf{x}$ である. ゆえに体 \mathbb{R} 上の線形空間としての線形写像 $T_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ に固有ベクトルは存在しない.

(2) 列ベクトル成分やスカラー倍に複素数を許す場合を考えよう. (1) における A について, 体 \mathbb{C} 上の線形空間としての線形写像 $T_A: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ は固有ベクトルを持つ. 実際, $\mathbf{x} = \begin{bmatrix} i \\ 1 \end{bmatrix}$ は A の固有値 i に関する固有ベクトルである:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ i \end{bmatrix} = i \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

練習 28.2.6. λ を正方形行列 A の固有値とする. 多項式 $\Psi(t)$ に対して, $\Psi(\lambda)$ が行列 $\Psi(A)$ の固有値になることを示せ.

略解: $\Psi(t) = at^2 + bt + c$ の場合について示そう ($\Psi(t)$ が一般の n 多項式の場合も以下と同様にして示せる). 仮定より, 固有値 λ に関する A の固有ベクトル $\mathbf{v} \neq \mathbf{0}$ が存在する. このとき,

$$\begin{aligned} \Psi(A)\mathbf{v} &= (aA^2 + bA + cE)\mathbf{v} = aA^2\mathbf{v} + bA\mathbf{v} + cE\mathbf{v} \\ &= a\lambda^2\mathbf{v} + b\lambda\mathbf{v} + c\mathbf{v} = (a\lambda^2 + b\lambda + c)\mathbf{v} = \Psi(\lambda)\mathbf{v}. \end{aligned}$$

ゆえに \mathbf{v} は $\Psi(A)$ の固有値 $\Psi(\lambda)$ に関する固有ベクトルである.

28.3 固有ベクトルからなる基底と行列の対角化

各ベクトルを固有ベクトルの線形結合で表すと f^n が容易に計算できることを前節で見た. これと同様に, 対角行列 B において B^n を求めることは容易である:

$$\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{bmatrix}, \quad \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_n^n \end{bmatrix}.$$

実は, 次の二つの問題:

- (i) 線形変換の固有ベクトルからなる基底を求める,
- (ii) 与えられた正方形行列 A に対して $P^{-1}AP$ が対角行列となるような可逆行列 P を探す,

⁴仮に $\lambda^n \neq 0$ とすると, その逆数倍を $\lambda^n \mathbf{x} = \mathbf{0}$ の両辺にほどこせば $\mathbf{x} = \mathbf{0}$ となり矛盾してしまう.

について論じることは次の議論によって同等であることが分かる。

(I) (i) を用いて (ii) を解く：

正方行列 A に対して、標準基底 e_1, \dots, e_n に関する $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ の表現行列は A である。ここで、 T_A の固有ベクトルからなる基底 p_1, \dots, p_n があるとすれば、基底 p_1, \dots, p_n に関する T_A の表現行列 B は例 28.2.2(2) により対角行列である。また、 e_1, \dots, e_n による p_1, \dots, p_n の変換行列は $P = [p_1, \dots, p_n]$ であり、このとき系 26.3.2 より $B = P^{-1}AP$ である。

(II) (ii) を用いて (i) を解く：

U の基底 u_1, \dots, u_n に関する f の表現行列を A とする。ある可逆行列 $P = [p_1, \dots, p_n]$ について

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \text{ となるならば, } [q_1, \dots, q_n] := [u_1, \dots, u_n]P \text{ で定められる}^5 \text{ 新たな基底}$$

q_1, \dots, q_n に関する f の表現行列は、系 26.3.2 により対角行列 $P^{-1}AP$ に等しい。表現行列の定義から $f(q_i) = \lambda_i q_i$ ($i = 1, \dots, n$) であり、ゆえに各 q_i は固有値 λ_i に関する固有ベクトルである。

$B = P^{-1}AP$ が対角行列であるとき、 B^n が容易に計算できることから、 A^n も次のように計算できる：

$$A^n = (PBP^{-1})^n = \underbrace{PBP^{-1} \cdot PBP^{-1} \cdot PBP^{-1} \cdot \dots \cdot PBP^{-1}}_n = PB^nP^{-1}. \quad (28.3.1)$$

定義 28.3.1. 正方行列 A に対して、 $P^{-1}AP$ が対角行列となるように可逆行列 P を取ることができるとき、 A は対角化可能である (**diagonalizable**) という。また、このような P や $P^{-1}AP$ の成分表示を求めるプロセスを A の対角化 (**diagonalization**) という。

A が対角化可能であるとき、対角行列 $P^{-1}AP$ を指す呼称がないと不便であるから、本書ではこれを A の対角化行列あるいは略して A の対角化と呼ぶことにしよう⁷。

行列の対角化可能性については次章で詳しく解説する。まずは線形変換の固有ベクトルをいかに求めるかを論じよう。

28.4 特性多項式

線形空間 U における恒等写像 $\text{id}_U \in \text{End}(U)$ を I と書くのであった。線形変換 $f : U \rightarrow U$ および $\lambda \in \mathbb{R}$ について、 $x \in U$ (ただし $x \neq \mathbf{0}$) が λ に関する固有ベクトルとなる条件は次のように言い換えることができる：

$$\begin{aligned} f(x) = \lambda x &\iff \lambda x - f(x) = \mathbf{0} \iff \lambda I(x) - f(x) = \mathbf{0} \\ &\iff (\lambda I - f)(x) = \mathbf{0} \iff x \in \text{Ker}(\lambda I - f). \end{aligned}$$

したがって、固有値 λ に関する f の固有ベクトルが存在することは、 $\text{Ker}(\lambda I - f) \neq \{\mathbf{0}\}$ と同値であり、このとき λ に関する固有ベクトル全体のなす集合は、 $\text{Ker}(\lambda I - f)$ から $\mathbf{0}$ を除いた集合に等しい。以上をまとめると、

命題 28.4.1. 線形空間 U および $f \in \text{End}(U)$, $\lambda \in \mathbb{R}$ について次は同値である。

- (1) λ は f の固有値である。
- (2) $\text{Ker}(\lambda I - f)$ は自明でない、すなわち $\text{Ker}(\lambda I - f) \neq \{\mathbf{0}\}$ 。
- (3) $\lambda I - f$ は単射でない。

⁵すなわち、 $q_i := [u_1, \dots, u_n]p_i$ である。

⁶命題 25.4.1(2) により q_1, \dots, q_n は U の基底である。

⁷これは本書に限る用法である。広く知られている対角化という用語は、あくまでもプロセスのことを指す。

とくに $U = \mathbb{R}^n$ および $f = T_A$ について, $(\lambda I - T_A)(\mathbf{x}) = (\lambda E - A)\mathbf{x}$ であり, $\text{Ker}(\lambda I - T_A)$ が自明でないことは, 方程式 $(\lambda E - A)\mathbf{x} = \mathbf{0}$ が自明でない解をもつことに他ならない. これは, 定理 6.2.2 により A が可逆でないこと, すなわち $|\lambda E - A| = 0$ と同値である.

また, 一般の n 次元線形空間 U においても, U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列を A とすれば, $\lambda I - f$ の表現行列は $\lambda E - A$ である (1 次多項式 $\Psi(t) = \lambda - t$ について系 26.2.5 を適用せよ). $\lambda I - f$ が単射でないことは $T_{\lambda E - A}$ が単射でないことに同値であり (命題 26.4.1), これは $|\lambda E - A| = 0$ と同値である.

以上の考察から次が成り立つ:

系 28.4.2. $n \in \mathbb{N}$ および $\lambda \in \mathbb{R}$ とする.

(1) n 次正方行列 A について, 次が成り立つ: λ は A の固有値である $\iff |\lambda E - A| = 0$.

(2) n 次元線形空間 U のある基底に関する $f \in \text{End}(U)$ の表現行列を A とすれば,

$$\lambda \text{ は } f \text{ の固有値である } \iff |\lambda E - A| = 0.$$

以上により, 固有値を求めるには t に関する方程式 $|tE - A| = 0$ を解けばよいことが分かった.

定義 28.4.3. 正方行列 A に対して多項式 $\Phi_A(t) := |tE - A|$ を A の特性多項式 (characteristic polynomial) あるいは固有多項式と呼ぶ. また, 有限次元線形空間 U において, その上の線形変換 $f \in \text{End}(U)$ の固有多項式 $\Phi_f(t)$ を, U のある基底に関する f の表現行列の固有多項式として定める. 次の命題により, $\Phi_f(t)$ は基底の選び方に依らずに決まる多項式である. 固有多項式の零点を求める方程式 $\Phi_A(t) = 0$ および $\Phi_f(t) = 0$ を特性方程式 (characteristic equation) という.

【補足】特性方程式の解, すなわち固有値のことを特性根 (characteristic root) あるいは特性解と呼ぶ場合もある.

命題 28.4.4. $f: U \rightarrow U$ を線形変換とし, 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による f の表現行列を A , 基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ による f の表現行列を B とすれば $\Phi_A = \Phi_B$.

Proof. 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P とすれば系 26.3.2 より $B = P^{-1}AP$ が成り立つ. このとき

$$\begin{aligned} \Phi_B(t) &= |tE - B| = |tP^{-1}EP - P^{-1}AP| = |P^{-1}(tE)P - P^{-1}AP| = |P^{-1}((tE)P - AP)| \\ &= |P^{-1}(tE - A)P| = |P^{-1}| \cdot |tE - A| \cdot |P| = |tE - A| = \Phi_A(t). \end{aligned}$$

□

上の証明から次の事実も導かれる:

系 28.4.5. 相似な正方行列の特性多項式は等しい.

例 28.4.6. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ について

$$\begin{aligned} \Phi_A(t) &= \begin{vmatrix} t-a & -b \\ -c & t-d \end{vmatrix} = (t-a)(t-d) - (-b)(-c) \\ &= t^2 - (a+d)t + (ad-bc) = t^2 - (\text{tr } A)t + \det A. \end{aligned}$$

例 28.4.7. 上三角行列 $A = \begin{bmatrix} a_1 & & * \\ & a_2 & * \\ O & \ddots & a_n \end{bmatrix}$ について, 例 11.1.2 より

$$\Phi_A(t) = \begin{vmatrix} t-a_1 & & * \\ & t-a_2 & * \\ O & \ddots & t-a_n \end{vmatrix} = (t-a_1)(t-a_2)\cdots(t-a_n).$$

すなわち、 A の対角成分は A の固有値である。

練習 28.4.8. n 次正方形行列 A について、 $\Phi_A(t)$ は n 次多項式であることを確認せよ。

解答例: $tE - A = [b_{ij}]$ の行列式 $|B| = \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)}$ の各項のうち、 t の n 次多項式が現れるのは $\sigma = \text{id}$ に関する項 $\text{sgn}(\text{id}) b_{11} b_{22} \cdots b_{nn} = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$ のみであり、他の置換に関する項はすべて $n - 1$ 次以下の多項式となる。ゆえにこれらの和は n 次多項式である。□

練習 28.4.9. 有限次元線形空間 U 上の線形変換 $f: U \rightarrow U$ について、その固有多項式 $\Phi_f(t)$ の次数は $\dim U$ に等しいことを示せ。

解答例: $n := \dim W$ とする。 W の基底を一つ取り、その基底に関する f の表現行列を A とすれば、 A は n 次正方形行列である。したがって $\Phi_f(t) = \Phi_A(t)$ は、練習 28.4.8 より n 次多項式である。□

練習 28.4.10. A を m 次正方形行列、 B を n 次正方形行列とする。 $m + n$ 次正方形行列 $X = \begin{bmatrix} A & * \\ O & B \end{bmatrix}$ について $\Phi_X(t) = \Phi_A(t)\Phi_B(t)$ を示せ。

解答例: 定理 12.5.1 より明らか。実際、

$$\Phi_X(t) = \begin{vmatrix} tE_m - A & * \\ O & tE_n - B \end{vmatrix} = |tE_m - A| \cdot |tE_n - B| = \Phi_A(t)\Phi_B(t).$$

□

練習 28.4.11. 正方形行列 A について $\Phi_{tA} = \Phi_A$ を示せ。

解答例: 転置行列と、もとの行列の行列式が等しいことから明らか。転置記号との混同をさけるため、多項式の変数を x として計算すると、 $\Phi_{tA}(x) = |xE - {}^tA| = |{}^t(xE - {}^tA)| = |{}^t(xE) - {}^t({}^tA)| = |xE - A| = \Phi_A(x)$ 。□

28.5 固有空間

線形変換の固有ベクトルをすべて求める方法を考えよう。28.4 節の冒頭で論じたように、線形変換 $f: U \rightarrow U$ および $\lambda \in \mathbb{R}$ について、 $\text{Ker}(\lambda I - f) \neq \{\mathbf{0}\}$ であるとき、 $\text{Ker}(\lambda I - f)$ から $\mathbf{0}$ ベクトルを除いた集合は f の固有値 λ に関する固有ベクトル全体に一致する。そこで、

定義 28.5.1. $\lambda \in \mathbb{R}$ を $f \in \text{End}(U)$ の固有値とする。このとき、 U の部分空間 $\text{Ker}(\lambda I - f)$ を f の固有値 λ に関する固有空間 (eigenspace) と呼び、本書ではこれを $W(\lambda, f)$ で表す。また、 n 次正方形行列 A について、 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ の固有値 λ における固有空間を $W(\lambda, A)$ と書く。すなわち、 $W(\lambda, A)$ は連立 1 次方程式 $(\lambda E - A)x = \mathbf{0}$ の解空間 $W_{\lambda E - A}$ である。

例題 28.5.2. 行列 $A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$ について、次の問いに答えよ。

(1) A の特性多項式 $\Phi_A(t)$ を求めよ。

解答例:

$$\begin{aligned} \Phi_A(t) &= |tE - A| = \left| \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} - \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix} \right| = \begin{vmatrix} t-8 & 10 \\ -5 & t+7 \end{vmatrix} \\ &= (t-8)(t+7) - (-5) \cdot 10 = t^2 - t - 56 + 50 = t^2 - t - 6 = (t-3)(t+2). \end{aligned}$$

(2) A の固有値をすべて求めよ。

解答例: 方程式 $\Phi_A(t) = 0$ を解いて $t = 3, -2$ 。系 28.4.2 により、 3 と -2 が A の固有値である。

(3) A の各固有値に関する固有空間を求めよ (固有空間をその基底による外延的表記で表せ).

解答例: A の固有値 $\lambda = -2, 3$ について連立 1 次方程式 $(\lambda E - A)\mathbf{x} = \mathbf{0}$ をそれぞれ解けばよい. いま $\Phi_A(\lambda) = 0$ に注意して $\lambda E - A$ の簡約化を行うと

$$\lambda E - A = \begin{bmatrix} \lambda - 8 & 10 \\ -5 & \lambda + 7 \end{bmatrix} \rightarrow \begin{bmatrix} -5 & \lambda + 7 \\ \lambda - 8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -\frac{\lambda+7}{5} \\ \lambda - 8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -\frac{\lambda+7}{5} \\ 0 & \frac{(\lambda+7)(\lambda-8)}{5} + 10 \end{bmatrix},$$

ここで, 最後の (2, 2)-成分は

$$\frac{(\lambda+7)(\lambda-8)}{5} + 10 = \frac{1}{5}((\lambda+7)(\lambda-8) + 50) = \frac{1}{5}\Phi_A(\lambda) = 0.$$

したがって $\lambda = -2, 3$ における $(\lambda E - A)\mathbf{x} = \mathbf{0}$ の解はそれぞれ次のようになる:

$$W(-2, A) = \left\{ r \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mid r \in \mathbb{R} \right\}, \quad W(3, A) = \left\{ r \begin{bmatrix} 2 \\ 1 \end{bmatrix} \mid r \in \mathbb{R} \right\}.$$

上の例題における各固有空間の基底による組 $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ は線形独立である. これは偶然ではなく, 一般論として命題 29.1.1 にて示す. $\mathbf{v}_1, \mathbf{v}_2$ は \mathbb{R}^2 の基底であり, この基底に関する T_A の表現行列を B とすれば例 28.2.2(2) (より具体的には例題 26.1.4(1)) より $B = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$ となる. また, 標準基底 $\mathbf{e}_1, \mathbf{e}_2$ による基底 $\mathbf{v}_1, \mathbf{v}_2$ に関する変換行列は $P = [\mathbf{v}_1, \mathbf{v}_2] = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ である. ゆえに系 26.3.2 より $B = P^{-1}AP$ を得る.

例題 28.5.3. $A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$ について, A^n を求めよ.

解答例: $B = P^{-1}AP$ とすれば, 式 (28.3.1) より $A^n = PB^nP^{-1}$ である. $P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ とすれば上で論じたことから $B = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$ である. また, P の逆行列を計算すれば $P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$ である. よって,

$$A^n = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (-2)^n & 0 \\ 0 & 3^n \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} (-2)^n & 2 \cdot 3^n \\ (-2)^n & 3^n \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -(-2)^n + 2 \cdot 3^n & 2(-2)^n - 2 \cdot 3^n \\ -(-2)^n + 3^n & 2(-2)^n - 3^n \end{bmatrix}.$$

練習 28.5.4. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ において $\Phi_A(t) = 0$ の解が λ_1, λ_2 であるとき, A の固有空間を求めよ.

解答例: $c \neq 0$ の場合と $c = 0$ の場合に分けて考える.

(1) $c \neq 0$ の場合:

$$\begin{aligned} \lambda_i E - A &= \begin{bmatrix} \lambda_i - a & -b \\ -c & \lambda_i - d \end{bmatrix} \rightarrow \begin{bmatrix} -c & \lambda_i - d \\ \lambda_i - a & -b \end{bmatrix} \\ &\rightarrow \begin{bmatrix} 1 & -(\lambda_i - d)/c \\ \lambda_i - a & -b \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -(\lambda_i - d)/c \\ 0 & -b + (\lambda_i - a)(\lambda_i - d)/c \end{bmatrix}. \end{aligned}$$

$\Phi_A(t) = (t-a)(t-d) - bc$ より, 上の行基本変形における最後の (2, 2)-成分は

$$-b + \frac{(\lambda_i - a)(\lambda_i - d)}{c} = \frac{-bc + (\lambda_i - a)(\lambda_i - d)}{c} = \frac{\Phi_A(\lambda_i)}{c} = \frac{0}{c} = 0.$$

ゆえに固有値 λ_i ($i = 1, 2$) に関する固有空間は, $W(\lambda_i, A) = \left\langle \begin{bmatrix} (\lambda_i - d)/c \\ 1 \end{bmatrix} \right\rangle$.

(2) $c = 0$ の場合: $\Phi(t) = (t - a)(t - d)$ より $\lambda_1 = a, \lambda_2 = d$ としてよい. b が 0 かどうかで更に場合分けをする.

- $b = 0$ ならば A は対角行列であり, a, d が A の固有値となる. $a = d$ の場合は T_A は各ベクトルを a 倍する写像に他ならず, 固有値 a に関する固有空間は全空間 $W(a, A) = \mathbb{R}^2$ である. $a \neq d$ の場合は, $W(\lambda_i, A) = \langle e_i \rangle$ ($i = 1, 2$) となる.
- $b \neq 0$ のとき, 固有値 a に関する固有空間は,

$$aE - A = \begin{bmatrix} 0 & -b \\ 0 & a - d \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 \\ 0 & a - d \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

より $W(a, A) = \langle e_1 \rangle$ である. 更に $a \neq d$ の場合について, 固有値 d に関する固有空間は,

$$dE - A = \begin{bmatrix} d - a & -b \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -b/(d - a) \\ 0 & 0 \end{bmatrix}$$

$$\text{より } W(d, A) = \left\langle \begin{bmatrix} b/(d - a) \\ 1 \end{bmatrix} \right\rangle.$$

28.6 一般の線形写像の固有空間

一般の線形写像の固有値も, 系 28.4.2 により特性方程式を解くことで求められる. f の固有ベクトルと, その表現行列の固有ベクトルがいかに対応づけられるかを確認しておこう.

命題 28.6.1. U を n 次元線形空間とし $f: U \rightarrow U$ を線形変換とする. 更に, U の基底 v_1, \dots, v_n による f の表現行列を A とすれば次が成り立つ.

- (1) 各 v_i を $e_i \in \mathbb{R}^n$ に対応させる線形同型を $F: U \rightarrow \mathbb{R}^n$ とする. このとき, $u \in U$ が f の固有値 λ に関する固有ベクトルであることと, $F(u)$ が A の固有値 λ に関する固有ベクトルであることは同値である.

【注意】 $x \in \mathbb{R}^n$ とすると, F は $[v_1, \dots, v_n]x$ を x に写す写像である (25.4 節). つまり, $[v_1, \dots, v_n]x$ が f の固有値 λ に関する固有ベクトルであることと, x が A の固有値 λ に関する固有ベクトルであることは同値である.

- (2) λ を f の固有値 (および A の固有値) とする. 列ベクトルの組 $x_1, \dots, x_k \in \mathbb{R}^n$ について, $y_i = [v_1, \dots, v_n]x_i$ ($i = 1, \dots, k$) とすれば, y_1, \dots, y_k が $W(\lambda, f)$ の基底であることと x_1, \dots, x_k が $W(\lambda, A)$ の基底であることは同値である.

- (3) λ を f の固有値 (および A の固有値) とすると, $\dim W(\lambda, f) = \dim W(\lambda, A)$.

Proof. (1): $g = \lambda I - f$ の表現行列は系 26.2.5 より $G = \lambda E - A$ である. また, 命題 26.4.2(2) の証明によれば, $\text{Ker } g$ と $\text{Ker } T_G$ は F によって 1 対 1 に対応し, これらは線形同型である. つまり, 「 $u \in \text{Ker } g \Leftrightarrow F(u) \in \text{Ker } T_G$ 」であり, この同値性は示すべき主張を導く.

(2) および (3): λ が固有値であることから, (1) で与えた g, G について, $W(\lambda, f) = \text{Ker } g$ および $W(\lambda, A) = \text{Ker } T_G$ であり, 対応 $F|_{W(\lambda, f)}: W(\lambda, f) \rightarrow W(\lambda, A)$ ($[v_1, \dots, v_n]x \mapsto x$) が線形同型であることから明らか. \square

【補足】 上の (1) の証明が抽象的と感じる場合は, 次のような同値変形を行って確認してもよい:

命題 28.6.1(1) の別証明. 【注意】に述べた同値性を示そう. $\mathbf{x} = {}^t[x_1, \dots, x_n] \neq \mathbf{0}$ とし, $\mathbf{u} = [\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x} = \sum_{i=1}^n x_i \mathbf{v}_i$ とおく. 表現行列の定義より $[f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_n]A$ に注意すると,

$$\begin{aligned} \mathbf{u} \text{ は } \lambda \text{ に関する } f \text{ の固有ベクトルである} &\iff f\left(\sum_{i=1}^n x_i \mathbf{v}_i\right) = \lambda \mathbf{u} \\ \iff \sum_{i=1}^n x_i f(\mathbf{v}_i) = \lambda [\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x} &\iff [f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)]\mathbf{x} = [\mathbf{v}_1, \dots, \mathbf{v}_n](\lambda \mathbf{x}) \\ \iff [\mathbf{v}_1, \dots, \mathbf{v}_n]A\mathbf{x} = [\mathbf{v}_1, \dots, \mathbf{v}_n](\lambda \mathbf{x}) & \\ \iff A\mathbf{x} = \lambda \mathbf{x} \iff \mathbf{x} \text{ は } \lambda \text{ に関する } A \text{ の固有ベクトルである.} & \\ (\uparrow \text{ここで } \mathbf{v}_1, \dots, \mathbf{v}_n \text{ の線形独立性および練習 25.3.1 を用いた.}) & \end{aligned}$$

□

例 28.6.2. 線形写像 $D: \mathbb{R}[x]_3 \rightarrow \mathbb{R}[x]_3$ を $D(p(x)) := \frac{d}{dx}p(x)$ と定める. 基底 $\mathbf{1}, x, x^2, x^3$ に関する D の

表現行列は, 例題 26.1.4(2) より $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ である. したがって $\Phi_D(t) = |tE - A| = t^4$ であり,

D の固有値は 0 のみである. また, A の固有空間は, 連立 1 次方程式 $(0E - A)\mathbf{x} = \mathbf{0}$ を解いて

$$W(0, A) = \left\{ r \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mid r \in \mathbb{R} \right\}$$

である. これに対応する $\mathbb{R}[x]_3$ の部分空間が D の固有空間 $W(0, D)$ となる. $W(0, A)$ の各元 $\begin{bmatrix} r \\ 0 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^4$

に対応する $\mathbb{R}[x]_3$ の元は命題 28.6.1 より

$$[\mathbf{1}, x, x^2, x^3] \begin{bmatrix} r \\ 0 \\ 0 \\ 0 \end{bmatrix} = r\mathbf{1} + 0x + 0x^2 + 0x^3 = r\mathbf{1} \quad (\text{定数関数})$$

である. つまり, $W(0, D) = \langle \mathbf{1} \rangle$.

第29章 固有空間分解と行列の対角化

前章において、線形変換 $f: U \rightarrow U$ の固有値および固有ベクトルを求める方法を与えた。また、 U の各ベクトル \mathbf{u} を固有ベクトルに分解すると $f(\mathbf{u})$ の計算が容易になることを見た。しかしながら、そのためには任意のベクトルが固有ベクトルの線形結合で書けるかどうかが問題になる。残念ながら一般の線形変換についてこのような分解ができるわけではなく、本章では、この分解が可能であるための必要十分条件を提示する。

ところで、固有値を求めるには特性方程式 $\Phi_f(t) = 0$ を解くことが前提となっている。つまり、特性方程式が実数解を持たない場合には、数の範囲を複素数にまで広げたほうが都合がよい場合がある。ただし、その際に行列の各成分が複素数であることを許すのか、あるいは実数に限って考えているのか、その立場を曖昧にしてはならない。

そこで本章以降においては、記号 \mathbb{K} で \mathbb{R} か \mathbb{C} のいずれかを表すとし、「線形空間」とは体 \mathbb{K} 上の線形空間のこととする。また、「行列」は \mathbb{K} の元を各成分に持つ行列を意味する。更に「多項式」は \mathbb{K} の元を係数に持つ多項式のことであり、特性方程式の解(すなわち線形写像の固有値)は \mathbb{K} の元であるとする。以上の設定のもとで、 $\mathbb{K} = \mathbb{R}$ の場合としてこれまでに述べてきたことは $\mathbb{K} = \mathbb{C}$ の場合においても成り立つ。

以降の内容において、実数または複素数のいずれかに限定するものについては、その都度、 \mathbb{K} がどちらであるか明確にした上で説明する。とくに $\mathbb{K} = \mathbb{R}$ とするとき、上の約束のもとでは、特性方程式の虚数解は固有値に含めない点に留意すること。

29.1 固有ベクトルからなる組の線形独立性

次の命題は、異なる固有値に関する固有ベクトルがどのように配置されているかを理解するうえで基本となる事実である。そこで、本章では3回にわたってこれを証明しよう。まずは固有ベクトルの定義さえ知っていれば済む証明を与える。

命題 29.1.1. $f: U \rightarrow U$ を線形変換とし、その相異なる固有値を $\lambda_1, \dots, \lambda_r$ とする。このとき、各固有空間 $W(\lambda_k, f)$ ごとに線形独立な一組 $\mathbf{u}_{k,1}, \dots, \mathbf{u}_{k,n_k}$ を選べば、これらをすべて集めたベクトルの組 $\mathcal{B} = \{\mathbf{u}_{k,i} \mid k = 1, \dots, r, i = 1, \dots, n_k\}$ は線形独立である。

Proof. 背理法により示す。そこで \mathcal{B} が線形従属であると仮定する。 \mathcal{B} の中からいくつかのベクトルを選び、 $\langle \mathcal{B} \rangle$ の基底を取ることができる(命題 18.3.1)。これを $\mathbf{b}_1, \dots, \mathbf{b}_\ell \in \mathcal{B}$ とする。 \mathcal{B} 自身は線形従属ゆえ、 $\mathbf{b}_1, \dots, \mathbf{b}_\ell$ のいずれでもない $\mathbf{b} \in \mathcal{B}$ が取れる。ここで、各 \mathbf{b}_j ($j = 1, \dots, \ell$) は固有値 γ_j に関する固有ベクトルであるとし(γ_j には重複がある可能性もある)、また \mathbf{b} を固有値 γ に関する固有ベクトルとする。上で取った基底による \mathbf{b} の線形結合表示を $\mathbf{b} = \sum_{j=1}^{\ell} c_j \mathbf{b}_j$ とし、この両辺に f をほどこすと、

$$\begin{aligned} f(\mathbf{b}) &= c_1 f(\mathbf{b}_1) + \dots + c_\ell f(\mathbf{b}_\ell) \\ \gamma \mathbf{b} &= c_1 \gamma_1 \mathbf{b}_1 + \dots + c_\ell \gamma_\ell \mathbf{b}_\ell \\ \gamma(c_1 \mathbf{b}_1 + \dots + c_\ell \mathbf{b}_\ell) &= c_1 \gamma_1 \mathbf{b}_1 + \dots + c_\ell \gamma_\ell \mathbf{b}_\ell \\ \mathbf{0} &= \sum_{j=1}^{\ell} c_j (\gamma_j - \gamma) \mathbf{b}_j. \end{aligned}$$

$\mathbf{b}_1, \dots, \mathbf{b}_\ell$ の線形独立性より, $c_j(\gamma_j - \gamma) = 0$ ($j = 1, \dots, \ell$) であり, とくに γ_j と γ が異なる固有値であれば $\gamma_j - \gamma \neq 0$ ゆえ $c_j = 0$ を得る. したがって, 線形結合 $\mathbf{b} = \sum_{j=1}^{\ell} c_j \mathbf{b}_j$ に現れる係数 c_j のうち, 0でないものは $\gamma_j = \gamma$ の場合, すなわち $\mathbf{b}_j \in W(\gamma, f)$ の場合に限る. さて, γ は $\lambda_1, \dots, \lambda_r$ のいずれかであるから, これを $\gamma = \lambda_{k_0}$ とすれば, $\mathbf{b}_j \in W(\gamma, f)$ をみたま \mathbf{b}_j および \mathbf{b} は, $\mathbf{u}_{k_0,1}, \dots, \mathbf{u}_{k_0,n_{k_0}}$ のいずれかである. 命題 17.2.5 より $\mathbf{u}_{k_0,1}, \dots, \mathbf{u}_{k_0,n_{k_0}}$ は線形従属となり, これは仮定 ($\mathbf{u}_{k_0,1}, \dots, \mathbf{u}_{k_0,n_{k_0}}$ の線形独立性) に反する. \square

系 29.1.2. A を n 次正方行列とし, その相異なる固有値を $\lambda_1, \dots, \lambda_r$ とすれば, $\sum_{k=1}^r \dim W(\lambda_k, A) \leq n$.

固有空間の次元の情報について, 上の系よりもさらに詳しいことが分かっている (系 29.1.6).

さて, 背理法による証明には, 命題が成立する様子がいま一つ掴めないところがある. そこで, 背理法を用いない命題 29.1.1 の証明について検討したい. このような考えに興味がない者は, 直ちに次節に進んで構わない.

補題 29.1.3. 線形変換 $f : U \rightarrow U$ の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とする. $\mathbf{v}_k \in W(\lambda_k, f)$ ($k = 1, 2, \dots, r$) かつ $\mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_r = \mathbf{0}$ ならば, $\mathbf{v}_1 = \mathbf{v}_2 = \dots = \mathbf{v}_r = \mathbf{0}$.

Proof. $\sum_{k=1}^r \mathbf{v}_k = \mathbf{0}$ の両辺に f をほどこすことで $\sum_{k=1}^r \lambda_k \mathbf{v}_k = \mathbf{0}$ を得る. この式に更に f をほどこせば, $\sum_{k=1}^r \lambda_k^2 \mathbf{v}_k = \mathbf{0}$ であり, この操作を順次繰り返すことで次の等式たちを得る:

$$\begin{aligned} \mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_r &= \mathbf{0}, \\ \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_r \mathbf{v}_r &= \mathbf{0}, \\ \lambda_1^2 \mathbf{v}_1 + \lambda_2^2 \mathbf{v}_2 + \dots + \lambda_r^2 \mathbf{v}_r &= \mathbf{0}, \\ &\vdots \\ \lambda_1^{r-1} \mathbf{v}_1 + \lambda_2^{r-1} \mathbf{v}_2 + \dots + \lambda_r^{r-1} \mathbf{v}_r &= \mathbf{0}. \end{aligned}$$

$$\text{つまり, } P = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{r-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_r & \lambda_r^2 & \dots & \lambda_r^{r-1} \end{bmatrix} \quad \text{とおけば} \quad [\mathbf{v}_1, \dots, \mathbf{v}_r]P = [\mathbf{0}, \dots, \mathbf{0}].$$

各 λ_k は相異なる固有値としていたゆえ, $1 \leq p < q \leq r$ ならば $\lambda_q - \lambda_p \neq 0$ である. したがってヴァンデルモンドの行列式 (定理 11.2.2) より $|P| \neq 0$. つまり P は可逆であり, $[\mathbf{v}_1, \dots, \mathbf{v}_r]P = [\mathbf{0}, \dots, \mathbf{0}]$ の両辺に右から P^{-1} をかけることで $[\mathbf{v}_1, \dots, \mathbf{v}_r] = [\mathbf{0}, \dots, \mathbf{0}]$ を得る. \square

補題 29.1.3 の証明では強力な公式 (ヴァンデルモンドの行列式) に議論を帰着させている. 公式を用いて示すということは, 議論の一部をブラックボックス化して処理することに他ならない. ブラックボックスを用いた証明は, 場合によっては, 定理が成立する背景はよく分からないものの正しい主張であることだけは分かった, という理解になりかねない. そこで, ブラックボックス化せずに上を示す方法について, 30 章で論じる.

命題 29.1.1 の別証明.

$$\sum_{k=1}^r \left(\sum_{i=1}^{n_k} c_{k,i} \mathbf{u}_{k,i} \right) = \mathbf{0}$$

と仮定し, $c_{k,i} = 0$ を示そう. $\mathbf{v}_k = \sum_{i=1}^{n_k} c_{k,i} \mathbf{u}_{k,i}$ とおけば $\sum_{k=1}^r \mathbf{v}_k = \mathbf{0}$ である. 補題 29.1.3 より $\mathbf{v}_1 = \mathbf{v}_2 = \dots = \mathbf{v}_r = \mathbf{0}$ であり, ゆえに, いずれの k についても $\sum_{i=1}^{n_k} c_{k,i} \mathbf{u}_{k,i} = \mathbf{0}$. これと $\mathbf{u}_{k,1}, \dots, \mathbf{u}_{k,n_k}$ の線形独立性を合わせて $c_{k,i} = 0$ を得る. \square

【補足】 補題 29.1.3 にある性質から命題 29.1.1 を証明する手法は, 直和分解なる概念を通して一般論として論じられる. 詳しくは 33.5 章を見よ.

任意のベクトルは, 二つ以上の固有値に関する共通の固有ベクトルにはならない. これは固有ベクトルの定義から直ちに分かることであるものの, 次のように示すこともできる.

系 29.1.4. 線形変換 $f : U \rightarrow U$ の相異なる二つの固有値 λ, γ について, $W(\lambda, f) \cap W(\gamma, f) = \{\mathbf{0}\}$.

Proof. $W := W(\lambda, f) \cap W(\gamma, f)$ とおく, $\{\mathbf{0}\} \subset W$ は明らか, $\mathbf{x} \in W$ とする. $\mathbf{v}_1 = \mathbf{x}$, $\mathbf{v}_2 = -\mathbf{x}$ とおけば $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}$ であり, $r = 2$ の場合として補題 29.1.3 を適用すれば $\mathbf{x} = \mathbf{0}$ を得る. \square

次の命題と系は, 後で学ぶ定理 34.1.1 から直ちに導かれる主張でもあり, ここであえて証明しなくてもよい. なお, これらの事実を知っていると行列の対角化可能性の判定において計算の量を減らせる場合がある.

命題 29.1.5 (発展). 有限次元線形空間 U 上の線形変換 $f : U \rightarrow U$ の特性多項式が $\Phi_f(t) = (t - \lambda)^k \Psi(t)$ と分解されるとき (ただし, $\lambda \in \mathbb{K}$, $\Psi(\lambda) \neq 0$ とする), $\dim W(\lambda, f) \leq k$ が成り立つ.

Proof. $n := \dim U$ とおく. 求めるべき不等式を背理法により示そう. そこで $\dim W(\lambda, f) \geq k + 1$ であると仮定する. このとき, 線形独立な組 $\mathbf{u}_1, \dots, \mathbf{u}_{k+1} \in W(\lambda, f)$ が取れる. これらの組にさらに U の元を加えることで, U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を作る事ができる (命題 18.3.2). $\mathbf{u}_1, \dots, \mathbf{u}_{k+1}$ が f の固有値 λ に関する固有ベクトルであることから, 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列は次の形になる:

$$B = \left(\begin{array}{ccc|ccc} \lambda & & & & & \\ & \ddots & & & & \\ & & \lambda & & & \\ \hline 0 & \cdots & 0 & & & \\ \vdots & \ddots & \vdots & & & \\ 0 & \cdots & 0 & & & \end{array} \right) \begin{array}{l} \left. \begin{array}{l} * \\ * \\ * \end{array} \right\} k+1 \\ \\ \left. \begin{array}{l} C \\ C \\ C \end{array} \right\} n - (k+1) \end{array} .$$

ゆえに $\Phi_f(t) = \Phi_B(t) = |tE - B| = (t - \lambda)^{k+1} \Phi_C(t)$. これに仮定を合わせると $(t - \lambda)^{k+1} \Phi_C(t) = (t - \lambda)^k \Psi(t)$ であり, この両辺を $(t - \lambda)^k$ で割って $(t - \lambda) \Phi_C(t) = \Psi(t)$ を得る¹. この両辺に $t = \lambda$ を代入することで $\Psi(\lambda) = 0$ が得られ, これは仮定に反する. 以上より, $\dim W(\lambda, f) \leq k$ でなければならない. \square

前命題より直ちに次の主張を得る.

系 29.1.6 (発展). n 次元線形空間 U 上の線形変換 $f : U \rightarrow U$ の特性多項式が, 相異なる $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ を用いて $\Phi_f(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r} \Psi(t)$ と分解されるとき (ただし, 各 $k = 1, \dots, r$ について $\Psi(\lambda_k) \neq 0$ とする), $\dim W(\lambda_k, f) \leq n_k$ が成り立つ.

【注意】 $\Phi_f(t)$ が \mathbb{K} 係数 1 次式の積に因数分解できるときは, $\Psi(t) = 1$ の場合に相当する.

先に述べた系 29.1.2 は, 上の系からも導かれる (上の系において $n_1 + n_2 + \cdots + n_r \leq n$ であることに注意せよ). この意味において, 系 29.1.6 のほうが系 29.1.2 よりも詳しい.

29.2 対角化可能条件

次の定理は正方行列の対角化可能性の必要十分条件を与えるとともに, その証明において対角化を求める方法も提示している. 証明の一部 ((2) \Leftrightarrow (3)) は 28.3 節で述べたことと重複するが, ここでは復習も兼ねて述べておこう.

定理 29.2.1. A を n 次正方行列とすれば次は同値.

(1) A の相異なるすべての固有値を $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ とすれば, $\sum_{k=1}^r \dim W(\lambda_k, A) = n$.

(注意: この条件は系 29.1.2 より $\sum_{k=1}^r \dim W(\lambda_k, A) \geq n$ と同値である.)

¹多項式の除算について, 詳しくは補題 29.4.9 をみよ.

(2) A の固有ベクトルからなる \mathbb{K}^n の基底が存在する.

(3) A は対角化できる. すなわち $P^{-1}AP$ が対角行列となるような可逆行列 P が存在する.

Proof. (1) \Rightarrow (2): 各 $\lambda = \lambda_1, \dots, \lambda_r$ について $n_\lambda = \dim W(\lambda, A)$ とおく. $W(\lambda, A)$ の基底を一つ取り, これらを集めて $\mathcal{B} = \{\mathbf{u}_{\lambda,i} \mid \lambda = \lambda_1, \dots, \lambda_r, i = 1, \dots, n_\lambda\}$ とすれば, 命題 29.1.1 より \mathcal{B} は線形独立である. また, 仮定より \mathcal{B} の元の個数は n に等しいゆえ \mathcal{B} は \mathbb{K}^n の基底となる.

(2) \Rightarrow (3): 列ベクトルの組 $\mathbf{p}_1, \dots, \mathbf{p}_n$ を A の固有ベクトルからなる \mathbb{K}^n の基底とし, \mathbf{p}_i は固有値 γ_i に関する固有ベクトルとする (つまり $A\mathbf{p}_i = \gamma_i\mathbf{p}_i$). $P := [\mathbf{p}_1, \dots, \mathbf{p}_n]$ と定めれば, 定理 17.3.6 より P は可逆である. このとき $P^{-1}AP$ が対角行列となることは 28.3 節 (I) で述べた通りであるが, これを計算によって確認してみよう. $P\mathbf{e}_i = \mathbf{p}_i$ および, この両辺に左から P^{-1} をかけて得られる $\mathbf{e}_i = P^{-1}\mathbf{p}_i$ に注意すれば

$$“P^{-1}AP \text{ の } i \text{ 列目}” = P^{-1}AP\mathbf{e}_i = P^{-1}A\mathbf{p}_i = P^{-1}\gamma_i\mathbf{p}_i = \gamma_i P^{-1}\mathbf{p}_i = \gamma_i\mathbf{e}_i.$$

ゆえに

$$P^{-1}AP = [\gamma_1\mathbf{e}_1, \dots, \gamma_n\mathbf{e}_n] = \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix}.$$

(3) \Rightarrow (2): いま,

$$P^{-1}AP = \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix} = B \quad (\text{したがって } A = PBP^{-1})$$

であると仮定し, $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ とおく. さらに, ベクトルの組 $\mathbf{q}_1, \dots, \mathbf{q}_n \in \mathbb{K}^n$ を次で定める:

$$[\mathbf{q}_1, \dots, \mathbf{q}_n] := [e_1, \dots, e_n]P = EP = P = [\mathbf{p}_1, \dots, \mathbf{p}_n].$$

このとき, $\mathbf{q}_1, \dots, \mathbf{q}_n$ は \mathbb{K}^n の基底であり, また \mathbf{q}_i が固有値 γ_i の固有ベクトルとなることは, 28.3 節 (II) で述べた通りである. あるいは, \mathbf{q}_i が固有ベクトルであることは次のような計算からも確認できる ($P\mathbf{e}_i = \mathbf{q}_i$ に注意する):

$$\begin{aligned} A\mathbf{q}_i &= PBP^{-1}\mathbf{q}_i = PBP^{-1}(P\mathbf{e}_i) = P\mathbf{B}\mathbf{e}_i \\ &= P \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix} \mathbf{e}_i = P(\gamma_i\mathbf{e}_i) = \gamma_i P\mathbf{e}_i = \gamma_i\mathbf{q}_i. \end{aligned}$$

(2) \Rightarrow (1): A の相異なるすべての固有値を $\lambda_1, \dots, \lambda_r$ とし, $\mathbf{q}_1, \dots, \mathbf{q}_n \in \mathbb{K}^n$ を A の固有ベクトルからなる基底とする. 各 $k = 1, \dots, r$ に対して, $\mathbf{q}_1, \dots, \mathbf{q}_n$ のうち $W(\lambda_k, A)$ に含まれるものの個数を m_k と置けば, $\dim W(\lambda_k, A) \geq m_k$ および $\sum_{k=1}^r m_k \geq n$ が成り立つ. ゆえに

$$\sum_{k=1}^r \dim W(\lambda_k, A) \geq \sum_{k=1}^r m_k \geq n.$$

上の式と系 29.1.2 を合わせて $\sum_{k=1}^r \dim W(\lambda_k, A) = n$ を得る. □

上の定理は, \mathbb{K}^n の各ベクトルが A の各固有空間の元によって分解できることと A が対角化可能であることが必要十分であることを述べている. 一般の行列については, 任意のベクトルが固有ベクトルに分解できるわけではない. これを補う方法として, 固有空間を少し広げた一般固有空間の元によって任意のベクトルを分解する理論がある. 詳しくは 33 および 34 章にて論じる.

系 29.2.2. n 次正方行列 A が n 個の相異なる固有値 $\lambda_1, \dots, \lambda_n$ を持つならば, A は対角化可能である.

Proof. 各固有値 λ_k の固有ベクトルは少なくとも一つは存在する. したがって $\dim W(\lambda, T_A) \geq 1$ であり, $\sum_{k=1}^n \dim W(\lambda_k, T_A) \geq n$. ゆえに前定理より A は対角化できる. \square

備考 29.2.3. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ において $\Phi_A(t) = 0$ の解が λ_1, λ_2 であるとしよう. $\lambda_1 \neq \lambda_2$ の場合: 系 29.2.2 より A は対角化可能である. 一方, $\lambda_1 = \lambda_2$ の場合は, A が対角化可能であるためには $\dim W(\lambda_1, A) = 2$, つまり $W(\lambda_1, A) = \mathbb{R}^2$ でなければならない. そのためには, A が元々対角行列でなければならないことが練習 28.5.4 から分かる.

ここで, 行列の対角化の方法についてまとめておこう.

まとめ (対角化の仕方)

n 次正方行列 A の対角化可能性の判定, および A の対角化は次の手順によって得られる:

- (1) 特性多項式 $\Phi_A(t) = |tE - A|$ を求め, 特性方程式 $\Phi_A(t) = 0$ を解く.
- (2) $\Phi(t)$ が \mathbb{R} 係数 1 次式の積に因数分解できない場合は, (3) 以下の計算を行うまでもなく \mathbb{R} 成分行列の範囲で A は対角化可能でないことが知られている (詳しくは命題 29.2.6). もちろん, \mathbb{C} 成分を許せば対角化できる場合もある ((3) 以下の計算を \mathbb{C} 成分について行って判定せよ).
- (3) $\Phi_A(t) = 0$ の相異なるすべての解 $\lambda_1, \dots, \lambda_k$ について, 各固有空間 $W(\lambda_k, A) = \text{Ker}(\lambda_k I - T_A)$ の次元を求める. $W(\lambda_k, A)$ は連立 1 次方程式 $(\lambda_k E - A)\mathbf{x} = \mathbf{0}$ の解空間のことであり, この基底は例題 18.3.3(2) と同様にして求められる.
- (4) 定理 29.2.1 により, $\sum_{k=1}^r \dim W(\lambda_k, A) = n$ ならば A は対角化可能であり, そうでなければ A は対角化可能でない.
- (5) A が対角化可能であるとき, (2) で求めた各固有空間 $W(\lambda_k, A)$ の基底を並べれば, これは n 個のベクトルの組 $\mathbf{p}_1, \dots, \mathbf{p}_n$ であり, 定理 29.2.1 における (2) \Rightarrow (3) の証明によれば $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ について $B = P^{-1}AP$ は対角行列となる.
- (6) このとき対角行列 B の (i, i) -成分は, \mathbf{p}_i に対応する固有値に等しい.

以上の手順から, A を対角化するための可逆行列 P の取り方には, 固有ベクトルからなる基底の選び方ぶんの自由度があることが分かる.

ところで, 方程式 $\Phi_A(t) = 0$ を解くには, n 次方程式の解の公式を用いるか, $\Phi_A(t)$ を上手く因数分解しなければならない. 一般に, 4 次以下の n 次方程式については根号と四則演算のみによる解の公式があり, 5 次以上についてはそのような公式は存在しないことが知られている. 3 次および 4 次方程式の解の公式の導出や, 5 次以上の方程式の解をいかに求めるかという問題は線形代数学の枠を超えるため, 本書では上の事実に言及するに留め, これ以上は論じないことにする.

例題 29.2.4. 次の行列が対角化可能かどうか答え, 対角化可能ならば対角化せよ.

$$(1) A = \begin{bmatrix} 5 & 6 & 0 \\ -1 & 0 & 0 \\ 1 & 2 & 2 \end{bmatrix}.$$

まず固有値を求め, 各固有値に関する固有空間の次元を求める.

$$\begin{aligned} \Phi_A(t) = |tE - A| &= \begin{vmatrix} t-5 & -6 & 0 \\ 1 & t & 0 \\ -1 & -2 & t-2 \end{vmatrix} = \begin{vmatrix} t-5 & -6 \\ 1 & t \end{vmatrix} (t-2) \quad (\text{定理 12.5.1 を用いた}) \\ &= ((t-5)t - (-6) \cdot 1)(t-2) = (t^2 - 5t + 6)(t-2) = (t-2)^2(t-3). \end{aligned}$$

特性方程式 $\Phi_A(\lambda) = 0$ を解いて, $\lambda = 3, 2$ が A の固有値である.

- $\lambda = 3$ の固有空間 $W(3, A)$: 方程式 $(3E - A)\mathbf{x} = \mathbf{0}$ を解くために $3E - A$ を簡約化すると

$$3E - A = \begin{bmatrix} -2 & -6 & 0 \\ 1 & 3 & 0 \\ -1 & -2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 0 \\ -2 & -6 & 0 \\ -1 & -2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに $(3E - A)\mathbf{x} = \mathbf{0}$ の解空間は, $W(3, A) = \left\langle \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \right\rangle$ であり, $\dim W(3, A) = 1$.

- $\lambda = 2$ の固有空間 $W(2, A)$:

$$2E - A = \begin{bmatrix} -3 & -6 & 0 \\ 1 & 2 & 0 \\ -1 & -2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに $(2E - A)\mathbf{x} = \mathbf{0}$ の解空間は, $W(2, A) = \left\langle \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle$ であり, $\dim W(2, A) = 2$.

以上より, $\dim W(3, A) + \dim W(2, A) = 3$ である. したがって A は対角化可能であり, 各固有空間の基底を並べた正方行列を

$$P = \begin{bmatrix} 3 & -2 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{とすれば} \quad P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

【注意】 これまでの議論により, P^{-1} を求めなくても $P^{-1}AP$ の成分表示が分かる.

$$(2) \quad A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 0 \\ 1 & 2 & 0 \end{bmatrix}.$$

$$\begin{aligned} \Phi_A(t) &= |tE - A| = \begin{vmatrix} t-1 & -3 & -2 \\ 0 & t+1 & 0 \\ -1 & -2 & t \end{vmatrix} = (-1) \begin{vmatrix} -3 & t-1 & -2 \\ t+1 & 0 & 0 \\ -2 & -1 & t \end{vmatrix} \quad (1 \text{ 列と } 2 \text{ 列の入れ替え}) \\ &= (-1)^2 \begin{vmatrix} t+1 & 0 & 0 \\ -1 & t-1 & -2 \\ -2 & -1 & t \end{vmatrix} = (t+1) \begin{vmatrix} t-1 & -2 \\ -1 & t \end{vmatrix} \\ &= (t+1) \left((t-1)t - (-2)(-1) \right) = (t+1)(t^2 - t - 2) = (t+1)^2(t-2). \end{aligned}$$

- $\lambda = 2$ の固有空間 $W(2, A)$:

$$2E - A = \begin{bmatrix} 1 & -3 & -2 \\ 0 & 3 & 0 \\ -1 & -2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & -2 \\ 0 & 3 & 0 \\ 0 & -5 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & -2 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに $(2E - A)\mathbf{x} = \mathbf{0}$ の解空間は, $W(2, A) = \left\langle \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\rangle$ であり, $\dim W(2, A) = 1$.

- $\lambda = -1$ の固有空間 $W(-1, A)$:

$$-E - A = \begin{bmatrix} -2 & -3 & -2 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに方程式 $(-E - A)\mathbf{x} = \mathbf{0}$ の解空間は, $W(-1, A) = \left\langle \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\rangle$ であり, $\dim W(-1, A) = 1$.

以上より, $\dim W(2, A) + \dim W(-1, A) \neq 3$ である. したがって A は対角化可能でない.

【補足】 3次正方行列 A の特性多項式が $\Phi_A(t) = (t - \lambda)(t - \gamma)^2$ (ただし $\lambda \neq \gamma$) である場合, 系 29.1.6 より $1 \leq \dim W(\lambda, A) \leq 1$ および $1 \leq \dim W(\gamma, A) \leq 2$ であることが分かる. つまり, $\dim W(\lambda, A) = 1$ となることは計算せずとも分かり, A が対角化可能かどうかは, $\dim W(\gamma, A)$ の値のみによって決まる. したがって, 先に $\dim W(\gamma, A)$ の次元から確認するほうが筋がよい.

ここで, 定期試験等で対角化を計算する際のポイントをまとめておく. 計算の精度に自信がない者は参考にとするとよいだろう.

A を n 次正方行列とする.

(1) 次のようにして, 固有値の一つがすぐに分かる場合がある. その場合は, $\Phi_A(t)$ の因数分解が楽になる.

(i) $\lambda E - A$ の各列 (あるいは各行) が 1 次従属になるような λ が山勘で見つかることがある. この λ は A の固有値である. つまり $\Phi_A(t)$ は $(t - \lambda)$ で割り切れる.

(ii) 定期試験では難度を下げるために, $A = [a_{ij}]$ の対角成分のいずれかが固有値になるように細工して出題されることがある. とくに $n = 3$ の場合は, たすきがけを用いて展開すると $\Phi_A(t) = (t - a_{11})(t - a_{22})(t - a_{33}) - (t$ の 1 次式) となるから, 対角成分の一つが固有値になる場合は第一項をすべて展開することなく $\Phi_A(t)$ の因数分解が得られる.

(2) 以下で「連立 1 次方程式 $(\lambda E - A)\mathbf{x} = \mathbf{0}$ の計算を間違えている」と述べる件については, $\lambda E - A$ の簡約化で変形ミスをしている場合が多い.

(3) 連立 1 次方程式 $(\lambda E - A)\mathbf{x} = \mathbf{0}$ の解が $\mathbf{x} = \mathbf{0}$ のみになった場合は, どこかで計算を間違えており, 次の二つの可能性が考えられる.

(i) 連立 1 次方程式の計算を間違えている.

(ii) 連立 1 次方程式の計算にミスがない場合は, λ は A の固有値ではない. したがって $\Phi_A(t)$ の計算を間違えている.

(4) $\Phi_A(t) = \prod_{k=1}^r (t - \lambda_k)^{n_k}$ (ここで各 λ_k は相異なる) とする. このとき連立 1 次方程式 $(\lambda_k E - A)\mathbf{x} = \mathbf{0}$ の解空間の次元は $n_k + 1$ 以上にはならない (系 29.1.6). そうなってしまった場合は, 連立 1 次方程式の計算を間違えている.

(5) $k_1 \neq k_2$ とすれば, 二つの連立 1 次方程式 $(\lambda_{k_1} E - A)\mathbf{x} = \mathbf{0}$ と $(\lambda_{k_2} E - A)\mathbf{x} = \mathbf{0}$ の共通解は $\mathbf{x} = \mathbf{0}$ のみである (系 29.1.4). $\mathbf{0}$ 以外に共通解が見つかるようであれば, 連立 1 次方程式の計算を間違えている.

(6) A の対角化 B が求まった際に, P^{-1} の成分表示を求めずに式 $B = P^{-1}AP$ の検算を行う方法としては, これと同値な式 $PB = AP$ が成り立つかどうかを確認すればよい. 行列 $PB = AP$ には, 各固有ベクトルを固有値倍した列が並ぶはずである. この検算で $PB \neq AP$ となった場合は, 一致しなかった列に対応する固有値 λ において, 連立 1 次方程式 $(\lambda E - A)\mathbf{x} = \mathbf{0}$ の計算を間違えている.

【注意】ただし, この検算は P が可逆ゆえに有効なのであり, 対角化の手順を間違えて P が可逆でなくなってしまう場合にこの検算を行っても無意味である. 実際, 可逆でない P について $PB = AP$ が成り立つことはいくらかもある. 例えば固有値 λ に関する固有ベクトル \mathbf{p} を一つとり, $P = [\mathbf{p}, \mathbf{p}, \dots, \mathbf{p}]$, $B = \lambda E$ とすれば, $PB = AP = [\lambda\mathbf{p}, \lambda\mathbf{p}, \dots, \lambda\mathbf{p}]$ が成り立つ.

さて, 任意の複素係数多項式は, 複素係数 1 次式の積に必ず因数分解できることが知られている (系 29.4.6). したがって複素正方行列 A について, その固有多項式は $\Phi_A(t) = \prod_{i=1}^n (t - \lambda_i)$ と因数分解される. このとき $\lambda_1, \dots, \lambda_n$ のことを A の重複を込めたすべての固有値と呼ぶ.

命題 29.2.5. n 次複素正方行列 A の重複を込めたすべての固有値を $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ とする. A が対角化可能ならば, $|A| = \lambda_1 \lambda_2 \cdots \lambda_n$.

Proof. $D = P^{-1}AP$ を A の対角化とすれば, $|D| = \lambda_1 \lambda_2 \cdots \lambda_n$ である. したがって, $|A| = |PDP^{-1}| = |P| \cdot |D| \cdot |P|^{-1} = |D| = \lambda_1 \lambda_2 \cdots \lambda_n$.

□

実は、上の主張は対角化できない任意の正方行列についても成り立つ (備考 34.1.5(3)(i)). 実正方行列の特性方程式が虚数解をもつ場合についても論じておこう.

命題 29.2.6. n 次正方行列 A について、特性多項式 $\Phi_A(t)$ が実係数 1 次式の積に因数分解できないならば、実成分行列の範囲において A は対角化可能でない.

Proof. 特性方程式の相異なるすべての実数解を $\lambda_1, \dots, \lambda_r$ とすると、 $\Phi_A(t) = (t - \lambda_1)^{n_1} \cdots (t - \lambda_r)^{n_r} \Psi(t)$ と分解される (ただし、各 $k = 1, \dots, r$ について $\Psi(\lambda_k) \neq 0$)². このとき仮定より $\deg \Psi(t) \geq 1$ である³.

$$n = \deg \Phi_A(t) = n_1 + n_2 + \cdots + n_k + \deg \Psi(t)$$

であることから (命題 15.2.4), とくに $\sum_{k=1}^r n_k < n$ である. したがって系 29.1.6 より, $\sum_{k=1}^r \dim W(\lambda_k, T_A) \leq \sum_{k=1}^r n_k < n$. すなわち, A は対角化可能でない. \square

例 28.2.5 に現れた行列について改めて考えてみよう.

練習 29.2.7. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ が対角化可能かどうか答えよ.

解答例: $\Phi_A(t) = t^2 + 1$ より方程式 $\Phi_A(t) = 0$ は実数解を持たない. したがって, A の固有ベクトルは \mathbb{R}^2 上には存在せず, 実数を成分とする行列 P を用いて $P^{-1}AP$ を対角行列にすることはできない. 一方, 成分に複素数を認める場合は $\Phi_A(t) = 0$ の解 $t = \pm i$ が A の固有値であり, その固有空間はそれぞれ $W(i, A) = \left\langle \begin{bmatrix} i \\ 1 \end{bmatrix} \right\rangle$, $W(-i, A) = \left\langle \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\rangle$ である⁴. ゆえに $P = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$ とすれば

$$P^{-1}AP = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

29.3 一般の線形変換の場合

これまでの内容を一般の線形空間の表現行列に関する主張に書き換えると次のようになる. この定理の証明の一部も 28.3 節で論じたことと重複するが, ここで改めて述べておく.

定理 29.3.1. n 次元線形空間 U 上の線形変換 $f: U \rightarrow U$ について次は同値である.

- (i) f の相異なるすべての固有値を $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ とすれば, $\sum_{k=1}^r \dim W(\lambda_k, f) = n$.
- (ii) f の固有ベクトルからなる U の基底が存在する.
- (iii) f の任意の表現行列は対角化可能である.
- (iv) f は対角化可能な表現行列を持つ.
- (v) f は対角行列となるような表現行列を持つ.

Proof. (i) \Leftrightarrow (ii): U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列を A とすれば, 上の条件 (i), (ii) はそれぞれ定理 29.2.1 における条件 (1), (2) と同値になる. 実際, (i) \Leftrightarrow (1) および (ii) \Leftrightarrow (2) は命題 28.6.1 による. 条件 (1), (2) の同値性より, 条件 (i), (ii) も同値である.

(i) \Rightarrow (iii): U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を一つ取り, この基底に関する f の表現行列を A とする. いま条件 (i) を仮定しているゆえ, 命題 28.6.1(3) より A は定理 29.2.1 における条件 (1) を満たし, したがって対角化できる.

²詳しくは系 29.4.4 を見よ.

³もし $\deg \Psi(t) = 0$, すなわち $\Psi(t)$ が定数ならば, $\Phi_A(t)$ の最高次の係数が 1 であることから $\Psi(t) = 1$ であり, これは仮定に反する.

⁴これらの計算は練習 28.5.4(1) から直ちに分かる.

(iii) \Rightarrow (i): U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を一つ取り, この基底に関する f の表現行列を A とする. 仮定より (iii) は対角化可能であり, したがって A は定理 29.2.1 における条件 (1) を満たす. 命題 28.6.1(3) より, これは f が (i) を満たすことに他ならない.

(iii) \Rightarrow (iv): 明らか.

(iv) \Rightarrow (v): 28.3 節 (II) で述べた通り.

(v) \Rightarrow (iii): U の基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ に関する f の表現行列が対角行列 B であるとする. U の任意の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列 A が対角化できることを示そう. 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P とすれば, 系 26.3.2 より $B = P^{-1}AP$ であり, したがって A は対角化可能である. \square

系 29.2.2 に対応する主張は次の通りである:

系 29.3.2. n 次元線形空間 U 上の線形変換 $f: U \rightarrow U$ について, 特性方程式 $\Phi_f(t) = 0$ が n 個の相異なる解 $\lambda_1, \dots, \lambda_n$ をもつならば, f の固有ベクトルからなる U の基底が存在する.

Proof. あらかじめ U の基底を一つ与えておき, その基底による f の表現行列を A とする. このとき $\Phi_A(t) = \Phi_f(t)$ および系 29.2.2 より A は対角化可能であり, したがって定理 29.3.1 より f の固有ベクトルからなる基底が存在する. \square

練習 29.3.3. 例 28.6.2 で定めた線形変換 $D: \mathbb{R}[x]_3 \rightarrow \mathbb{R}[x]_3$ ($D(p(x)) := \frac{d}{dx}p(x)$) は, 対角行列となるような表現行列を持つかどうか答えよ.

解答例: D の固有値は 0 のみであり, その固有空間は 1 次元である. $\dim W(0, D) \neq \dim \mathbb{R}[x]_3 = 4$ より, 対角行列となるような D の表現行列は存在しない. \square

例題 26.1.4(2) の備考で述べたように, 上の D (の表現行列) は冪零である. 実は, 一般論として次が成り立つ:

命題 29.3.4 (発展). 対角化可能な冪零行列は O に限る.

Proof. n 次冪零行列 A が対角化できるならば, A の固有ベクトル $\mathbf{p}_1, \dots, \mathbf{p}_n$ からなる \mathbb{R}^n の基底がある. 例 28.2.3 より A の固有値は 0 のみゆえ, A の対角化は $P^{-1}AP = O$ である. この両辺に左から P を, 右から P^{-1} をかけて $A = O$. \square

29.4 多項式と方程式の解に関する基本的性質 (付録)

\mathbb{K} 係数の多項式, および n 次方程式の解に関するいくつかの性質について簡単に解説しておく. 多項式に関する約数のことを整数の場合と区別して, ここでは因子と呼ぶことにしよう:

定義 29.4.1. \mathbb{K} 係数多項式 Φ, ξ について, Φ が ξ によって割り切れるとき, すなわち, ある別の \mathbb{K} 係数多項式 Ψ を用いて, $\Phi(x) = \xi(x)\Psi(x)$ と書けるとき, ξ を Φ の因子 (**factor**) とよぶ. また, Φ は ξ を因子に持つという.

命題 29.4.2. 多項式 $\Phi(x)$ が 1 次式 $(x - \lambda)$ を因子に持つならば, $(x - \lambda)$ を因子に持たない多項式 $\Psi(x)$ を用いて $\Phi(x) = (x - \lambda)^m \Psi(x)$ と書ける.

Proof. Φ の次数を n とすれば, 仮定より $n - 1$ 次多項式 Ψ_1 を用いて $\Phi(x) = (x - \lambda)\Psi_1(x)$ と書ける. ここで $\Psi_1(x)$ が $(x - \lambda)$ を因子に持たなければ主張を得る. そうでない場合は, $n - 2$ 次多項式 Ψ_2 を用いて $\Psi_1(x) = (x - \lambda)\Psi_2(x)$ と書ける. ここで $\Psi_2(x)$ が $(x - \lambda)$ を因子に持たなければ $\Phi(x) = (x - \lambda)^2\Psi_2(x)$ として主張を得る. そうでない場合は $n - 2$ 次多項式 Ψ_3 を用いて $\Psi_2(x) = (x - \lambda)\Psi_3(x)$ と書ける. この手順を順次繰り返して得られる多項式 $\Psi_1, \Psi_2, \Psi_3, \dots$ は次数が一つずつ減っていく. k 回目の過程 (ただし $k < n$) において $(x - \lambda)$ を因子に持たない Ψ_k が現れれば, $\Phi(x) = (x - \lambda)^k \Psi_k(x)$ となり主張を得る. あるいは n 回目まで繰り返すことができた場合は Ψ_n が次数 0 の定数 C となり, $\Phi(x) = C(x - \lambda)^n$ を得る. \square

定理 29.4.3 (因数定理). \mathbb{K} 係数多項式 $\Phi(x)$ について, および $\lambda \in \mathbb{K}$ について,

$$\Phi(\lambda) = 0 \iff \Phi(x) \text{ は } (x - \lambda) \text{ を因子に持つ.}$$

Proof. (\Leftarrow) は明らかゆえ (\Rightarrow) のみ示す. $\Phi(x)$ を $x - \lambda$ で割った商を $\Psi(x)$, 余りを $R(x)$ とする. いま $x - \lambda$ が 1 次式であることから, 商は $n - 1$ 次式, 余りは 0 次式, すなわち定数 $r \in \mathbb{K}$ である. つまり, $\Phi(x) = (x - \lambda)\Psi(x) + r$ と書ける. この両辺に $x = \lambda$ を代入し $r = 0$ を得る. 以上より $\Phi(x) = (x - \lambda)\Psi(x)$. \square

系 29.4.4. $\Phi(x)$ を多項式とする. 相異なる数 $\lambda_1, \dots, \lambda_r$ がそれぞれ方程式 $\Phi(x) = 0$ の解ならば, $\lambda_1, \dots, \lambda_r$ のいずれも因子に持たない多項式 $\Psi(x)$ を用いて $\Phi(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_r)^{n_r} \Psi(x)$ と書ける.

Proof. 次を満たす多項式 $\Psi_k(x)$ ($k = 1, \dots, r$) の存在を k に関する帰納法により示そう:

- (1) $\Phi(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k} \Psi_k(x)$,
- (2) $\Psi_k(x)$ は $\lambda_1, \dots, \lambda_k$ のいずれも因子に持たない.

まず $k = 1$ の場合を示す. 因数定理より $\Phi(x)$ は $(x - \lambda_1)$ を因子に持つ. ゆえに命題 29.4.2 より, λ_1 を因子に持たない多項式 $\Psi_1(x)$ を用いて $\Phi(x) = (x - \lambda_1)^{n_1} \Psi_1(x)$ と書ける.

次に $k < r$ を満たす $k \in \mathbb{N}$ について多項式 $\Psi_k(x)$ が上の (1) と (2) を満たすと仮定し, $\Psi_{k+1}(x)$ の存在を示そう. (1) の両辺に λ_{k+1} を代入すると $0 = (\lambda_{k+1} - \lambda_1)^{n_1} \cdots (\lambda_{k+1} - \lambda_k)^{n_k} \Psi_k(\lambda_{k+1})$ であり, 仮定より各 $\lambda_{k+1} - \lambda_i$ ($i = 1, \dots, k$) は 0 でない. ゆえに $\Psi_k(\lambda_{k+1}) = 0$, すなわち $\Psi_k(x)$ は $(x - \lambda_{k+1})$ を因子に持つ. したがって命題 29.4.2 より, λ_{k+1} を因子に持たない多項式 $\Psi_{k+1}(x)$ を用いて $\Psi_k(x) = (x - \lambda_{k+1})^{n_{k+1}} \Psi_{k+1}(x)$ と書ける. このとき, $\Psi_k(x)$ は $\lambda_1, \dots, \lambda_k$ を因子にもたないことから, $\Psi_{k+1}(x)$ も $\lambda_1, \dots, \lambda_k$ を因子にもたない. まとめれば, $\Psi_{k+1}(x)$ は $\lambda_1, \dots, \lambda_{k+1}$ を因子にもたない. また, $\Phi(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k} \Psi_k(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k} (x - \lambda_{k+1})^{n_{k+1}} \Psi_{k+1}(x)$ である. 以上により $\Psi_{k+1}(x)$ が得られた.

こうして得られた $\Psi_r(x)$ が求めるべき多項式 $\Psi(x)$ である. \square

n 次多項式 $\Phi(x)$ に関する方程式 $\Phi(x) = 0$ を n 次方程式という. 次は n 次方程式論において基本となる定理である. この定理を示すには \mathbb{R}^2 の位相に関する知識, あるいは微分可能な複素数値関数についての知識が必要となるため本書では証明を割愛する⁵.

定理 29.4.5 (代数学の基本定理). 任意の \mathbb{C} 係数 n 次多項式 $\Phi(x)$ について (ただし $n \in \mathbb{N}$), 方程式 $\Phi(x) = 0$ は複素数の解を必ず持つ.

代数学の基本定理を認めると, 任意の複素係数多項式が 1 次式の積に因数分解されることが分かる:

系 29.4.6. 任意の \mathbb{C} 係数 n 次多項式 $\Phi(x)$ は, 定数 $C, \lambda_1, \dots, \lambda_r$ を用いて $\Phi(x) = C(x - \lambda_1)^{n_1} \cdots (x - \lambda_r)^{n_r}$ と因数分解できる.

【注意】 上の定数 C は, $\Phi(x)$ の最高次の係数に等しい.

Proof. n 次方程式 $\Phi(x) = 0$ の相異なるすべての複素解を $\lambda_1, \dots, \lambda_r$ とおく (代数学の基本定理より $r \geq 1$ である). このとき系 29.4.4 より, $\lambda_1, \dots, \lambda_r$ のいずれも因子に持たない多項式 $\Psi(x)$ を用いて

$$\Phi(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_r)^{n_r} \Psi(x)$$

と書ける. このとき, 方程式 $\Psi(x) = 0$ は解を持たない. 何故なら, 仮に $\Psi(\lambda) = 0$ を満たす $\lambda \in \mathbb{C}$ があるとすれば, 上式に $x = \lambda$ を代入することで $\Phi(\lambda) = 0$ を得る. すなわち λ は方程式 $\Phi(x) = 0$ の解ゆえ, $\lambda_1, \dots, \lambda_r$ のいずれかに等しい. しかしながらこれは $\Psi(x)$ が $\lambda_1, \dots, \lambda_r$ を因子に持たないことに矛盾する. 以上より, 方程式 $\Psi(x) = 0$ は解を持たず, ゆえに $\Psi(x)$ は 0 次式, すなわち定数 $C \neq 0$ でなければならない. \square

⁵例えば, 巻末の文献 [6] に代数学の基本定理の証明がある.

複素数を成分とする行列について、系 29.2.2 を次のように言い換えてもよい。

系 29.4.7. n 次正方行列 A における特性方程式 $\Phi_A(t) = 0$ が \mathbb{C} において重解を持たないならば、 A は複素数の範囲で対角化可能である。

Proof. 系 29.4.6 より $\Phi_A(t) = (t - \lambda_1) \cdots (t - \lambda_n)$ と因数分解できる。仮定より、 $\lambda_1, \dots, \lambda_n$ は相異なる複素数ゆえ系 29.2.2 より A は対角化できる。□

実数係数の n 次多項式における複素解とその共役についても述べておこう。複素数 $z = a + bi$ (ただし $a, b \in \mathbb{R}$) に対して、 $\bar{z} := a - bi$ を z の共役 (conjugate) という。 $z + \bar{z}$ および $z\bar{z}$ は共に実数である。また、 $\overline{z + w} = \bar{z} + \bar{w}$ および $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$, $\overline{z^n} = \bar{z}^n$ が成り立つ。実数 r について $\bar{r} = r$ である。

命題 29.4.8. \mathbb{R} 係数の n 次方程式 $\Phi(x) = 0$ が複素解 $z \in \mathbb{C}$ を持つならば、その共役 \bar{z} も $\Phi(x) = 0$ の解である。

Proof. $\Phi(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ とおく。 $\Phi(z) = 0$ と仮定し、 $\Phi(\bar{z}) = 0$ を導こう。

$$\begin{aligned} \Phi(\bar{z}) &= \bar{z}^n + a_{n-1}\bar{z}^{n-1} + \cdots + a_1\bar{z} + a_0 = \overline{z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0} \\ &= \overline{z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0} = \overline{\Phi(z)} = \bar{0} = 0. \end{aligned}$$

□

次の命題は、多項式環 $\mathbb{R}[x]$ および $\mathbb{C}[x]$ が整域をなすことを意味する。この命題は、多項式の連続性を用いて証明することもできる。しかしここでは代数的な証明を与えよう。

補題 29.4.9. \mathbb{K} 係数多項式 P, Q, R について $P(x)Q(x) = P(x)R(x)$ が成り立つならば、 $Q = R$ 。

【補足】 もちろん定数 $a \in \mathbb{K}$ において $P(a) \neq 0$ が成り立つならば、 $P(a)Q(a) = P(a)R(a)$ の両辺に $P(a)$ の逆数をかけることで $Q(a) = R(a)$ を得る。しかし、この方法では $P(x) = 0$ を満たす $x \in \mathbb{K}$ についても $Q(x) = R(x)$ となるかどうかは分からない。この補題は、そのような x についても $Q(x) = R(x)$ であることを述べている。

Proof. $n = \max\{\deg P, \deg Q\}$ とおく。系 29.4.6 により、 $P(x) = 0$ を満たす $x \in \mathbb{K}$ は有限個しかない。そこで $P(x_i) \neq 0$ ($i = 0, \dots, n$) を満たす相異なる $n + 1$ 個の定数 x_0, \dots, x_n を決めておく。各 $i = 0, \dots, n$ について、 $P(x_i)Q(x_i) = P(x_i)R(x_i)$ ゆえ $Q(x_i) = R(x_i)$ である。よって命題 15.2.6 より $P = Q$ 。□

次は 34 章において、主定理の証明で用いる。この補題は証明の中で代数学の基本定理を用いるものの、 $\mathbb{K} = \mathbb{R}$ の場合についても成り立つ。

補題 29.4.10 (発展). 相異なる $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ および、因数分解された多項式 $\Phi(x) = (x - \lambda_1)^{n_1}(x - \lambda_2)^{n_2} \cdots (x - \lambda_r)^{n_r}$ が与えられているとする。 Φ が次の条件を満たす \mathbb{K} 係数多項式 ξ, η をもちいて $\Phi(x) = \xi(x)\eta(x)$ と表されるならば、 $\xi(x) = (x - \lambda_1)^{n_1}$ である。

- (i) 各 $i = 2, 3, \dots, r$ について、 $\xi(x)$ は $(x - \lambda_i)$ を因子に持たない。
- (ii) $\eta(x)$ は $(x - \lambda_1)$ を因子に持たない。
- (iii) $\xi(x)$ の最高次の係数は 1 である。

Proof. $0 = \Phi(\lambda_1) = \xi(\lambda_1)\eta(\lambda_1)$ であり、因数定理と条件 (ii) より $\eta(\lambda_1) \neq 0$ であるから $\xi(\lambda_1) = 0$ 。ふたたび因数定理により $\xi(x)$ は $(x - \lambda_1)$ を因子に持ち、ゆえに命題 29.4.2 により $(x - \lambda_1)$ を因子に持たない多項式 $\zeta(x)$ を用いて $\xi(x) = (x - \lambda_1)^m \zeta(x)$ と書ける。これから $m = n_1$ および $\zeta(x) = 1$ を示そう。

いま、多項式 $\Phi(x)$ は次の二通りに表せている：

$$(x - \lambda_1)^{n_1}(x - \lambda_2)^{n_2} \cdots (x - \lambda_r)^{n_r} = (x - \lambda_1)^m \zeta(x) \eta(x). \quad (29.4.1)$$

ここで仮に $m < n_1$ として矛盾を導こう. 上式の両辺を $(x - \lambda_1)^m$ で割れば

$$(x - \lambda_1)^{n_1 - m} (x - \lambda_2)^{n_2} \cdots (x - \lambda_r)^{n_r} = \zeta(x)\eta(x).$$

この左辺に $x = \lambda_1$ を代入すれば 0 になる. 一方で, $\zeta(x)$ と $\eta(x)$ は $(x - \lambda_1)$ を因子に持たないことから上式の右辺に λ_1 を代入した値は 0 にはならず, これは不合理である. 次に $m > n_1$ を仮定し, 式 (29.4.1) の両辺を $(x - \lambda_1)^{n_1}$ で割れば

$$(x - \lambda_2)^{n_2} \cdots (x - \lambda_r)^{n_r} = (x - \lambda_1)^{m - n_1} \zeta(x)\eta(x).$$

上式の両辺に $x = \lambda_1$ を代入すれば, 左辺は 0 ではなく右辺は 0 となり, やはりこれも不合理である. ゆえに $m = n_1$ でなければならない.

次に, ζ の次数が 1 以上ならば矛盾が生じることを示そう. $\xi(x) = (x - \lambda_1)^{n_1} \zeta(x)$ であり, ζ の定め方および (i) より $\lambda_1, \dots, \lambda_r$ はいずれも方程式 $\zeta(x) = 0$ の解ではない. ζ の次数が 1 以上ならば, 代数学の基本定理により方程式 $\zeta(x) = 0$ は複素解 λ を持つ. そこで $x = \lambda$ を式 (29.4.1) の両辺に代入すれば, 左辺は 0 にならず右辺は 0 になり, 不合理を得る. ゆえに ζ の次数は 0, つまり $\zeta(x)$ は定数である. ξ の最高次の係数が 1 であること, および $\xi(x) = (x - \lambda_1)^{n_1} \zeta(x)$ から, $\zeta(x) = 1$ でなければならない. 以上より, $\xi(x) = (x - \lambda_1)^{n_1}$. □

第30章 ケーリー-ハミルトンの定理

ケーリー-ハミルトンの定理とは、次の公式のことを指す：

定理 30.0.11 (ケーリー-ハミルトン).

- (i) n 次正方行列 A について、 $\Phi_A(A) = O$.
- (ii) 有限次元線形空間上の線形変換 $f : U \rightarrow U$ について、 $\Phi_f(f) = \mathbf{0}_U$.

この定理の主張はあまりにも唐突である。果たして、どのように考えればこの主張を予想するに至るのだろうか。このような疑問を初学者がもつことは当然のことと思う。しかし、この問いに答えるのは容易ではない。そこで本書では、この公式の証明を三通りの方法で紹介し、この疑問への回答の代わりとする。

30.1 まえおき

ケーリー-ハミルトンの定理を次のように誤解して解釈してはならない。

ケーリー-ハミルトンの定理の間違った理解

一見すると次の式変形によってケーリー-ハミルトンの定理が示されたかに思えるが、これは正しくない：

$$\Phi_A(t) = |tE - A| \text{ に } t = A \text{ を代入して、 } \Phi_A(A) = |AE - A| = |O| = 0.$$

そもそも上の式変形における左辺 $\Phi_A(A)$ は n 次正方行列であり、右辺の 0 は数であるから、これでは行列と数という異なる概念を等式で結んでしまったことになる。勘違いの原因は、 \mathbb{K} の元と $M_n(\mathbb{K})$ の元を混同したことにある。あるいは、多項式 $\Phi_A(t)$ に正方行列 X を代入したものが $\Phi_A(X)$ であると誤認してしまったと言ってもよい。 $\Phi_A(X)$ の正確な定義は、定義 21.5.1 で述べたように、 $\Phi_A(t)$ における変数の冪 t^n を X^n に置き換え、定数項 a を aE に置き換えた式で表される行列のことであった。

さて、ケーリー-ハミルトンの定理の証明には、大きく分けて次の三通りの方法が知られている。

- (I) 予備知識をあまり必要とせず、ごり押し、あるいは巧みな計算技法を駆使して証明する方法。例えば $n = 2$ の場合は練習 21.5.3 のようにして示せる。このような方法は定理の正しさは理解できるものの、どういった背景で定理が成立するのかを理解するのは難しい。
- (II) 相似な行列への変形を通して (例えば対角化や上三角化、ジョルダン標準形など)、比較的計算しやすい行列の問題に帰着させたいうで、行列の成分計算によって証明する方法。これを通して計算を簡単にする手法を学ぶことが出来る。しかしながら (I) と同様に、計算による証明から背景を理解するのは容易ではない場合がある。
- (III) 一般固有ベクトルへの分解まで学ぶと、この定理の主張はほとんど自明に見えるようになり、もはや証明は不要になる。すなわち、この定理が成り立つことを自然と想像できるようになる。

議論を複雑にすることなく上記の (I) から (III) の違いを味わってもらうために、本章では、対角化可能な行列に対するケーリー-ハミルトンの定理の証明を、これら三つの立場から検討する。また、章末で (I) の立場から一般の場合の証明を与える。(III) の立場による証明は 34.2 節を見よ。

ケーリー-ハミルトンの定理に限らず、線形代数学におけるいくつかの定理は、上の三つの立場からそれぞれに違った証明ができる場合が多い。例えば前章で紹介した系 29.1.6 の証明は上の (I) の立場に相当する。(III) の立場による系 29.1.6 の証明は命題 34.4.5 の直後に述べた。そこで余力ある読者は、(II) の立場による系 29.1.6 の証明を考えてみるとよい。このように、様々な証明のスタイルに習熟できれば、線形代数学への理解がより深まるであろう。

30.2 (i) と (ii) の同値性

定理の証明について考察する前に、二つの主張 (i) と (ii) が本質的に同じものであり、したがって片方のみを示せば十分であることを確認しておく。

命題 30.2.1. 次は同値である。

(i) n 次正方行列 A について、 $\Phi_A(A) = O$ 。

(ii) 有限次元線形空間上の線形変換 $f : U \rightarrow U$ について、 $\Phi_f(f) = \mathbf{0}_U$ 。

Proof. (i) \Rightarrow (ii): U のある基底に関する f の表現行列を A とすれば、 $\Phi_f(A) = \Phi_A(A) = O$ 。系 26.2.5 により線形変換 $\Phi_f(f)$ の表現行列は $\Phi_f(A) = O$ であり、したがって $\Phi_f(f)$ は任意のベクトルを $\mathbf{0}_U$ に写す定値写像である。

(ii) \Rightarrow (i): 行列 A から定まる自然な線形写像 $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ に対して (ii) を適用すれば $\Phi_A(T_A) = \Phi_{T_A}(T_A) = \mathbf{0}_{\mathbb{R}^n}$ 。つまり、線形変換 $\Phi_A(T_A)$ はすべてのベクトルを零ベクトルにうつす定置写像である。標準基底に関する T_A の表現行列は A だから (例 26.1.2)、標準基底に関する $\Phi_A(T_A)$ の表現行列は $\Phi_A(A)$ であり ($U = \mathbb{R}^n$, $f = T_A$ について系 26.2.5 を適用せよ)、これは $T_{\Phi_A(A)} = \Phi_A(T_A)$ を意味する (系 26.2.2(2))。行列 $\Phi_A(A)$ の第 j 列は $\Phi_A(A)e_j = T_{\Phi_A(A)}(e_j) = \Phi_A(T_A)(e_j) = \mathbf{0}_{\mathbb{R}^n}$ 。ゆえに、 $\Phi_A(A)$ は零行列である。 \square

30.3 対角化可能な行列の場合

命題 30.3.1. n 次正方行列 A が対角化可能であるとき、 $\Phi_A(A) = O$ 。

(I) の立場における証明. A の対角化可能性から、 \mathbb{K}^n の各元は A の固有ベクトルの線形結合で表せる。したがって、各固有ベクトル \mathbf{v} について $\Phi(A)\mathbf{v} = \mathbf{0}$ を示せば十分である (練習 18.1.4)。 \mathbf{v} を固有値 λ に関する固有ベクトルとし、 $\Phi_A(t) = \sum_{k=0}^n a_k t^k$ (ただし $a_n = 1$) とする。 $A^k \mathbf{v} = \lambda^k \mathbf{v}$ および $\Phi_A(\lambda) = 0$ に注意すれば

$$\Phi_A(A)\mathbf{v} = \left(\sum_{k=0}^n a_k A^k \right) \mathbf{v} = \sum_{k=0}^n (a_k A^k \mathbf{v}) = \sum_{k=0}^n (a_k \lambda^k \mathbf{v}) = \left(\sum_{k=0}^n a_k \lambda^k \right) \mathbf{v} = \Phi_A(\lambda)\mathbf{v} = 0\mathbf{v} = \mathbf{0}.$$

\square

(II) の立場における証明. A の対角化を $P^{-1}AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ とする。このとき、系 28.4.5 より

$\Phi_A(t) = \Phi_{P^{-1}AP}(t) = \prod_{i=1}^n (t - \lambda_i)$ である. $P^{-1}\Phi_A(A)P = O$ が示されれば, $\Phi_A(A) = O$ を得る.

$$\begin{aligned}
 P^{-1}\Phi_A(A)P &= P^{-1} \left(\prod_{i=1}^n (A - \lambda_i E) \right) P \\
 &= P^{-1}(A - \lambda_1 E)P \cdot P^{-1}(A - \lambda_2 E)P \cdots P^{-1}(A - \lambda_n E)P \\
 &= (P^{-1}AP - P^{-1}(\lambda_1 E)P) \cdot (P^{-1}AP - P^{-1}(\lambda_2 E)P) \cdots (P^{-1}AP - P^{-1}(\lambda_n E)P) \\
 &= (P^{-1}AP - \lambda_1 E) \cdot (P^{-1}AP - \lambda_2 E) \cdots (P^{-1}AP - \lambda_n E) \\
 &= \begin{bmatrix} 0 & & & \\ & * & & \\ & & \ddots & \\ & & & * \end{bmatrix} \begin{bmatrix} * & & & \\ & 0 & & \\ & & \ddots & \\ & & & * \end{bmatrix} \cdots \begin{bmatrix} * & & & \\ & * & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \ddots & \\ & & & * \end{bmatrix} \cdots \begin{bmatrix} * & & & \\ & * & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} = \cdots = O.
 \end{aligned}$$

□

$P^{-1}AP$ が上三角行列となる時, これを A の上三角化という. 上三角化を持つ行列 A に対するケーリー-ハミルトンの定理は, 上の (II) の立場における証明と同様にして示すことができる (各自確かめよ). なお, 任意の行列は, 複素数の範囲で上三角化を持つことが知られており (備考 34.1.5 あるいは定理 39.6.3), これは一般の場合のケーリー-ハミルトンの定理の証明を意味する.

(III) の立場における命題 30.3.1 の証明は, 次に紹介する補題 29.1.3 の別証と関係している. この別証からは, 固有ベクトルが写像される様子がはっきりと見えるであろう. また, これと写像の合成を上手く適用することでベクトルを消去していることが分かる.

補題 29.1.3 (再掲). f の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とする. $\mathbf{u}_k \in W(\lambda_k, f)$ ($k = 1, \dots, r$) かつ $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ ならば, $\mathbf{u}_1 = \dots = \mathbf{u}_r = \mathbf{0}$.

Proof. $(f - \lambda_k I)(\mathbf{u}_k) = f(\mathbf{u}_k) - \lambda_k \mathbf{u}_k = \lambda_k \mathbf{u}_k - \lambda_k \mathbf{u}_k = \mathbf{0}$ に注意する. いまから, $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ の両辺に線形変換 $(f - \lambda_k I)$ ($k = 2, \dots, r$) を順次ほどこすことで $\mathbf{u}_2, \dots, \mathbf{u}_r$ の項が消えて \mathbf{u}_1 の項のみが残ることを見よう. $r - 1$ 次多項式 $\Psi_1(t) = (t - \lambda_2) \cdots (t - \lambda_r)$ は, 各 $k = 2, \dots, r$ について別の多項式 $\Theta_k(t)$ を用いて $\Psi_1(t) = \Theta_k(t)(t - \lambda_k)$ と分解できる. このとき,

$$\begin{aligned}
 \Psi_1(f)(\mathbf{u}_k) &= \left(\Theta_k(f)(f - \lambda_k I) \right) (\mathbf{u}_k) = \Theta_k(f) \circ (f - \lambda_k I)(\mathbf{u}_k) \\
 &= \Theta_k(t) \left((f - \lambda_k I)(\mathbf{u}_k) \right) = \Theta_k(t)(\mathbf{0}) = \mathbf{0}
 \end{aligned}$$

である. ゆえに $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ の両辺に線形変換 $\Psi_1(f)$ をほどこせば

$$\begin{aligned}
 \Psi_1(f)(\mathbf{u}_1 + \dots + \mathbf{u}_r) &= \mathbf{0} \\
 \sum_{k=1}^r \Psi_1(f)(\mathbf{u}_k) &= \mathbf{0} \\
 \Psi_1(f)(\mathbf{u}_1) + \sum_{k=2}^r \Psi_1(f)(\mathbf{u}_k) &= \mathbf{0} \\
 \Psi_1(f)(\mathbf{u}_1) &= \mathbf{0}.
 \end{aligned}$$

また, 各 $k = 2, \dots, r$ について $(f - \lambda_k I)(\mathbf{u}_1) = \lambda_1 \mathbf{u}_1 - \lambda_k \mathbf{u}_1 = (\lambda_1 - \lambda_k) \mathbf{u}_1$ である. これを順次繰り返

すことで

$$\begin{aligned}
 \mathbf{0} &= \Psi_1(f)(\mathbf{u}_1) = \left((f - \lambda_2 I) \cdots (f - \lambda_r I) \right) (\mathbf{u}_1) \\
 &= (f - \lambda_2 I) \circ \cdots \circ (f - \lambda_{r-1} I) \circ (f - \lambda_r I) (\mathbf{u}_1) \\
 &= (f - \lambda_2 I) \circ \cdots \circ (f - \lambda_{r-1} I) \left((\lambda_1 - \lambda_r) \mathbf{u}_1 \right) \\
 &= (f - \lambda_2 I) \circ \cdots \circ (f - \lambda_{r-2} I) \left((\lambda_1 - \lambda_{r-1})(\lambda_1 - \lambda_r) \mathbf{u}_1 \right) \\
 &= \cdots = (\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_r) \mathbf{u}_1.
 \end{aligned}$$

各 λ_k は相異なる数であったゆえ $(\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_r) \neq 0$ であり, したがって $\mathbf{u}_1 = \mathbf{0}$ である.

次に $\mathbf{u}_1 + \cdots + \mathbf{u}_r = \mathbf{0}$ の両辺に $(f - \lambda_k I)$ ($k = 1, 3, 4, \dots, r$) を順次ほどこせば, 同様の理由により $\mathbf{u}_2 = \mathbf{0}$ を得る. これらと類似の議論を $\mathbf{u}_3, \dots, \mathbf{u}_r$ についても行い, $\mathbf{u}_1 = \cdots = \mathbf{u}_r = \mathbf{0}$ を得る. \square

上の証明の鍵は, 各線形変換 $(f - \lambda_k I)$ が互いに可換になることにある. 実際, 線形写像 $\Psi_1(f) = (f - \lambda_2 I) \cdots (f - \lambda_r I)$ における合成の順番を入れ替えて, $\Psi_1(f)$ を様々な形 $\Theta_k(f)(f - \lambda_k I)$ に表せたことが利いている. いまと同様に上手く可換性を用いることで, 次の命題が得られる.

命題 30.3.2. n 次正方行列 A が対角化可能であるとき, A の相異なるすべての固有値を $\lambda_1, \dots, \lambda_r$ とすれば, 多項式 $\phi(t) = (t - \lambda_1) \cdots (t - \lambda_r)$ について $\phi(A) = O$.

Proof. 仮定より \mathbb{K}^n の各元は A の固有ベクトルの線形結合で表せることから, 各固有ベクトル $\mathbf{v} \in W(\lambda_k, A)$ について $\phi(A)\mathbf{v} = \mathbf{0}$ を示せば十分である (練習 18.1.4). $\phi(t)$ は, $r-1$ 次多項式 $\Theta(t)$ を用いて $\phi(t) = \Theta(t)(t - \lambda_k)$ と書けること, および $(A - \lambda_k E)\mathbf{v} = A\mathbf{v} - \lambda_k \mathbf{v} = \mathbf{0}$ から

$$\phi(A)\mathbf{v} = \Theta(A)(A - \lambda_k E)\mathbf{v} = \Theta(A)\mathbf{0} = \mathbf{0}.$$

以上より $\phi(A) = O$. \square

以上が (III) の立場における命題 30.3.1 の証明の本質的部分である.

(III) の立場における命題 30.3.1 の証明. A の相異なるすべての固有値を $\lambda_1, \dots, \lambda_r$ とし, $\phi(t) := (t - \lambda_1) \cdots (t - \lambda_r)$ とすれば, 別の \mathbb{K} 係数多項式 Ψ をもちいて $\Phi_A(t) = \Psi(t)\phi(t)$ と書ける (詳しくは系 29.4.4 を見よ). ゆえに $\Phi_A(A) = \Psi(A)\phi(A) = \Psi(A)O = O$. \square

これまでの議論をまとめておこう. $\Phi_A(A) = O$ をいうには, $\Phi_A(A)$ の各列 $\Phi_A(A)\mathbf{e}_1, \dots, \Phi_A(A)\mathbf{e}_n$ が零ベクトルになることをいえばよい. A が対角化可能な場合は任意のベクトル (とくに標準ベクトル $\mathbf{e}_1, \dots, \mathbf{e}_n$) が固有ベクトルたちに分解可能であり, 各固有ベクトル \mathbf{v} について $\Phi_A(A)\mathbf{v} = \mathbf{0}$ となることから $\Phi_A(A)\mathbf{e}_i = \mathbf{0}$ も導かれ, すなわち $\Phi_A(A) = O$ となるのである. また, $\Phi_A(A)\mathbf{v} = \mathbf{0}$ の証明は, (I) の立場における証明でみたように $\Phi_A(t)$ の形を考慮せずに強引に $\Phi_A(t) = \sum_{k=0}^n a_k t^k$ と置いて示すこともできるのであるが, 命題 30.3.2 の証明のように因数分解の形まで考慮すると, 本質的には $\Phi_A(A)$ の積表示の一部である $\phi(A) = (A - \lambda_1 E) \cdots (A - \lambda_r E)$ さえほどこせば消えてしまうことまで分かった.

なお, A が一般の正方行列の場合は, 任意のベクトルを固有ベクトルたちに分解できない代わりに, 一般固有ベクトルと呼ばれるベクトルたちに分解できることが知られている. この分解を認めれば, 命題 30.3.2 の証明と類似する戦略により, ケーリー-ハミルトンの定理は直ちに導かれるのである (34.2 節をみよ).

もちろん, 本章の段階で一般固有ベクトルを持ち出すのは適当ではない. そこで最後に, (I) の立場における定理の証明を述べておく.

30.4 一般の場合

231 ページのコラムにおける勘違いを思い出そう. この誤解の元は, tE における t を A に置き換える際に, これをスカラー倍としての積ではなく行列の積と解釈してしまったことにある. t を A に置き

換えたのちも、この積をスカラー倍と解釈すれば誤解は生じない。すなわち、 $\Phi_A(t) = \det(tE - A)$ という表示の段階で t を A に置き換えるのであれば、

$$\Phi_A(A) = \det \begin{pmatrix} A - a_{11}E & -a_{12}E & \cdots & -a_{1n}E \\ -a_{21}E & A - a_{22}E & \cdots & -a_{2n}E \\ \vdots & \vdots & \vdots & \vdots \\ -a_{n1}E & -a_{n2}E & \cdots & A - a_{nn}E \end{pmatrix}$$

なる行列を考察すればよい。ここで、上の行列式の中に現れているものは、行列を成分とする行列である。このような表示が正当化されるかどうかも含めて、定理の証明の前に $\text{End}(U)$ の元を成分とする行列について補足しておこう。

$\text{End}(U)$ の間には和と積が定義されていたから、 \mathbb{K} を成分とする行列の場合と同様にして $\text{End}(U)$ 成分の行列の和、積、行列式などが定義できる。例えば、 $X = [g_{ij}]$ を $\text{End}(U)$ 成分の n 次正方行列とすれば、その行列式

$$\det X := \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) (g_{1\sigma(1)} \circ g_{2\sigma(2)} \circ \cdots \circ g_{n\sigma(n)}) \quad (30.4.1)$$

は $\text{End}(U)$ の元、すなわち何らかの線形変換である。

さて、 \mathbb{K} 成分の行列について証明したことを並行して論じることで、 $\text{End}(U)$ 成分の行列に関する多くの命題が得られる。ここで注意すべきことは、行列式の性質などの証明において各成分どうしの積の可換性 ($ab = ba$) を用いた部分がいくつかあり、可換性が成り立つとは限らない $\text{End}(U)$ 成分の行列においては、これらを適用できない点である。しかし、互いに可換な元のみを成分とする特別な行列のみを考えるのであれば、その限りではない。例えば、各 $g_{ij} \in \text{End}(U)$ が互いに可換であるとき、正方行列 $X = [g_{ij}]$ について、その余因子行列を \tilde{X} とすれば、13.2 節で論じたことと同様にして

$$\tilde{X}X = \begin{bmatrix} \det X & & \\ & \ddots & \\ & & \det X \end{bmatrix} = X\tilde{X} \quad (30.4.2)$$

が成り立つ。

さらに証明の準備として、ここでは計算を見やすくする観点から代入記号の用法を既存のものから次の通りに変更する：

- ベクトル $\mathbf{u} \in U$ を線形写像 $g: U \rightarrow U$ に代入したものを $(\mathbf{u})g$ と書くことにする。普段 $(h \circ g)(\mathbf{u})$ あるいは $(hg)(\mathbf{u})$ と書いていたものは、ここでは $((\mathbf{u})g)h$ と書くため、これを $(\mathbf{u})(g \circ h)$ あるいは $(\mathbf{u})(gh)$ と書く。ここで、式 (30.4.1) に現れた写像の合成 $g_{1\sigma(1)} \circ g_{2\sigma(2)} \circ \cdots \circ g_{n\sigma(n)}$ を新旧どちらの用法とみなすか混乱するかもしれない。ここ限りの約束として、この写像に \mathbf{u} を代入したものは、左側から順に合成して得られるもののこと、すなわち

$$(\cdots (((\mathbf{u})g_{1\sigma(1)})g_{2\sigma(2)}) \cdots)g_{n\sigma(n)}$$

のことに定める。なお、ここでは互いに可換な線形変換しか考えないことから、上式は普段の合成記号の用法による $(g_{1\sigma(1)} \circ \cdots \circ g_{n\sigma(n)})(\mathbf{u})$ と (意味は異なるものの) 値自体は等しい。すなわち、各 g_{ij} が可換であるとき、式 (30.4.1) の合成は新旧どちらの意味で考えても結局は一致する。

- 各 $g, h \in \text{End}(U)$ について、いま上で定めた新しい合成の記法 $(\mathbf{u})(gh) := (\mathbf{u})(g \circ h)$ によって g と h の積 gh を定める。なお、和 $g + h$ とスカラー倍 rg ($r \in \mathbb{K}$) についてはこれまでと同様に定める (すなわち $(\mathbf{u})(g + h) := (\mathbf{u})g + (\mathbf{u})h$, $(\mathbf{u})(rg) := r(\mathbf{u})g$)。また、 $\text{End}(U)$ の元を成分にもつ行列たちの積の計算は、各成分ごとの積と和をいま定めた意味で行うとする。更に、 U の元 \mathbf{u} と、 $\text{End}(U)$ の元 g の積を $\mathbf{u}g := (\mathbf{u})g$ と定める。これにより、 U の元を成分にもつ行ベクトルと $\text{End}(U)$ の元を成分にもつ行列の積が導入される。

定理 30.0.11 の証明. (ii) を示す. U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による f の表現行列を $A = [a_{ij}]$ とすれば,

$$\begin{bmatrix} (\mathbf{u}_1)f & (\mathbf{u}_2)f & \cdots & (\mathbf{u}_n)f \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} A$$

である. この行ベクトルは, 上で導入したベクトルと行列の積として, 次のように書き換えられる. ここで, $\text{End}(U)$ の元を成分とする対角行列の空白部分には $\text{End}(U)$ の零ベクトル (すなわち, すべてのベクトルを $\mathbf{0}_U$ に写す線形写像) が現れているとする.

$$\begin{aligned} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} f & & & \\ & f & & \\ & & \ddots & \\ & & & f \end{bmatrix} &= \begin{bmatrix} (\mathbf{u}_1)f & (\mathbf{u}_2)f & \cdots & (\mathbf{u}_n)f \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} A \\ &= \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} a_{11}I & a_{12}I & \cdots & a_{1n}I \\ a_{21}I & a_{22}I & \cdots & a_{2n}I \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}I & a_{n2}I & \cdots & a_{nn}I \end{bmatrix}. \end{aligned}$$

さて, U の元を成分にもつ行ベクトル \mathbf{x} , および $\text{End}(U)$ の元を成分にもつ行列 Y, Z について分配律 $\mathbf{x}Y - \mathbf{x}Z = \mathbf{x}(Y - Z)$ が成り立つことが容易に示され, そこで上式を移項したのちに分配律を適用すれば,

$$\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} f - a_{11}I & -a_{12}I & \cdots & -a_{1n}I \\ -a_{21}I & f - a_{22}I & \cdots & -a_{2n}I \\ \vdots & \vdots & \vdots & \vdots \\ -a_{n1}I & -a_{n2}I & \cdots & f - a_{nn}I \end{bmatrix} = \begin{bmatrix} \mathbf{0}_U & \mathbf{0}_U & \cdots & \mathbf{0}_U \end{bmatrix}.$$

上式に現れる $\text{End}(U)$ 成分の行列を X とおけば, X の各成分は互いに可換である (命題 21.5.6). また, 結合律 $(\mathbf{x}Y)Z = \mathbf{x}(YZ)$ が成り立つことも容易に示され, したがって上の両辺に右から X の余因子行列 \tilde{X} をかければ, 式 (30.4.2) より

$$\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \det X & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \det X \end{bmatrix} = \begin{bmatrix} \mathbf{0}_U & \mathbf{0}_U & \cdots & \mathbf{0}_U \end{bmatrix}. \quad (30.4.3)$$

を得る. 一方,

$$\begin{aligned} \det X &= \text{“多項式 } \Phi_A(t) = |tE - A| \text{ の展開式における } t^n \text{ を } f^n \text{ に置き換えた写像”} \\ &= \Phi_A(f) = \Phi_f(f) \end{aligned}$$

である. ゆえに式 (30.4.3) は, 各 $j = 1, \dots, n$ について $(\mathbf{u}_j)(\Phi_f(f)) = \mathbf{0}_U$ であることを意味する. すなわち, 線形変換 $\Phi_f(f)$ はすべての元を $\mathbf{0}_U$ に写す写像である. \square

第31章 斉次形線形漸化式

固有ベクトルへの分解による応用として、本章および次章を通して線形漸化式の一般項の導出、および定数係数線形常微分方程式の解法を斉次形の場合に限って解説する。本章と次章における議論の鍵は、シフト作用素 S や微分作用素 D の固有ベクトルの形があらかじめ分かっているため、適切な設定のもとで固有値さえ求めれば直ちに固有ベクトルへの分解、すなわち一般項や一般解の表示が得られるところにある。

本章では初めに簡単な2次の漸化式の解法について紹介し、後半で一般の高次の場合について論じる。なお、これまでに得た知識から結論が出せるのは、特性方程式が重解を持たない場合に限られている。特性方程式が重解を持つ場合は、34章で述べる一般固有ベクトルへの分解を通して一般項の表示が得られる(詳しくは34.3節を見よ)。

31.1 線形漸化式と固有値

既知の数 $a_1, a_0 \in \mathbb{K}$ によって定められる2次の漸化式

$$x_{n+2} + a_1 x_{n+1} + a_0 x_n = 0 \quad (31.1.1)$$

について論じよう。漸化式(31.1.1)を満たす数列全体のなす部分空間を $F \subset \mathbb{K}^{\mathbb{N}}$ と置き、 $S: \mathbb{K}^{\mathbb{N}} \rightarrow \mathbb{K}^{\mathbb{N}}$ を $S((x_n)_{n \in \mathbb{N}}) := (x_{n+1})_{n \in \mathbb{N}}$ で定めるシフト作用素とする。 S や F に関する基本的な事実は例20.4.4に述べた通りである。とくに、 $S(F) \subset F$ であり、 $S|_F$ は F 上の線形変換と見なせるのであった。まず、 F の次元を確認する。

命題 31.1.1. $\dim F = 2$.

Proof. 初項が1で第2項が0の F の元を \mathbf{u}_1 、初項が0で第1項が1の F の元を \mathbf{u}_2 とすれば、任意の F の元は $\mathbf{u}_1, \mathbf{u}_2$ の線形結合で書ける。実際、 $\mathbf{v} \in F$ の初項を p 、第2項を q とすれば、 $\mathbf{v} = p\mathbf{u}_1 + q\mathbf{u}_2$ が成り立つ。何故ならこの両辺の初項と第2項は等しく、漸化式(31.1.1)を満たす数列は初項と第2項さえ決めれば第3項以降の値は自動的に決まってしまうからである。 \square

S の固有ベクトルは次のように直ちに分かる:

命題 31.1.2. $S: \mathbb{K}^{\mathbb{N}} \rightarrow \mathbb{K}^{\mathbb{N}}$ の固有値 $\lambda \in \mathbb{K}$ に関する固有ベクトルは公比 λ なる等比数列に限る。

Proof. 初項が0でない公比 λ の等比数列が固有ベクトルとなることは明らかである。一方、数列 $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ を S の固有値 λ に関する固有ベクトルとすれば、 $S(\mathbf{x}) = \lambda \mathbf{x}$ 、つまり $(x_{n+1})_{n \in \mathbb{N}} = (\lambda x_n)_{n \in \mathbb{N}}$ が成り立つ。各 $n \in \mathbb{N}$ について $x_{n+1} = \lambda x_n$ となるゆえ \mathbf{x} は公比 λ の等比数列である。 \square

漸化式(31.1.1)を満たす数列の一般項は次の方針に沿って導出される:

一般項の解法の方針

$S|_F: F \rightarrow F$ の固有ベクトルは等比数列のみであり、初項1公比 λ の等比数列の一般項は λ^{n-1} と良く分かっている。したがって、 $S|_F$ の固有値 λ_1, λ_2 が求まり、 $\lambda_1 \neq \lambda_2$ ならば、二つの数列 λ_1^{n-1} および λ_2^{n-1} は F の基底となる。つまり、漸化式を満たす数列の一般項は、 $r, s \in \mathbb{K}$ を用いて $r\lambda_1^{n-1} + s\lambda_2^{n-1}$ と書ける。

$S|_F$ の固有値を求めよう. S の固有ベクトルは等比数列に限ることから, 仮に数列 $x_n = t^{n-1}$ が漸化式 (31.1.1) を満たすとすれば,

$$\text{各 } n \in \mathbb{N} \text{ について, } t^{n+1} + a_1 t^n + a_0 t^{n-1} = 0.$$

上の条件は $n = 1$ の場合に限った次の条件と同値である:

$$t^2 + a_1 t + a_0 = 0.$$

上式の左辺に現れる多項式 $\Phi(t) = t^2 + a_1 t + a_0$ は漸化式 (31.1.1) の特性多項式と呼ばれ, 方程式 $\Phi(t) = 0$ をこの漸化式の特性方程式という¹. 特性方程式の解を公比とする等比数列が $S|_F$ の固有ベクトルであり, したがって我々は次を得る:

命題 31.1.3. 漸化式 (31.1.1) の特性方程式が重解を持たず, その解を λ_1, λ_2 とすれば, この漸化式を満たす任意の数列の一般項は $r, s \in \mathbb{K}$ を用いて $r\lambda_1^{n-1} + s\lambda_2^{n-1}$ と表される.

特性方程式が重解を持つ場合, その固有値に関する $S|_F$ の固有空間の次元は 1 である. つまり, F の任意の元を固有ベクトルの線形結合で表すことはできない. これと定理 29.3.1 から, $S|_F$ の表現行列が対角化できないことが分かる. この場合の漸化式の一般項は, 一般固有ベクトルへの分解として表す (定理 34.3.4).

例題 31.1.4. 次の漸化式の一般項を求めよ: $a_{n+2} = a_{n+1} + a_n$.

解答例: 特性多項式は $\Phi(t) = t^2 - t - 1$ であり, $\Phi(t) = 0$ の解は

$$\alpha = \frac{1 + \sqrt{5}}{2}, \quad \beta = \frac{1 - \sqrt{5}}{2}.$$

固有値 α, β に関する固有ベクトルをそれぞれ

$$\mathbf{x}^\alpha = (1, \alpha, \alpha^2, \alpha^3, \alpha^4, \dots), \quad \mathbf{x}^\beta = (1, \beta, \beta^2, \beta^3, \beta^4, \dots)$$

とおけば, 漸化式 $a_{n+2} = a_{n+1} + a_n$ を満たす数列 $\mathbf{x} := (x_n)_{n \in \mathbb{N}}$ は上の等比数列の線形結合で書ける. すなわち, 実数 p, q を用いて

$$\mathbf{x} = p\mathbf{x}^\alpha + q\mathbf{x}^\beta, \quad \text{とくに第 } n \text{ 項について } x_n = p\alpha^{n-1} + q\beta^{n-1}.$$

p, q を x_1, x_2 について解こう. 上式に $n = 1, 2$ を代入した連立 1 次方程式

$$\begin{cases} x_1 = \alpha + \beta, \\ x_2 = p\alpha + q\beta \end{cases}$$

を解くと,

$$p = \frac{x_2 - x_1\beta}{\alpha - \beta}, \quad q = \frac{x_1\alpha - x_2}{\alpha - \beta}.$$

ゆえに求める数列の一般項は

$$x_n = \frac{x_2 - x_1\beta}{\alpha - \beta} \alpha^{n-1} + \frac{x_1\alpha - x_2}{\alpha - \beta} \beta^{n-1}, \quad \text{ただし } \alpha = \frac{1 + \sqrt{5}}{2}, \beta = \frac{1 - \sqrt{5}}{2}.$$

特に, $x_1 = 0, x_2 = 1$ を上式に代入すると ($\alpha - \beta = \sqrt{5}$ に注意する)

$$p = \frac{1}{\sqrt{5}}, \quad q = -\frac{1}{\sqrt{5}}.$$

したがってフィボナッチ数列 $0, 1, 1, 2, 3, 5, 8, 13, 21, \dots$ の一般項は次で与えられる:

$$x_n = p\alpha^{n-1} + q\beta^{n-1} = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^{n-1} - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^{n-1}.$$

¹実は, Φ は $S|_F$ の特性多項式 $\Phi_{S|_F}$ に一致する. 詳しくは 31.3 節をみよ.

31.2 複素数列のなかの実数列

いま漸化式 (31.1.1) に現れる係数 a_1, a_0 が実数であり, しかしながら特性方程式が実数解を持たない場合を考えよう. 特性方程式の複素数解を α, β とすれば $\beta = \bar{\alpha}$ が成り立つ (命題 29.4.8). このとき命題 31.1.3 で表される数列は複素数列である. この中で実数列となるものはどのような形になるか, ここで検討しておこう.

補題 31.2.1. p, q を複素数とし, $\alpha \in \mathbb{C}$ を実数でないとする. 数列 $x_n = p\alpha^{n-1} + q\bar{\alpha}^{n-1}$ が実数列となるための必要十分条件は $p = \bar{q}$.

Proof. 複素数 $p\alpha^{n-1} + q\bar{\alpha}^{n-1}$ を実部と虚部に分解し, 虚部が常に 0 となる条件を調べよう. 実数 $r, \theta, a_1, a_2, b_1, b_2$ を用いて

$$\alpha = re^{i\theta}, \quad \bar{\alpha} = re^{-i\theta}, \quad p = a_1 + b_1i, \quad q = a_2 + b_2i$$

と書ける. ただし $r > 0$ とし, 必要があれば α と $\bar{\alpha}$ の役割を入れ替えることで, $0 < \theta < \pi$ としてよい.

$$\begin{aligned} x_n &= pr^{n-1}e^{i(n-1)\theta} + qr^{n-1}e^{i(n-1)\theta} \\ &= pr^{n-1}(\cos(n-1)\theta + i\sin(n-1)\theta) + qr^{n-1}(\cos(n-1)\theta - i\sin(n-1)\theta) \\ \frac{x_n}{r^{n-1}} &= (p+q)\cos(n-1)\theta + (p-q)i\sin(n-1)\theta \\ &= ((a_1+a_2) + (b_1+b_2)i)\cos(n-1)\theta + ((a_1-a_2) + (b_1-b_2)i)i\sin(n-1)\theta \\ &= ((a_1+a_2)\cos(n-1)\theta - (b_1-b_2)\sin(n-1)\theta) \\ &\quad + i((b_1+b_2)\cos(n-1)\theta + (a_1-a_2)\sin(n-1)\theta). \end{aligned}$$

したがって, $p = \bar{q}$ (つまり $b_2 = -b_1$ および $a_2 = a_1$) ならば上式は実数となる. 逆に, 上式の虚部

$$(b_1 + b_2)\cos(n-1)\theta + (a_1 - a_2)\sin(n-1)\theta$$

が $n \in \mathbb{N}$ によらず 0 となるためには, 数列の組 $\cos(n-1)\theta, \sin(n-1)\theta$ の線形独立性より $b_1 + b_2 = 0$ かつ $a_1 - a_2 = 0$ でなければならない. 実際, $n = 1$ を代入すれば $(b_1 + b_2)\cos 0 = 0$ より $b_1 + b_2 = 0$ であり, $n = 2$ を代入すれば $(a_1 - a_2)\sin \theta = 0$. いま $0 < \theta < \pi$ としているゆえ $\sin \theta \neq 0$. ゆえに $a_1 - a_2 = 0$ を得る. \square

$p = \bar{q}$ のとき, 上の補題の証明によれば,

$$x_n = r^{n-1}(2a_1\cos(n-1)\theta - 2b_1\sin(n-1)\theta)$$

が成り立つ. $2a_1, -2b_1$ を改めて C_1, C_2 と置きなおすことで, 漸化式 (31.1.1) を満たす実数列 x_n の一般項は次のように表せる:

$$x_n = r^{n-1}(C_1\cos(n-1)\theta + C_2\sin(n-1)\theta) \quad (C_1, C_2 \text{ は任意定数}).$$

ここで, r および θ は特性根 α の絶対値および偏角である.

31.3 高次の線形漸化式と表現行列 (発展)

次に, より高次の斉次形線形漸化式

$$x_{n+k} + a_{k-1}x_{n+k-1} + a_{k-2}x_{n+k-2} + \cdots + a_1x_{n+1} + a_0x_n = 0 \quad (31.3.1)$$

について論じよう. 基本的な考え方は先ほどの 2 次の場合と変わらない. 上式を移項して, 次の式を考えても良い:

$$x_{n+k} = -a_{k-1}x_{n+k-1} - a_{k-2}x_{n+k-2} - \cdots - a_1x_{n+1} - a_0x_n. \quad (31.3.2)$$

漸化式 (31.3.2) と漸化式 (16.3.2) では係数の符号が異なることに注意せよ. 上の漸化式を満たす数列全体のなす部分空間を $F_k \subset \mathbb{K}^{\mathbb{N}}$ と置き, 本節ではこれを $F = F_k$ と略記する. 既に述べたように $S(F) \subset F$ であり, $S|_F$ は F 上の線形変換と見なせる.

命題 31.3.1. $\dim F = k$.

Proof. 次のような k 個の F の元からなる組 $\mathbf{u}_1, \dots, \mathbf{u}_k$ を考えよう.

$$\begin{aligned}\mathbf{u}_1 &= (1, 0, \dots, 0, -a_0, a_0 a_{k-1}, \dots), \\ \mathbf{u}_2 &= (0, 1, \dots, 0, -a_1, -a_0 + a_1 a_{k-1}, \dots), \\ &\vdots \\ \mathbf{u}_k &= (0, 0, \dots, 1, -a_{k-1}, a_{k-2} + a_{k-1}^2, \dots).\end{aligned}$$

すなわち、各 \mathbf{u}_i は第 k 項までのうち第 i 項が 1 でそれ以外の項が 0 であり、 $k+1$ 項より先は漸化式 (31.3.2) によって順次さだめられる数列である。これらが F の基底となることを確認しよう。まず、第 k 項までの値から、 $\mathbf{u}_1, \dots, \mathbf{u}_k$ の線形独立性はすぐに分かる。次に、各 $\mathbf{v} = (v_1, v_2, v_3, \dots) \in F$ に対して、

$$\mathbf{v} = \sum_{i=1}^k v_i \mathbf{u}_i$$

が成り立つ。実際、上式の両辺の各項について、第 k 項までが等しいことは明らかである。また、 F の各元は、第 k 項までの値を決めてしまえば以降の項の値は漸化式 (31.3.2) によって自動的に決まってしまう。つまり上式は第 k 項以降も一致しなければならない。この事実の形式的な証明を補題 31.3.2 として与えておこう。以上により、 $\mathbf{u}_1, \dots, \mathbf{u}_k$ は F の基底となる。□

数列 $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ に対して、 \mathbf{x} の第 k 項までに限った有限列を $\mathbf{x}|_k$ と書く。つまり $\mathbf{x}|_k = (x_1, \dots, x_k)$ である。

補題 31.3.2. 各 $\mathbf{u}, \mathbf{v} \in F$ について、 $\mathbf{u}|_k = \mathbf{v}|_k \implies \mathbf{u} = \mathbf{v}$.

Proof. $\mathbf{u} = (u_1, u_2, \dots)$, $\mathbf{v} = (v_1, v_2, \dots)$ と置く。各 $n \in \mathbb{N}$ について $u_{k+n} = v_{k+n}$ を示せばよい。これを n に関する帰納法により示そう。 $k+n-1$ 以下の項が一致している (つまり $\mathbf{u}|_{k+n-1} = \mathbf{v}|_{k+n-1}$) と仮定すと、漸化式 (31.3.2) より

$$\begin{aligned}u_{n+k} &= a_{k-1}u_{n+k-1} + a_{k-2}u_{n+k-2} + \dots + a_1u_{n+1} + a_0u_n \\ &= a_{k-1}v_{n+k-1} + a_{k-2}v_{n+k-2} + \dots + a_1v_{n+1} + a_0v_n = v_{n+k}.\end{aligned}$$

つまり \mathbf{u}, \mathbf{v} の第 $k+n$ 項も一致する。□

31.1 節で述べたことと同様に、 $S|_F$ の固有値を求めるには、

$$\Phi(t) = t^k + a_{k-1}t^{k-1} + \dots + a_1t + a_0 \quad (31.3.3)$$

によって定まる k 次方程式 $\Phi(t) = 0$ を解けばよい。理解を深めるためにここでは少しよりみちをして、上の $\Phi(t)$ が $S|_F$ の特性多項式 $\Phi_{S|_F}(t)$ に一致することを見よう。命題 31.3.1 で与えた F の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ について、それぞれを S に代入すると

$$\begin{aligned}S(\mathbf{u}_1) &= (0, \dots, 0, -a_0, a_0 a_{k-1}, \dots) = -a_0 \mathbf{u}_k, \\ S(\mathbf{u}_2) &= (1, \dots, 0, -a_1, -a_0 + a_1 a_{k-1}, \dots) = \mathbf{u}_1 - a_1 \mathbf{u}_k, \\ &\vdots \\ S(\mathbf{u}_i) &= \mathbf{u}_{i-1} - a_{i-1} \mathbf{u}_k, \\ &\vdots \\ S(\mathbf{u}_k) &= (0, \dots, 1, -a_{k-1}, a_{k-2} + a_{k-1}^2, \dots) = \mathbf{u}_{k-1} - a_{k-1} \mathbf{u}_k.\end{aligned}$$

したがって基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する S の表現行列は

$$A = \begin{bmatrix} 0 & 1 & & & O \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ O & & & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{k-2} & -a_{k-1} \end{bmatrix}. \quad (31.3.4)$$

$S|_F$ の特性方程式 $\Phi_{S|_F}(t) := \Phi_A(t)$ は例 13.1.4 により次の式で表される:

$$\Phi_{S|_F}(t) = |tE - A| = \begin{vmatrix} t & -1 & & & O \\ & t & -1 & & \\ & & \ddots & \ddots & \\ O & & & t & -1 \\ a_0 & a_1 & \cdots & a_{k-2} & t + a_{k-1} \end{vmatrix} = t^k + a_{k-1}t^{k-1} + \cdots + a_1t + a_0.$$

定義 31.3.3. 漸化式 (31.3.1) の左辺において, 各 x_{n+i} ($i = 0, \dots, k$) を t^i に置き換えることによって得られる多項式 $\Phi(t)$ (式 (31.3.3) のこと) を, この漸化式の特性多項式と呼ぶ. 上の議論により, これは $S|_F$ の特性多項式 $\Phi_{S|_F}(t)$ に等しい. 方程式 $\Phi(t) = 0$ を, この漸化式の特性方程式と呼ぶ.

$S|_F$ がちょうど k 個の相異なる固有値を持つとき, それらに対応する k 個の固有ベクトル (すなわち固有値を公比とする等比数列) からなる組は命題 29.1.1 により線形独立であり, したがって F の基底となる:

命題 31.3.4. 初項 1 公比 λ の等比数列を \mathbf{x}^λ とする (つまり $\mathbf{x}^\lambda = (1, \lambda, \lambda^2, \lambda^3, \dots)$). 漸化式 (31.3.1) の特性方程式が互いに異なる k 個の解 $\lambda_1, \dots, \lambda_k$ を持つならば, この漸化式を満たす任意の数列 $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ は $\mathbf{x} = \sum_{i=1}^k r_i \mathbf{x}^{\lambda_i}$ と表せる. このとき, \mathbf{x} の一般項の表示は $x_n = \sum_{i=1}^k r_i \lambda_i^{n-1}$ である.

第32章 斉次形線形常微分方程式

定数係数線形常微分方程式の解法を斉次形の場合に限って解説する. 20.4節で紹介した漸化式と微分方程式の類似性, とくに例 20.4.6 で与えた対応関係から, 微分方程式の解法は前章で論じた漸化式の一般項の導出法と並行して論じられることが示唆される. ただし漸化式の場合と違い, 微分方程式を解くには微分積分学あるいは複素関数論の知識が必要である. 本章では高校数学における微積分法の知識のみで理解できる部分については証明を与え, それを超える部分については証明を略した. また, 関数の定義域は \mathbb{R} としているが, これを一般の开区間 I としても同様に議論が進められるだろう.

32.1 線形常微分方程式と固有値

前章と同様に, まず初めに既知の数 $a_1, a_0 \in \mathbb{R}$ によって定められる簡単な2次の微分方程式

$$y''(x) + a_1 y'(x) + a_0 y(x) = 0 \quad (32.1.1)$$

の解法について紹介し, 後半で一般の高次の場合について論じる.

方程式 (32.1.1) を満たす関数全体のなす $C^\infty(\mathbb{R})$ の部分空間を W とし, $D : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ を $D(y) = y'$ で定める微分作用素とする. D や W に関する基本的な事実は例 20.4.5 に述べた通りである. とくに $D(W) \subset W$ であり, $D|_W$ は W の線形変換と見なせる.

W の次元が2であることはどのようにして理解できるだろうか. 方程式 (32.1.1) の両辺を n 回微分して移項することにより,

$$y^{(n+2)}(x) = -a_1 y^{(n+1)}(x) - a_0 y(x)^{(n)}$$

を得る. つまり, W に含まれる関数は, $y(x)$ および $y'(x)$ の値さえ決めれば, それ以降の n 次導関数 $y^{(2)}(x), y^{(3)}(x), y^{(4)}(x), \dots$ の値も上の漸化式によって自動的に決まってしまう. すなわち, 数列 $y^{(n)}(x)$ は漸化式 (31.1.1) を満たす. 話を分かりやすくするために W の元が原点 $a = 0$ においてテイラー展開できるとしよう. すると, いまの考察から, W 内の二つの関数が一致するか否かは原点における関数の値と微分係数がそれぞれ一致するか否かで決定できる. とくに, $(u_1(0), u_1'(0)) = (1, 0)$ を満たす関数 $u_1(x)$ および, $(u_2(0), u_2'(0)) = (0, 1)$ を満たす関数 $u_2(x)$ が W の中に存在すれば, u_1, u_2 が W の基底となる. この事実の厳密な証明は割愛する.

次に, 微分作用素 D の固有ベクトル (これを固有関数と呼ぶ場合もある) を求めよう.

命題 32.1.1. 微分作用素 $D : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ ($D(y) := y'$) において, D の固有ベクトルは $ae^{\lambda x}$ の形のものに限られる. ただし, $a \neq 0$ は定数である.

Proof. $y = f(x)$ が D の固有値 λ に関する固有ベクトルであるならば,

$$\frac{dy}{dx} = \lambda y \quad (32.1.2)$$

を満たす. これは変数分離型の微分方程式ゆえ, 上式を変形した $\frac{1}{y} dy = \lambda dx$ の両辺に積分記号を付加することで解ける. こうした手法が正当化される理由をここでは振り返ってみよう.

$y(x)$ が0に値を取らないと仮定し, 式 (32.1.2) を変形した式 $\frac{1}{y} \frac{dy}{dx} = \lambda$ において, これらを x の関数とみなして両式の不定積分をとれば

$$\int \frac{1}{y} \frac{dy}{dx} dx = \int \lambda dx = \lambda x + C_1. \quad (C_1 \text{ は積分定数})$$

また、上式の左辺は置換積分公式により次のように変形される:

$$\int \frac{1}{y} \frac{dy}{dx} dx = \int \frac{1}{y} dy = \log |y| + C_2. \quad (C_2 \text{ は積分定数})$$

以上より積分定数をまとめれば、 $\log |y| = \lambda x + C$ である。つまり、 $|y| = e^{\lambda x + C} = e^C \cdot e^{\lambda x}$ 。ゆえに各 $x \in \mathbb{R}$ について $y(x) = e^C \cdot e^{\lambda x}$ または $y(x) = -e^C \cdot e^{\lambda x}$ である。ここで、 \mathbb{R} の連結性から次が成り立つ:

- 補題 32.1.2.** (1) $y(x)$ がある x_0 について正の値を取るならば、すべての $x \in \mathbb{R}$ について $y(x) > 0$ 、
 (2) $y(x)$ がある x_1 について負の値を取るならば、すべての $x \in \mathbb{R}$ について $y(x) < 0$ 。

補題の証明. $y(x_0) > 0$ であるとし、(1) を背理法により示そう。仮に $y(x) \leq 0$ を満たす $x \in \mathbb{R}$ があるとしよう。このとき $y(x) \neq 0$ であるから $y(x) < 0$ である。このとき中間値の定理により、 $y(a) = 0$ を満たす $a \in \mathbb{R}$ が x_0 と x の間に存在する。これは関数 $y(x)$ が 0 に値をとらないことに反する。したがって常に $y(x) > 0$ でなければならない。(2) も類似の論法で示すことができる。□

以上より、 $y(x)$ は $x \in \mathbb{R}$ の位置によらずに $e^C \cdot e^{\lambda x}$ となるか、あるいは $-e^C \cdot e^{\lambda x}$ となる。□

【補足】 上の証明では $y(x)$ が 0 に値をとらないことを仮定していた。一方で、0 値定数関数は式 (32.1.2) を満たす。もちろんこれは $C^\infty(\mathbb{R})$ の零ベクトルゆえ D の固有ベクトルではない。また、定数関数を除いて 0 に値をとる関数で式 (32.1.2) を満たすものは存在しない。なぜなら、そのような関数は、0 に値をとらない部分については上の証明にあるように $\pm e^C \cdot e^{\lambda x}$ なる形をしており、この式を拡張して 0 に値をとる関数をつくると不連続関数になってしまうからである。

一般解の解法の方針

$D|_W : W \rightarrow W$ の固有ベクトルは指数関数 $e^{\lambda x}$ の定数倍のみである。したがって、 $D|_W$ の固有値 λ_1, λ_2 が求まり、 $\lambda_1 \neq \lambda_2$ ならば、 $e^{\lambda_1 x}, e^{\lambda_2 x}$ は W の基底となる。

$D|_W$ の固有値を求めよう。仮に $y(x) = e^{tx} \in W$ とすれば、

$$\begin{aligned} (e^{tx})'' + a_1 (e^{tx})' + a_0 e^{tx} &= 0 \\ (t^2 + a_1 t + a_0) e^{tx} &= 0. \end{aligned}$$

$e^{tx} \neq 0$ より $t^2 + a_1 t + a_0 = 0$ を得る。多項式 $\Phi(t) = t^2 + a_1 t + a_0$ および方程式 $\Phi(t) = 0$ をそれぞれ、微分方程式 (32.1.1) の特性多項式、特性方程式と呼ぶ。 $D|_W$ の固有値は特性方程式の解に等しい。

【補足】 先の考察によれば数列 $(y^{(n-1)}(0))_{n \in \mathbb{N}} = (t^{n-1})_{n \in \mathbb{N}}$ は漸化式 (31.1.1) を満たす。この事実からも $t^2 + a_1 t + a_0 = 0$ が得られる。

結局、漸化式の場合と類似した次の事実を得る:

命題 32.1.3. 微分方程式 (32.1.1) の特性方程式が重解を持たず、その解を $\lambda_1, \lambda_2 \in \mathbb{R}$ とすれば、この微分方程式の一般解は $r, s \in \mathbb{R}$ を用いて $y(x) = r e^{\lambda_1 x} + s e^{\lambda_2 x}$ と表される。

例題 32.1.4. 微分方程式 $\frac{d^2}{dx^2} y(x) = \frac{d}{dx} y(x) + y(x)$ を解け。

解答例: 特性多項式は $\Phi(t) = t^2 - t - 1$ であり、 $\Phi(t) = 0$ の解は

$$\lambda_1 = \frac{1 + \sqrt{5}}{2}, \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

ゆえに方程式の一般解は

$$y(x) = C_1 e^{\frac{1+\sqrt{5}}{2}x} + C_2 e^{\frac{1-\sqrt{5}}{2}x} \quad (r_1, r_2 \text{ は任意定数}).$$

特性方程式が重解を持つ場合は、漸化式で論じたことと同様に W の任意の元を固有ベクトルの線形結合で表すことはできない。この場合の一般解は、一般固有ベクトルを用いて表示する (定理 34.3.3)。

32.2 特性多項式が複素解をもつ場合における実数解

31.2 節と類似の議論が微分方程式においても成立することを補足しておこう。複素関数に関する微積分を論じることにより、微分方程式 (32.1.1) の特性方程式が複素解 $\lambda, \bar{\lambda}$ を持つとき、この微分方程式の複素関数としての解は命題 32.1.3 における r, s を複素数として取ることで得られることが知られている。この解が実数を代入した際に実数を与えるための条件を導くことで、我々は微分方程式 (32.1.1) の実数関数としての解を得る。

補題 32.2.1. c_1, c_2 および λ を複素数の定数とし、 $y(x) := c_1 e^{\lambda x} + c_2 e^{\bar{\lambda} x}$ とする。関数 $y: \mathbb{R} \rightarrow \mathbb{C}$ が実数値関数になるための必要十分条件は、 $c_2 = \bar{c}_1$ 。

Proof. 実数 $a, b, \alpha_1, \beta_1, \alpha_2, \beta_2$ を用いて $\lambda, \bar{\lambda}, c_1, c_2$ を

$$\lambda = a + bi, \quad \bar{\lambda} = a - bi, \quad c_1 = \alpha_1 + \beta_1 i, \quad c_2 = \alpha_2 + \beta_2 i$$

とおく。オイラーの公式 $e^{i\theta} = \cos \theta + i \sin \theta$ を用いて $y(x)$ を三角関数に分解すれば、

$$\begin{aligned} y(x) &= c_1 e^{(a+bi)x} + c_2 e^{(a-bi)x} = c_1 e^{ax} e^{ibx} + c_2 e^{ax} e^{-ibx} \\ &= c_1 e^{ax} (\cos(bx) + i \sin(bx)) + c_2 e^{ax} (\cos(bx) - i \sin(bx)) \\ &= e^{ax} \left((c_1 + c_2) \cos(bx) + (c_1 - c_2) i \sin(bx) \right) \\ &= e^{ax} \left(((\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) i) \cos(bx) + ((\alpha_1 - \alpha_2) + (\beta_1 - \beta_2) i) i \sin(bx) \right) \\ &= e^{ax} \left(((\alpha_1 + \alpha_2) \cos(bx) - (\beta_1 - \beta_2) \sin(bx)) + i((\beta_1 + \beta_2) \cos(bx) + (\alpha_1 - \alpha_2) \sin(bx)) \right). \end{aligned}$$

したがって、任意の実数 x について $y(x)$ が実数になることは、その虚部

$$e^{ax} \left((\beta_1 + \beta_2) \cos(bx) + (\alpha_1 - \alpha_2) \sin(bx) \right)$$

が常に 0 であることと同値である。

いまの考察から、 $c_2 = \bar{c}_1$ を満たすときに $y(x)$ が実数値関数となることはすぐに分かる。逆に、 $y(x)$ が実数値関数となるとき $c_2 = \bar{c}_1$ となることを示そう。 $e^{ax} \neq 0$ より、これは

$$\text{各 } x \in \mathbb{R} \text{ について、} \quad (\beta_1 + \beta_2) \cos(bx) + (\alpha_1 - \alpha_2) \sin(bx) = 0$$

を意味する。関数空間において $\cos(bx)$ と $\sin(bx)$ は線形独立であることから、 $\beta_1 + \beta_2 = 0$, $\alpha_1 - \alpha_2 = 0$ を得る¹。つまり、 $c_2 = \bar{c}_1$ である。また、このとき

$$y(x) = 2\alpha_1 e^{ax} \cos(bx) - 2\beta_1 e^{ax} \sin(bx). \quad (32.2.1)$$

□

式 (32.2.1) について、 $r_1 = 2\alpha$, $r_2 = -2\beta$ と置きなおすことで次を得る：

系 32.2.2. 実数を係数とする微分方程式 (32.1.1) の特性方程式が実数でない複素数による解 $\lambda, \bar{\lambda}$ を持つとする。このとき $\lambda = a + bi$ (ただし $a, b \in \mathbb{R}$) とすれば、この微分方程式の実数解は次で与えられる：

$$y(x) = r_1 e^{ax} \cos(bx) + r_2 e^{ax} \sin(bx) \quad (r_1, r_2 \text{ は任意定数}).$$

練習 32.2.3. 微分方程式 $y'' + 2y' + 3y = 0$ を満たす実数値関数を求めよ。

解答例: 特性方程式 $\lambda^2 + 2\lambda + 3 = 0$ を解くと、 $\lambda = -1 \pm \sqrt{2}i$ である。よってこの微分方程式の実数関数としての解は

$$y(x) = r_1 e^{-x} \cos(\sqrt{2}x) + r_2 e^{-x} \sin(\sqrt{2}x).$$

¹例えば $x = 0$ を代入することで $\beta_1 + \beta_2 = 0$ を得る。 $x = \frac{\pi}{2b}$ を代入することで $\alpha_1 - \alpha_2 = 0$ を得る。

32.3 高次の線形常微分方程式 (発展)

高次の常微分方程式

$$y^{(k)}(x) + a_{k-1}y^{(k-1)}(x) + a_{k-2}y^{(k-2)}(x) + \cdots + a_1y^{(1)}(x) + a_0y^{(0)}(x) = 0. \quad (32.3.1)$$

について解説する. 上の方程式の解空間を $W_k \subset C^\infty(\mathbb{R})$ と置き, 本節ではこれを $W = W_k$ と略記する. 既に述べたように $D(W) \subset W$ であり, $D|_W$ は W の線形変換と見なせる.

$\dim W = k$ を示すために必要となる事実を挙げておこう. 次の定理の証明は微分方程式論の専門書に譲る.

定理 32.3.1. $a \in \mathbb{R}$ を一つ固定する. $z(x) \in W$ が $(z(a), z'(a), \dots, z^{(k-1)}(a)) = \mathbf{0}$ を満たすならば $z(x) \equiv 0$, すなわち $z(x)$ は 0 値定数関数である.

上の事実から, W の元は $k-1$ 次以下の微分係数によって決定されることが分かる. すなわち,

系 32.3.2. (1) $a \in \mathbb{R}$ および $\xi(x), \eta(x) \in W$ について,

$$(\xi(a), \xi'(a), \dots, \xi^{(k-1)}(a)) = (\eta(a), \eta'(a), \dots, \eta^{(k-1)}(a)) \implies \xi(x) = \eta(x).$$

(2) $\dim W \leq k$.

Proof. 線形写像 $G: W \rightarrow \mathbb{R}^k$ を $G(\xi) := (\xi(a), \xi'(a), \dots, \xi^{(k-1)}(a))$ と定める. 定理 32.3.1 によれば $\text{Ker } G$ は自明であり, したがって定理 20.3.4 より G は単射である. これは (1) の主張にほかならない. また, 命題 22.4.5(1) より $\dim W \leq \dim \mathbb{R}^k = k$. \square

練習 32.3.3. 上の結果を用いて命題 32.1.1 を示せ.

解答例: $\xi(x)$ を D の固有値 λ に関する D の固有ベクトルとする. $A := \xi(0)$ とし, $\eta(x) = Ae^{\lambda x}$ とおく. このとき, ξ, η はともに微分方程式 $y'(x) - \lambda y(x) = 0$ の解であり, $\xi(0) = \eta(0)$ を満たす. ゆえに系 32.3.2(1) より $\xi(x) = \eta(x)$. \square

我々は最終的には, 微分方程式の解の公式として解空間 W における線形独立な k 個の関数の組を与える (定理 34.3.3). このことから $\dim W \geq k$ が分かり, 先に示した $\dim W \leq k$ と合せて $\dim W = k$ を得る. ここでは, 31.3 節で与えた部分空間 $F \subset \mathbb{R}^N$ のかりそめの基底 $\mathbf{u}_1, \dots, \mathbf{u}_k$ に対応する $W \subset C^\infty(\mathbb{R})$ の基底がどのような関数であるのか検討しよう.

いま, 定義域上の点 $a \in \mathbb{R}$ を一つ取って決めておく. この a は定義域上のどの点でも構わない. とくに $a = 0$ として考えると以降の式はいくぶんか楽になる. 微分方程式 (32.3.1) の両辺を n 階微分することで漸化式 (31.3.1) を得る. すなわち, 各 $y(x) \in W$ について, 点 a での n 階微分係数の列を

$$\mathbf{y} = (y(a), y'(a), y''(a), y'''(a), \dots, y^{(n)}(a), \dots)$$

とすれば, 数列 \mathbf{y} は 31.3 節における F の元である. 次の等式を満たす W の元からなる k 個の関数の組 $u_1(x), \dots, u_k(x)$ を考える:

$$\begin{aligned} (u_1(a), u_1'(a), \dots, u_1^{(k-1)}(a)) &= (1, 0, \dots, 0) \\ (u_2(a), u_2'(a), \dots, u_2^{(k-1)}(a)) &= (0, 1, \dots, 0) \\ &\vdots \\ (u_k(a), u_k'(a), \dots, u_k^{(k-1)}(a)) &= (0, 0, \dots, 1) \end{aligned}$$

つまり, 各関数 $u_i(x)$ は, $k-1$ 階までの点 a における微分係数のうち, i 階微分係数が 1 でそれ以外が 0 となる関数である.

【注意】 上式を満足するような関数 $u_1(x), \dots, u_k(x)$ の存在性は明らかではない. しかし天下りの言え, 最後に我々は解の公式としての W の基底を得ることから, それらの線形結合を上手く取ることにより上の条件を満た

す関数たちを構成することができる。ここでは $u_1(x), \dots, u_k(x)$ の存在性を構成的な立場から補足しておこう。例えば数列 $(b_n)_{n \in \mathbb{N}}$ が漸化式 (31.3.1) を満たすとき、関数 y を

$$y(x) := \sum_{n=0}^{\infty} \frac{b_n}{n!} (x-a)^n$$

とおこう。すると上式の右辺が項別微分可能であることが分かり (詳細は解析学の専門書に譲る)、上式の i 階導関数

$$y^{(i)}(x) = \sum_{n=i}^{\infty} \frac{n(n-1)\cdots(n-(i-1)) \cdot b_n}{n!} (x-a)^{n-i} = \sum_{n=i}^{\infty} \frac{b_n}{(n-i)!} (x-a)^{n-i} = \sum_{n=0}^{\infty} \frac{b_{n+i}}{n!} (x-a)^n$$

に $x = a$ を代入することで $(y(a), y'(a), y''(a), y'''(a), \dots) = (b_1, b_2, b_3, b_4, \dots)$ を得る。また、項別微分可能性から、この $y(x)$ が式 (32.3.1) を満たすことも分かる。実際、 b_n が式 (31.3.1) を満たすことから $b_{n+k} + a_{k-1}b_{n+k-1} + a_{k-2}b_{n+k-2} + \dots + a_1b_{n+1} + a_0b_n = 0$ であり、

$$\begin{aligned} y^{(k)}(x) + a_{k-1}y^{(k-1)}(x) + a_{k-2}y^{(k-2)}(x) + \dots + a_1y^{(1)}(x) + a_0y^{(0)}(x) \\ = \sum_{n=0}^{\infty} \frac{b_{n+k} + a_{k-1}b_{n+k-1} + a_{k-2}b_{n+k-2} + \dots + a_1b_{n+1} + a_0b_n}{n!} (x-a)^n = 0. \end{aligned}$$

次の二つの補題により、 $u_1(x), \dots, u_k(x)$ は W の基底となる。

補題 32.3.4. $u_1(x), \dots, u_n(x)$ は線形独立である。

Proof. 系 32.3.2 の証明で与えた写像 G は $u_i(x)$ を e_i にうつす。 e_i の線形独立性より $u_i(x)$ も線形独立である (命題 20.1.7). \square

補題 32.3.5. 各 $y(x) \in W$ に対して、 $y(x) = \sum_{i=1}^k y^{(i-1)}(a)u_i(x)$.

Proof. $y(x)$ と $\sum_{i=1}^k y^{(i-1)}(a)u_i(x)$ の点 a における $k-1$ 階まで微分係数が等しいことは明らかである。さらに系 32.3.2(1) から、これらは同じ関数であることが分かる。 \square

基底 $u_1(x), \dots, u_n(x)$ に関する $D: W \rightarrow W$ の表現行列を求めよう。そのためには、各 $D(u_1), \dots, D(u_n)$ を u_1, \dots, u_n の線形結合で表示した際の係数を見ればよい。

$$\begin{aligned} u_1 \text{ の点 } a \text{ における微分係数列: } & (1, 0, \dots, 0, -a_0, a_0a_{k-1}, \dots), \\ u_2 \text{ の点 } a \text{ における微分係数列: } & (0, 1, \dots, 0, -a_1, -a_0 + a_1a_{k-1}, \dots), \\ & \vdots \\ u_k \text{ の点 } a \text{ における微分係数列: } & (0, 0, \dots, 1, -a_{k-1}, a_{k-2} + a_{k-1}^2, \dots), \end{aligned}$$

であり、 $D(u_i)$ の微分係数の列は u_i のそれを左にずらしたもののゆえ

$$\begin{aligned} D(u_1) \text{ の点 } a \text{ における微分係数列: } & (0, \dots, 0, -a_0, a_0a_{k-1}, \dots), \\ D(u_2) \text{ の点 } a \text{ における微分係数列: } & (1, \dots, 0, -a_1, -a_0 + a_1a_{k-1}, \dots), \\ & \vdots \\ D(u_k) \text{ の点 } a \text{ における微分係数列: } & (0, \dots, 1, -a_{k-1}, a_{k-2} + a_{k-1}^2, \dots). \end{aligned}$$

したがって補題 32.3.5 より

$$D(u_1) = -a_0u_k, D(u_2) = u_1 - a_1u_k, \dots, D(u_i) = u_{i-1} - a_{i-1}u_k, \dots, D(u_k) = u_{k-1} - a_{k-1}u_k.$$

上式から、 $D|_W$ の表現行列は 31.3 節で求めた $S|_F$ のそれと一致し、式 (31.3.4) で与えられる行列 A となる。また、 $D|_W$ の特性多項式は式 (31.3.3) で与えた $\Phi(t)$ に等しい。

定義 32.3.6. 微分方程式 (32.3.1) の左辺において、各 $y^{(i)}(x)$ ($i = 0, \dots, k$) を t^i に置き換えた多項式 $\Phi(t) = t^k + a_{k-1}t^{k-1} + \dots + a_1t + a_0$ を、この微分方程式の特性多項式と呼ぶ。これまでの議論により、これは線形変換 $D|_W$ の特性多項式に等しい。 k 次方程式 $\Phi(t) = 0$ を、この微分方程式の特性方程式と呼ぶ。

命題 31.3.4 と同様にして、次が成り立つ:

命題 32.3.7. 微分方程式 (32.3.1) の特性多項式を Φ とする。 k 次方程式 $\Phi(t) = 0$ が互いに異なる k 個の解 $\lambda_1, \dots, \lambda_k$ を持つとすれば、この微分方程式一般解は $\sum_{i=1}^k C_i e^{\lambda_i x}$ である。

複素数値関数としての微分方程式 (32.3.1) の解も上と同様の表示が得られることが知られている。

よりみち (ロンスキー行列式)

本節で与えた関数の組 u_1, \dots, u_k の線形独立性は、これらの微分係数を並べた列の線形独立性から導かれた (補題 32.3.4)。ここで用いた議論をより一般的な場合に適用しよう。補題 32.3.4 の証明にあるように、 $k-1$ 階微分可能な関数の組 $f_1(x), f_2(x), \dots, f_k(x)$ が線形独立かどうかを示すには、これらの $k-1$ 階以下の微分係数を並べた k 次ベクトル

$$\begin{bmatrix} f_1(a) \\ f_1'(a) \\ f_1''(a) \\ \vdots \\ f_1^{(k-1)}(a) \end{bmatrix}, \begin{bmatrix} f_2(a) \\ f_2'(a) \\ f_2''(a) \\ \vdots \\ f_2^{(k-1)}(a) \end{bmatrix}, \dots, \begin{bmatrix} f_k(a) \\ f_k'(a) \\ f_k''(a) \\ \vdots \\ f_k^{(k-1)}(a) \end{bmatrix}$$

の線形独立性を示せば十分である。上の k 次ベクトルの線形独立性は、定理 17.3.6 により次で与える行列式の値が 0 でないことと同値である:

$$W(f_1, f_2, \dots, f_k)(a) = \begin{vmatrix} f_1(a) & f_2(a) & \dots & f_k(a) \\ f_1'(a) & f_2'(a) & \dots & f_k'(a) \\ \vdots & \vdots & \dots & \vdots \\ f_1^{(k-1)}(a) & f_2^{(k-1)}(a) & \dots & f_k^{(k-1)}(a) \end{vmatrix}.$$

行列式 $W(f_1, f_2, \dots, f_k)(a)$ を関数の組 f_1, f_2, \dots, f_k のロンスキー行列式 (**Wronskian**) という。

【補足】 組 f_1, f_2, \dots, f_k が線形独立であるにもかかわらず、それらの定義域上の各点 a において常にロンスキー行列式が消える場合もある。例えば、 $f(x) = x^2$, $g(x) = x|x|$ と定めれば、組 f, g は線形独立である。しかしながら

$$W(f, g)(a) = \begin{vmatrix} a^2 & a|a| \\ 2a & 2|a| \end{vmatrix} = 2a^2|a| - 2a^2|a| = 0.$$

第33章 不変部分空間と冪零部分空間(発展)

これまで、線形変換 $f: U \rightarrow U$ の表現行列を与える際に、上手く基底を選んでより複雑でない表現行列を得る方法について論じてきた。定理 29.2.1(1) の条件のもとでは、表現行列は対角行列に取れる。一方で、定理 29.2.1(1) の条件を満たさない線形写像に対して、どこまで表現行列を簡単にできるのだろうか。基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による f の表現行列を A としよう。 A は次で定義される行列であった。

$$A \text{ の第 } j \text{ 列} = \text{線形結合 } f(\mathbf{u}_j) = \sum_{i=1}^n a_{ij} \mathbf{u}_i \text{ に現れる係数を並べた列ベクトル.}$$

したがって、 A をより簡単な行列にせよという課題は、 $f(\mathbf{u}_j) = \sum_{i=1}^n a_{ij} \mathbf{u}_i$ に現れる係数の多くをいかに 0 にできるかという問題に帰着される。本章では、この問題への自然なアプローチとして不変部分空間および冪零部分空間の概念が導かれることを見る。

33.1 不変部分空間

U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$ とする。このとき、 j 番目のベクトル \mathbf{u}_j として $f \in \text{End}(U)$ の固有ベクトルを選ぶと、 $f(\mathbf{u}_j) = \lambda \mathbf{u}_j$ ゆえ $f(\mathbf{u}_j)$ は \mathbf{u}_j 自身のみによる線形結合で書けて、したがって表現行列の第 j 列は簡単な形 λe_j になる。ここで、 \mathbf{u}_j が固有ベクトルであるための条件が次のように書き換えられることに注意しよう。

$$f(\langle \mathbf{u}_j \rangle) \subset \langle \mathbf{u}_j \rangle.$$

そこで上の条件を一般化し、 U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の一部として次の条件を満たす組 $\mathbf{u}_\ell, \mathbf{u}_{\ell+1}, \dots, \mathbf{u}_{\ell+k}$ を考える：

$$f(\langle \mathbf{u}_\ell, \mathbf{u}_{\ell+1}, \dots, \mathbf{u}_{\ell+k} \rangle) \subset \langle \mathbf{u}_\ell, \mathbf{u}_{\ell+1}, \dots, \mathbf{u}_{\ell+k} \rangle.$$

すると、各 $f(\mathbf{u}_{\ell+j})$ ($j = 0, \dots, k$) は $\mathbf{u}_\ell, \mathbf{u}_{\ell+1}, \dots, \mathbf{u}_{\ell+k}$ のみによる線形結合で書けるゆえ、表現行列の第 ℓ 列から第 $\ell+k$ 列は比較的簡単な成分になる (実際、 $\ell, \ell+1, \dots, \ell+k$ 成分以外は 0 となる)。この着想を一般的な立場から述べようとすれば次の定義に至る：

定義 33.1.1. 線形変換 $f: U \rightarrow U$ に対して、 $f(W) \subset W$ を満たす U の部分空間 W のことを f の不変部分空間 (invariant subspace) という。このとき、 W は f -不変であるともいう。

U がいくつかの f -不変部分空間に分解できるならば、次の命題に述べるような表現行列が得られる。これまでの考察からこの命題の主張は明らかであるが、一応証明を述べておこう。

命題 33.1.2. U を有限次元線形空間とし、 $f \in \text{End}(U)$ とする。各 W_1, \dots, W_r が U の f -不変部分空間であり、 W_γ ($\gamma = 1, \dots, r$) の基底 $\mathbf{u}_{\gamma,1}, \dots, \mathbf{u}_{\gamma,n_\gamma}$ をそれぞれ一つ選び、これらをすべて集めたベクトルの組 $\mathcal{B} = \{ \mathbf{u}_{\gamma,i} \mid \gamma = 1, \dots, r, i = 1, \dots, n_\gamma \}$ が U の基底になるとする。このとき、基底 \mathcal{B} に関する f の表現行列 A は次の形になる：

$$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{bmatrix}.$$

ここで、各 A_γ ($\gamma = 1, \dots, r$) はサイズ $\dim W_\gamma = n_\gamma$ の正方行列である。

Proof. $\dim U = n$ とする. \mathcal{B} が U の基底となることから $n = \sum_{\gamma=1}^r n_\gamma$ である. そこで $n_0 = 0$ と置けば, 各 $j = 1, \dots, n$ は次のように書ける:

$$j = n_0 + n_1 + n_2 + \dots + n_{\gamma-1} + k. \quad (\text{ただし, } 1 \leq \gamma \leq r \text{ かつ } 1 \leq k \leq n_\gamma)$$

我々が示すべき事は A の第 j 列目の成分のうち第 $n_0 + \dots + n_{\gamma-1} + 1$ 成分から第 $n_0 + \dots + n_{\gamma-1} + n_\gamma$ 成分のほかがすべて 0 になることである. 基底 \mathcal{B} における $j = n_0 + \dots + n_{\gamma-1} + k$ 番目のベクトルは $\mathbf{u}_{\gamma,k} \in W_\gamma$ であり, W_γ が f -不変なことから $f(\mathbf{u}_{\gamma,k}) \in W_\gamma$ である. つまり, 基底 \mathcal{B} の線形結合で $f(\mathbf{u}_{\gamma,k})$ を表示した際に現れる係数のうち 0 でないものは $\mathbf{u}_{\gamma,1}, \dots, \mathbf{u}_{\gamma,n_\gamma}$ の係数に限られる. したがって, A の第 j 列目に現れる成分のうち 0 でないものは $\mathbf{u}_{\gamma,1}, \dots, \mathbf{u}_{\gamma,n_\gamma}$ に対応する成分, すなわち第 $n_0 + \dots + n_{\gamma-1} + 1$ 成分から第 $n_0 + \dots + n_{\gamma-1} + n_\gamma$ 成分に限られる. \square

備考 33.1.3. 上の命題における各 A_γ は, 基底 $\mathbf{u}_{\gamma,1}, \dots, \mathbf{u}_{\gamma,n_\gamma}$ に関する $f|_{W_\gamma} : W_\gamma \rightarrow W_\gamma$ の表現行列に等しい.

例 33.1.4. U が $f \in \text{End}(U)$ の固有ベクトルからなる基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を持つとき, 各 $W_j = \langle \mathbf{u}_j \rangle$ は U の f -不変部分空間である. これらに対して前命題を適用すれば, 各 A_j は $(1, 1)$ -行列であり, f の表現行列は対角行列となる.

33.2 冪零部分空間

本章の始めに提示した問題を前節とは別の視点から論じよう. 前節では, 望ましい基底の性質を導きだし, その性質のもとで表現行列が比較的簡単になることを見た. 本節ではこれとは逆の方向から検討する. すなわち, 表現行列が簡単な形をしていると仮定し, そのときに基底が満たすべき性質は何かを調べていく.

さて, 論ずべきことは U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を上手く取ることによって $T(\mathbf{u}_j) = \sum_{i=1}^n a_{ij} \mathbf{u}_i$ に現れる係数の多くをいかに 0 にできるかであった. そして, 最も都合が良い場合とは, \mathbf{u}_j が固有ベクトルであるとき, すなわち $f(\mathbf{u}_j)$ が \mathbf{u}_j 自身の線形結合で書ける場合であった. これが望めないとするならば, 次に考える最も単純な形は, \mathbf{u}_j 自身と, 基底をなす別のもう一つのベクトル \mathbf{u}_i との線形結合によって $f(\mathbf{u}_j)$ が表せる場合であろう. ここで更に踏み込んで, 各 $f(\mathbf{u}_j)$ が, \mathbf{u}_j と一つ隣のベクトル \mathbf{u}_{j-1} の線形結合で書ける場合, すなわち

$$f(\mathbf{u}_j) = s_{j-1} \mathbf{u}_{j-1} + r_j \mathbf{u}_j \quad (j = 2, \dots, n)$$

となる場合を考えよう. また, 簡単のために f が体 \mathbb{C} 上の線形空間における線形変換である場合を考えるとすれば, 代数学の基本定理により f の固有ベクトルは必ず存在する. ゆえに基底の並びにおける最初の \mathbf{u}_1 は固有ベクトルである (つまり $f(\mathbf{u}_1) = r_1 \mathbf{u}_1$) としてよい. このとき f の表現行列は次のようになる.

$$A = \begin{bmatrix} r_1 & s_1 & & & \\ & r_2 & s_2 & & \\ & & \ddots & \ddots & \\ & & & r_{n-1} & s_{n-1} \\ & & & & r_n \end{bmatrix}.$$

ここで, $s_{j-1} = 0$ ならば \mathbf{u}_j は固有ベクトルである. 一方 $s_{j-1} \neq 0$ の場合は \mathbf{u}_j の代わりに $\frac{1}{s_{j-1}} \mathbf{u}_j$ を基底として取れば,

$$f\left(\frac{1}{s_{j-1}} \mathbf{u}_j\right) = \frac{1}{s_{j-1}} f(\mathbf{u}_j) = \frac{1}{s_{j-1}} (s_{j-1} \mathbf{u}_{j-1} + r_j \mathbf{u}_j) = \mathbf{u}_{j-1} + r_j \left(\frac{1}{s_{j-1}} \mathbf{u}_j\right)$$

であるから $s_{j-1} = 1$ の場合が本質的である¹. また, A は上三角行列であるから, 各 r_j は A の固有値 (つまり f の固有値) になっている (例 28.4.7). そこで, 列 r_1, \dots, r_n が各固有値ごとに行儀よく並んで

¹このとき, 次に並んでいるベクトル \mathbf{u}_{j+1} に関する式 $f(\mathbf{u}_{j+1}) = s_j \mathbf{u}_j + r_{j+1} \mathbf{u}_{j+1}$ について, いま \mathbf{u}_j を置き換えたから s_j を $s_{j-1} s_j$ に置き換える必要がある. 更に同様の議論を適用し, 必要ならば \mathbf{u}_{j+1} を置き換えることで, $s_j = 0$ または $s_j = 1$ とできる. これを順次繰り返せば各 $s_{j-1}, s_j, \dots, s_{n-1}$ を 0 または 1 に置き換えられる.

いるような更に特別な場合を検討しよう. すなわち, f の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とし, U の基底 $\{\mathbf{u}_{\lambda_k, j} \mid k = 1, \dots, r, j = 1, \dots, n_k\}$ に関する表現行列 A が次のようになる場合である:

$$A = \begin{bmatrix} A_{\lambda_1} & & & \\ & A_{\lambda_2} & & \\ & & \ddots & \\ & & & A_{\lambda_r} \end{bmatrix}, \quad A_{\lambda_k} = \begin{bmatrix} \lambda_k & s_{\lambda_k, 1} & & \\ & \lambda_k & s_{\lambda_k, 2} & \\ & & \ddots & \ddots \\ & & & \lambda_k & s_{\lambda_k, n_k-1} \\ & & & & \lambda_k \end{bmatrix}. \quad (33.2.1)$$

ここで各 A_{λ_k} は n_k 次正方行列であり, $s_{\lambda_k, j} = 0$ または $s_{\lambda_k, j} = 1$ である.

【補足】 各固有値 λ_k の固有ベクトルは少なくとも一つは存在する. そこで A_{λ_k} の第 1 列目が固有ベクトルに対応する列となるような, 更に特別な場合のみをここでは考えている. このとき, A_{λ_k} における (1, 1)-成分の一つ上の成分は 0 であり, A は上式の形になる.

このとき, $\mathbf{u}_{\lambda_k, j}$ ($j = 2, \dots, n_k$) が固有ベクトルでなければ $s_{\lambda_k, j-1} = 1$ ゆえ $f(\mathbf{u}_{\lambda_k, j}) = \mathbf{u}_{\lambda_k, j-1} + \lambda_k \mathbf{u}_{\lambda_k, j}$ である. ゆえに,

$$\mathbf{u}_{\lambda_k, j-1} = f(\mathbf{u}_{\lambda_k, j}) - \lambda_k \mathbf{u}_{\lambda_k, j} = (f - \lambda_k I)(\mathbf{u}_{\lambda_k, j}).$$

つまり, $\mathbf{u}_{\lambda_k, j}$ に線形写像 $f - \lambda_k I$ を繰り返すことで $\mathbf{u}_{\lambda_k, j-1}, \mathbf{u}_{\lambda_k, j-2}, \dots$ が次々と得られ, これを続けると最後には λ_k に関する固有ベクトル $\mathbf{u}_{\lambda_k, j-\ell}$ を得る². また,

$$(f - \lambda_k I)(\mathbf{u}_{\lambda_k, j-\ell}) = f(\mathbf{u}_{\lambda_k, j-\ell}) - \lambda_k \mathbf{u}_{\lambda_k, j-\ell} = \lambda_k \mathbf{u}_{\lambda_k, j-\ell} - \lambda_k \mathbf{u}_{\lambda_k, j-\ell} = \mathbf{0}$$

であるから

$$\mathbf{u}_{\lambda_k, j} \xrightarrow{f - \lambda_k I} \mathbf{u}_{\lambda_k, j-1} \xrightarrow{f - \lambda_k I} \mathbf{u}_{\lambda_k, j-2} \xrightarrow{f - \lambda_k I} \dots \xrightarrow{f - \lambda_k I} \mathbf{u}_{\lambda_k, j-\ell} \xrightarrow{f - \lambda_k I} \mathbf{0}.$$

すなわち,

$$\mathbf{u}_{\lambda_k, j} \in \text{Ker}(f - \lambda_k I)^{\ell+1}. \quad (33.2.2)$$

こうして我々は冪零部分空間の概念に至る:

定義 33.2.1. 有限次元線形空間 U 上の線形変換 $g: U \rightarrow U$ に対して, 部分空間の増大列 $\text{Ker } g \subset \text{Ker } g^2 \subset \text{Ker } g^3 \subset \dots$ は下の補題により十分大きい N について, $\text{Ker } g^N = \text{Ker } g^{N+1} = \text{Ker } g^{N+2} = \dots$ を満たす. このとき, $\text{Ker } g^N$ を g の冪零部分空間と呼ぼう³.

定義 33.2.2. 線形変換 $f: U \rightarrow U$ の固有値 λ に対して, 線形変換 $(f - \lambda I): U \rightarrow U$ の冪零部分空間を, f の固有値 λ に関する一般固有空間 (generalized eigenspace) あるいは広義固有空間という. 本書では, これを記号 $\widetilde{W}(\lambda, f)$ で表す. 更に, 正方行列 A について, $\widetilde{W}(\lambda, T_A)$ を $\widetilde{W}(\lambda, A)$ とも表す. 零ベクトルでない $\widetilde{W}(\lambda, f)$ の元を, f の固有値 λ に関する一般固有ベクトルあるいは広義固有ベクトルと呼ぶ. また, 零ベクトルでない $\widetilde{W}(\lambda, A)$ の元を A の固有値 λ に関する一般固有ベクトルと呼ぶ.

【補足】 (1) $\widetilde{W}(\lambda, f) = \text{Ker}(\lambda I - f) = \text{Ker}(f - \lambda I)$ より, f の固有値 λ に関する固有空間は, 固有値 λ に関する一般固有空間の部分空間である.

(2) n 次正方行列 A について, $\mathbf{x} \in \mathbb{R}^n$ が A の固有値 λ に関する一般固有ベクトルであることは, 十分大きな $N \in \mathbb{N}$ について $(A - \lambda E)^N \mathbf{x} = \mathbf{0}$ が成り立つことにほかならない ($(T_A - \lambda T_E)^N(\mathbf{x}) = (A - \lambda E)^N \mathbf{x}$ に注意する).

以下では $g \in \text{End}(U)$ を一般の線形変換として扱うものの, $g = f - \lambda I$ のことと考えて読むと何を論じているのかイメージが湧くことと思う.

補題 33.2.3. 有限次元線形空間 U 上の線形変換 $g: U \rightarrow U$ において次が成り立つ.

$$(1) \text{Ker } g \subset \text{Ker } g^2 \subset \text{Ker } g^3 \subset \dots$$

²少なくとも $\mathbf{u}_{\lambda_k, 1}$ は λ_k に関する固有ベクトルであるから, $f - \lambda_k I$ を j 回ほどこす間に λ_k に関する固有ベクトルが必ず得られる.

³この名称は本書のみで通じるものである. $\text{Ker } g^{\dim U}$ と記せばよいから, この空間に一般的な名称は与えられていない.

(2) $\text{Ker } g^n = \text{Ker } g^{n+1}$ ならば, $\text{Ker } g^n = \text{Ker } g^{n+1} = \text{Ker } g^{n+2} = \text{Ker } g^{n+3} = \dots$.

(3) $\dim U = N$ ならば, $\text{Ker } g^N = \text{Ker } g^{N+1} = \text{Ker } g^{N+2} = \dots$.

Proof. (1): $\mathbf{u} \in \text{Ker } g^n$ とすれば, $g^n(\mathbf{u}) = \mathbf{0}$. よって, $g^{n+1}(\mathbf{u}) = g(g^n(\mathbf{u})) = g(\mathbf{0}) = \mathbf{0}$ ゆえ $\mathbf{u} \in \text{Ker } g^{n+1}$.

(2): $\text{Ker } g^{n+1} = \text{Ker } g^{n+2}$ を示せば,あとは帰納的に各 $\text{Ker } g^{n+m}$ がすべて一致することが分かる. $\text{Ker } g^{n+1} \subset \text{Ker } g^{n+2}$ は (1) で示したゆえ $\text{Ker } g^{n+2} \subset \text{Ker } g^{n+1}$ を示そう. 各 $\mathbf{u} \in \text{Ker } g^{n+2}$ について, $\mathbf{0} = g^{n+2}(\mathbf{u}) = g^{n+1}(g(\mathbf{u}))$ より $g(\mathbf{u}) \in \text{Ker } g^{n+1} = \text{Ker } g^n$. よって, $g(\mathbf{u}) \in \text{Ker } g^n$ ゆえ $g^n(g(\mathbf{u})) = \mathbf{0}$. つまり, $\mathbf{u} \in \text{Ker } g^{n+1}$ である.

(3): g が単射ならば $\{\mathbf{0}\} = \text{Ker } g = \text{Ker } g^2 = \text{Ker } g^3 = \dots$ である. そこで, 単射でない場合を考えよう. このとき $\dim \text{Ker } g \geq 1$ である. $\text{Ker } g^{n+1}$ が $\text{Ker } g^n$ よりも真に大きくなる時, 命題 22.4.1 よりこれらの次元も真に大きくなる. 仮に, すべての $n = 1, \dots, N$ について $\text{Ker } g^{n+1}$ が $\text{Ker } g^n$ よりも真に大きくなるとすれば,

$$\dim \text{Ker } g^{N+1} \geq \dim \text{Ker } g^N + 1 \geq (\dim \text{Ker } g^{N-1} + 1) + 1 \geq \dots \geq \dim \text{Ker } g + N \geq 1 + N$$

となり $\dim U = N$ に矛盾してしまう. したがって, $n = 1, \dots, N$ のいずれかにおいて $\text{Ker } g^{n+1} = \text{Ker } g^n$ となる必要があり, それ以降は (2) よりすべて一致する. \square

【補足】上の補題 (3) において, 多くの場合は $\dim U$ よりも小さい K について, g^K 以降の核が等しくなる.

これまでの議論をまとめると, 式 (33.2.1) のような形の表現行列を得るためには, 式 (33.2.2) により, 少なくとも一般固有ベクトルからなる U の基底が取れる必要があることが分かった. 実は, 体 \mathbb{C} 上の線形空間における任意の線形変換について, これが可能である (定理 34.1.1).

一般固有空間が不変部分空間であることを確認しておこう.

補題 33.2.4. $f, g : U \rightarrow U$ を可換な線形変換とする (すなわち $g \circ f = f \circ g$). このとき, $W = \text{Ker } g$ は f -不変部分空間である.

Proof. 各 $\mathbf{u} \in W$ について, $f(\mathbf{u}) \in W$ を示したい. そのためには $g(f(\mathbf{u})) = \mathbf{0}$ を言えばよい. g と f が可換になること, および $g(\mathbf{u}) = \mathbf{0}$ から,

$$g(f(\mathbf{u})) = g \circ f(\mathbf{u}) = f \circ g(\mathbf{u}) = f(g(\mathbf{u})) = f(\mathbf{0}) = \mathbf{0}.$$

\square

線形変換 g は自身の冪 g^n と可換である. また, 線形変換 f および $(f - \lambda I)^n, f - \gamma I$ はそれぞれ可換である (命題 21.5.6). これらの事実と先の補題から次を得る.

系 33.2.5. $f, g : U \rightarrow U$ を線形変換とする. 各 $\lambda, \gamma \in \mathbb{K}$ について次が成り立つ.

(1) $g : U \rightarrow U$ の冪零部分空間 W は g -不変部分空間である.

(2) $f - \lambda I$ の冪零部分空間 \widetilde{W} は f -不変である. すなわち, f の固有値 λ に関する一般固有空間 $\widetilde{W}(\lambda, f)$ は f -不変である.

(3) $f - \lambda I$ の冪零部分空間 \widetilde{W} は $(f - \gamma I)$ -不変である.

線形写像を表現行列によってデータ化するという観点からは, 次が成り立つ. これは命題 28.6.1 の一般化であり, 証明の戦略も変わらない.

命題 33.2.6. U を n 次元線形空間とし $f : U \rightarrow U$ を線形変換とする. 更に, U の基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ による f の表現行列を A とすれば次が成り立つ.

- (1) 各 \mathbf{v}_i を $\mathbf{e}_i \in \mathbb{R}^n$ に対応させる線形同型を $F : U \rightarrow \mathbb{R}^n$ とする. このとき, $\mathbf{u} \in U$ が f の固有値 λ に関する一般固有ベクトルであることと, $F(\mathbf{u})$ が A の固有値 λ に関する一般固有ベクトルであることは同値である.
- (2) λ を f の固有値 (および A の固有値) とする. 列ベクトルの組 $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ について, $\mathbf{y}_i = [\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x}_i$ ($i = 1, \dots, k$) とすれば, $\mathbf{y}_1, \dots, \mathbf{y}_k$ が $\widetilde{W}(\lambda, f)$ の基底であることと $\mathbf{x}_1, \dots, \mathbf{x}_k$ が $\widetilde{W}(\lambda, A)$ の基底であることは同値である.
- (3) λ を f の固有値 (および A の固有値) とすると, $\dim \widetilde{W}(\lambda, f) = \dim \widetilde{W}(\lambda, A)$.

Proof. (1): 各 $k \in \mathbb{N}$ について $[(f - \lambda I)^k(\mathbf{u}) = \mathbf{0} \Leftrightarrow (A - \lambda E)^k F(\mathbf{u}) = \mathbf{0}]$ を示せば, これは (1) を意味する. さて, $g = (f - \lambda I)^k$ の表現行列は系 26.2.5 より $G = (A - \lambda E)^k$ である. また, 命題 26.4.2(2) の証明によれば, $\text{Ker } g$ と $\text{Ker } T_G$ は F によって 1 対 1 に対応し, これらは線形同型である. つまり, $[\mathbf{u} \in \text{Ker } g \Leftrightarrow F(\mathbf{u}) \in \text{Ker } T_G]$ であり, これが示したかったことである.

(2) および (3): 十分大きな $N \in \mathbb{N}$ を取れば⁴, $\widetilde{W}(\lambda, f) = \text{Ker}(f - \lambda I)^N$ かつ $\widetilde{W}(\lambda, A) = \text{Ker}(T_A - \lambda T_E)^N$ が成り立つ. そこで, $k := N$ として (1) で与えた対応 $F|_{\widetilde{W}(\lambda, f)} : \widetilde{W}(\lambda, f) \rightarrow \widetilde{W}(\lambda, A)$ ($[\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x} \mapsto \mathbf{x}$) を考える. この $F|_{\widetilde{W}(\lambda, f)}$ が線形同型であることから (2) および (3) は直ちに得られる. \square

33.3 微分作用素とシフト作用素の一般固有ベクトル

ここで関数空間の微分作用素と数列空間のシフト作用素について, どのようなベクトルが一般固有ベクトルとなるか見ておこう. ここで述べる例は, 線形漸化式の一般項や線形常微分方程式の一般解の構造を理解するうえで助けとなるものである (詳しくは 34.3 節を見よ).

まず, 一般論として次の補題を用意する. これは, 例 17.2.3 で述べた手法の一般化に他ならない.

補題 33.3.1. 線形変換 $g : U \rightarrow U$ および $\mathbf{u} \in U$ について, $g^n(\mathbf{u}) \neq \mathbf{0}$ かつ $g^{n+1}(\mathbf{u}) = \mathbf{0}$ ならば, $n + 1$ 個のベクトルからなる組 $\mathbf{u}, g(\mathbf{u}), g^2(\mathbf{u}), \dots, g^n(\mathbf{u})$ は線形独立である.

Proof. 各 $k = 0, \dots, n$ について, $k + 1$ 個の組 $g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-k}(\mathbf{u})$ が線形独立であることを k に関する帰納法で示そう. $k \leq n$ を満たす自然数 k について, k 個の組 $g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-(k-1)}(\mathbf{u})$ が線形独立であると仮定し, $k + 1$ 個の組 $g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-k}(\mathbf{u})$ の線形独立性を示す. 線形関係 $\sum_{i=0}^k r_i g^{n-i}(\mathbf{u}) = \mathbf{0}$ を仮定し, この両辺に g をほどこせば,

$$\begin{aligned} r_0 g^{n-0+1}(\mathbf{u}) + r_1 g^{n-1+1}(\mathbf{u}) + r_2 g^{n-2+1}(\mathbf{u}) + \dots + r_k g^{n-k+1}(\mathbf{u}) &= \mathbf{0} \\ r_1 g^n(\mathbf{u}) + r_2 g^{n-1}(\mathbf{u}) + \dots + r_k g^{n-(k-1)}(\mathbf{u}) &= \mathbf{0}. \end{aligned}$$

$g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-(k-1)}(\mathbf{u})$ は線形独立であったゆえ $r_1 = r_2 = \dots = r_k = 0$. ゆえに $r_0 g^n(\mathbf{u}) = \mathbf{0}$ であり, $g^n(\mathbf{u}) \neq \mathbf{0}$ より $r_0 = 0$. 以上より $k + 1$ 個の組 $g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-k}(\mathbf{u})$ は線形独立である. \square

次の例において $\mathbb{K} = \mathbb{C}$ の場合は複素関数についての知識が必要である. ここでは, 複素数値関数においても実数値関数の場合と同様の微分公式が満たされることを既知の事実であるとして話を進める.

例 33.3.2. $D : C^\infty(\mathbb{K}) \rightarrow C^\infty(\mathbb{K})$ を微分作用素 $D(y) = y'$ とし, また $\lambda \in \mathbb{K}$ とする.

- (1) 非負整数 n について $x^n e^{\lambda x} \in \text{Ker}(D - \lambda I)^{n+1}$ かつ $x^n e^{\lambda x} \notin \text{Ker}(D - \lambda I)^n$.
- (2) $e^{\lambda x}, x e^{\lambda x}, x^2 e^{\lambda x}, \dots, x^n e^{\lambda x}$ は線形独立である.

⁴実際には $N := n$ とすればよい (補題 33.2.3(3)).

Proof. (1): n に関する帰納法で示す. $n = 0$ について $e^{\lambda x}$ は固有値 λ の固有ベクトルゆえ $e^{\lambda x} \in \text{Ker}(D - \lambda I)$. 次に, $x^{n-1}e^{\lambda x} \in \text{Ker}(D - \lambda I)^n$ かつ $x^{n-1}e^{\lambda x} \notin \text{Ker}(D - \lambda I)^{n-1}$ を仮定して, $x^n e^{\lambda x} \in \text{Ker}(D - \lambda I)^{n+1}$ および $x^n e^{\lambda x} \notin \text{Ker}(D - \lambda I)^n$ を示そう.

$$\begin{aligned}(D - \lambda I)(x^n e^{\lambda x}) &= (x^n e^{\lambda x})' - \lambda x^n e^{\lambda x} = (nx^{n-1}e^{\lambda x} + x^n \lambda e^{\lambda x}) - \lambda x^n e^{\lambda x} \\ &= nx^{n-1}e^{\lambda x} \in \text{Ker}(D - \lambda I)^n.\end{aligned}$$

ゆえに $x^n e^{\lambda x} \in \text{Ker}(D - \lambda I)^{n+1}$ である. また, $nx^{n-1}e^{\lambda x} \notin \text{Ker}(D - \lambda I)^{n-1}$ ゆえ $x^n e^{\lambda x} \notin \text{Ker}(D - \lambda I)^n$ である⁵.

(2): $g = D - \lambda I$ および $\mathbf{u} = x^n e^{\lambda x}$ について前補題を適用すると, $N + 1$ 個の組

$$x^n e^{\lambda x}, nx^{n-1}e^{\lambda x}, n(n-1)x^{n-2}e^{\lambda x}, \dots, n!e^{\lambda x}$$

は線形独立である. ゆえに, これらにスカラー倍をほどこした組 $x^n e^{\lambda x}, x^{n-1}e^{\lambda x}, x^{n-2}e^{\lambda x}, \dots, e^{\lambda x}$ も線形独立である. \square

別解. (2) は, 点 $a = 0$ におけるロンスキー行列式が消えないことから導かれる. 実際, e^x のテイラー展開を用いれば $x^k e^{\lambda x} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} x^{n+k}$ と表示できる. ゆえに関数 $f(x) = x^k e^{\lambda x}$ は $i = 0, \dots, k-1$ について $f^{(i)}(0) = 0$ であり, $f^{(k)}(0) = k!$ となる. したがって組 $e^{\lambda x}, x e^{\lambda x}, x^2 e^{\lambda x}, \dots, x^n e^{\lambda x}$ に関する点 $a = 0$ におけるロンスキー行列は, 対角成分に 0 を持たない下三角行列であり, ゆえに可逆である. \square

例 33.3.3. $S : \mathbb{K}^{\mathbb{N}} \rightarrow \mathbb{K}^{\mathbb{N}}$ をシフト作用素 $S((x_n)_{n \in \mathbb{N}}) := (x_{n+1})_{n \in \mathbb{N}}$ とする. $\lambda \neq 0$ および非負整数 N について, 数列 $\mathbf{x}_N^\lambda \in \mathbb{R}^{\mathbb{N}}$ を次で定める.

$$\mathbf{x}_N^\lambda := (1, 2^N \lambda, 3^N \lambda^2, 4^N \lambda^3, \dots) \quad (\text{つまり } \mathbf{x}_N^\lambda = (n^N \lambda^{n-1})_{n \in \mathbb{N}}).$$

また, $\lambda = 0$ の場合については, 第 $N + 1$ 座標が 1 でそれ以外の座標がすべて 0 の数列を \mathbf{x}_N^0 と定める.

(1) $\mathbf{x}_N^\lambda \in \text{Ker}(S - \lambda I)^{N+1}$ かつ $\mathbf{x}_N^\lambda \notin \text{Ker}(S - \lambda I)^N$.

(2) $\mathbf{x}_0^\lambda, \mathbf{x}_1^\lambda, \dots, \mathbf{x}_N^\lambda$ は線形独立である.

Proof. $\lambda = 0$ の場合は明らかゆえ証明は省略する. $\lambda \neq 0$ とし, 煩雑にならぬよう \mathbf{x}_N^λ を \mathbf{x}_N と略そう.

(1): N に関する帰納法で示す. $N = 0$ の場合, \mathbf{x}_0 は初項 1 公比 λ の等比数列であり, これは S の λ に関する固有ベクトルであるから $\mathbf{x}_0 \in \text{Ker}(S - \lambda I)$ である. 次に各 $k = 0, \dots, N - 1$ について $\mathbf{x}_k \in \text{Ker}(S - \lambda I)^{k+1}$ かつ $\mathbf{x}_k \notin \text{Ker}(S - \lambda I)^k$ を仮定して, $\mathbf{x}_N \in \text{Ker}(S - \lambda I)^{N+1}$ および $\mathbf{x}_N \notin \text{Ker}(S - \lambda I)^N$ を示そう.

$$\begin{aligned}(S - \lambda I)(\mathbf{x}_N) &= (S - \lambda I)\left((n^N \lambda^{n-1})_{n \in \mathbb{N}}\right) = \left((n+1)^N \lambda^n\right)_{n \in \mathbb{N}} - \lambda \left(n^N \lambda^{n-1}\right)_{n \in \mathbb{N}} \\ &= \left(\left((n+1)^N - n^N\right) \lambda^n\right)_{n \in \mathbb{N}} = \lambda \left(\left((n+1)^N - n^N\right) \lambda^{n-1}\right)_{n \in \mathbb{N}}.\end{aligned}$$

ここで, $(n+1)^N - n^N$ は n に関する $N - 1$ 次の多項式であるから $(n+1)^N - n^N = \sum_{k=0}^{N-1} a_k n^k$ と書ける (ここで $a_{N-1} \neq 0$). ゆえに

$$\begin{aligned}\lambda \left(\left((n+1)^N - n^N\right) \lambda^{n-1}\right)_{n \in \mathbb{N}} &= \lambda \left(\left(\sum_{k=0}^{N-1} a_k n^k\right) \lambda^{n-1}\right)_{n \in \mathbb{N}} = \lambda \left(\sum_{k=0}^{N-1} a_k n^k \lambda^{n-1}\right)_{n \in \mathbb{N}} \\ &= \lambda \sum_{k=0}^{N-1} (a_k n^k \lambda^{n-1})_{n \in \mathbb{N}} = \lambda \sum_{k=0}^{N-1} a_k \mathbf{x}_k \in \text{Ker}(S - \lambda I)^N.\end{aligned}$$

⁵何故なら, 仮に $\mathbf{u} = x^n e^{\lambda x} \in \text{Ker}(D - \lambda I)^n$ と仮定すれば $(D - \lambda I)(\mathbf{u}) \in \text{Ker}(D - \lambda I)^{n-1}$ となり, これは $(D - \lambda I)(\mathbf{u}) = nx^{n-1}e^{\lambda x} \notin \text{Ker}(D - \lambda I)^{n-1}$ に矛盾してしまう.

以上より, $(S - \lambda I)(\mathbf{x}_N) \in \text{Ker}(S - \lambda I)^N$ ゆえ $\mathbf{x}_N \in \text{Ker}(S - \lambda I)^{N+1}$. また, 帰納法の仮定により各 $k = 0, \dots, N-2$ について $(S - \lambda I)^{N-1}(\mathbf{x}_k) = \mathbf{0}$ および $(S - \lambda I)^{N-1}(\mathbf{x}_{N-1}) \neq \mathbf{0}$ であったから,

$$\begin{aligned} (S - \lambda I)^N(\mathbf{x}_N) &= (S - \lambda I)^{N-1}((S - \lambda I)(\mathbf{x}_N)) = (S - \lambda I)^{N-1} \left(\lambda \sum_{k=0}^{N-1} a_k \mathbf{x}_k \right) \\ &= \lambda \sum_{k=0}^{N-1} a_k (S - \lambda I)^{N-1}(\mathbf{x}_k) = \lambda a_{N-1} (S - \lambda I)^{N-1}(\mathbf{x}_{N-1}) \neq \mathbf{0}. \end{aligned}$$

ゆえに $\mathbf{x}_N \notin \text{Ker}(S - \lambda I)^N$.

(2): $g = S - \lambda I$ について前補題を適用すれば, $\mathbf{x}_N, g(\mathbf{x}_N), \dots, g^N(\mathbf{x}_N)$ は線形独立である. また, (1) の証明で行った計算によれば

$$\langle \mathbf{x}_N, g(\mathbf{x}_N), \dots, g^N(\mathbf{x}_N) \rangle \subset \langle \mathbf{x}_N, \mathbf{x}_{N-1}, \dots, \mathbf{x}_0 \rangle$$

である. したがって

$$N + 1 = \dim \langle \mathbf{x}_N, g(\mathbf{x}_N), \dots, g^N(\mathbf{x}_N) \rangle \leq \dim \langle \mathbf{x}_N, \mathbf{x}_{N-1}, \dots, \mathbf{x}_0 \rangle \leq N + 1$$

より $\dim \langle \mathbf{x}_N, \mathbf{x}_{N-1}, \dots, \mathbf{x}_0 \rangle = N + 1$. これと命題 22.3.3(3) を合わせて, $\mathbf{x}_N, \mathbf{x}_{N-1}, \dots, \mathbf{x}_0$ の線形独立性を得る. \square

別解. $\mathbf{x}_0^\lambda, \mathbf{x}_1^\lambda, \dots, \mathbf{x}_N^\lambda$ の第 $N+1$ 項までを並べたベクトルの列からなる $N+1$ 次正方行列 $[\mathbf{x}_0^\lambda|_{N+1}, \mathbf{x}_1^\lambda|_{N+1}, \dots, \mathbf{x}_N^\lambda|_{N+1}]$ の可逆性からも (2) は導かれる. 実際, $\lambda = 0$ の場合はこの行列は単位行列であり, そうでない場合は各 $\mathbf{x}_k^\lambda|_{N+1}$ を列ベクトルとみなして行列式をとれば, ヴァンデルモンドの行列式 (定理 11.2.2) より

$$\begin{aligned} \det \left(\mathbf{x}_0^\lambda|_{N+1}, \mathbf{x}_1^\lambda|_{N+1}, \dots, \mathbf{x}_N^\lambda|_{N+1} \right) &= \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \lambda & 2\lambda & 2^2\lambda & \cdots & 2^N\lambda \\ \lambda^2 & 3\lambda^2 & 3^2\lambda^2 & \cdots & 3^N\lambda^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \lambda^N & (N+1)\lambda^N & (N+1)^2\lambda^N & \cdots & (N+1)^N\lambda^N \end{vmatrix} \\ &= \lambda^{1+2+\cdots+N} \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2^2 & \cdots & 2^N \\ 1 & 3 & 3^2 & \cdots & 3^N \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & (N+1) & (N+1)^2 & \cdots & (N+1)^N \end{vmatrix} = \lambda^{1+2+\cdots+N} \prod_{1 \leq i < j \leq N+1} (j-i) \neq 0. \end{aligned}$$

\square

33.4 冪零部分空間と安定部分空間への分解

冪零部分空間と対になる概念として, 安定部分空間が定義される:

定義 33.4.1. 有限次元線形空間 U 上の線形変換 $g: U \rightarrow U$ に対して, 部分空間の減少列 $\text{Im } g \supset \text{Im } g^2 \supset \text{Im } g^3 \supset \dots$ は次の補題により十分大きい N について $\text{Im } g^N = \text{Im } g^{N+1} = \text{Im } g^{N+2} = \dots$ を満たす. この部分空間 $\text{Im } g^N$ を g の安定部分空間 と呼ぼう⁶.

補題 33.4.2. 線形空間 U 上の線形変換 $g: U \rightarrow U$ において次が成り立つ.

- (1) $\text{Im } g \supset \text{Im } g^2 \supset \text{Im } g^3 \supset \dots$.
- (2) $\text{Im } g^n = \text{Im } g^{n+1}$ ならば, $\text{Im } g^n = \text{Im } g^{n+1} = \text{Im } g^{n+2} = \dots$.
- (3) $\dim U = N$ ならば, $\text{Im } g^N = \text{Im } g^{N+1} = \text{Im } g^{N+2} = \dots$.

⁶これも本書でしか通じない名称である.

Proof. (1): 各 $g^{n+1}(\mathbf{u}) \in \text{Im } g^{n+1}$ について, $g^{n+1}(\mathbf{u}) = g^n(g(\mathbf{u})) \in \text{Im } g^n$. よって, $\text{Im } g^{n+1} \subset \text{Im } g^n$.

(2): $\text{Im } g^{n+1} = \text{Im } g^{n+2}$ さえ示せば, あとは帰納的に各 $\text{Im } g^{n+m}$ (ただし $m \in \mathbb{N}$) がすべて一致することが分かる. $\text{Im } g^{n+1} \supset \text{Im } g^{n+2}$ は (1) で示したゆえ $\text{Im } g^{n+1} \subset \text{Im } g^{n+2}$ を示そう. 各 $g^{n+1}(\mathbf{u}) \in \text{Im } g^{n+1}$ について, $g^{n+1}(\mathbf{u}) = g(g^n(\mathbf{u}))$ である. また, $g^n(\mathbf{u}) \in \text{Im } g^n = \text{Im } g^{n+1}$ ゆえ, ある $\mathbf{v} \in U$ を用いて $g^n(\mathbf{u}) = g^{n+1}(\mathbf{v})$ と書ける. よって, $g^{n+1}(\mathbf{u}) = g(g^n(\mathbf{u})) = g(g^{n+1}(\mathbf{v})) = g^{n+2}(\mathbf{v}) \in \text{Im } g^{n+2}$.

(3): 補題 33.2.3(3) と類似の議論を部分空間の減少列に対して適用すればよい. g が全射ならば $U = \text{Im } g = \text{Im } g^2 = \text{Im } g^3 = \dots$ である. そこで, 全射でない場合を考えよう. このとき練習 22.4.7(2) より $\dim U \geq \dim \text{Im } g + 1$ である. $\text{Im } g^{n+1}$ が $\text{Im } g^n$ よりも真に小さくなる時, 命題 22.4.1 よりそれらの次元も真に小さくなる. 仮に, すべての $n = 1, \dots, N$ について $\text{Im } g^{n+1}$ が $\text{Im } g^n$ よりも真に小さくなるとすれば,

$$\dim U \geq \dim \text{Im } g + 1 \geq (\dim \text{Im } g^2 + 1) + 1 \geq \dots \geq \dim \text{Im } g^{N+1} + N + 1 \geq N + 1$$

となり $\dim U = N$ に矛盾してしまう. したがって, $n = 1, \dots, N$ のいずれかにおいて $\text{Im } g^{n+1} = \text{Im } g^n$ となる必要があり, それ以降は (2) よりすべて一致する. \square

【補足】 上の (3) の証明, および補題 33.2.3(3) の証明において, g が全射 (あるいは単射) であるか否かの場合分けは不要である. 実際, $g^0 = \text{id}_U$ に注意し, すべての $n = 0, \dots, N$ に対して部分空間の減少 (増大) 列が真に小さく (大きく) なる時に矛盾が生じることを示せばよい.

任意の線形変換 $g : U \rightarrow U$ について, U の各元は冪零部分空間の元と安定部分空間の元に分解される. これを示すために, いくつかの事実について確認しよう.

命題 33.4.3. 有限次元線形空間上の線形変換 $g : U \rightarrow U$ の安定部分空間 V において, 次が成り立つ.

(1) $g(V) = V$. とくに V は g -不変部分空間である.

(2) $g|_V : V \rightarrow V$ は線形同型である.

Proof. 自然数 N を十分大きく取り, $V = \text{Im } g^N = \text{Im } g^{N+1} = \dots$ であるとする.

(1): $g(V) = g(\text{Im } g^N) = \text{Im } g^{N+1} = V$.

(2): (1) より $g|_V : V \rightarrow V$ は全射であり, ゆえに命題 22.4.8 より同型である. \square

補題 33.2.4 および系 33.2.5 と類似の事実が安定部分空間においても成り立つ.

補題 33.4.4. $f, g : U \rightarrow U$ を可換な線形変換とする (すなわち $g \circ f = f \circ g$). このとき $V = \text{Im } g$ は f -不変部分空間である.

Proof. 各 $\mathbf{v} \in V$ について, $f(\mathbf{v}) \in V$ を示したい. $\mathbf{v} \in V = \text{Im } g$ より, $g(\mathbf{u}) = \mathbf{v}$ を満たす $\mathbf{u} \in U$ が存在する. このとき, f と g の可換性より

$$f(\mathbf{v}) = f(g(\mathbf{u})) = f \circ g(\mathbf{u}) = g \circ f(\mathbf{u}) = g(f(\mathbf{u})) \in g(U) = V.$$

\square

系 33.2.5 と同様の理由により, 次を得る:

系 33.4.5. 有限次元線形空間上の線形変換 $f : U \rightarrow U$ および $\lambda \in \mathbb{K}$ に対して, $f - \lambda I$ の安定部分空間は f -不変部分空間である.

次の主張は, 環論における Fitting の補題に相当する. 本書において次元定理が応用される最初の例となろう.

定理 33.4.6. 線形変換 $g : U \rightarrow U$ の冪零部分群を W , 安定部分空間を V とすれば次が成り立つ:

(1) $W \cap V = \{\mathbf{0}\}$.

(2) w_1, \dots, w_k を W の基底, v_1, \dots, v_m を V の基底とすれば, これらを合わせた $k+m$ 個のベクトルからなる組は U の基底である.

Proof. 自然数 N を十分大きく取り, $W = \text{Ker } g^N$ かつ $V = \text{Im } g^N$ が満たされているとする.

(1): $u \in W \cap V$ とすれば, $u \in W$ より $g^N(u) = \mathbf{0}$ である. 命題 33.4.3 より $g^N|_V : V \rightarrow V$ は単射であり, これと $u \in \text{Ker } g^N|_V$ を合わせれば $u = \mathbf{0}$ を得る.

(2): まず線形独立性を示そう. $\sum_{i=1}^k c_i w_i + \sum_{j=1}^m r_j v_j = \mathbf{0}$ を仮定し, この式を次のように変形する:

$$W \ni \sum_{i=1}^k c_i w_i = -\sum_{j=1}^m r_j v_j \in V.$$

上の等式が表すベクトルは $W \cap V$ の元であるから, これらは (1) より零ベクトルに等しい. つまり $\sum_{i=1}^k c_i w_i = \mathbf{0}$ かつ $\sum_{j=1}^m r_j v_j = \mathbf{0}$ であり, 組 w_1, \dots, w_k および v_1, \dots, v_m の線形独立性より $c_i = 0, r_j = 0$ を得る. $k+m$ 個のベクトルの組 $w_1, \dots, w_k, v_1, \dots, v_m$ が U を生成することをいうには, 命題 22.3.3(2) より $\dim U = k+m$ を示せばよい. $g^N : U \rightarrow U$ に対して次元定理を適用すると,

$$\dim U = \dim \text{Ker } g^N + \dim \text{Im } g^N = \dim W + \dim V = k+m.$$

□

33.5 直和分解 (発展)

先の定理 33.4.6 や固有空間分解において, いくつかの部分空間の基底たちを並べることで全空間の基底を与えるという操作を行った. このような分解のことを直和分解という.

定義 33.5.1. 線形空間 V の部分空間 W_1, \dots, W_n が次の条件を満たすとき, V は W_1, \dots, W_n たちによって直和分解されるという.

- (i) 和集合 $W_1 \cup W_2 \cup \dots \cup W_n$ は V を生成する.
- (ii) 各 $u_i \in W_i$ について, $u_1 + u_2 + \dots + u_n = \mathbf{0} \implies u_1 = u_2 = \dots = u_n = \mathbf{0}$.

またこのとき,

$$V = \bigoplus_{i=1}^n W_i \quad \text{あるいは} \quad V = W_1 \oplus W_2 \oplus \dots \oplus W_n$$

と表す.

【補足】 形式上, 自明な分解 $U = U \oplus \{\mathbf{0}_U\}$ も直和分解とみなす.

例 33.5.2. 平面 \mathbb{R}^2 における x 軸のなす集合 $X = \{(x, 0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$ および y 軸のなす集合 $Y = \{(0, y) \in \mathbb{R}^2 \mid y \in \mathbb{R}\}$ は \mathbb{R}^2 の部分空間であり, $\mathbb{R}^2 = X \oplus Y$ が成り立つ.

各 W_i が部分空間であることから, 上の条件 (i) は次と同値である:

- 任意の $v \in V$ は, $w_i \in W_i$ を用いて $v = w_1 + \dots + w_n$ と表せる.

また, 条件 (ii) は, 文献によっては次のように置き換えられる.

命題 33.5.3. 線形空間 V の部分空間 W_1, \dots, W_r について, 次は同値である:

- (1) 各 $u_i \in W_i$ について, $u_1 + \dots + u_r = \mathbf{0} \implies u_1 = \dots = u_r = \mathbf{0}$.
- (2) 任意の増大列 $1 \leq m_1 < m_2 < \dots < m_\ell \leq r$ および任意の零ベクトルでない $u_{m_i} \in W_{m_i}$ について, $u_{m_1}, \dots, u_{m_\ell}$ は線形独立である.

(3) $v \in V$ が各 $w_i \in W_i$ の和として表されるならば、その表し方は一意的である。

更に V が有限次元ならば、これらは次の条件とも同値である：

(4) $u_{i,1}, u_{i,2}, \dots, u_{i,n_i}$ を W_i の基底とすれば、これらを全て集めた

$$\mathcal{B} = \{ u_{i,k} \mid i = 1, \dots, r, k = 1, \dots, n_i \}$$

は線形独立なベクトルの組である。

Proof. (1) \Rightarrow (2): $u_{m_i} \in W_{m_i}$ および $u_{m_i} \neq \mathbf{0}$, $\sum_{j=1}^{\ell} c_j u_{m_j} = \mathbf{0}$ を仮定する。(1) より $c_j u_{m_j} = \mathbf{0}$ であり、これと $u_{m_j} \neq \mathbf{0}$ から $c_j = 0$ を得る。

(2) \Rightarrow (3): $\sum_{i=1}^r u_i = \sum_{i=1}^r w_i$ (ただし $u_i, w_i \in W_i$) を仮定し、各 $v_i = u_i - w_i \in W_i$ が零ベクトルになることを示そう。いま $\sum_{i=1}^r v_i = \mathbf{0}$ である。仮に零ベクトルにならない v_i があるとし、それらをすべて列挙したものを $v_{m_1}, \dots, v_{m_\ell}$ とすれば、(2) よりこれは線形独立である。一方、 $\sum_{j=1}^{\ell} v_{m_j} = \mathbf{0}$ であり、これは非自明な線形関係ゆえ、 $v_{m_1}, \dots, v_{m_\ell}$ の線形独立性に反する。

(3) \Rightarrow (1): $u_1 + \dots + u_n = \mathbf{0}$ とする。零ベクトルは $\mathbf{0} \in W_i$ をもちいて $\mathbf{0} = \sum_{i=1}^r \mathbf{0}$ と書ける。(3) より W_i の元の和として $\mathbf{0}$ を表す方法は一通りしかないことから、 $u_i = \mathbf{0}$ を得る。

(1) \Rightarrow (4): 29.1 節で述べた命題 29.1.1 の別証明と同様にして示される $\sum_{i=1}^r \sum_{k=1}^{n_i} c_{i,k} u_{i,k} = \mathbf{0}$ とする。 $w_i := \sum_{k=1}^{n_i} c_{i,k} u_{i,k}$ とおけば $w_i \in W_i$ であり、 $\sum_{i=1}^r w_i = \mathbf{0}$ 。ゆえに (1) より $w_i = \mathbf{0}$ となる。したがって $\sum_{k=1}^{n_i} c_{i,k} u_{i,k} = \mathbf{0}$ であり、 $u_{i,1}, u_{i,2}, \dots, u_{i,n_i}$ の線形独立性から $c_{i,k} = 0$ 。

(4) \Rightarrow (1): 各 $w_i \in W_i$ について、 $w_1 + \dots + w_r = \mathbf{0}$ を仮定しよう。このとき、(4) より $w_i := \sum_{k=1}^{n_i} c_{i,k} u_{i,k}$ と書ける。すなわち、 $\sum_{i=1}^r \sum_{k=1}^{n_i} c_{i,k} u_{i,k} = \mathbf{0}$ であり、 \mathcal{B} の線形独立性から $c_{i,k} = 0$ 。つまり $w_i = \mathbf{0}$ である。□

したがって、次の (2) あるいは (3), (4) を直和分解の定義としてもよい。本書では主に (4) を用いている。

系 33.5.4. 線形空間 V の部分空間 W_1, \dots, W_r について、次は同値である：

$$(1) V = \bigoplus_{i=1}^r W_i.$$

(2) 各 $v \in V$ が $w_i \in W_i$ の和として表され、任意の増大列 $1 \leq m_1 < m_2 < \dots < m_\ell \leq r$ および任意の零ベクトルでない $u_{m_i} \in W_{m_i}$ について、 $u_{m_1}, \dots, u_{m_\ell}$ は線形独立である。

(3) 各 $v \in V$ が $w_i \in W_i$ の和として表され、かつその表し方は一意的である。

更に V が有限次元ならば、これらは次の条件とも同値である：

(4) $u_{i,1}, u_{i,2}, \dots, u_{i,n_i}$ を W_i の基底とすれば、これらを全て集めた

$$\mathcal{B} = \{ u_{i,k} \mid i = 1, \dots, r, k = 1, \dots, n_i \}$$

は V の基底である。

例 33.5.5. U を有限次元線形空間とする。

(1) 線形変換 $g : U \rightarrow U$ の冪零部分空間を W 、安定部分空間を V とすれば、 $U = W \oplus V$ (定理 33.4.6)。

(2) 線形変換 $f : U \rightarrow U$ の表現行列が対角化可能であるとき、 U は f の各固有空間に直和分解される (定理 29.3.1)。

第34章 一般固有空間分解とその応用(発展)

表現行列として対角行列を取れる線形変換においては、各ベクトルが固有ベクトルの線形結合で表されるのであった。これに対して一般の線形変換については、その特性多項式が因数分解できるならば各ベクトルを一般固有ベクトルの線形結合で書くことができる。本章ではまず、この事実から導かれる主張や応用について言及し、最後に証明を与える。

34.1 一般固有空間分解

次の定理の証明は本章の後半で与えるとして、まずはこの定理から導かれる基本的事実について解説しよう。

定理 34.1.1 (一般固有空間分解). 有限次元線形空間上の線形変換 $f : U \rightarrow U$ の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とし, f の特性多項式が

$$\Phi_f(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$$

と因数分解されているとする。このとき, $\dim \widetilde{W}(\lambda_k, f) = n_k$ ($k = 1, \dots, r$) が成り立つ。また, 各 $\widetilde{W}(\lambda_k, f)$ の基底 $\mathbf{u}_{\lambda_k, 1}, \dots, \mathbf{u}_{\lambda_k, n_k}$ をそれぞれ一組えらび, これらをすべて集めたベクトルの組

$$\mathcal{B} = \{ \mathbf{u}_{\lambda_k, j} \mid k = 1, \dots, r, j = 1, \dots, n_k \}$$

を取れば, \mathcal{B} は U の基底になる。

【補足】 上は, 直和分解 $U = \bigoplus_{k=1}^r \widetilde{W}(\lambda_k, f)$ が成立することを主張している。

体として \mathbb{C} を取る場合は系 29.4.6 により定理 34.1.1 の前提が必ず満たされ, したがって次が導かれる。

系 34.1.2. U を体 \mathbb{C} 上の有限次元線形空間とする。任意の線形変換 $f : U \rightarrow U$ について, U は f の一般固有ベクトルからなる基底を持つ。

表現行列が対角化可能であるとき, 一般固有空間と固有空間は一致する:

系 34.1.3. $f : U \rightarrow U$ の表現行列が対角化可能であるとき, f の各固有値 λ について $W(\lambda, f) = \widetilde{W}(\lambda, f)$ 。

Proof. 背理法で示す。ある固有値 λ_j について $W(\lambda_j, f) \neq \widetilde{W}(\lambda_j, f)$ であると仮定すれば, $\dim W(\lambda_j, f) < \dim \widetilde{W}(\lambda_j, f)$ である。このとき, 定理 29.3.1 から

$$\dim U = \sum_{k=1}^r \dim W(\lambda_k, f) < \sum_{k=1}^r \dim \widetilde{W}(\lambda_k, f) = \sum_{k=1}^r n_k = \dim U.$$

$\dim U < \dim U$ が導かれ, これは不合理である。 □

冪零変換 $g = (f - \lambda_k I)|_{\widetilde{W}(\lambda_k, f)} : \widetilde{W}(\lambda_k, f) \rightarrow \widetilde{W}(\lambda_k, f)$ に対して $\dim \widetilde{W}(\lambda_k, f) = n_k$ および補題 33.2.3(3) を適用することで次を得る。

系 34.1.4. 定理 34.1.1 の前提のもとで, $\widetilde{W}(\lambda_k, f) = \text{Ker}(f - \lambda_k I)^{n_k}$ 。

備考 34.1.5. (1) 定理 34.1.1 の設定のもとで、次のような $\widetilde{W}(\lambda_k, f)$ の基底の選び方を考えよう:

$\mathbf{u}_{\lambda_k,1}, \dots, \mathbf{u}_{\lambda_k,n_k}$ の選び方として、まず $\text{Ker}(f - \lambda_k I)$ の基底を選び、それに $\text{Ker}(f - \lambda_k I)^2$ の元を付け加えて $\text{Ker}(f - \lambda_k I)^2$ の基底とし、更に $\text{Ker}(f - \lambda_k I)^3$ の元を付け加えて $\text{Ker}(f - \lambda_k I)^3$ の基底とし… という操作を繰り返して $\widetilde{W}(\lambda_k, f)$ の基底 $\mathbf{u}_{\lambda_k,1}, \dots, \mathbf{u}_{\lambda_k,n_k}$ を得る. ここで $\mathbf{u}_{\lambda_k,j} \in \text{Ker}(f - \lambda_k I)^\ell$ とすれば,

$$\begin{aligned}(f - \lambda_k I)(\mathbf{u}_{\lambda_k,j}) &\in \text{Ker}(f - \lambda_k I)^{\ell-1} \\ f(\mathbf{u}_{\lambda_k,j}) - \lambda_k \mathbf{u}_{\lambda_k,j} &\in \text{Ker}(f - \lambda_k I)^{\ell-1} \\ f(\mathbf{u}_{\lambda_k,j}) &\in \text{Ker}(f - \lambda_k I)^{\ell-1} + \lambda_k \mathbf{u}_{\lambda_k,j}.\end{aligned}$$

上の最後の式は、 $f(\mathbf{u}_{\lambda_k,j})$ が $\mathbf{u}_{\lambda_k,1}, \mathbf{u}_{\lambda_k,2}, \dots, \mathbf{u}_{\lambda_k,j}$ の線形結合で書けることを述べている. このとき、この基底に関する $f|_{\widetilde{W}(\lambda_k, f)} : \widetilde{W}(\lambda_k, f) \rightarrow \widetilde{W}(\lambda_k, f)$ の表現行列 A_k は上三角行列となる. これと命題 33.1.2 および備考 33.1.3 を合わせれば、 f の表現行列を上三角行列に取れることが導かれる.

(2) X を n 次正方行列とする. 系 34.1.2 により \mathbb{C}^n は $T_X : \mathbb{C}^n \rightarrow \mathbb{C}^n$. このとき、(1) で与えた基底に関する T_X の表現行列は上三角である. つまり、任意の正方行列は複素数の範囲において、ある上三角行列と相似になる. なお、(エルミート) 内積に関する文脈からも直交行列 (ユニタリ行列) による行列の上三角化が論じられる.

(3) (2) より任意の複素正方行列は上三角行列と相似であり、上三角行列の対角成分には、その固有値が重複を込めて並んでいる (例 28.4.7). また、相似な行列のトレースと固有多項式は等しかった (命題 26.3.5 および系 28.4.5). したがって、複素数を成分とする正方行列 A の固有多項式が $\Phi_A(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$ と因数分解されるとき、

(i) A の重複を込めたすべての固有値の積は、 A の行列式に等しい. すなわち、

$$|A| = \lambda_1^{n_1} \lambda_2^{n_2} \cdots \lambda_r^{n_r}.$$

(ii) A のトレースは、重複を込めたすべての固有値の和に等しい. すなわち、

$$\text{tr } A = \sum_{k=1}^r \lambda_k n_k.$$

備考 34.1.6. 実は、巧妙に $\widetilde{W}(\lambda_k, f)$ の基底を選ぶことにより、表現行列を式 (33.2.1) のような形にできることが知られており、これをジョルダン標準形 (Jordan normal form) と言う.

次の定理は練習 28.2.6 の内容をさらに精査したものである.

定理 34.1.7 (フロベニウス). 線形変換 $f : U \rightarrow U$ の固有多項式が $\Phi_f(t) = \prod_{i=1}^n (t - \gamma_i)$ と因数分解されているとする (つまり $\gamma_1, \dots, \gamma_n$ は重複を込めた f の固有値である). このとき任意の多項式 $\Psi(t)$ について、線形変換 $\Psi(f)$ の特性多項式は

$$\Phi_{\Psi(f)}(t) = \prod_{i=1}^n (t - \Psi(\gamma_i)).$$

とくに、 $\lambda \in \mathbb{K}$ を f の固有値とすれば、 $\dim \widetilde{W}(\lambda, f) \leq \dim \widetilde{W}(\Psi(\lambda), \Psi(f))$ である.

Proof. 備考 34.1.5(1) により、 U の基底を上手くとることで f の表現行列 A を上三角行列にできる. このとき、この基底における $\Psi(f)$ の表現行列は $\Psi(A)$ である (系 26.2.5). また A が上三角行列であることから、その固有多項式は例 28.4.7 のように計算され、したがって A の対角成分は必要があれば順番を入れ替えて $\gamma_1, \dots, \gamma_n$ であるとしてよい. このとき A^k も上三角であり、その対角成分は $\gamma_1^k, \dots, \gamma_n^k$ となる¹. したがって $\Psi(A)$ も上三角であり、その対角成分は $\Psi(\gamma_1), \dots, \Psi(\gamma_n)$ である. ゆえに $\Phi_{\Psi(f)}(t) = \Phi_{\Psi(A)}(t) = \prod_{i=1}^n (t - \Psi(\gamma_i))$. \square

¹定義 3.3.1 の直前にある計算を見よ.

【補足】写像 $\Psi : \{\gamma_1, \dots, \gamma_n\} \rightarrow \{\Psi(\gamma_1), \dots, \Psi(\gamma_n)\}$ が単射である場合は $\dim \widetilde{W}(\lambda, f) = \dim \widetilde{W}(\Psi(\lambda), \Psi(f))$ が成り立つ。

34.2 ケーリー-ハミルトンの定理 (再論)

一般固有空間への分解を用いたケーリー-ハミルトンの定理の証明を紹介する。証明の筋書きが命題 30.3.2 と類似していることを確認してほしい。

定理 34.2.1. n 次正方形行列 A について $\Phi_A(A) = O$.

Proof. 複素数の範囲において $\Phi_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$ と因数分解しよう。ここで、 $\lambda_1, \dots, \lambda_r$ は相異なる A の固有値である。 $T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ に対して系 34.1.2 を適用すれば、 \mathbb{C}^n の各元は一般固有ベクトルの線形結合で表せる。したがって、各 $v \in \widetilde{W}(\lambda_k, A)$ について $\Phi_A(A)v = \mathbf{0}$ を示せば十分である (練習 18.1.4)。系 34.1.4 より $\widetilde{W}(\lambda_k, A) = \text{Ker}(T_A - \lambda_k I)^{n_k}$ 、つまり、

$$(A - \lambda_k E)^{n_k} v = (T_A - \lambda_k I)^{n_k}(v) = \mathbf{0}$$

である。 $\Phi_A(t)$ は、 $n - n_k$ 次多項式 $\Theta(t)$ を用いて $\Phi_A(t) = \Theta(t)(t - \lambda_k)^{n_k}$ と書けることから、

$$\Phi_A(A)v = \Theta(A)(A - \lambda_k E)^{n_k} v = \Theta(A)\mathbf{0} = \mathbf{0}.$$

以上より $\Phi_A(A) = O$. □

A が実数を成分とする正方形行列であり、実数の範囲において A の特性多項式が 1 次式の積に因数分解されない場合や、特に特性方程式が実数解を持たない場合 (つまり固有ベクトルが存在しない場合) においても上の証明は有効である。この事実は、実数に限った話題であっても複素数に範囲を広げておくことで理解が容易になる可能性を示唆している。

34.3 線形漸化式と線形常微分方程式 (再論)

線形漸化式や線形常微分方程式の特性多項式が重解を持つ場合は、任意のベクトルを固有ベクトルに分解することはできなかった。この場合における一般解の表示は、一般固有空間の基底による線形結合表示によって得られる。

常微分方程式 (32.3.1) および線形漸化式 (31.3.1) の特性多項式は $\Phi(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1t + a_0$ であった。関数 y が常微分方程式 (32.3.1) を満たすことは、次の三つの同値な式に言い換えられる:

$$\begin{aligned} D^k(y) + a_{k-1}D^{k-1}(y) + a_{k-2}D^{k-2}(y) + \cdots + a_1D(y) + a_0I(y) &= \mathbf{0}, \\ \left(D^k + a_{k-1}D^{k-1} + a_{k-2}D^{k-2} + \cdots + a_1 + a_0I \right)(y) &= \mathbf{0}, \\ \Phi(D)(y) &= \mathbf{0}. \end{aligned}$$

同様にして数列 $x \in \mathbb{K}^{\mathbb{N}}$ が線形漸化式 (31.3.1) を満たすことは次を満たすことと同値である:

$$\Phi(S)(x) = \mathbf{0}.$$

常微分方程式 (32.3.1) の解空間を W 、漸化式 (31.3.1) を満たす数列空間を F とする。

補題 34.3.1. 常微分方程式 (32.3.1) の特性多項式が $\Phi(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$ と因数分解されるとする。このとき、各 $k = 1, \dots, r$ について、例 33.3.2 で与えた一般固有ベクトルによる n_k 個の線形独立な組 $e^{\lambda_k x}, x e^{\lambda_k x}, x^2 e^{\lambda_k x}, \dots, x^{n_k-1} e^{\lambda_k x}$ は $D|_W : W \rightarrow W$ の固有値 λ_k に関する一般固有空間 $\widetilde{W}(\lambda_k, D)$ の基底である。

Proof. 定理 34.1.1 より $\dim \widetilde{W}(\lambda_k, S) = n_k$ ゆえ, $e^{\lambda_k x}, xe^{\lambda_k x}, \dots, x^{n_k-1}e^{\lambda_k x} \in W$ さえ示せばよい. つまり, 各 $y_i(x) = x^i e^{\lambda_k x}$ ($i = 0, \dots, n_k - 1$) について $\Phi(D)(y_i) = \mathbf{0}$ を示せばよい. $n - n_k$ 次多項式 $\Theta(t)$ を用いて $\Phi(t) = \Theta(t)(t - \lambda_k)^{n_k}$ と書けば, $\Phi(D) = \Theta(D)(D - \lambda_k I)^{n_k}$ である. 例 33.3.2(1) より $y_i \in \text{Ker}(D - \lambda_k I)^{n_k}$ ゆえ, $\Phi(D)(y_i) = \Theta(D)(D - \lambda_k I)^{n_k}(y_i) = \Theta(D)(\mathbf{0}) = \mathbf{0}$. \square

いまと類似の議論で次を得る.

補題 34.3.2. 線形漸化式 (31.3.1) の特性多項式が $\Phi(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \dots (t - \lambda_r)^{n_r}$ と因数分解されるとする. このとき, 各 $k = 1, \dots, r$ について, 例 33.3.3 で与えた一般固有ベクトルによる n_k 個の線形独立な組 $\mathbf{x}_0^{\lambda_k}, \dots, \mathbf{x}_{n_k-1}^{\lambda_k}$ は $S|_F : F \rightarrow F$ の固有値 λ_k に関する一般固有空間 $\widetilde{W}(\lambda_k, S)$ の基底である.

定理 34.1.1 の後半より次が従う:

定理 34.3.3. 常微分方程式 (32.3.1) の特性多項式が $\Phi(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \dots (t - \lambda_r)^{n_r}$ であるとする. このとき, 次の集合 \mathcal{B} は常微分方程式 (32.3.1) の解空間の基底となる:

$$\mathcal{B} = \{ e^{\lambda_k x}, xe^{\lambda_k x}, x^2 e^{\lambda_k x}, \dots, x^{n_k-1} e^{\lambda_k x} \mid k = 1, \dots, r \}.$$

つまり, 常微分方程式 (32.3.1) の一般解は \mathcal{B} に現れる関数の線形結合として表示できる.

定理 34.3.4. 線形漸化式 (31.3.1) の特性多項式が $\Phi(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \dots (t - \lambda_r)^{n_r}$ であるとする. このとき, この漸化式を満たす数列は次の基底による線形結合で表される:

$$\mathcal{B} = \{ \mathbf{x}_0^{\lambda_k}, \dots, \mathbf{x}_{n_k-1}^{\lambda_k} \mid k = 1, \dots, r \}.$$

ここで, 各 $\mathbf{x}_N^{\lambda_k}$ は例 33.3.3 で与えた数列を指す. とくに, この漸化式を満たす数列の一般項は, \mathcal{B} に現れる数列の一般項の線形結合として表示できる.

34.4 定理 34.1.1 の証明

最後に主定理の証明を述べよう. 本節では定理 34.1.1 にある仮定が満たされていること, すなわち次を前提とする:

- 線形変換 $f : U \rightarrow U$ の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とし, f の特性多項式が $\Phi_f(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \dots (t - \lambda_r)^{n_r}$ と因数分解されているとする.

次の補題において, $\widetilde{W}(\lambda, f)$ が $(f - \gamma I)$ -不変部分空間であることに注意しておく (系 33.2.5).

補題 34.4.1. f の固有値 λ について次が成り立つ.

- (1) $\gamma \neq \lambda$ とすれば, $\widetilde{W}(\lambda, f)$ は固有値 γ に関する固有ベクトルを含まない.
- (2) $\gamma \neq \lambda$ とすれば, $(f - \gamma I)|_{\widetilde{W}(\lambda, f)} : \widetilde{W}(\lambda, f) \rightarrow \widetilde{W}(\lambda, f)$ は線形同型である.
- (3) 多項式 $\Psi(t) = (t - \delta_1) \dots (t - \delta_m)$ が $t - \lambda$ を因子に含まないとすれば, λ に関する一般固有ベクトル $\mathbf{v} \neq \mathbf{0}$ について $\Psi(f)(\mathbf{v}) \neq \mathbf{0}$.

Proof. (1): $\widetilde{W}(\lambda, f) = \text{Ker}(f - \lambda I)^N$ が成り立つよう十分大きな自然数 N を与えれば, 各 $\mathbf{v} \in \widetilde{W}(\lambda, f)$ について $(f - \lambda I)^N(\mathbf{v}) = \mathbf{0}$ が成り立つ. 仮にこの \mathbf{v} が γ に関する固有ベクトルであるとすれば, $(f - \lambda I)(\mathbf{v}) = \gamma \mathbf{v} - \lambda \mathbf{v} = (\gamma - \lambda)\mathbf{v}$ であり, さらに $(f - \lambda I)$ をほどこすことで, $(f - \lambda I)^N(\mathbf{v}) = (\gamma - \lambda)^N \mathbf{v}$ を得る. ゆえに $(\gamma - \lambda)^N \mathbf{v} = \mathbf{0}$ となり, $\gamma - \lambda \neq 0$ ゆえ $\mathbf{v} = \mathbf{0}$. これは \mathbf{v} が固有ベクトル (つまり $\mathbf{v} \neq \mathbf{0}$) であることに反する.

(2): 単射性を示せばよい. 各 $\mathbf{v} \in \widetilde{W}(\lambda, f)$ について, $(f - \gamma I)(\mathbf{v}) = \mathbf{0}$ とすれば \mathbf{v} は零ベクトルであるか固有値 γ に関する固有ベクトルである. (1) より後者は否定され, ゆえに $\mathbf{v} = \mathbf{0}$. すなわち, $(f - \gamma I)|_{\widetilde{W}(\lambda, f)}$ の核は自明であり, これは単射である.

(3): 同型写像 $(f - \delta_i I)|_{\widetilde{W}(\lambda, f)}$ を順次ほどこすことで主張を得る. \square

定理 34.1.1 の前半を示そう.

命題 34.4.2. $\dim \widetilde{W}(\lambda_k, f) = n_k$ ($k = 1, \dots, r$).

Proof. $W = \widetilde{W}(\lambda_k, f)$ とおく. W は線形写像 $(f - \lambda_k I) : U \rightarrow U$ の冪零部分空間である. W の基底を $\mathbf{w}_1, \dots, \mathbf{w}_d$ とする. W は f -不変であり, この基底に関する $f|_W : W \rightarrow W$ の表現行列を A とする. 次に, V を $(f - \lambda_k I)$ の安定部分空間とし, その基底を $\mathbf{v}_1, \dots, \mathbf{v}_m$ とする. V も f -不変であり, この基底に関する $f|_V : V \rightarrow V$ の表現行列を B としよう. 定理 33.4.6 より U は W と V に直和分解される. すなわち, $\mathbf{w}_1, \dots, \mathbf{w}_d$ および $\mathbf{v}_1, \dots, \mathbf{v}_m$ を合わせた $d + m$ 個のベクトルからなる組は U の基底となる. この基底に関する $f : U \rightarrow U$ の表現行列を X とすれば, 命題 33.1.2 より,

$$X = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$$

となる. ゆえに特性多項式は次のように分解される:

$$\Phi_f(t) = \Phi_X(t) = \Phi_A(t)\Phi_B(t) = \Phi_{f|_W}(t)\Phi_{f|_V}(t).$$

いまから $\Phi_{f|_W}(t) = (t - \lambda_k)^{n_k}$ を示そう. そのためには補題 29.4.10 より次の二点を確認すればよい:

- λ_k を除く f の固有値 γ について, $\Phi_{f|_W}(t)$ は $(t - \gamma)$ を因子に持たない.

Proof. $\Phi_{f|_W}(t)$ が $(t - \gamma)$ を因子に持つとすれば γ は $f|_W$ の固有値であり, ゆえに γ に関する $f|_W$ の固有ベクトル $\mathbf{v} \in W$ が存在する. しかしこれは補題 34.4.1(1) に矛盾してしまう. \square

- $\Phi_{f|_V}(t)$ は $(t - \lambda_k)$ を因子に持たない.

Proof. 仮に $\Phi_{f|_V}(t)$ が $(t - \lambda_k)$ を因子に持つとすれば, V は λ_k に関する $f|_V$ の固有ベクトル $\mathbf{v} \neq \mathbf{0}$ を含む. このとき $(f - \lambda_k I)(\mathbf{v}) = \mathbf{0}$ であり, これは $(f - \lambda_k I) : V \rightarrow V$ の単射性 (命題 33.4.3(2)) に反する. \square

以上により, $\Phi_{f|_W}(t) = (t - \lambda_k)^{n_k}$ であり, したがって $\dim W = n_k$ である (練習 28.4.9). \square

次の主張はケーリー-ハミルトンの定理の項目で紹介した補題 29.1.3 の一般化にほかならない. 証明方法も大して違いはない.

補題 34.4.3. $\mathbf{u}_k \in \widetilde{W}(\lambda_k, f)$ ($k = 1, \dots, r$) かつ $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ ならば, $\mathbf{u}_1 = \dots = \mathbf{u}_r = \mathbf{0}$.

Proof. 系 34.1.4 より $\widetilde{W}(\lambda_k, f) = \text{Ker}(f - \lambda_k I)^{n_k}$, つまり $(f - \lambda_k I)^{n_k}(\mathbf{u}_k) = \mathbf{0}$ である. いまから, $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ の両辺に $(f - \lambda_k I)^{n_k}$ ($k = 2, \dots, r$) を順次ほどこすことで $\mathbf{u}_2, \dots, \mathbf{u}_r$ の項が消えて \mathbf{u}_1 に関する項のみが残ることを見よう.

多項式 $\Psi_1(t) = (t - \lambda_2)^{n_2} \dots (t - \lambda_r)^{n_r}$ は, 別の多項式 $\Theta_k(t)$ を用いて $\Psi_1(t) = \Theta_k(t)(t - \lambda_k)^{n_k}$ ($k = 2, \dots, r$) と書ける. つまり $\Psi_1(f) = \Theta_k(f)(f - \lambda_k I)^{n_k}$ である. $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ の両辺に線形変換 $\Psi_1(f)$ をほどこすと

$$\begin{aligned} \Psi_1(f)(\mathbf{u}_1 + \dots + \mathbf{u}_r) &= \mathbf{0} \\ \sum_{k=1}^r \Psi_1(f)(\mathbf{u}_k) &= \mathbf{0} \\ \Psi_1(f)(\mathbf{u}_1) + \sum_{k=2}^r \Theta_k(f)(f - \lambda_k I)^{n_k}(\mathbf{u}_k) &= \mathbf{0} \\ \Psi_1(f)(\mathbf{u}_1) &= \mathbf{0}. \quad (\text{ここで } (f - \lambda_k I)^{n_k}(\mathbf{u}_k) = \mathbf{0} \text{ を用いた}) \end{aligned}$$

ゆえに補題 34.4.1(3) より $\mathbf{u}_1 = \mathbf{0}$ でなければならない.

次に $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ の両辺に $(f - \lambda_k I)^{n_k}$ ($k = 1, 3, 4, \dots, r$) を順次ほどこすことで $\mathbf{u}_2 = \mathbf{0}$ を得る. これらと類似の操作を順次繰り返す, $\mathbf{u}_1 = \dots = \mathbf{u}_r = \mathbf{0}$ を得る. \square

定理 34.1.1 の後半は次の通りである.

命題 34.4.4. 各 $\widetilde{W}(\lambda_k, f)$ の基底 $\mathbf{u}_{k,1}, \dots, \mathbf{u}_{k,n_k}$ をそれぞれ一組えらび, これらをすべて集めたベクトルの組

$$\mathcal{B} = \{ \mathbf{u}_{k,j} \mid k = 1, \dots, r, j = 1, \dots, n_k \}$$

を取れば, \mathcal{B} は U の基底になる.

Proof. 命題 34.4.2 より

$$\sum_{k=1}^r \dim \widetilde{W}(\lambda_k, f) = \sum_{k=1}^r n_k = \dim U$$

であるから, $\dim U$ 個のベクトルからなる集合 \mathcal{B} の線形独立性さえ示せば, 命題 22.3.3(2) より \mathcal{B} は U の基底となる. 線形独立性は, 先の補題および命題 33.5.3 における (1) と (4) の同値性から導かれる. \square

本節では簡単のため固有多項式が 1 次式の積に因数分解される場合に限って論じた. 一方で, $\mathbb{K} = \mathbb{R}$ の場合で特性多項式が \mathbb{R} 上で 1 次式の積に因数分解されない場合についても, 実固有値の一般固有空間の次元について, 命題 34.4.2 と同様の公式が成り立つ (命題 34.4.5).

線形空間 U において, それを体 \mathbb{K} 上の線形空間と見なしていることを強調するとき, その次元を $\dim_{\mathbb{K}} U$ と書く. 例えば, $\dim_{\mathbb{R}} \mathbb{C} = 2$ (例えば組 $1, i$ が基底となる) および $\dim_{\mathbb{C}} \mathbb{C} = 1$ である. また, 一般固有空間についても考えている体 \mathbb{K} を強調する場合は $\widetilde{W}_{\mathbb{K}}(\lambda_k, A)$ などと書くことにする.

命題 34.4.5. (1) 実係数 n 次正方行列 A の固有多項式が, 相異なる $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ を用いて $\Phi_A(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \dots (t - \lambda_r)^{n_r} \Psi(t)$ と因数分解される時 (ただし, 各 $k = 1, \dots, r$ について $\Psi(\lambda_k) \neq 0$ とする), $\dim_{\mathbb{R}} \widetilde{W}_{\mathbb{R}}(\lambda_k, A) = n_k$ が成り立つ.

(2) U を体 \mathbb{R} 上の有限次元線形空間とする. 線形変換 $f: U \rightarrow U$ の特性多項式が, 相異なる $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ を用いて $\Phi_A(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \dots (t - \lambda_r)^{n_r} \Psi(t)$ と因数分解される時 (ただし, 各 $k = 1, \dots, r$ について $\Psi(\lambda_k) \neq 0$ とする), $\dim_{\mathbb{R}} \widetilde{W}_{\mathbb{R}}(\lambda_k, f) = n_k$ が成り立つ.

Proof. (1): $m_k := \dim_{\mathbb{R}} \widetilde{W}_{\mathbb{R}}(\lambda_k, A)$ とおく. 示すべきことは $m_k = n_k$ である. さて, $\Phi_A(t)$ は, 複素数係数の多項式とみなせば 1 次式の積に因数分解され, このとき命題 34.4.2 より $\dim_{\mathbb{C}} \widetilde{W}_{\mathbb{C}}(\lambda_k, A) = n_k$ である. $\widetilde{W}_{\mathbb{R}}(\lambda_k, A) = \widetilde{W}_{\mathbb{C}}(\lambda_k, A) \cap \mathbb{R}^n$ である点に注意せよ. いまから両方の不等式 $m_k \leq n_k$ および $m_k \geq n_k$ を示そう.

$\widetilde{W}_{\mathbb{R}}(\lambda_k, A)$ の基底 $\mathbf{u}_1, \dots, \mathbf{u}_{m_k} \in \mathbb{R}^n$ を取れば, これらは複素ベクトルとみなしても線形独立である. すなわち, $\widetilde{W}_{\mathbb{C}}(\lambda_k, A)$ は m_k 個の線形独立なベクトル $\mathbf{u}_1, \dots, \mathbf{u}_{m_k}$ を含むことから, $\dim_{\mathbb{C}} \widetilde{W}_{\mathbb{C}}(\lambda_k, A) \geq m_k$.

$\widetilde{W}_{\mathbb{C}}(\lambda_k, A)$ は方程式 $(A - \lambda_k E)^n \mathbf{x} = \mathbf{0}$ の解空間に等しい. 行列 $(A - \lambda_k E)^n$ の各成分が実数であることから, $(A - \lambda_k E)^n$ の簡約化の各成分も実数である. したがって, 掃き出し法により $\widetilde{W}_{\mathbb{C}}(\lambda_k, A)$ の基底 $\mathbf{v}_1, \dots, \mathbf{v}_{n_k}$ を得れば², 各 \mathbf{v}_i の成分はすべて実数であり, ゆえに $\mathbf{v}_i \in \widetilde{W}_{\mathbb{C}}(\lambda_k, A) \cap \mathbb{R}^n = \widetilde{W}_{\mathbb{R}}(\lambda_k, A)$. つまり, $\widetilde{W}_{\mathbb{R}}(\lambda_k, A)$ は n_k 個の線形独立なベクトル $\mathbf{v}_1, \dots, \mathbf{v}_{n_k}$ を含む. ゆえに $\dim_{\mathbb{R}} \widetilde{W}_{\mathbb{R}}(\lambda_k, A) \geq n_k$.

(2): いま示した (1) と命題 33.2.6(3) による. \square

命題 34.4.2 および 34.4.5 から, 直ちに系 29.1.6 が導かれる.

²18.3 節で紹介した方法のこと.

第VI部

計量ベクトル空間

第35章 内積と正規直交基底

\mathbb{R}^n の標準ベクトルたちは互いに直交している。本書ではこれまで、この明らかな事実を用いることなく議論を進めてきた。ここで、 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ のなす角 (とくに直交性) やベクトルの長さについて分析し、幾何的な側面をやや強めた線形代数を紹介しよう。そしてそれは、内積を用いた線形代数と言い換えることができる。

35.1 標準的な内積とユークリッド・ノルム

定義 35.1.1. 二つの実 n 次列ベクトル $\mathbf{x} = {}^t(x_1, \dots, x_n)$ および $\mathbf{y} = {}^t(y_1, \dots, y_n)$ に対して定まる次の数

$$\langle \mathbf{x}, \mathbf{y} \rangle := {}^t\mathbf{x}\mathbf{y} = x_1y_1 + \dots + x_ny_n$$

を \mathbf{x} と \mathbf{y} の標準的な内積と呼ぶ。

本書では \mathbf{x} と \mathbf{y} で生成される線形部分空間のことも $\langle \mathbf{x}, \mathbf{y} \rangle$ と書いていた。そこで、第 VI 部では、二つのベクトルによる表示 $\langle \mathbf{x}, \mathbf{y} \rangle$ に限り、これは内積を指すとす。ただし、 $\langle x_1, \dots, x_n \rangle$ と表示した場合は、たとえ $n = 2$ の場合であっても部分空間を指すと約束しよう。なお、他の文献においては、内積の記号として $\mathbf{x} \cdot \mathbf{y}$, (\mathbf{x}, \mathbf{y}) , $\langle \mathbf{x} | \mathbf{y} \rangle$ などが用いられている。

備考 35.1.2. 本書では言葉のあやとして、行ベクトルどうし、あるいは行ベクトルと列ベクトルの内積を次の意味で既に用いている：

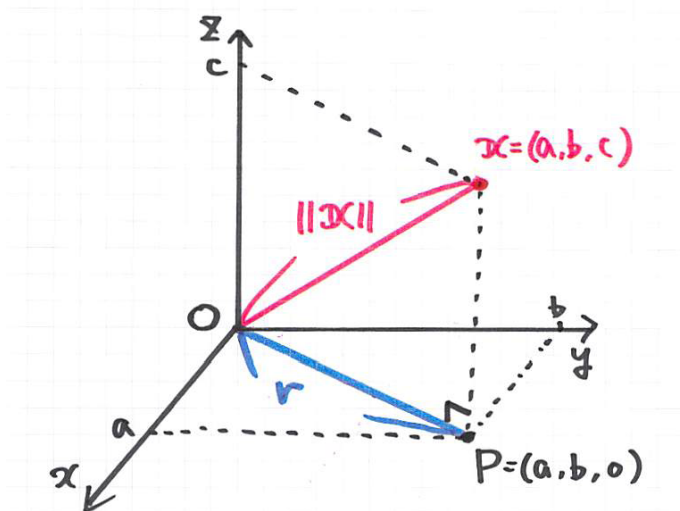
- $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ をともに行ベクトルとするとき、これらの内積を $\mathbf{x} {}^t\mathbf{y}$ と定める。
- $\mathbf{x} \in \mathbb{R}^n$ を行ベクトル、 $\mathbf{y} \in \mathbb{R}^n$ を列ベクトルとするとき、これらの内積を $\mathbf{x}\mathbf{y}$ と定める。

第 VI 部では主に列ベクトルどうしの内積を考えるものの、稀に上の意味で「内積」という語句を用いることがある。

下図における $\mathbf{x} = {}^t(a, b, c) \in \mathbb{R}^3$ の長さを $\|\mathbf{x}\|$ とすると、ピタゴラスの定理により

$$\|\mathbf{x}\|^2 = r^2 + c^2 = (a^2 + b^2) + c^2 = \langle \mathbf{x}, \mathbf{x} \rangle$$

である。ゆえに $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ が成り立つ。



この考察から類推されることとして、 \mathbb{R}^n において次の定義を与える:

定義 35.1.3. n 次列ベクトル $\mathbf{x} = {}^t(x_1, \dots, x_n) \in \mathbb{R}^n$ に対して、次で定まる 0 以上の実数

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_1^2 + \dots + x_n^2}$$

を \mathbf{x} の標準的なノルムあるいはユークリッド・ノルム (**Euclidean norm**) という.

とくに誤解がない限り、本書では標準的な内積やノルムのことをそれぞれ略して内積 (**inner product**) およびノルム (**norm**) と呼ぶ. $\|\mathbf{x}\|$ は、 \mathbb{R}^n における $\mathbf{0}$ と \mathbf{x} の間の距離 (あるいはベクトル \mathbf{x} の長さ) であると考えられる. そこで、ノルムをしばしば \mathbf{x} の長さ (**length**) とも言い表す.

ここで、内積とノルムの性質をまとめておこう. 次は、定義から直ちに導かれる性質ゆえ証明は略す.

命題 35.1.4. $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ および $r \in \mathbb{R}$ について次が成り立つ.

- (1) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$,
- (2) $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$, $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
- (3) $\langle r\mathbf{x}, \mathbf{y} \rangle = r\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, r\mathbf{y} \rangle$,
- (4) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$,
- (5) $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$.

【備考】 (2) の後半は、前半と (1) から導ける. 実際、 $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{z}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{z}, \mathbf{x} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$. (3) の後半の等号にも同様のことがいえる.

本書では、次節において余弦定理の復習をしたのち、35.3 節以降では内積の定義まで戻ることなく、内積に関する計算はいずれも上の性質のみを用いて行っていく. これにより、内積の抽象化が可能になる (詳しくは 35.6 節を見よ).

次の性質のうち、(3) 以外は明らかである (証明略). もちろん (3) は「三角形の二辺の長さの和は残す一辺の長さ以上である」を意味し、したがって明らかに思う読者も多いであろう. しかしながら、これを図示によらず内積の定義 (あるいは命題 35.1.4 に挙げた性質) と計算のみで示そうとすれば、それは明らかではない. (3) は劣加法性 (あるいは三角不等式) とも呼ばれ、この形式的証明は 38.1 節で与える.

命題 35.1.5. $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ および $r \in \mathbb{R}$ について次が成り立つ.

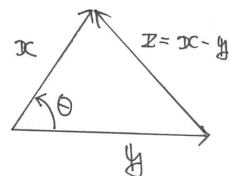
- (1) $\|r\mathbf{x}\| = |r| \cdot \|\mathbf{x}\|$,
- (2) $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$.
- (3) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

35.2 余弦定理と内積

内積とノルムの関係を、二つのベクトルのなす角を通して見てみよう.

命題 35.2.1 (余弦定理). $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ のなす角を $\theta \in [0, \pi]$ とし、 $\mathbf{z} = \mathbf{x} - \mathbf{y}$ とすれば、次が成り立つ:

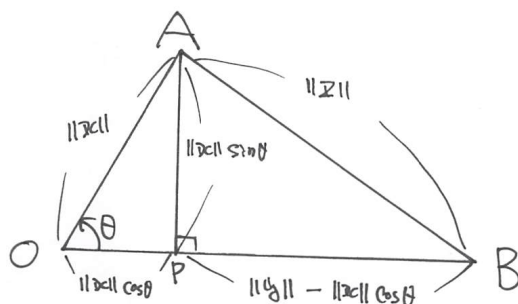
$$\|\mathbf{z}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta.$$



Proof. ベクトル \mathbf{x}, \mathbf{y} が指す位置をそれぞれ点 A, B , 原点を O とする. また点 O, B を結ぶ直線を L とする. 点 A から, 直線 L に向けて引いた垂線の足を P とする. 直角三角形 APB に対してピタゴラスの定理を適用すれば

$$\begin{aligned}\|\mathbf{z}\|^2 &= (\|\mathbf{x}\| \sin \theta)^2 + (\|\mathbf{y}\| - \|\mathbf{x}\| \cos \theta)^2 \\ &= \|\mathbf{x}\|^2 \sin^2 \theta + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos \theta + \|\mathbf{x}\|^2 \cos^2 \theta \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos \theta.\end{aligned}$$

下図は θ が鋭角な場合である. θ が鈍角な場合も上の議論は成り立つ. □



上の公式を変形すると,

$$\begin{aligned}\|\mathbf{x}\|\|\mathbf{y}\| \cos \theta &= \frac{1}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{z}\|^2) = \frac{1}{2} \left(\sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 - \sum_{i=1}^n (x_i - y_i)^2 \right) \\ &= \frac{1}{2} \sum_{i=1}^n 2x_i y_i = \sum_{i=1}^n x_i y_i = \langle \mathbf{x}, \mathbf{y} \rangle.\end{aligned}$$

したがって次を得る:

命題 35.2.2. $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ のなす角を $\theta \in [0, \pi]$ とすると,

$$(1) \langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\| \cos \theta,$$

$$(2) \cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}},$$

$$(3) \langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2).$$

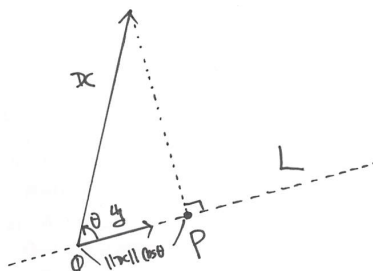
上式は, 内積の情報だけから $\cos \theta$ が得られること, すなわち \mathbf{x} と \mathbf{y} のなす角度 θ が分かることを述べている.

【補足】 命題 35.2.2 から $\cos \theta = \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{z}\|^2}{2\|\mathbf{x}\|\|\mathbf{y}\|}$ が得られるため, ベクトルの長さからも角 θ の情報が得られるように見える. しかしながら, “ベクトルの長さ” という概念を抽象化していくと (これをノルムという¹), 一般のノルムからは角の情報は得られない (角について論じてあまり有意義でない) ことが知られている. すなわち, 角 θ を知るための情報の本質は, 長さではなく内積にある.

本書では行列とベクトルの積 $A\mathbf{x}$ のことを, しばしば線形写像 T_A に \mathbf{x} を代入したもの, とみなしていた. この特別な場合として, $A = {}^t\mathbf{y}$ とすれば, 内積 $\langle \mathbf{y}, \mathbf{x} \rangle = {}^t\mathbf{y}\mathbf{x}$ も \mathbf{x} に何らかの操作をほどこす対応とみなせる. この操作の幾何学的意味は次の通りである:

¹ノルムの定義は 295 ページのコラム内にある.

$\|\mathbf{y}\| = 1$ (つまり \mathbf{y} の長さが 1) のとき, 命題 35.2.2 より $\langle \mathbf{y}, \mathbf{x} \rangle = \|\mathbf{x}\| \cos \theta$ である. したがって, $\mathbf{0}$ と \mathbf{y} を結ぶ直線 L に点 \mathbf{x} を射影した点 (点 \mathbf{x} から L に向けて引いた垂線の足) を P とすれば, L を数直線とみなしたときの点 P の値は $\langle \mathbf{y}, \mathbf{x} \rangle$ に等しい.



したがって, 点 P の位置を指すベクトルは

$$\langle \mathbf{y}, \mathbf{x} \rangle \mathbf{y} \tag{35.2.1}$$

で与えられる. つまり内積 $\langle \mathbf{y}, \mathbf{x} \rangle$ は, \mathbf{y} 方向に \mathbf{x} がどの程度伸びているかを表す量である. とくに, もっとも極端な場合として \mathbf{x} と \mathbf{y} が互いに直交するとき, ベクトル \mathbf{x} は \mathbf{y} 方向にまったく伸びていない. すなわち, $\langle \mathbf{y}, \mathbf{x} \rangle = 0$ となる. 以後, 式 (35.2.1) は何度も用いられるであろう.

35.3 正規直交基底

$\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ とする. \mathbf{x} と \mathbf{y} のなす角を θ とすれば, \mathbf{x} と \mathbf{y} が直交するとき $\cos \theta = 0$, つまり命題 35.2.2 より $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ である. そこで角 θ を表に出さない形で, あらためて直交性を定義しよう:

定義 35.3.1. $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ が $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ を満たすとき, \mathbf{x} と \mathbf{y} は直交する (orthogonal) という.

なお, \mathbf{x}, \mathbf{y} のいずれかが $\mathbf{0}$ の場合も $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ である. 長さをもつベクトルと零ベクトルが直交していると見るのはいささか滑稽かもしれないが, 以降では言葉のあやとして, \mathbf{x}, \mathbf{y} のいずれかが $\mathbf{0}$ の場合も含めて $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ が成り立つとき, \mathbf{x} と \mathbf{y} は直交すると呼ぼう.

命題 35.3.2. 組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ が互いに直交するならば, 線形独立である.

Proof. $\sum_{j=1}^n r_j \mathbf{u}_j = \mathbf{0}$ とする. 各 \mathbf{u}_i と $\mathbf{0}$ の内積を計算する. $i \neq j$ のときに $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ となることに注意すれば,

$$0 = \langle \mathbf{u}_i, \mathbf{0} \rangle = \left\langle \mathbf{u}_i, \sum_{j=1}^n r_j \mathbf{u}_j \right\rangle = \sum_{j=1}^n r_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle = r_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle.$$

$\mathbf{u}_i \neq \mathbf{0}$ ゆえ $\langle \mathbf{u}_i, \mathbf{u}_i \rangle \neq 0$ であるから, 上式は $r_i = 0$ を意味する. □

ベクトル $\mathbf{x} \in \mathbb{R}^n$ の長さがちょうど 1 のとき, $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2 = 1^2 = 1$ である. 長さが 1 で, 互いに直交するベクトルからなる基底を正規直交基底という:

定義 35.3.3. 部分空間 $W \subset \mathbb{R}^n$ の基底 $\mathbf{u}_1, \dots, \mathbf{u}_k \in W$ が $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$ を満たすとき, これを W の正規直交基底 (orthonormal basis) という.

部分空間 $W \subset \mathbb{R}^n$ において, 正規直交基底は必ず存在する (定理 35.4.1).

例 35.3.4. 標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{R}^n$ について次が成り立つ.

(1) $\mathbf{e}_1, \dots, \mathbf{e}_n$ は \mathbb{R}^n の正規直交基底である.

(2) $\langle \mathbf{e}_j, \mathbf{x} \rangle$ は、 \mathbf{x} の第 j 成分に等しい。したがって、各 $\mathbf{x} \in \mathbb{R}^n$ について、 $\mathbf{x} = \sum_{j=1}^n \langle \mathbf{e}_j, \mathbf{x} \rangle \mathbf{e}_j$ が成り立つ。実際、 $\mathbf{x} = {}^t(x_1, \dots, x_n)$ と成分表示すれば、

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} = \langle \mathbf{e}_1, \mathbf{x} \rangle \mathbf{e}_1 + \cdots + \langle \mathbf{e}_n, \mathbf{x} \rangle \mathbf{e}_n.$$

前節のコラムで述べた考察 (式 (35.2.1)) から、上の (2) に相当する性質が一般の正規直交基底についても成り立つことが示唆される。すなわち、

命題 35.3.5. 部分空間 $W \subset \mathbb{R}^n$ の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_k \in W$ および $\mathbf{x} \in W$ について、

$$\mathbf{x} = \sum_{i=1}^k \langle \mathbf{u}_i, \mathbf{x} \rangle \mathbf{u}_i.$$

Proof. $\mathbf{u}_1, \dots, \mathbf{u}_k$ が W を生成することから、 $\mathbf{x} = \sum_{j=1}^k r_j \mathbf{u}_j$ と書ける。このとき、各 $i = 1, \dots, k$ について

$$\langle \mathbf{u}_i, \mathbf{x} \rangle = \left\langle \mathbf{u}_i, \sum_{j=1}^k r_j \mathbf{u}_j \right\rangle = \sum_{j=1}^k r_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle = r_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle = r_i \cdot 1^2 = r_i.$$

□

備考 35.3.6. $\mathbf{u}_1, \dots, \mathbf{u}_n$ を \mathbb{R}^n の正規直交基底とし、 $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ を線形変換とする。このとき前命題より $f(\mathbf{u}_j) = \sum_{i=1}^n \langle \mathbf{u}_i, f(\mathbf{u}_j) \rangle \mathbf{u}_i$ であるから、この基底に関する f の表現行列を $A = [a_{ij}]$ とすれば、 $a_{ij} = \langle \mathbf{u}_i, f(\mathbf{u}_j) \rangle$ である。

35.4 シュミットの直交化法

有限個のベクトルで生成される線形空間において、その基底を取るアルゴリズムがあることを既に学んだ (命題 18.3.2)。こうして得られた基底をさらに変形して正規直交基底を作る手法があり、そのアルゴリズムをグラム-シュミットの直交化法 (**Gram-Schmidt orthonormalization**) という。本書ではこれを「シュミットの直交化法」と略記する。

$\mathbf{x} \neq \mathbf{0}$ と向きが等しい長さ 1 のベクトルは、 $\frac{1}{\|\mathbf{x}\|} \mathbf{x}$ で与えられる。実際、 $\left\| \frac{1}{\|\mathbf{x}\|} \mathbf{x} \right\| = \frac{1}{\|\mathbf{x}\|} \|\mathbf{x}\| = 1$ である。ベクトルを伸び縮みさせて長さを 1 にするこの操作を、これ以降で何度か用いる。

定理 35.4.1 (シュミットの直交化法)。部分空間 $W \subset \mathbb{R}^n$ の基底 $\mathbf{u}_1, \dots, \mathbf{u}_k \in W$ が与えられているとする。さらに、次のように再帰的に $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_k$ および $\mathbf{v}_1, \dots, \mathbf{v}_k$ を定める。

- $\tilde{\mathbf{v}}_1 := \mathbf{u}_1$ および $\mathbf{v}_1 := \frac{1}{\|\tilde{\mathbf{v}}_1\|} \tilde{\mathbf{v}}_1$ と定義する。
- 既に $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}$ が定まっているとき、

$$\tilde{\mathbf{v}}_i := \mathbf{u}_i - \sum_{j=1}^{i-1} \langle \mathbf{v}_j, \mathbf{u}_i \rangle \mathbf{v}_j$$
 および $\mathbf{v}_i := \frac{1}{\|\tilde{\mathbf{v}}_i\|} \tilde{\mathbf{v}}_i$ と定義する。

こうして得られた組 $\mathbf{v}_1, \dots, \mathbf{v}_k$ は W の正規直交基底である。

定理の証明の前に、 $n = 3$ の場合に上の $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ がどのようなベクトルになるか図示により確認しておこう。

(1) \mathbf{u}_1 を伸び縮みさせることで長さ 1 のベクトル \mathbf{v}_1 を得る。

- (2) 次に u_2 を用いて v_1 と直交するベクトルを探したい。そこで u_2 の v_1 方向への射影を考える。これは 35.2 節のコラムで述べた考察 (式 (35.2.1)) を通して、ベクトル $\langle v_1, u_2 \rangle v_1$ として得られる (図 35.1)。このとき、 u_2 と $\langle v_1, u_2 \rangle v_1$ の終点を結ぶベクトル $\tilde{v}_2 = u_2 - \langle v_1, u_2 \rangle v_1$ は v_1 に直交する。このベクトルを伸び縮みさせて長さを 1 にしたものが v_2 である。
- (3) いま、長さ 1 の直交するベクトル v_1, v_2 が得られている。今度は u_3 を用いてこれらと直交するベクトルを探したい。そこで、 u_3 の v_1 および v_2 方向への射影をそれぞれ考え、それらの和をとったベクトル $\sum_{j=1}^2 \langle v_j, u_3 \rangle v_j$ を考えれば、これは v_1, v_2 で張られる平面への u_3 の射影となる (図 35.2)²。このとき $\tilde{v}_3 = u_3 - \sum_{j=1}^2 \langle v_j, u_3 \rangle v_j$ は v_1, v_2 に直交し、これを伸び縮みさせて v_3 を得る。

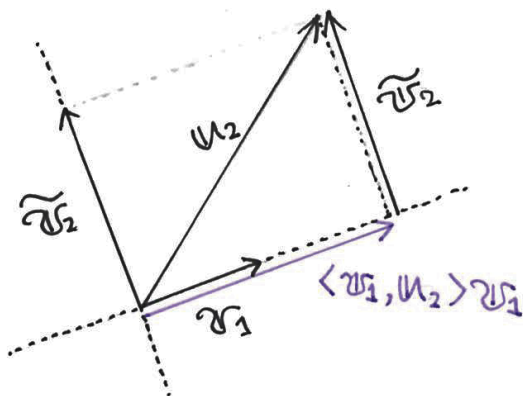


図 35.1: \tilde{v}_2 の探し方

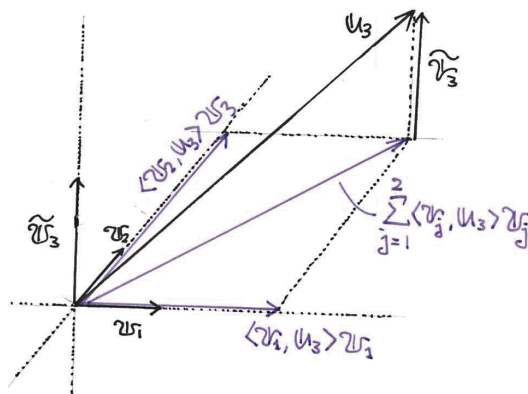


図 35.2: \tilde{v}_3 の探し方

上では図示によって直交する様子を示唆したものの、これは計算によっても確かめられることである。詳しくは次の証明を見よ。

定理 35.4.1 の証明. はじめに各 $i = 1, \dots, k$ について、 i に関する次の条件が成り立つことを帰納法により示そう。

- (1) _{i} $v_1, \dots, v_i \in \langle u_1, \dots, u_i \rangle$,
- (2) _{i} v_1, \dots, v_i は互いに直交する (ゆえに線形独立である)。

(1)₁ と (2)₁ が成立することは明らか。次に (1) _{$i-1$} と (2) _{$i-1$} が成り立つと仮定し、(1) _{i} と (2) _{i} を示そう。(1) _{$i-1$} より $v_1, \dots, v_{i-1} \in \langle u_1, \dots, u_{i-1} \rangle$ であり、これと \tilde{v}_i の定義から $\tilde{v}_i \in \langle u_1, \dots, u_i \rangle$ が成り立つ。また、 $\tilde{v}_i \notin \langle u_1, \dots, u_{i-1} \rangle$ ゆえ³、とくに $\tilde{v}_i \neq \mathbf{0}$ である。つまり $\|\tilde{v}_i\| \neq 0$ ゆえ、 v_i を定理の主張にあるように定義することができる。このとき (1) _{i} が成り立つことは明らか。

組 v_1, \dots, v_{i-1}, v_i が互いに直交することを示そう。そのためには、 \tilde{v}_i が各 v_1, \dots, v_{i-1} と直交することを示せば十分である。各 $l = 1, \dots, i-1$ について、

$$\begin{aligned} \langle v_l, \tilde{v}_i \rangle &= \left\langle v_l, u_i - \sum_{j=1}^{i-1} \langle v_j, u_i \rangle v_j \right\rangle = \langle v_l, u_i \rangle - \left\langle v_l, \sum_{j=1}^{i-1} \langle v_j, u_i \rangle v_j \right\rangle \\ &= \langle v_l, u_i \rangle - \sum_{j=1}^{i-1} \langle v_j, u_i \rangle \langle v_l, v_j \rangle = \langle v_l, u_i \rangle - \langle v_l, u_i \rangle \langle v_l, v_l \rangle \\ &= \langle v_l, u_i \rangle - \langle v_l, u_i \rangle \|v_l\|^2 = \langle v_l, u_i \rangle - \langle v_l, u_i \rangle \cdot 1^2 = 0. \end{aligned}$$

² $\mathbf{a} = (a, b, c)$ のとき、 \mathbf{a} を xy 平面に射影した点が $\mathbf{s} = (a, b, 0) = a(1, 0, 0) + b(0, 1, 0)$ と書けることを思い出そう。

³ 仮に $\tilde{v}_i \in \langle u_1, \dots, u_{i-1} \rangle$ とすれば $u_i = \tilde{v}_i + \sum_{j=1}^{i-1} \langle v_j, u_i \rangle v_j \in \langle u_1, \dots, u_{i-1} \rangle$ が導かれ、これは u_1, \dots, u_k の線形独立性に反する。

以上により, v_1, \dots, v_k は互いに直交し, ゆえに線形独立である (命題 35.3.2). また $\dim W = k$ ゆえ, これらは W の基底となる (命題 22.3.3). \square

上の定理において, あらかじめ与えられていた基底のうち最初のいくつか u_1, \dots, u_r (ただし $r \leq k$) について, これらが長さ 1 で互いに直交している場合は, 各 $i = 1, \dots, r$ について $\tilde{v}_i = u_i$ (したがって $v_i = u_i$) となる. ゆえに次が成り立つ:

系 35.4.2. 部分空間 $W \subset \mathbb{R}^n$ および, 互いに直交する長さ 1 のベクトルの組 $u_1, \dots, u_r \in W$ が与えられているとする. このとき, いくつかの W の元をこれらに加えて, W の正規直交基底にできる.

Proof. 組 u_1, \dots, u_r は線形独立ゆえ (命題 35.3.2), いくつかの W の元をこれらに加えて W の基底が取れる (命題 18.3.2). こうして得られた基底についてシュミットの直交化法を適用すれば, 求める正規直交基底を得る. \square

備考 35.4.3. 定理 35.4.1 における \tilde{v}_i の定義に現れたベクトル $\sum_{j=1}^{i-1} \langle v_j, u_i \rangle v_j$ は, 部分空間 $\langle v_1, \dots, v_{i-1} \rangle$ への直交射影 (定義 38.3.1(2)) に等しい.

35.5 直交補空間

斉次形連立 1 次方程式 $Ax = \mathbf{0}$ の解空間 W_A は, A のすべての行ベクトル (の転置) と直交するベクトル全体からなる集合である. とくに, W_A の各元は, A の行ベクトル (の転置) で生成される部分空間の各元とも直交する (補題 35.5.3). この例を抽象化し, 与えられたベクトル (の集合) と直交するベクトル全体について考えよう.

定義 35.5.1. 部分空間 $W \subset \mathbb{R}^n$ が与えられているとき, W のすべての元と直交するベクトルからなる集合を W の直交補空間 (orthogonal complement) とよび, これを W^\perp で表す. すなわち,

$$W^\perp = \{x \in \mathbb{R}^n \mid \text{各 } w \in W \text{ について } \langle x, w \rangle = 0\}.$$

W の直交補空間は, W の補集合ではない (互いに $\mathbf{0}$ を共有している). また, W^\perp は \mathbb{R}^n の部分空間である (各自確認せよ). さらに集合 $W \subset \mathbb{R}^n$ が部分空間でない場合において, 上式でもって W^\perp を定義しても W^\perp は \mathbb{R}^n の部分空間となる (補題 35.5.3 より, これは $\langle W \rangle$ の直交補空間 $\langle W \rangle^\perp$ に等しい).

例 35.5.2. (1) x 軸, y 軸, z 軸からなる座標空間 \mathbb{R}^3 において, x 軸と直交するベクトル全体の集合は yz 平面に一致する. また, yz 平面と直交するベクトル全体の集合は x 軸に一致する. すなわち, 標準ベクトル $e_1, e_2, e_3 \in \mathbb{R}^3$ について, $W = \{re_1 \mid r \in \mathbb{R}\}$, $V = \{re_2 + se_3 \mid r, s \in \mathbb{R}\}$ とすれば, $W^\perp = V$ および $V^\perp = W$. とくに $(W^\perp)^\perp = W$ である. 等式 $(W^\perp)^\perp = W$ は, 一般の部分空間 $W \subset \mathbb{R}^n$ においても成り立つ (命題 35.5.4(5)).

(2) 自明な部分空間 $\{\mathbf{0}\}$, $\mathbb{R}^n \subset \mathbb{R}^n$ について, $\{\mathbf{0}\}^\perp = \mathbb{R}^n$ および $(\mathbb{R}^n)^\perp = \{\mathbf{0}\}$ である.

補題 35.5.3. $x, w_1, \dots, w_k \in \mathbb{R}^n$ とする.

(1) x が各 w_1, \dots, w_k と直交するならば, x は $\langle w_1, \dots, w_k \rangle$ の任意の元とも直交する.

(2) $x \in \langle w_1, \dots, w_k \rangle^\perp \iff$ 各 $i = 1, \dots, k$ について $\langle x, w_i \rangle = 0$.

Proof. (1): $y \in \langle w_1, \dots, w_k \rangle$ とすれば, $y = \sum_{i=1}^k r_i w_i$ と書ける. このとき $\langle x, y \rangle = \langle x, \sum_{i=1}^k r_i w_i \rangle = \sum_{i=1}^k r_i \langle x, w_i \rangle = 0$.

(2): (\Rightarrow) は直交補空間の定義より明らか. (\Leftarrow) は (1) による. \square

命題 35.5.4. 部分空間 $W \subset \mathbb{R}^n$ について次が成り立つ.

- (1) $\mathbf{w}_1, \dots, \mathbf{w}_k$ を W の正規直交基底とする. いくつかのベクトルをこれらに加えた組 $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_1, \dots, \mathbf{v}_{n-k}$ が \mathbb{R}^n の正規直交基底となるならば, $\mathbf{v}_1, \dots, \mathbf{v}_{n-k}$ は W^\perp の正規直交基底である.
 【備考】 上の $\mathbf{v}_1, \dots, \mathbf{v}_{n-k}$ が実際に取れることは系 35.4.2 が保証している.
- (2) $\dim W^\perp = n - \dim W$.
- (3) $W \oplus W^\perp = \mathbb{R}^n$. すなわち, $\mathbf{w}_1, \dots, \mathbf{w}_k$ を W の基底, $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ を W^\perp の基底とすれば, これらを合わせた組 $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_1, \dots, \mathbf{v}_\ell$ は \mathbb{R}^n の基底である.
- (4) 部分空間 $V \subset \mathbb{R}^n$ が次の条件を満たすならば, $V = W^\perp$ である.
- (i) $W \oplus V = \mathbb{R}^n$,
 - (ii) $\mathbf{w} \in W$ かつ $\mathbf{v} \in V \implies \langle \mathbf{w}, \mathbf{v} \rangle = 0$.
 (すなわち, W の各元と V の各元はつねに直交する.)
- (5) $(W^\perp)^\perp = W$.

Proof. (1): $\mathbf{v}_1, \dots, \mathbf{v}_{n-k}$ が W^\perp を生成することを示せばよい. まず $\mathbf{v}_1, \dots, \mathbf{v}_{n-k} \in W^\perp$ を確認すべきだが, これは補題 35.5.3(1) による. 次に, 各 $\mathbf{x} \in W^\perp$ が $\mathbf{v}_1, \dots, \mathbf{v}_{n-k}$ の線形結合で書けることを示そう. \mathbf{x} が各 $\mathbf{w}_1, \dots, \mathbf{w}_k$ と直交することに注意すれば, 命題 35.3.5 より

$$\mathbf{x} = \sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i + \sum_{j=1}^{n-k} \langle \mathbf{v}_j, \mathbf{x} \rangle \mathbf{v}_j = \sum_{i=1}^k 0 \mathbf{w}_i + \sum_{j=1}^{n-k} \langle \mathbf{v}_j, \mathbf{x} \rangle \mathbf{v}_j = \sum_{j=1}^{n-k} \langle \mathbf{v}_j, \mathbf{x} \rangle \mathbf{v}_j.$$

(2): (1) より明らか.

(3): 系 33.5.4 における (4) と (3) の同値性から, \mathbb{R}^n の各元が W の元と W^\perp の元の和として, 一意に表せることを言えばよい. しかしこれは (1) より明らかである. 実際, (1) で取った基底について, $\mathbf{x} = \sum_{i=1}^k s_i \mathbf{w}_i$, $\mathbf{y} = \sum_{j=1}^{n-k} t_j \mathbf{v}_j$, $\mathbf{x}' = \sum_{i=1}^k s'_i \mathbf{w}_i$, $\mathbf{y}' = \sum_{j=1}^{n-k} t'_j \mathbf{v}_j$, $\mathbf{x} + \mathbf{y} = \mathbf{x}' + \mathbf{y}'$ とすれば, $(\mathbf{x} - \mathbf{x}') + (\mathbf{y} - \mathbf{y}') = \mathbf{0}$, すなわち $\sum_{i=1}^k (s_i - s'_i) \mathbf{w}_i + \sum_{j=1}^{n-k} (t_j - t'_j) \mathbf{v}_j = \mathbf{0}$ であり, 基底の線形独立性から $s_i = s'_i$ および $t_j = t'_j$, つまり $\mathbf{x} = \mathbf{x}'$ および $\mathbf{y} = \mathbf{y}'$ を得る.

(4): W および V の正規直交基底をとり, それぞれ $\mathbf{w}_1, \dots, \mathbf{w}_k$ および $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ とする. このとき (i) より, これらを合わせた $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_1, \dots, \mathbf{v}_\ell$ は \mathbb{R}^n の基底であり, また (ii) より, これは \mathbb{R}^n の正規直交基底である. したがって (1) より $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ は W^\perp の基底でもあり, $V = \langle \mathbf{v}_1, \dots, \mathbf{v}_\ell \rangle = W^\perp$.

(5): (3) および定義より次が成り立つ: (i) $W^\perp \oplus W = \mathbb{R}^n$ および, (ii) W^\perp の元と W の元は直交する. よって (4) を適用すれば, W は W^\perp の直交補空間である. \square

\mathbb{R}^n の任意の部分空間が, 連立 1 次方程式の解空間 W_A の形に表せることを確認しておく.

命題 35.5.5. (1) $\mathbf{w}_1, \dots, \mathbf{w}_m \in \mathbb{R}^n$ を列ベクトルとし, $A = {}^t[\mathbf{w}_1, \dots, \mathbf{w}_m]$ とすれば (これは (m, n) -行列である), $W_A = \langle \mathbf{w}_1, \dots, \mathbf{w}_m \rangle^\perp$.

(2) $V \subset \mathbb{R}^n$ を部分空間とすれば, $V = W_A$ を満たす行列 A が存在する.

Proof. (1): $\mathbf{x} \in W_A \iff$ 各 $i = 1, \dots, m$ について $\langle \mathbf{w}_i, \mathbf{x} \rangle = 0 \iff \mathbf{x} \in \langle \mathbf{w}_1, \dots, \mathbf{w}_m \rangle^\perp$ (補題 35.5.3(2)).

(2): V^\perp の基底を一つ取り, これを $\mathbf{w}_1, \dots, \mathbf{w}_m$ とする. $A := {}^t[\mathbf{w}_1, \dots, \mathbf{w}_m]$ とすれば, (1) より $W_A = \langle \mathbf{w}_1, \dots, \mathbf{w}_m \rangle^\perp = (V^\perp)^\perp = V$. \square

35.6 内積空間 (発展)

一般の線形空間 U において内積は定義されない. しかしながら, U に付加的な数学的構造があり, それを用いて標準的な内積と類似する性質を持つ概念を定義できる場合がある.

定義 35.6.1. U を体 \mathbb{R} 上の線形空間とする. 各 $\mathbf{x}, \mathbf{y} \in U$ に対して定められる関数 $\langle \mathbf{x}, \mathbf{y} \rangle$ が命題 35.1.4 のすべての性質をみたすとき, これを U 上の内積 (inner product) という. このとき, 空間と内積の組 $(U, \langle \cdot, \cdot \rangle)$ を内積空間 (inner product space) あるいは計量ベクトル空間という.

【補足】 英文における数学用語 metric は, 「計量」あるいは「距離」と和訳される. metric という語句は, ベクトル \mathbf{x} と \mathbf{y} の間の距離 $\|\mathbf{x} - \mathbf{y}\|$ や, あるいは距離を定めるためのより本質的な概念である内積のことを指す. metric を和訳する際, 前者を「距離」, 後者を「計量」と訳するのが慣例となっている.

どんな内積を考えているか誤解がない場合は, $(U, \langle \cdot, \cdot \rangle)$ を略して U と表す. 標準的な内積とそれ以外の内積をここでは区別し, 前者を $\langle \cdot, \cdot \rangle_{\mathbb{R}}$ と書くことにしよう. 組 $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\mathbb{R}})$ は内積空間である.

標準的な内積とは異なる \mathbb{R}^n 上の内積の例は 37.2 節で与える (例 37.2.5).

例 35.6.2. $X = [a, b] \subset \mathbb{R}$ を有界閉区間とし, X 上の実数値連続関数全体の集合を $C(X)$ とする. $C(X)$ は \mathbb{R} 上の線形空間であり, 各 $f, g \in C(X)$ に対して $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ と定めれば, これは $C(X)$ 上の内積となる.

内積空間 $(U, \langle \cdot, \cdot \rangle)$ の $\mathbf{0}$ でないベクトル \mathbf{x}, \mathbf{y} に対して, 命題 35.2.2(2) を念頭に, \mathbf{x} と \mathbf{y} のなす角 $\theta \in [0, \pi]$ の余弦を $\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}}$ と定めることで⁴, U の元の間にあたかも角が定まっていると見なすことができる. とくに, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{0}$ のとき, \mathbf{x} と \mathbf{y} は直交すると言う. これにより, 角 (とくに直交性) を用いた幾何的な議論の類推が U においても可能になる.

さて, 本章の 35.3 節以降で述べた $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\mathbb{R}})$ に関する性質は, 有限次元内積空間 $(U, \langle \cdot, \cdot \rangle)$ においてもすべて成り立つ. その証明も $\langle \cdot, \cdot \rangle_{\mathbb{R}}$ を $\langle \cdot, \cdot \rangle$ に置き換えるだけで直ちに得られる. 何故なら, これまでの証明において通常の内積の定義まで戻ること一度もなく, 命題 35.1.4 にある性質のみを用いて論じてきたからである. とくに, U においてもシュミットの直交化法は有効であり, また部分空間 $W \subset U$ について $(W^\perp)^\perp = W$ が成り立つ. なお, 無限次元内積空間においては $W \subset (W^\perp)^\perp$ であり, 等号が成り立たないことがある.

定義 35.6.3. 内積空間 $(U, \langle \cdot, \cdot \rangle)$ が与えられているとき, 各 $\mathbf{x} \in U$ に対して $\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ で定められる関数 $\|\cdot\| : U \rightarrow [0, \infty)$ を内積 $\langle \cdot, \cdot \rangle$ から定まるノルムという.

次の主張のうち, 劣加法性の証明は 38.1 節で与える.

命題 35.6.4. 内積から定まるノルム $\|\cdot\|$ は, 命題 35.1.5 に挙げたすべての性質を満たす.

25 章と 26 章において, 線形空間や線形写像の情報を \mathbb{R}^n の情報に還元する手法を論じた. これと同じことが内積空間においても適用できる. 次の命題は, 多くの文献では自明なものとして暗黙理に用いられるものであるが, 初学者の利便を考えてここに記す.

命題 35.6.5. U を体 \mathbb{R} 上の内積空間とし, $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の正規直交基底とする. このとき, 各 \mathbf{u}_i ($i = 1, \dots, n$) を $\mathbf{e}_i \in \mathbb{R}^n$ に写す線形同型 $F : U \rightarrow \mathbb{R}^n$ について次が成り立つ:

(1) F は内積を保つ. すなわち,

- (i) 各 $\mathbf{u}, \mathbf{v} \in U$ について $\langle \mathbf{u}, \mathbf{v} \rangle = \langle F(\mathbf{u}), F(\mathbf{v}) \rangle_{\mathbb{R}}$.
- (ii) 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}} = \langle F^{-1}(\mathbf{x}), F^{-1}(\mathbf{y}) \rangle$.

(2) 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する線形変換 $f, g : U \rightarrow U$ の表現行列を A, B とすれば,

- (i) 各 $\mathbf{u}, \mathbf{v} \in U$ について, $\langle f(\mathbf{u}), g(\mathbf{v}) \rangle = \langle AF(\mathbf{u}), BF(\mathbf{v}) \rangle_{\mathbb{R}}$,
- (ii) 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について, $\langle A\mathbf{x}, B\mathbf{y} \rangle_{\mathbb{R}} = \langle f(F^{-1}(\mathbf{x})), g(F^{-1}(\mathbf{y})) \rangle$,
- (iii) 各 $\mathbf{u}, \mathbf{v} \in U$ について, $\langle f(\mathbf{u}), f(\mathbf{v}) \rangle = \langle AF(\mathbf{u}), AF(\mathbf{v}) \rangle_{\mathbb{R}}$,

⁴ただし, この値が $[-1, 1]$ 区間に収まることを確認しておく必要がある. これは定理 38.1.2 による.

(iv) 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について, $\langle A\mathbf{x}, A\mathbf{y} \rangle_{\mathbb{R}} = \langle f(F^{-1}(\mathbf{x})), f(F^{-1}(\mathbf{y})) \rangle$,

(v) 各 $\mathbf{u}, \mathbf{v} \in U$ について, $\langle f(\mathbf{u}), \mathbf{v} \rangle = \langle AF(\mathbf{u}), F(\mathbf{v}) \rangle_{\mathbb{R}}$,

(vi) 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について, $\langle A\mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}} = \langle f(F^{-1}(\mathbf{x})), F^{-1}(\mathbf{y}) \rangle$.

Proof. (1): (i): $\mathbf{u} = \sum_{i=1}^n r_i \mathbf{u}_i$, $\mathbf{v} = \sum_{j=1}^n s_j \mathbf{u}_j$ とおくと, $F(\mathbf{u}) = {}^t(r_1, \dots, r_n)$, $F(\mathbf{v}) = {}^t(s_1, \dots, s_n)$ ゆえ $\langle F(\mathbf{u}), F(\mathbf{v}) \rangle_{\mathbb{R}} = \sum_{i=1}^n r_i s_i$. 一方,

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= \left\langle \sum_{i=1}^n r_i \mathbf{u}_i, \mathbf{v} \right\rangle = \sum_{i=1}^n r_i \langle \mathbf{u}_i, \mathbf{v} \rangle = \sum_{i=1}^n r_i \left\langle \mathbf{u}_i, \sum_{j=1}^n s_j \mathbf{u}_j \right\rangle = \sum_{i=1}^n r_i \left(\sum_{j=1}^n s_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n r_i s_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \sum_{i=1}^n \sum_{j=1}^n r_i s_j \delta_{ij} = \sum_{i=1}^n r_i s_i. \end{aligned}$$

ゆえに F は内積を保つ.

(ii): $\mathbf{u} := F^{-1}(\mathbf{x})$ および $\mathbf{v} := F^{-1}(\mathbf{y})$ として (i) を適用すると, $\mathbf{x} = F(\mathbf{u})$, $\mathbf{y} = F(\mathbf{v})$ ゆえ, $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}} = \langle F(\mathbf{u}), F(\mathbf{v}) \rangle_{\mathbb{R}} = \langle \mathbf{u}, \mathbf{v} \rangle = \langle F^{-1}(\mathbf{x}), F^{-1}(\mathbf{y}) \rangle$.

(2): $G = F$ として 193 ページの可換図式 (26.1.2) を書けば,

$$\begin{array}{ccc} U & \xrightarrow{f} & U & & U & \xrightarrow{g} & U \\ F \downarrow & & \downarrow F & & F \downarrow & & \downarrow F \\ \mathbb{R}^n & \xrightarrow{T_A} & \mathbb{R}^n & & \mathbb{R}^n & \xrightarrow{T_B} & \mathbb{R}^n \end{array}$$

であり, $F \circ f = T_A \circ F$ および $F \circ g = T_B \circ F$ に注意する. また T_A, T_B の定義から, $T_A(F(\mathbf{u})) = AF(\mathbf{u})$ および $T_B(F(\mathbf{u})) = BF(\mathbf{u})$ である.

(i): (1) より $\langle f(\mathbf{u}), g(\mathbf{v}) \rangle = \langle F(f(\mathbf{u})), F(g(\mathbf{v})) \rangle_{\mathbb{R}} = \langle T_A(F(\mathbf{u})), T_B(F(\mathbf{v})) \rangle_{\mathbb{R}} = \langle AF(\mathbf{u}), BF(\mathbf{v}) \rangle_{\mathbb{R}}$.

(ii): $\mathbf{u} = F^{-1}(\mathbf{x})$, $\mathbf{v} = F^{-1}(\mathbf{y})$ とおいて (i) を適用すると, $\mathbf{x} = F(\mathbf{u})$, $\mathbf{y} = F(\mathbf{v})$ ゆえ, $\langle A\mathbf{x}, B\mathbf{y} \rangle_{\mathbb{R}} = \langle AF(\mathbf{u}), BF(\mathbf{v}) \rangle_{\mathbb{R}} = \langle f(\mathbf{u}), g(\mathbf{v}) \rangle = \langle f(F^{-1}(\mathbf{x})), g(F^{-1}(\mathbf{y})) \rangle$.

(iii) と (iv), および (v) と (vi) は, (i) と (ii) の特別な場合 ($g = f$ あるいは $g = \text{id}_U$) ゆえ明らか. \square

第36章 直交行列と上三角化

行列 A と相似な行列の一つを B とすれば, ある基底による T_A の表現行列が B であるとみなすことができる. 本章では, 正規直交基底を取り換えることで表現行列を簡単な形にする方法, とくに固有値が実数に限る場合について, 表現行列を上三角にするための正規直交基底の取り方について論じる. その第一の応用として, 対称行列が直交行列を用いて対角化可能であることが分かる.

36.1 直交行列とその性質

\mathbb{R}^n 上の正規直交基底を並べた行列を直交行列 (orthogonal matrix) という. すなわち, 次の同値な条件のいずれか (したがってすべて) を満たす行列のことである:

命題 36.1.1. n 次実正方行列 A について, 次は同値である.

- (1) A の各列は \mathbb{R}^n 上の正規直交基底をなす,
- (2) A の各行は \mathbb{R}^n 上の正規直交基底をなす,
- (3) ${}^tA = A^{-1}$,
- (4) 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について $\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$.

Proof. tAA の (i, j) -成分は, A の i 列と j 列の内積に等しい. したがって (1) \Leftrightarrow (3) は明らかである (【注意】 (1) \Rightarrow (3) の証明では, ${}^tAA = E$ が直ちに分かる. さらにこの両辺の転置をとり, $A{}^tA = E$ を得る.) また, $A{}^tA$ の (i, j) -成分は, A の i 行と j 行の内積に等しい. ゆえに (2) \Leftrightarrow (3) も明らかである.

(1) \Rightarrow (4): 標準ベクトル $\mathbf{e}_1, \dots, \mathbf{e}_n$ について $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$, および,

$$\langle A\mathbf{e}_i, A\mathbf{e}_j \rangle = A \text{ の } i \text{ 列目と } j \text{ 列目の内積} = \delta_{ij}$$

に注意する. $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ が標準ベクトルの場合について等式が成り立つことから, それらの線形結合で書ける一般のベクトル $\mathbf{x} = {}^t(x_1, \dots, x_n)$ および $\mathbf{y} = {}^t(y_1, \dots, y_n)$ についても等式 $\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ が成り立つ. その形式的計算は次の通りである:

$$\begin{aligned} \langle A\mathbf{x}, A\mathbf{y} \rangle &= \left\langle A \sum_{i=1}^n x_i \mathbf{e}_i, A \sum_{j=1}^n y_j \mathbf{e}_j \right\rangle = \left\langle \sum_{i=1}^n x_i A\mathbf{e}_i, \sum_{j=1}^n y_j A\mathbf{e}_j \right\rangle = \sum_{i=1}^n x_i \left\langle A\mathbf{e}_i, \sum_{j=1}^n y_j A\mathbf{e}_j \right\rangle \\ &= \sum_{i=1}^n x_i \left(\sum_{j=1}^n y_j \langle A\mathbf{e}_i, A\mathbf{e}_j \rangle \right) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \delta_{ij} = \sum_{i=1}^n x_i y_i = \langle \mathbf{x}, \mathbf{y} \rangle. \end{aligned}$$

(4) \Rightarrow (1): \mathbf{x}, \mathbf{y} が標準ベクトルの場合に (4) を適用すれば, A の i 列目と j 列目の内積が $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$ に一致することが分かる. つまり (1) が成り立つ. \square

例 36.1.2. 13 ページのコラムで述べた θ 回転を表す行列 $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ は直交行列である.

備考 36.1.3 (発展). (1) 二点 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ の間の距離と, これらを $-\mathbf{y}$ 方向に平行移動した $\mathbf{x} - \mathbf{y}$ と $\mathbf{0}$ の間の距離は等しい. すなわち, \mathbf{x} と \mathbf{y} の間の距離は $\|\mathbf{x} - \mathbf{y}\|$ で与えられる. これを \mathbf{x} と \mathbf{y} のユークリッド距離 (Euclidean metric) という.

- (2) 命題 36.1.1 の条件 (4) は $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ($T_A(\mathbf{x}) = A\mathbf{x}$) が内積を保つ線形写像であることを意味する。また、ユークリッド距離は内積を通して定義されることから、条件 (4) は $\|\mathbf{x} - \mathbf{y}\| = \|T_A(\mathbf{x}) - T_A(\mathbf{y})\|$ を導く¹。すなわち、 T_A は距離も保つ写像である。

命題 36.1.4. P, Q を n 次直交行列とすれば次が成り立つ。

- (1) P^{-1} も直交行列である。 (2) PQ も直交行列である。

Proof. 命題 36.1.1 の条件 (3) を確認しよう。(1): ${}^t(P^{-1}) = {}^t(P) = P = (P^{-1})^{-1}$ 。(2): ${}^t(PQ) = {}^tQ {}^tP = Q^{-1}P^{-1} = (PQ)^{-1}$. \square

次の話題は本書の以降では用いない。

備考 36.1.5 (よりみち). 与えられた行列 P を直交行列 Q と上三角行列 R の積に表すとき、これを **QR 分解 (QR decomposition)** という。これはシュミットの直交化法から直ちに導かれる事実である。

- (1) 可逆行列の場合: $P = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ を n 次可逆行列とする。 \mathbb{R}^n の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ についてシュミットの直交化法を適用して正規直交基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ を得れば、定理 35.4.1 における証明で挙げた条件 (1) _{i} と (2) _{i} から、各 $i = 1, \dots, n$ について $\langle \mathbf{u}_1, \dots, \mathbf{u}_i \rangle = \langle \mathbf{v}_1, \dots, \mathbf{v}_i \rangle$ となる。とくに、 $\mathbf{u}_i \in \langle \mathbf{v}_1, \dots, \mathbf{v}_i \rangle$ であり、したがって正規直交基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ による基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合表示に現れる行列 (基底の変換行列) R は上三角である。いま、

$$[\mathbf{u}_1, \dots, \mathbf{u}_n] = [\mathbf{v}_1, \dots, \mathbf{v}_n]R$$

が成り立っている。ここで $Q = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ は直交行列であり、上式は $P = QR$ を意味する。すなわち、任意の可逆行列は直交行列と上三角行列の積に分解できる。

- (2) 一般の場合: $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ を (m, n) -行列とする。 m 次列ベクトルの組 $\mathbf{p}_1, \dots, \mathbf{p}_n$ の中から順番にベクトルを選んで $\langle \mathbf{p}_1, \dots, \mathbf{p}_n \rangle$ の基底 $\mathbf{p}_{n_1}, \dots, \mathbf{p}_{n_k}$ を取り²、これを改めて $\mathbf{u}_1, \dots, \mathbf{u}_k$ とおく。もし $k < m$ の場合は、これに加える形で \mathbb{R}^m の基底 $\mathbf{u}_1, \dots, \mathbf{u}_m$ を取る。このとき、 $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ の $A = [\mathbf{u}_1, \dots, \mathbf{u}_m]$ による線形結合表示に現れる行列を B とすれば (つまり $P = AB$)、 (m, n) -行列 B の形は、 $m < n$ および $m = n$, $m > n$ のそれぞれの場合において、次のようになる:

$$\left[\begin{array}{cccc|c} * & * & \cdots & * & \\ & * & \cdots & * & \\ & & \ddots & \vdots & * \\ O & & & * & \\ & & & * & \end{array} \right], \quad \left[\begin{array}{ccc} * & & \\ & * & * \\ O & \ddots & * \end{array} \right], \quad \left[\begin{array}{ccc} * & & * \\ & * & * \\ O & \ddots & * \\ \hline O & & \end{array} \right]. \quad (36.1.1)$$

さらに (1) により、 m 次可逆行列 A は、 m 次直交行列 Q と m 次上三角行列 \tilde{R} を用いて、 $A = Q\tilde{R}$ と書ける。このとき、 $P = AB = Q\tilde{R}B$ であり、 (m, n) -行列 $R = \tilde{R}B$ は、やはり式 (36.1.1) と同じ形をしている。こうして得られる $P = QR$ も QR 分解と呼ばれる。

- (3) $m > n$ の場合: ${}^tQ_{m,n}Q_{m,n} = E_n$ を満たす (m, n) -行列 $Q_{m,n}$ と n 次上三角行列 R_n を用いて $P = Q_{m,n}R_n$ と分解できる。実際、上の (2) の方法で $P = QR$ が得られるとき、 $k = m - n$ として、

$$Q = \left[Q_{m,n} \mid Q_{m,k} \right] \text{ および } R = \left[\begin{array}{c} R_n \\ O_{k,n} \end{array} \right] \text{ と分割すれば,}$$

$$P = \left[Q_{m,n} \mid Q_{m,k} \right] \left[\begin{array}{c} R_n \\ O_{k,n} \end{array} \right] = Q_{m,n}R_n + Q_{m,k}O_{k,n} = Q_{m,n}R_n.$$

また、 $Q_{m,n}$ の各列は互いに直交するゆえ ${}^tQ_{m,n}Q_{m,n} = E_n$ である。

【注意】 上の (1) と (2) いずれにおいても、直交行列 Q さえ求まれば上三角行列 R もすぐに求まる。実際、 $R = Q^{-1}P = {}^tQP$ とすればよい。

¹詳しくは命題 38.2.2 をみよ。

²命題 18.3.1 の証明で述べたように順番に取る。

36.2 直交行列による上三角化

正方行列 A が与えられているとき、可逆行列 P を用いて $P^{-1}AP$ を上三角行列にすることを、 A の上三角化 (triangulation) という。 $P^{-1}AP$ が上三角となるとき、本書では、これを A の上三角化と呼ぶ。ジョルダン標準形は上三角化の一例である。

体 \mathbb{R} 上の線形空間について論じているとき、29章の冒頭で述べた約束によれば、線形変換や正方行列 A の固有値はすべて実数である。この約束に反する形で、いわば言葉のあやとして、本章に限り次の表現を使うことを許されたい:

定義 36.2.1. 実正方行列 A の特性多項式 $\Phi_A(t) = 0$ が虚数解を持たないとき、 A の固有値はすべて実数であるという。

次の証明は、直交行列による行列の上三角化の計算手順を兼ねている。なお、対称行列の対角化 (上三角化) については、このような複雑な手順を踏む必要はない。対称行列の対角化にのみ興味がある者は、次の定理を事実として認めただうえで本節を飛ばして次節に進んでもよい。

定理 36.2.2 (シューア分解). A を n 次実正方行列とする。 A の固有値がすべて実数ならば、 $P^{-1}AP$ が上三角行列となるような直交行列 P が存在する。

Proof. n に関する帰納法で示す。 $n = 1$ のときは明らかである。 $n - 1$ 次の行列について定理の主張が成り立つと仮定する。 n 次正方行列 A の固有値がすべて実数であるとし、その固有値の一つ $\lambda \in \mathbb{R}$ および対応する固有ベクトル $\mathbf{x} \in \mathbb{R}^n$ をとる。スカラー倍することで $\|\mathbf{x}\| = 1$ であるとしてよい。さて、シュミットの直交化法により正規直交基底 $\mathbf{x}, \mathbf{q}_2, \dots, \mathbf{q}_n$ を取り、 $Q = [\mathbf{x}, \mathbf{q}_2, \dots, \mathbf{q}_n]$ とおく。 Q は直交行列であり、 $Q^{-1}AQ$ は次の形になる:

$$Q^{-1}AQ = \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \right) \begin{array}{c} \\ \\ B \\ \\ \end{array}.$$

【注意】 $Q^{-1}AQ$ は、基底 $\mathbf{x}, \mathbf{q}_2, \dots, \mathbf{q}_n$ に関する T_A の表現行列であること、および $A\mathbf{x} = \lambda\mathbf{x}$ より上の表示を得る。あるいは、左辺の1列目 $= Q^{-1}AQ\mathbf{e}_1 = Q^{-1}A\mathbf{x} = Q^{-1}\lambda\mathbf{x} = \lambda Q^{-1}\mathbf{x} = \lambda\mathbf{e}_1$ と計算で確認してもよい。ここで最後の等式において $Q\mathbf{e}_1 = \mathbf{x}$ より $\mathbf{e}_1 = Q^{-1}\mathbf{x}$ であることを用いた。

A と $Q^{-1}AQ$ の固有多項式は等しく (系 28.4.5), ゆえに $\Phi_A(t) = \Phi_{Q^{-1}AQ}(t) = (t - \lambda)\Phi_B(t)$ である。仮定より、 $\Phi_A(t)$ は実係数1次式に因数分解できるから $\Phi_B(t)$ もそうであり、ゆえに帰納法の仮定より $n - 1$ 次正方行列 B は上三角化できる。すなわち、 $D = R^{-1}BR$ が上三角となるような $n - 1$ 次直交行列 R が存在する。このとき

$$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \right) \begin{array}{c} \\ \\ R \\ \\ \end{array} \text{ は直交行列であり、したがって } P := Q \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \right) \begin{array}{c} \\ \\ R \\ \\ \end{array}$$

も直交行列である (命題 36.1.4). また, $P^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R^{-1} & \\ 0 & & & \end{array} \right) Q^{-1}$ であり³,

$$\begin{aligned} P^{-1}AP &= \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R^{-1} & \\ 0 & & & \end{array} \right) Q^{-1}AQ \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R & \\ 0 & & & \end{array} \right) \\ &= \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R^{-1} & \\ 0 & & & \end{array} \right) \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array} \right) \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R & \\ 0 & & & \end{array} \right) \\ &= \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & R^{-1}B & \\ 0 & & & \end{array} \right) \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R & \\ 0 & & & \end{array} \right) = \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & R^{-1}BR & \\ 0 & & & \end{array} \right) = \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & D & \\ 0 & & & \end{array} \right) \end{aligned}$$

は上三角行列である. □

上三角化の計算は, 一般には上の証明と同様にして行うことになる. 噛み砕いて説明すれば次のようになる:

まず $Q_n^{-1}AQ_n = \left(\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & B_{n-1} & \\ 0 & & & \end{array} \right)$ となる λ_1 および直交行列 Q_n と B_{n-1} を求める⁴. ここで,

$R_{n-1}^{-1}B_{n-1}R_{n-1}$ を上三角にするような直交行列 R_{n-1} が求まれば, 上の証明にあるように,

$P := Q_n \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R_{n-1} & \\ 0 & & & \end{array} \right)$ とすることで A の上三角化 $P^{-1}AP$ が求まる. そこで, R_{n-1} を求めるために,

$Q_{n-1}^{-1}B_{n-1}Q_{n-1} = \left(\begin{array}{c|ccc} \lambda_2 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & B_{n-2} & \\ 0 & & & \end{array} \right)$ となる λ_2 および直交行列 Q_{n-1} と B_{n-2} を求める. R_{n-1} を得る

には $R_{n-2}^{-1}B_{n-2}R_{n-2}$ を上三角にするような直交行列 R_{n-2} を求めて, $R_{n-1} := Q_{n-1} \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R_{n-2} & \\ 0 & & & \end{array} \right)$

とすればよい. そこで, R_{n-2} を求めるために…というふうに計算を繰り返していくしかない. これを続けていくと最終的には B_2 が得られたのちに $Q_2^{-1}B_2Q_2$ を上三角にするような 2×2 の直交行列 $Q_2 = R_2 = [r_{ij}]$

³これは $P^{-1} = {}^tP$ および $R^{-1} = {}^tR$, $Q^{-1} = {}^tQ$ による. あるいは実際に P との積が E になることを確認してもよい.

⁴ Q_n を求める方法は証明の中で既にかいた通りである. なお, $Q_n^{-1} = {}^tQ_n$ ゆえ, Q_n^{-1} を改めて求める必要はない点に注意せよ. また, 行列の積 $Q_n^{-1}AQ_n$ を計算することで B_{n-1} が求められる.

が求まり, これと既に得られている Q_n, Q_{n-1}, \dots, Q_3 を用いて,

$$R_3 = Q_3 \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & r_{11} & r_{12} \\ 0 & r_{21} & r_{22} \end{array} \right), \quad R_{k+1} = Q_{k+1} \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R_k & \\ 0 & & & \end{array} \right), \quad P = Q_n \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R_{n-1} & \\ 0 & & & \end{array} \right)$$

と再帰的に R_4, R_5, \dots, R_{n-1} および P が求められるのである.

なお, 固有空間の次元が 1 を超える場合は, 上で説明した手順の一部を次の方法で略せる場合がある.

備考 36.2.3. 定理 36.2.2 の証明において, $\dim W(\lambda, A) = m$ の場合は, $W(\lambda, A)$ の正規直交基底として $\mathbf{x}_1, \dots, \mathbf{x}_m$ をとり, これに加える形で \mathbb{R}^n の正規直交基底 $\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{q}_{m+1}, \dots, \mathbf{q}_n$ を取れば, これらを並べた直交行列 Q について次式を得る:

$$Q^{-1}AQ = \left(\begin{array}{c|c} \lambda E_m & * \\ \hline O & B \end{array} \right),$$

ここで, B は $n - m$ 次正方行列である. したがって, B を上三角化する直交行列 R が得られれば, 直交行列 $P = Q \left(\begin{array}{c|c} E_m & O \\ \hline O & R \end{array} \right)$ について, $P^{-1}AP$ は上三角となる. このように A の固有空間で次元が 1 を超えるものがある場合は, その正規直交基底を取ることににより, 上の再帰的計算のステップを短くできる. また, B の上三角化においても同様のことが言える.

36.3 直交行列による対称行列の対角化

定義 36.3.1. ${}^tA = A$ をみたす正方行列を対称行列 (symmetric matrix) という.

本節では, 実対称行列が直交行列によって対角化可能であることを示す.

対称行列とは何か

A が対称行列であるとき、線形変換 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ がどのような意味を持っているかを、対称行列の定義それ自体から直接に説明することは(少なくとも筆者には)容易ではない。しかしながら本節で学ぶ事実: 「実対称行列 A は、ある直交行列 P によって対角化可能である」を認めれば、線形変換 T_A は P の各列の方向に伸び縮みする写像であると説明することができる。一方、直交行列により対角化される行列は対称行列に限られる(系 36.3.6 (2))。別の言い方をすれば、対称行列を表現行列とするような線形変換 $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ は次のようにして構成されるものに限られる:

あらかじめ \mathbb{R}^n に正規直交基底(つまり直交座標) $\mathbf{p}_1, \dots, \mathbf{p}_n$ を与えておき、各座標軸 $\langle \mathbf{p}_i \rangle$ 上の点については f で λ_i 倍されると定め ($f(\mathbf{p}_i) = \lambda_i \mathbf{p}_i$)、座標軸以外の点の f による行き先は基底の行き先から線形性を通して自動的に定まるものとする。このようにして f を定めると、標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ による f の表現行列 A は対称行列である。実際、 $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ は直交行列であり、基底 $\mathbf{p}_1, \dots, \mathbf{p}_n$ による $f = T_A$ の表現行列 $P^{-1}AP$ は、 f の定め方から $\lambda_1, \dots, \lambda_n$ を対角成分とする対角行列であり、系 36.3.6 (2) より A は対称行列である。

他方で、行列の各成分を何らかの情報(データ)とみなす立場からは、対称行列が様々な分野に顔を出すであろうことが想像できる。何故なら、二つのパラメータ i, j に関する何らかデータを a_{ij} とするとき、 $a_{ij} = a_{ji}$ を満たすことがしばしば起こるからである。例えば、有限個の位置情報 $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^n$ が与えられているとき、 \mathbf{x}_i と \mathbf{x}_j のユークリッド距離を a_{ij} とすれば $a_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\| = \|\mathbf{x}_j - \mathbf{x}_i\| = a_{ji}$ である。この a_{ij} の対称性は距離の対称性から来しているわけだが、一般に、与えられた対象から二つを取り出してデータを作るとき、二つを選ぶ順番によらないデータは対称となる。しばしば我々は対称な関数(データ)を好んで扱う傾向にあり、それゆえに対称行列の応用の範囲は広い。

ちなみに大学初年次で学ぶ数学に現れる対称行列の例としては、微積分においては C^2 級関数のヘッセ行列がある。あるいはこれと関係が深いものとして、2次形式を定める正方行列も対称である(としてよい)。これらの関係については37章で紹介しよう。

実対称行列の固有値がすべて実数であることを示すために、複素数と複素行列について補足しておく。複素数 z の共役を \bar{z} と書く。各 $z, w \in \mathbb{C}$ について次が成り立つのであった:

$$\overline{z+w} = \bar{z} + \bar{w}, \quad \overline{zw} = \bar{z}\bar{w}, \quad (\bar{z})^n = \overline{z^n}, \quad \bar{z}z \geq 0.$$

複素 (m, n) -行列 $A = [a_{ij}]$ に対して、 (m, n) -行列 \bar{A} を $\bar{A} := [\bar{a}_{ij}]$ と定める。このとき、次が成り立つ:

命題 36.3.2. $\overline{A+B} = \bar{A} + \bar{B}$, $\overline{AB} = \bar{A}\bar{B}$, $\overline{A^n} = \bar{A}^n$.

補題 36.3.3. n 次複素正方行列 A の固有値の一つを $\lambda \in \mathbb{C}$ 、その固有ベクトルの一つを $\mathbf{x} \in \mathbb{C}^n$ とする。このとき、

- (1) $\bar{\lambda}$ は \bar{A} の固有値であり、 $\bar{\mathbf{x}}$ は $\bar{\lambda}$ に関する \bar{A} の固有ベクトルである。
- (2) A が実正方行列ならば、 $\bar{\mathbf{x}}$ は固有値 $\bar{\lambda}$ に関する A の固有ベクトルである。

Proof. (1): $\mathbf{x} \neq \mathbf{0}$ ゆえ $\bar{\mathbf{x}} \neq \mathbf{0}$ である。また、 $\bar{A}\bar{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$ 。(2): $\bar{A} = A$ と (1) より明らか。□

次の証明は計算が少々ややこしく感じるかもしれない。39章でエルミート内積の性質についていくつか学ぶと、証明は明快になる(定理 39.6.5 の証明を見よ)。

定理 36.3.4. 実対称行列 A の固有値はすべて実数である。すなわち、 $\Phi_A(t)$ は実係数1次式の積に因数分解できる。

Proof. $\lambda \in \mathbb{C}$ を複素行列としての A の固有値とし、 $\lambda = \bar{\lambda}$ を示そう。 λ に関する A の固有ベクトル $\mathbf{z} = {}^t(z_1, \dots, z_n) \in \mathbb{C}^n$ を取れば、前補題 (2) により $\bar{\mathbf{z}} \in \mathbb{C}^n$ は、 A の固有値 $\bar{\lambda}$ に関する A の固有ベクトルである(つまり $A\bar{\mathbf{z}} = \bar{\lambda}\bar{\mathbf{z}}$)。複素列ベクトル \mathbf{a}, \mathbf{b} について $\langle \mathbf{a}, \mathbf{b} \rangle := {}^t\bar{\mathbf{a}}\mathbf{b}$ と定めると(エルミート内積)、

$$\lambda \langle \mathbf{z}, \mathbf{z} \rangle = \lambda {}^t\bar{\mathbf{z}}\mathbf{z} = {}^t\bar{\mathbf{z}}(\lambda\mathbf{z}) = {}^t\bar{\mathbf{z}}A\mathbf{z} = {}^t\bar{\mathbf{z}}{}^tA\mathbf{z} = {}^t(A\bar{\mathbf{z}})\mathbf{z} = {}^t(\bar{\lambda}\bar{\mathbf{z}})\mathbf{z} = \bar{\lambda}{}^t\bar{\mathbf{z}}\mathbf{z} = \bar{\lambda}\langle \mathbf{z}, \mathbf{z} \rangle.$$

$z \neq 0$ ゆえ $\langle z, z \rangle = \bar{z}_1 z_1 + \cdots + \bar{z}_n z_n (\geq 0)$ は正数であり, したがって $\lambda = \bar{\lambda}$. すなわち λ は実数である. □

補題 36.3.5. n 次対称行列 A について次が成り立つ.

- (1) n 次正方行列 P について, ${}^t P A P$ は対称行列である.
- (2) n 次直交行列 P について, $P^{-1} A P$ は対称行列である.

Proof. (1) : ${}^t({}^t P A P) = {}^t P {}^t A {}^t({}^t P) = {}^t P A P$. (2) : $P^{-1} = {}^t P$ および (1) より明らか. □

系 36.3.6. 実正方行列 A について次は同値である.

- (1) A は対称行列である.
- (2) A は, ある直交行列 P を用いて対角化可能である.

Proof. (1) \Rightarrow (2): 定理 36.2.2 および 36.3.4 より, A は直交行列 P を用いて上三角化できる. このとき $P^{-1} A P$ は上三角かつ対称ゆえ対角行列である.

(2) \Rightarrow (1): $D = P^{-1} A P$ を対角行列とすれば, $A = P D P^{-1} = P D {}^t P$ であり, このとき ${}^t A = {}^t(P D {}^t P) = {}^t({}^t P) {}^t D {}^t P = P D P^{-1} = A$. すなわち, A は対称である. □

次は, 直交行列による対称行列の対角化可能性を認めれば, 直ちに分かることでもある (補題 39.7.8).

問題 36.3.7. A を実対称行列, $\lambda, \gamma \in \mathbb{R}$ を A の異なる固有値とすれば, $W(\lambda, A)$ の各元と $W(\gamma, A)$ の各元が直交することを示せ.

解答例: $\mathbf{x} \in W(\lambda, A)$, $\mathbf{y} \in W(\gamma, A)$ とすれば, $\lambda \langle \mathbf{x}, \mathbf{y} \rangle = \langle A \mathbf{x}, \mathbf{y} \rangle = \langle {}^t(A \mathbf{x}), \mathbf{y} \rangle = \langle {}^t \mathbf{x}, {}^t A \mathbf{y} \rangle = \langle \mathbf{x}, {}^t A \mathbf{y} \rangle = \langle \mathbf{x}, A \mathbf{y} \rangle = \langle \mathbf{x}, \gamma \mathbf{y} \rangle = \gamma \langle \mathbf{x}, \mathbf{y} \rangle$. ゆえに $(\lambda - \gamma) \langle \mathbf{x}, \mathbf{y} \rangle = 0$ であり, $\lambda - \gamma \neq 0$ より $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

上の計算でも用いたこととして, 一般に次が成り立つ:

命題 36.3.8. $A \in M_n(\mathbb{R})$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について, $\langle A \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, {}^t A \mathbf{y} \rangle$.

ここで直交行列による対称行列の対角化の手順をまとめておく. 系 36.3.6(1) では対角化可能性の証明において上三角化を用いたものの, 実際の計算においては上三角化の節で述べた複雑な手順を踏む必要はなく, 次のようにすればよい.

直交行列による対称行列の対角化の手順

A を n 次実対称行列とする.

- (1) A の固有値をすべて求める. 定理 36.3.4 より固有値はすべて実数である. したがって固有多項式は $\Phi_A(t) = \prod_{i=1}^k (t - \lambda_i)^{n_i}$ なる形をしている. ここで, 各 $\lambda_i \in \mathbb{R}$ は相異なるとする.
- (2) 各固有値 λ_i について $W(\lambda_i, A)$ の基底 $\mathbf{u}_1, \dots, \mathbf{u}_{n_i}$ を求め, さらにシュミットの直交化法を用いて, $W(\lambda_i, A)$ の正規直交基底を得る.
- (3) 各固有値ごとに求めた正規直交基底をすべて並べたものを $\mathbf{p}_1, \dots, \mathbf{p}_n$ とすれば (A は対角化可能であるから, ちょうど n 個分のベクトルが並ぶ), これは \mathbb{R}^n の正規直交基底になる (問題 36.3.7). ゆえに $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ は直交行列である.
- (4) このとき $D = P^{-1} A P$ は対角行列であり, その (i, i) -成分は \mathbf{p}_i に対応する固有値に等しい. なお P^{-1} の成分表示は, $P^{-1} = {}^t P$ ゆえ直ちに求まる.

例題 36.3.9. 次の実対称行列 A について, 直交行列を用いて対角化せよ.

$$(1) A = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix}.$$

解答例. $\Phi_A(t) = t^2 - 1/4$ ゆえ, 固有値は $\pm 1/2$ である.

- $\lambda = 1/2$ の固有空間:

$$\frac{1}{2}E - A = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} \xrightarrow{\text{簡約化}} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

したがって $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ が $W(1/2, A)$ の基底であり, $\frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ がその正規直交基底である.

- $\lambda = -1/2$ の固有空間:

$$-\frac{1}{2}E - A = \begin{bmatrix} -1/2 & -1/2 \\ -1/2 & -1/2 \end{bmatrix} \xrightarrow{\text{簡約化}} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

したがって $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ が $W(-1/2, A)$ の基底であり, $\frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ がその正規直交基底である.

以上より $P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ は直交行列であり, $P^{-1}AP = \begin{bmatrix} 1/2 & 0 \\ 0 & -1/2 \end{bmatrix}$.

$$(2) \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

解答例. A の成分表示から固有ベクトルが何かすぐに気づけることが望ましいが, ここでは上で与えた手順に従って解いてみよう.

$$\Phi_A(t) = \begin{vmatrix} t-1 & -1 & 0 \\ -1 & t-1 & 0 \\ 0 & 0 & t-2 \end{vmatrix} = \begin{vmatrix} t-1 & -1 \\ -1 & t-1 \end{vmatrix} (t-2) = ((t-1)^2 - (-1)^2)(t-2) = t(t-2)^2.$$

よって固有値は 0 と 2 である.

- $\lambda = 2$ の固有空間:

$$0E - A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{簡約化}} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに $W(2, A)$ の基底として $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ が取れる. これらの長さを 1 にした正規直交基底

は $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ である.

- $\lambda = 0$ の固有空間:

$$0E - A = \begin{bmatrix} -1 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \xrightarrow{\text{簡約化}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

したがって $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ が $W(0, A)$ の基底であり, $\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ がその正規直交基底である.

以上より, $P = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \end{bmatrix}$ とおけばこれは直交行列であり,

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

第37章 2次形式の分類(発展)

本章の主題は、2次形式と呼ばれる特殊な2次多項式を変数変換によって簡単な形に変形し、その挙動を分類することである。この分類には、直交行列を用いた対称行列の対角化が主要な役割を果たす。

特殊な多項式を分類することの意義に疑念を抱く初学者がいるかもしれない。しかし、本章で紹介する手法は、多変数関数の極値判定や2次曲線の分類において、本質的に用いられるものである(37.6節および37.5節)。

37.1 対称行列が定める2次形式

定義 37.1.1. 次数¹がちょうど2の項のみからなる $\mathbf{x} = (x_1, \dots, x_n)$ の多項式を2次形式 (quadratic form) と呼ぶ。すなわち、次式で表される n 変数2次多項式のことである:

$$f(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j. \quad (\text{ただし, 各 } a_{ij} \in \mathbb{R} \text{ は定数})$$

上では変数ベクトルに列ベクトルを用いたが、行列の積による表示との関係で、これ以降では変数ベクトル \mathbf{x} に列ベクトルを用いる。ただし、 $f({}^t(x_1, \dots, x_n))$ という表示に限っては、転置記号を略して $f(x_1, \dots, x_n)$ と書くことと約束する。

任意の2次形式は、次のように行列を用いて表すことができる:

定義 37.1.2. n 次実正方行列 $A = [a_{ij}]$ に対して $q_A(\mathbf{x}) := {}^t \mathbf{x} A \mathbf{x}$ と定め、これを A が定める2次形式という。

この式が2次形式であることは、次のように確認できる:

$$\begin{aligned} {}^t \mathbf{x} A \mathbf{x} &= [x_1, x_2, \dots, x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= [x_1, x_2, \dots, x_n] \begin{bmatrix} \sum_{j=1}^n a_{1j} x_j \\ \sum_{j=1}^n a_{2j} x_j \\ \vdots \\ \sum_{j=1}^n a_{nj} x_j \end{bmatrix} = \sum_{i=1}^n \left(x_i \sum_{j=1}^n a_{ij} x_j \right) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j. \end{aligned}$$

いまの計算を改めて命題として述べておこう。

命題 37.1.3. n 次実正方行列 $A = [a_{ij}]$ について $q_A(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$.

(i, j) -成分と (j, i) -成分の値を均等に分け直すことにより、上の A としては対称行列のみを考えれば十分である。実際、

命題 37.1.4. 任意の2次形式は、対称行列 B を用いて q_B の形に表せる。

¹ n 変数単項式 $ax_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ の次数を $\sum_{i=1}^n k_i$ と定める。一般に、次数がちょうど k の項のみからなる多項式を k 次形式 (k -form) という。ここで、0 に値を取る定数多項式は0次形式であるが、これを k 次形式でもあると形式的に約束する。

Proof. 2次形式 $f(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j$ に対して, $b_{ij} := \frac{a_{ij} + a_{ji}}{2}$ とおくと, n 次正方行列 $B = [b_{ij}]$ は対称である. このとき

$$\begin{aligned} f(\mathbf{x}) &= \sum_{1 \leq i < j \leq n} (a_{ij}x_i x_j + a_{ji}x_j x_i) + \sum_{i=1}^n a_{ii}x_i^2 = \sum_{1 \leq i < j \leq n} (a_{ij} + a_{ji})x_i x_j + \sum_{i=1}^n \frac{a_{ii} + a_{ii}}{2} x_i^2 \\ &= \sum_{1 \leq i < j \leq n} 2b_{ij}x_i x_j + \sum_{i=1}^n b_{ii}x_i^2 = \sum_{1 \leq i < j \leq n} (b_{ij}x_i x_j + b_{ji}x_j x_i) + \sum_{i=1}^n b_{ii}x_i^2 = q_B(\mathbf{x}). \end{aligned}$$

□

以下, 2次形式 q_A について論じるとき, とくに断りが無い場合は A は実対称行列であると約束する.

例 37.1.5. $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ が定める 2次形式は,

$$\begin{aligned} q_A \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + bxy + byx + cy^2 \\ &= ax^2 + 2bxy + cy^2. \end{aligned}$$

37.2 2次形式の分類

2次形式 q の性質として, $q(\mathbf{0}) = 0$ が常に成り立つ. またベクトルを伸び縮みさせても代入した値の正負は変わらない:

補題 37.2.1. 2次形式 q , および $\mathbf{x} \in \mathbb{R}^n$, $t \in \mathbb{R} \setminus \{0\}$ について, $q(\mathbf{x})$ と $q(t\mathbf{x})$ は共に 0 であるか, そうでなければこれらの符号は等しい.

Proof. $q(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j$ とすれば, $q(t\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}t^2 x_i x_j = t^2 q(\mathbf{x})$. $t^2 > 0$ ゆえ, これらは共に 0 であるか, あるいは符号は等しい. □

定義 37.2.2. q を 2次形式とする.

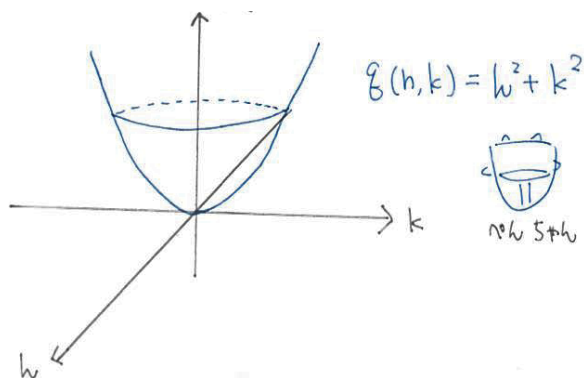
- (1) q が $\mathbf{x} = \mathbf{0}$ において唯一の最小値を取るとき, q は正定値 (positive definite) あるいは正の定符号であるという.
- (2) q が $\mathbf{x} = \mathbf{0}$ において最小値を取るとき, q は半正定値 (positive-semidefinite) であるという. q が正定値ならば半正定値である.
- (3) q が $\mathbf{x} = \mathbf{0}$ において唯一の最大値を取るとき, q は負定値 (negative definite) あるいは負の定符号であるという. これは $-q$ が正定値であることと同値である.
- (4) q が $\mathbf{x} = \mathbf{0}$ において最大値を取るとき, q は半負定値 (negative-semidefinite) であるという. これは $-q$ が半正定値であることと同値である. q が負定値ならば半負定値である.
- (5) q が上のいずれでもないとき, q は不定値 (indefinite) あるいは不定符号であるという. つまり半正定値でも半負定値でもないこと, すなわち, 次を満たすことである:

$$q(\mathbf{x}_1) < 0 \text{ および } q(\mathbf{x}_2) > 0 \text{ を満たす } \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n \text{ が存在する.}$$

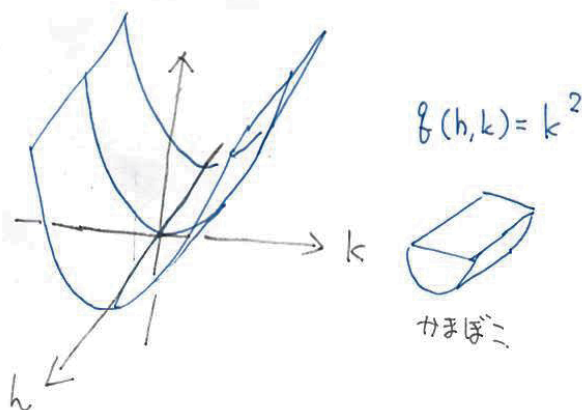
q が正定値であるとき, q は $\mathbf{x} = \mathbf{0}$ 以外の点で必ず正の値を取る. また q が半正定値であるとき, 任意の $\mathbf{x} \in \mathbb{R}^n$ について $q(\mathbf{x}) \geq 0$ である.

例 37.2.3. 2変数の場合の例を挙げよう.

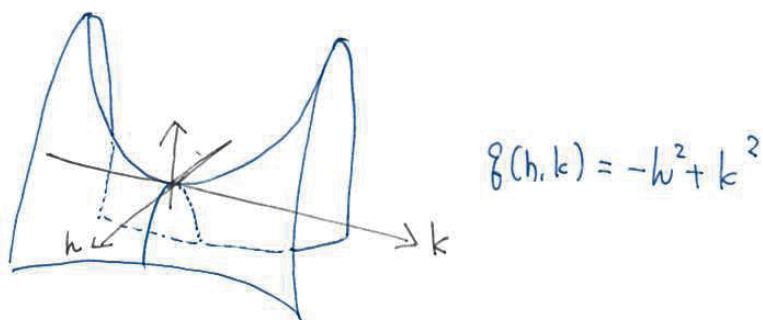
(1) 正定値 (正の定符号)



(2) 半正定値



(3) 不定値 (不定符号)



3変数以上の場合においても, 次の形の2次形式においては分類が容易にできる.

例 37.2.4. 2次形式 $q(x_1, \dots, x_n) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$ について次が成り立つ.

- (1) $\lambda_1, \dots, \lambda_n$ がすべて正数である $\iff q$ は正定値である.
- (2) $\lambda_1, \dots, \lambda_n$ がすべて0以上である $\iff q$ は半正定値である.
- (3) $\lambda_1, \dots, \lambda_n$ がすべて負数である $\iff q$ は負定値である.
- (4) $\lambda_1, \dots, \lambda_n$ がすべて0以下である $\iff q$ は半負定値である.

(5) $\lambda_1, \dots, \lambda_n$ の中に正数のものと負数のものがそれぞれ存在する $\iff q$ は不定符号である.

Proof. (1): (\Rightarrow): $\mathbf{x} = {}^t(x_1, \dots, x_n) \neq \mathbf{0}$ とすれば, $x_i \neq 0$ を満たす $i = 1, \dots, n$ が取れる. このとき, $q(\mathbf{x}) \geq \lambda_1 x_i^2 > 0$. (\Leftarrow): 標準ベクトルについて, $0 < q(\mathbf{e}_i) = \lambda_i$.

(2) から (4) は, いまと同様の議論で示せるゆえ略す. また, (2) と (4) のいずれでもない場合として, (5) の同値性も直ちに導かれる. \square

$q(x_1, \dots, x_n) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$ なる形の 2 次形式を対角形式 (diagonal form) という. 次節において, 任意の 2 次形式が適当な変数変換により対角形式に書き換えられることを示す.

例 37.2.5 (よりみち). A を n 次実対称行列とし, かつ q_A が正定値であるとする. このとき, 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について $\langle \mathbf{x}, \mathbf{y} \rangle_A := {}^t \mathbf{x} A \mathbf{y}$ と定めれば, これは \mathbb{R}^n における内積となる. とくに, $A = E$ の場合が標準的な内積である. 逆に, \mathbb{R}^n 上の任意の内積は, この形に表すことができる. 実際, $\langle \cdot, \cdot \rangle$ を \mathbb{R}^n の内積とすれば, $b_{ij} := \langle \mathbf{e}_i, \mathbf{e}_j \rangle$ を成分とする n 次正方行列 $B = [b_{ij}]$ は対称であり, $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle_B$ が成り立つ (標準ベクトルについて一致するから, それらの線形結合で書ける任意のベクトルについても一致する). また, $\mathbf{x} \neq \mathbf{0}$ ならば $0 < \langle \mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle_B = {}^t \mathbf{x} B \mathbf{x} = q_B(\mathbf{x})$ ゆえ, q_B は正定値である.

備考 37.2.6 (よりみち). 補題 37.2.1 により, 微積分における極値の概念を用いれば, 次のような言い換えもできる. これらは多変数関数の極値判定において基礎となる事実である.

- q は正定値である $\iff q$ は $\mathbf{x} = \mathbf{0}$ で狭義の極小値²をとる,
- q は半正定値である $\iff q$ は $\mathbf{x} = \mathbf{0}$ で極小値をとる,
- q は負定値である $\iff q$ は $\mathbf{x} = \mathbf{0}$ で狭義の極大値をとる,
- q は半負定値である $\iff q$ は $\mathbf{x} = \mathbf{0}$ で極大値をとる,
- q は不定符号である $\iff q$ は $\mathbf{x} = \mathbf{0}$ で極値をとらない.

任意の 2 次形式 q において, q の各偏導関数 $\frac{\partial}{\partial x_i} q(\mathbf{x})$ は $\mathbf{x} = \mathbf{0}$ においていずれも 0 となる³. 上にあるように q が不定符号の場合, q は $\mathbf{x} = \mathbf{0}$ で極値をとらない. このように, 微分は消えるものの極値ではない点のことを鞍点 (saddle point) という (例 37.2.3(3) の図は鞍の形に似ている).

37.3 2 次形式と固有値

与えられた 2 次形式 q_A が前節で与えた分類のいずれに属するかは, 対称行列 A の固有値を調べることですぐに分かる.

命題 37.3.1. A を実対称行列, $\mathbf{y} \in \mathbb{R}^n$ とする.

(1) P を n 次実正方行列とし, $\mathbf{x} = P\mathbf{y}$ とおけば, $q_A(\mathbf{x}) = q_{{}^t P A P}(\mathbf{y})$.

【注意】 ${}^t P A P$ は対称行列である (補題 36.3.5).

(2) A は, ある直交行列 P を用いて対角化できる (系 36.3.6). このとき, ${}^t P A P = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ とす

れば, 変数変換 $\mathbf{x} = P\mathbf{y}$ により (ここで $\mathbf{x} = {}^t(x_1, \dots, x_n)$, $\mathbf{y} = {}^t(y_1, \dots, y_n)$ とする),

$$q_A(x_1, \dots, x_n) = q_{{}^t P A P}(y_1, \dots, y_n) = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2.$$

²点 \mathbf{x} の十分近くに f の定義域を制限したとき, 制限した定義域において f が \mathbf{x} で (唯一の) 最小値を取るとき, f は \mathbf{x} で (狭義の) 極小値を取るといふ. 極大値も同様に定められる.

³ $q(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$ のとき, $\frac{\partial}{\partial x_i} q(\mathbf{x}) = 2a_{ii} x_i + (a_{ij} + a_{ji}) x_j$ である.

Proof. (1): $q_A(\mathbf{x}) = {}^t\mathbf{x}A\mathbf{x} = {}^t(P\mathbf{y})A(P\mathbf{y}) = {}^t\mathbf{y} {}^tPAP\mathbf{y} = q_{{}^tPAP}(\mathbf{y})$.

(2): (1) より明らか. 実際,

$$q_A(\mathbf{x}) = q_{{}^tPAP}(\mathbf{y}) = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

□

命題 37.3.2. A を実対称行列, P を実可逆行列とし, $B = {}^tPAP$ とすれば

- (1) q_A は正定値である $\iff q_B$ は正定値である,
- (2) q_A は半正定値である $\iff q_B$ は半正定値である,
- (3) q_A は負定値である $\iff q_B$ は負定値である,
- (4) q_A は半負定値である $\iff q_B$ は半負定値である,
- (5) q_A は不定符号である $\iff q_B$ は不定符号である.

Proof. (1): (\Rightarrow): q_A を正定値とする. 各 $\mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ について, $\mathbf{x} := P\mathbf{y}$ とおけば P は可逆ゆえ $\mathbf{x} \neq \mathbf{0}$ であり, 命題 37.3.1(1) より $0 < q_A(\mathbf{x}) = q_B(\mathbf{y})$. つまり q_B は正定値である. (\Leftarrow): q_B を正定値とする. $Q := P^{-1}$ とおけば, ${}^tQBQ = {}^tQ({}^tPAP)Q = {}^t(PQ)A(PQ) = A$ である. 各 $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ について $\mathbf{y} := Q\mathbf{x}$ とおけば $\mathbf{y} \neq \mathbf{0}$ であり, 命題 37.3.1(1) より $0 < q_B(\mathbf{y}) = q_{{}^tQBQ}(\mathbf{x}) = q_A(\mathbf{x})$. つまり q_A は正定値である.

(2) および (3), (4) は上と同様の論法で示せるゆえ略す. また, q_A および q_B が (1) から (4) のいずれでもない場合として, (5) の同値性を得る. □

以上により, q_A が (半) 正 (負) 定値かどうかを分析するには, 命題 37.3.1(2) で得られた $q_B(y_1, \dots, y_n) = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$ がそうであるかを見ればよい. そして, それは A の各固有値 λ_i の正負の情報さえ分かれば明らかなのであった (例 37.2.4). 以上をまとめると,

系 37.3.3. n 次実対称行列 A の重複を込めたすべての固有値を $\lambda_1, \dots, \lambda_n$ とすれば次が成り立つ.

- (1) $\lambda_1, \dots, \lambda_n$ はすべて正数である $\iff q_A$ は正定値である.
- (2) $\lambda_1, \dots, \lambda_n$ はすべて 0 以上である. $\iff q_A$ は半正定値である.
- (3) $\lambda_1, \dots, \lambda_n$ はすべて負数である $\iff q_A$ は負定値である.
- (4) $\lambda_1, \dots, \lambda_n$ はすべて 0 以下である. $\iff q$ は半負定値である.
- (5) $\lambda_1, \dots, \lambda_n$ の中に正数のものと負数のものがそれぞれ一つ以上ある. $\iff q$ は不定符号である.

37.4 正定値性の判別法

前節の分類を行うには, A の固有値を求める必要があった. ここでは行列式の計算を通して固有値を求めずに q_A が正 (負) 定値か否かを判定する方法を紹介する. 次の補題により, 正定値かどうかの判別が本質的である.

補題 37.4.1. A を実対称行列とする.

- (1) q_A は (半) 正定値である $\iff q_{(-A)}$ は (半) 負定値である.
- (2) q_A は (半) 負定値である $\iff q_{(-A)}$ は (半) 正定値である.

(3) q_A は不定符号である $\iff q_{(-A)}$ は不定符号である.

Proof. 次の計算により $-q_A = q_{(-A)}$ であることから, いずれの主張も明らかである. $A = [a_{ij}]$ とおくと, $-A = [-a_{ij}]$ であり,

$$-q_A(x_1, \dots, x_n) = -\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \sum_{i=1}^n \sum_{j=1}^n (-a_{ij}) x_i x_j = q_{(-A)}(x_1, \dots, x_n).$$

□

定義 37.4.2. n, k を $k \leq n$ を満たす自然数とする. n 次正方形行列 $A = [a_{ij}]$ に対して, その左上の $k \times k$ 成分からなる k 次正方形行列を A の k 次主小行列とよび, これを A_k と書く. すなわち,

$$A_k := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}.$$

A が実対称行列ならば A_k もそうであり, ゆえに A_k も直交行列により対角化可能である. また, $|A_k|$ の重複をこめたすべての固有値 $\lambda_1, \dots, \lambda_k$ はいずれも実数であり, $|A_k| = \lambda_1 \lambda_2 \cdots \lambda_k$ が成り立つ (命題 29.2.5).

定理 37.4.3. n 次対称行列 A について次が成り立つ:

$$q_A \text{ は正定値である. } \iff \text{各 } k = 1, \dots, n \text{ について, } |A_k| > 0.$$

Proof. (\Rightarrow): 各 $k = 1, \dots, n$ および $(x_1, \dots, x_k) \neq \mathbf{0}$ について,

$$q_{A_k}(x_1, \dots, x_k) = q_A(x_1, \dots, x_k, 0, \dots, 0) > 0$$

ゆえ q_{A_k} は正定値である. よって, $|A_k|$ の重複をこめたすべての固有値を $\lambda_1, \dots, \lambda_k$ とすれば, これらはすべて正であり (系 37.3.3(1)), したがって $|A_k| = \lambda_1 \lambda_2 \cdots \lambda_k > 0$.

(\Leftarrow): A のサイズ n に関する帰納法で示す. $n = 1$ のときは, $A = [a]$ (1×1 行列) とすれば $q_A(x) = ax^2$ および $\det A = a$ より (\Leftarrow) は明らか. 次に, $(n-1)$ 次対称行列について (\Leftarrow) が成り立つと仮定する. n 次対称行列 $A = [a_{ij}]$ が $|A_1|, \dots, |A_n| > 0$ を満たすとし, q_A の正定値性を示そう. さて, 次の 2 次形式を考える:

$$a_{11}q_A(x_1, \dots, x_n) - (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)^2. \quad (37.4.1)$$

ここで, 上式の x_1 に関する項は相殺されて消えてしまう. 実際, $a_{11}q_A(x_1, \dots, x_n)$ における x_1^2 の項は $a_{11}^2 x_1^2$ であり, また $x_1 x_i$ (ただし $i > 2$) に関する同類項をまとめると $a_{11}(a_{1i} x_1 x_i + a_{i1} x_i x_1) = 2a_{11} a_{i1} x_1 x_i$ である. これらは上式の第 2 項を展開した際に現れる x_1^2 および $x_1 x_i$ に関する項とそれぞれ一致している. ゆえに上式は $n-1$ 個の変数 x_2, \dots, x_n に関する 2 次形式とみなせる. そこでこれを $Q(x_2, \dots, x_n)$ と置こう. さらに 2 次形式 Q を定める $n-1$ 次対称行列を

$$B = \begin{bmatrix} b_{22} & b_{23} & \cdots & b_{2n} \\ b_{32} & b_{33} & \cdots & b_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n2} & b_{n3} & \cdots & b_{nn} \end{bmatrix}$$

とおく. なお A の対称性より $Q(x_2, \dots, x_n)$ における $x_i x_j$ の係数と $x_j x_i$ の係数は一致している. したがって B の各係数は,

$$b_{ij} = Q(x_2, \dots, x_n) \text{ における } x_i x_j \text{ の係数} = a_{11} a_{ij} - a_{1i} a_{1j}.$$

ここで, A_k ($k = 2, \dots, n$) の第 2 行から第 k 行までをそれぞれ a_{11} 倍した行列の行列式を考えると,

$$a_{11}^{k-1}|A_k| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{11}a_{21} & a_{11}a_{22} & \cdots & a_{11}a_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ a_{11}a_{k1} & a_{11}a_{k2} & \cdots & a_{11}a_{kk} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ 0 & & & \\ \vdots & [a_{11}a_{ij} - a_{i1}a_{1j}] & & \\ 0 & & & \end{vmatrix} = a_{11}|B_{k-1}|.$$

ここで, 上の二つ目の等式において, 各 $i = 2, \dots, k$ について 1 行目の a_{i1} 倍を i 行目から引いている. また最後の等式では $a_{i1} = a_{1i}$ を用いた. 仮定より $a_{11} = |A_1| > 0$ および $|A_k| > 0$ であるから, 上式より $|B_1|, |B_2|, \dots, |B_{n-1}| > 0$ を得る. したがって帰納法の仮定により, B が定める 2 次形式 Q は正定値である. さて, $Q(x_2, \dots, x_n)$ は式 (37.4.1) によって定義されていた. これを移項すれば

$$a_{11}q_A(x_1, \dots, x_n) = Q(x_2, \dots, x_n) + (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)^2.$$

この右辺の第 2 項は半正定値であり, これと $a_{11} > 0$ を合わせれば, q_A の正定値性が導かれる. \square

次のアルゴリズムにより, 固有値を求めずとも q_A が正 (負) 定値か否かを判別できる.

(1) A を実対称行列とする. まず $|A| = |A_n|$ を計算する.

- $|A| = 0$ の場合: A の固有値の中に 0 が含まれている. したがって q_A は正定値でも負定値でもない. q_A は半正定値, 半負定値, 不定符号のいずれかとなるが, これらのいずれになるかは A の 0 以外の固有値の符号を調べる必要がある.
- $|A| > 0$ の場合: (2) へ.
- $|A| < 0$ の場合: 定理 37.4.3 より q_A は正定値ではない. (3) へ.

(2) 定理 37.4.3 を用いて q_A が正定値かどうかを判定する. 正定値でない場合は (3) へ.

(3) いま, q_A は正定値ではない. 次に, q_{-A} が正定値か否かを, 定理 37.4.3 を用いて判定する.

- q_{-A} が正定値ならば q_A は負定値である (補題 37.4.1).
- q_{-A} が正定値でなければ q_A は正定値でも負定値でもない. $|A| \neq 0$ より A の固有値に 0 は含まれず, したがって A の固有値はいくつかの正数といくつかの負数からなる. ゆえに q_A は不定符号である.

2 次対称行列の場合は $|A| = 0$ かつ q_A が不定符号となることはない (次の命題を見よ). したがって, 行列式と対角成分の符号のみから, 次のように完全な分類ができる. とくに, 次の (1) と (2) は微積分の極値判定においてよく用いられる事実である.

命題 37.4.4. 実対称行列 $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ について次が成り立つ:

(1) $|A| > 0$ のとき:

- $a > 0 \implies q_A$ は正定値.
- $a < 0 \implies q_A$ は負定値.

(2) $|A| < 0 \implies q_A$ は不定符号.

(3) $|A| = 0$ のとき:

- $a > 0$ または $c > 0 \implies q_A$ は正定値ではない半正定値.
- $a < 0$ または $c < 0 \implies q_A$ は負定値ではない半負定値.

- 上記のいずれにも当てはまらない場合は $A = O$ (つまり $q_A = 0$).

Proof. A の重複度をこめた固有値を λ_1, λ_2 とすれば, $|A| = \lambda_1 \lambda_2$ である.

(1) $|A| > 0$ のとき: $\lambda_1, \lambda_2 \neq 0$ であり, かつこれらの正負は等しい.

- $a > 0$ の場合: 定理 37.4.3 より q_A は正定値である.
- $a < 0$ の場合: $|-A| = (-1)^2|A| = |A|$ が成り立つ. ゆえに定理 37.4.3 より q_{-A} は正定値であり, したがって q_A は負定値である.

(2) $|A| < 0$ のとき: $\lambda_1, \lambda_2 \neq 0$ であり, かつこれらの正負は異なる. つまり q_A は不定符号である.

(3) $|A| = 0$ のとき: $\lambda_1 \lambda_2 = 0$ である. λ_1, λ_2 のうち少なくとも一方は 0 であり, ゆえに q_A は正定値でも負定値でもない. もう片方の固有値が 0 以上ならば半正定値, 0 以下ならば半負定値である. つまり q_A は半正定値または半負定値のいずれかとなる.

- $a > 0$ または $c > 0$ の場合: $a > 0$ の場合を考えよう. このとき

$$q_A(1, 0) = a1^2 + 2b1 \cdot 0 + c0^2 = a > 0.$$

つまり q_A は半負定値にはなり得ない. したがって半正定値である. $c > 0$ の場合は $q_A(0, 1) = c$ より同様の結論を得る.

- $a < 0$ または $c < 0$ の場合: $q_A(1, 0) = a$ および $q_A(0, 1) = c$ より q_A は半正定値にはなり得ない. ゆえに半負定値である.
- 上記のいずれにも当てはまらない場合: これは $|A| = ac - b^2 = 0, a \leq 0, c \leq 0, a \geq 0, c \geq 0$ が全て満たされている場合に相当する. $a = c = 0$ であり, $ac - b^2 = 0$ から $b = 0$. 以上より $A = O$ である.

□

37.5 2次曲線(よりみち)

126 ページのコラムで述べた変数変換は, 次のようにして得られる. 関数 $y = \frac{1}{x}$ のグラフ, すなわち曲線 $xy = 1$ について考える. この左辺は変数 x, y に関する 2 次形式であるから, これを変数変換により $\lambda_1 X^2 + \lambda_2 Y^2$ の形に書き換えよう.

$$A = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix} \quad \text{とすれば} \quad q_A(x, y) = xy$$

である. 直交行列による A の対角化は例題 36.3.9(1) より,

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad {}^t P A P = \begin{bmatrix} 1/2 & 0 \\ 0 & -1/2 \end{bmatrix}. \quad \text{したがって, } \begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} X \\ Y \end{bmatrix} \quad \text{と変数変換すれば,}$$

命題 37.3.1 より, $xy = \frac{1}{2}X^2 - \frac{1}{2}Y^2$ を得る.

これと同様の書き換えが一般の 2 次曲線についてもいえる. いまから 2 次曲線⁴

$$ax^2 + 2bxy + cy^2 + 2(\alpha x + \beta y) + \gamma = 0 \quad (\text{ただし, } a, b, c, \alpha, \beta, \gamma \in \mathbb{R} \text{ は定数}) \quad (37.5.1)$$

⁴1 次項の係数を 2 倍しているのは, 3 変数の 2 次形式の特別な場合と見なすことによる. $B = \begin{bmatrix} a & b & \alpha \\ b & c & \beta \\ \alpha & \beta & \gamma \end{bmatrix}$ とおくと,

$q_B(x, y, 1) = ax^2 + 2bxy + cy^2 + 2(\alpha x + \beta y) + \gamma$ が成り立つ. この事実から, 平面上の任意の 2 次曲線は, 2 次形式から定まる \mathbb{R}^3 上の 2 次曲面 $q_B(x, y, z) = 0$ を平面 $z = 1$ で切ったときに現れる切り口と合同であることが分かる.

を変数変換によって簡単な形に書き換えてみよう. 対称行列 $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ の重複を含めた固有値を λ_1, λ_2 とする. ここで, いずれの固有値も 0 であるとする, $q_A = 0$ ゆえ $ax^2 + 2bxy + cy^2 = 0$ となり, 式 (37.5.1) は直線を表す⁵. そこで少なくとも一つの固有値は 0 でないとしてよい. ここでは $\lambda_1 \neq 0$ とする. さて, A の対称性より $P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ となるような直交行列 P が存在する. そこで, $\begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} X \\ Y \end{bmatrix}$ と変数変換しよう. このとき 1 次式の項について

$$2(\alpha x + \beta y) = 2 \begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} \alpha & \beta \end{bmatrix} P \begin{bmatrix} X \\ Y \end{bmatrix}$$

となるから, 命題 37.3.1 より式 (37.5.1) は次のように書き換えられる:

$$\lambda_1 X^2 + \lambda_2 Y^2 + 2(AX + BY) + \gamma = 0, \quad \text{ただし,} \quad \begin{bmatrix} A & B \end{bmatrix} := \begin{bmatrix} \alpha & \beta \end{bmatrix} P.$$

さらに $X = \tilde{X} - A/\lambda_1$ と平行移動することにより, 上式の X の 1 次の項が消去され,

$$\lambda_1 \tilde{X}^2 + \lambda_2 Y^2 + 2BY + \gamma - \frac{A^2}{\lambda_1} = 0 \quad (37.5.2)$$

を得る. また, $\lambda_2 \neq 0$ の場合は, 平行移動 $Y = \tilde{Y} - B/\lambda_2$ によって Y の 1 次の項も消去され,

$$\lambda_1 \tilde{X}^2 + \lambda_2 \tilde{Y}^2 + \gamma - \frac{A^2}{\lambda_1} - \frac{B^2}{\lambda_2} = 0.$$

一方, $\lambda_2 = 0$ の場合は, 式 (37.5.2) において $Y = \tilde{Y} - \frac{1}{2B} \left(\gamma - \frac{A^2}{\lambda_1} \right)$ と平行移動すれば,

$$\lambda_1 \tilde{X}^2 + 2B\tilde{Y} = 0.$$

ここで変数記号に改めて x, y を用いるとし, 上に現れた定数をまとめれば,

$$\begin{cases} \lambda_1 x^2 + \lambda_2 y^2 + \tilde{\gamma} = 0 & (\lambda_2 \neq 0 \text{ のとき}), \\ \lambda_1 x^2 + 2By = 0 & (\lambda_2 = 0 \text{ のとき}). \end{cases}$$

以上により, 任意の 2 次曲線は, 直交行列による座標変換 (これを直交変換という) と平行移動により, 上の二種の 2 次曲線のいずれかと合同⁶ になることが分かった (定数の値によっては, 上式を満たす点 (x, y) の集合が空集合, あるいは一点, 直線になることもあり得る). 上の二式を更に変形することにより, 楕円, 双曲線, 放物線などの標準形が得られることは, 周知のとおりである.

\mathbb{R}^3 上の 2 次曲面 (さらには \mathbb{R}^n 上の 2 次超曲面) についても 2 次の項をまとめることで, 上と類似する項の消去を行うことができる⁷.

37.6 多変数関数の極値判定 (よりみち)

多変数関数の極値判定と 2 次形式の関係について述べる. $U \subset \mathbb{R}^n$ を開集合とし, $f: U \rightarrow \mathbb{R}$ を C^∞ 級関数とする.

次で定める n 次正方行列 $H_f(\mathbf{x})$ を f のヘッセ行列という:

$$H_f(\mathbf{x}) := [f_{x_i x_j}(\mathbf{x})] = \begin{bmatrix} f_{x_1 x_1}(\mathbf{x}) & f_{x_1 x_2}(\mathbf{x}) & \cdots & f_{x_1 x_n}(\mathbf{x}) \\ f_{x_2 x_1}(\mathbf{x}) & f_{x_2 x_2}(\mathbf{x}) & \cdots & f_{x_2 x_n}(\mathbf{x}) \\ \vdots & \vdots & \cdots & \vdots \\ f_{x_n x_1}(\mathbf{x}) & f_{x_n x_2}(\mathbf{x}) & \cdots & f_{x_n x_n}(\mathbf{x}) \end{bmatrix}.$$

⁵ただし $\alpha, \beta = 0$ の場合は直線ではない.

⁶直交変換と平行移動はいずれも等長変換 (備考 38.2.3) ゆえ, これらの写像による図形 $G \subset \mathbb{R}^2$ の像は G と合同である.

⁷例えば文献 [5] を見よ.

f が C^2 級であることから, $H_f(\mathbf{x})$ は対称行列である.

点 $\mathbf{a} \in U$ における $\mathbf{h} = {}^t(h_1, \dots, h_n)$ 方向の 2 次の方向微分係数は, 次のように書き下すことができる:

$$\begin{aligned} \left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} \right)^2 f(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{a}} &= \sum_{i=1}^n \sum_{j=1}^n h_i h_j \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{a}} \\ &= \sum_{i=1}^n \sum_{j=1}^n h_i h_j f_{x_i x_j}(\mathbf{a}) = q_{H_f(\mathbf{a})}(h_1, \dots, h_n). \end{aligned}$$

したがって, テイラーの定理を点 $\mathbf{a} \in U$ において適用すれば,

$$\begin{aligned} f(\mathbf{a} + \mathbf{h}) &= \sum_{k=1}^2 \frac{1}{k!} \left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} \right)^k f(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{a}} + R_3 \\ &= f(\mathbf{a}) + (\text{grad } f)(\mathbf{a}) \cdot \mathbf{h} + \frac{1}{2} q_{H_f(\mathbf{a})}(h_1, \dots, h_n) + R_3, \end{aligned}$$

ここでドット \cdot は内積⁸, R_3 は剰余項を表す. $\mathbf{a} + \mathbf{h}$ が \mathbf{a} に十分近いときに R_3 は微小な数となる (漸近展開). さて, 点 \mathbf{a} が f の極値をとる点の候補であるとき (つまり $(\text{grad } f)(\mathbf{a}) = \mathbf{0}$ のとき), 上式は

$$f(\mathbf{a} + \mathbf{h}) = (\text{定数}) + \frac{1}{2} q_{H_f(\mathbf{a})}(h_1, \dots, h_n) + (\text{微小部分})$$

となり, f が \mathbf{a} で極値を取るかどうかは, $q_{H_f(\mathbf{a})}(h_1, \dots, h_n)$ の挙動に大きく左右されることが分かる. そこで多変数の微積分法では, まず 2 次形式 $q_{H_f(\mathbf{a})}$ の挙動を本章で紹介する形で調べたうえで, それとともに微小部分の挙動が本当に無視できる程度かどうかを分析することによって, f の極値判定を行うのである. その詳しい議論については微積分の教科書を参照してもらうとして, その結論だけ述べれば次の通りである. $\mathbf{a} \in U$ を $(\text{grad } f)(\mathbf{a}) = \mathbf{0}$ 満たす点とすれば,

- $q_{H_f(\mathbf{a})}$ が正 (負) 定値を取れば, f は \mathbf{a} で極小 (大) となる.
- $q_{H_f(\mathbf{a})}$ が不定符号ならば, f は \mathbf{a} で極値を取らない.

とくに, 2 変数関数 $f(x, y)$ の場合は, $q_{H_f(a,b)}$ の挙動が $|H_f(a, b)|$ と, $H_f(a, b)$ の (1, 1)-成分 $f_{xx}(a, b)$ の符号から判定できる (命題 37.4.4). なお, ヘッセ行列式は

$$|H_f(a, b)| = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2$$

である. この値 (あるいはその -1 倍) を通して極値判定を行う手法は, 2 変数のみを扱う微積分の教科書にも必ずといっていいほど紹介されている.

⁸微積分では, 内積の記号に $\langle \cdot, \cdot \rangle$ を用いると式が長くなってしまいうためドットがよく用いられる.

第38章 内積空間上の線形変換

体 \mathbb{R} 上の内積空間におけるいくつかの線形変換と、その基本的性質を紹介する。特に幾何的な側面が強い射影について述べる。

38.1 ピタゴラスの定理とコーシー-シュワルツの不等式

\mathbb{R}^n においてピタゴラスの定理や三角不等式が成立することに疑いの余地はないかもしれない¹。しかしながら、一般の内積空間 $(U, \langle \cdot, \cdot \rangle)$ においてピタゴラスの定理が成り立つことは明らかではない。そこで、内積から定まるノルムの定義に基づいて次を示そう。

定理 38.1.1 (ピタゴラスの定理). 内積空間 $(U, \langle \cdot, \cdot \rangle)$ 上の二つのベクトル $\mathbf{x}, \mathbf{y} \in U$ が直交するとき、次が成り立つ:

- (1) $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2$.
- (2) $\|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$ であり、等号成立は $\mathbf{x} = \mathbf{0}$ のときに限る.
- (3) $\|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\|$ であり、等号成立は $\mathbf{y} = \mathbf{0}$ のときに限る.

Proof. (1): $\|\mathbf{x} - \mathbf{y}\|^2 = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - 0 - 0 + \langle \mathbf{y}, \mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$. (2): (1) より $\|\mathbf{y}\|^2 \leq \|\mathbf{x} - \mathbf{y}\|^2$ であるから、 $\|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$ である。また、 $\|\mathbf{y}\| = \|\mathbf{x} - \mathbf{y}\|$ が成立するならば $\|\mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2$ 、すなわち $\langle \mathbf{y}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle$ であり、したがって $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ 。つまり $\mathbf{x} = \mathbf{0}$ である。

(3): (2) において \mathbf{x} と \mathbf{y} の立場を入れかえれば、 $\|\mathbf{x}\| \leq \|\mathbf{y} - \mathbf{x}\| = \|\mathbf{x} - \mathbf{y}\|$. □

ベクトル \mathbf{x} の \mathbf{y} 方向への射影を \mathbf{u} とすれば、明らかに $\|\mathbf{u}\| \leq \|\mathbf{x}\|$ である (図 38.1)。 \mathbf{x} と \mathbf{y} のみを用いてこの不等式を表したものをコーシー-シュワルツの不等式という。

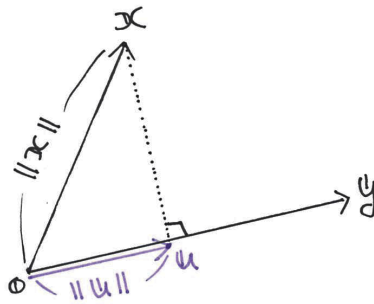


図 38.1: $\|\mathbf{u}\| \leq \|\mathbf{x}\|$

定理 38.1.2 (コーシー-シュワルツの不等式). U を内積空間とすれば、各 $\mathbf{x}, \mathbf{y} \in U$ について

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$$

が成り立つ。また、 $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ の場合における等号成立は \mathbf{x} と \mathbf{y} のいずれかが他方のスカラー倍であるときに限る。

¹ $n \geq 4$ の場合は図示が難しいことから、 \mathbb{R}^4 の場合についても疑うべきかもしれない。なお、我々は \mathbb{R}^n においてピタゴラスの定理が成り立つことを前提として、標準的なノルムを定義したのであった。

Proof. $\mathbf{x} = \mathbf{0}$ または $\mathbf{y} = \mathbf{0}$ のときは等号が成立するゆえ、 $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ の場合を考える。

$$\mathbf{v} := \frac{1}{\|\mathbf{y}\|}\mathbf{y}, \quad \mathbf{u} := \langle \mathbf{v}, \mathbf{x} \rangle \mathbf{v} = \left\langle \frac{1}{\|\mathbf{y}\|}\mathbf{y}, \mathbf{x} \right\rangle \frac{1}{\|\mathbf{y}\|}\mathbf{y} = \frac{\langle \mathbf{y}, \mathbf{x} \rangle}{\|\mathbf{y}\|^2}\mathbf{y}$$

とすれば \mathbf{u} は \mathbf{x} の \mathbf{y} 方向への射影であり、 \mathbf{u} と $\mathbf{u} - \mathbf{x}$ は直交する (これはシュミットの直交化法 (定理 35.4.1) の証明の中で既に示したことである²⁾). 前定理 (3) より $\|\mathbf{u}\| \leq \|\mathbf{u} - (\mathbf{u} - \mathbf{x})\| = \|\mathbf{x}\|$. また,

$$\|\mathbf{u}\| = \left\| \frac{\langle \mathbf{y}, \mathbf{x} \rangle}{\|\mathbf{y}\|^2}\mathbf{y} \right\| = \frac{|\langle \mathbf{y}, \mathbf{x} \rangle|}{\|\mathbf{y}\|^2}\|\mathbf{y}\| = \frac{|\langle \mathbf{y}, \mathbf{x} \rangle|}{\|\mathbf{y}\|}.$$

したがって、 $|\langle \mathbf{y}, \mathbf{x} \rangle| \leq \|\mathbf{y}\| \cdot \|\mathbf{x}\|$ である。なお等号成立は前定理より $\mathbf{u} - \mathbf{x} = \mathbf{0}$, つまり $\mathbf{u} = \mathbf{x}$ のときであり、ゆえに \mathbf{x} は \mathbf{y} のスカラー倍である。逆に、 \mathbf{x} と \mathbf{y} のいずれかが他方のスカラー倍であるときに等号が成立することはすぐに分かる。実際、 $|\langle \mathbf{x}, t\mathbf{x} \rangle| = |t\langle \mathbf{x}, \mathbf{x} \rangle| = |t| \cdot |\langle \mathbf{x}, \mathbf{x} \rangle| = \|\mathbf{x}\| \cdot |t|\|\mathbf{x}\| = \|\mathbf{x}\| \cdot \|t\mathbf{x}\|$. \square

命題 38.1.3 (三角不等式). U を内積空間とすれば、 $\mathbf{x}, \mathbf{y}, \mathbf{z} \in U$ について次が成り立つ。

- (1) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$,
- (2) $\|\mathbf{x} - \mathbf{z}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\|$.

Proof. (1): 不等式 $\langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$ より得られる。実際、

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\| &= \sqrt{\langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle} \\ &\leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle + 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| + \langle \mathbf{y}, \mathbf{y} \rangle} = \sqrt{\|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| + \|\mathbf{y}\|^2} \\ &= \sqrt{(\|\mathbf{x}\| + \|\mathbf{y}\|)^2} = \|\mathbf{x}\| + \|\mathbf{y}\|. \end{aligned}$$

(2): (1) を用いると、 $\|\mathbf{x} - \mathbf{z}\| = \|(\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z})\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\|$. \square

本書では、ノルムに関する不等式は定理 38.1.1 のみを用い、三角不等式を直接に用いることはない。しかしながら、微分積分学において三角不等式が何度も用いられることは読者にも周知の事実であり、ここで改めてその重要性を述べる必要はなからう。

よりみち (線形空間と極限)

一般に、線形空間 U に何らかの方法で各ベクトル \mathbf{x} の長さ $\|\mathbf{x}\|$ (これをノルムという、その定義は下を見よ) が与えられると、二点 $\mathbf{x}, \mathbf{y} \in U$ の間の距離 (metric) を $\|\mathbf{x} - \mathbf{y}\|$ と定めることで、 U 上に極限概念が導入できる。すなわち、変数ベクトル $\mathbf{x} \in U$ が点 $\mathbf{a} \in U$ に限りなく近づくことを、 $\|\mathbf{x} - \mathbf{a}\|$ が 0 に限りなく近づくことと定義するのである。無限次元の線形空間では、有限回の演算のみで代数的に処理できる部分は限られており、より詳しいことを分析するには極限操作の助け (解析学の知識) が必須となる。この解析学と線形代数学が融合した分野を関数解析学 (function analysis) という。無限次元の関数空間を調べる動機の一例として、本書ではフーリエ級数展開を挙げた (39.8 節)。

定義 38.1.4. 線形空間 U 上の関数 $\|\cdot\| : U \rightarrow [0, \infty)$ が命題 35.1.5 で与えたすべての条件を満たすとき、これを U 上のノルム (norm) という。また、ノルムが与えられている線形空間をノルム空間 (norm space) という。

\mathbb{R}^n 上のノルムとして、例えば $\|(x_1, \dots, x_n)\| = \sum_{i=1}^n |x_i|$ がある。このノルムは、 \mathbb{R}^n 上の内積から定められるような、いかなるノルムとも異なる。何故なら、中線定理 (命題 38.2.1(4)) が成立しないからである (標準ベクトルについて成り立たない)。

²⁾定理 35.4.1 の証明では、 $\mathbf{u}_1 := \mathbf{y}$, $\mathbf{u}_2 := \mathbf{x}$ として、 $\mathbf{v}_1 = \mathbf{v}$ と $\tilde{\mathbf{v}}_2 = \mathbf{x} - \mathbf{u}$ が直交することを示した。したがって、これらをスカラー倍した \mathbf{u} と $\mathbf{u} - \mathbf{x}$ も直交する。

38.2 直交変換

35.2節では、余弦定理を通して内積とノルムが相互に翻訳できることを見た(命題35.2.2(3)). これは一般の内積空間においても成り立つ. 次の(4)は中線定理または平行四辺形の法則(parallelogram law)と呼ばれる.

命題 38.2.1. 体 \mathbb{R} 上の内積空間 U の各元 $x, y \in U$ において, 次が成り立つ

$$(1) \langle x, y \rangle = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2).$$

$$(2) \langle x, y \rangle = \frac{1}{2}(\|x\|^2 + \|y\|^2 - \|x - y\|^2).$$

$$(3) \langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2).$$

$$(4) \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

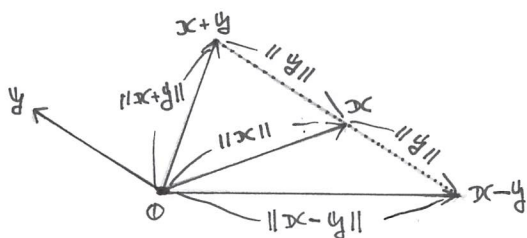


図 38.2: 中線定理

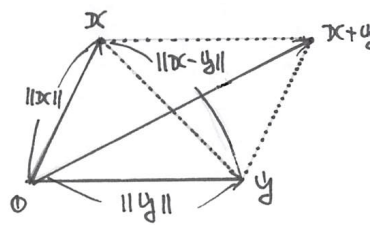


図 38.3: 平行四辺形の法則

【補足】 次の証明に現れる $\overline{\langle x, y \rangle}$ は $\langle x, y \rangle$ の共役な複素数を表す. ただし, いまは $\langle x, y \rangle \in \mathbb{R}$ ゆえ $\overline{\langle x, y \rangle} = \langle x, y \rangle$ であり, この記法は無視してよい. あとでエルミート内積について同様のことを論じる際に必要となるものである.

Proof. (1): $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2$.
これを移項して2で割れば,

$$\frac{\langle x, y \rangle + \overline{\langle x, y \rangle}}{2} = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2).$$

$\langle x, y \rangle$ は実数ゆえ, 上の左辺は $\langle x, y \rangle$ に等しい.

(2): $\|x - y\|^2$ を展開することで得られる(計算略).

(3): (1) と (2) の右辺の平均をとったもの.

(4): (1) と (2) の右辺の2倍を等号で結び, 移項したもの. □

以上により内積は, その内積から定まるノルムによる式に書き換えられる. これにより, 次の(1)から(3)の同値性を得る:

命題 38.2.2. 体 \mathbb{R} 上の有限次元内積空間 U における線形変換 $f: U \rightarrow U$ について次は同値である:

$$(1) \text{ 各 } x, y \in U \text{ について } \langle f(x), f(y) \rangle = \langle x, y \rangle.$$

$$(2) \text{ 各 } x \in U \text{ について } \|f(x)\| = \|x\|.$$

$$(3) \text{ 各 } x, y \in U \text{ について } \|f(x) - f(y)\| = \|x - y\|.$$

(4) U の任意の正規直交基底に関する f の表現行列は直交行列である.

(5) f の表現行列が直交行列となるような, U の正規直交基底が少なくとも一つ存在する.

Proof. (1) \Rightarrow (2): $\|f(\mathbf{x})\| = \sqrt{\langle f(\mathbf{x}), f(\mathbf{x}) \rangle} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \|\mathbf{x}\|$.

(2) \Rightarrow (1): 命題 38.2.1 より,

$$\begin{aligned}\langle f(\mathbf{x}), f(\mathbf{y}) \rangle &= \frac{1}{4}(\|f(\mathbf{x}) + f(\mathbf{y})\|^2 - \|f(\mathbf{x}) - f(\mathbf{y})\|^2) \\ &= \frac{1}{4}(\|f(\mathbf{x} + \mathbf{y})\|^2 - \|f(\mathbf{x} - \mathbf{y})\|^2) \\ &= \frac{1}{4}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2) = \langle \mathbf{x}, \mathbf{y} \rangle.\end{aligned}$$

(2) \Rightarrow (3): $\|f(\mathbf{x}) - f(\mathbf{y})\| = \|f(\mathbf{x} - \mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$.

(3) \Rightarrow (2): (2) は, (3) において $\mathbf{y} = \mathbf{0}$ とした特別な場合である.

(1) \Rightarrow (4): $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の正規直交基底とし, この基底に関する f の表現行列を A とする. 命題 35.6.5 における線形同型 $F: U \rightarrow \mathbb{R}^n$ によって翻訳すれば, 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について $\langle A\mathbf{x}, A\mathbf{y} \rangle_{\mathbb{R}} = \langle f(F^{-1}(\mathbf{x})), f(F^{-1}(\mathbf{y})) \rangle = \langle F^{-1}(\mathbf{x}), F^{-1}(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}}$. ゆえに A は直交行列である (命題 36.1.1).

(4) \Rightarrow (5): 明らか.

(5) \Rightarrow (1): 仮定より, f の表現行列 A が直交行列となるような, U の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が存在する. 命題 35.6.5 における線形同型 $F: U \rightarrow \mathbb{R}^n$ によって翻訳すれば, 各 $\mathbf{u}, \mathbf{v} \in U$ について, $\langle f(\mathbf{u}), f(\mathbf{v}) \rangle = \langle AF(\mathbf{u}), AF(\mathbf{v}) \rangle_{\mathbb{R}} = \langle F(\mathbf{u}), F(\mathbf{v}) \rangle_{\mathbb{R}} = \langle \mathbf{u}, \mathbf{v} \rangle$. ここで, 二つ目の等号において A が直交行列であることを用いた (命題 36.1.1(4)). \square

上の条件のいずれか (したがってすべて) を満たす U 上の線形変換を直交変換 (orthogonal transformation) という. 普段は基底の取り方に依存しない (1) を定義とすることが多い.

備考 38.2.3 (発展). 上の条件 (3) を満たす写像を等長変換³(isometry) という. 等長変換は, 必ずしも内積が定まらない一般のノルム空間 (あるいはより広義に, 線形部分空間とは限らないノルム空間の部分集合) においても定義できる概念である. 例えば, ノルム空間における平行移動は線形写像ではないが ($\mathbf{0}$ を $\mathbf{0}$ に写さない), 等長変換である.

38.3 直交射影とその鏡映

次に述べる直交射影は, シュミットの直交化法において用いていた (備考 35.4.3). 定義 38.3.1 にある三条件の同値性の証明は 38.4 節で与える. 余力のある者は, 先に自分なりの証明を考えてみるとよい.

定義 38.3.1. U を有限次元内積空間とし, $W \subset U$ を部分空間とする. 写像 $p: U \rightarrow W$ に関する次の条件は同値であり (補題 38.4.3), このような写像を W への直交射影 (orthogonal projection) あるいは正射影という.

- (1) $p|_W = \text{id}_W$ かつ, 各 $\mathbf{x}, \mathbf{y} \in U$ について $\langle p(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, p(\mathbf{y}) \rangle$,
- (2) W の任意の正規直交基底 $\mathbf{w}_1, \dots, \mathbf{w}_k$ について, $p(\mathbf{x}) = \sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i$.
- (3) $p(\mathbf{x}) = \sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i$ を満たす W の正規直交基底 $\mathbf{w}_1, \dots, \mathbf{w}_k$ が存在する.
- (4) 各 $\mathbf{x} \in U$ について, $\mathbf{x} - p(\mathbf{x}) \in W^\perp$.

上の条件のいずれを定義にしてもよいのであるが, 基底の取り方に依存しない (1) をここでは定義とする. このような写像が少なくとも存在することは, (3) より分かる. また (3) からは p が線形写像であることも分かる. さらに, (1) より $\text{Im } p = W$ であり, (3) より $p|_{W^\perp} = \mathbf{0}$, すなわち $\text{Ker } p = W^\perp$ である⁴. なお, W への直交射影が唯一つしかないことは, 次の議論により導かれる: あらかじめ W の正規直交基

³等長変換に全射性を要求する流儀もある.

⁴ $\text{Ker } p \subset W^\perp$ の証明は次の通り: 各 $\mathbf{x} \in \text{Ker } p$ は, $\mathbf{w} \in W$ と $\mathbf{v} \in W^\perp$ を用いて $\mathbf{x} = \mathbf{w} + \mathbf{v}$ と書ける (命題 35.5.4). ゆえに $\mathbf{0} = p(\mathbf{x}) = p(\mathbf{w}) + p(\mathbf{v}) = \mathbf{w} + \mathbf{0} = \mathbf{w}$. したがって $\mathbf{x} = \mathbf{w} + \mathbf{v} = \mathbf{0} + \mathbf{v} = \mathbf{v} \in W^\perp$.

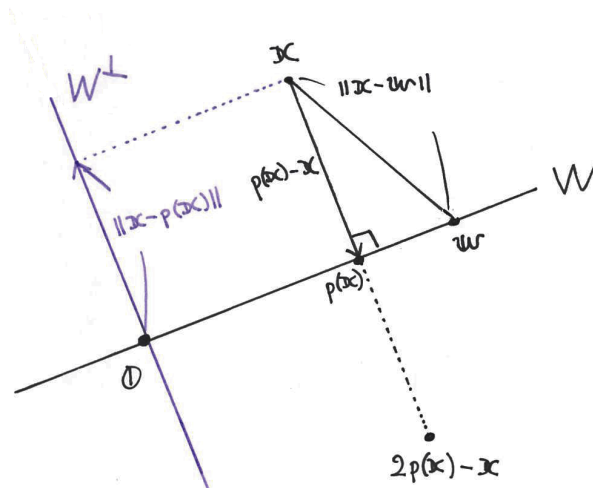


図 38.4: 直交射影とその鏡映

底 w_1, \dots, w_k を一つ決めておき, $p_1(x) := \sum_{i=1}^k \langle w_i, x \rangle w_i$ と定める. また, p_2 を W への任意の直交射影 (つまり (1) から (4) のいずれかを満たす写像) とする. このとき, 条件の同値性より p_2 は (2) を満たし, したがっていま取っておいた基底 w_1, \dots, w_k について $p_2(x) = \sum_{i=1}^k \langle w_i, x \rangle w_i$ が成り立つ. つまり $p_2 = p_1$ であり, W へのいかなる直交射影も p_1 に等しいことが分かる.

条件 (4) によれば, x と $p(x)$ を結ぶ直線は部分空間 W と直交する. すなわち $p(x)$ は, x から W へ向けて引いた垂線の足に等しい (図 38.4). 直交射影の唯一性は, x から W へ向かう垂線が唯一つに限ることを言っている. (4) を満たす写像の唯一性 (すなわち直交射影の唯一性) は次の命題からも分かる:

命題 38.3.2. $W \subset U$ を部分空間とし, そこへの直交射影を $p: U \rightarrow W$ とする. このとき, 各 $x \in U$ について, $p(x)$ は W の元の中で最も x に近い点である. すなわち, 次が成り立つ:

$$\text{各 } w \in W \text{ について, } \|x - p(x)\| \leq \|x - w\|.$$

また, この不等式の等号成立は $w = p(x)$ のときに限る. つまり, W の元の中で x に最も近い点は $p(x)$ 唯一つである.

Proof. $w \in W$ とすれば $w - p(x) \in W$ であるから, 直交射影の性質 (3) より $w - p(x)$ と $x - p(x)$ は直交する. ゆえに定理 38.1.1(3) より

$$\|x - p(x)\| \leq \|(x - p(x)) - (w - p(x))\| = \|x - w\|.$$

また, 等号成立は $w - p(x) = \mathbf{0}$ のとき, つまり $w = p(x)$ のときに限る. □

備考 38.3.3. (1) $p: U \rightarrow W$ が W への直交射影であるとき, $q(x) = x - p(x)$ は W^\perp への直交射影である. 実際, 定義 38.3.1(4) がすぐに確認できる: $x - q(x) = x - (x - p(x)) = p(x) \in W = (W^\perp)^\perp$.

(2) 直交補空間 W^\perp の正規直交基底を v_1, \dots, v_ℓ とすれば, W^\perp への直交射影 $x - p(x)$ は, 定義 38.3.1(4) により, $x - p(x) = \sum_{i=1}^{\ell} \langle v_i, x \rangle v_i$ と書ける. これを移項して次式を得る.

$$p(x) = x - \sum_{i=1}^{\ell} \langle v_i, x \rangle v_i. \quad (38.3.1)$$

すなわち, W への直交射影 $p: U \rightarrow W$ は, W の直交補空間の基底を用いて表すこともできる.

さて, $p(\mathbf{x})$ をさらに $p(\mathbf{x}) - \mathbf{x}$ 方向に平行移動させた位置にあたる $2p(\mathbf{x}) - \mathbf{x}$ を考えれば, この点と \mathbf{x} は W について対称な点となる (図 38.4). そこで, $r(\mathbf{x}) := 2p(\mathbf{x}) - \mathbf{x}$ で定められる線形変換 $r: U \rightarrow U$ を直交射影 p による鏡映 (refraction) とよぶ. 鏡映を直交補空間の正規直交基底 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ を用いて表す場合は, 式 (38.3.1) より次のようになる:

$$r(\mathbf{x}) = 2p(\mathbf{x}) - \mathbf{x} = \mathbf{x} - 2 \sum_{i=1}^{\ell} \langle \mathbf{v}_i, \mathbf{x} \rangle \mathbf{v}_i.$$

直交射影による鏡映は直交変換である:

命題 38.3.4. 直交射影 $p: U \rightarrow W$ による鏡映変換 $r: U \rightarrow U$ は直交変換である.

Proof. W の正規直交基底を $\mathbf{w}_1, \dots, \mathbf{w}_k$, また W^\perp の正規直交基底を $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ とすれば, $r(\mathbf{x}) = 2p(\mathbf{x}) - \mathbf{x}$ は, $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ を $\mathbf{w}_1, \dots, \mathbf{w}_k, -\mathbf{v}_{k+1}, \dots, -\mathbf{v}_n$ に写す線形写像である (これは $p|_W = \text{id}_W$ および $p|_{W^\perp} = \mathbf{0}$ による). したがって基底 $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ に関する r の表現行列は対角成分が ± 1 からなる対角行列, つまり直交行列である. ゆえに r は直交変換である (命題 38.2.2). \square

【備考】 $p|_W = \text{id}_W$ より $p^2 = p$ であり (詳しくは命題 38.5.4(1) を見よ), これと $\langle p(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, p(\mathbf{y}) \rangle$ を用いて $\langle r(\mathbf{x}), r(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ を導いても良い. 実際, $\langle r(\mathbf{x}), r(\mathbf{y}) \rangle = \langle 2p(\mathbf{x}) - \mathbf{x}, 2p(\mathbf{y}) - \mathbf{y} \rangle = 4\langle p(\mathbf{x}), p(\mathbf{y}) \rangle - 2\langle p(\mathbf{x}), \mathbf{y} \rangle - 2\langle \mathbf{x}, p(\mathbf{y}) \rangle + \langle \mathbf{x}, \mathbf{y} \rangle = 4\langle \mathbf{x}, p(p(\mathbf{y})) \rangle - 4\langle \mathbf{x}, p(\mathbf{y}) \rangle + \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$.

定義 38.3.1(1) の前半の条件 $p|_W = \text{id}_W$ は内積を用いずとも定義できる性質である. 一般に, この様な性質を満たす線形写像を射影とよぶ (詳しくは 38.5 節で述べる). 一方, 定義 38.3.1(1) の後半の条件「各 $\mathbf{x}, \mathbf{y} \in U$ について $\langle p(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, p(\mathbf{y}) \rangle$ 」は対称行列との関係が深い. そこで次節でこれについて論じたうえで, 定義 38.3.1 の条件の同値性を示そう.

38.4 対称行列の性質

定義 38.3.1(1) の後半の条件は, p の表現行列が対称であることに他ならない (命題 38.4.2).

補題 38.4.1. n 次実正方行列 A について, 次は同値である.

- (1) A は対称行列である, (2) 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について, $\langle \mathbf{x}, A\mathbf{y} \rangle = \langle A\mathbf{x}, \mathbf{y} \rangle$.

Proof. (1) \Rightarrow (2): 命題 36.3.8 より, $\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, {}^t A\mathbf{y} \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle$.

(2) \Rightarrow (1): A の (i, j) -成分は, 標準基底を用いて $\langle \mathbf{e}_i, A\mathbf{e}_j \rangle$ とかけることから, $\langle \mathbf{e}_i, A\mathbf{e}_j \rangle = \langle \mathbf{e}_j, A\mathbf{e}_i \rangle$ を示せば $A = {}^t A$ を得る. (2) より, $\langle \mathbf{e}_j, A\mathbf{e}_i \rangle = \langle A\mathbf{e}_j, \mathbf{e}_i \rangle = \langle \mathbf{e}_i, A\mathbf{e}_j \rangle$. \square

命題 38.4.2. 体 \mathbb{R} 上の有限次元内積空間 U における線形変換 $f: U \rightarrow U$ について次は同値である:

- (1) 各 $\mathbf{x}, \mathbf{y} \in U$ について $\langle f(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, f(\mathbf{y}) \rangle$,
 (2) U の任意の正規直交基底に関する f の表現行列は対称である.
 (3) f の表現行列が対称となるような U 正規直交基底が少なくとも一つ存在する.

Proof. (1) \Rightarrow (2): U の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を任意に取り, その表現行列を A とする. 命題 35.6.5 における線形同型 $F: U \rightarrow \mathbb{R}^n$ によって翻訳すれば, 各 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ について $\langle A\mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}} = \langle f(F^{-1}(\mathbf{x})), F^{-1}(\mathbf{y}) \rangle = \langle F^{-1}(\mathbf{x}), f(F^{-1}(\mathbf{y})) \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle_{\mathbb{R}}$. ゆえに, 前補題より A は対称である.

(2) \Rightarrow (3): 明らか.

(3) \Rightarrow (1): 仮定より, f の表現行列 A が対称となるような, U の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が存在する. 命題 35.6.5 における線形同型 $F: U \rightarrow \mathbb{R}^n$ によって翻訳すれば, 各 $\mathbf{u}, \mathbf{v} \in U$ について, $\langle f(\mathbf{u}), \mathbf{v} \rangle = \langle AF(\mathbf{u}), F(\mathbf{v}) \rangle_{\mathbb{R}} = \langle F(\mathbf{u}), AF(\mathbf{v}) \rangle_{\mathbb{R}} = \langle \mathbf{u}, f(\mathbf{v}) \rangle$. \square

【補足】上の条件のいずれか(したがってすべて)を満たす線形変換(正規直交基底による表現行列が必ず対称行列になる変換)を「対称変換」と呼んでもよいが、その場合は、鏡映変換におけるベクトルの位置の対称性と混同せぬよう注意すること。なお、この変換は、複素内積空間におけるエルミート変換(自己随伴変換)と呼ぶものである。

定義 38.3.1 で述べた条件の同値性を示そう。

補題 38.4.3. 定義 38.3.1 で与えた線形性を仮定しない写像 $p: U \rightarrow W$ に関する次の条件は同値である。

- (1) $p|_W = \text{id}_W$ かつ、各 $\mathbf{x}, \mathbf{y} \in U$ について $\langle p(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, p(\mathbf{y}) \rangle$,
- (2) W の任意の正規直交基底 $\mathbf{w}_1, \dots, \mathbf{w}_k$ について、 $p(\mathbf{x}) = \sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i$.
- (3) $p(\mathbf{x}) = \sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i$ を満たす W の正規直交基底 $\mathbf{w}_1, \dots, \mathbf{w}_k$ が存在する。
- (4) 各 $\mathbf{x} \in U$ について、 $\mathbf{x} - p(\mathbf{x}) \in W^\perp$.

Proof. (1) \Rightarrow (2): $p|_W = \text{id}_W$ より $p(\mathbf{w}_i) = \mathbf{w}_i$ である。これと後半の性質を合わせると $\langle \mathbf{w}_i, p(\mathbf{x}) \rangle = \langle p(\mathbf{w}_i), \mathbf{x} \rangle = \langle \mathbf{w}_i, \mathbf{x} \rangle$ 。したがって、 W の元 $p(\mathbf{x})$ を正規直交基底に分解すれば(命題 35.3.5), $p(\mathbf{x}) = \sum_{i=1}^k \langle \mathbf{w}_i, p(\mathbf{x}) \rangle \mathbf{w}_i = \sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i$ 。

(2) \Rightarrow (3): 明らか。

(3) \Rightarrow (1): $p(\mathbf{x}) = \sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i$ の線形性および $p(\mathbf{w}_i) = \mathbf{w}_i$ から $p|_W = \text{id}_W$ である。次に、 $\mathbf{w}_1, \dots, \mathbf{w}_k$ にいくつかの元を加えて U の正規直交基底 $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ を取る。このとき、各 $i = k+1, \dots, n$ について $p(\mathbf{v}_i) = \mathbf{0}$ ゆえ $p(\mathbf{x})$ の表現行列は対角成分に 1 と 0 が並ぶ対称行列である。したがって命題 38.4.2 より $\langle p(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, p(\mathbf{y}) \rangle$ 。

(3) \Rightarrow (4): $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ を W^\perp の正規直交基底とすれば $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_1, \dots, \mathbf{v}_\ell$ は U の正規直交基底であり(命題 35.5.4), $\mathbf{x} = \sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i + \sum_{i=1}^\ell \langle \mathbf{v}_i, \mathbf{x} \rangle \mathbf{v}_i = p(\mathbf{x}) + \sum_{i=1}^\ell \langle \mathbf{v}_i, \mathbf{x} \rangle \mathbf{v}_i$ 。ゆえに $\mathbf{x} - p(\mathbf{x}) = \sum_{i=1}^\ell \langle \mathbf{v}_i, \mathbf{x} \rangle \mathbf{v}_i \in W^\perp$ 。

(4) \Rightarrow (2): 仮定より、 $i = 1, \dots, k$ について $\langle \mathbf{w}_i, \mathbf{x} - p(\mathbf{x}) \rangle = 0$ である。すなわち、 $\langle \mathbf{w}_i, \mathbf{x} \rangle = \langle \mathbf{w}_i, p(\mathbf{x}) \rangle$ であり、これらを係数とする線形結合をとれば、

$$\sum_{i=1}^k \langle \mathbf{w}_i, \mathbf{x} \rangle \mathbf{w}_i = \sum_{i=1}^k \langle \mathbf{w}_i, p(\mathbf{x}) \rangle \mathbf{w}_i.$$

上の右辺は $p(\mathbf{x})$ に等しい(命題 35.3.5)。 □

38.5 射影変換(発展)

本章では主に内積空間上の線形変換について述べてきたが、本節に限り、 U は内積を持つことを仮定しない一般の線形空間とする。

定義 38.5.1. 線形空間 U とその部分空間 $W \subset U$ が与えられているとき、 $p|_W = \text{id}_W$ を満たす線形写像 $p: U \rightarrow W$ を W への射影(**projection**)という。また、 $r(\mathbf{x}) = 2p(\mathbf{x}) - \mathbf{x}$ で定められる線形変換 $r: U \rightarrow U$ を、射影 p から定まる鏡映(**refraction**)という。

ここで定義した射影は第 i 座標への射影 $\text{pr}_i: \mathbb{R}^n \rightarrow \mathbb{R}$ とは異なる概念ゆえ注意すること。

例 38.5.2. (1) $a \in \mathbb{R}$ を定数とし、 $A = \begin{bmatrix} 1 & a \\ 0 & 0 \end{bmatrix}$ とすれば、 $T_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ は第 1 座標軸 $W = \langle \mathbf{e}_1 \rangle$ への射影である。とくに、 $a = 0$ のとき、標準的な内積について T_A は直交射影となる。

(2) W の基底 $\mathbf{w}_1, \dots, \mathbf{w}_k$ にいくつかの元を加えて U の基底 $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ を取っておく。このとき各 $\mathbf{x} \in U$ は、 $\mathbf{x} = \sum_{i=1}^k r_i \mathbf{w}_i + \sum_{i=k+1}^n r_i \mathbf{v}_i$ と一意的に書ける。そこで、この \mathbf{x} を $\sum_{i=1}^k r_i \mathbf{w}_i \in W$ に写す写像を $p: U \rightarrow W$ と定めれば、これは W への射影である。このように、 $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ の取り方に応じて様々な W への射影を考えることができる。

$p: U \rightarrow W$ が $p|_W = \text{id}_W$ を満たすとき、 $p \circ p = p$ (つまり $p^2 = p$) が成り立つ (命題 38.5.4(1) の証明を見よ). この事実に着目し、部分空間を表に出さずに射影を定義することもできる:

定義 38.5.3. 線形変換 $p: U \rightarrow U$ が $p \circ p = p$ を満たすとき、これを射影変換という.

この射影変換は、部分空間への射影と本質的には同じものである:

命題 38.5.4. (1) W への射影 $p: U \rightarrow W$ が与えられているとき、これを $p: U \rightarrow U$ と見れば、 p は射影変換である.

(2) 射影変換 $p: U \rightarrow U$ において $W := \text{Im } p$ とおけば、線形写像 $p: U \rightarrow W$ は W への射影である.

(3) 射影変換 $p: U \rightarrow U$ から定まる鏡映変換を $r: U \rightarrow U$ とすれば $p = p \circ r$. すなわち、各 $\mathbf{x} \in U$ について \mathbf{x} と $r(\mathbf{x})$ は p によって同じ位置に射影される.

Proof. (1): 各 $\mathbf{x} \in U$ について、 $\mathbf{w} := p(\mathbf{x}) \in W$ とおくと、 $p(\mathbf{w}) = \mathbf{w}$ ゆえ $p(p(\mathbf{x})) = p(\mathbf{w}) = \mathbf{w} = p(\mathbf{x})$. すなわち $p^2 = p$ である.

(2): $\mathbf{w} \in W = \text{Im } p$ を取れば、像の定義から $p(\mathbf{x}) = \mathbf{w}$ を満たす $\mathbf{x} \in U$ が存在する. このとき、 $p(p(\mathbf{x})) = p(\mathbf{x})$ ゆえ $p(\mathbf{w}) = \mathbf{w}$. よって $p|_W = \text{id}_W$ である.

(3): $p^2 = p$ に注意すれば、 $p(r(\mathbf{x})) = p(2p(\mathbf{x}) - \mathbf{x}) = 2p^2(\mathbf{x}) - p(\mathbf{x}) = 2p(\mathbf{x}) - p(\mathbf{x}) = p(\mathbf{x})$. \square

次は、直交射影の場合については既に論じていたことである.

命題 38.5.5. $p: U \rightarrow U$ を射影変換とし、線形変換 $q: U \rightarrow U$ を $q(\mathbf{x}) := \mathbf{x} - p(\mathbf{x})$ と定めれば、次が成り立つ.

(1) $q \circ p = \mathbf{0}$. さらに $\text{Ker } q = \text{Im } p$.

(2) q も射影変換である.

(3) $p \circ q = \mathbf{0}$. さらに $\text{Ker } p = \text{Im } q$.

(4) $\text{Ker } p \oplus \text{Im } p = U$.

Proof. (1): $q(p(\mathbf{x})) = p(\mathbf{x}) - p(p(\mathbf{x})) = p(\mathbf{x}) - p(\mathbf{x}) = \mathbf{0}$. ゆえに $\text{Im } p \subset \text{Ker } q$ である. 一方、 $\mathbf{x} \in \text{Ker } q$ とすれば、 $\mathbf{x} - p(\mathbf{x}) = \mathbf{0}$. つまり $\mathbf{x} = p(\mathbf{x}) \in \text{Im } p$. ゆえに $\text{Ker } q \subset \text{Im } p$ である.

(2): $q(q(\mathbf{x})) = q(\mathbf{x} - p(\mathbf{x})) = q(\mathbf{x}) - q(p(\mathbf{x})) = q(\mathbf{x}) - \mathbf{0} = q(\mathbf{x})$.

(3): q が射影変換であること、および $p(\mathbf{x}) = \mathbf{x} - q(\mathbf{x})$ に注意すれば、 p と q の立場を入れかえることにより (1) から直ちに導かれる.

(4): 各 $\mathbf{x} \in U$ に対して、 $\mathbf{x} = q(\mathbf{x}) + p(\mathbf{x})$ であること、および $q(\mathbf{x}) \in \text{Im } q = \text{Ker } p$ 、 $p(\mathbf{x}) \in \text{Im } p$ より、 $\text{Ker } p \cup \text{Im } p$ は U を生成する. また、 $\mathbf{z} + \mathbf{w} = \mathbf{0}$ かつ $\mathbf{z} \in \text{Ker } p$ 、 $\mathbf{w} \in \text{Im } p$ を仮定すれば、 $\mathbf{z} = -\mathbf{w} \in \text{Ker } p \cap \text{Im } p$ である. このとき $\mathbf{z} \in \text{Ker } p$ より $p(\mathbf{z}) = \mathbf{0}$. また $\mathbf{z} \in \text{Im } p$ および $p|_{\text{Im } p} = \text{id}_{\text{Im } p}$ より (命題 38.5.4(2))、 $p(\mathbf{z}) = \mathbf{z}$. ゆえに $\mathbf{z} = \mathbf{0}$ および $\mathbf{w} = -\mathbf{z} = \mathbf{0}$ を得る. 以上より $U = \text{Ker } p \oplus \text{Im } p$. \square

直交射影とその鏡映において、表現行列を対角行列に取れることは既に述べた通りである (補題 38.4.3 および命題 38.3.4 の証明を見よ). 一般の射影についても、これと同様の事実が成り立つ:

命題 38.5.6. 射影 $p: U \rightarrow U$ について次が成り立つ.

(1) p の固有値は 0 または 1 である.

(2) p の固有値が 0 のみならば、 $p = \mathbf{0}$.

(3) p の固有値が 1 のみならば、 $p = I$.

(4) $\dim U < \infty$ のとき、 p の表現行列は対角化可能である.

(5) 鏡映 $r = 2p - I$ の固有値は 1 または -1 である.

(6) $\dim U < \infty$ のとき, r の表現行列は対角化可能である.

Proof. (1): λ を p の固有値とすれば, $p(\mathbf{x}) = \lambda\mathbf{x}$ を満たす固有ベクトル $\mathbf{x} \in U \setminus \{\mathbf{0}\}$ が存在する. このとき, $\lambda\mathbf{x} = p(\mathbf{x}) = p^2(\mathbf{x}) = \lambda^2\mathbf{x}$ ゆえ $(\lambda^2 - \lambda)\mathbf{x} = \mathbf{0}$. ここで $\mathbf{x} \neq \mathbf{0}$ より $\lambda^2 - \lambda = 0$. つまり λ は 0 と 1 のいずれかに等しい.

(2): $p|_{\text{Im } p} = \text{id}_{\text{Im } p}$ であるから (命題 38.5.4(2)), $\mathbf{0}$ でない $\text{Im } p$ の元は固有値 1 の固有ベクトルである. いま, 1 は固有値ではないから, ゆえに $\text{Im } p$ は $\mathbf{0}$ 以外の元を含まず, $\text{Im } p = \{\mathbf{0}\}$. つまり, p はすべてのベクトルを $\mathbf{0}$ に写す写像であり, ゆえに $p = \mathbf{0}$.

(3): $\mathbf{0}$ でない $\text{Ker } p$ の元は固有値 0 の固有ベクトルである. いま, 0 は固有値ではないから, ゆえに $\text{Ker } p$ は $\mathbf{0}$ 以外の元を含まず, $\text{Ker } p = \{\mathbf{0}\}$. 命題 38.5.5(4) より $U = \text{Ker } p \oplus \text{Im } p = \{\mathbf{0}\} \oplus \text{Im } p = \text{Im } p$ であり, ゆえに $p = p|_U = p|_{\text{Im } p} = \text{id}|_{\text{Im } p} = \text{id}_U$.

(4): $p = \mathbf{0}, I$ の場合は明らかゆえ, そうでない場合を考える. このとき, (2) と (3) より 0 と 1 はいずれも p の固有値である. $\text{Ker } p = W(0, p)$ および $\text{Im } p = \text{Ker } q = W(1, p)$ に注意すれば, $U = \text{Ker } p \oplus \text{Im } p = W(0, p) \oplus W(1, p)$. つまり $\dim U = \dim W(0, p) + \dim W(1, p)$ ゆえ p は対角化可能である.

(5): λ を r の固有値, \mathbf{x} をその固有ベクトルの一つとする (つまり $r(\mathbf{x}) = \lambda\mathbf{x}$). $p = \frac{1}{2}(r + I)$ より, $p(\mathbf{x}) = \frac{1}{2}(r(\mathbf{x}) + \mathbf{x}) = \frac{1}{2}(\lambda + 1)\mathbf{x}$. つまり \mathbf{x} は, p の固有値 $\frac{1}{2}(\lambda + 1)$ に関する固有ベクトルでもあり, (1) より $\frac{1}{2}(\lambda + 1)$ は 0 または 1 に等しい. したがって λ は -1 か 1 のいずれかである.

(6): U の基底を決めておき, その基底による p の表現行列を A とする. このとき, r の表現行列は $B = 2A - E$ である. $D = P^{-1}AP$ を A の対角化とすれば, $P^{-1}BP = P^{-1}(2A - E)P = P^{-1}(2A)P - P^{-1}EP = 2D - E$ は B の対角化である. \square

【備考】 $\text{Im } p$ と $\text{Ker } p$ の基底をそれぞれ $\mathbf{w}_1, \dots, \mathbf{w}_k$ および $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ とすれば, 命題 38.5.5(4) より $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_1, \dots, \mathbf{v}_\ell$ は U の基底である. このとき p は, $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{u}_1, \dots, \mathbf{u}_\ell$ を $\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{0}, \dots, \mathbf{0}$ に写す. したがって, この基底による p の表現行列は 1 と 0 が対角成分に並ぶ対角行列である. また, r は上の基底を $\mathbf{w}_1, \dots, \mathbf{w}_k, -\mathbf{v}_1, \dots, -\mathbf{v}_\ell$ に写す. よって, この基底による r の表現行列は ± 1 が対角成分に並ぶ対角行列である.

射影変換 $p: U \rightarrow U$ の表現行列を A とすれば, 基底の取り方によらずに $A^2 = A$ が成り立つ. 一般に, $A^2 = A$ を満たす行列のことを冪等行列 (**idempotent matrix**) という.

第39章 エルミート内積

体 \mathbb{C} 上の線形空間の内積について概観しよう. 本章で紹介する概念を通して, 例えば「実正方行列の直交行列による上三角化」が複素行列に対してどのように一般化されるかを理解することができる. また, エルミート内積の応用例として, フーリエ級数の展開公式が自然と導けるようになる.

39.1 エルミート内積とノルム

$z = {}^t(z_1, \dots, z_n), w = {}^t(w_1, \dots, w_n) \in \mathbb{C}^n$ について, z と w の内積を次のように安直に定めるわけにはいかない:

$$z \cdot w = z_1 w_1 + \dots + z_n w_n.$$

このように定めてしまうと, 一般には $z \cdot z$ が正数にはならず, ベクトルの長さを $\|z\| = \sqrt{z \cdot z}$ と定めるのに適さない. そこで, $z \in \mathbb{C}$ について

$$\bar{z}z \geq 0, \quad |z| = \sqrt{\bar{z}z}$$

であることから類推し, 複素ベクトルにおける内積を次のように定める:

定義 39.1.1. 列ベクトル $z, w \in \mathbb{C}^n$ に対して定められる複素数 $\langle z, w \rangle := \bar{z}w$ を標準的なエルミート内積という. すなわち, $z = {}^t(z_1, \dots, z_n), w = {}^t(w_1, \dots, w_n)$ と成分表示すれば,

$$\langle z, w \rangle = \bar{z}_1 w_1 + \dots + \bar{z}_n w_n.$$

【注意】 $x, y \in \mathbb{R}^n$ について, これらの標準的な内積と標準的なエルミート内積は等しい. そこで, 標準的なエルミート内積を単に標準的な内積と呼ぶ場合もある. すなわち, \mathbb{C}^n の標準的な内積を上のように定義し, その \mathbb{R}^n への制限として, \mathbb{R}^n 上の標準的な内積が定められるのである.

以下, 誤解がない範囲において, 標準的なエルミート内積を単にエルミート内積 (あるいは内積) と呼ぶ.

備考 39.1.2. 「エルミート内積」という語句を次の意味でも用いる:

- $z, w \in \mathbb{C}^n$ をともに行ベクトルとするとき, これらのエルミート内積を $z {}^t w$ と定める.
- $z \in \mathbb{C}^n$ を行ベクトル, $z \in \mathbb{C}^n$ を列ベクトルとするとき, これらのエルミート内積を $z w$ と定める.

次は難しくないゆえ証明は略す.

命題 39.1.3. $z, w, u \in \mathbb{C}^n$ および $c \in \mathbb{C}$ について次が成り立つ.

- (1) $\langle z, w \rangle = \overline{\langle w, z \rangle}$,
- (2) $\langle z, w + u \rangle = \langle z, w \rangle + \langle z, u \rangle, \quad \langle z + w, u \rangle = \langle z, u \rangle + \langle w, u \rangle$,
- (3) $\langle cz, w \rangle = \bar{c} \langle z, w \rangle, \quad \langle z, cw \rangle = c \langle z, w \rangle$,
- (4) $\langle z, z \rangle \geq 0$,
- (5) $\langle z, z \rangle = 0 \iff z = \mathbf{0}$.

【備考】 (2) の後半は, 前半と (1) から導ける. 実際, $\langle z + w, u \rangle = \overline{\langle u, z + w \rangle} = \overline{\langle u, z \rangle + \langle u, w \rangle} = \overline{\langle u, z \rangle} + \overline{\langle u, w \rangle} = \langle z, u \rangle + \langle w, u \rangle$. また, (3) の一方は他方と (1) から導ける.

備考 39.1.4. エルミート内積を $\langle z, w \rangle := {}^t z \bar{w} = z_1 \bar{w}_1 + \cdots + z_n \bar{w}_n$ と定める流儀もある. その場合, 上の性質 (3) は次に置き換わるゆえ注意せよ:

$$(3) \langle cz, w \rangle = c \langle z, w \rangle, \quad \langle z, cw \rangle = \bar{c} \langle z, w \rangle.$$

\mathbb{R}^n の内積に関する性質が一般化されて論じられたように, エルミート内積も次のように一般化して論じることができる. エルミート内積の実用的な例は 39.8 節で与える (フーリエ級数展開).

定義 39.1.5. (1) U を体 \mathbb{C} 上の線形空間とする. 各 $x, y \in U$ に対して定められる関数 $\langle x, y \rangle$ が命題 39.1.3 のすべての性質をみたすとき, これを U 上のエルミート内積 (Hermitian inner product) (あるいは単に内積) という. このとき, 空間と内積の組 $(U, \langle \cdot, \cdot \rangle)$ を (複素) 内積空間あるいはユニタリ空間 (unitary space), エルミート空間 (Hermitian space)¹ などという.

(2) 体 \mathbb{C} 上の内積空間 $(U, \langle \cdot, \cdot \rangle)$ が与えられているとき, 各 $x \in U$ に対して $\|x\| := \sqrt{\langle x, x \rangle}$ で定められる関数 $\|\cdot\| : U \rightarrow [0, \infty)$ を内積 $\langle \cdot, \cdot \rangle$ から定まるノルムという.

本書では, 標準的なエルミート内積であることを強調するとき, これを $\langle \cdot, \cdot \rangle_{\mathbb{C}}$ と書く.

39.2 複素内積空間における直交性

本節では, とくに断りがなくとも U を有限次元複素内積空間とする. $x, y \in U$ が $\langle x, y \rangle = 0$ を満たすとき, x と y は直交するという. x と y が直交するとき, $\langle y, x \rangle = \overline{\langle x, y \rangle} = \overline{0} = 0$ ゆえ y と x も直交する. 次に挙げる主張は, すべて参考に挙げた命題と同様にして証明することができる.

命題 39.2.1 (参考: 命題 35.3.2). ベクトルの組 $u_1, \dots, u_n \in U \setminus \{0\}$ が互いに直交するならば, 線形独立である.

定義 39.2.2. 部分空間 $W \subset U$ の基底 $u_1, \dots, u_k \in W$ が $\langle u_i, u_j \rangle = \delta_{ij}$ を満たすとき, これを W の正規直交基底という.

命題 39.2.3 (参考: 命題 35.3.5). 部分空間 $W \subset U$ の正規直交基底 $u_1, \dots, u_k \in W$ および $x \in W$ について,

$$x = \sum_{i=1}^k \langle u_i, x \rangle u_i.$$

備考 39.2.4 (参考: 備考 35.3.6). u_1, \dots, u_n を U の正規直交基底とし, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ を線形変換とする. このとき前命題より $f(u_j) = \sum_{i=1}^n \langle u_i, f(u_j) \rangle u_i$ であるから, この基底に関する f の表現行列を $A = [a_{ij}]$ とすれば, $a_{ij} = \langle u_i, f(u_j) \rangle$ である.

定理 39.2.5 (参考: 定理 35.4.1). 部分空間 $W \subset U$ の基底 $u_1, \dots, u_k \in W$ が与えられているとする. さらに, 次のように再帰的に $\tilde{v}_1, \dots, \tilde{v}_k$ および v_1, \dots, v_k を定める.

- $\tilde{v}_1 := u_1$ および $v_1 := \frac{1}{\|\tilde{v}_1\|} u_1$ と定義する.
- 既に v_1, \dots, v_{i-1} が定まっているとき,

$$\tilde{v}_i := u_i - \sum_{j=1}^{i-1} \langle v_j, u_i \rangle v_j$$
 および $v_i := \frac{1}{\|\tilde{v}_i\|} \tilde{v}_i$ と定義する.

こうして得られた組 v_1, \dots, v_k は W の正規直交基底である.

¹ 「エルミート空間」は本書の定義よりも弱い意味で用いられることもある. その流儀では, $\langle x, y \rangle$ が命題 39.1.3 の (1) から (3) を満たすとき, これをエルミート形式と呼び, エルミート形式が与えられた複素線形空間をエルミート空間とよぶ.

系 39.2.6 (参考: 系 35.4.2). 部分空間 $W \subset U$ および, 互いに直交する長さ 1 のベクトルの組 $w_1, \dots, w_r \in W$ が与えられているとする. このとき, いくつかの W の元をこれらに加えて, W の正規直交基底にできる.

定義 39.2.7. 部分空間 $W \subset U$ に対して定められる次の集合を W の直交補空間と呼ぶ:

$$W^\perp = \{x \in U \mid \text{各 } w \in W \text{ について } \langle x, w \rangle = 0\}.$$

補題 39.2.8 (参考: 補題 35.5.3). $x, w_1, \dots, w_k \in U$ とする.

- (1) x が各 w_1, \dots, w_k と直交するならば, x は $\langle w_1, \dots, w_k \rangle$ の任意の元とも直交する.
- (2) $x \in \langle w_1, \dots, w_k \rangle^\perp \iff$ 各 $i = 1, \dots, k$ について $\langle x, w_i \rangle = 0$.

命題 39.2.9 (参考: 命題 35.5.4). 部分空間 $W \subset U$ に対して次が成り立つ.

- (1) w_1, \dots, w_k を W の正規直交基底とする. いくつかの U の元をこれらに加えた組 $w_1, \dots, w_k, v_1, \dots, v_{n-k}$ が \mathbb{C}^n の正規直交基底となるならば, v_1, \dots, v_{n-k} は W^\perp の正規直交基底である.
- (2) $\dim W^\perp = n - \dim W$.
- (3) $W \oplus W^\perp = U$. すなわち, w_1, \dots, w_k を W の基底, v_1, \dots, v_r を W^\perp の基底とすれば, これらを合わせた組 $w_1, \dots, w_k, v_1, \dots, v_r$ は U の基底である.
- (4) 部分空間 $V \subset U$ が次の条件を満たすならば, $V = W^\perp$ である.
 - (i) $W \oplus V = U$,
 - (ii) $w \in W$ かつ $v \in V \iff \langle w, v \rangle = 0$.
(すなわち, W の元と V の元はつねに直交する.)
- (5) $(W^\perp)^\perp = W$.

線形変換とエルミート内積に関する情報は, 次のように \mathbb{C}^n 上の性質に帰着される.

命題 39.2.10 (参考: 命題 35.6.5). U を有限次元複素内積空間とし, u_1, \dots, u_n を U の正規直交基底とする. このとき, 各 u_i ($i = 1, \dots, n$) を $e_i \in \mathbb{C}^n$ に写す線形同型 $F: U \rightarrow \mathbb{C}^n$ について次が成り立つ:

- (1) F は内積を保つ. すなわち,
 - (i) 各 $u, v \in U$ について $\langle u, v \rangle = \langle F(u), F(v) \rangle_{\mathbb{C}}$.
 - (ii) 各 $x, y \in \mathbb{C}^n$ について $\langle x, y \rangle_{\mathbb{C}} = \langle F^{-1}(x), F^{-1}(y) \rangle$.
- (2) 基底 u_1, \dots, u_n に関する線形変換 $f, g: U \rightarrow U$ の表現行列を A, B とすれば,
 - (i) 各 $u, v \in U$ について, $\langle f(u), g(v) \rangle = \langle AF(u), BF(v) \rangle_{\mathbb{C}}$,
 - (ii) 各 $x, y \in \mathbb{R}^n$ について, $\langle Ax, By \rangle_{\mathbb{C}} = \langle f(F^{-1}(x)), g(F^{-1}(y)) \rangle$.
 - (iii) 各 $u, v \in U$ について, $\langle f(u), f(v) \rangle = \langle AF(u), AF(v) \rangle_{\mathbb{C}}$,
 - (iv) 各 $x, y \in \mathbb{R}^n$ について, $\langle Ax, Ay \rangle_{\mathbb{C}} = \langle f(F^{-1}(x)), f(F^{-1}(y)) \rangle$.
 - (v) 各 $u, v \in U$ について, $\langle f(u), v \rangle = \langle AF(u), F(v) \rangle_{\mathbb{C}}$,
 - (vi) 各 $x, y \in \mathbb{R}^n$ について, $\langle Ax, y \rangle_{\mathbb{C}} = \langle f(F^{-1}(x)), F^{-1}(y) \rangle$.

Proof. 命題 35.6.5(1) の証明における式変形が $\left\langle \sum_{i=1}^n r_i u_i, v \right\rangle = \sum_{i=1}^n \bar{r}_i \langle u_i, v \rangle$ と変更される点に注意すれば, あとは類似する戦略ですべて示せる. □

39.3 エルミート内積から定まるノルムの性質

ピタゴラスの定理, コーシー-シュワルツの不等式, 三角不等式は複素内積空間においても成り立つ.

定理 39.3.1 (参考: 定理 38.1.1). 複素内積空間 U における二つのベクトル $\mathbf{x}, \mathbf{y} \in U$ が直交するとき, 次が成り立つ:

- (1) $\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2$.
- (2) $\|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$ であり, 等号成立は $\mathbf{x} = \mathbf{0}$ のときに限る.
- (3) $\|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\|$ であり, 等号成立は $\mathbf{y} = \mathbf{0}$ のときに限る.

定理 39.3.2 (参考: 定理 38.1.2). U を複素内積空間とすれば, 各 $\mathbf{x}, \mathbf{y} \in U$ について $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$ が成り立つ. また, $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ の場合における等号成立は \mathbf{x} と \mathbf{y} のいずれかが他方のスカラー倍であるときに限る.

複素数 z の実部を $\Re z$, 虚部を $\Im z$ で表す².

命題 39.3.3 (参考: 命題 38.1.3). U を複素内積空間とすれば, $\mathbf{x}, \mathbf{y}, \mathbf{z} \in U$ について次が成り立つ.

- (1) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$,
- (2) $\|\mathbf{x} - \mathbf{z}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\|$.

Proof. $\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} = 2\Re \langle \mathbf{x}, \mathbf{y} \rangle \leq 2|\langle \mathbf{x}, \mathbf{y} \rangle| \leq 2\|\mathbf{x}\| \cdot \|\mathbf{y}\|$ に注意すれば, 命題 38.1.3 と同様にして示せる. □

39.4 随伴行列と随伴変換

36 章において転置行列および直交行列, 対称行列が演じていた役回りは, 本章の文脈では, それぞれ次の概念に置き換わる.

定義 39.4.1. A を複素行列とする.

- \overline{A} を A の随伴 (adjoint) あるいは共役転置 (conjugate transpose), 転置共役 (transjugate) などよび, これを A^* で表す (転置行列のエルミート版).
- $A^{-1} = A^*$ を満たす正方行列をユニタリ行列 (unitary matrix) という (直交行列のエルミート版).
- $A^* = A$ を満たす正方行列をエルミート行列 (Hermitian matrix) あるいは自己随伴行列 (self-adjoint matrix) という (対称行列のエルミート版).

練習 39.4.2. 実正方行列 A について次を示せ.

- (1) A の転置行列は A の随伴行列に等しい.
- (2) A が直交行列であることと, A がユニタリ行列であることは同値である.
- (3) A が対称行列であることと, A がエルミート行列であることは同値である.

次は本章において断りなく用いる.

備考 39.4.3. (1) 列ベクトルの標準的なエルミート内積は, $\langle \mathbf{z}, \mathbf{w} \rangle_{\mathbb{C}} = \mathbf{z}^* \mathbf{w}$ と表すことができる.

² z の実部と虚部をそれぞれ $\Re z, \Im z$ と書く流儀もある. 後者を線形写像の像と混同しないように.

(2) $(AB)^* = B^*A^*$. 実際, $(AB)^* = \overline{(AB)} = \overline{tB^tA} = \overline{tB}^t \overline{tA} = B^*A^*$.

(3) $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$ および $A \in M_n(\mathbb{C})$ について, $\langle A\mathbf{z}, \mathbf{w} \rangle = \langle \mathbf{z}, A^*\mathbf{w} \rangle$. 実際, $\langle A\mathbf{z}, \mathbf{w} \rangle = (A\mathbf{z})^* \mathbf{w} = \mathbf{z}^* A^* \mathbf{w} = \langle \mathbf{z}, A^* \mathbf{w} \rangle$.

(4) A^*A の (i, j) -成分は, A の i 列と j 列のエルミート内積 $\langle A\mathbf{e}_i, A\mathbf{e}_j \rangle$ に等しい. また, AA^* の (i, j) -成分は, A の j 行目と i 行目のエルミート内積に等しい (i と j の順が入れ替わる点に注意せよ).

連立1次方程式を複素ベクトルに関する直交性の条件とみなすことで次が成り立つ.

命題 39.4.4 (参考: 命題 35.5.5). (1) $\mathbf{w}_1, \dots, \mathbf{w}_m \in \mathbb{C}^n$ を列ベクトルとし, $A = [\mathbf{w}_1, \dots, \mathbf{w}_m]^*$ とすれば (これは (m, n) -行列である), $W_A = \langle \mathbf{w}_1, \dots, \mathbf{w}_m \rangle^\perp$.

(2) $V \subset \mathbb{C}^n$ を部分空間とすれば, $V = W_A$ を満たす行列 A が存在する.

直交行列や対称行列に対応する線形変換が定義されたのと同様にして, ユニタリ行列やエルミート行列に対応する線形変換が定義できる. その議論は次節以降にまわすとして, そのまえに, 表現行列の転置行列や随伴行列に対応する線形変換について述べておく.

定義 39.4.5. 有限次元複素内積空間 U 上の線形変換 $f : U \rightarrow U$ に対して, 次の性質を満たす線形変換 $f^* : U \rightarrow U$ が唯一つ存在する.

$$\text{各 } \mathbf{u}, \mathbf{v} \in U \text{ について, } \langle f(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, f^*(\mathbf{v}) \rangle. \quad (39.4.1)$$

この f^* を f の随伴あるいは共役³という.

【注意】

(1) 定義より, $(f^*)^* = f$ である.

(2) 記号 f^* は, f の双対写像 $f^* : U^* \rightarrow U^*$ に対しても用いられるため注意を要する. 線形変換 f の双対と f の随伴は全く別のものであり, 双対を ${}^t f$, 随伴を f^* と書いて区別する流儀もある.

Proof. あらかじめ, U の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を与えておく. もし, $f^* : U \rightarrow U$ が上の性質を満たすとすれば, 各 $\mathbf{v} \in U$ について, 命題 39.2.3 より, $f^*(\mathbf{v}) = \sum_{i=1}^n \langle \mathbf{u}_i, f^*(\mathbf{v}) \rangle \mathbf{u}_i = \sum_{i=1}^n \langle f(\mathbf{u}_i), \mathbf{v} \rangle \mathbf{u}_i$ となる. したがって,

$$\text{各 } \mathbf{v} \in U \text{ に対して, } f^*(\mathbf{v}) := \sum_{i=1}^n \langle f(\mathbf{u}_i), \mathbf{v} \rangle \mathbf{u}_i$$

と定めるしかない. すなわち, 条件 (39.4.1) を満たす f^* があるとすれば, それは上式で定められるもの唯一つである. あとは, 上式で定義した f^* が線形性と条件 (39.4.1) を満たすことを確認すればよい.

f^* の線形性は定義からすぐ分るゆえ略す. f^* が条件 (39.4.1) を満たすことを示そう. f の表現行列を $A = [a_{ij}]$ とすれば $a_{ij} = \langle \mathbf{u}_i, f(\mathbf{u}_j) \rangle$ である (備考 39.2.4). また,

$$f^*(\mathbf{u}_j) := \sum_{i=1}^n \langle f(\mathbf{u}_i), \mathbf{u}_j \rangle \mathbf{u}_i = \sum_{i=1}^n \overline{\langle \mathbf{u}_j, f(\mathbf{u}_i) \rangle} \mathbf{u}_i = \sum_{i=1}^n \overline{a_{ji}} \mathbf{u}_i.$$

したがって, f^* の表現行列は A^* である. さて, 各 $\mathbf{u}, \mathbf{v} \in U$ に対して, 命題 39.2.10 で与えた同型 $F : U \rightarrow \mathbb{C}^n$ において $\mathbf{x} := F(\mathbf{u})$, $\mathbf{y} := F(\mathbf{v})$ とおけば, $\langle f(\mathbf{u}), \mathbf{v} \rangle = \langle A\mathbf{x}, \mathbf{y} \rangle_{\mathbb{C}} = \langle \mathbf{x}, A^*\mathbf{y} \rangle_{\mathbb{C}} = \langle \mathbf{u}, f^*(\mathbf{v}) \rangle$. \square

上の証明において, 次の (1) も示していた:

系 39.4.6. (1) U の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する $f : U \rightarrow U$ の表現行列を A とすれば, $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f^* の表現行列は A^* である.

(2) 正方行列 A について, $(T_A)^* = T_{A^*}$.

³ \bar{f} に相当するものはない. 詳しくは本節の後半で述べる.

Proof. (2): 標準基底に関する T_A の表現行列が A であるから, (1) により, 標準基底に関する $(T_A)^*$ の表現行列は A^* である. ゆえに $(T_A)^* = T_{A^*}$ である (系 26.2.2). \square

なお, 体 \mathbb{R} 上の内積空間において次もいえるのであるが, 本書では略した. 証明も上と同じである.

命題 39.4.7. 体 \mathbb{R} 上の有限次元内積空間 U 上の線形変換 $f : U \rightarrow U$ に対して, 次の性質を満たす線形変換 $f^* : U \rightarrow U$ が唯一つ存在する.

$$\text{各 } \mathbf{u}, \mathbf{v} \in U \text{ について, } \langle f(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, f^*(\mathbf{v}) \rangle.$$

また, U の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列を A とすれば, f^* の表現行列は ${}^t A$ である.

改めて U を複素内積空間とする. このとき, 行列 A に対して \bar{A} が考えられたのに対して, $f \in \text{End}(U)$ に対して $\bar{f} : U \rightarrow U$ は定まらない. それは次の理由による: ある正規直交基底に関する $f, g \in \text{End}(U)$ の表現行列がそれぞれ A および \bar{A} であるとし, 別の正規直交基底に関する f の表現行列が B であるとする. このとき, その別に基底に関する g の表現行列は \bar{B} になるとは限らない (次の例をみよ).

例 39.4.8. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ とし, $f = T_A, g = T_{\bar{A}} = T_A : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ とすれば, 標準基底に関する f, g

の表現行列は, それぞれ A および $\bar{A} = A$ である. 一方で, $\mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -i \\ 1 \end{bmatrix}$ とすれば, これらは A の固有ベクトルからなる正規直交基底であり (例 29.2.7), この基底における f の表現行列は $B = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$ である. したがって $g = f$ の表現行列も B であり, これは \bar{B} に一致しない.

また, $f \in \text{End}(U)$ に対して, 転置に相当するような ${}^t f \in \text{End}(U)$ も定まらない. 何故なら, もし定まるとすれば, $\bar{f} = {}^t(f^*)$ として \bar{f} が定められるからである.

備考 39.4.9 (よりみち). 複素内積空間 U の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ と線形変換 $f : U \rightarrow U$ が与えられているとし, この基底に関する f の表現行列が $A = [a_{ij}]$ であるとする (つまり $a_{ij} = \langle \mathbf{u}_i, f(\mathbf{u}_j) \rangle$). このとき, f の双対変換 ${}^t f : U^* \rightarrow U^*$ の, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の双対基底に関する表現行列は ${}^t A$ である. そこで, $\mathbf{u}_1, \dots, \mathbf{u}_n$ とその双対基底を無理矢理に同一視 (つまり U と U^* を同一視) することにより, ${}^t f : U^* \rightarrow U^*$ に対応するような U 上の線形変換を考えることはできる. この写像を $g : U \rightarrow U$ としよう. すなわち g は

$$\mathbf{u}_j \xrightarrow{g} \sum_{i=1}^n \langle \mathbf{u}_j, f(\mathbf{u}_i) \rangle \mathbf{u}_i \quad (= \sum_{i=1}^n a_{ji} \mathbf{u}_i) \quad (39.4.2)$$

を満たす線形変換であり (基底の行き先が決まれば, 他のベクトルの行き先も決まる), このとき $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する g の表現行列は ${}^t A$ である. なお, この写像を

$$g(\mathbf{x}) = \sum_{i=1}^n \langle \mathbf{x}, f(\mathbf{u}_i) \rangle \mathbf{u}_i \quad (39.4.3)$$

と書けない点に注意せよ (エルミート内積を考えているゆえ右辺は線形写像でない). 式 (39.4.2) によって定義される写像は, 正規直交基底の取り方に依存する写像である (例 39.4.10).

なお, U が体 \mathbb{R} 上の内積空間である場合はこの限りではなく, 式 (39.4.3) によって線形変換 $g : U \rightarrow U$ を定めることができ, この定義は正規直交基底の取り方に依存しない. 事実, g は命題 39.4.7 で定めた f^* に一致する.

例 39.4.10 (よりみち). $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ とし, $f = T_A : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ とおく. このとき, 標準基底 $\mathbf{u}_i = \mathbf{e}_i$ ($i = 1, 2$) について式 (39.4.2) で定められる線形変換を g_1 とすれば, $g_1(\mathbf{e}_1) = \mathbf{e}_2$ である (実は $g_1 = T_A$ で

ある). 一方, $\mathbf{u}_1 = \mathbf{e}_1, \mathbf{u}_2 = i\mathbf{e}_2$ とした場合, この $\mathbf{u}_1, \mathbf{u}_2$ も \mathbb{C}^2 の正規直交基底であり, この基底について式 (39.4.2) で定められる線形変換を g_2 とすれば,

$$g_2(\mathbf{e}_1) = \langle \mathbf{e}_1, f(\mathbf{u}_1) \rangle \mathbf{u}_1 + \langle \mathbf{e}_1, f(\mathbf{u}_2) \rangle \mathbf{u}_2 = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle \mathbf{e}_1 + \langle \mathbf{e}_1, i\mathbf{e}_1 \rangle i\mathbf{e}_2 = \mathbf{0} + i^2 \mathbf{e}_2 = -\mathbf{e}_2.$$

つまり, 式 (39.4.2) で定められる写像は正規直交基底の取り方に依存する.

このように, $f \in \text{End}(U)$ に対して無理矢理 $\bar{f}, {}^t f \in \text{End}(U)$ を定めようとする, それは基底の取り方に依存するものになってしまう. n 次複素正方行列 A に対して, $\text{End}(\mathbb{C}^n) (\simeq M_n(\mathbb{C}))$ の元としてあたかも \bar{A} や ${}^t A$ が自然に定まるように思えるのは, 標準基底が先天的に与えられていると我々が認識していることによるのかもしれない.

いずれにせよ, 本節を通して \bar{A} よりも A^* のほうがより自然な概念であることが分かった. この観点から, 次のような呼称も考えられる:

備考 39.4.11. $z \in \mathbb{C}$ を $(1, 1)$ -行列とみれば, $\bar{z} = z^*$ である. すなわち, 複素数の共役とは 1 次元の随伴を意味する. そこで改めて, $A \in M_n(\mathbb{C})$ に対する $A^* \in M_n(\mathbb{C})$ や $f \in \text{End}(U)$ に対する $f^* \in \text{End}(U)$ を共役 (**conjugate**) と呼ぶ流儀もある (\bar{A} のことではないので注意を要する).

39.5 ユニタリ行列

これから, 36 章で論じた諸性質の複素行列版について述べる. その証明の多くは, 標準的な内積および, 転置行列, 直交行列, 対称行列をそれぞれ, 標準的なエルミート内積および, 随伴行列, ユニタリ行列, エルミート行列に置き換えるだけで済む. そのような命題については対応する体 \mathbb{R} 上の内積空間に関する主張の命題番号を参考として挙げるにとどめ, 証明は略す.

命題 39.5.1 (参考: 命題 36.1.1). n 次複素正方行列 A について, 次は同値である.

- (1) A の各列は \mathbb{C}^n 上の正規直交基底をなす,
- (2) A の各行は \mathbb{C}^n 上の正規直交基底をなす,
- (3) A はユニタリ行列である. すなわち, $A^* = A^{-1}$,
- (4) 各 $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ について $\langle A\mathbf{x}, A\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$.

Proof. (2) \Leftrightarrow (3) のみ示す. AA^* の (i, j) -成分を z_{ij} とおくと, 備考 39.4.3(4) より z_{ij} は, A の j 行目と i 行目のエルミート内積に等しい. したがって, (2) を仮定すれば $z_{ij} = \delta_{ji} = \delta_{ij}$ より, $AA^* = E$. さらにこの両辺の随伴を取り $A^*A = E$ を得る. また, (3) を仮定すれば $AA^* = E$ より $z_{ij} = \delta_{ij} = \delta_{ji}$. すなわち, A の j 行目と i 行目のエルミート内積は δ_{ji} に等しい. つまり A の各行は \mathbb{C}^n 上の正規直交基底をなす. \square

複素数 z に対して, $z + \bar{z}$ および $z - \bar{z}$ はそれぞれ実数と純虚数であり,

$$z = \frac{z + \bar{z}}{2} + \frac{z - \bar{z}}{2}$$

が成り立つ. すなわち, $\frac{z + \bar{z}}{2} \in \mathbb{R}$ は z の実部 $\Re z$ に, $\frac{z - \bar{z}}{2}(-i) \in \mathbb{R}$ は z の虚部 $\Im z$ にそれぞれ等しい.

命題 39.5.2 (参考: 命題 38.2.1). 複素内積空間 U の各元 $\mathbf{x}, \mathbf{y} \in U$ において, 次が成り立つ

- (1) $\Re \langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{2}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2)$, (2) $\Re \langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2)$,
- (3) $\Re \langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2)$, (4) $\Im \langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}(\|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2)$,
- (5) $\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$.

Proof. (4)のみ示す. $\Im z = \Re(-iz)$ に注意すれば, (3) より $\Im\langle \mathbf{x}, \mathbf{y} \rangle = \Re(-i\langle \mathbf{x}, \mathbf{y} \rangle) = \Re\langle \mathbf{x}, -i\mathbf{y} \rangle = \frac{1}{4}(\|\mathbf{x} - i\mathbf{y}\|^2 - \|\mathbf{x} + i\mathbf{y}\|^2) = \frac{1}{4}(\|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2)$. \square

上の (3) と (4) より, 次の分極公式 (**polarization identity**) を得る:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2) + \frac{1}{4}(\|\mathbf{x} + i\mathbf{y}\|^2 - \|\mathbf{x} - i\mathbf{y}\|^2)i.$$

この等式は, 次の (1) から (3) の同値性を導く.

命題 39.5.3 (参考: 命題 38.2.2). 有限次元複素内積空間 U 上の線形変換 $f : U \rightarrow U$ について次は同値である:

- (1) 各 $\mathbf{x}, \mathbf{y} \in U$ について $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$.
- (2) 各 $\mathbf{x} \in U$ について $\|f(\mathbf{x})\| = \|\mathbf{x}\|$.
- (3) 各 $\mathbf{x}, \mathbf{y} \in U$ について $\|f(\mathbf{x}) - f(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$.
- (4) U の任意の正規直交基底に関する f の表現行列はユニタリ行列である.
- (5) f の表現行列がユニタリ行列となるような, U の正規直交基底が少なくとも一つ存在する.

上の性質のいずれか (したがってすべて) を満たす線形変換を **ユニタリ変換** という. ユニタリ変換は等長変換である.

39.6 エルミート行列の性質

$A^* = A$ を満たす行列をエルミート行列と呼ぶのであった. これが対称行列のエルミート版に相当することを確認しよう.

補題 39.6.1 (参考: 補題 38.4.1). n 次複素正方行列 A について, 次は同値である.

- (1) A はエルミート行列である,
- (2) 各 $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ について, $\langle \mathbf{x}, A\mathbf{y} \rangle = \langle A\mathbf{x}, \mathbf{y} \rangle$.

Proof. (1) \Rightarrow (2): $\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^*\mathbf{y} \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle$.

(2) \Rightarrow (1): A の (i, j) -成分は, 標準基底を用いて $\langle \mathbf{e}_i, A\mathbf{e}_j \rangle$ とかけることから, $\langle \mathbf{e}_j, A\mathbf{e}_i \rangle = \overline{\langle \mathbf{e}_i, A\mathbf{e}_j \rangle}$ を示せば $A = A^*$ を得る. (2) より, $\langle \mathbf{e}_j, A\mathbf{e}_i \rangle = \langle A\mathbf{e}_j, \mathbf{e}_i \rangle = \overline{\langle \mathbf{e}_i, A\mathbf{e}_j \rangle}$. \square

命題 39.6.2 (参考: 命題 38.4.2). 有限次元複素内積空間 U 上の線形変換 $f : U \rightarrow U$ について次は同値である:

- (1) 各 $\mathbf{x}, \mathbf{y} \in U$ について $\langle f(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, f(\mathbf{y}) \rangle$,
- (2) U の任意の正規直交基底に関する f の表現行列はエルミート行列である.
- (3) f の表現行列がエルミート行列となるような, U の正規直交基底が少なくとも一つ存在する.
- (4) $f^* = f$.

Proof. (1) から (3) の同値性は命題 38.4.2 と同様にして示せる. また, (1) \Leftrightarrow (4) は随伴変換の定義より明らかである. \square

上のいずれか（したがってすべて）の性質を満たす線形変換 $f: U \rightarrow U$ をエルミート変換あるいは自己随伴作用素, 自己共役作用素⁴などという。

固有値がすべて実数となるような実正方行列は, 直交行列を用いて上三角化可能なのであった (定理 36.2.2). 複素行列の固有値はもちろんすべて複素数であり, したがって定理 36.2.2 は次の主張に書き換えられる:

定理 39.6.3 (参考: 定理 36.2.2). 任意の複素正方行列 A は, あるユニタリ行列 P を用いて上三角化できる。

直交行列を用いた対称行列の対角化の議論をエルミート行列の場合に拡張しよう。

補題 39.6.4 (参考: 補題 36.3.5). A を n 次エルミート行列, P を n 次ユニタリ行列とすれば, $P^{-1}AP$ もエルミート行列である。

上三角エルミート行列は実対角行列になることから, 次の定理が示唆される:

定理 39.6.5 (参考: 定理 36.3.4). 複素エルミート行列の固有値はすべて実数である。

Proof. $\lambda \in \mathbb{C}$ を A の固有値, $z \in \mathbb{C}^n$ をその固有ベクトルとすれば,

$$\lambda \langle z, z \rangle = \langle z, \lambda z \rangle = \langle z, Az \rangle = \langle Az, z \rangle = \langle \lambda z, z \rangle = \bar{\lambda} \langle z, z \rangle.$$

$z \neq \mathbf{0}$ ゆえ $\langle z, z \rangle > 0$ であり, したがって $\lambda = \bar{\lambda}$. すなわち λ は実数である。□

系 39.6.6 (参考: 系 36.3.6). 複素正方行列 A について次は同値である。

- (1) A はエルミート行列である。
- (2) A は, あるユニタリ行列 P を用いて, 実対角行列に対角化可能である。

Proof. (1) \Rightarrow (2): 系 36.3.6 と同様にして示せる。

(2) \Rightarrow (1): $D = P^{-1}AP$ を実対角行列とすれば, $A = PDP^{-1} = PDP^*$ ゆえ, $A^* = PD^*P^* = PDP^* = A$. つまり A はエルミート行列である。□

39.7 正規行列

ユニタリ行列によって対角化可能な行列はエルミート行列に限るわけではない。すなわち, 固有値に虚数を含む場合を前節では考えてなかった。

例 39.7.1. 90度回転を表す行列 $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ はユニタリ行列 $P = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$ によって対角化可能である (例 29.2.7)。

複素正方行列 A がユニタリ行列 $P = [p_1, \dots, p_n]$ により対角化されるとき, 正規直交基底 p_1, \dots, p_n に関する T_A の表現行列 D は対角行列である。また, この基底に関する $(T_A)^* = T_{A^*}$ の表現行列は D^* であり, D と D^* は対角行列ゆえ $DD^* = D^*D$ が成り立つ。したがって $T_A \circ T_{A^*} = T_{A^*} \circ T_A$, つまり $AA^* = A^*A$ でなければならない。このような性質を持つ行列 A を正規行列という:

定義 39.7.2. 複素正方行列 A が $AA^* = A^*A$ を満たすとき, A を正規行列 (normal matrix) という。また, 線形変換 $f: U \rightarrow U$ が $f \circ f^* = f^* \circ f$ を満たすとき, f を正規変換 (normal transformation) という。

⁴「自己共役」は $A = \bar{A}$ と $A = A^*$ のいずれの意味にも受け取れる可能性があるものの, ふつうは前者を「 A は実行列である」という。

例 39.7.3. (1) ユニタリ行列 P は, $PP^* = E = P^*P$ を満たすゆえ正規である. とくに直交行列⁵は正規である.

(2) エルミート行列 P は, $PP^* = P^2 = P^*P$ を満たすゆえ正規である. とくに実対称行列は正規である.

(3) $A^* = -A$ を満たす行列を歪エルミート行列 (skew-Hermitian matrix) といい, これは正規である. また, ${}^tA = -A$ を満たす行列を歪対称行列 (skew-symmetric matrix) という. 実歪対称行列は歪エルミート行列ゆえ正規である.

(4) 対角行列は正規である.

本書をここまで読み進めてきた読者には次は明らかであろう.

問題 39.7.4. 有限次元複素内積空間 U 上の線形変換 $f: U \rightarrow U$ について, 次の同値性を示せ.

(1) U の任意の正規直交基底に関する f の表現行列は正規行列である.

(2) f の表現行列が正規行列となるような, U の正規直交基底が少なくとも一つ存在する.

(3) f は正規変換である.

解答例: (2) \Rightarrow (3): 仮定より, f の表現行列 A が正規となるような U の正規直交基底が存在する. この基底に関する f^* および ff^* , f^*f の表現行列は, それぞれ A^* , AA^* , A^*A である (系 39.4.6 および命題 26.2.3). A は正規ゆえ $AA^* = A^*A$ であり, ゆえに $ff^* = f^*f$ (系 26.2.2(1)).

補題 39.7.5. 正規行列 A について次が成り立つ.

(1) P をユニタリ行列とすれば, $P^{-1}AP$ は正規行列である.

(2) A が上三角行列ならば, A は対角行列である.

Proof. (1): $P^* = P^{-1}$ および $(P^{-1})^* = P$ に注意すれば, $(P^{-1}AP)(P^{-1}AP)^* = P^{-1}APP^*A^*(P^{-1})^* = P^{-1}AA^*P$. また $(P^{-1}AP)^*(P^{-1}AP) = P^*A^*(P^{-1})^*P^{-1}AP = P^{-1}A^*AP$ であり, A は正規ゆえこれらは等しい.

(2): A のサイズ n に関する帰納法により示す. $n = 1$ の場合は明らかである. $n - 1$ の場合に主張が成り立つと仮定し, $A = [a_{ij}]$ を n 次上三角正規行列とする. このとき, AA^* および A^*A は次のようになる.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{bmatrix} \begin{bmatrix} \overline{a_{11}} & & & \\ \overline{a_{12}} & \overline{a_{22}} & & \\ \vdots & & \ddots & \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{bmatrix}, \quad \begin{bmatrix} \overline{a_{11}} & & & \\ \overline{a_{12}} & \overline{a_{22}} & & \\ \vdots & & \ddots & \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{bmatrix}.$$

これらの (1,1)-成分を比較すれば, $a_{11}\overline{a_{11}} + a_{12}\overline{a_{12}} + \cdots + a_{1n}\overline{a_{1n}} = \overline{a_{11}}a_{11}$. すなわち,

$$a_{12}\overline{a_{12}} + \cdots + a_{1n}\overline{a_{1n}} = 0, \quad \text{つまり} \quad \|(a_{12}, \dots, a_{1n})\|^2 = 0$$

であり, $a_{12} = a_{13} = \cdots = a_{1n} = 0$ を得る. また, $B = [a_{ij}]_{\substack{i=2, \dots, n \\ j=2, \dots, n}}$ とおけば B は上三角であり,

$$A = \left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array} \right), \quad AA^* = \left(\begin{array}{c|ccc} a_{11}\overline{a_{11}} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & BB^* & \\ 0 & & & \end{array} \right), \quad A^*A = \left(\begin{array}{c|ccc} \overline{a_{11}}a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B^*B & \\ 0 & & & \end{array} \right).$$

上式と A の正規性より, B も正規であることが分かる. 帰納法の仮定により B は対角行列であり, したがって A もそうである. \square

⁵直交行列は, 実正方形行列に対してのみ定義される概念である.

以上により, 系 39.6.6 をさらに一般化した主張として, 次が得られる:

系 39.7.6. 複素正方行列 A について次は同値である.

- (1) A は正規行列である.
- (2) A は, あるユニタリ行列 P を用いて対角化可能である.

Proof. (1) \Rightarrow (2): 定理 39.6.3 より, A はユニタリ行列 P を用いて上三角化できる. このとき $P^{-1}AP$ は上三角かつ正規ゆえ対角行列である (補題 39.7.5).

(2) \Rightarrow (1): 定義 39.7.2 の直前で既に述べたことであるが, 改めて計算により確認しておこう. P をユニタリ行列, $D = P^{-1}AP$ を対角行列とすれば, $A = PDP^{-1} = PDP^*$ であり, このとき

$$\begin{aligned} AA^* &= (PDP^*)(PD^*P^*) = PDD^*P^*, \\ A^*A &= (PD^*P^*)(PDP^*) = PD^*DP^*. \end{aligned}$$

D は正規ゆえ (例 39.7.3(4)), これらは等しい. □

系 39.7.7. 正規変換の各固有空間は互いに直交する.

Proof. U の正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を一つ与えておく. この基底に関する f の表現行列 A は正規であり, ゆえにユニタリ行列 $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ を用いて対角化できる. このとき, $\mathbf{p}_1, \dots, \mathbf{p}_n$ は $T_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ の固有ベクトルからなる \mathbb{C}^n の正規直交基底である. よって, 命題 39.2.10 で与えた同型 $F: U \rightarrow \mathbb{C}^n$ において, $F^{-1}(\mathbf{p}_1), \dots, F^{-1}(\mathbf{p}_n)$ は f の固有ベクトルからなる U の正規直交基底である (命題 28.6.1). したがって, 次の補題 39.7.8 より主張を得る. □

補題 39.7.8. 線形変換 $f: U \rightarrow U$ において, f の固有ベクトルからなる正規直交基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が存在するならば, f の各固有空間は互いに直交する.

Proof. λ, γ を f の異なる固有値とし, 各 $\mathbf{x} \in W(\lambda, f)$ および各 $\mathbf{y} \in W(\gamma, f)$ が互いに直交することを示そう. $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中からいくつかを選んで $W(\lambda, f)$ および $W(\gamma, f)$ の基底を取り, これらをそれぞれ $\mathbf{v}_1, \dots, \mathbf{v}_k \in W(\lambda, f)$, $\mathbf{w}_1, \dots, \mathbf{w}_\ell \in W(\gamma, f)$ とする⁶. 仮定より, 各 \mathbf{v}_i ($i = 1, \dots, k$) は $\mathbf{w}_1, \dots, \mathbf{w}_\ell$ のいずれとも直交するゆえ, $W(\gamma, f)$ の各元とも直交し (補題 35.5.3 または 39.2.8), とくに \mathbf{y} とも直交する. 一方, いま示したように, \mathbf{y} は $\mathbf{v}_1, \dots, \mathbf{v}_k$ のいずれとも直交するゆえ, $W(\lambda, f)$ の各元とも直交し, とくに \mathbf{x} とも直交する. □

39.8 フーリエ級数展開 (発展)

関数を多項式の無限和に分解する式を冪級数展開というのであった. 一方で, 指数関数 (ないし正弦・余弦関数) に関数を分解する式として, フーリエ級数と呼ばれる展開式がある. 関数空間における正規直交基底に相当するものが何かさえ把握しておけば, フーリエ展開公式が自然と導けることを見てみよう. 本節ではオイラーの公式 $e^{i\theta} = \cos \theta + i \sin \theta$ を既知とする. とくに各 $n \in \mathbb{Z}$ について $e^{2n\pi i} = 1$ および $e^{(2n+1)\pi i} = -1$, また $e^{-i\theta} = \overline{e^{i\theta}}$ である.

定義 39.8.1. 閉区間 $[-\pi, \pi]$ を定義域とする複素数値連続関数全体の集合を $C^*([-\pi, \pi])$ と書く. $C^*([-\pi, \pi])$ には次のエルミート内積が定められる:

$$\langle f, g \rangle := \int_{-\pi}^{\pi} \overline{f(t)}g(t) dt. \quad (39.8.1)$$

⁶ $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中から $W(\lambda, f)$ に含まれるものをすべて選び, これを $\mathbf{v}_1, \dots, \mathbf{v}_k$ とすれば, それは $W(\lambda, f)$ の基底になる. 実際, 仮にこれらの線形結合で書けない $\mathbf{z} \in W(\lambda, f)$ があるとすれば, $\mathbf{z}, \mathbf{v}_1, \dots, \mathbf{v}_k$ は線形独立であり, 命題 29.1.1 より $\mathbf{z}, \mathbf{u}_1, \dots, \mathbf{u}_n$ も線形独立である. しかしこれは $\mathbf{u}_1, \dots, \mathbf{u}_n$ が U の基底であることに反する.

$C^*([-\pi, \pi])$ において互いに直交する長さ 1 のベクトルの例を挙げよう. 各 $n \in \mathbb{Z}$ に対して

$$v_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$$

とおく.

命題 39.8.2. $v_n(x)$ たちは長さ 1 で互いに直交する.

Proof. 各 $n \in \mathbb{Z}$ について,

$$\langle v_n, v_n \rangle = \int_{-\pi}^{\pi} \overline{\frac{1}{\sqrt{2\pi}} e^{int}} \frac{1}{\sqrt{2\pi}} e^{int} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-int} e^{int} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} dt = 1.$$

ゆえに $\|v_n(x)\| = \sqrt{\langle v_n, v_n \rangle} = 1$ である. また, $m, n \in \mathbb{Z}$ かつ $m \neq n$ のとき,

$$\begin{aligned} \langle v_n, v_m \rangle &= \int_{-\pi}^{\pi} \overline{\frac{1}{\sqrt{2\pi}} e^{int}} \frac{1}{\sqrt{2\pi}} e^{imt} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-int} e^{imt} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)t} dt \\ &= \frac{1}{2\pi} \left[\frac{1}{i(n-m)} e^{i(n-m)t} \right]_{-\pi}^{\pi} = \frac{1}{2\pi} \frac{1}{i(n-m)} \left(e^{i(n-m)\pi} - e^{i(n-m)(-\pi)} \right). \end{aligned}$$

$$e^{i(n-m)\pi} - e^{i(n-m)(-\pi)} = \begin{cases} 1 - 1 & (m-n \text{ が偶数}), \\ -1 - (-1) & (m-n \text{ が奇数}). \end{cases}$$

ゆえに $\langle v_n, v_m \rangle = 0$. □

定義 39.8.3. $T > 0$ を定数とする. $f: \mathbb{R} \rightarrow \mathbb{C}$ が次の性質を満たすとき, 周期 T の関数という:

$$\text{各 } x \in \mathbb{R} \text{ について } f(x+T) = f(x).$$

例 39.8.4. 各 $n \in \mathbb{Z}$ について, $v_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$ を \mathbb{R} 上の関数とみなすとき, これは周期 2π の連続関数である.

周期 2π の関数は, 区間 $[-\pi, \pi]$ 上のグラフの形が分かれば, この区間の外のグラフも復元できる. したがって, 周期 2π の連続関数 f の定義域を制限して $C^*([-\pi, \pi])$ の元とみなしても f の情報は失われない. 実は, 周期 2π の連続関数 $f: [-\pi, \pi] \rightarrow \mathbb{C}$ は, ある解釈のもとで各 v_n たちに分解されることが知られている. すなわち, 命題 35.3.5 の類推として, 次が成り立つ:

$$f = \sum_{n=-\infty}^{\infty} \langle v_n, f \rangle v_n. \quad (39.8.2)$$

ここで関数 $\langle v_n, f \rangle v_n$ に $x \in [-\pi, \pi]$ を代入した値を書き下せば,

$$\langle v_n, f \rangle v_n(x) = \int_{-\pi}^{\pi} \overline{\frac{1}{\sqrt{2\pi}} e^{int}} f(t) dt \cdot \frac{1}{\sqrt{2\pi}} e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt \cdot e^{inx}.$$

そこで, $z_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt$ とおけば,

$$f(x) \sim \sum_{n=-\infty}^{\infty} z_n e^{inx}$$

を得る. 上の右辺を $f(x)$ のフーリエ級数展開という. ここで, 上で等号とは異なる記号 \sim を用いており, これには技術的な理由がある. 簡単に言うと, 一部の x について右辺が収束しなかったり, あるいは収束しても右辺が連続関数とは限らない (とくに一部の x について $f(x)$ とは値が異なる場合がある) ことに

起因する。詳しくは関数空間における等号 (39.8.2) の意味を学ぶ必要があり、本書の範囲で説明するのは難しい⁷。実際には、ほとんどいたるところの $x \in [-\pi, \pi]$ において、上式の \sim は等号に変えてもよい。

ここまでの議論をまとめると、 $B = \{v_n \mid n \in \mathbb{Z}\}$ がある解釈のもとで関数空間の正規直交基底になっているという立場で見れば、フーリエ級数展開公式は直ちに得られる。したがって、この公式は特に改めて覚える必要のない代物なのである。

なお、読者の中には三角関数による展開のほうが馴染みがあるかもしれない。これは、指数関数 e^{inx} をオイラーの公式を通して三角関数に分解するだけで自然と得られる。これを確認してみよう。

各 $n \in \mathbb{N} \cup \{0\}$ について、 z_n と z_{-n} は共役の関係にあるから、

$$a_n := z_n + z_{-n}, \quad b_n := (z_n - z_{-n})i$$

は共に実数である。このとき、各 $n \in \mathbb{N}$ について

$$\begin{aligned} z_n e^{inx} + z_{-n} e^{-inx} &= z_n (\cos nx + i \sin nx) + z_{-n} (\cos(-nx) + i \sin(-nx)) \\ &= z_n (\cos nx + i \sin nx) + z_{-n} (\cos nx - i \sin nx) \\ &= (z_n + z_{-n}) \cos nx + (z_n - z_{-n})i \sin nx = a_n \cos nx + b_n \sin nx. \end{aligned}$$

また、 $a_0 = z_0 + z_0 = 2z_0$ ゆえ、 $z_0 e^{0x} = z_0 = a_0/2$ である。以上より、

$$f(x) \sim \sum_{n=-\infty}^{\infty} z_n e^{inx} \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx).$$

上の右辺も $f(x)$ のフーリエ級数展開 (あるいは三角級数展開) という。さらに、 a_n と b_n の値は次のように書き下せる:

$$\begin{aligned} a_n &= z_n + z_{-n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (e^{-inx} + e^{inx}) f(t) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos(-nx) + i \sin(-nx) + \cos nx + i \sin nx) f(t) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} 2 \cos(nx) f(t) dt = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(nx) f(t) dt. \\ b_n &= (z_n - z_{-n})i = \frac{i}{2\pi} \int_{-\pi}^{\pi} (e^{-inx} - e^{inx}) f(t) dt \\ &= \frac{i}{2\pi} \int_{-\pi}^{\pi} (\cos(-nx) + i \sin(-nx) - \cos nx - i \sin nx) f(t) dt \\ &= \frac{i}{2\pi} \int_{-\pi}^{\pi} (-2i) \sin(nx) f(t) dt = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(nx) f(t) dt. \end{aligned}$$

なお、 $v_n(x)$ や $\cos nx$, $\sin nx$ は周期 2π の関数であるから、これらを用いて周期 2π の関数 $f(x)$ を展開した式は、定義域を $[-\pi, \pi]$ に制限せずとも成り立つ。以上を公式としてまとめておこう:

定理 39.8.5 (フーリエ級数展開). 周期 2π の連続関数 $f: \mathbb{R} \rightarrow \mathbb{C}$ について、

$$f(x) \sim \sum_{n=-\infty}^{\infty} z_n e^{inx} \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

$$\text{ここで、} \quad z_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(nx) f(t) dt, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(nx) f(t) dt.$$

⁷式 (39.8.2) における等号の意味は「 $f = g \Leftrightarrow \|f - g\| = 0 \Leftrightarrow \int_{-\pi}^{\pi} (f(x) - g(x))(f(x) - g(x)) = 0$ 」, つまり積分によって判断されるものである。 f, g に積分可能だが連続でない関数を許すと、関数として $f(x) \neq g(x)$ である (つまり $f(x) \neq g(x)$ を満たす $x \in [-\pi, \pi]$ が存在する) もの、 $\|f - g\| = 0$ となる場合が起こり得る。

上の公式は、式 (39.8.2) さえ認めれば、あとは形式的な計算を書き下すだけで得られるのであった。式 (39.8.2) は内積空間上の極限に関する命題であるから、その証明は本書の枠を大きく超えるものである。関数解析学においてヒルベルト空間の一般論、およびストーン-ワイエルシュトラスの近似定理を学べば、下の証明を理解できるであろう。

備考 39.8.6 (発展). (1) 内積空間がノルムについて完備であるとき⁸、これをヒルベルト空間 (**Hilbert space**) という。 \mathbb{R}^n や \mathbb{C}^n は有限次元ヒルベルト空間である。例えば微分積分学では、各 $z \in \mathbb{C}$ に対して $e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}$ の右辺が収束することを示す際に \mathbb{C} の完備性 (すなわち \mathbb{R}^2 の完備性) を用いる。つまり完備性は、オイラーの公式を成立させるために必要な概念である。

(2) ノルム空間 U の部分集合 D において、 U の各元を D の元からなる列の極限として表すことができるとき、 D を U の稠密部分集合という。また、可算な稠密部分集合をもつ空間を可分であるという。例えば \mathbb{Q} は \mathbb{R} の稠密部分集合であり、したがって \mathbb{R} は可分である。

(3) 無限次元ヒルベルト空間 H において、正規直交基底を極限和に拡張した概念が上手く機能することが知られている。例えば H が可分であるとき、任意の $f \in H$ について式 (39.8.2) が成立するように、互いに直交する長さ 1 のベクトルたちの集合 $B = \{v_n \mid n \in \mathbb{Z}\} \subset H$ を上手く取れることが知られている。この B を H の正規直交基底という。

定理 39.8.7 (ストーン-ワイエルシュトラス). コンパクトなハウスドルフ空間 X 上の複素数値連続関数全体 $C^*(X)$ の単位的部分多元環 R が共役について閉じており、かつ X 上の任意の二点を分離する⁹ならば、一様収束位相において R は $C^*(X)$ の稠密部分集合である。

式 (39.8.2) の証明 \mathbb{R} を定義域とする周期 2π の複素数値連続関数全体の集合を $C^*(S^1)$ とおく¹⁰。 e^{ix} と $\overline{e^{ix}} = e^{-ix}$ 、および $\mathbf{1}$ (定数関数) の和とスカラー倍および積を有限回繰り返して得られる $C^*(S^1)$ の部分環を R とおく。このとき、 R は共役について閉じており、また $B = \{e^{inx}/\sqrt{2\pi} \mid n \in \mathbb{Z}\}$ の線形結合で書ける関数全体に等しい。 $(-\pi, \pi]$ 上の関数 e^{ix} を S^1 上の関数とみなせば、これは恒等写像であり、ゆえに単射である。したがって定理 39.8.7 により、一様収束位相において R は $C^*(S^1)$ の稠密部分集合となる。すなわち、 $C^*(S^1)$ の各元 f に対して、 f に一様収束する関数列 $f_n \in R$ が存在する。内積から定まるノルム (L_2 -ノルム) における収束 (L_2 -収束) は一様収束よりも弱いことから、この f_n は f に L_2 -収束する。すなわち、 B を正規直交基底とするヒルベルト空間 (L_2 -ノルムに関する $R = \langle B \rangle$ の完備化) は $C^*(S^1)$ を含む。ゆえに $C^*(S^1)$ の各元に対して、式 (39.8.2) による分解が成り立つ。 \square

⁸任意のコーシー列が収束列となると、そのようなノルム空間を完備であるという。

⁹すなわち次が成り立つことである: 各 $x, y \in X$ (ただし $x \neq y$) に応じて、 $f(x) \neq f(y)$ を満たす $f \in R$ が存在する。

¹⁰ $(-\pi, \pi]$ と単位円 $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$ は $T(\theta) = e^{i\theta}$ によって 1 対 1 に対応する。この対応を通して、 \mathbb{R} 上の周期 2π の連続関数と S^1 上の連続関数を同一視することができる。ただし、上の証明において関数に元を代入する際は $(-\pi, \pi]$ の元として代入している点に注意せよ。

関連図書

- [1] 齋藤正彦, 『線形代数入門』, 東京大学出版会.
- [2] 三宅敏恒, 『線形代数学 初歩からジョルダン標準形へ』, 培風館.
- [3] 笠原皓司, 『線形代数と固有値問題 スペクトル分解を中心に』, 現代数学社.
- [4] 川久保勝夫, 『線形代数学』, 日本評論社.
- [5] 木村達雄・竹内光弘・宮本雅彦・森田純, 『明解 線形代数』, 日本評論社
- [6] 嶺幸太郎, 『微分積分学の試練 実数の連続性と ε - δ 』, 日本評論社.

索引

- $*$, 24
- $:=$, 9
- $[\mathbf{u}]$, 174
- $\mathbf{0}$, 104
- $\mathbf{1}$, 119
- \mathbb{C} , 9
- \cap , 100
- \cup , 100
- deg, 105
- δ_{ij} , 21
- det, 70
- dim, 160
- \emptyset , 96
- $\text{End}(U)$, 155
- \equiv , 134
- $\text{Hom}(U, V)$, 155
- id_X , 58, 135
- $\Im z$, 306
- Im, 145
- \in , 9
- inv, 67
- Ker, 146
- $\langle \cdot, \cdot \rangle$, 265, 273, 303, 304
- $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$, 123
- $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}}$, 273
- $\langle A \rangle$, 129
- $\|\cdot\|$, 266, 273, 304
- $|A|$, 70
- \mathfrak{A}_n, A_n , 61
- \mathfrak{S}_n, S_n , 61
- max, 44
- min, 44
- \mathbb{N} , 8
- \notin , 9
- null, 160
- \oplus , 256
- \bar{A} , 280
- \bar{z} , 229
- $\Phi_A(t)$, 213
- \prod , 21
- \mathbb{Q} , 8
- \mathbb{R} , 9
- $\mathbb{R}[x]$, 106
- $\mathbb{R}^{\mathbb{N}}$, 105
- \mathbb{R}^n , 9
- rank, 44, 160
- $\Re z$, 306
- \setminus , 100
- sgn, 60
- \simeq , 153
- \subset , 96
- \sum , 21
- $\text{tr } A$, 148
- $\mathbf{u} + A$, 170
- \tilde{A} , 91
- $\widetilde{W}(\lambda, A)$, 250
- $\widetilde{W}(\lambda, f)$, 250
- ${}^t A$, 24
- ${}^t f$, 202
- A^* , 306
- a_{ij}^* , 91
- A^{-1} , 37
- A_k , 289
- $C(X)$, 108
- $C^\infty(I)$, 115
- $f(A)$, 133
- $f|_A$, 136
- f^* , 202, 307
- f^{-1} , 136
- $f^{-1}(B)$, 133
- $f^{-1}(b)$, 133
- K^n , 109
- $M_{m,n}(\mathbb{R})$, 105
- $M_{m,n}(K)$, 109
- $M_n(\mathbb{R})$, 105
- $M_n(K)$, 109
- q_A , 284
- U/W , 174
- $W(\lambda, A)$, 214
- $W(\lambda, f)$, 214
- W^\perp , 271, 305
- W_A , 113

X/\simeq , 177
 QR 分解, 276
 安定部分空間, 254
 鞍点, 287
 1 次

- 関係, 117
- 結合, 117
- 従属, 118
- 独立, 118

 1 対 1 の対応, 135
 一般

- 固有空間, 250
- 固有ベクトル, 250
- 線形群, 63

 イデアル, 116
 因子, 227
 因数定理, 228
 ヴァンデルモンドの行列式, 78
 上三角

- 化, 277
- 行列, 24

 上への写像, 134
 n 次方程式, 228
 エルミート

- 行列, 306
- 空間, 304
- 内積, 304
- 変換, 311
- 標準的な—内積, 303
- 歪—行列, 312

 オイラーの公式, 209
 外延的定義, 97
 階数, 44

- 線形写像の—, 160

 カヴァリエリの原理, 71
 ガウス・ジョルダンの消去法, 35
 可換

- 環, 20
- 図式, 193
- 性, 151

 可逆行列, 37
 核, 146
 拡大係数行列, 31
 拡張, 136
 重ね合わせの原理, 47
 可算, 140

- 非—, 140

 型, 15
 合併集合, 100
 かなめ, 32
 可分, 316
 加法定理, 13
 環, 20

- 可換—, 20
- 多元—, 20
- 多項式—, 106

 関係, 154
 関数, 10

- 周期—, 314
- 線形汎—, 201
- デルタ—, 206
- 比例—, 10

 完全系列, 182
 完備, 316
 簡約

- 化, 32
- 行列, 31

 奇置換, 60
 基底, 124

- 正規直交—, 268, 304, 316
- 双対—, 202
- ハメル—, 129
- 標準—, 124

 基本

- 行列, 38
- ベクトル, 117
- 行—変形, 31

 逆

- 行列, 37
- 元, 63
- 写像, 136
- 像, 133
- 置換, 58
- 左—行列, 167
- 右—行列, 167

 行, 15

- 階段形, 32
- 簡約形, 32
- 基本変形, 31
- のかなめ, 32

- の主成分, 31
- ベクトル, 9
- 鏡映, 299, 300
- 共通部分, 100
- 共役
 - 行列の—転置, 306
 - 線形変換の—, 307
 - 複素数の—, 229
- 行列, 15
 - 上三角—, 24
 - エルミート—, 306
 - 可逆—, 37
 - 拡大係数—, 31
 - 簡約—, 31
 - 基本—, 38
 - 逆—, 37
 - が等しい, 16
 - の階数, 44
 - の型, 15
 - の簡約化, 32
 - の共役転置, 306
 - のサイズ, 15
 - の随伴, 306
 - のスカラー倍, 16
 - の成分, 15
 - の積, 17
 - のランク, 44
 - の和, 16
 - の割り算, 38
 - 係数—, 29
 - 恒等—, 18
 - 三角—, 25
 - 自己随伴—, 306
 - 下三角—, 25
 - 実—, 109
 - 主小—, 289
 - 正規—, 311
 - 正則—, 37
 - 正方—, 18
 - 対称—, 279
 - 単位—, 18
 - 直交—, 275
 - 転置—, 24
 - 左逆—, 167
 - 表現—, 191
 - 複素—, 109
 - 冪等—, 302
 - ヘッセ—, 54, 292
- 変換—, 190
- 右逆—, 167
- ヤコビ—, 55
- ユニタリ—, 306
- 余因子—, 91
- 零—, 18
- 歪エルミート—, 312
- 歪対称—, 312
- 行列式, 70
 - ヴァンデルモンドの—, 78
 - ヘッセ—, 54
 - ヤコビ—, 55
 - ロンスキー—, 247
- 距離, 295
 - ユークリッド—, 275
- 空間
 - 安定部分—, 254
 - 一般固有—, 250
 - エルミート—, 304
 - 計量ベクトル—, 273
 - 広義固有—, 250
 - 固有—, 214
 - 商—, 174
 - 線形—, 104
 - 双対—, 200
 - 直交補—, 271, 305
 - 内積—, 273, 304
 - ノルム—, 295
 - ヒルベルト—, 316
 - 不変部分—, 248
 - 冪零部分—, 250
 - ベクトル—, 104
 - ユークリッド—, 9
 - ユニタリ—, 304
- 空集合, 96
- 偶置換, 60
- グラム-シュミットの直交化法, 269
- クラメルの公式, 54, 93
- クロネッカーのデルタ, 21
- 群, 63
 - 一般線形—, 63
 - 交代—, 61
 - 対称—, 61
- 係数, 117
 - 行列, 29
- 計量ベクトル空間, 273
- ケーリー-ハミルトンの定理, 158, 231

結合律, 19
 元, 8
 逆—, 63
 単位—, 63
 広義固有空間, 250
 合成, 10
 交代
 —群, 61
 —性, 75, 84
 恒等
 —行列, 18
 —写像, 135
 —置換, 58
 —的に等しい, 134
 コーシー-シュワルツの不等式, 294
 互換, 58
 固有
 一般—空間, 250
 一般—ベクトル, 250
 広義—空間, 250
 広義—ベクトル, 250
 —空間, 214
 —多項式, 213
 —値, 210
 —ベクトル, 210
 固有値, 210
 —はすべて実数, 277
 重複を含めたすべての—, 225
 最小
 —値, 44
 —の部分空間, 131
 サイズ, 15
 最大値, 44
 差集合, 100
 作用素, 142
 自己共役—, 311
 自己随伴—, 311
 シフト—, 148
 線形—, 142
 微分—, 149
 サラスの方法, 71
 三角
 上—化, 277
 上—行列, 24
 —級数展開, 315
 —行列, 25
 —不等式, 295, 306
 下—行列, 25
 次元, 160
 —公式, 172
 無限—, 160
 有限—, 160
 自己
 —共役作用素, 311
 —準同型, 155
 —随伴行列, 306
 —随伴作用素, 311
 次数, 105, 284
 退化—, 161
 下三角行列, 25
 実
 —行列, 109
 —ベクトル, 109
 シフト作用素, 148
 自明
 —でない組, 118
 —でない線形関係, 118
 —な線形写像, 142
 —な部分空間, 111
 射影, 148
 —変換, 301
 正—, 297
 直交—, 297
 部分空間への—, 300
 写像, 10
 逆—, 136
 恒等—, 135
 —が等しい, 134
 —の拡張, 136
 —の合成, 10
 —の制限, 136
 商—, 174
 線形—, 11
 シューア分解, 277
 終域, 133
 周期関数, 314
 集合, 8
 合併—, 100
 空—, 96
 差—, 100
 —が等しい, 97
 —の共通部分, 100
 部分—, 96
 和—, 100

主小行列, 289
 主成分, 31
 シュミットの直交化法, 269
 巡回置換, 58
 準同型, 151, 155, 157
 自己—, 155
 多元環—, 157
 商
 —空間, 174
 —写像, 174
 —集合, 177
 ジョルダン
 ガウス・—の消去法, 35
 —標準形, 259
 推移律, 153
 随伴
 行列の—, 306
 自己—行列, 306
 線形変換の—, 307
 スカラー, 16
 行列の—倍, 16
 —律, 19
 ベクトルの—倍, 9
 $\text{Hom}(U, V)$ の元における—倍, 155
 正規
 —行列, 311
 —直交基底, 268, 304, 316
 —変換, 311
 制限, 136
 斉次形, 46
 正射影, 297
 生成
 —される, 123
 —する, 124, 129
 正則行列, 37
 正定値, 285
 半—, 285
 正の定符号, 285
 成分, 15
 主—, 31
 対角—, 18
 正方行列, 18
 積
 $\text{End}(U)$ の元における—, 156
 行列の—, 17
 置換の—, 57
 内—, 25
 跡, 148
 線形
 一般—群, 63
 自明な—写像, 142
 —回帰数列, 115
 —関係, 117
 —空間, 104
 —結合, 117
 —作用素, 142
 —写像, 11, 142
 —従属, 118
 —常微分方程式, 115
 —性, 11, 142
 —漸化式, 115
 —同型, 152, 153
 —独立, 118, 129
 —独立な最大個数, 162
 —汎関数, 201
 —部分空間, 111
 —変換, 155
 体 K 上の—空間, 109
 多重—性, 72, 83
 全射, 81, 134
 全単射, 135
 像, 133
 逆—, 133
 相似, 197
 双対
 基底の—, 201
 線形写像の—, 202
 —基底, 202
 —空間, 200
 体, 20, 108
 — K 上の線形空間, 109
 第1同型定理, 181
 対角
 —化, 212
 —化可能, 212
 —形式, 287
 —成分, 18
 —和, 148
 退化次数, 161
 対称
 —行列, 279
 —群, 61
 —性, 62
 —律, 153

反一性, 74
 歪一行列, 312
 歪一性, 74
 代数, 20
 —学の基本定理, 228
 多元環, 20, 156
 —準同型, 157
 多項式, 105
 固有—, 213
 —環, 106
 —の次数, 105
 特性—, 158, 213, 238, 241, 243, 247
 多重線形性, 72, 83
 たすきがけ, 71
 単位
 —行列, 18
 —元, 18, 20, 63
 短完全列, 182
 単射, 81, 134
 置換, 56
 奇—, 60
 逆—, 58
 偶—, 60
 恒等—, 58
 巡回—, 58
 —の積, 57
 —の符号, 60
 中線定理, 296
 稠密, 316
 重複を込めたすべての固有値, 225
 超平面, 72
 直和分解, 256
 直交
 グラム-シュミットの—化法, 269
 正規—基底, 268, 304, 316
 —行列, 275
 —射影, 297
 —する, 268, 304
 —変換, 297
 —補空間, 271, 305
 定義域, 10
 デルタ
 クロネッカーの—, 21
 —関数, 206
 転置
 行列の—共役, 306
 線形写像の—, 202
 —行列, 24
 転倒数, 67
 同型, 152, 153, 157
 第1—定理, 181
 同次形, 46
 同値
 —関係, 154
 —類, 176
 等長変換, 297
 特性
 —多項式, 158, 213, 238, 241, 243, 247
 —方程式, 213, 238, 241, 243, 247
 閉じている, 111
 トレース, 148
 内積, 25, 273
 エルミート—, 304
 —から定まるノルム, 273, 304
 —空間, 273, 304
 標準的なエルミート—, 303
 標準的な—, 265
 内包的定義, 97
 長さ, 266
 2次形式, 284
 任意定数, 34
 —の個数, 36, 162
 ノルム, 295
 内積から定まる—, 273, 304
 —空間, 295
 標準的な—, 266
 ユークリッド・—, 266
 はきだし法, 35
 ハメル基底, 129
 パリティ, 60
 反射律, 153
 半正定値, 285
 半負定値, 285
 非可算, 140
 ピタゴラスの定理, 294
 等しい
 行列が—, 16
 恒等的に—, 134
 集合が—, 97
 微分作用素, 149
 表現行列, 191

標準
 —基底, 124
 —的なエルミート内積, 303
 —的な内積, 265
 —的なノルム, 266
 —ベクトル, 117
 ヒルベルト空間, 316
 比例関数, 10

 フィボナッチ数列, 114
 フーリエ級数展開, 315
 複素
 —行列, 109
 —ベクトル, 109
 符号, 60
 正の定—, 285
 不定—, 285
 負の定—, 285
 不定
 —値, 285
 —符号, 285
 負定値, 285
 半—, 285
 負の定符号, 285
 部分空間, 111
 安定—, 254
 最小の—, 131
 自明な—, 111
 線形—, 111
 不変—, 248
 冪零—, 250
 部分集合, 96
 不変
 —部分空間, 248
 —量, 161
 分解
 QR—, 276
 シューア—, 277
 直和—, 256
 分配律, 19

 平行四辺形の法則, 296
 冪等行列, 302
 冪零, 210
 —部分空間, 250
 ベクトル, 9
 一般固有—, 250
 n 次—, 15
 基本—, 117
 行—, 9
 計量—空間, 273
 広義固有—, 250
 固有—, 210
 実—, 109
 標準—, 117
 複素—, 109
 —空間, 104
 —空間の公理, 104
 —のスカラー倍, 9
 —の長さ, 266
 —の和, 9
 列—, 9
 ヘッセ行列, 54, 292
 変換
 エルミート—, 311
 射影—, 301
 正規—, 311
 直交—, 297
 等長—, 297
 —行列, 190
 ユニタリ—, 310

 方程式
 n 次—, 228
 線形常微分—, 115
 特性—, 213, 238, 241, 243, 247
 連立1次—, 29

 右逆行列, 167

 無限次元, 160

 ヤコビ行列, 55

 ユークリッド
 —距離, 275
 —空間, 9
 —ノルム, 266
 有限次元, 160
 ユニタリ
 —行列, 306
 —空間, 304
 —変換, 310

 余因子, 89
 —行列, 91
 —展開, 89
 要素, 8
 余弦定理, 266

ランク

- 行列の—, 44
- 線形写像の—, 160

零

- 行列, 18
- 元, 20, 104
- ベクトル, 104

列, 15

- 基本変形, 42
- ベクトル, 9

連立1次方程式, 29

ロンスキー行列式, 247

和

- 行列の—, 16
- ベクトルの—, 9
- $\text{Hom}(U, V)$ の元における—, 155
- 集合, 100

歪

- エルミート行列, 312
- 対称行列, 312
- 対称性, 74, 83