線形代数学 講義ノート

まえがき

これは大学１年次を対象にした線形代数学の講義ノートである．前半部分では連立１次方程式の解法と行列式の計算を主に扱う．後半は線形空間の抽象論の初歩を踏まえた上で, 行列の対角化までを目標に定めている．

本論では, 簡単な計算演習はある程度こなせるものの, 線形代数学で扱う数学的諸概念の意図が分からずに悩み苦しんでいる人を主な読者として想定し, 次の点を重んじた:

- 「何故学ぶのか」「どうして新しい概念を導入するのか」といったそもそも論をできる限り展開する．
- 天下り的な定義の導入はなるべく控える．
- 線形代数学の枠組みの外にある数学にも言及する．

本論の構成について説明しておこう．本論は約３０の節で構成され, 各節は講義１回分の内容に相当している．各節の冒頭には, 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．各項の題目や命題に「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．

担当: 嶺 幸太郎

目 次

1　線形代数学とは何か

始めに, 線形代数学とはどんな理論であり, 学問全体の中でどういう位置を占めているかについて簡単に述べておきたい. そのためには次の二つの視点から眺めるのがよいように思う. 一つは大学教育における線形代数学の役割について, もう一つは数学の諸分野の中での線形代数学の占める位置についてである. 前者においては線形代数学に限定せずに, 大学における数学教育の役割について考えよう. また後者を論じる前提として, 数学の分野がどのように大別されているかを概説する. すなわち, 「線形代数学」という言葉はいったん忘れて, まずはこれらの大枠において俯瞰する.

1.1　大学で学ぶということ

大学で学ぶとはどういう事か考えてみれば, そもそも大学とは何かという「大学論」を避けては通れない. すなわち, 大学の歴史やその存在意義について問わねばならなくなる. そこで, 西洋中世における大学がどんな機関であったのか少しだけ振り返っておこう.

大学とは何を学ぶ場所であったのだろうか. それは高度知識人になるための教養を学ぶ場であった. ここでいう「教養」とは何かという問いはこれまた難しいのだが, 簡単にまとめれば, 物事の在り方を正確に把握し適切な判断を下すための分析力と行動力のことを指す. したがって教養とは広い知識だけをいうのではなく, それを活用する能力のことも含んでいる. そして, 大学で教養を修めた者は, 次に専門職教育機関 (大学院[1]) で専門課程を学び, それぞれの職業に就くのである. 教養が必要とされる専門職は当初は神学・法学・医学に関するものが主であった. のちに社会が専門化・多様化したことにより, 様々な職業において高度な知識人が求められるようになり, 教養を必要とする者の多くが大学に足を運んだのである. 大学とは, どのような職に就くにせよ必要となる基礎学力を身につける場であると見なされていた.

なお, 中世の教養教育はリベラル・アーツあるいは自由七科とも呼ばれる. ここで自由という言葉には, 偏見なく自らの判断で物事を見定めることができるという意味が込められている. また, 七科とは具体的な学科のことであり, 言語に関わる三学（文法・論理・修辞）と数理科学に関わる四科（算術・幾何・天文・音楽）を指す. これらの教養を身につけていることが当時の知識人のあいだの常識であった.

1.2　数学教育の役割

自由七科のうち半分が数学的な学科であった事情を検討しよう[2]. 教養教育の目的は先に述べた通りであり, その中でも特に基盤となる力が論理的思考力である. そして, この能力を獲得する一つの指針を与えるものとして数学が尊重されていた. その背景には, あまたある学問の中でも特に厳密性が高い理論であると数学がみなされ, 論理展開のモデルとなったという事情がある. もちろん, 数学以外の学科を学ぶことでも論理的思考力を高めることはできる. しかしながら数学では, 余分な情報を落として単純化した問題を扱う傾向が強いことから, 論理によって決着がつく様子を初学者が把握しやすいという側面がある. ここに教育的効果における数学の優位性があり, 論証の方法論を学ぼうと思えば数学はうってつけの学問といえる. この事情は現代においても同じであり, 教養としての数学に確かな価値が認められている.

一方で, 科学が発達した近代以降においては, その技術的側面からも数学を学ぶ必要が生じている. 科学の多くは数学を用いて表現されるため, サイエンスの言語としてまず学ばねばならない学科となったのである. もちろん, 将来専攻する分野によっては, 数式を活用する機会は少ないという人もいるだろう. それでも実験やアンケート調査といった観測を行う分野を専攻するのであれば, 観測結果にどの程度の信憑性があるかを知るために少なくとも統計学を学ぶ必要がある. 自然科学系のどの専門学科に進むにせよ学ぶべき基礎学科であるという数学の立場は, 中世における教養教育の役割を彷彿させるものがある.

[1]大学院は professional school および graduate school の和訳に相当し, ここでは professional school を指す.
[2]あるいは修辞学を除く六科目すべてが数学的とも考えられる.

1.3 この講義が目指すところ

かつての大学では, 自由七科に含まれる部分についての数学が学ばれていた. ここで視点を現代に戻し, これから大学で学ぶことになる数学に話を向けよう. 現代においても大学生に教えるべき数学内容の共通見解が得られており, したがって, どこの大学に行っても誰が教えても殆ど同じ内容が扱われることになる. そして, 多くの講義は学習の効率性の名のもとでパッケージ化されている. 特に理工系学科の 1 年次生向けの講義についてそれは顕著であり, ほとんどの大学において, 微積分学 (解析学の初歩) と線形代数学 (行列と行列式の理論) を二つの柱とした初年度カリキュラムが組まれている.

個人的見解ではあるが, こうしたパッケージ化には次のような功罪があるように思う:

(功) パッケージされた部分に限れば, そこで述べられる内容は完成しきった理論と見なせる. これは, 本に書いてあることを書き写すだけでも講義として成立することを意味し, これにより担当の講師は別の仕事に多くの時間を割くことができるようになった. また, 扱うべき内容が決まっていることから参考書を書く側の工夫としては, いかにして読者の用途に応じるかを重視するようになり, 数学書の様々な差別化が図られることになる. 実際, 講義用のものや自習に適したもの, あるいは分かりやすさに特化したものなど多種多様な要請に応じた参考書が数多く出版されており, 学習する側の環境はこの上なく整っている.

(罪) 一度パッケージ化されると担当講師はノルマの消化のみを念頭に授業を進めがちになり, そのため広い視野を持って論ずる機会は失われやすい. そうなれば内容が縦割りになり, 本来は有機的な繋がりを持つ各分野がそれぞれ独立・分断されたものであるかのように受講者は感じてしまう. また, パッケージ化された教育が長い間続くと, 下手をすれば, それらを学ぶべきことは理由なく当然であるといった錯覚に陥ってしまう可能性がある. これはある種の思考停止であり, 思考停止してしまえば学問の意義が問われることはない. これでは社会の変化に応じた数学教育は行えないし, あるいは学生の学習意欲の低下に繋がることもあるだろう.

本論では上のような事情を踏まえて, 次に重点を置いている（これは「はじめに」にも述べた):

- 「何故学ぶのか」「どうして新しい概念を導入するのか」といったそもそも論をできる限り展開する.
- 天下り的な定義の導入はなるべく控える.
- 線形代数学の枠組みの外にある数学にも言及する.

このため本論は, 全体にわたって非常に冗長な内容となっている. これは, 枝葉末節を切り捨て簡潔な記述を好む数学の美意識を貶めるものである. しかしながら読者が数学に対して広く深い理解を得ることを優先し, あえてこのような記述をしている.

1.4 数学概論

数学における線形代数学の位置を知るために, 数学の各分野について簡単に触れておこう. 伝統的には, 数学は代数・解析・幾何の三分野に大別される. もちろん, これは大まかな分類であり, 数学の各分野はすべてこのいずれかに属するというわけではない. むしろ数学は, この三分野の考え方を駆使して, 分野に囚われることなく縦横無尽に研究されていると考えたほうがよい. これから三分野の簡単な説明を与えるが, ここで述べていることは, 予備知識の無い初学者には想像しにくい部分もあるだろう. しかしながら, 数学をある程度学び終えた後でもう一度読めば, より深い理解が得られるはずである.

代数 代数学とは, 簡単に説明すれば四則演算の技法を高める学問のことである. 例えば小中学校で学んだ計算練習などがこれに相当する. しかしながら代数学の神髄は, <u>二つの異なる世界を一つの見方で繋げること</u> にあると言えるだろう. これについては本節の後半 (1.6 および 1.7 項) で例を挙げて説明する. さて, 代数学の中には数論・群論・体論・環論などがある. 数論は素数の性質を調べる分

野で, 最大公約数や最小公倍数の問題を扱った経験のある一般の人にとって最もなじみある数学である. 群論とは, 図形をはじめとする様々な数学的構造の対称性を研究する分野であり, 体論とは四則演算が成立する世界をいくつも考えだし, それらの間にある関係を群論を通して調べるものである. 環論の説明を予備知識なしに述べるのは難しい. しかし, 線形代数学がこれら四分野のどこに入るかあえて問うてみれば, 環論に属するというのが答えである.

解析 高校で学んだ微分や積分などに現れる極限操作を扱う数学を解析学という. もちろん複素数の範囲も含めた関数の解析も扱われる (関数論). 自然科学の法則 (物理法則だけとは限らない) を記述する式の多くは微分積分の記号を用いて表され, これらは微分 (積分) 方程式と呼ばれる. こうした方程式の解を探す手法や, 解が満たす性質を研究することが解析学の主な目的であり, したがって実用的な諸科学と最も関係の深い分野とも言えるだろう. 解析学において厳密な論理展開を行うには ε-δ 論法の会得はもちろんのこと, 関数の列に関する収束・発散の精緻な議論（関数解析学）が必須であり, そのための基礎としてルベーグ積分というキーワードがあることを覚えておくとよい. また積分論は, 統計学の基礎となる確率論 (大数の法則や中心極限定理) とも深い関わりを持っている.

幾何 空間図形を扱う数学を幾何学といい, 現代の幾何学は微分幾何学と位相幾何学（トポロジー）に大別されている. 前者は面積や体積など量的な調べかたを下地にした幾何学であり, 後者は図形の持っている性質の違いに着目する幾何学である. 図形の性質とは, 例えば円上の 1 点を切り離してもまだ繋がったままだが, 線分で同じことを考えると二つに分離されてしまうといった具合である. いずれの幾何学も多様体の構造を調べることが念頭にある. 大航海時代以前の人々は, 世界の形が平らで海が無限に広がっているのか, それとも球の形をしているか, あるいは第三の可能性はあるか (例えばドーナツ型など）という問題に挑むには, 限られた観測結果と頭の中の想像を頼りに結論づけるしかなかった. 多様体論とはこれの多次元版に相当する. つまり, 宇宙の形の可能性について追求する数学である.

いま挙げた分野は数学の一部であり, ここにすべてが列挙されたわけではない. 例えば複数の分野にまたがる数学 (応用数学を含む) については何一つ述べていない. このほか, 数学それ自体を問う分野もある. その一つは数学の歴史を紐解く数学史に関する分野であり, もう一つは数学における証明の厳密性や正当性について論じる数学基礎論 (数理論理学) である.

最後に, 大学初年度に学ぶ微分積分学と線形代数学について述べておく. これらが理工系分野で求められる基本的な素養の一つといえるかどうか, 読者自身で検討されることを望む.

微分積分 微分とは曲線の傾きのことを指す. これは局所的な変化量を意味しており, 微分法では傾きの分析を通して物事の変化を捉える手法が論じられる. 一方, 積分 (定積分) は図形の面積や体積に相当する概念である. 細かく切り分けた図形の面積を足し合わせることで求積したように, 積分法では情報の集積や総合の仕方が扱われる. 一見すると, 微分と積分は全く関係のない概念のように思えるものの, これらは微分積分学の基本定理を通して関連付けられている. 歴史的視点で振り返れば, ニュートンによる力学への導入によって微分積分学が今日の形で確立され, 様々な物理法則を方程式を通して説明することが可能になった. しかし, 変化を記述したり情報を集積・総合するといった行為は物理学の専売特許ではない. これらは人間が行う分析・考察において基本的であり, したがって, 微積分の手法は多様な分野に応用できる可能性を秘めている.

線形代数 線形写像（線形関数）を分析するための理論を線形代数学という. 線形写像とは, 大雑把にいえば比例関数の多変数化に相当し, ゆえに線形写像は定数関数の次に単純な関数といえる. 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関数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. そして事実, 我々人間が考える数学的対象の多くに線形性が潜んでいるのである.

それでは, 線形写像の詳しい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復習しよう.

1.5 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, この講義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[3]:

- $\mathbb{N}$: 自然数[4](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[5]. .
- $\mathbb{Q}$: 有理数 (rational number) 全体のなす集合のこと.
- $\mathbb{R}$: 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[6].

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 1.5.1. $5 \in \mathbb{N}, \quad -1 \notin \mathbb{N}, \quad \dfrac{12}{13} \in \mathbb{Q}, \quad \sqrt{2} \notin \mathbb{Q}, \quad$ ぶどう $\notin \{$ りんご, みかん, 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, "並び順を込めた意味" というのは, (x, y) と (y, x) は違うものと見なすということである[7]. $\mathbb{R}^2$ は平面上の点全体に一致している. 同様にして, 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 $(x_1, x_2, \cdots, x_n)$ たち全体からなる集合を $\mathbb{R}^n$ と表し, これを**$\boldsymbol{n}$次元ユークリッド空間** (**$\boldsymbol{n}$-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. 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 $(x_1, \cdots, x_n), (y_1, \cdots, y_n) \in \mathbb{R}^n$ および実数 $r \in \mathbb{R}$ に対して, ベクトルの**和** $(x_1, \cdots, x_n) + (y_1, \cdots, y_n)$ およびベクトルの**スカラー倍** (**scalar multiplication**) $r(x_1, \cdots, x_n)$ を次で定める:

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

上で用いた記号 := は, 新たな概念である左辺を右辺によって定義するという意味を表すものである.

ベクトルを一文字で表す場合は, $\boldsymbol{x} = (x_1, \cdots, x_n)$ のように太字で書く. 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. $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ (ただし $n = 2, 3$) に対して, $\mathbf{0}$, $\boldsymbol{x}$, $\boldsymbol{x} + \boldsymbol{y}$, $\boldsymbol{y}$ 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. 我々は三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, 次節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**

---

[3] 太字 $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}, \mathbf{C}$ を用いる場合もある.

[4] 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.

[5] 記号 $\mathbb{Z}$ はドイツ語で数を意味する Zahlen に由来する.

[6] 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.

[7] 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.

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．

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

1.6 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．

定義 1.6.1. 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:

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

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

例 1.6.2. 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 $f : \mathbb{R} \to \mathbb{R}$ 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 1.6.3. 二つの写像 $f : X \to Y$ および $g : Y \to Z$ 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 $g\bigl(f(x)\bigr)$ を対応させる写像を f と g の**合成** (**composition**) と呼び, 記号 $g \circ f : X \to Z$ 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 1.6.4. ある実数 a を用いて $f(x) := ax$ と定められる関数 $f : \mathbb{R} \to \mathbb{R}$ を**比例関数**という[8]．

例 1.6.5. 関数 $g : \mathbb{R} \to \mathbb{R}$ および $f : \mathbb{R} \to \mathbb{R}$ 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, $g \circ f(x) = g\bigl(f(x)\bigr) = g(bx) = a(bx) = (ab)x$.

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ろうか．

$$\boxed{\begin{array}{c} \text{実数全体} \\ r \in \mathbb{R} \end{array}} \quad \longleftrightarrow \quad \boxed{\begin{array}{c} \text{比例関数全体} \\ f(x) = rx \end{array}}$$

[8]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」は本論でのみ通じる用語である．

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 1.6.5). あまりに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, これこそが 1.4 項の代数学の説明で述べた "二つの異なる世界を一つの見方で繋げること" の一例なのである. 今の話では実数をただの数と思うだけでなく, 比例関数である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. これを次項で詳しく述べよう.

備考 1.6.6. 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.

1.7 そして線形代数学へ

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 1.7.1. 次の性質 (i) および (ii) を持つ写像 $f : \mathbb{R}^n \to \mathbb{R}^m$ を**線形写像** (**linear map**) という:

(i) すべての $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ に対して, $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$,

(ii) すべての $\boldsymbol{x} \in \mathbb{R}^n$ および各 $r \in \mathbb{R}$ について, $f(r\boldsymbol{x}) = rf(\boldsymbol{x})$.

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.

練習 1.7.2. 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 □ 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後期で扱う. ここでは $\mathbb{R}^2$ の間の線形写像を挙げよう.

例題 1.7.3. あらかじめ実数 a, b, c, d を与えておき, 写像 $f : \mathbb{R}^2 \to \mathbb{R}^2$ 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) $\boldsymbol{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ および $\boldsymbol{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 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:

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

(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 $\boldsymbol{a} = \begin{pmatrix} x \\ y \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} x' \\ y' \end{pmatrix}$ に対して,

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

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 $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ および $r \in \mathbb{R}$ に対して,

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

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, $g : \mathbb{R}^2 \to \mathbb{R}^2$ 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 $g : \mathbb{R}^2 \to \mathbb{R}^2$ 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.

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:

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

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, 前項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 1.7.4. $\mathbb{R}^2$ 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 1.7.5. $\mathbb{R}^2$ 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 $g, f : \mathbb{R}^2 \to \mathbb{R}^2$ 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 $m\times n$ 行列全体の間の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とは線形写像を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数値化することで, 写像を視覚化するというのが行列理論のねらいである. 前期の講義では, 行列に関する計算を理解することを目標とする. また後期においては, 行列を駆使した線形写像の分析について学ぶ. 今後の講義において, 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が, 行列理論の目的が線形写像の分析にあるということを念頭において学習してもらいたい.

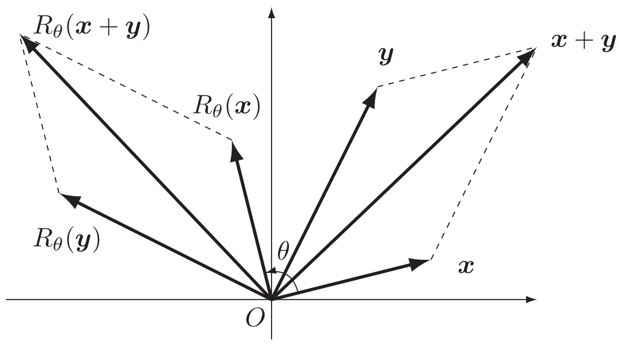
練習 1.7.6. 二つの線形写像 $\xi, \eta : \mathbb{R}^2 \to \mathbb{R}^2$ を $\xi\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$, $\eta\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} Fx+Gy \\ Hx+Iy \end{pmatrix}$ で定める. このとき $\xi\circ\eta\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} (aF+bH)x+(aG+bI)y \\ (cF+dH)x+(cG+dI)y \end{pmatrix}$ となることを確かめよ.

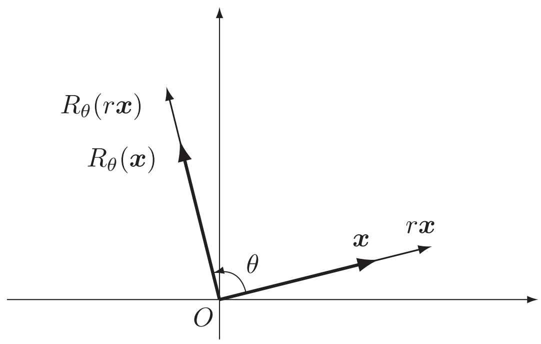
よりみち (加法定理).

$\mathbb{R}^2$ の各ベクトル $\boldsymbol{x}$ に対して, 原点 O を中心に $\boldsymbol{x}$ を θ 回転させたベクトルを対応させる写像 $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ を考える. これは明らかに線形写像である. 実際, 定義 1.7.1 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 $R_\theta(\boldsymbol{x}+\boldsymbol{y})$ は二つのベクトル $R_\theta(\boldsymbol{x})$, $R_\theta(\boldsymbol{y})$ で張られる平行四辺形の頂点となる. これは $R_\theta(\boldsymbol{x}+\boldsymbol{y}) = R_\theta(\boldsymbol{x}) + R_\theta(\boldsymbol{y})$ 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 $\boldsymbol{x}$ と $r\boldsymbol{x}$ における長さの比と, これらを R_θ で写像した $R_\theta(\boldsymbol{x})$ と $R_\theta(r\boldsymbol{x})$ における長さの比は共に $1:r$ である. 更に, $\boldsymbol{x}$ と $r\boldsymbol{x}$ は平行ゆえ θ 回転後の $R_\theta(\boldsymbol{x})$ と $R_\theta(r\boldsymbol{x})$ も平行である. 以上のことから $R_\theta(r\boldsymbol{x}) = rR_\theta(\boldsymbol{x})$ 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:

$$
\begin{aligned}
\cos(\alpha+\beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta, \\
\sin(\alpha+\beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta,
\end{aligned}
$$

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.

2　行列の演算

この節では行列 (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．前節で見たように, 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:

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

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 $i = 1, \cdots, m$ および $j = 1, \cdots, n$ 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[9],

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

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

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

と書く．更に (i, j) の動く範囲に誤解がないとき, $A = [a_{ij}]$ 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．

---

[9] (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．

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:

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

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:

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

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, $ax + by$ という形の数を成分にもつことを前節で見た. 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 $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ の間の積 $A\boldsymbol{x}$ を, やや天下り的ではあるものの定義してしまおう. $A\boldsymbol{x}$ は次で定義される m 次列ベクトルである:

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

ここで A の列の数と $\boldsymbol{x}$ の成分数が等しく, A の行の数と $A\boldsymbol{x}$ 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
交通費 (円)	1000	500	0
おにぎり (個)	3	2	1
⋮	⋮	⋮	⋮

	斎藤	田端	嶺	⋯
大人	4	1	2	⋯
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,

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

通常は, 積の記号 $\cdot$ (ドット) を略して AB と書くことが多い.

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[10]. 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.

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.

$$\left[\begin{array}{ccc} 1 & 0 & 1 \\ \hline 3 & 2 & 0 \\ \hline 1 & 1 & 2 \end{array}\right] \left[\begin{array}{c|c|c} 0 & 3 & 1 \\ 2 & 0 & 1 \\ 0 & 2 & 0 \end{array}\right] = \begin{bmatrix} 1\cdot 0+0\cdot 2+1\cdot 0 & 1\cdot 3+0\cdot 0+1\cdot 2 & 1\cdot 1+0\cdot 1+1\cdot 0 \\ 3\cdot 0+2\cdot 2+0\cdot 0 & 3\cdot 3+2\cdot 0+0\cdot 2 & 3\cdot 1+2\cdot 1+0\cdot 0 \\ 1\cdot 0+1\cdot 2+2\cdot 0 & 1\cdot 3+1\cdot 0+2\cdot 2 & 1\cdot 1+1\cdot 1+2\cdot 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 5 & 1 \\ 4 & 9 & 5 \\ 2 & 7 & 2 \end{bmatrix}.$$

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.

[10] 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.

(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, $O_{mn} = [0]_{\substack{i=1,\cdots,m,\\ j=1,\cdots,n}}$ 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$\boldsymbol{n}$**次正方行列** (**square matrix of order $\boldsymbol{n}$**) という.
- n 次正方行列 $A = [a_{ij}]$ において (i,i)-成分 a_{ii} $(i = 1, \cdots, n)$ 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. 本論では 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[11].

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

- $(-1)A$ 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.

(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) $(AB)C = A(BC)$.

[11] 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, 代数演算においてほどこして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.

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) からも導ける:

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

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．

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

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

□

<u>(16) $A(B+C)=AB+AC$ の証明.</u>　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, 次節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．

$$\text{(a) } (A+2E)(A+E) = A^2+3A+2E, \qquad \text{(b) } (A+B)^2 = A^2+2AB+B^2.$$

(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. 例えば, 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 節以降で論じる. 群については 7 節のコラムで紹介する.

3 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, いくつかの補足次項を説明する. 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 $a_1, \cdots, a_n$ たちの和 $a_1 + a_2 + \cdots + a_n$ を

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

と表す. 積 $a_1 a_2 \cdots a_n$ については

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

と表す. また, a_{ij} $(i = 1, \cdots, m,\ j = 1, \cdots, n)$ たちの総和を $\displaystyle\sum_{\substack{i=1,\cdots,m,\\ j=1,\cdots,n}} a_{ij}$ 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 $\cdots$ または $i = n$"

  注: この条件文 $P(i)$ を略して, 我々は "$i = 1, \cdots, n$" 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[12].

[12] 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.

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

  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 $P(i)$ を略して “$i,j=1,\cdots,n$” 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 $\displaystyle\sum_{P(i)} a_i$ と書く. 積 $\displaystyle\prod_{P(i)} a_i$ についても同様に定める.

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

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

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

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

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

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

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

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, $\displaystyle\sum_i$ あるいは $\displaystyle\prod_i$ と略記することがある. 本論ではこのような曖昧な表記は行わない.

例 3.2.4.

$$
\begin{aligned}
\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\right) &= \sum_{i=1}^{m}(a_{i1}+a_{i2}\cdots+a_{in}) \\
&= (a_{11}+a_{12}\cdots+a_{1n})+(a_{21}+a_{22}\cdots+a_{2n})+\cdots+(a_{m1}+a_{m2}\cdots+a_{mn}) \\
&= \sum_{\substack{i=1,\cdots,m,\\ j=1,\cdots,n}} a_{ij} \\
&= (a_{11}+a_{21}\cdots+a_{m1})+(a_{12}+a_{22}\cdots+a_{m2})+\cdots+(a_{1n}+a_{2n}\cdots+a_{mn}) \\
&= \sum_{j=1}^{n}(a_{1j}+a_{2j}\cdots+a_{mj}) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}\right).
\end{aligned}
$$

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.

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

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

$(AB)C$ 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 $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$ 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 $\Big((h \circ g) \circ f\Big)(x) = \Big(h \circ (g \circ f)\Big)(x)$ となることを示そう. まず $h \circ g$ を ϕ, $f(x)$ を y とおくことで

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

次に $\psi = g \circ f$ とおくことで

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

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, ユークリッド空間の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.

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

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 $\mathlarger{*}$ 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:

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

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)-行列の対角成分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. これは形式的には次のように定義されるものである.

定義 3.4.1. (m,n)-行列 $A = [a_{ij}]$ に対して, $b_{kh} := a_{hk}$ (ただし $k = 1, \cdots, n$, $h = 1, \cdots, m$) を成分とする (n,m)-行列 $[b_{kh}]$ を A の**転置行列** (**transposed matrix**) といい, ${}^tA$ 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) $${}^t\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix} = \begin{bmatrix} 1 & 6 & 11 \\ 2 & 7 & 12 \\ 3 & 8 & 13 \\ 4 & 9 & 14 \\ 5 & 10 & 15 \end{bmatrix}.$$

(3) 上三角行列の転置行列を**下三角行列** (**lower triangular matrix**) と呼ぶ.

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:

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

(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 $\boldsymbol{x}, \boldsymbol{y}$ に対して, これらの**内積**を $\boldsymbol{x} \cdot {}^t\boldsymbol{y}$ で定め, 列ベクトル $\boldsymbol{x}, \boldsymbol{y}$ に対して, これらの**内積**を ${}^t\boldsymbol{x} \cdot \boldsymbol{y}$ 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[13]．いまは, A の各行 (あるいは列) どうし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．

例 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, $[b_{ij}]_{\substack{i=1,\cdots,n,\\ j=1,\cdots,m}}$ となる．すなわち,

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

上の右辺と左辺では 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 $A = [a_{ij}]_{\substack{i=1,\cdots,m,\\ j=1,\cdots,n}}$ に対して, 式 (3.4.1) の右辺における i, j の動く範囲を入れ替えた表示に相当する $B = [a_{ji}]_{\substack{i=1,\cdots,m,\\ j=1,\cdots,n}}$ は A の転置行列とは限らない．A の転置行列は (n, m)-行列でなければならないが, B は (m, n)-行列であり, これは一般には (n, m)-行列ではない．例外は $m = n$ の場合, すなわち A が正方行列の場合であり, この場合に限り B は A の転置行列に一致する．この点について, 混乱を避けるため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

$$
a_{11} = 1, \quad a_{12} = 2, \quad a_{13} = 3, \quad a_{21} = 4, \quad a_{22} = 5, \quad a_{23} = 6
$$

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, $B = [a_{ji}]_{\substack{i=1,2,\\ j=1,2,3}}$ と定めようとすれば, A, B の成分表示は次のようになる:

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

つまり, a_{31} と a_{32} は未だ定めていないゆえ, B は定義そのものが不適切であることが分かる．A が正方行列の場合はこのような不具合は生じない．

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

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

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

[13] 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．

である. ${}^t\!B\,{}^t\!A$ および ${}^t(AB)$ の各成分をそれぞれ計算しよう.

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

一方で,

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

以上より, ${}^t\!B\,{}^t\!A$ 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.

$$
\left[\begin{array}{cc|c|cc} 1 & 2 & 3 & 4 & 5 \\ \hline 6 & 7 & 8 & 1 & 0 \\ 11 & 12 & 13 & 0 & 1 \end{array}\right]
\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 6 & 7 \\ \hline 1 & 2 \\ 0 & 1 \end{array}\right]
= \left[\begin{array}{c}
\begin{bmatrix} 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 3\begin{bmatrix} 6 & 7 \end{bmatrix} + \begin{bmatrix} 4 & 5 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \\
\begin{bmatrix} 6 & 7 \\ 11 & 12 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 8 \\ 13 \end{bmatrix}\begin{bmatrix} 6 & 7 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}
\end{array}\right]
$$

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

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[14]. 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.

[14] 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「線形代数学 斉藤雅彦 著 (東京大学出版会)」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, $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ を n 次行ベクトル, $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_\ell$ を n 次列ベクトルとして

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

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 $z_{ik} = \boldsymbol{a}_i \boldsymbol{b}_k$). AB の分割として, 次のような表し方がある:

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

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

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 $A = {}^t[{}^t\boldsymbol{a}_1, \cdots, {}^t\boldsymbol{a}_m]$, $B = [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_\ell]$ 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.

4 連立 1 次方程式

本節の目標は, 連立 1 次方程式と行列の関係を見極め, 方程式の解全体の集合がどのような形になるか理論立てて理解することにある. 本論において「方程式を解く」とは, すべての解を出しつくすことを意味する. これは, あてずっぽうで解の一つや二つを見つけるという話ではない. 理論的な見地から, 解はこれら以外にはありえないということまで我々は示さねばならない.

4.1 導入

一般に, 方程式の解は無限にたくさんあるかもしれないし, 一つも存在しないということもあり得る. この事実を次の三つの方程式を例に確かめてみよう.

$$(1)\ \begin{cases} x+y=1, \\ 2x+2y=4. \end{cases} \qquad (2)\ \begin{cases} x+y=1, \\ 2x+2y=2. \end{cases} \qquad (3)\ \begin{cases} x+y=1, \\ x-y=1. \end{cases}$$

方程式 (1) の解は存在しない. 何故なら, xy 平面において直線 $x+y=1$ と直線 $2x+2y=4$ は平行ゆえ交わらないことが図を描けば分かるからである. 図の助けを借りずに論証する場合は次のようになるだろう. 図は理解を助けるうえで重要ではあるが, 錯覚している可能性が残されていることを忘れてはならない.

<u>解の非存在証明.</u>　仮に方程式 (1) に解 $(x,y)=(a,b)$ が存在すると仮定すると, 上段の式を 2 倍することで $2a+2b=2$ を得る. また, 下段の式からは $2a+2b=4$ を得る. 以上より $4=2$ となる. これは明らかにおかしい. ゆえに方程式 (1) の解は存在しない. □

方程式 (2) の上下の式は互いに定数倍した関係にあるから, 二つの式の意味は同じであり, ゆえに直線 $x+y=1$ 上の点 $(a,1-a)$ が方程式 (2) の解となる. a は任意の実数をとり得るため, この方程式の解は無数にある. 方程式 (3) の解は二つの直線 $x+y=1$ および $x-y=1$ の交点である. すなわち, 解は唯一つ $(x,y)=(1,0)$ のみである.

いまの例によって, 解の形には色々な可能性があり得ることが分かった.

4.2 連立 1 次方程式と行列

いま, 実数 a_{ij} および b_i (ただし $i=1,\cdots,m,\ j=1,\cdots,n$) が既知の数であるとしよう. このとき n 個の変数 $x_1,x_2,\cdots,x_n$ に関する連立方程式

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m, \end{cases} \tag{4.2.1}$$

を解くことと, 行列に関する次の関係式

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

を満たす n 次列ベクトル $\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ 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 $\boldsymbol{b}$ が与えられたとき, $A\boldsymbol{x} = \boldsymbol{b}$ を満たす未知のベクトル $\boldsymbol{x}$ 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. このとき, 列ベクトル $\boldsymbol{a}$ が方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解であると仮定すれば (すなわち $A\boldsymbol{a} = \boldsymbol{b}$ を満たす), この両辺に左から B をかけることで $BA\boldsymbol{a} = B\boldsymbol{b}$ を得る. 左辺を計算すると $BA\boldsymbol{a} = E\boldsymbol{a} = \boldsymbol{a}$ ゆえ $\boldsymbol{a} = B\boldsymbol{b}$ 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 $\boldsymbol{x} = B\boldsymbol{b}$ のみであることが分かった. また, 実際に $\boldsymbol{x} = B\boldsymbol{b}$ はこの方程式の解である. 何故なら, $A\boldsymbol{x} = AB\boldsymbol{b} = E\boldsymbol{b} = \boldsymbol{b}$ ゆえ $A\boldsymbol{x} = \boldsymbol{b}$ を満たすからである. 以上より次を得る.

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 $A\boldsymbol{x} = \boldsymbol{b}$ は唯一解 $\boldsymbol{x} = B\boldsymbol{b}$ を持つ.

この命題は理論的にはすっきりとしているものの, いくつかの欠点がある. 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ことである. 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, そうとも知らずに上の命題に準じた方向で挑んでしまったら, 計算が徒労に終わってしまうのだろうか. 実は, 議論を進めていくと, これらの問題はすべて繋がっていることが分かる.

4.4 行列の行基本変形

A を (m, n)-行列とする. このとき $A\boldsymbol{x} = \boldsymbol{b}$ 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, 係数が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) は方程式の意味を変えない. 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す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 (left elementary transformation)** 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 $\boldsymbol{b}$ を並べた, $(m,n+1)$-行列 $[A|\boldsymbol{b}]$ を考えよう.

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

これを連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ 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 $[A|\boldsymbol{b}]$ 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 $[A'|\boldsymbol{b}']$ としよう. このとき, n 次列ベクトル $\boldsymbol{a}$ について, $\boldsymbol{a}$ が方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解であることと, $\boldsymbol{a}$ が方程式 $A'\boldsymbol{x} = \boldsymbol{b}'$ 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 $[A|\boldsymbol{b}]$ 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 $[A|\boldsymbol{b}]$ に有限回の行基本変形をほどこしたものを $[A'|\boldsymbol{b}']$ とすれば, 方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解全体と方程式 $A'\boldsymbol{x} = \boldsymbol{b}'$ の解全体は一致する.

さて, 連立方程式 $A\boldsymbol{x} = \boldsymbol{b}$ を解くには, 行基本変形を用いてより簡単な連立方程式 $A'\boldsymbol{x} = \boldsymbol{b}'$ 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, $[A|\boldsymbol{b}]$ を運よく $[E|\boldsymbol{b}']$ 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, $\boldsymbol{x} = \boldsymbol{b}'$ 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, それは簡約行列と呼ばれている.

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.

$$\begin{bmatrix} \underline{2} & 0 & 1 & 5 & 4 \\ 0 & 0 & \underline{1} & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \underline{1} \end{bmatrix}.$$

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, $j_1 < j_2 < j_3 < \cdots$ 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[15], 簡約形

[15]参考書によっては, 行階段形に条件 (II) も仮定する.

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 $j_1 \leq j_2 \leq j_3 \leq \cdots$ 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) $j_1 < j_2 < j_3 < \cdots$ 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.

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

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

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 項で述べよう (定理 6.1.1).

4.6 連立 1 次方程式の解法

簡約行列 $[B|\boldsymbol{b}]$ における連立 1 次方程式 $B\boldsymbol{x} = \boldsymbol{b}$ の解法を具体例を通して説明する. $[B|\boldsymbol{b}]$ 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 $[B|\boldsymbol{b}]$ の最後の列 $\boldsymbol{b}$ に主成分が含まれる場合である. 例えば,

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

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 $\boldsymbol{x}$ は存在しないゆえ, したがって方程式 $B\boldsymbol{x} = \boldsymbol{b}$ の解は存在しない.

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

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

このとき, <u>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</u> 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 $\boldsymbol{x}$ が方程式 $B\boldsymbol{x} = \boldsymbol{b}$ 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

$$(\sharp_1) \begin{cases} 0x_1 + 1x_2 + 3x_3 + 0x_4 + 2x_5 = 4, \\ 0x_1 + 0x_2 + 0x_3 + 1x_4 + 1x_5 = 5, \end{cases}$$

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.

$$(\sharp_2) \begin{cases} x_2 = 4 - 3x_3 - 2x_5 & = 4 - 3c_2 - 2c_3, \\ x_4 = 5 - x_5 & = 5 - c_3. \end{cases}$$

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 $B\boldsymbol{x} = \boldsymbol{b}$ の解 $\boldsymbol{x}$ が存在するとすれば, ベクトル $\boldsymbol{x}$ は次の形になっていることが分かった:

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

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 $\boldsymbol{x}$ を定めてみよう. このとき $\boldsymbol{x}$ が方程式 $B\boldsymbol{x} = \boldsymbol{b}$ を満たすことは明らかである. 何故なら, 上でもって $\boldsymbol{x}$ を定めるということは, 式 $(\sharp_2)$ が満たされるように定めたことを意味し, 式 $(\sharp_2)$ を移項すれば元々の連立方程式 $(\sharp_1)$ が得られるからである. これまでの議論により, 方程式 $B\boldsymbol{x} = \boldsymbol{b}$ の解は存在し, しかも解の形は次のもので出しつくされていることが分かった:

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

また, 解の形を次のように整理しておくと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 $\boldsymbol{a}_1, \boldsymbol{a}_2 \in \mathbb{R}^3$ を取って固定しよう. 次に $\boldsymbol{y} = c\boldsymbol{a}_1 + d\boldsymbol{a}_2$ 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 $\boldsymbol{a}_0 \in \mathbb{R}^3$ を取り, $\boldsymbol{x} = \boldsymbol{a}_0 + c\boldsymbol{a}_1 + d\boldsymbol{a}_2$ 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 $\boldsymbol{a}_0$ 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 $A\boldsymbol{x} = \boldsymbol{b}$ の解全体を W としよう. 本項で与えた連立 1 次方程式の解法によれば, 固定された n 次列ベクトル $\boldsymbol{a}_0, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$ および任意定数 $c_1, \cdots, c_k$ を用いて, W の元は次のような形で表示される:

$$\boldsymbol{x} = \boldsymbol{a}_0 + c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_k\boldsymbol{a}_k. \tag{4.6.1}$$

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

$$\boldsymbol{y} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_k\boldsymbol{a}_k.$$

例 4.6.1. 上で解いた連立 1 次方程式 $\begin{bmatrix} 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$ 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, $\begin{bmatrix} 0 \\ 4 \\ 0 \\ 5 \\ 0 \end{bmatrix}$ 方向に平行移動した集合 W に一致する.

本節で与えた連立 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 $A\boldsymbol{x} = \boldsymbol{b}$ の解法は次のとおりである. まず拡大係数行列 $[A|\boldsymbol{b}]$ の簡約化 $[B|\boldsymbol{b}']$ を求める. このとき, $A\boldsymbol{x} = \boldsymbol{b}$ の解全体と $B\boldsymbol{x} = \boldsymbol{b}'$ の解全体は一致する. $[B|\boldsymbol{b}']$ の最後の列 $\boldsymbol{b}'$ 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 $B\boldsymbol{x} = \boldsymbol{b}'$ 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, 方程式の解は唯一つである. また, 方程式の解をベクトルの和に分解しておくと, 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[16]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, 線形写像の準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, 準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, 後期の講義において, 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. その際に, 講義中にいちいち復習する余裕はないゆえ, 本節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

本節において, 連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ 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 $A\boldsymbol{x} = \boldsymbol{b}$ 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

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 2 & 2 \end{array}\right] \xrightarrow{\text{簡約化}} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

の解として, 次の四つの形が挙げられる:

$$(1)\ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \qquad (2)\ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -2 \\ 2 \end{bmatrix},$$

$$(3)\ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad (4)\ \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + d \begin{bmatrix} 1 \\ -1 \end{bmatrix} + e \begin{bmatrix} 2 \\ -2 \end{bmatrix},$$

ただし, a, b, c, d, e はそれぞれ任意定数である. (1) は本節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, 本節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.

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, 本節で与えた解法において任意定数の水増しはないか.

[16] 同次形の方程式は 6.3 項で扱う.

(ii) 任意定数の水増しの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, 後期で学ぶ線形空間論を踏まえた上でなされる (22.2 項で述べる). 未定義語の羅列になることを承知で述べると, 線形独立性と呼ばれる概念を通して (i) の判定がなされ, とくに本節で与えた解法において任意定数の水増しはない. (ii) は正しく, 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ことが分かる. しかし, その定義の妥当性の議論 (すなわち (i) と (ii) の証明) は後期まで待たねばなない. 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」とは本節で与えた解法 (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.

5 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, 前節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 19.4 項にて, 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と, $B = C$ 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, $\left(A^{-1}\right)^{-1} = B^{-1} = A$ が成り立つ.

(3) 各 $X_1, X_2, \cdots, X_k$ 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と, $AB = O$ 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).

5.2 行基本変形再考

(m,n)-行列 A の行基本変形について再考する. まず次のような特別な正方行列を考えよう.

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, この講義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.

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

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

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

(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.

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

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こと.

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:

$$\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & r & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ ra_{i1} & ra_{i2} & \dots & ra_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

(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:

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & & 1 & \\ & & & \ddots & & \\ & & 1 & & 0 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

(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:

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & r & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \left[\begin{array}{c|c|c} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \vdots & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{array}\right] = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \vdots & \vdots \\ ra_{i1} + a_{j1} & \dots & ra_{in} + a_{jn} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

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 $X_1, \cdots, X_k$ 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}, $\cdots$, 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. これを次項で見ていこう.

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, また前項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．その理由は次節で述べる定理 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． □

本項では, $BA = E$ 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, 本項とは別の方法で $DA = E$ 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 節を見よ).

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)

本節では, 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てきた. 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.

6　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. 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は, 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 $\boldsymbol{0}$ に限る. $\boldsymbol{x} = \boldsymbol{0}$ の場合, $\boldsymbol{0}$ に基本変形をいくらほどこしても変化せず, $\boldsymbol{0}$ の簡約化は $\boldsymbol{0}$ 自身以外にありえない. また, $\boldsymbol{x} \neq \boldsymbol{0}$ の場合, $\boldsymbol{x}$ を $\boldsymbol{0}$ に行基本変形することはできない. 何故なら, もし $\boldsymbol{0}$ が $\boldsymbol{x}$ の簡約化であるならば, 命題 5.2.4 により $\boldsymbol{0}$ を行基本変形して $\boldsymbol{x} \neq \boldsymbol{0}$ 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:

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

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

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

$[B|\boldsymbol{b}], [C|\boldsymbol{c}]$ はともに $[A|\boldsymbol{a}]$ を行基本変形を繰り返して得られる行列である．ゆえに $[B|\boldsymbol{b}]$ に行基本変形を繰り返すことで $[C|\boldsymbol{c}]$ 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 $X[B|\boldsymbol{b}] = [C|\boldsymbol{c}]$ となる (命題 5.2.4)．X を次のように分割する:

$$X = \begin{bmatrix} P_{n,n} & Q_{n,m-n} \\ R_{m-n,n} & S_{m-n,m-n} \end{bmatrix} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix},$$

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．$X[B|\boldsymbol{b}]$ を計算すると次のようになる．

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

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

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

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

$\boldsymbol{c}$ が主成分を含まない場合は, $\boldsymbol{b}$ と $\boldsymbol{c}$ 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 $[B|\boldsymbol{b}] = [C|\boldsymbol{c}]$ 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 $\operatorname{rank} A$ 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．

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

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, $\operatorname{rank} A = \operatorname{rank} C$ 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 $A\boldsymbol{x} = \boldsymbol{b}$ の解の形は, 拡大係数行列 $[A|\boldsymbol{b}]$ 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 $A\boldsymbol{x} = \boldsymbol{b}$ において次が成り立つ.

(1) $A\boldsymbol{x} = \boldsymbol{b}$ の解が存在するための必要十分条件は $\operatorname{rank}[A|\boldsymbol{b}] = \operatorname{rank} A$ である.

(2) $A\boldsymbol{x} = \boldsymbol{b}$ の解が存在するとき, 解の任意定数の個数は $n - \operatorname{rank} A$ である.

(3) $A\boldsymbol{x} = \boldsymbol{b}$ の解が唯一つであるための必要十分条件は, $\operatorname{rank}[A|\boldsymbol{b}] = \operatorname{rank} A = n$ 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) $\operatorname{rank} A = n$,　(2) A の簡約化は E である,

(3) 任意の n 次列ベクトル $\boldsymbol{b}$ について方程式 $A\boldsymbol{x} = \boldsymbol{b}$ は唯一つの解を持つ,

(4) 方程式 $A\boldsymbol{x} = \boldsymbol{0}$ は唯一つの解 $\boldsymbol{0}$ 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)⇒(2) : $\operatorname{rank} A = n$ とすれば, A の簡約化 B の主成分の個数は n である. B は $n \times n$ 行列であるゆえ, $B = E$ となる.

(2)⇒(3) : $[A|\boldsymbol{b}]$ の簡約化を $[E|\boldsymbol{b}']$ とする. 命題 4.4.1 より方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解全体と $E\boldsymbol{x} = \boldsymbol{b}'$ の解全体は等しい. 方程式 $E\boldsymbol{x} = \boldsymbol{b}'$ が唯一の解 $\boldsymbol{x} = \boldsymbol{b}'$ を持つことは明らかであり, したがって $A\boldsymbol{x} = \boldsymbol{b}$ の解も $\boldsymbol{x} = \boldsymbol{b}'$ 唯一つである. (別証明: $\operatorname{rank} A = \operatorname{rank} E = n$ に注意する. $\operatorname{rank}[E|\boldsymbol{b}'] = n$ ゆえ $\operatorname{rank}[A|\boldsymbol{b}] = \operatorname{rank}[E|\boldsymbol{b}'] = n = \operatorname{rank} A$. よって, 命題 6.2.1(3) より方程式は唯一つの解を持つ.)

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

(4)⇒(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 $0 = n - \operatorname{rank} A$. ゆえに $\operatorname{rank} A = n$.

(2)⇒(5) : 5.3 項で論じた逆行列の求め方より得る.

(5)⇒(4) : これは命題 4.3.1 の特別な場合に相当する.

$\boldsymbol{x} = \boldsymbol{0}$ が方程式 $A\boldsymbol{x} = \boldsymbol{0}$ の解であることは明らかである. また, 解の一意性も直ちに分かる. 実際, $A\boldsymbol{x} = \boldsymbol{0}$ の両辺に左から A^{-1} を掛けることで $\boldsymbol{x} = \boldsymbol{0}$ 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, $A\boldsymbol{x} = \boldsymbol{0}$ なる方程式を**同次形**あるいは**斉次形**の方程式という[17]. 同次形の方程式は $\boldsymbol{x} = \boldsymbol{0}$ 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 $A\boldsymbol{x} = \boldsymbol{0}$ の拡大係数行列 $[A|\boldsymbol{0}]$ の簡約化は $[B|\boldsymbol{0}]$ であるから $\operatorname{rank}[A|\boldsymbol{0}] = \operatorname{rank}[B|\boldsymbol{0}] = \operatorname{rank} B = \operatorname{rank} A$. 確かに $\operatorname{rank}[A|\boldsymbol{0}] = \operatorname{rank} A$ は成り立っている.

[17] 同次および斉次はともに homogeneous の訳語である.

4.6 項で述べた連立 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 $k = n - \operatorname{rank} A$ 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 $[B|\mathbf{0}]$ なる形になり, ゆえに式 (4.6.1) において $\boldsymbol{a}_0 = \mathbf{0}$ となる．したがって解の表示は

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

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．

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 6.3.2. 方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解 $\boldsymbol{a}$ 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 $A\boldsymbol{x} = \mathbf{0}$ の任意の解 $\boldsymbol{z}$ に対し, $\boldsymbol{a} + \boldsymbol{z}$ は方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解である．

(2) 方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の任意の解 $\boldsymbol{y}$ は, 方程式 $A\boldsymbol{x} = \mathbf{0}$ のとある解 $\boldsymbol{z}$ を用いて $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{z}$ と表せる．

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

(2) : $\boldsymbol{y}$ を $A\boldsymbol{x} = \boldsymbol{b}$ の解とする．ここで $\boldsymbol{z} := \boldsymbol{y} - \boldsymbol{a}$ とおこう．このとき $\boldsymbol{z}$ は $A\boldsymbol{x} = \mathbf{0}$ の解である．何故なら, $A\boldsymbol{z} = A(\boldsymbol{y} - \boldsymbol{a}) = A\boldsymbol{y} - A\boldsymbol{a} = \boldsymbol{b} - \boldsymbol{b} = \mathbf{0}$ ゆえ $\boldsymbol{z}$ は $A\boldsymbol{z} = \mathbf{0}$ を満たすからである．また, $\boldsymbol{z}$ の定め方から $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{z}$ 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, $A\boldsymbol{x} = \boldsymbol{b}$ の解全体の集合 W は H を $\boldsymbol{a}$ 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 項にて説明していた．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だろう: 方程式 $A\boldsymbol{x} = \mathbf{0}$ の解法は, 方程式 $A\boldsymbol{x} = \boldsymbol{b}$ 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 $A\boldsymbol{x} = \mathbf{0}$ の解を求めておき, 更に, 何らかの方法で $A\boldsymbol{x} = \boldsymbol{b}$ の解 $\boldsymbol{a}$ を一つでよいから見つけてくる．すると, $A\boldsymbol{x} = \boldsymbol{b}$ の解は, $A\boldsymbol{x} = \mathbf{0}$ の解全体を $\boldsymbol{a}$ 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, 次の命題を与えよう．

命題 6.3.3. $\alpha(x), \beta(x), \gamma(x), \delta(x)$ 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[18]を考える．

$$\text{(I)}\ \ \alpha(x)f''(x) + \beta(x)f'(x) + \gamma(x)f(x) = \delta(x), \quad \text{(II)}\ \ \alpha(x)f''(x) + \beta(x)f'(x) + \gamma(x)f(x) = 0.$$

方程式 (I) の解 $a(x)$ を一つとって固定しよう．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 $y(x)$ は, 微分方程式 (II) のとある解 $z(x)$ を用いて $y(x) = a(x) + z(x)$ と表せる．

Proof. (1) : $z(x)$ を (II) の解とすると,

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

ゆえに, $a(x) + z(x)$ は微分方程式 (I) の解である．

[18] 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,

$$\begin{aligned}\alpha(x)(y(x)-a(x))'' + \beta(x)(y(x) - a(x))' + \gamma(x)(y(x) - a(x))\\ &= \Big(\alpha(x)y''(x) + \beta(x)y'(x) + \gamma(x)y(x)\Big) - \Big(\alpha(x)a''(x) + \beta(x)a'(x) + \gamma(x)a(x)\Big)\\ &= \delta(x) - \delta(x) = 0.\end{aligned}$$

ゆえ $z(x)$ は (II) を満たすからである．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． □

命題 6.3.2 と 6.3.3 の証明がパラレルであることに着目せよ．これはベクトル $\boldsymbol{x}$ に対して $A\boldsymbol{x}$ を対応させる操作と, 関数 $f(x)$ に対して関数 $\alpha(x)f''(x) + \beta(x)f'(x) + \gamma(x)f(x)$ 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．このように, 似たような証明を何度も繰り返し行う手間を省くために, 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．

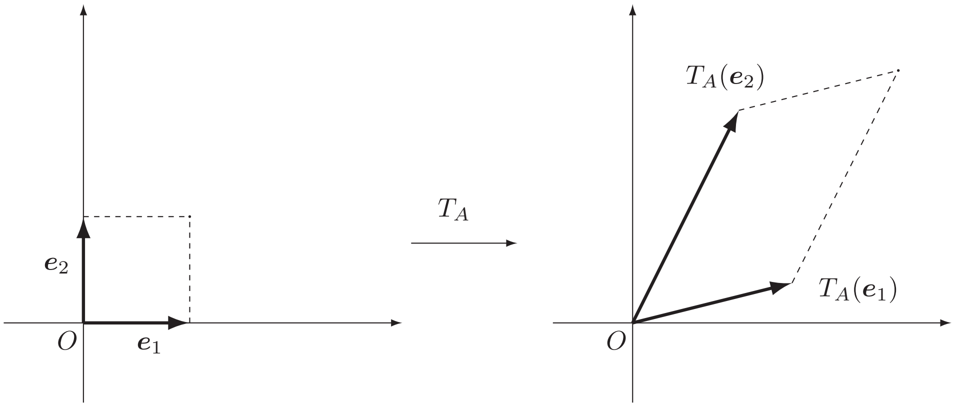
7 行列式とは何か

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, 本節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見る方法で, 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 $\mathbb{R}^n$ 上の線形変換の面積拡大率

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ とし, 写像 $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ を $T_A(\boldsymbol{x}) := A\boldsymbol{x}$ で定める (21.3 項以降において, 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.



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

単位正方形とは $\boldsymbol{e}_1, \boldsymbol{e}_2$ 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 $T_A(\boldsymbol{e}_1)$ と $T_A(\boldsymbol{e}_2)$ 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 $T_A(\boldsymbol{e}_1)$ を $T_A(\boldsymbol{e}_2)$ 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. $\boldsymbol{e}_1, \boldsymbol{e}_2$ 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, $T_A(\boldsymbol{e}_1), T_A(\boldsymbol{e}_2)$ 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 1 の平行四辺形の面積は, 原点と $T_A(\boldsymbol{e}_1) + T_A(\boldsymbol{e}_2)$ 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,

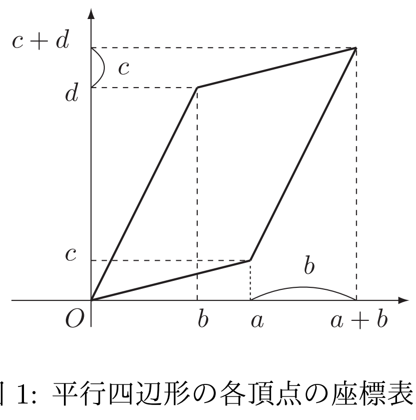


図 1: 平行四辺形の各頂点の座標表示

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

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 $T_A : \mathbb{R}^n \to \mathbb{R}^n$ を $T_A(\boldsymbol{x}) := A\boldsymbol{x}$ 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[19]:

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

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 節で述べることとしよう.

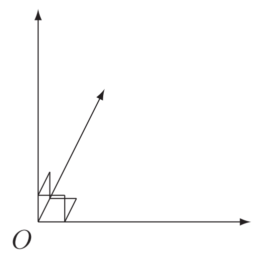
さて, 行列式が体積拡大率を意味するのであれば, T_E は元を動かさない恒等写像ゆえ $|E| = 1$ である. また, $|AB|$ は合成写像 $T_A \circ T_B$ の体積拡大率であるから, これは $|A| \cdot |B|$ に等しい. つまり $|AB| = |A||B|$ となる. 更に A が可逆であるとき, $|A||A^{-1}| = |AA^{-1}| = |E| = 1$ より $|A^{-1}| = |A|^{-1}$, とくに $|A| \neq 0$ である. 逆に, $|A| \neq 0$ ならば A が可逆であることも後に示され, したがって $|A|$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

[19] 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 節で説明する.

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 項で述べる) 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　クラメルの公式

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,

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

ここで, 上式の最後の $(1,3)$-成分は

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

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 $A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ 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 節で述べる). 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, 実際に解を求めるための公式としてはあまり適さないであろう.

7.3 微積分学における行列式

本節の最後に, 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, 解析系の講義に譲ろう.

- ヘッセ行列式

  C^2 級関数 $f:\mathbb{R}^2\to\mathbb{R}$ の極値判定において行列式が現れる. 1 変数関数のときもそうであったように, 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ない. 2 変数関数の極値判定も 1 変数の場合と同様に局所的に凸関数になるかどうかをテーラーの定理の 2 次の項を調べることでなされる. その具体的議論 (極値判定と 2 次形式の分類との関係) は省略するが, 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):

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

  の値によって判定される. C^2 級ゆえ $f_{xy}=f_{yx}$ に注意せよ. なお, 3 変数以上の場合は行列式の値のみでは判定できず, 2 次形式の分類を詳しく見る必要があり, そこではヘッセ行列の固有値[20]が

[20]固有値は後期の講義で扱う.

調べられる．ヘッセ行列式は, その性質から判別式 (discriminant) とも呼ばれている．

- 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

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

が与えられているとする．このとき, D 上の 2 変数関数 $f : D \to \mathbb{R}$ 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

$$\iint_D f(x,y)\,dxdy$$

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:

$$\iint_D f(x,y)\,dxdy = \iint_K f(\phi(u,v), \psi(u,v))|J(u,v)|\,dudv.$$

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 の訳語である．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 (クラメルの公式) を得るために考え出されたものである．一方, matrix という用語は行列式の理論の中で生まれたのちに, 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．

8　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となる置換について述べる．置換とはその名の通り置き変え, 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．

置換に関係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．しかしながら本論において置換は, 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

n 個の元からなる集合 X_n を考える．ここで, X の構成要素は何でも構わないものの, 記述の簡素化のため $X_n := \{ 1, 2, \cdots, n \}$ とする．

定義 8.1.1. X_n から X_n 自身への写像 $\sigma : X_n \to X_n$ において, $\sigma(1), \cdots, \sigma(n)$ 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 $\sigma(1), \sigma(2), \cdots, \sigma(n)$ 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 $\sigma : X_3 \to X_3$ 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

$\sigma(1) = k_1$, $\sigma(2) = k_2$, $\cdots$, $\sigma(n) = k_n$ なる置換 $\sigma : X_n \to X_n$ 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 $k_1, k_2, \cdots, k_n$ は文字列 $1, 2, \cdots, n$ 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 $\sigma : X_4 \to X_4$ 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:

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

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 $\sigma : X_4 \to X_4$ および $\tau : X_3 \to X_3$ 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 $\sigma : X_n \to X_n$ 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 $i = n+1, n+2, \cdots, m$ について $\sigma(i) = i$ で定められる置換 $\sigma : X_m \to X_m$ でもあると見なす.

8.3 置換の積

二つの置換 $\sigma, \tau : X_n \to X_n$ に対して, これらの合成写像 $\sigma \circ \tau : X_n \to X_n$ 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. 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 $\sigma, \tau, \upsilon : X_n \to X_n$ について $\sigma(\tau\upsilon) = (\sigma\tau)\upsilon$ が成り立つ．そこで, これらの積を略して $\sigma\tau\upsilon$ 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 節をみよ).

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, $\sigma\sigma^{-1} = \mathrm{id} = \sigma^{-1}\sigma$ 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 $\sigma^0 := \mathrm{id}$ 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 $\sigma^m = \sigma^1 = \sigma = \sigma \circ \mathrm{id} = \sigma \circ \sigma^0 = \sigma \circ \sigma^{m-1}$ と確認することができる．ゆえに $\sigma^{-1} \circ \sigma^m = \sigma^{-1} \circ (\sigma \circ \sigma^{m-1}) = (\sigma^{-1} \circ \sigma) \circ \sigma^{m-1} = \mathrm{id} \circ \sigma^{m-1} = \sigma^{m-1}$ となり, 求める等式を得る． □

8.4 巡回置換とその表示

定義 8.4.1. X_n の元のうち $k_1, k_2, \cdots, k_r$ 以外の元は動かさず, $\sigma(k_1) = k_2, \sigma(k_2) = k_3, \cdots, \sigma(k_{r-1}) = k_r$, $\sigma(k_r) = k_1$ 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 $(k_1, k_2, \cdots, k_r)$ 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:

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

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

練習 8.4.3. $\sigma : X_5 \to X_5$ 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[21]. 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, $1 \mapsto 4 \mapsto 2 \mapsto 1 \mapsto 4 \mapsto 2 \mapsto \cdots$ 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,

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

つまり $\sigma^\ell(p) = \sigma^\ell(1)$ であり, σ^ℓ が置換であることから $\sigma^\ell(1), \cdots, \sigma^\ell(n)$ 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 を繰り返し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 $3 \mapsto 6 \mapsto 5 \mapsto 7 \mapsto 3 \mapsto \cdots$ を得る. 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 回施すことで $\sigma^{-j}(a_i) \in A$ を得る. ところが, $3 = \sigma^{-j}(x) = \sigma^{-j}(a_i) \in A$ となってしまい, 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. □

[21] 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.

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

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:

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

ここで, 記号 $\overset{(k_i,k_j)}{\longleftrightarrow}$ 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, $\mathrm{sgn}(\sigma) := (-1)^m$ と定め, これを σ の**符号** (**signature** または **sign**) という. $\mathrm{sgn}(\sigma) = 1$ なる置換 σ を**偶置換** (**even permutation**) と呼び, $\mathrm{sgn}(\sigma) = -1$ 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: $\mathrm{id}_{X_n} = (1,2)(1,2)$.

次の例 (練習 8.4.3 と同じ置換) からも分かるように, 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, 次節にまわそう.

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

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

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.**

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

練習 **8.5.5.** 次を示せ.

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

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

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

$(\Leftarrow)$: 逆置換の定義より明らかである. □

(2) 置換 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. このとき,

$$\tau := (p_\ell, q_\ell)\cdots(p_2, q_2)(p_1, q_1)$$

とおけば, $\tau = \sigma^{-1}$ であることが次の計算により確かめられる:

$$\begin{aligned}\tau\sigma &= (p_\ell, q_\ell)\cdots(p_2, q_2)\underline{(p_1, q_1)\cdot(p_1, q_1)}(p_2, q_2)\cdots(p_\ell, q_\ell) \\ &= (p_\ell, q_\ell)\cdots(p_2, q_2)\cdot\underline{\mathrm{id}}\cdot(p_2, q_2)\cdots(p_\ell, q_\ell) \\ &= (p_\ell, q_\ell)\cdots(p_2, q_2)\cdot(p_2, q_2)\cdots(p_\ell, q_\ell) = \cdots = (p_\ell, q_\ell)(p_\ell, q_\ell) = \mathrm{id}.\end{aligned}$$

ゆえに (1) より $\tau = \sigma^{-1}$ である. □

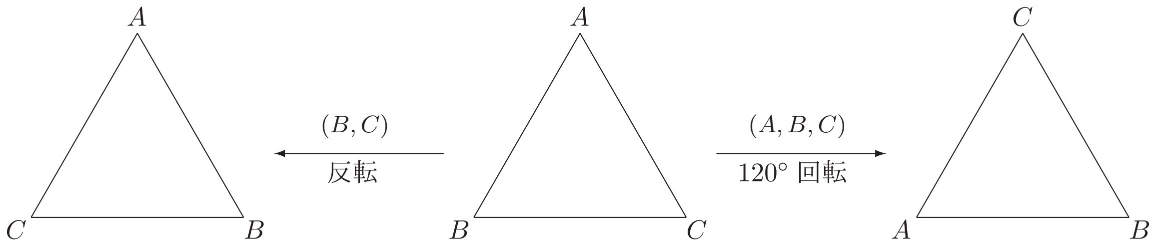
発展 (対称性と群)

本節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 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．

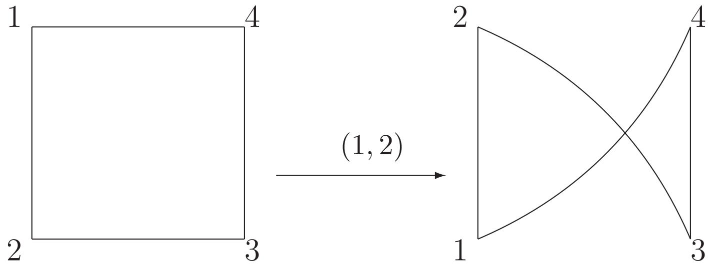


他に考えられる操作は 240° 回転に相当する巡回置換 (A,C,B), 360° 回転 (結果的にこれは頂点を全く動かさない操作 id に等しい), B,C のいずれかを固定した反転に相当する互換 (A,C) および (A,B) であり, 次の計 6 つである．

$$\{\,\mathrm{id},\ (B,C),\ (A,B),\ (A,B,C),\ (A,C,B),\ (A,C)\,\}$$

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, $\mathfrak{S}_4$ の元のうち次の計 8 つに数え上げられる:

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

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だろうか．式 x^2+xy+y^2+zw の対称性は $\{\,\mathrm{id},\ (x,y),\ (z,w),\ (x,y)(z,w)\,\}$ で与えられ, ジャンケンの対称性は $\{\,\mathrm{id},\ (1,2,3),\ (1,3,2)\,\}$ 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 の各元の間に演算 (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.

9 置換の符号について

本節では, まずはじめに前節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節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, 本節の 9.1 項は不要である. また, 9.2 項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も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:

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

例 9.1.1. (1) $\sigma = (1,2)(3,4) \in \mathfrak{S}_n$ 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 $\mathrm{id} \in \mathfrak{S}_n$ において, $s(\mathrm{id}) = \displaystyle\prod_{1 \le i < j \le n} \frac{j-i}{j-i} = 1$.

$s(\sigma)$ が符号を意味することをこれから見ていく[22]. 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. $k < \ell$ 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 $\ell - k$ が現れている. また, 分子においても, $\sigma^{-1}(\ell)$ と $\sigma^{-1}(k)$ のどちらが大きいかは分からないが, $\sigma(\sigma^{-1}(\ell)) - \sigma(\sigma^{-1}(k)) = \ell - k$ あるいは $\sigma(\sigma^{-1}(k)) - \sigma(\sigma^{-1}(\ell)) = k - \ell$ のうちいずれか一方が 1 回だけ現れる. したがって, 分子 ($\ell - k$ または $k - \ell$) と分母 ($\ell - k$) 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. □

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 $s(\tau\sigma) = s(\tau)s(\sigma)$.

[22] したがって, $s(\sigma)$ でもって置換の符号を定義してもよい.

Proof. 次のように計算できる.

$$s(\tau\sigma) = \prod_{1\le i<j\le n} \frac{\tau\sigma(j)-\tau\sigma(i)}{j-i} = \left(\prod_{1\le i<j\le n} \frac{\tau\sigma(j)-\tau\sigma(i)}{\sigma(j)-\sigma(i)}\right)\left(\prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i}\right)$$

$$= \left(\prod_{\substack{1\le i<j\le n,\\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \prod_{\substack{1\le i<j\le n,\\ \sigma(j)<\sigma(i)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)}\right)\cdot s(\sigma)$$

$$= \left(\prod_{\substack{1\le i<j\le n,\\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \prod_{\substack{1\le i<j\le n,\\ \sigma(j)<\sigma(i)}} \frac{\tau(\sigma(i))-\tau(\sigma(j))}{\sigma(i)-\sigma(j)}\right)\cdot s(\sigma)$$

↑二つ目の $\prod$ において, 分母・分子それぞれに -1 を掛けた

$$= \left(\prod_{1\le k<\ell\le n} \frac{\tau(\ell)-\tau(k)}{\ell-k}\right)\cdot s(\sigma) = s(\tau)s(\sigma).$$

↑この等式については後述

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. □

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

補題 9.1.4. $1\le k<\ell\le n$ 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 $k<j\le n$ かつ $j\neq\ell$,
- $i=\ell$ であり $j\neq k$ の場合, このとき $\ell<j\le n$,
- $j=k$ であり $i\neq\ell$ の場合, このとき $1\le i<k$,
- $j=\ell$ であり $i\neq k$ の場合, このとき $1\le i<\ell$ 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:

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

↑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 $\sigma_1, \cdots, \sigma_k$ 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 $\mathrm{sgn}(\sigma) = s(\sigma)$.

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 に対応する並べ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, $\cdots$, 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 $l_1 \cdots l_n$ を文字列 $\sigma(l_1)\cdots\sigma(l_n)$ 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, $\cdots$, 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 $l_4l_1l_6l_2l_7l_5l_3$ を文字列 $l_1l_2l_3l_4l_5l_6l_7$ 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 $l_{\sigma(1)}\cdots l_{\sigma(n)}$ を文字列 $l_1\cdots l_n$ に並べ替える操作, あるいは文字列 $l_1\cdots l_n$ を文字列 $l_{\sigma^{-1}(1)}\cdots l_{\sigma^{-1}(n)}$ 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 $123\cdots n$ 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 $k_1k_2k_3\cdots k_n$ 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 $\cdots$ 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.

$$
\begin{aligned}
(\natural)\quad 1234567 &\overset{(3,4)}{\longleftrightarrow} 1243567 \overset{(2,4)}{\longleftrightarrow} 1423567 \overset{(1,4)}{\longleftrightarrow} 4123567 \overset{(5,6)}{\longleftrightarrow} 4123657 \\
&\overset{(3,6)}{\longleftrightarrow} 4126357 \overset{(2,6)}{\longleftrightarrow} 4162357 \overset{(5,7)}{\longleftrightarrow} 4162375 \overset{(3,7)}{\longleftrightarrow} 4162735 \overset{(3,5)}{\longleftrightarrow} 4162753.
\end{aligned}
$$

ここで, 記号 $\overset{(i,j)}{\longleftrightarrow}$ 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, $k_{i+1},\cdots,k_n$ の中にある 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 $\mathrm{inv}(k_1,\cdots,k_n)$ で与えられる:

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

例えば $\mathrm{inv}(4,1,6,2,7,5,3) = 3+0+3+0+2+1 = 9$ であり, 確かに σ は 9 つの互換の積に分解されている. $\mathrm{inv}(k_1,\cdots,k_n)$ 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.

$$
\begin{aligned}
(\sharp)\quad 4162753 &\overset{(6,7)}{\longleftrightarrow} 4162735 \overset{(5,6)}{\longleftrightarrow} 4162375 \overset{(6,7)}{\longleftrightarrow} 4162357 \overset{(3,4)}{\longleftrightarrow} 4126357 \\
&\overset{(4,5)}{\longleftrightarrow} 4123657 \overset{(5,6)}{\longleftrightarrow} 4123567 \overset{(1,2)}{\longleftrightarrow} 1423567 \overset{(2,3)}{\longleftrightarrow} 1243567 \overset{(3,4)}{\longleftrightarrow} 1234567.
\end{aligned}
$$

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 $(\sharp)$ における記号 $\overset{(i,j)}{\longleftrightarrow}$ 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 $\mathrm{inv}(k_1, \cdots, k_n)$ 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:

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

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 $\mathrm{sgn}(\sigma)$ 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 $\mathrm{id} \in \mathfrak{S}_n$ において, $\tilde{s}(\mathrm{id}) = (-1)^{\mathrm{inv}(1, \cdots, n)} = (-1)^0 = 1$

(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 $Z = 1, \cdots, i-1, j, i+1, \cdots, j-1, i, j+1, \cdots, n$ 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 $\mathrm{inv}(Z) = (j-i) + (j-i-1) = 2(j-i) - 1$. 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 $X = k_1, \cdots, k_n$ と $Y = k_1, \cdots, k_{i-1}, k_j, k_{i+1}, \cdots, k_{j-1}, k_i, k_{j+1}, \cdots, k_n$ における転倒数の差 $\mathrm{inv}(Y) - \mathrm{inv}(X)$ 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, $\mathrm{inv}(X)$ および $\mathrm{inv}(Y)$ 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 $\mathrm{inv}(Y)-\mathrm{inv}(X)$ にあたるから

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

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 $\mathrm{inv}(Y)-\mathrm{inv}(X)$ は奇数となる. したがって,

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

□

上の定理は $\mathrm{inv}(X)$ と $\mathrm{inv}(Y)$ 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 $\mathrm{inv}(Y)-\mathrm{inv}(X)$ 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, $\mathrm{inv}(X)$ や $\mathrm{inv}(Y)$ 自身, すなわち絶対的な量よりも, これらの間の関係 $\mathrm{inv}(Y)-\mathrm{inv}(X)$, 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 $\mathrm{inv}(k_1, \cdots, k_n)$ 個の互換の積で書ける. $\ell = \mathrm{inv}(k_1, \cdots, k_n)$, $\tau = (p_1, q_1)\cdots(p_\ell, q_\ell)$ と置き, 補題 9.3.2 を ℓ 回適用すると,

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

□

10　行列式の定義と性質

いよいよ行列式の定義に入ろう. 本節では, 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かどうか, 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節まで待ってほしい. 11.3 項まで読めば, 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が場合について, 行列式を実際に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. $\mathfrak{S}_1 = \{\mathrm{id}\}$ および $\operatorname{sgn}(\mathrm{id}) = 1$ より,

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

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

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

(3) $n = 3$ の場合. 例 8.5.4 より

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

である. したがって,

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

本節および次節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 節で行おう.

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.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$(-)$ $(+)$

3 次の場合はやや複雑になるが, 次のように分けよう.

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

上の図式をもとに次の展開式を得る:

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

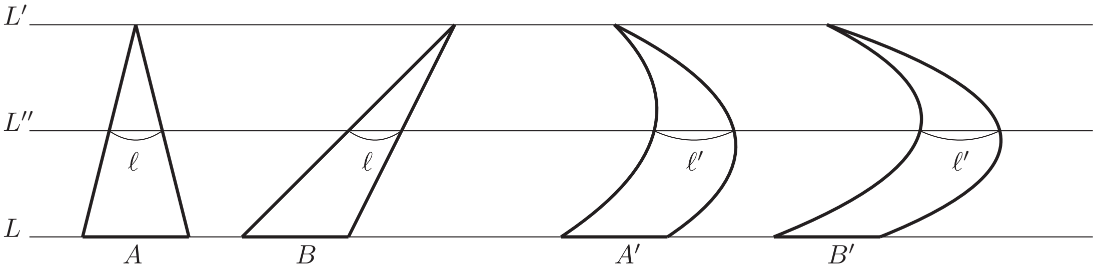
なお, 4 次以上の行列式には, この様な方法は使えない.

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, 本論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うものの, 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. しかし残念なことに, これに答えるのは容易ではない. 何故なら, この問いに答えるには, 論理的な曖昧さを排除するために, そもそも n 次元の図形およびその体積とは何か, という図形や体積の定義について論じる必要が生じるからである. 体積の議論が非常に面倒なものであることは, 微積分学において定積分の定義に一苦労したことを思い返せば容易に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を一つ挙げておこう. 例えば, 二つの連続関数のグラフ $f, g : [a, b] \to \mathbb{R}$ 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. また, 高次元の図形については, 重積分と累次積分 (逐次積分) の一致がカヴァリエリの原理を示唆している.

ところで, 今の例は図形の体積が重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. 実のところ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, 積分を別の視点から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. ルベーグ積分に対して, 大学初年次の微積分で学ぶ定積分はリーマン積分と呼ばれている. 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:

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

(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:

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

(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.

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

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 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ 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うえで, 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

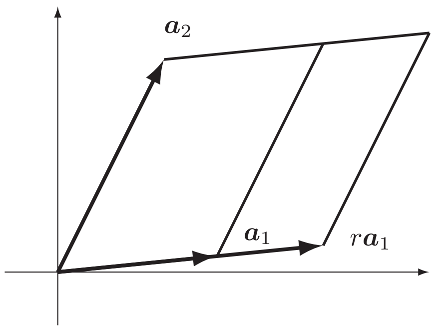


図 2: ベクトルのスカラー倍

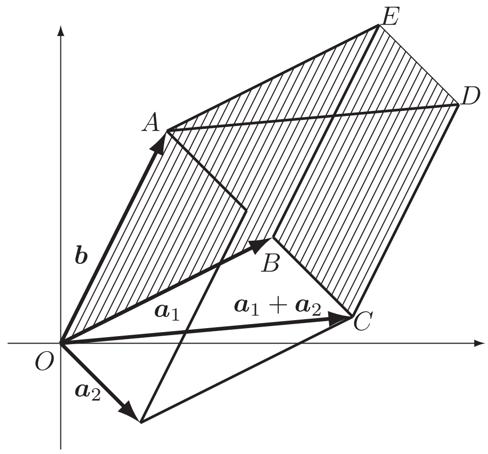


図 3: ベクトルの和

ルを r 倍すれば面積が r 倍されることは明らかである (図 2). 和については, 図 3 を見よ. $\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b}$ で張られる平行四辺形 $AOCD$ と六角形 $AOBCDE$ の面積は等しい. これは, $\boldsymbol{b}$ 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 $\boldsymbol{b}$ 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 $AOBCDE$ の面積は平行四辺形 $AOBE$ の面積 $\det(\boldsymbol{a}_1, \boldsymbol{b})$ と平行四辺形 $EBCD$ の面積の和に等しい. 平行四辺形 $EBCD$ は $\boldsymbol{a}_2, \boldsymbol{b}$ で張られる平行四辺形を $\boldsymbol{a}_1$ 方向に平行移動したものであるから, その面積は $\det(\boldsymbol{a}_2, \boldsymbol{b})$ である. 以上より, $\det(\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b}) = \det(\boldsymbol{a}_1, \boldsymbol{b}) + \det(\boldsymbol{a}_2, \boldsymbol{b})$ 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 節にまわそう.

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 $\boldsymbol{x}_k$ $(k = 1, \cdots, n)$ を n 次列ベクトルとするとき, $i < j$ について,

(1) $\det[\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_j, \boldsymbol{x}_{i+1}, \cdots, \boldsymbol{x}_{j-1}, \boldsymbol{x}_i, \boldsymbol{x}_{j+1}, \cdots, \boldsymbol{x}_n] = -\det[\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n]$,

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

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 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ 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**交代的** (**alternating**) であると呼ばる．一般に, 交代性は歪対称性から直ちに導くことができる[23]:

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 (または行) の何倍かを別の行 (または列) に加えても行列式は変わらない．すなわち, 各 $\boldsymbol{x}_k$ $(k = 1, \cdots, n)$ を n 次列ベクトルとするとき, $i \neq j$ について

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

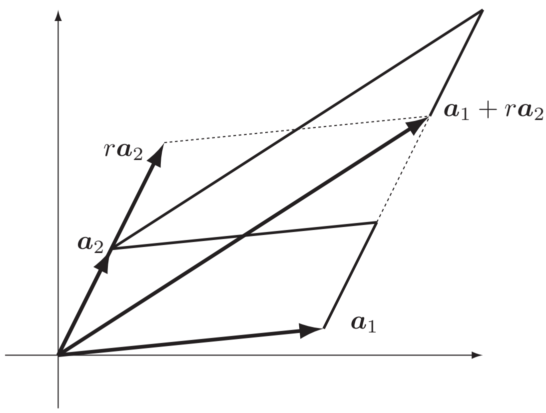
(2) $\det {}^t[\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{j-1}, \boldsymbol{x}_j + r\boldsymbol{x}_i, \boldsymbol{x}_{j+1}, \cdots, \boldsymbol{x}_n] = \det {}^t[\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n]$.

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

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

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 $\det(\boldsymbol{a}_1 + r\boldsymbol{a}_2, \boldsymbol{a}_2) = \det(\boldsymbol{a}_1, \boldsymbol{a}_2)$ 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．



[23] 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).

11　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るのではないだろうか. そして, この実感が妥当であるゆえんを 11.3 項において解説する.

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. 証明は次節にまわす (定理 12.5.1).

命題 11.1.1.

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

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 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ 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,

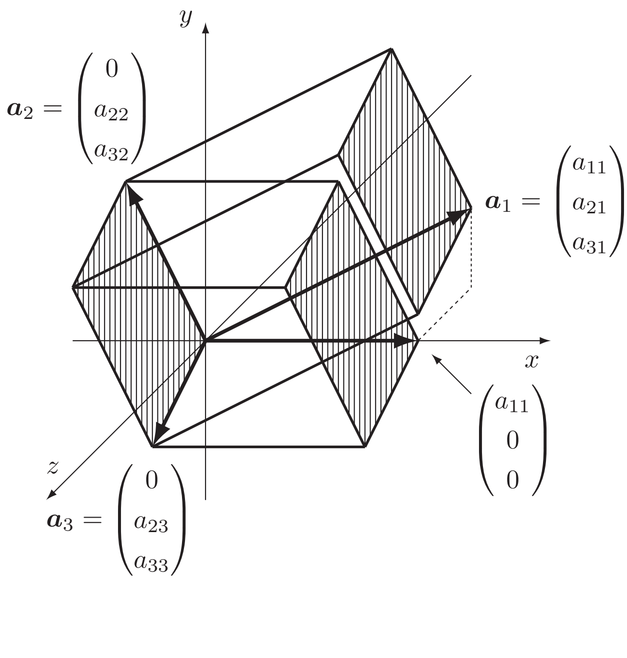
$$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \cdots, \boldsymbol{a}_n = \begin{bmatrix} 0 \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} \quad \text{としよう.}$$

このとき, $n-1$ 個のベクトル $\boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 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,

$$\boldsymbol{a}'_2 = \begin{bmatrix} a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \cdots, \boldsymbol{a}'_n = \begin{bmatrix} a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} \text{ で張られる } \mathbb{R}^{n-1}$$

上の図形と合同であり, ゆえにその $n-1$ 次元体積は $\det[\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_n]$ である. したがって, 図形 D は底面積 $\det[\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_n]$ および高さ a_{11} なる角柱であり, ゆえに体積 $V(D)$ は $a_{11} \det[\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_n]$ に一致する.

$\boldsymbol{a}_1$ 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 $\boldsymbol{a}_1$ 方向に歪ませた形になる. 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, $V(D) = a_{11} \det[\boldsymbol{a}'_2, \cdots, \boldsymbol{a}'_n]$ である.



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:

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

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.

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

(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)**.**

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

Proof. $|A|=|{}^tA|$ 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 $\begin{bmatrix}1&x_1\\1&x_2\end{bmatrix}=x_2-x_1$ 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.

$$\begin{bmatrix}1&1&\dots&1\\x_1&x_2&\dots&x_n\\{x_1}^2&{x_2}^2&\dots&{x_n}^2\\\vdots&\vdots&\vdots&\vdots\\{x_1}^{n-1}&{x_2}^{n-1}&\dots&{x_n}^{n-1}\end{bmatrix}=\begin{bmatrix}1&1&\dots&1\\0&x_2-x_1&\dots&x_n-x_1\\0&{x_2}^2-x_1x_2&\dots&{x_n}^2-x_1x_n\\\vdots&\vdots&\vdots&\vdots\\0&{x_2}^{n-1}-x_1{x_2}^{n-2}&\dots&{x_n}^{n-1}-x_1{x_n}^{n-2}\end{bmatrix}\begin{matrix}\\1\text{ 行}\times x_1\text{ を引く}\\2\text{ 行}\times x_1\text{ を引く}\\\vdots\\\text{n-1 行}\times x_1\text{ を引く}\end{matrix}$$

$$= \begin{bmatrix} x_2 - x_1 & x_3 - x_1 & \dots & x_n - x_1 \\ x_2(x_2 - x_1) & x_3(x_3 - x_1) & \dots & x_n(x_n - x_1) \\ \vdots & \vdots & \vdots & \vdots \\ {x_2}^{n-2}(x_2 - x_1) & {x_3}^{n-2}(x_3 - x_1) & \dots & {x_n}^{n-2}(x_n - x_1) \end{bmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)\cdots(x_n - x_1) \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_2 & x_3 & \dots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ {x_2}^{n-2} & {x_3}^{n-2} & \dots & {x_n}^{n-2} \end{bmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)\cdots(x_n - x_1) \prod_{2 \le i < j \le n} (x_j - x_i) \qquad \text{(ここで帰納法の仮定を用いた)}$$

$$= \prod_{1 \le i < j \le n} (x_j - x_i).$$

□

11.3 体積との関係

前節と本節を通して, 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 $1^n = 1$ であると考えるだろう. 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[24]. 次の性質すべてを満たす写像 $D : M_n(\mathbb{R}) \to \mathbb{R}$ を取れば, D は行列式に一致する:

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

(1) 多重線形性：

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

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

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

(3) 正規化：$D(E) = 1$.

この定理の証明も次節で与えよう (定理 12.4.1).

[24] $M_n(\mathbb{R})$ 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.

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:

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

ここで, P_n は n 個の文字 $1,\cdots,n$ 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 $D : M_n(\mathbb{R}) \to \mathbb{R}$ 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.

12　行列式の性質（証明）

10 および 11 節で述べた命題の証明を本節で与える．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 項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, $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$.

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

補足. 上の補題は練習 8.5.5(2) からも導くことができる.

補題 12.1.2. 写像 $I : \mathfrak{S}_n \to \mathfrak{S}_n$ を $I(\sigma) := \sigma^{-1}$ と定めれば, I 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, $I(\sigma)$ も重複なく $\mathfrak{S}_n$ 全体を動く.

Proof. 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 は**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, これを**全射** (**surjective**) であるという．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.

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 $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$ (補題 12.1.1) から,

$$|{}^tA| = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \tag{12.1.1}$$

$$= \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma^{-1}) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}. \tag{12.1.2}$$

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 $\sigma(1), \cdots, \sigma(n)$ を小さい順に並べ替える:

$$\sigma(k_1) = 1,\ \sigma(k_2) = 2,\ \cdots,\ \sigma(k_n) = n.$$

つまり,

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

となり, ゆえに各 $i = 1, \cdots, n$ 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:

$$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{\sigma(k_1)k_1} a_{\sigma(k_2)k_2} \cdots a_{\sigma(k_n)k_n} = a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)}. \tag{12.1.3}$$

写像 $I : \mathfrak{S}_n \to \mathfrak{S}_n$ 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,

$$\begin{aligned}|{}^tA| &= \sum_{\sigma\in\mathfrak{S}_n} \mathrm{sgn}(\sigma^{-1}) a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)} \\ &= \sum_{\sigma\in\mathfrak{S}_n} \mathrm{sgn}(I(\sigma)) a_{1,I(\sigma)(1)} a_{2,I(\sigma)(2)} \cdots a_{n,I(\sigma)(n)}.\end{aligned}$$

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:

$$\sum_{\tau\in\mathfrak{S}_n} \mathrm{sgn}(\tau) a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)} = |A|.$$

□

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 $F : \mathfrak{S}_n \to \mathfrak{S}_n$ 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

$$\begin{aligned}\sigma_1\tau\tau^{-1} &= \sigma_2\tau\tau^{-1} \\ \sigma_1\,\mathrm{id} &= \sigma_2\,\mathrm{id} \\ \sigma_1 &= \sigma_2.\end{aligned}$$

(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, $F(\sigma) = F(\varphi\tau^{-1}) = \varphi\tau\tau^{-1} = \varphi\,\mathrm{id} = \varphi$. □

上の補題において互換 $\tau = (p, q)$ (ただし $p \neq q$) を考えると, $F : \mathfrak{S}_n \to \mathfrak{S}_n$ 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 $\dfrac{n!}{2}$ 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 $j = 1, \cdots, n$ 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, $F : \mathfrak{S}_n \to \mathfrak{S}_n$ 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,

$$(\mathrm{i})\ k \neq p, q \implies F(\sigma)(k) = \sigma(k), \quad (\mathrm{ii})\ F(\sigma)(p) = \sigma(q), \quad (\mathrm{iii})\ F(\sigma)(q) = \sigma(p).$$

また, $\mathrm{sgn}(F(\sigma)) = \mathrm{sgn}(\sigma(p,q)) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}((p,q)) = -\mathrm{sgn}(\sigma)$ である．これらを用いると

$$\begin{aligned}
|B| &= \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) b_{1\sigma(1)} \cdots b_{p\sigma(p)} \cdots b_{q\sigma(q)} \cdots b_{n\sigma(n)} \\
&= \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{q\sigma(p)} \cdots a_{p\sigma(q)} \cdots a_{n\sigma(n)} \\
&= \sum_{\sigma \in \mathfrak{S}_n} \big(-\mathrm{sgn}(F(\sigma)) \big) a_{1,F(\sigma)(1)} \cdots a_{q,F(\sigma)(q)} \cdots a_{p,F(\sigma)(p)} \cdots a_{n,F(\sigma)(n)} \\
&= -\sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(F(\sigma)) a_{1,F(\sigma)(1)} \cdots a_{p,F(\sigma)(p)} \cdots a_{q,F(\sigma)(q)} \cdots a_{n,F(\sigma)(n)}.
\end{aligned}$$

$\uparrow$ 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:

$$-\sum_{\varphi \in \mathfrak{S}_n} \mathrm{sgn}(\varphi) a_{1\varphi(1)} \cdots a_{p\varphi(p)} \cdots a_{q\varphi(q)} \cdots a_{n\varphi(n)} = -|A|.$$

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． □

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．本項および次項では写像 $F : M_n(\mathbb{R}) \to \mathbb{R}$ 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 $X = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n]$ 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 $F(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)$ 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 $i = 1, \cdots, n$ について,
  (i) $F(\boldsymbol{x}_1, \cdots, r\boldsymbol{x}, \cdots, x_n) = rF(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_i, \cdots, \boldsymbol{x}_n)$,
  (ii) $F(\boldsymbol{x}_1, \cdots, \boldsymbol{y}_i + \boldsymbol{z}_i, \cdots, \boldsymbol{x}_n) = F(\boldsymbol{x}_1, \cdots, \boldsymbol{y}_i, \cdots, \boldsymbol{x}_n) + F(\boldsymbol{x}_1, \cdots, \boldsymbol{z}_i, \cdots, \boldsymbol{x}_n)$.
- 歪対称性: 各 $i < j$ について,
  $$F(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_j, \boldsymbol{x}_{i+1}, \cdots, \boldsymbol{x}_{j-1}, \boldsymbol{x}_i, \boldsymbol{x}_{j+1}, \cdots, \boldsymbol{x}_n) = -F(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n).$$

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:

多重線形性 (2): $$F\left(\boldsymbol{a}_1, \cdots, \sum_{k=1}^{\ell} r_k \boldsymbol{b}_k, \cdots, \boldsymbol{a}_n\right) = \sum_{k=1}^{\ell} r_k F(\boldsymbol{a}_1, \cdots, \boldsymbol{b}_k, \cdots, \boldsymbol{a}_n).$$

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 (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.

$$
\begin{aligned}
&F\left(\boldsymbol{a}_1, \cdots, \sum_{k=1}^{\ell+1} r_k \boldsymbol{b}_k, \cdots, \boldsymbol{a}_n\right) \\
&= F\left(\boldsymbol{a}_1, \cdots, \left(\sum_{k=1}^{\ell} r_k \boldsymbol{b}_k\right) + r_{\ell+1}\boldsymbol{b}_{\ell+1}, \cdots, \boldsymbol{a}_n\right) \\
&= F\left(\boldsymbol{a}_1, \cdots, \sum_{k=1}^{\ell} r_k \boldsymbol{b}_k, \cdots, \boldsymbol{a}_n\right) + F\left(\boldsymbol{a}_1, \cdots, r_{\ell+1}\boldsymbol{b}_{\ell+1}, \cdots, \boldsymbol{a}_n\right) && \text{(多重線形性 (i))} \\
&= \sum_{k=1}^{\ell} r_k F(\boldsymbol{a}_1, \cdots, \boldsymbol{b}_k, \cdots, \boldsymbol{a}_n) + r_{\ell+1} F\left(\boldsymbol{a}_1, \cdots, \boldsymbol{b}_{\ell+1}, \cdots, \boldsymbol{a}_n\right) && \text{(帰納法の仮定と多重線形性 (ii))} \\
&= \sum_{k=1}^{\ell+1} r_k F(\boldsymbol{a}_1, \cdots, \boldsymbol{b}_k, \cdots, \boldsymbol{a}_n).
\end{aligned}
$$

□

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, $\boldsymbol{a}_i = \boldsymbol{a}_j$ (ただし $i < j$) ならば

$$F(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_j, \cdots, , \boldsymbol{a}_n) = 0.$$

歪対称性を用いて列の入れ替えを何度か行うと, 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

$$F(\boldsymbol{x}_{\sigma(1)}, \cdots, \boldsymbol{x}_{\sigma(n)}) = \mathrm{sgn}(\sigma) F(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n).$$

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, $F(x_{\sigma(1)}, \cdots, x_{\sigma(n)}) = \mathrm{sgn}(\sigma) F(x_1, \cdots, x_n)$ 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 $F(x_{\tau(1)}, \cdots, x_{\tau(n)}) = \mathrm{sgn}(\tau) F(x_1, \cdots, x_n)$ 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 $F(x_{\sigma(1)}, \cdots, x_{\sigma(n)}) = \mathrm{sgn}(\sigma) F(x_1, \cdots, x_n)$ 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

$$
\begin{aligned}
F(x_{\sigma(1)}, \cdots, x_{\sigma(n)}) &= F(x_{\tau(i,j)(1)}, \cdots, x_{\tau(i,j)(i)}, \cdots, x_{\tau(i,j)(j)}, \cdots, x_{\tau(i,j)(n)}) \\
&= F(x_{\tau(1)}, \cdots, x_{\tau(j)}, \cdots, x_{\tau(i)}, \cdots, x_{\tau(n)}) \\
&= -F(x_{\tau(1)}, \cdots, x_{\tau(i)}, \cdots, x_{\tau(j)}, \cdots, x_{\tau(n)}) && \text{(歪対称性を用いた)} \\
&= -\mathrm{sgn}(\tau) F(x_1, \cdots, x_n) = \mathrm{sgn}(\sigma) f(x_1, \cdots, x_n). && \text{(帰納法の仮定を用いた)}
\end{aligned}
$$

□

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 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ 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

$$
\boldsymbol{e}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \boldsymbol{e}_n := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \qquad \in \mathbb{R}^n
$$

を与えておく．このとき, 各 $i = 1, \cdots, n$ において

$$\boldsymbol{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \vdots \\ a_{ni} \end{bmatrix} = \begin{bmatrix} a_{1i} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{2i} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_{ni} \end{bmatrix} = a_{1i}\boldsymbol{e}_1 + \cdots + a_{ni}\boldsymbol{e}_n = \sum_{k_i=1}^{n} a_{k_i i}\boldsymbol{e}_{k_i} \tag{12.4.1}$$

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 $F : M_n(\mathbb{R}) \to \mathbb{R}$ 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, $A = [a_{ij}] = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ 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 (2) を各列に対して適用しながら展開すると,

$$\begin{aligned}
F(A) &= F(\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n) = F\Big(\sum_{k_1=1}^{n} a_{k_1 1}\boldsymbol{e}_{k_1}, \sum_{k_2=1}^{n} a_{k_2 2}\boldsymbol{e}_{k_2}, \cdots, \sum_{k_n=1}^{n} a_{k_n n}\boldsymbol{e}_{k_n}\Big) \\
&= \sum_{k_1=1}^{n} a_{k_1 1} F\Big(\boldsymbol{e}_{k_1}, \sum_{k_2=1}^{n} a_{k_2 2}\boldsymbol{e}_{k_2}, \cdots, \sum_{k_n=1}^{n} a_{k_n n}\boldsymbol{e}_{k_n}\Big) \\
&= \sum_{k_1=1}^{n} \left(a_{k_1 1} \sum_{k_2=1}^{n} a_{k_2 2} F\Big(\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \sum_{k_n=1}^{n} a_{k_n n}\boldsymbol{e}_{k_n}\Big) \right) \\
&= \cdots = \sum_{k_1=1}^{n} a_{k_1 1} \left(\sum_{k_2=1}^{n} a_{k_2 2} \left(\cdots \left(\sum_{k_{n-1}=1}^{n} a_{k_{n-1} n-1} \left(\sum_{k_n=1}^{n} a_{k_n n} F(\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \boldsymbol{e}_{k_n}) \right) \right) \cdots \right) \right) \\
&= \sum_{k_1=1}^{n} \left(\sum_{k_2=1}^{n} a_{k_1 1} a_{k_2 2} \left(\cdots \left(\sum_{k_{n-1}=1}^{n} a_{k_{n-1} n-1} \left(\sum_{k_n=1}^{n} a_{k_n n} F(\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \boldsymbol{e}_{k_n}) \right) \right) \cdots \right) \right) \\
&= \cdots = \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_{n-1}=1}^{n} \sum_{k_n=1}^{n} a_{k_1 1} a_{k_2 2} \cdots a_{k_{n-1} n-1} a_{k_n n} F(\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \boldsymbol{e}_{k_n}) \\
&= \sum_{1 \le k_1, k_2, \cdots . k_n \le n} a_{k_1 1} a_{k_2 2} \cdots a_{k_n n} F(\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \boldsymbol{e}_{k_n}).
\end{aligned}$$

いま得られた総和の各項において, $\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \boldsymbol{e}_{k_n}$ 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 $F(\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \boldsymbol{e}_{k_n}) = 0$ 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, $\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \boldsymbol{e}_{k_n}$ はいずれも異なる単位列ベクトルであり, とくに $k_1, k_2, \cdots, k_n$ 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

$$\begin{aligned}
a_{k_1 1} a_{k_2 2} \cdots a_{k_n n} F(\boldsymbol{e}_{k_1}, \boldsymbol{e}_{k_2}, \cdots, \boldsymbol{e}_{k_n}) &= a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} F(\boldsymbol{e}_{\sigma(1)}, \boldsymbol{e}_{\sigma(2)}, \cdots, \boldsymbol{e}_{\sigma(k_n)}) \\
&= a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \operatorname{sgn}(\sigma) F(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n) \\
&= \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} F(E).
\end{aligned}$$

以上より,

$$\begin{aligned}
F(A) &= \sum_{\sigma \in \mathfrak{S}_n} \big(\operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} F(E) \big) \\
&= \left(\sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \right) F(E) = |{}^t A| \cdot F(E) = |A| \cdot F(E).
\end{aligned}$$

補足. 上の二行目にある括弧内の式 $\sum_{\sigma\in\mathfrak{S}_n}\mathrm{sgn}(\sigma)a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}$ が $|{}^tA|$ に等しいことは式 (12.1.1) による. □

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

$$|Y| = \sum_{\sigma\in\mathfrak{S}_n}\mathrm{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{r\sigma(r)}a_{r+1\sigma(r+1)}\cdots a_{n\sigma(n)} \tag{12.5.1}$$

である. ここで, 総和に現れる各項

$$\mathrm{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{r\sigma(r)}a_{r+1\sigma(r+1)}\cdots a_{n\sigma(n)} \tag{12.5.2}$$

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, $1\le i\le r$ かつ $r+1\le j\le n$ ならば $a_{ij}=0$ である. よって $\{\sigma(1),\cdots,\sigma(r)\}$ 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,

$$\{\sigma(1),\cdots,\sigma(r)\}=\{1,\cdots,r\}$$

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

$$\{\sigma(r+1),\cdots,\sigma(n)\}=\{r+1,\cdots,n\}$$

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 $i=r+1,\cdots,n$ があれば Y の $(i,\sigma(i))$-成分は E_s において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 $i=r+1,\cdots,n$ 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:

$$\sigma=\begin{pmatrix} 1 & \cdots & r & r+1 & \cdots & n \\ k_1 & \cdots & k_r & r+1 & \cdots & n \end{pmatrix},\qquad \text{ただし } k_1,\cdots,k_r \text{ は } r \text{ 以下の数による順列.}$$

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,

$$\begin{aligned}\begin{vmatrix} A & O_{r,s} \\ * & E \end{vmatrix} = |Y| &= \sum_{\sigma\in\mathfrak{S}_r}\mathrm{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{r\sigma(r)}a_{r+1,r+1}\cdots a_{n,n} \\ &= \sum_{\sigma\in\mathfrak{S}_r}\mathrm{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{r\sigma(r)}\cdot 1\cdots 1 = |A|.\end{aligned}$$

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 $X=[\boldsymbol{x}_1,\cdots\boldsymbol{x}_s]$ たち全体を定義域とする関数 $F:M_s(\mathbb{R})\to\mathbb{R}$ を

$$F(\boldsymbol{x}_1,\cdots\boldsymbol{x}_s) := \begin{vmatrix} A & O_{r,s} \\ * & X \end{vmatrix}$$

と定めよう. すると, 行列式の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, これまでの議論より $F(E)=|Y|=|A|$ 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. □

例 12.5.2. $A_1, \cdots, A_n$ を正方行列とすると (各々のサイズは異なっていても構わない),

$$\begin{vmatrix} A_1 & & & \\ & A_2 & * & \\ & & \ddots & \\ \Large O & & & A_n \end{vmatrix} = |A_1| \cdot \begin{vmatrix} A_2 & & & \\ & A_3 & * & \\ & & \ddots & \\ \Large O & & & A_n \end{vmatrix} = \cdots = |A_1| \cdot |A_2| \cdots |A_n|.$$

12.6 $|AB| = |A| \cdot |B|$ の証明

一般に, (m, n)-行列 A および (n, ℓ)-行列 $B = [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_\ell]$ に対して, $AB = [A\boldsymbol{b}_1, \cdots, A\boldsymbol{b}_\ell]$ である (例 3.5.1). この事実に注意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 $X = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n]$ たち全体を定義域とする写像 $F : M_n(\mathbb{R}) \to \mathbb{R}$ を $F(\boldsymbol{x}_1, \cdots \boldsymbol{x}_n) := |AX|$ と定める. このとき,

$$F(\boldsymbol{x}_1, \cdots \boldsymbol{x}_n) = |A[\boldsymbol{x}_1, \cdots \boldsymbol{x}_n]| = |A\boldsymbol{x}_1, \cdots, A\boldsymbol{x}_n|$$

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:

$$\begin{aligned} F(\boldsymbol{x}_1, \cdots, \boldsymbol{b}_i + \boldsymbol{c}_i, \cdots, \boldsymbol{x}_n) &= |A\boldsymbol{x}_1, \cdots, A(\boldsymbol{b}_i + \boldsymbol{c}_i), \cdots, A\boldsymbol{x}_n| = |A\boldsymbol{x}_1, \cdots, A\boldsymbol{b}_i + A\boldsymbol{c}_i, \cdots, A\boldsymbol{x}_n| \\ &= |A\boldsymbol{x}_1, \cdots, A\boldsymbol{b}_i, \cdots, A\boldsymbol{x}_n| + |A\boldsymbol{x}_1, \cdots, A\boldsymbol{c}_i, \cdots, A\boldsymbol{x}_n| \\ &= F(\boldsymbol{x}_1, \cdots, \boldsymbol{b}_i, \cdots, \boldsymbol{x}_n) + F(\boldsymbol{a}_1, \cdots, \boldsymbol{c}_i, \cdots, \boldsymbol{x}_n). \end{aligned}$$

$$\begin{aligned} F(\boldsymbol{x}_1, \cdots, r\boldsymbol{x}_i, \cdots, \boldsymbol{x}_n) &= |A\boldsymbol{x}_1, \cdots, A(r\boldsymbol{x}_i), \cdots, A\boldsymbol{x}_n| = |A\boldsymbol{x}_1, \cdots, r(A\boldsymbol{x}_i), \cdots, A\boldsymbol{x}_n| \\ &= r|A\boldsymbol{x}_1, \cdots, A\boldsymbol{x}_i, \cdots, A\boldsymbol{x}_n| = rF(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_i, \cdots, \boldsymbol{x}_n). \end{aligned}$$

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. □

次は 7.1 項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. □

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. その証明は次節で与えよう (定理 13.2.3).

13　余因子展開とクラメルの公式

11 節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. 本節では, 余因子展開を通して, 2 次正方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 $n-1$ 次正方行列を A_{ij} とかく. すなわち,

$$A = \left(\begin{array}{ccccc} & & a_{1j} & & \\ & A & \vdots & B & \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ & C & \vdots & D & \\ & & a_{lj} & & \end{array}\right) \quad \text{のとき} \quad A_{ij} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

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 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ とすれば, 式 12.4.1 と同様にして, 各 $\boldsymbol{a}_j$ は次のように書ける:

$$\boldsymbol{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} = a_{1j}\boldsymbol{e}_1 + a_{2j}\boldsymbol{e}_2 + \cdots + a_{nj}\boldsymbol{e}_n.$$

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:

$$\begin{aligned} |A| &= |\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{j-1}, a_{1j}\boldsymbol{e}_1 + a_{2j}\boldsymbol{e}_2 + \cdots + a_{nj}\boldsymbol{e}_n, \boldsymbol{a}_{j+1}, \cdots, \boldsymbol{a}_n| \\ &= a_{1j}\,|\boldsymbol{a}_1, \cdots, \boldsymbol{e}_1, \cdots, \boldsymbol{a}_n| + a_{2j}\,|\boldsymbol{a}_1, \cdots, \boldsymbol{e}_2, \cdots, \boldsymbol{a}_n| + \cdots + a_{nj}\,|\boldsymbol{a}_1, \cdots, \boldsymbol{e}_n, \cdots, \boldsymbol{a}_n|. \end{aligned} \tag{13.1.1}$$

更に, 式 13.1.1 に現れる i 番目の項を計算すると

$$
\begin{aligned}
a_{ij}\begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & 1 & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix}
&= a_{ij}(-1)^{i-1}\begin{vmatrix} a_{i1} & \cdots & 1 & \cdots & a_{in} \\ a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} \qquad (i\text{ 行を } 1\text{ 行へ移動させた}) \\
&= a_{ij}(-1)^{i-1}(-1)^{j-1}\begin{vmatrix} 1 & a_{i1} & \cdots & \cdots & a_{in} \\ 0 & a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} \qquad (j\text{ 列を } 1\text{ 列へ移動させた}) \\
&= a_{ij}(-1)^{i+j-2}\left|\begin{array}{c|ccc} 1 & a_{i1} & \cdots & a_{in} \\ \hline 0 & & & \\ \vdots & & \Large A_{ij} & \\ 0 & & & \end{array}\right| = a_{ij}(-1)^{i+j}|A_{ij}|.
\end{aligned}
$$

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,

$$\begin{vmatrix} t & -1 & & & \\ & t & -1 & & \Large O \\ & & \ddots & \ddots & \\ & \Large O & & t & -1 \\ a_0 & a_1 & \cdots & a_{k-2} & t + a_{k-1} \end{vmatrix} = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0.$$

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:

$$\begin{vmatrix} t & -1 \\ a_0 & t + a_1 \end{vmatrix} = t(t + a_1) - a_0 \cdot (-1) = t^2 + a_1 t + a_0.$$

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:

$$
\begin{aligned}
(\text{左辺}) &= t\begin{vmatrix} t & -1 & & \\ & \ddots & \ddots & \\ & & t & -1 \\ a_1 & \cdots & a_{k-2} & t+a_{k-1} \end{vmatrix} + (-1)^{k+1}a_0 \begin{vmatrix} -1 & & & \\ t & -1 & & \\ & \ddots & \ddots & \\ & & t & -1 \end{vmatrix} \\
&= t(t^{k-1} + a_{k-1}t^{k-2} + \cdots + a_2 t + a_1) + (-1)^{k+1}a_0 \cdot (-1)^{k-1} = (\text{右辺}).
\end{aligned}
$$

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. □

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.

$$
\begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \vdots \\ \boldsymbol{c}_n \end{bmatrix}
\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix}
= \begin{bmatrix} \boldsymbol{c}_1\boldsymbol{a}_1 & \boldsymbol{c}_1\boldsymbol{a}_2 & \cdots & \boldsymbol{c}_1\boldsymbol{a}_n \\ \boldsymbol{c}_2\boldsymbol{a}_1 & \boldsymbol{c}_2\boldsymbol{a}_2 & \cdots & \boldsymbol{c}_2\boldsymbol{a}_n \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{c}_n\boldsymbol{a}_1 & \boldsymbol{c}_n\boldsymbol{a}_2 & \cdots & \boldsymbol{c}_n\boldsymbol{a}_n \end{bmatrix} = E.
$$

すなわち,

$$
\boldsymbol{c}_j\boldsymbol{a}_i = \begin{cases} 1 & j = i \text{ のとき}, \\ 0 & j \neq i \text{ のとき}. \end{cases} .
$$

上式を満たす行ベクトル $\boldsymbol{c}_j$ 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 $\boldsymbol{b}$ 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 $i = 1, \cdots, n$ について $B_{ij} = A_{ij}$ であるから,

$$
\begin{array}{c}
\begin{array}{ccccc} \phantom{a_{11}} & & j & & \phantom{a_{1n}} \\ & & \text{列} & & \end{array} \\
\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & b_i & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}
\end{array}
= (-1)^{1+j}b_1|A_{1j}| + (-1)^{2+j}b_2|A_{2j}| + \cdots + (-1)^{n+j}b_n|A_{nj}| .
$$

この右辺は次の二つのベクトル

$$
\tilde{\boldsymbol{a}}_j = \begin{bmatrix} (-1)^{1+j}|A_{1j}| \\ (-1)^{2+j}|A_{2j}| \\ \vdots \\ (-1)^{n+j}|A_{nj}| \end{bmatrix} \qquad \text{および} \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}
$$

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り, ここでは

$\tilde{\boldsymbol{a}}_j$ を転置して

$$
\begin{array}{c}
\begin{array}{ccccc} & & j\text{列} & & \end{array} \\
\begin{vmatrix}
a_{11} & \cdots & b_1 & \cdots & a_{1n} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_{i1} & \cdots & b_i & \cdots & a_{in} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_{n1} & \cdots & b_n & \cdots & a_{nn}
\end{vmatrix}
\end{array}
= \begin{bmatrix} (-1)^{1+j}|A_{1j}| & (-1)^{2+j}|A_{2j}| & \cdots & (-1)^{n+j}|A_{nj}| \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}
$$

$$
= {}^t\tilde{\boldsymbol{a}}_j \boldsymbol{b} \tag{13.2.1}
$$

を得る．とくに $\boldsymbol{b}$ として A の各列 $\boldsymbol{a}_i$ 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，

$$
{}^t\tilde{\boldsymbol{a}}_j \boldsymbol{a}_i = \begin{cases} |A| & j = i \text{ のとき}, \\ 0 & j \neq i \text{ のとき}. \end{cases} \tag{13.2.2}
$$

そこで，行ベクトル ${}^t\tilde{\boldsymbol{a}}_1, {}^t\tilde{\boldsymbol{a}}_2, \cdots, {}^t\tilde{\boldsymbol{a}}_n$ を成分とする行列を $\widetilde{A}$ としよう:

$$
\widetilde{A} := \begin{bmatrix} {}^t\tilde{\boldsymbol{a}}_1 \\ {}^t\tilde{\boldsymbol{a}}_2 \\ \vdots \\ {}^t\tilde{\boldsymbol{a}}_n \end{bmatrix} = \begin{bmatrix}
(-1)^{1+1}|A_{11}| & (-1)^{2+1}|A_{21}| & \cdots, & (-1)^{n+1}|A_{n1}| \\
(-1)^{1+2}|A_{12}| & (-1)^{2+2}|A_{22}| & \cdots, & (-1)^{n+2}|A_{n2}| \\
\vdots & \vdots & \vdots & \vdots \\
(-1)^{1+n}|A_{1n}| & (-1)^{2+n}|A_{2n}| & \cdots, & (-1)^{n+n}|A_{nn}|
\end{bmatrix}.
$$

ここで，$\widetilde{A} = [a^*_{ij}]$ 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 $\widetilde{A}A = |A|E$ を得る:

$$
\widetilde{A}A = \begin{bmatrix} {}^t\tilde{\boldsymbol{a}}_1 \\ {}^t\tilde{\boldsymbol{a}}_2 \\ \vdots \\ {}^t\tilde{\boldsymbol{a}}_n \end{bmatrix} [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n] = \begin{bmatrix} |A| & & & \\ & |A| & & \\ & & \ddots & \\ & & & |A| \end{bmatrix} = |A|E.
$$

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，$A\widetilde{A} = |A|E$ を示せ．

定義 13.2.2. n 次正方行列 $A = [a_{ij}]$ に対して，本項で定義した $\widetilde{A} = [a^*_{ij}]$ (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 $\widetilde{A}$ 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，本論もこれに従った．

これまでの議論により，$\widetilde{A}A = A\widetilde{A} = |A|E$ 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 $A^{-1} = \dfrac{1}{|A|}\widetilde{A}$ である．

Proof. $\dfrac{1}{|A|}\widetilde{A}$ が A の逆行列であることは直ちに確かめられる:

$$
\frac{1}{|A|}\widetilde{A} \cdot A = \frac{1}{|A|} \cdot |A|E = E, \qquad A \cdot \frac{1}{|A|}\widetilde{A} = \frac{1}{|A|}(A\widetilde{A}) = \frac{1}{|A|} \cdot |A|E = E.
$$

□

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 13.2.4. 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． □

例 13.2.5. 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, $\widetilde{A} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$. 更に $|A| = a_{11}a_{22} - a_{12}a_{21} \neq 0$ のとき,

$$
A^{-1} = \frac{1}{|A|}\widetilde{A} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.
$$

発展 (無限次元の行列).

本節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. 次元の概念は後期に入ってから扱うゆえ, 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ことができるゆえ, 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:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \end{bmatrix}.$$

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:

$$BA = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & & \ddots \end{bmatrix} = E, \quad AB = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & & \ddots \end{bmatrix}.$$

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 19.6 項を, 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 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ (ただし各 $\boldsymbol{a}_j$ 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 $A\boldsymbol{x} = \boldsymbol{b}$ の唯一解 $\boldsymbol{x} = {}^t[x_1, \cdots, x_n]$ は, $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ 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:

$$x_j = \frac{|\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{j-1}, \boldsymbol{b}, \boldsymbol{a}_{j+1}, \cdots, \boldsymbol{a}_n|}{|A|} \qquad (j = 1, \cdots, n).$$

Proof. A の余因子行列を $\widetilde{A}$ とする. 定理 13.2.3 より $A^{-1} = \dfrac{1}{|A|}\widetilde{A}$ であるから,

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A^{-1}\boldsymbol{b} = \frac{1}{|A|}\widetilde{A}\boldsymbol{b} = \frac{1}{|A|}\begin{bmatrix} {}^t\tilde{\boldsymbol{a}}_1 \\ {}^t\tilde{\boldsymbol{a}}_2 \\ \vdots \\ {}^t\tilde{\boldsymbol{a}}_n \end{bmatrix}\boldsymbol{b} = \frac{1}{|A|}\begin{bmatrix} {}^t\tilde{\boldsymbol{a}}_1\boldsymbol{b} \\ {}^t\tilde{\boldsymbol{a}}_2\boldsymbol{b} \\ \vdots \\ {}^t\tilde{\boldsymbol{a}}_n\boldsymbol{b} \end{bmatrix}.$$

したがって, 各成分 x_j は,

$$x_j = \frac{1}{|A|}\,{}^t\tilde{\boldsymbol{a}}_j\boldsymbol{b} = \frac{1}{|A|}\cdot\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & b_i & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}.$$

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. $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ と置けば

$$\boldsymbol{b} = A\boldsymbol{x} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n$$

である. 行列式 $|\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{j-1}, \boldsymbol{b}, \boldsymbol{a}_{j+1}, \cdots, \boldsymbol{a}_n|$ を式 13.1.1 と同様にして展開すると,

$$\begin{aligned} &|\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{j-1}, \boldsymbol{b}, \boldsymbol{a}_{j+1}, \cdots, \boldsymbol{a}_n| \\ &\quad = |\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{j-1}, x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n, \boldsymbol{a}_{j+1}, \cdots, \boldsymbol{a}_n| \\ &\quad = x_1\,|\boldsymbol{a}_1, \cdots, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n| + \cdots + x_j\,|\boldsymbol{a}_1, \cdots, \boldsymbol{a}_j, \cdots, \boldsymbol{a}_n| + \cdots + x_n\,|\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n, \cdots, \boldsymbol{a}_n|. \end{aligned}$$

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 $\boldsymbol{a}_i$ 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,

$$|\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{j-1}, \boldsymbol{b}, \boldsymbol{a}_{j+1}, \cdots, \boldsymbol{a}_n| = x_j|A|.$$

これを移項して求める等式を得る. □

14　集合概念の基礎

数学の歴史に集合が現れるのは 19 世紀の後半であり, それ以前の数学は集合を用いずに記述されていた. 確かに, これまでに扱ってきた内容は連立 1 次方程式の解法や行列式の計算など式変形を主体とする数学であり, ことさらに集合概念を押し出す必要はなかった. もしかすると, これ以降に学ぶ内容についても集合を用いずに議論を展開することが, あるいは可能かもしれない. しかし, この立場に固執すれば, 今後, より複雑な概念が縦横無尽に現れる中で, 定義や命題をやや曖昧に述べざるをえなかったり, あるいは証明において数学的に重要ではない些細な部分には目をつぶるような判断力や数学的センスを読者に要求することになるだろう. しかし, それでは多くの読者を路頭に迷わせることになってしまう.

現代数学において集合を用いた表現が市民権を得たのは, その記法を用いると厳密に述べやすいことにつきる. 数学は一部の選ばれた者のみに許された学問ではなく, 万人に許される学であるとする立場において, 数学的なセンスを問わずに誤解なく伝わる集合による表現はかかせない. 本論もこの立場に身を置き, 以降では集合と写像を用いた記述を採用する. そこで本節と 19 節では, 先に 1.5 および 1.6 項で述べた集合と写像に関する概念の発展として, これらのより高度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のみを取り上げるゆえ, それ以外の部分, 例えば合併集合や共通部分, 補集合といった集合演算などの扱いについては集合論の入門的な参考書を参照されたい.

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) $\{$ りんご, スイカ $\} \subset \{$ りんご, みかん, 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 1.5 項を参照せよ.

(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:

$$\text{集合 } A, B \text{ について,} \qquad A = B \iff A \subset B \text{ かつ } B \subset A. \tag{14.1.1}$$

例 14.1.6. 集合 $A = \{ \text{りんご, スイカ} \}$ と集合 $B = \{ \text{りんご, スイカ, スイカ} \}$ 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, $\{ \text{りんご, スイカ 1, スイカ 2} \}$ と書けばよい．

例 14.1.7. 本節以前の議論においても, 式 (14.1.1) の左向き「$A \subset B$ かつ $B \subset A \Longrightarrow A = B$」を暗黙裡のうちに何度か用いていた．

(1) 例 1.7.3 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 $f : \mathbb{R}^2 \to \mathbb{R}^2$ 全体の集合 Y が一致することを確認した．例 1.7.3 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 $A\boldsymbol{x} = \boldsymbol{b}$ の解全体の集合 X と一点からなる集合 $Y = \{B\boldsymbol{b}\}$ が一致することを述べている．まず, $\boldsymbol{a}$ を方程式の解 (つまり $\boldsymbol{a} \in X$) とすれば, $\boldsymbol{a}$ は $B\boldsymbol{b}$ に一致すること (つまり $\boldsymbol{a} \in Y$) を示した．これは $X \subset Y$ を示すことに相当している．次に $B\boldsymbol{b}$ 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 項) 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 $A\boldsymbol{x} = \boldsymbol{b}$ の解全体の集合 X と $\boldsymbol{a}_0 + c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + c_3\boldsymbol{a}_3$ 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 $B \subset A \Longrightarrow A = B$」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し, 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[25]を**正多面体**という．

[25] より正確には, ここでは凸多面体のみを考えている．

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 が指すもの全体の集合を与える.

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, 本論では "|" 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, $a_1, a_2, a_3, \ldots$ 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:

$$\{ 2,\ 4,\ 6,\ 8,\ 10,\ 12,\ \cdots \}, \quad \text{あるいは} \quad \{ 2n \mid n \in \mathbb{N} \}.$$

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 $f : \mathbb{R} \to \mathbb{R}$ 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[26]. 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[27].

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 $\{ x \in \mathbb{N} \mid x = 2m$ を満たす $m \in \mathbb{N}$ が存在する $\}$ は正の偶数全体の集合である.

[26] P の例は 3.2 項を参照せよ.

[27] 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.

(2) (m,n)-行列 A および m 次列ベクトル $\boldsymbol{b}$ に対して集合 $\{\, \boldsymbol{x} \in \mathbb{R}^n \,|\, A\boldsymbol{x} = \boldsymbol{b} \,\}$ は連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解全体からなる集合である.

(3) $\mathbb{R}^2$ の部分集合 $\{\, \boldsymbol{x} \in \mathbb{R}^2 \mid \boldsymbol{x} = (x,y)$ と置くと, $x > 0$ 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.

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): $[a,b] := \{\, x \in \mathbb{R} \,|\, a \leq x \leq b \,\}$, $(a,b) := \{\, x \in \mathbb{R} \,|\, a < x < b \,\}$, $[a,\infty) = \{\, x \in \mathbb{R} \,|\, x \geq a \,\}$.

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 $a < x < b \Longrightarrow x \in I$」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 $F : \mathbb{R}^4 \to M_2(\mathbb{R})$ を $F(a,b,c,d) := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ と定めておく. $M_2(\mathbb{Z}) = \{\, F(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathbb{Z}^4 \,\}$.

(2) 上の $\boldsymbol{x}$ や $F(\boldsymbol{x})$ を具体的に成分表示した表記: $M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \ (a,b,c,d) \in \mathbb{Z}^4 \right\}$.

(3) $M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| \ a,b,c,d \in \mathbb{Z} \right\}$, $\quad M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}) \middle| \ a,b,c,d \in \mathbb{Z} \right\}$.

(4) $M_2(\mathbb{Z}) = \left\{ A \in M_2(\mathbb{R}) \ \middle| \ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{をみたす } a,b,c,d \in \mathbb{Z} \text{ が存在する} \right\}$.

(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見なすことができる[28].

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. 本論もこれに準じる.

練習 14.3.2. 次の集合の記法は外延的か内包的か答えよ.

$$(1)\quad \{\, x \in \mathbb{Z} \mid x \text{ は奇数} \,\}, \qquad (2)\quad \{\, (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \,\}, \qquad (3)\quad \{\, (x, \sin x) \mid x \in \mathbb{R} \,\},$$

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 $f : \mathbb{R} \to \mathbb{R}$ 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) $X = \left\{ \, \boldsymbol{x} \in \mathbb{R}^5 \;\middle|\; \begin{bmatrix} 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix} \, \right\}$.

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,

$$X = \left\{ \begin{bmatrix} 0 \\ 4 \\ 0 \\ 5 \\ 0 \end{bmatrix} + c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix} \;\middle|\; c_1, c_2, c_3 \in \mathbb{R} \right\}.$$

14.4 和集合と共通部分

ここで紹介する記号は, 本論では, 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.

[28] 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 $f : \mathbb{N} \to \mathbb{R}$ を用いることで, f に各自然数を代入した値と見なせる.

事実 14.4.2. 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.

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 $A_1, A_2 \cdots, A_n$ に対して,

- $A_1 \cup \cdots \cup A_n := \{\, x \in X \mid i = 1, \cdots, n \text{ のいずれかにおいて } x \in A_i \text{ が成り立つ} \,\}$.
- $A_1 \cap \cdots \cap A_n := \{\, x \in X \mid \text{各 } i = 1, \cdots, n \text{ について } x \in A_i \,\}$.

上で定めた和集合および共通部分をそれぞれ $\displaystyle\bigcup_{i=1}^{n} A_i$, $\displaystyle\bigcap_{i=1}^{n} A_i$ 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. 本項では, 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[29]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[30].

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 ZFC[31]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 (詳しい理由は本節末のコラムを見よ). 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.

---

[29] $A \Rightarrow A$ 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.

[30] 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.

[31] 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.

よりみち (集合とは何か).

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．

しかしそうなると，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，大概は，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．

15　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 項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 (命題 6.3.2 と 6.3.3 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

$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in V$, $r, s \in \mathbb{R}$ とする．

I. 各元 $\boldsymbol{a}, \boldsymbol{b} \in V$ に対して和 $\boldsymbol{a} + \boldsymbol{b} \in V$ が定まっており, 次の性質を満たす:

(1) $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$,　(2) $(\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c})$,　(3) $\boldsymbol{a} + \mathbf{0} = \mathbf{0} + \boldsymbol{a} = \boldsymbol{a}$.

II. 各元 $\boldsymbol{a} \in V$ および $r \in \mathbb{R}$ に対してスカラー倍 $r \cdot \boldsymbol{a} \in V$ が定まっており, 次の性質を満たす:

(4) $r \cdot (s \cdot \boldsymbol{a}) = (rs) \cdot \boldsymbol{a}$,　(5) $(r + s) \cdot \boldsymbol{a} = (r \cdot \boldsymbol{a}) + (s \cdot \boldsymbol{a})$,　(6) $r \cdot (\boldsymbol{a} + \boldsymbol{b}) = (r \cdot \boldsymbol{a}) + (r \cdot \boldsymbol{b})$,

(7) $1 \cdot \boldsymbol{a} = \boldsymbol{a}$,　(8) $0 \cdot \boldsymbol{a} = \mathbf{0}$.

上の性質を満たす $\mathbf{0}$ のことを V の**零元**または**零ベクトル**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 $r \cdot \boldsymbol{a}$ を $r\boldsymbol{a}$ 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[32]．命題 2.4.1 から次が従う:

(1) $\mathbb{R}^n$ における原点 $\mathbf{0} := (0, \cdots, 0)$ を零元と見なすことにより, 1.5 項で定めた $\mathbb{R}^n$ 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．

(2) (m, n)-行列全体のなす集合を $M_{m,n}(\mathbb{R})$ と書く．2.2 節で述べたように $M_{m,n}(\mathbb{R})$ には和とスカラー倍が定義されている．零行列 $O_{m,n}$ を零元と見なすことにより, $M_{m,n}(\mathbb{R})$ 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．

[32] 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 2 節最後のコラムを見よ)．

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節で更に詳しく述べる．

例 15.2.1. $W_{[1,-1]} := \{ (x, y) \in \mathbb{R}^2 \mid x - y = 0 \}$ とすれば, W は直線 $y = x$ 上の点全体を表す．$W_{[1,-1]}$ 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 $W_{[1,-1]}$ に含まれること, および $W_{[1,-1]}$ の各元どうしの和や W の元のスカラー倍が再び $W_{[1,-1]}$ 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, $\boldsymbol{a} = (a_1, a_2) \in W$, $\boldsymbol{b} = (b_1, b_2) \in W$, $r \in \mathbb{R}$ と仮定し, $\boldsymbol{a}+\boldsymbol{b}, r\boldsymbol{a} \in W$ を示そう．$\boldsymbol{a}+\boldsymbol{b} = (a_1 + b_1, a_2 + b_2)$ 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．$\boldsymbol{a}, \boldsymbol{b} \in W$ 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, $\boldsymbol{a}+\boldsymbol{b} \in W$ である．また $r\boldsymbol{a} = (ra_1, ra_2)$ 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 $(ra_1) - (ra_2) = r(a_1 - a_2) = r0 = 0$ と直ちに分かる． □

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 $(x_1, x_2, x_3, \cdots)$ 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:

- $(x_1, x_2, x_3, \cdots) + (y_1, y_2, y_3, \cdots) := (x_1 + y_1, x_2 + y_2, x_3 + y_3, \cdots)$,
- $r(x_1, x_2, x_3, \cdots) := (rx_1, rx_2, rx_3, \cdots)$.

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 $(x_1, x_2, x_3, \cdots)$ 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:

$$(x_n)_{n\in\mathbb{N}} + (y_n)_{n\in\mathbb{N}} = (x_n + y_n)_{n\in\mathbb{N}}, \qquad r(x_n)_{n\in\mathbb{N}} := (rx_n)_{n\in\mathbb{N}}.$$

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 $a_0, a_1, \ldots, a_m$ を用いて

$$f(x) = \sum_{i=0}^{m} a_i x^i \quad \Big(= a_m x^m + a_{m-1}x^{m-1} + \cdots + a_2 x^2 + a_1 x + a_0 \Big)$$

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 $x^0 = 1$ と定めている．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$\boldsymbol{n}$**次多項式 (polynomial of dgree** $\boldsymbol{n}$**)** という．

多項式の次数が n 以下であるとき, これを$\boldsymbol{n}$**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[33]．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．

[33] 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であれば次のようになる: 多項式 $\sum_{i=0}^{m} a_i x^i \in \mathbb{R}[x]$ に対して数列 $(a_0, a_1, \cdots, a_m, 0, 0, \cdots) \in \mathbb{R}^{\mathbb{N}}$ を対応させる写像を $T : \mathbb{R}[x] \to \mathbb{R}^{\mathbb{N}}$ とし, 二つの多項式 f, g が等しいとは, $T(f) = T(g)$ が成り立つことと定める．

例 15.2.4. 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**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, 形式的な式と見なす立場と, x を変数とする関数と見なす立場がある. 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 $i=0,\cdots,n$ 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であろう. 次の命題はその根拠となる.

命題 15.2.5. $n+1$ 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) 各 $x \in I$ について $p(x)=q(x)$, (2) 各 $i=0,\ldots,n$ について $a_i=b_i$.

Proof. (2)⇒(1) は明らかゆえ (1)⇒(2) を示す. $p(x)=q(x)$ より関数 $p(x)-q(x)=\sum_{i=0}^{n}(a_i-b_i)x^n$ 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 $t_1,\cdots,t_{n+1} \in I$ 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.

$$\begin{cases} {t_1}^n(a_n-b_n)+{t_1}^{n-1}(a_{n-1}-b_{n-1})+\cdots+t_1(a_1-b_1)+1\cdot(a_0-b_0) & =0 \\ {t_2}^n(a_n-b_n)+{t_2}^{n-1}(a_{n-1}-b_{n-1})+\cdots+t_2(a_1-b_1)+1\cdot(a_0-b_0) & =0 \\ & \vdots \\ {t_n}^n(a_n-b_n)+{t_n}^{n-1}(a_{n-1}-b_{n-1})+\cdots+t_n(a_1-b_1)+1\cdot(a_0-b_0) & =0 \\ {t_{n+1}}^n(a_n-b_n)+{t_{n+1}}^{n-1}(a_{n-1}-b_{n-1})+\cdots+t_{n+1}(a_1-b_1)+1\cdot(a_0-b_0) & =0 \end{cases} \tag{15.2.1}$$

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:

$$A := \begin{bmatrix} {t_1}^n & {t_1}^{n-1} & \cdots & t_1 & 1 \\ {t_2}^n & {t_2}^{n-1} & \cdots & t_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ {t_n}^n & {t_n}^{n-1} & \cdots & t_n & 1 \\ {t_{n+1}}^n & {t_{n+1}}^{n-1} & \cdots & t_{n+1} & 1 \end{bmatrix}.$$

すると式 15.2.1 は, $\boldsymbol{x} = {}^t(a_n-b_n,\cdots,a_0-b_0)$ が $A\boldsymbol{x}=\mathbf{0}$ 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 $|A| = \pm\prod_{1\le i<j\le n+1}(t_j-t_i)$ 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 $A\boldsymbol{x}=\mathbf{0}$ の解は唯一解 $\boldsymbol{x}=\mathbf{0}$ のみである. したがって $a_i-b_i=0$ $(i=0,\cdots,n)$. □

多項式を関数と見なす立場からは, 次のような一般化が考えられる.

例 15.2.6. $\mathbb{R}$ 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 $(f+g): I \to \mathbb{R}$ 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 $(rf): I \to \mathbb{R}$ 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 $f+g, rf \in C(I)$) の証明は解析系の講義に譲ろう[34]. $C(I)$ 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 15.2.4 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 15.2.4 における和とスカラー倍の定義と, 例 15.2.6 におけるそれは一致する.

例 15.2.7. (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. そこで, 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を定義することができる. 例えば $\mathbb{R}^2$ の点列 $\boldsymbol{x}_n$ が点 $\boldsymbol{x} \in \mathbb{R}^2$ に収束することは, $\boldsymbol{x}_n$ と $\boldsymbol{x}$ の距離[35]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 15.2.6 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 15.2.6 と同様に和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見なすことができる. 実際, 数列空間の各元 $(x_1, x_2, \cdots)$ は $x(n) := x_n$ なる関数 $x : \mathbb{N} \to \mathbb{R}$ に対応し, 逆に関数 $y : \mathbb{N} \to \mathbb{R}$ は実数列 $y_n := y(n)$ (すなわち $(y(1), y(2), \cdots)$) 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[36]. $\mathbb{N}$ を定義域とする関数と数列の違いは, $x(1), x(2), \cdots$ と書くか, あるいは $x_1, x_2, \cdots$ 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 15.2.8 (発展)**.** $C(X)$ には次のようにして積も定めることができる.

$$f, g \in C(X) \text{ に対して関数 } fg \text{ を次で定める: 各 } x \in X \text{ について, } (fg)(x) := f(x)f(g).$$

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 におけるにおける (1) から (18) すべての性質を満たす[37]. 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.

実は, いかなる線形空間も, 強引に積を導入して多元環とみなすことができる. しかしながらここでは, 線形代数学における行列のもつ性質, すなわち $M_n(\mathbb{R})$ の数学的な捉え方をいかに昇華するかという文脈における多元環についての言及に留めておこう.

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

[34] 収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.

[35] $\boldsymbol{x}_n = (a_n, b_n)$ と $\boldsymbol{x} = (a, b)$ 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.

[36] 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.

[37] 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.

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. ここで, 行列に現れる成分を有理数に限ったり, あるいは複素数も認めるといった状況を考えよう.

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[38].

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.

(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うえで $1-\sqrt{2}$ 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. □

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 $\boldsymbol{a},\boldsymbol{b}\in V$ に対して和 $\boldsymbol{a}+\boldsymbol{b}\in V$ が定まっており, 次の性質を満たす:

(1) $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$, (2) $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$, (3) $\boldsymbol{a}+\mathbf{0}=\mathbf{0}+\boldsymbol{a}=\boldsymbol{a}$.

II. 各元 $\boldsymbol{a}\in V$ および $r\in K$ に対してスカラー倍 $r\cdot\boldsymbol{a}\in V$ が定まっており, 次の性質を満たす:

(4) $r\cdot(s\cdot\boldsymbol{a})=(rs)\cdot\boldsymbol{a}$, (5) $(r+s)\cdot\boldsymbol{a}=(r\cdot\boldsymbol{a})+(s\cdot\boldsymbol{a})$, (6) $r\cdot(\boldsymbol{a}+\boldsymbol{b})=(r\cdot\boldsymbol{a})+(r\cdot\boldsymbol{b})$,

(7) $1\cdot\boldsymbol{a}=\boldsymbol{a}$, (8) $0\cdot\boldsymbol{a}=\mathbf{0}$.

本論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る. この点についてより詳しいことは固有値の項目で述べよう.

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 $(x_1,\cdots,x_n)$ 全体からなる集合を K^n とする. 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. これは K 上の線形空間である.

[38] 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.

発展 (体を線形空間と見る)

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も分かる. 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が必要なはずである. 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.

16　いろいろな線形部分空間

例 15.2.1 における $W_{[1,-1]}$ と $\mathbb{R}^2$ の関係や, 例 15.2.4 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．

本節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．

$$\text{(i) } \mathbf{0}_V \in W, \quad \text{(ii) } \boldsymbol{a}, \boldsymbol{b} \in W \Longrightarrow \boldsymbol{a} + \boldsymbol{b} \in W, \quad \text{(iii) } \boldsymbol{a} \in W, r \in \mathbb{R} \Longrightarrow r\boldsymbol{a} \in W.$$

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．本論では, これらを**部分空間**と略称で述べる[39].

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．

部分空間の例は本節の後半で述べる．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, あと一歩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, $\boldsymbol{w} = (x_1, y_1)$, $\boldsymbol{v} = (x_2, y_2)$ とし, $\boldsymbol{w}, \boldsymbol{v} \in W$ 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 $\boldsymbol{w} + \boldsymbol{v} = (x_1 + x_2, y_1 + y_2) \in W$．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) を満たさない反例を一つ挙げればよい．例えば $\boldsymbol{a} = (1, 0)$, $r = -1$ とすれば $\boldsymbol{a} \in W$, $r \in \mathbb{R}$ である．しかしながら $r\boldsymbol{a} = -1(1, 0) = (-1, 0)$ 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 $r\boldsymbol{a} \notin W$．□

(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, $\boldsymbol{w} = (a, b) \in W$, $r \in \mathbb{R}$ 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

[39]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．

たがって $r\boldsymbol{w} = (rx, ry) \in W$ である．W が (ii) を満たさないことを示そう．例えば, $\boldsymbol{a} = (1, 0)$, $\boldsymbol{b} = (0, 1)$ とすれば $\boldsymbol{a}, \boldsymbol{b} \in W$ である．$\boldsymbol{a} + \boldsymbol{b} = (1, 0) + (0, 1) = (1, 1)$ 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 $\boldsymbol{a} + \boldsymbol{b} \notin W$. □

(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．もし仮に (ii) を満たさないとすれば, それは $\boldsymbol{w}, \boldsymbol{v} \in W$ であるにも関わらず $\boldsymbol{w} + \boldsymbol{v} \notin W$ 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 $\boldsymbol{w} \in W$ 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. □

よりみち (前提が偽なる命題).

例 16.1.3 (3) において空集合が条件 (ii) を満たすことの説明として「(*) 前提が満たされない命題は, いかなる結論が述べられていても正しい」という論理の原則を持ち出すことが多い．いまの例では, 前提となる $\boldsymbol{a}, \boldsymbol{b} \in W$ が成立しないゆえ (ii) は真であるという考え方である．しかしながら, この論理の原則を盲目的に認める立場に立って学ぶのであれば, それは迷信を信仰しているに等しい. 原則 (*) が認められるゆえんは何か, 自らの言葉で咀嚼することが学習者に望まれている.

ところで, 前提が満たされない議論があることを踏まえれば, 反例の存在を論じる際にも注意が必要なことがわかる．何故なら, 反例を構成するための手順を説明したつもりでいても, そのような手順を踏める対象が現実には存在しない可能性があるからである．つまり反例の存在証明においては, その構成手順を提示するのみでは不十分であり, 具体的な例を挙げる必要がある.

さて, 原則 (*) は次のようにして説明される．ここでは三種類の説明を挙げておく.

命題 16.1.4. F を偽なる条件とする．任意の条件 P について「F ならば P」は成立する.

Proof. 「F ならば P」を示すために F を仮定しよう．すると「F または P」であることが認められる．すなわち, F と P のうち少なくともいずれか一方が成立することになる．ところが F は偽なる条件であったゆえ成立せず, したがって, もう一方の条件である P が成り立たねばならない．以上より P が導かれた. □

Proof. 背理法により示す．仮に「F ならば P」が成り立たないとすれば, それは F が成り立つにもかかわらず P が不成立であることを意味する．とくに F が成立し, これは F が偽であることに反する．ゆえに「F ならば P」は成り立つ. □

Proof. 「F ならば P」と同値な対偶命題「(P でない) ならば (F でない)」について考えよう．この命題は結論が正しい命題ゆえ真である．ゆえに, もとの命題「F ならば P」も真である. □

論理の原則に更に踏み込んで, 背理法による論法や対偶の同値性が認められるのは何故だろうか. そこには, 条件「A」とその否定「A でない」において, 一方が成立しなければもう一方は成立すると考える立場 (これを**排中律**という) が背景にある.

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 16.1.5. 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:

$$\text{(iv)} \qquad r, s \in \mathbb{R},\ \boldsymbol{a}, \boldsymbol{b} \in W \Longrightarrow r\boldsymbol{a} + s\boldsymbol{b} \in W.$$

Proof. (ii) と (iii) を仮定して (iv) を示そう. $r, s \in \mathbb{R}$, $\boldsymbol{a}, \boldsymbol{b} \in W$ とすれば, (iii) より $r\boldsymbol{a}, s\boldsymbol{b} \in W$ である. これに (ii) を適用し $r\boldsymbol{a} + s\boldsymbol{b} \in W$ 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. □

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 16.1.6. W を線形空間 V の部分空間とすれば, 各 $\boldsymbol{v}_i \in W$ および $r_i \in \mathbb{R}$ について $\sum_{i=1}^{\ell} r_i \boldsymbol{v}_i \in W$.

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. $\boldsymbol{v}_i \in W$, $r_i \in \mathbb{R}$ $(i = 1, \ldots, \ell + 1)$ とし, $\boldsymbol{u} = \sum_{i=1}^{\ell+1} r_i \boldsymbol{v}_i$ とすれば, $\boldsymbol{u} = (\sum_{i=1}^{\ell} r_i \boldsymbol{v}_i) + r_{\ell+1}\boldsymbol{v}_{\ell+1}$ と書ける, 帰納法の仮定より $\sum_{i=1}^{\ell} r_i \boldsymbol{v}_i \in W$ であり, 性質 (iii) より $r_{\ell+1}\boldsymbol{v}_{\ell+1} \in W$ である. よって, 性質 (ii) より $(\sum_{i=1}^{\ell} r_i \boldsymbol{v}_i) + r_{\ell+1}\boldsymbol{v}_{\ell+1} \in W$, すなわち $\boldsymbol{u} \in W$ である. □

17 節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.

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[40]でないベクトル $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$ を取ると (ただし $\boldsymbol{a}, \boldsymbol{b} \neq \boldsymbol{0}$), $W = \{r_1\boldsymbol{a} + r_2\boldsymbol{b} \mid r_1, r_2 \in \mathbb{R}\}$ によって $\mathbb{R}^3$ 上の原点 $\boldsymbol{0}$ および $\boldsymbol{a}, \boldsymbol{b}$ 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 $\boldsymbol{0} = r_1\boldsymbol{a} + r_2\boldsymbol{b}$ と表せる. ゆえに $\boldsymbol{0} \in W$ である.

(iv): $\boldsymbol{x}, \boldsymbol{y} \in W$ とすれば $x_1, x_2, y_1, y_2 \in \mathbb{R}$ を用いて $\boldsymbol{x} = x_1\boldsymbol{a} + x_2\boldsymbol{b}$, $\boldsymbol{y} = y_1\boldsymbol{a} + y_2\boldsymbol{b}$ と書ける. 各 $r, s \in \mathbb{R}$ について

$$r\boldsymbol{x} + s\boldsymbol{y} = r(x_1\boldsymbol{a} + x_2\boldsymbol{b}) + s(y_1\boldsymbol{a} + y_2\boldsymbol{b}) = (rx_1 + sy_1)\boldsymbol{a} + (rx_2 + sy_2)\boldsymbol{b}$$

である. すなわち, $r\boldsymbol{x} + s\boldsymbol{y}$ は $r_1 = rx_1 + sy_1$, $r_2 = rx_2 + sy_2$ とすることで $r_1\boldsymbol{a} + r_2\boldsymbol{b}$ と表せる. ゆえに $r\boldsymbol{x} + s\boldsymbol{y} \in W$ である. □

例 16.2.1 の証明において, $\boldsymbol{a}, \boldsymbol{b}$ が平行でないことは用いられていない. $\boldsymbol{a}, \boldsymbol{b}$ 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 $\{\boldsymbol{0}\}$) 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$ について $W = \{\sum_{i=1}^{n} r_i\boldsymbol{u}_i \mid r_1, \ldots, r_n \in \mathbb{R}\}$ 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[41].

[40] 線形空間における $\boldsymbol{0}$ でない二つのベクトル $\boldsymbol{a}, \boldsymbol{b}$ が平行であるとは, $r\boldsymbol{a} = \boldsymbol{b}$ を満たす実数 r が存在することと定める.

[41] 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.

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:

$$W_A := \{\, \boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} = \boldsymbol{0} \,\}.$$

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): $\boldsymbol{x} = \boldsymbol{0} \in \mathbb{R}^n$ とすれば, $\boldsymbol{x}$ は条件 $A\boldsymbol{x} = \boldsymbol{0}$ を満たす. ゆえに $\boldsymbol{0} \in W_A$ である.

(iv): $\boldsymbol{a}, \boldsymbol{b} \in W_A$, $r, s \in \mathbb{R}$ とする. このとき, W_A の定義から $A\boldsymbol{a} = \boldsymbol{0}$, $A\boldsymbol{b} = \boldsymbol{0}$ が成り立つ. このとき $A(r\boldsymbol{a} + s\boldsymbol{b}) = r(A\boldsymbol{a}) + s(A\boldsymbol{b}) = r\boldsymbol{0} + s\boldsymbol{0} = \boldsymbol{0}$ ゆえ, $\boldsymbol{x} = r\boldsymbol{a} + s\boldsymbol{b}$ は条件 $A\boldsymbol{x} = \boldsymbol{0}$ を満たす. ゆえに $r\boldsymbol{a} + s\boldsymbol{b} \in W_A$ である. □

上の W_A は斉次形連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{0}$ の解全体の空間に一致する. そこで, W_A は方程式 $A\boldsymbol{x} = \boldsymbol{0}$ 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 $\boldsymbol{x} = \boldsymbol{0}$ のみからなる空間 $\{\boldsymbol{0}\}$ となる.

解空間の幾何的な意味を検討しよう. 列ベクトル $\boldsymbol{x} \in \mathbb{R}^n$ が $A\boldsymbol{x} = \boldsymbol{0}$ を満たすということは, A の各行と $\boldsymbol{x}$ 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 $\boldsymbol{a}, \boldsymbol{b}$ と共に直交するベクトル $\boldsymbol{c} \in \mathbb{R}^3$ を一つ取ると[42], $\boldsymbol{c}$ の成分を横に並べた $(1,3)$-行列を A とすれば, $W = W_A$ となることが予想される. 詳しい説明は内積空間の節で述べよう.

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) $W = \{\boldsymbol{0}\}$.　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) $W = \{\, r\boldsymbol{y} \mid r \in \mathbb{R} \,\}$, ただし $\boldsymbol{y} = \begin{bmatrix} a \\ b \end{bmatrix} \neq \boldsymbol{0}$ とする (W は原点 $\boldsymbol{0}$ と $\boldsymbol{y}$ を通る直線上の点全体を表す).

解答例: $\boldsymbol{z} = \begin{bmatrix} -b \\ a \end{bmatrix}$ は $\boldsymbol{y}$ に直交するベクトルである. $\boldsymbol{z}$ 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.

($W \subset W_A$): $\boldsymbol{x} \in W$ を勝手に取れば, 実数 r を用いて $\boldsymbol{x} = r\boldsymbol{y}$ と書ける. このとき,

$$A\boldsymbol{x} = A(r\boldsymbol{y}) = rA\boldsymbol{y} = r\begin{bmatrix} -b & a \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = r(-ba + ab) = r0 = 0.$$

ゆえに $\boldsymbol{x} \in W_A$ である.

($W_A \subset W$): $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in W_A$ を勝手に取れば, $A\boldsymbol{x} = \boldsymbol{0}$ が成り立っている. すなわち, $-bx_1 + ax_2 = 0$ である. $\boldsymbol{y} \neq \boldsymbol{0}$ 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, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ra \\ rb \end{bmatrix} = r\boldsymbol{y}$. ゆえに $\boldsymbol{x} \in W$ 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 $\boldsymbol{x} \in W$ を得る. □

[42] このような $\boldsymbol{c}$ 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, $\boldsymbol{a}$ と $\boldsymbol{b}$ の外積と呼ばれるベクトル $\boldsymbol{a} \times \boldsymbol{b}$ が, この例の一つに相当する.

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:

$$x_{n+2} = x_{n+1} + x_n. \qquad (n \in \mathbb{N}) \tag{16.3.1}$$

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 $\mathbf{0} = (0, 0, 0, \cdots)$ 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 $\boldsymbol{x} = (x_n)_{n\in\mathbb{N}}$ および $\boldsymbol{y} = (y_n)_{n\in\mathbb{N}}$ 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 $r\boldsymbol{x} + s\boldsymbol{y} = (rx_n + sy_n)_{n\in\mathbb{N}}$ が漸化式 (16.3.1) を満たすことに他ならない. ゆえに $r\boldsymbol{x} + s\boldsymbol{y} \in F$. □

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

$$a_1 = 0,\ a_2 = 1,\ a_3 = 1,\ a_4 = 2,\ a_5 = 3,\ a_6 = 5,\ a_7 = 8,\ a_9 = 13,\ a_{10} = 21, \cdots$$

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 $a_0, a_1, \cdots, a_{k-1}$ 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:

$$x_{n+k} = a_{k-1}x_{n+k-1} + a_{k-2}x_{n+k-2} + \cdots + a_1 x_{n+1} + a_0 x_n \qquad (n \in \mathbb{N}) \tag{16.3.2}$$

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**と呼び, これを満たす数列のことを**線形回帰数列**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) $\mathbb{R}$ 上の開区間 I を定義域とする C^∞-級関数[43]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 15.2.6 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:

$$y^{(2)}(x) = y^{(1)}(x) + y^{(0)}(x) \tag{16.3.3}$$

ここで, $y^{(n)}(x) = \dfrac{d^n}{dx^n}y(x)$ 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

$$rf^{(2)}(x) + sg^{(2)}(x) = \left(rf^{(1)}(x) + sg^{(1)}(x)\right) + \left(rf^{(0)}(x) + sg^{(0)}(x)\right)$$
$$(rf(x) + sg(x))^{(2)} = (rf(x) + sg(x))^{(1)} + (rf(x) + sg(x))^{(0)}$$

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 $rf + sg \in W$. □

[43] 何回でも微分できる関数のことを C^∞-関数という

練習 16.3.4. 自然数 $k \geq 2$, および k 個の実数 $a_0, a_2, \cdots, a_{k-1}$ 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.

$$y^{(k)}(x) = a_{k-1}y^{(k-1)}(x) + a_{k-2}y^{(k-2)}(x) + \cdots + a_1 y^{(1)}(x) + a_0 y^{(0)}(x). \tag{16.3.4}$$

方程式 (16.3.4) は実数係数の**線形常微分方程式**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

$$y^{(n+k)}(x) = a_{k-1}y^{(n+k-1)}(x) + a_{k-2}y^{(n+k-2)}(x) + \cdots + a_1 y^{(n+1)}(x) + a_0 y^{(n)}(x)$$

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はなく, 線形回帰数列の一般項の解法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.

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.

(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 $f : I \to \mathbb{R}$ 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.

17 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. 本節で述べることは, 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節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 $\boldsymbol{v} \in V$ が, $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n \in V$ および $a_1, a_2, \cdots, a_n \in \mathbb{R}$ を用いて

$$\boldsymbol{v} = a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + \cdots + a_n\boldsymbol{u}_n = \sum_{i=1}^{n} a_i\boldsymbol{u}_i$$

と書けるとき, $\boldsymbol{v}$ は $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の**線形結合**[44](**linear combination**) で書けるという. また, 零ベクトルを線形結合で表す式 $\sum_{i=1}^{n} a_i\boldsymbol{u}_i = \boldsymbol{0}$ が成立するとき, この等式を $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による**線形関係** (または **1 次関係**) と呼ぶ.

線形結合の標語的な解釈

$\boldsymbol{v}$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合で書かれるとは, $\boldsymbol{v}$ の情報が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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, $\boldsymbol{v}$ の情報を得るには, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の情報と線形結合に現れる係数 $a_1, \cdots, a_n$ の値さえ分かってい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:

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \boldsymbol{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

実際, 各 $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$ について $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_n \end{bmatrix} = a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n = \sum_{i=1}^{n} a_i\boldsymbol{e}_i$ 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 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 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 $\boldsymbol{v}$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合で書けるとしよう. $\boldsymbol{v}$ の情報を $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ たちの情報に還元する際に, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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.

要不要論

[44]線形結合は **1 次結合**とも呼ばれる.

(1) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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, $\boldsymbol{u}_n$ の情報が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ に分解されるとき, $\boldsymbol{v}$ の情報は $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ の情報だけから復元できることが示唆され, $\boldsymbol{u}_n$ は不要となる. 実際, $\boldsymbol{v}$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合で書けており, 更に $\boldsymbol{u}_n$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ の線形結合で書けるならば, $\boldsymbol{v}$ は $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ の線形結合で書ける:

$\boldsymbol{v} = \sum_{i=1}^{n} a_i \boldsymbol{u}_i$, $\boldsymbol{u}_n = \sum_{i=1}^{n-1} b_i \boldsymbol{u}_i$ と書けるならば,

$$\boldsymbol{v} = \sum_{i=1}^{n} a_i \boldsymbol{u}_i = \left(\sum_{i=1}^{n-1} a_i \boldsymbol{u}_i\right) + a_n \boldsymbol{u}_n = \left(\sum_{i=1}^{n-1} a_i \boldsymbol{u}_i\right) + a_n \left(\sum_{i=1}^{n-1} b_i \boldsymbol{u}_i\right) = \sum_{i=1}^{n-1} (a_i + a_n b_i) \boldsymbol{u}_i.$$

(2) 線形結合に現れる係数 a_i のうちのいくつかが 0 の場合.

このとき, $a_i = 0$ に対応する $\boldsymbol{u}_i$ は不要となる. 例えば, $\boldsymbol{a} = {}^t(5, 3, 0) \in \mathbb{R}^3$ を標準ベクトルに分解すると, $\boldsymbol{a} = 5\boldsymbol{e}_1 + 3\boldsymbol{e}_2 + 0\boldsymbol{e}_3 = 5\boldsymbol{e}_1 + 3\boldsymbol{e}_2$ ゆえ $\boldsymbol{e}_3$ 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 $\boldsymbol{e}_3$ 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, $\boldsymbol{e}_3$ が必要になることもあろう. 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, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 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 $\boldsymbol{u}_n$ が不要ということは, 線形空間 V のいかなる元も $\boldsymbol{u}_n$ を用いずに $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ 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 $\boldsymbol{u}_n$ が不要になる場合, $\boldsymbol{u}_n$ はとくに V の元であるから $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ 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, (1) の最後の式は, $\boldsymbol{u}_n$ を用いない線形結合になっている.

ベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$ が**線形独立**[45](**linearly independent**) であるとは,

$$a_1, \cdots, a_n \in \mathbb{R},\ a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + \cdots + a_n\boldsymbol{u}_n = \boldsymbol{0} \implies a_1 = a_2 = \cdots = a_n = 0$$

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 $a_1, \cdots, a_n$ を用いて, $a_1\boldsymbol{u}_1 + \cdots + a_n\boldsymbol{u}_n = \boldsymbol{0}$ 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 $a_1, \cdots, a_n$ のことを**自明でない**組という. 実数の組 $a_1, \cdots, a_n$ が自明でないことは, $(a_1, \cdots, a_n) \neq \boldsymbol{0}$ であることに等しい. また, 自明でない実数の組 $a_1, \cdots, a_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ は線形独立である.

Proof. $a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n = \boldsymbol{0}$ 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. □

[45] 線形独立 (線形従属) は **1 次独立** (**1 次従属**) とも呼ばれる.

(2) m 次列ベクトル $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が線形独立であることと, m 次行ベクトル ${}^t\boldsymbol{u}_1, \cdots, {}^t\boldsymbol{u}_n$ が線形独立であることは同値である. 実際, 式 $\sum_{i=1}^n a_i \boldsymbol{u}_i = O_{m,1}$ と式 $\sum_{i=1}^n a_i\, {}^t\boldsymbol{u}_i = O_{1,m}$ 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) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_n = \mathbf{0}$ の場合, 自明でない実数の組 $0, \cdots, 0, 1$ を係数とする線形関係 $0\boldsymbol{u}_1 + \cdots + 0\boldsymbol{u}_{n-1} + 1\boldsymbol{u}_i = \mathbf{0}$ が成立する. また, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の中に重複がある場合も線形従属である. 例えば $\boldsymbol{u}_1 = \boldsymbol{u}_2$ の場合, 自明でない実数の組 $1, -1, 0, \ldots, 0$ を係数とする線形関係 $1\boldsymbol{u}_1 + (-1)\boldsymbol{u}_2 + 0\boldsymbol{u}_3 \cdots + 0\boldsymbol{u}_i = \mathbf{0}$ 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 15.2.5), 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. □

練習 17.2.4. $n+1$ 個の組 $x^n, x^{n-1}, \cdots, x, \mathbf{1} \in \mathbb{R}[x]$ が線形独立であることを帰納法を用いて示せ.

次の命題は前項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, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の中に不要なものがあることと $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形従属性は同値である. また, それらの否定を取り, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の中に不要なものがないことと $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$ について次の (1) と (2) は同値である.

(1) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は線形従属である,

(2) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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)⇒(2): $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が線形従属であるとすれば, 自明でない実数の組 $a_1, \cdots, a_n$ を用いて $a_1\boldsymbol{u}_1 + \cdots + a_n\boldsymbol{u}_n = \mathbf{0}$ とできる. このとき $a_1, \cdots, a_n$ 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 $a_n\boldsymbol{u}_n = -a_1\boldsymbol{u}_1 - \cdots - a_n\boldsymbol{u}_n$ を得る. $a_n \neq 0$ ゆえこの両辺を a_n で割れば $\boldsymbol{u}_n = \left(-\frac{a_1}{a_n}\right)\boldsymbol{u}_1 + \cdots + \left(-\frac{a_1}{a_{n-1}}\right)\boldsymbol{u}_{n-1}$. ゆえに $\boldsymbol{u}_n$ は, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ の線形結合で書ける. $a_i \neq 0$ $(i = 1, \cdots, n-1)$ である場合も同様にして, $\boldsymbol{u}_i$ が他の $n-1$ 個の線形結合で書けることが示される.

(2)⇒(1): $\boldsymbol{u}_n$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ 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 $a_1, \cdots, a_{n-1}$ を用いて $\boldsymbol{u}_n = a_1\boldsymbol{u}_1 + \cdots + a_{n-1}\boldsymbol{u}_{n-1}$ と書ける. これを移項して $a_1\boldsymbol{u}_1 + \cdots + a_{n-1}\boldsymbol{u}_{n-1} + (-1)\boldsymbol{u}_n = \mathbf{0}$ を得る. 自明でない線形関係が得られたゆえ, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は線形従属である. $\boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形従属性を得る. □

例 17.2.6. (1) $\mathbb{R}^2$ の $\mathbf{0}$ でない列ベクトル $\boldsymbol{u}, \boldsymbol{v}$ 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 $\boldsymbol{u} = r\boldsymbol{v}$ と書けること, すなわち $\boldsymbol{u}, \boldsymbol{v}$ 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 $\boldsymbol{u}, \boldsymbol{v}$ が線形独立であるとは, $\boldsymbol{u}, \boldsymbol{v}$ が原点を通る同一直線上にないこと, つまり $\boldsymbol{u}, \boldsymbol{v}$ 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 $A = [\boldsymbol{u}, \boldsymbol{v}]$ 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 $\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\boldsymbol{u}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\boldsymbol{u}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ は線形従属である. これは例えば, $\boldsymbol{u}_3 = \frac{2}{5}\boldsymbol{u}_1 + \frac{1}{5}\boldsymbol{u}_2$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$ について次を示せ.

(1) n 個のベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が線形独立ならば, そこから一つ取り除いた $n-1$ 個の組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ も線形独立である.

解答例: $a_1\boldsymbol{u}_1+\cdots+a_{n-1}\boldsymbol{u}_{n-1} = \boldsymbol{0}$ とすれば, $a_n = 0$ とおくことで $a_1\boldsymbol{u}_1+\cdots+a_{n-1}\boldsymbol{u}_{n-1}+a_n\boldsymbol{u}_n = \boldsymbol{0}$ を得る. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形独立性より $a_1 = \cdots = a_{n-1} = a_n = 0$ であり, 特に $a_1 = \cdots = a_{n-1} = 0$. ゆえに $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ は線形独立である. □

(2) $n-1$ 個のベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}$ が線形従属ならば, そこに新たなベクトル $\boldsymbol{u}_n$ を加えた n 個のベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n-1}, \boldsymbol{u}_n$ も線形従属である.

解答例: (2) は (1) の対偶にほかならない. □

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 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ が線形独立であることと, (m, n)-行列 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ に関する斉次形連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{0}$ が唯一解を持つことは同値である.

Proof. $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ が線形独立であるとする. n 次列ベクトル $\boldsymbol{x} = {}^t(r_1, \cdots, r_n)$ が $A\boldsymbol{x} = \boldsymbol{0}$ を満たすならば,

$$r_1\boldsymbol{a}_1 + \cdots + r_n\boldsymbol{a}_n = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]\begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} = A\boldsymbol{x} = \boldsymbol{0}.$$

$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ の線形独立性より $r_1 = \cdots = r_n = 0$ である. つまり $\boldsymbol{x} = \boldsymbol{0}$ であり, 方程式 $A\boldsymbol{x} = \boldsymbol{0}$ の解は自明なものに限る.

次に, 方程式 $A\boldsymbol{x} = \boldsymbol{0}$ が自明な解しか持たないと仮定する. このとき $r_1\boldsymbol{a}_1 + \cdots + r_n\boldsymbol{a}_n = \boldsymbol{0}$ ならば $\boldsymbol{x} = {}^t(r_1, \cdots, r_n)$ について $A\boldsymbol{x} = r_1\boldsymbol{a}_1 + \cdots + r_n\boldsymbol{a}_n = \boldsymbol{0}$ となる. つまり $A\boldsymbol{x} = \boldsymbol{0}$ であり, この方程式は自明な解しか持たないゆえ $\boldsymbol{x} = \boldsymbol{0}$ を得る. すなわち $r_1 = \cdots = r_n = 0$ である. 以上より $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ は線形独立である. □

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 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ を行基本変形により $B = [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n]$ に変形できるとする. $1 \le n_1, n_2, \cdots, n_\ell \le n$ および $i = 1, \cdots, n$ に対して次が成り立つ.

(1) $\sum_{k=1}^{\ell} r_k\boldsymbol{a}_{n_k} = \boldsymbol{0} \iff \sum_{k=1}^{\ell} r_k\boldsymbol{b}_{n_k} = \boldsymbol{0}$.

(2) $\boldsymbol{a}_i = \sum_{k=1}^{\ell} r_k\boldsymbol{a}_{n_k}$ と書ける $\iff \boldsymbol{b}_i = \sum_{k=1}^{\ell} r_k\boldsymbol{b}_{n_k}$ と書ける

Proof. (1): A を B に行基本変形できることから, $A' = [\boldsymbol{a}_{n_1}, \cdots, \boldsymbol{a}_{n_\ell}]$ を $B' = [\boldsymbol{b}_{n_1}, \cdots, \boldsymbol{b}_{n_\ell}]$ に行基本変形できる. また, $[A'|\boldsymbol{0}]$ を $[B'|\boldsymbol{0}]$ に行基本変形できる. ゆえに命題 4.4.1 より $A'\boldsymbol{x} = \boldsymbol{0}$ の解と $B'\boldsymbol{x} = \boldsymbol{0}$ の解は一致しており, $\boldsymbol{x} = {}^t(r_1, \cdots, r_n)$ について

$$\textstyle\sum_{k=1}^{\ell} r_k\boldsymbol{a}_{n_k} = \boldsymbol{0} \iff \boldsymbol{x} \text{ は } A'\boldsymbol{x} = \boldsymbol{0} \text{ の解} \iff \boldsymbol{x} \text{ は } B'\boldsymbol{x} = \boldsymbol{0} \text{ の解} \iff \sum_{k=1}^{\ell} r_k\boldsymbol{b}_{n_k} = \boldsymbol{0}.$$

(2): $\boldsymbol{a}_i = \sum_{k=1}^{\ell} r_k\boldsymbol{a}_{n_k}$ とすれば, $\sum_{k=1}^{\ell} r_k\boldsymbol{a}_{n_k} + (-1)\boldsymbol{a}_i = \boldsymbol{0}$ である. $\ell+1$ 個のベクトルの組 $\boldsymbol{a}_{n_1}, \cdots \boldsymbol{a}_{n_\ell}, \boldsymbol{a}_i$ について (1) を適用し, $\sum_{k=1}^{\ell} r_k\boldsymbol{b}_{n_k} + (-1)\boldsymbol{b}_i = \boldsymbol{0}$ を得る. すなわち, $\boldsymbol{b}_i = \sum_{k=1}^{\ell} r_k\boldsymbol{b}_{n_k}$. 逆も同様に示される. □

上の (1) は組 $\boldsymbol{a}_{n_1}, \ldots, \boldsymbol{a}_{n_k}$ が線形独立 (あるいは線形従属) であることと組 $\boldsymbol{b}_{n_1}, \ldots, \boldsymbol{b}_{n_k}$ 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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_5$ 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.

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \boldsymbol{a}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{a}_5 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ -3 \\ -1 \end{bmatrix}.$$

解答例: 与えられた列ベクトルを並べた行列 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_5]$ の簡約化を $B = [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_5]$ とすれば,

$$A = \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 & -3 \\ 1 & 1 & 0 & 0 & -1 \end{bmatrix} \longrightarrow B = \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$\boldsymbol{b}_1, \cdots, \boldsymbol{b}_5$ 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 $\operatorname{rank} A = 3 \neq 5$ より方程式 $A\boldsymbol{x} = \boldsymbol{0}$ が自明でない解をもつことからも分かる), ゆえに $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_5$ も線形従属である. 行列 B の成分を見れば $\boldsymbol{b}_3 = -\boldsymbol{b}_1 + \boldsymbol{b}_2$, $\boldsymbol{b}_5 = -2\boldsymbol{b}_1 + \boldsymbol{b}_2 - \boldsymbol{b}_4$ と書けることが分かる. したがって $\boldsymbol{a}_3 = -\boldsymbol{a}_1 + \boldsymbol{a}_2$, $\boldsymbol{a}_5 = -2\boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{a}_4$ である.

注意 1: B の成分を見れば, $\boldsymbol{b}_2 = 2\boldsymbol{b}_1 + \boldsymbol{b}_4 + \boldsymbol{b}_5$ とも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 $\boldsymbol{b}_3$, $\boldsymbol{b}_5$ に限るというわけではない. 上で $\boldsymbol{b}_3$, $\boldsymbol{b}_5$ を取り上げたのは, これ以外の主成分を含む列 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_4$ が標準ベクトルであることから, $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_4$ が線形独立であること, および $\boldsymbol{b}_3$, $\boldsymbol{b}_5$ が $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_4$ 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 $\boldsymbol{b}_1, \boldsymbol{b}_4, \boldsymbol{b}_5$ も線形独立であり, これ以外の $\boldsymbol{b}_2, \boldsymbol{b}_3$ を $\boldsymbol{b}_1, \boldsymbol{b}_4, \boldsymbol{b}_5$ 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 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_4$ 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節の基底概念を通して説明される.

注意 2: あらかじめ $\boldsymbol{a}_4$, $\boldsymbol{a}_5$ 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 $\boldsymbol{a}_4, \boldsymbol{a}_5$ を先頭にした行列 $[\boldsymbol{a}_4, \boldsymbol{a}_5, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3]$ について同様の計算を行えばよい.

注意 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 $A\boldsymbol{x} = \boldsymbol{0}$ 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 $\operatorname{rank}[A|\boldsymbol{0}] = \operatorname{rank} A = n$ である. 階数の定義から A がいかなる行列であろうと $\operatorname{rank}[A|\boldsymbol{0}] = \operatorname{rank} A$ であり, したがって条件 $\operatorname{rank} A = n$ が $A\boldsymbol{x} = \boldsymbol{0}$ 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 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ が線形独立であることと, $\operatorname{rank}[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n] = n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は線形従属である.

Proof. (m, n)-行列 $A = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]$ の階数は A の行の数 m 以下である (式 6.1.1). ゆえに $\operatorname{rank} A \leq m < n$ であり, とくに $\operatorname{rank} A \neq n$. 系 17.3.4 より $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は線形従属である. □

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, (2) A の各行は線形独立である.

Proof. (1)⇔(2): 定理 6.2.2 および命題 17.3.1 から直ちに得られる:

$$A \text{ は可逆} \iff A\boldsymbol{x} = \boldsymbol{0} \text{ は唯一解をもつ} \iff A \text{ の各列は線形独立}.$$

(1)⇔(3): いま示した (1) と (2) の同値性および $|A| = |{}^tA|$ より得られる:

$$\begin{aligned} A \text{ は可逆} &\iff |A| \neq 0 \iff |{}^tA| \neq 0 \iff {}^tA \text{ は可逆} \\ &\iff {}^tA \text{ の各列は線形独立} \iff A \text{ の各行は線形独立}. \end{aligned}$$

□

18　基底

17.1 項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節で与えた．本節では要不要論の (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 $\boldsymbol{u}_1, \cdots .\boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$ において, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ と書く．すなわち, $\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle := \{ \sum_{i=1}^n a_i \boldsymbol{u}_i | a_1, \cdots, a_n \in \mathbb{R} \}$ である．次の命題により, これは V の部分空間となる．$\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ は, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$ において, $\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ 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．$\boldsymbol{0}$ は $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ たちの線形結合で書けるゆえ $\boldsymbol{0} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ である．次に, $\boldsymbol{x}, \boldsymbol{y} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ とすれば $\boldsymbol{x} = \sum_{i=1}^n a_i \boldsymbol{u}_i$, $\boldsymbol{y} = \sum_{i=1}^n b_i \boldsymbol{u}_i$ と書ける．このとき, 各 $r, s \in \mathbb{R}$ に対して $r\boldsymbol{x} + s\boldsymbol{y} = \sum_{i=1}^n (ra_i + sb_i)\boldsymbol{u}_i$ ゆえ $r\boldsymbol{x} + s\boldsymbol{y}$ も $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ たちの線形結合で書ける．よって $r\boldsymbol{x} + s\boldsymbol{y} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ 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 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell \in V$ が組 $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \in V$ の線形結合で書けるならば, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell$ の線形結合で書ける元は $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ の線形結合で書ける．すなわち,

$$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell \in \langle \boldsymbol{w}_1, \cdots, \boldsymbol{w}_n \rangle \implies \langle \boldsymbol{v}_1, \cdots, \boldsymbol{v}_\ell \rangle \subset \langle \boldsymbol{w}_1, \cdots, \boldsymbol{w}_n \rangle.$$

Proof. $W = \langle \boldsymbol{w}_1, \cdots, \boldsymbol{w}_n \rangle$ とおく．W は V の部分空間であり, 命題 16.1.6 より W は線形結合について閉じている．仮定より $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell \in W$ であるから, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell$ の線形結合で書ける元は W に含まれる．ゆえに $\langle \boldsymbol{v}_1, \cdots, \boldsymbol{v}_\ell \rangle \subset \langle \boldsymbol{w}_1, \cdots, \boldsymbol{w}_n \rangle$．□

上の証明では, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell$ の線形結合で書いたときに現れる係数と $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ 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 項を見よ)．

練習 18.1.4. $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ とし, $\langle \boldsymbol{v}_1, \cdots, \boldsymbol{v}_k \rangle = \mathbb{R}^n$ であるとする．また, A を (m, n)-行列とする．このとき, 各 $i = 1, \cdots, k$ について $A\boldsymbol{v}_i = \boldsymbol{0}$ 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 $A\boldsymbol{e}_j$ である．ここで, $\boldsymbol{e}_j \in \mathbb{R}^n = \langle \boldsymbol{v}_1, \cdots, \boldsymbol{v}_k \rangle$ より $\boldsymbol{e}_j$ は $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$ の線形結合で書ける．つまり $\boldsymbol{e}_j = \sum_{i=1}^k r_i \boldsymbol{v}_i$ と表せる．ゆえに $A\boldsymbol{e}_j = A\left(\sum_{i=1}^k r_i \boldsymbol{v}_i\right) = \sum_{i=1}^k A(r_i \boldsymbol{v}_i) = \sum_{i=1}^k r_i A\boldsymbol{v}_i = \sum_{i=1}^k r_i \boldsymbol{0} = \boldsymbol{0}$.

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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が $\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle = V$ を満たすとき, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 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 $\mathbb{R}^n$ の基底である.

(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 $1 \leq i \leq m$, $1 \leq n \leq n$), mn 個の行列の組 E_{ij} $(i = 1, \ldots, m,\ j = 1, \ldots, n)$ 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 $x^n, x^{n-1}, \cdots, x, \boldsymbol{1}$ 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 $\{\boldsymbol{0}\}$ 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 $\boldsymbol{v}$ について, 基底による $\boldsymbol{v}$ の線形結合表示に現れる係数を $\boldsymbol{v}$ の位置情報と見なすのである. とくに $\mathbb{R}^n$ の各元を標準基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 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. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$ が線形独立であるとし, $\boldsymbol{v} \in V$ 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, $\boldsymbol{v} = \sum_{i=1}^n a_i \boldsymbol{u}_i = \sum_{i=1}^n b_i \boldsymbol{u}_i$ ならば $a_i = b_i$ $(i = 1, \cdots, n)$ である.

Proof. $\sum_{i=1}^n a_i \boldsymbol{u}_i = \sum_{i=1}^n b_i \boldsymbol{u}_i$ を移項すると線形関係 $\sum_{i=1}^n (a_i - b_i)\boldsymbol{u}_i = \boldsymbol{0}$ を得る. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形独立性より $a_1 - b_1 = \cdots = a_n - b_n = 0$. つまり $a_i = b_i$ $(i = 1, \cdots, n)$ 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れ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論の後半において重要な役割を担い, その一般論は 25 節で述べる.

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, 本論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 $\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\boldsymbol{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ は線形独立である. $\mathbb{R}^2$ の任意のベクトル $\boldsymbol{v}$ は標準ベクトル $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ の線形結合でかけ, また $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ は $\boldsymbol{u}_1, \boldsymbol{u}_2$ の線形結合で書ける. 実際, $\boldsymbol{e}_1 = \boldsymbol{u}_1 - \boldsymbol{u}_2$, $\boldsymbol{e}_2 = \boldsymbol{u}_2$ である. このことは, $\boldsymbol{v}$ が $\boldsymbol{u}_1, \boldsymbol{u}_2$ 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 $\boldsymbol{u}_1, \boldsymbol{u}_2$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in V$ が線形独立であるとし, さらに $\boldsymbol{v} \in V$ とすれば次が成り立つ.

(1) $\boldsymbol{v}$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合で表せないならば, ベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n, \boldsymbol{v}$ は線形独立である.

(2) ベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n, \boldsymbol{v}$ は線形従属ならば, $\boldsymbol{v}$ は $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\sum_{i=1}^n a_i\boldsymbol{u}_i + a_{n+1}\boldsymbol{v} = \boldsymbol{0}$ 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 $a_{n+1} \neq 0$ ならば移項により $\boldsymbol{v} = \sum_{i=i}^n \frac{-a_i}{a_{n+1}}\boldsymbol{u}_i$ となり，これは $\boldsymbol{v}$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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，$\sum_{i=1}^n a_i\boldsymbol{u}_i = \boldsymbol{0}$ を得る．$\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形独立性より $a_1 = \cdots = a_n = 0$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合で書けないベクトル $\boldsymbol{v} \in \mathbb{R}^n$ があるとすれば，$n+1$ 個の組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n, \boldsymbol{v} \in \mathbb{R}^n$ 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 $\boldsymbol{v} \in \mathbb{R}^n$ は $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{b} \in \mathbb{R}^n$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合で表せることを示そう．$\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は線形独立ゆえ，定理 17.3.6 より正方行列 $A = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]$ は可逆である．したがって連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ は唯一解 $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ を持つ．$A^{-1}\boldsymbol{b} = {}^t[x_1, \cdots, x_n]$ と成分表示すれば，$x_1\boldsymbol{u}_1 + \cdots + x_n\boldsymbol{u}_n = \boldsymbol{b}$ である．

よりみち (反比例グラフと双曲線).

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 次形式の正規化を通して理解されるのであるが, ここでは結論だけを述べると

$$\begin{bmatrix} x \\ y \end{bmatrix} = A \begin{bmatrix} X \\ Y \end{bmatrix}, \qquad \text{ただし} \quad A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{18.2.1}$$

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

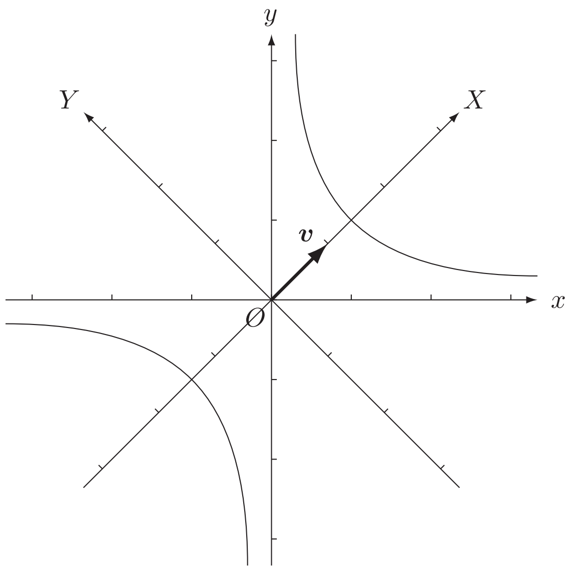
$$xy = \left(\frac{1}{\sqrt{2}}X - \frac{1}{\sqrt{2}}Y\right)\left(\frac{1}{\sqrt{2}}X + \frac{1}{\sqrt{2}}Y\right) = \frac{1}{2}X^2 - \frac{1}{2}Y^2.$$

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,

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{のとき,} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{のとき,} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

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 $\boldsymbol{v}$ 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 $\boldsymbol{v}$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ によって生成されているとすれば, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_{n_1}$ として $\boldsymbol{u}_1$ を取る. この $\boldsymbol{u}_{n_1}$ が V を生成するならば, $\boldsymbol{u}_{n_1}$ は V の基底である. そうでない場合は $\boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ のうち $\boldsymbol{u}_{n_1}$ のスカラー倍で表せないものがある. 何故なら, もし $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m \in \langle \boldsymbol{u}_{n_1} \rangle$ とすれば命題 18.1.3 より $V = \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle \subset \langle \boldsymbol{u}_{n_1} \rangle$ であり, V は $\boldsymbol{u}_{n_1}$ によって生成されてしまう. そこで, $\boldsymbol{u}_{n_1}$ のスカラー倍で表せないベクトルを仮に $\boldsymbol{u}_{n_2}$ とすれば補題 18.2.5(1) より $\boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2}$ は線形独立である. $\boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2}$ 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 $\boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2}$ を除いたベクトルのうちいずれかは $\boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2}$ の線形結合で書けない. 何故なら, もし $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ すべてが $\boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2}$ の線形結合で書けるとすると, 命題 18.1.3 より $V = \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle \subset \langle \boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2} \rangle$ であり, V は $\boldsymbol{u}_{n_1} \boldsymbol{u}_{n_2}$ によって生成されてしまう. ゆえに $\boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2}$ の線形結合で表せないベクトルがあり, これを $\boldsymbol{u}_{n_3}$ 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 $\boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2}, \boldsymbol{u}_{n_3}, \cdots, \boldsymbol{u}_{n_k}$ (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 $\boldsymbol{u}_1, \ldots \boldsymbol{u}_m$ 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 $\boldsymbol{u}_{n_1}, \boldsymbol{u}_{n_2}, \boldsymbol{u}_{n_3}, \cdots, \boldsymbol{u}_{n_k}$ 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 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m \in V$ 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, $V = \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ と表せる. このとき,

$$V = \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_m \rangle$$

でもあることに注意して, $n+m$ 個の組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ に対して命題 18.3.1 の証明を適用しよう. このとき先の証明において始めに $\boldsymbol{u}_1$ を選ぶところで, 代わりに $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ を選び取ってしまえばよい. すなわち, $\boldsymbol{u}_{n_i} := \boldsymbol{v}_i$ $(i = 1, \cdots, m)$ としたうえで, 命題 18.3.1 の証明で述べた手続きを進めれば求める基底が得られる. □

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 $A = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_5]$ について (例題 17.3.3 と同じもの), 次の問いに答えよ.

$$\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \boldsymbol{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \boldsymbol{a}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{a}_5 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ -3 \\ -1 \end{bmatrix}.$$

(1) $\langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_5 \rangle$ の基底を求めよ.

解答例: $A = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_5]$ の簡約化を $B = [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_5]$ とする (簡約化は例題 17.3.3 で行った). このとき $\langle \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_4 \rangle = \langle \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4, \boldsymbol{b}_5 \rangle$ である. よって命題 17.3.2 より $\langle \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_4 \rangle = \langle \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5 \rangle$ となる. また, 組 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_4$ 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 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_4$ も線形独立であり, これらは $\langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_5 \rangle$ の基底となる.

(2) 方程式 $A\boldsymbol{x} = \boldsymbol{0}$ 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 $B\boldsymbol{x}=\mathbf{0}$ 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:

$$W_A = \{\, c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 \mid c_1, c_2 \in \mathbb{R} \,\}. \quad (\text{ただし}, \ \boldsymbol{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \\ 1 \end{bmatrix})$$

ゆえに $W_A = \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$ である. $\boldsymbol{u}_1, \boldsymbol{u}_2$ 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 $\boldsymbol{u}_1, \boldsymbol{u}_2$ の $1, 2, 4$ 行を目隠しして

$$\boldsymbol{u}_1 = \begin{bmatrix} * \\ * \\ 1 \\ * \\ 0 \end{bmatrix}, \quad \boldsymbol{u}_2 = \begin{bmatrix} * \\ * \\ 0 \\ * \\ 1 \end{bmatrix}$$

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 $\boldsymbol{u}_1, \boldsymbol{u}_2$ 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 $A = [\boldsymbol{a}_1, \cdots \boldsymbol{a}_n]$ について次が成り立つ.

(1) $\mathbb{R}^m$ の部分空間 $\langle \boldsymbol{a}_1, \cdots \boldsymbol{a}_n \rangle$ は $\operatorname{rank} A$ 個のベクトルからなる基底を持つ.

(2) $A\boldsymbol{x}=\mathbf{0}$ の解空間 $W_A \subset \mathbb{R}^n$ は $n - \operatorname{rank} A$ 個のベクトルからなる基底を持つ.

Proof. $k = \operatorname{rank} A$ とおき, A の簡約化を $B = [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n]$, B の各列のうち主成分を持つ列を $\boldsymbol{b}_{n_1}, \ldots, \boldsymbol{b}_{n_k}$ とし, 主成分を持たない列を $\boldsymbol{b}_{r_1}, \ldots, \boldsymbol{b}_{r_{n-k}}$ とする.

(1): ベクトルの組 $\boldsymbol{b}_{n_1}, \ldots, \boldsymbol{b}_{n_k}$ 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 $\boldsymbol{b}_{n_1}, \ldots, \boldsymbol{b}_{n_k}$ の線形結合で書ける. 命題 17.3.2 より組 $\boldsymbol{a}_{n_1}, \ldots, \boldsymbol{a}_{n_k}$ は線形独立であり, A の各列は $\boldsymbol{a}_{n_1}, \ldots, \boldsymbol{a}_{n_k}$ の線形結合で書ける. すなわち, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \langle \boldsymbol{a}_{n_1}, \ldots, \boldsymbol{a}_{n_k} \rangle$. これと命題 18.1.3 を合わせて $\langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \rangle \subset \langle \boldsymbol{a}_{n_1}, \ldots, \boldsymbol{a}_{n_k} \rangle$ を得る. 以上より $\boldsymbol{a}_{n_1}, \ldots, \boldsymbol{a}_{n_k}$ は $\langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \rangle$ の基底となる.

(2): 連立 1 次方程式 $A\boldsymbol{x}=\mathbf{0}$ 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 $\boldsymbol{x} = \sum_{j=1}^{n-k} c_j \boldsymbol{u}_j$ と書ける. つまり $W_A = \langle \boldsymbol{u}_1, \cdots \boldsymbol{u}_{n-k} \rangle$ である. ベクトルの組 $\boldsymbol{u}_1, \cdots \boldsymbol{u}_{n-k}$ が線形独立であることを示すために, 各 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-k}$ の r_j 成分 $(j = 1, \cdots, n-k)$ に注目しよう. B の第 r_j 列 $\boldsymbol{b}_{r_j}$ 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, $\boldsymbol{u}_j$ の r_j 成分は 1 であり, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-k}$ のうち $\boldsymbol{u}_j$ 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 $\boldsymbol{b}_{r_1} = \boldsymbol{b}_3$, $\boldsymbol{b}_{r_2} = \boldsymbol{b}_5$ となる. 確かに $\boldsymbol{u}_1$ の $r_1 = 3$ 成分は 1, $r_2 = 5$ 成分は 0 であり, $\boldsymbol{u}_2$ 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 $\sum_{j=1}^{n-k} c_j \boldsymbol{u}_j = \mathbf{0}$ を与えると, 各 $j = 1, \cdots, n-k$ 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 $c_j = 0$ $(j = 1, \cdots, n-k)$ を得る. すなわち, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-k}$ 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. $\square$

本項や 17.3.1 項では, 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ に関する条件を $\mathbb{R}^n$ の標準基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 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. そこで次節からしばらくの間, 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

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.

例 18.4.1. 多項式環 $\mathbb{R}[x]$ において, 有限個の多項式の組 $f_1, \cdots, f_n$ が $\mathbb{R}[x]$ を生成することはない. なぜなら, 各多項式 $f_1, \cdots, f_n$ の中で最も高い次数を m とすれば, $m+1$ 次多項式を $f_1, \cdots, f_n$ 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 $\boldsymbol{a} \in A$ が取れる. このとき $\mathbf{0}_V = 0\boldsymbol{a}$ 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, $\boldsymbol{x}, \boldsymbol{y} \in \langle A \rangle$ とすれば, $\boldsymbol{x}, \boldsymbol{y}$ は A の元の線形結合で書ける. ここで, $\boldsymbol{x}, \boldsymbol{y}$ 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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ を用いて $\boldsymbol{x} = \sum_{i=1}^n x_i \boldsymbol{a}_i$, $\boldsymbol{y} = \sum_{i=1}^n y_i \boldsymbol{a}_i$ と書くことができる. このとき, 各 $r, s \in \mathbb{R}$ について $r\boldsymbol{x} + r\boldsymbol{y} = \sum_{i=1}^n (rx_i + sy_i)\boldsymbol{a}_i$ であるから $r\boldsymbol{x} + r\boldsymbol{y}$ は A の元の線形結合で書ける. ゆえに $r\boldsymbol{x} + s\boldsymbol{y} \in \langle A \rangle$. 以上より $\langle A \rangle$ は部分空間である. □

ベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および集合 $A = \{ \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \}$ について, 次が成り立つ:

- $\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle = \langle A \rangle$.
- ベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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, $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3 \in \mathbb{R}^2$ を $\boldsymbol{u}_1 := \boldsymbol{e}_1, \boldsymbol{u}_2 := \boldsymbol{e}_2, \boldsymbol{u}_3 := \boldsymbol{e}_2$ と定めれば, ベクトルの組 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ は線形独立ではない. しかしながら, $A = \{ \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3 \} = \{ \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_2 \} = \{ \boldsymbol{e}_1, \boldsymbol{e}_2 \}$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{v} \in \langle W \rangle$ 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 16.1.6) ゆえ $\boldsymbol{v} \in W$.

(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 $\boldsymbol{e}_n$ と書けば, $A = \{\, \boldsymbol{e}_n \mid n \in \mathbb{N} \,\}$ 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 $\boldsymbol{x} = (1, 1, \ldots)$ 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[46]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, 本論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[47]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.

[46]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.
[47]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.

発展 (最小の部分空間).

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 $\{\mathbf{0}\}$ を生成することになり, $\emptyset$ を $\{\mathbf{0}\}$ 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 $\bigcap$ 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,

$$\bigcap \mathcal{W} := \{\, \boldsymbol{v} \in V \mid \text{各 } W \in \mathcal{W} \text{ について } \boldsymbol{v} \in W \,\}$$

$$(\text{つまり}, \boldsymbol{v} \in \bigcap \mathcal{W} \iff \text{各 } W \in \mathcal{W} \text{ について } \boldsymbol{v} \in W).$$

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. 何故なら, 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 $\mathbf{0}_V$ を含むゆえ $\mathbf{0}_V \in U$ である. 次に $\boldsymbol{x}, \boldsymbol{y} \in U$ とすれば $r\boldsymbol{x} + s\boldsymbol{y} \in U$ となることを示そう. そのためには各 $W \in \mathcal{W}$ について $r\boldsymbol{x} + s\boldsymbol{y} \in W$ を示せばよい. $\boldsymbol{x}, \boldsymbol{y} \in U$ より各 $W \in \mathcal{W}$ において $\boldsymbol{x}, \boldsymbol{y} \in W$ であり, W が部分空間であることから $r\boldsymbol{x} + s\boldsymbol{y} \in W$ を得る. したがって $r\boldsymbol{x} + s\boldsymbol{y} \in U$. 部分空間となるための条件 (i) と (iv) が示されたゆえ, W 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 $U \subset \langle A \rangle$. □

19　写像概念の基礎

写像に関する概念のいくつかを述べる．これらの必要性は 14 節の冒頭で述べた通りである．本節では概念をひたすら提示することに終始するゆえ, 読者はやや退屈に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項末で述べた．

A を (m,n)-行列 A とする．次で定められる写像 T_A は本論全体を通して何度も論じられる:

$$T_A : \mathbb{R}^n \to \mathbb{R}^m, \qquad T_A(\boldsymbol{x}) := A\boldsymbol{x}.$$

19.1　像と逆像

集合 X から集合 Y への写像 $f : X \to Y$ 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) とも呼ばれる．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, 本論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, 本論では像と呼ぶ:

定義 19.1.1. 写像 $f : X \to Y$ 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$\boldsymbol{f}$**による**$\boldsymbol{A}$**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$\boldsymbol{f}$**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) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$.　解答: $f(\mathbb{R}) = [0, \infty)$.

(2) $f : [0,2] \to \mathbb{R}$, $f(x) = x^2$.　解答: $f([0,2]) = [0,4]$.

(3) $f : [-2,2] \to [0,4]$, $f(x) = x^2$.　解答: $f([-2,2]) = [0,4]$.

(4) $f : [0,2] \to [0,4]$, $f(x) = x^2$.　解答: $f([0,2]) = [0,4]$.

定義 19.1.3. 写像 $f : X \to Y$ 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$\boldsymbol{f}$**による**$\boldsymbol{B}$**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 $f : \mathbb{R} \to \mathbb{R}$ 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 $f, g : X \to Y$ 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

$$\text{各 } x \in X \text{ について } f(x) = g(x) \tag{19.1.1}$$

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 $\boldsymbol{f}$ と $\boldsymbol{g}$ 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 $f(x) = g(x)$ とも書く. 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 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 $f : U \to V$ のことを $f = \mathbf{0}_V$ と書く.

定値写像 $f : X \to Y$ $(f = b)$ 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, $\operatorname{Im} f$ 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 $\operatorname{Ker} f$ が用いられる. 行列 A の階数が $\operatorname{Im} T_A$ の次元に一致すること, および同次形連立 1 次方程式 $A\boldsymbol{x} = \mathbf{0}$ の解空間が $\operatorname{Ker} T_A$ 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[48] 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, 2 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. $f : X \to Y$ 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[49].

- **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** (あるいは **1 対 1 の対応, bijection**) であるという.

---

[48] 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.

[49] 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

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 を的の位置 (標的) を表す集合とし, f を弓であると考える. いま, 弓 f を一つ固定し, 写像 $f : X \to Y$ 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 $f : X \to X$ 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) $f : \mathbb{R} \to \mathbb{R}$ 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. □

(2) $f : [0, 2] \to \mathbb{R}$ 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. □

(3) $f : [-2, 2] \to [0, 4]$ 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. □

(4) $f : [0, 2] \to [0, 4]$ 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. □

定義 19.2.4. 写像 $f : X \to Y$ 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 $f|_A : A \to Y$ と書く. また, 部分集合を定義域とする写像 $g : A \to Y$ に対して, 新たに定めた $\widetilde{g} : X \to Y$ が $\widetilde{g}|_A = g$ を満たすとき, $\widetilde{g}$ は g の**拡張 (extension)** であるという.

任意の写像 $f : X \to Y$ に対して, 適当な部分集合 $A \subset X$ に f を制限すれば $f|_A$ は単射となる. 例えば, A を一点集合とすればよい. あるいはもっと極端に $A = \emptyset$ とすれば何も元を対応させない単射 $f|_\emptyset$ を得る. ほかにも, 各 $y \in f(X)$ に対して $f^{-1}(y)$ の元を一つだけ取ってこれを x_y とし, $A = \{\, x_y \mid y \in f(X) \,\}$ とすれば $f|_A$ は, その像が f の像に一致する単射となる[50]. 一方, $f : X \to Y$ の終域を $f(X)$ に置き換えれば, 写像 $f : X \to f(X)$ は全射となる. とくに単射 $f : X \to Y$ において終域を置き換えた写像 $f : X \to f(X)$ は全単射である. このことから, 単射のことを 1 対 1 写像と呼ぶ文献もある[51].

練習 19.2.5. 写像 $f : X \to Y$ 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 $f^{-1}\Big(f(A)\Big) = A$.

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 $a \in f^{-1}\Big(f(A)\Big)$. つまり $A \subset f^{-1}\Big(f(A)\Big)$ である. 次に, 任意に $\alpha \in f^{-1}\Big(f(A)\Big)$ 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 $f^{-1}\Big(f(A)\Big) \subset A$. □

(2) f が全射ならば, 部分集合 $B \subset Y$ について $f\Big(f^{-1}(B)\Big) = B$.

解答例: $\beta \in f\Big(f^{-1}(B)\Big)$ 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 $f\Big(f^{-1}(B)\Big) \subset B$. 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 $b = f(x') \in f\Big(f^{-1}(B)\Big)$. したがって $B \subset f\Big(f^{-1}(B)\Big)$. □

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, $f : \mathbb{R} \to \mathbb{R}$ を $f(x) := x^2$ と定めれば, $A = \{2\}$ および $B = [-5, 3]$ について

$$f^{-1}\Big(f(A)\Big) = f^{-1}\Big(f(\{2\})\Big) = f^{-1}(4) = \{2, -2\} \neq A,$$
$$f\Big(f^{-1}(B)\Big) = f\Big(f^{-1}([-5, 3])\Big) = f([0, \sqrt{3}]) = [0, 3] \neq B.$$

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 $A\boldsymbol{x} = \boldsymbol{0}$ 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 および命題 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, 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.

19.3 逆写像はいつ定まるか

具体例を通して逆写像 (逆関数) の存在性について論じよう. 次は関数 $f(x) = 2^x$ を表に記したものである.

x	$\cdots$	-4	-3	-2	-1	0	1	2	3	4	$\cdots$
$f(x)$	$\cdots$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	1	2	4	8	16	$\cdots$

ここで上表の上段と下段を入れ替えた次のような対応 $g(x)$ を考える.

[50] 公理的集合論において, $f|_A(A) = f(X)$ を満たす $A \subset X$ の存在を主張する命題は選択公理と呼ばれる (正確には選択公理と同値な主張となる).
[51] 単射を「1 対 1 写像」と呼び, 全単射を「1 対 1 対応」と呼んで区別している. 「写像」と「対応」を同義語として用いる立場では, このような呼称は誤解を与えるかもしれない.

x	$\cdots$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	1	2	4	8	16	$\cdots$
$g(x)$	$\cdots$	-4	-3	-2	-1	0	1	2	3	4	$\cdots$

これは置換に対して逆置換を対応させる手順と同等のことを行っている．この表により定められる関数 $g(x)$ を $f(x)$ の**逆関数**と呼ぶ．ちなみに，この $g(x)$ は対数関数と呼ばれ，$\log_2 x$ と書くのであった．

注意すべきことは，逆関数は常に定まるわけではないということである．例えば $f(x) = x^2$ について考えよう．これを表にすると次のようになる:

x	$\cdots$	-4	-3	-2	-1	0	1	2	3	4	$\cdots$
$f(x)$	$\cdots$	16	9	4	1	0	1	4	9	16	$\cdots$

先ほどと同様に上下を入れ替えた対応を考えようとすれば次の表が得られる:

x	$\cdots$	16	9	4	1	0	1	4	9	16	$\cdots$
$g(x)$	$\cdots$	-4	-3	-2	-1	0	1	2	3	4	$\cdots$

しかし，上の表を満足するような関数 $g(x)$ を定めることはできない．なぜなら，上の表では $g(16)$ の値が -4 と 4 の二重に指定されており，一つの値に定まらないからである．さらに $g(9)$, $g(4)$, $g(1)$ についても同様のことが言える．このような状況が生じる背景には，もとの関数 f が次の性質を持つことにある:

(*)　　$f(x_1) = f(x_2)$ を満たすような定義域の二つの元 x_1, x_2 が存在する．

性質 (*) は f が単射でないことと同値である．逆関数を定めるためには上の条件 (*) が成立しない必要がある．すなわち，少なくとも f は単射でなければならない．

単射でない関数においては定義域を制限することで単射にし，制限した関数の逆関数を考えることがある[52]．例えば次の二つの関数 f_1, f_2 を考えよう．

x	0	1	2	3	4	$\cdots$
$f_1(x)$	0	1	4	9	16	$\cdots$

x	$\cdots$	-4	-3	-2	-1	0
$f_2(x)$	$\cdots$	16	9	4	1	0

これらはともに $x \mapsto x^2$ なる対応であるが，定義域を f_1 では 0 以上の数に制限しており，f_2 では 0 以下の数に制限している．このため f_1, f_2 は条件 (*) を満たさず，したがって次の表により逆関数 g_1, g_2 がそれぞれ定められる:

x	0	1	4	9	16	$\cdots$
$g_1(x)$	0	1	2	3	4	$\cdots$

x	$\cdots$	16	9	4	1	0
$g_2(x)$	$\cdots$	-4	-3	-2	-1	0

g_1, g_2 を式で書けば，$g_1(x) = \sqrt{x}$, $g_2(x) = -\sqrt{x}$ である．

19.4　逆写像とその性質

前項における考察から，次のことが分かった:

- 任意の関数に対して逆関数が定まるわけではない．
- 逆関数が定まらない場合でも，定義域を制限した関数においては逆関数が定まる場合がある．
- 同じ式で定義される関数であっても，定義域の定め方次第で逆関数の式は異なる．

[52] 例えば，三角関数の逆関数はこのようにして定められる．

さて, 前項の議論において $f: X \to Y$ の終域 Y に関してはあえて言及していなかった. f の逆関数の定義域が終域全体となるためには, 当然 f が全射である必要がある. したがって, $f: X \to Y$ が逆関数 $g: X \to Y$ を持つためには f が全単射でなければならない. 写像 $f: X \to Y$ が全射かつ単射であるとは, 次を満たすことであった:

(※)　　いかなる $y \in Y$ に対しても, $f(x_y) = y$ を満たす $x_y \in X$ が唯一つ存在する.

ここで, あらためて逆写像の正確な定義を与えておこう:

定義 19.4.1. 写像 $f: X \to Y$ が全単射であるとき, 上の (※) を f は満たす. そこで, 各 $y \in Y$ に対して上の x_y (つまり f に代入すると y になる定義域の元のこと) を対応させる写像 $g: Y \to X$ $(g(y) := x_y)$ を f の**逆写像** (**inverse map**) とよび f^{-1} と書く.

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. この事実の確認は前項における議論で十分であるが, 命題 19.4.4 にて一般論として改めて述べよう. 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 $f: X \to Y$ 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る. 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 $f^{-1}: Y \to X$ に y を代入した値のことである. 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 19.4.3).

例 19.4.2. (1) 写像 $f: \mathbb{R} \to \mathbb{R}$ 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 $f: [0, \infty) \to [0, \infty)$ 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 $f: (-\infty, 0] \to [0, \infty)$ 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関数に代入する変数も文字 x を用いることになる. 例えば $f(x) = ax + b$ の逆関数は $f^{-1}(x) = \frac{1}{a}x - \frac{b}{a}$ と書かれる. この場合, $f: X \to Y$ の逆関数 $f^{-1}: Y \to X$ 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.

練習 19.4.3. $f: X \to Y$ 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 $p = f^{-1}\bigl(f(p)\bigr)$ ゆえ, B による f^{-1} の像 I に p は含まれる. □

次は定義 19.4.1 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 19.4.4. $f: X \to Y$ が全単射であることと, 次の性質 (1), (2) を満たす $g: Y \to X$ 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) いかなる $x \in X$ についても $g \circ f(x) = x$ が成り立つ (すなわち $g \circ f = \mathrm{id}_X$).

(2) いかなる $y \in f(X)$ についても $f \circ g(y) = y$ が成り立つ (すなわち $f \circ g = \mathrm{id}_Y$).

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 $x \in X$ を勝手に取る. $y := f(x) \in Y$ とおこう. 逆写像の定義よれば, $g(y)$ 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 $g(y) = x$ である. つまり, $g \circ f(x) = g((f(x)) = g(y) = x$. 次に (2) を示すために $y \in Y$ を勝手に取る. 逆写像の定義によれば, $g(y)$ は f に代入すると y になる元のことである. つまり $f(g(y)) = y$, すなわち $f \circ g(y) = y$.

次に条件 (1) と (2) を満たす g が存在するとき, f が全単射であることを示そう. ここでは対偶をとり, f の単射性と全射性のいずれか一方でも欠けると (1) かつ (2) を満たす g が作れないことを導く. f が単射でないとすると, ある $y \in f(X)$ について $f(x_1) = f(x_2) = y$ を満たす二つの異なる元 $x_1, x_2 \in X$ が存在する. 条件 (1) を満たす g を定義しようとするとき, $g(y)$ は x_1 と定めるべきだろうか, それとも x_2 と定めるべきだろうか. 前者を採用すると $g \circ f(x_2) = g(f(x_2)) = g(y) = x_1$ となり, g は (1) を満たさない. 後者を採用しても $g \circ f(x_1) = x_2$ となり, やはり (1) が満たされない. もちろん, x_1, x_2 以外の点 x_3 を採用しても $g \circ f(x_1) = x_3$ となり, やはり (1) は満たされない. すなわち, (1) を満たすように g を定めることは出来ないことが分かる. f が全射でない場合は, f に代入した値にはなり得ない $y \in Y$ が存在しており, このとき $g : Y \to X$ をどう定義するにしても条件 (2) を満たすことはない. 何故なら, $f \circ g(y) = f(g(y))$, つまり $f \circ g(y)$ は f に $g(y)$ を代入した値である. ゆえに, f に代入した値にはなり得ない元 y と $f \circ g(y)$ は異なり, $f \circ g(y) \neq y$.

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, $f \circ g(y) = y = f \circ f^{-1}(y)$ となる. すなわち $f(g(y)) = f(f^{-1}(y))$ であり, f の単射性より $g(y) = f^{-1}(y)$. つまり, いかなる Y の元についても, g および f^{-1} で写した値が等しいゆえ $g = f^{-1}$ である. □

上の証明の第二段落では, 対偶をとらずに (1) と (2) から直接 f の全単射性を示すこともできる (練習 19.4.7). また, そのほうが証明はエレガントである. にも関わらず, ここでは教育的配慮から対偶による証明を採用した. 上の証明の第二段落は, 全単射でない写像が逆写像 (すなわち (1) と (2) を満たす写像) を持たない理由の具体的な説明になっている.

備考 19.4.5. $f : X \to X$ が全単射であることの必要十分条件は, $f \circ g = \mathrm{id}_X = g \circ f$ を満たす $g : X \to X$ 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 19.4.6. 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 19.4.4 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 19.4.7 (発展)**.** 命題 19.4.4 の (1) および (2) を満たす g が存在するならば f が全単射となることを直接証明せよ.

Proof. 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 $y = f \circ g(y) = f(g(y)) = f(x)$. すなわち, f に代入すると y となる元 $x \in X$ が存在する. □

グラフの対称性についても述べておこう. $A, B \subset \mathbb{R}$ および関数 $f : A \to B$ に対して, 次で与えられる $\mathbb{R}^2$ の部分集合を f の**グラフ**という:

$$\Gamma_f := \{ (x, f(x)) \mid x \in A \}.$$

命題 19.4.8 (発展)**.** X, Y を $\mathbb{R}$ の部分集合, $f : X \to Y$ を全単射とし, 更に関数 $y = x$ が表す $\mathbb{R}^2$ の対角線を L とする. このとき, f のグラフと f^{-1} のグラフは, 直線 L を軸に線対称である.

Proof. 直線 L を軸に点 $(x, y) \in \mathbb{R}^2$ と対称な点は $(y, x) \in \mathbb{R}^2$ である. したがって, 次の同値性を示せばよい:

$$(x, y) \in \Gamma_f \iff (y, x) \in \Gamma_{f^{-1}}$$

$(\Rightarrow)$ を示すために $(x, y) \in \Gamma_f$ を取れば, $y = f(x)$ である. ゆえに $f^{-1}(y) = x$ であり $(y, x) = (y, f^{-1}(y))$ は f^{-1} のグラフ上の点である. したがって $(y, x) \in \Gamma_{f^{-1}}$.

いま全単射 $h : Y \to X$ に関する命題「$(b, a) \in \Gamma_h \implies (a, b) \in \Gamma_{h^{-1}}$」を示したと言ってもよい. この命題に $h = f^{-1}$, $(b, a) = (y, x)$ を適用することで, 「$(y, x) \in \Gamma_{f^{-1}} \implies (x, y) \in \Gamma_f$」を得る. □

関数 $f : \mathbb{R} \to \mathbb{R}$ のグラフ上の点 $(\alpha, f(\alpha))$ における接線 ℓ を表す関数が $g(x) = ax + b$ で与えられているとき, 逆関数 f^{-1} のグラフ上の点 $(f(\alpha), \alpha)$ の接線 ℓ' は, 直線 L を軸に ℓ と線対称である. このとき直線 ℓ' の式は g^{-1} で表される. 例 19.4.2 によれば g^{-1} の式は $g^{-1}(x) = \frac{1}{a}x - \frac{b}{a}$ であり, ゆえに ℓ' の傾きは $1/a$ となる. このことから点 $\beta = f(\alpha)$ における逆関数の微分公式 $(f^{-1})'(\beta) = \frac{1}{f'(\alpha)}$ が示唆される.

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に示すことは意外に難しい:

練習 19.4.9. $T_A : \mathbb{R}^n \to \mathbb{R}^m$ 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 $T_A(\boldsymbol{x}) = \mathbf{0}$ を満たす $\boldsymbol{x} \in \mathbb{R}^n$ は $\boldsymbol{x} = \mathbf{0}$ に限る. すなわち, 方程式 $A\boldsymbol{x} = \mathbf{0}$ は唯一解 $\boldsymbol{x} = \mathbf{0}$ を持つ. したがって命題 6.2.1(3) より $\operatorname{rank} A = n$ であり, これと式 (6.1.1) $\operatorname{rank} A \leq m$ を合わせて $n \leq m$ を得る.

各 $\boldsymbol{e}_i \in \mathbb{R}^m$ $(i = 1, \cdots, m)$ に対して, T_A の全射性より $A\boldsymbol{x}_i = \boldsymbol{e}_i$ を満たす $\boldsymbol{x}_i \in \mathbb{R}^n$ が存在する. ここで $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m$ 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 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m \in \mathbb{R}^n$ は線形従属でなければならない. $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m \in \mathbb{R}^n$ の線形独立性については各自で確かめよ (命題 20.1.7(3)).

19.5 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 19.5.1. 写像 $f : X \to Y$ および $g : Y \to Z$ が与えられているとする.

(1) f, g が共に単射ならば $g \circ f : X \to Z$ も単射である.

(2) f, g が共に全射ならば $g \circ f : X \to Z$ も全射である.

(3) f, g が共に全単射ならば $g \circ f : X \to Z$ 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. □

練習 19.4.7 で行った議論は次のように分解できる.

命題 19.5.2 (発展)**.** 写像 $f : X \to Y$ および $g : Y \to Z$ が与えられているとする.

(1) $g \circ f : X \to Z$ が単射ならば f も単射である.

(2) $g \circ f : X \to Z$ 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 19.5.3. 写像 $f : X \to Y$ および $g : Y \to Z$ 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 $(f^{-1} \circ g^{-1}) \circ (g \circ f) = \mathrm{id}_X$ が容易に確かめられる. 実際,

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ \big(g^{-1} \circ (g \circ f)\big) = f^{-1} \circ \big((g^{-1} \circ g) \circ f\big) = f^{-1} \circ \big(\mathrm{id}_Y \circ f\big) = f^{-1} \circ f = \mathrm{id}_X .$$

同様にしておよび $(g \circ f) \circ (f^{-1} \circ g^{-1}) = \mathrm{id}_Z$ も示され, ゆえに命題 19.4.4 より $f^{-1} \circ g^{-1}$ は $g \circ f$ の逆写像である. □

備考 19.5.4. 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 19.5.3 と類似の性質が成り立つ.

命題 19.5.5. 写像 $f : X \to Y$ および $g : Y \to Z$, 部分集合 $C \subset Z$ に対して, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$.

Proof. $x \in (g \circ f)^{-1}(C) \iff g \circ f(x) \in C \iff g(f(x)) \in C \iff f(x) \in g^{-1}(C) \iff x \in f^{-1}(g^{-1}(C))$. □

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:

$$(h \circ g \circ f)^{-1}(D) = \big(h \circ (g \circ f)\big)^{-1}(D) = (g \circ f)^{-1}\big(h^{-1}(D)\big) = f^{-1}\Big(g^{-1}\big(h^{-1}(D)\big)\Big).$$

以降で学ぶこと

19.2 で予告したことと命題 19.5.2 を合わせると直ちに次が導かれる:

命題 19.5.6. (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 19.5.7. 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 $B\boldsymbol{x} = \boldsymbol{0}$ が唯一解をもつことを示そう. ベクトル $\boldsymbol{x}$ を方程式 $B\boldsymbol{x} = \boldsymbol{0}$ の解とすれば, $AB\boldsymbol{x} = A(B\boldsymbol{x}) = A\boldsymbol{0} = \boldsymbol{0}$ ゆえ $\boldsymbol{x}$ は $AB\boldsymbol{x} = \boldsymbol{0}$ 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 $AB\boldsymbol{x} = \boldsymbol{0}$ の解は唯一である. ゆえに $\boldsymbol{x} = \boldsymbol{0}$. 以上より, $B\boldsymbol{x} = \boldsymbol{0}$ 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. □

19.6 無限集合 (発展)

定義 8.1.1 をよく読むと, 置換を単射 $f : X_n \to X_n$ 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 19.6.1. X, Y をともに n 点からなる集合とする. 写像 $f : X \to Y$ 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. $X = \{ x_1, , \cdots, x_n \}$, $Y = \{ y_1, \cdots, y_n \}$ とおいて証明しよう.

(1)⇒(2): f は単射ゆえ $f(x_1), \cdots, f(x_n)$ の中に重複はない. もし f が全射でないと仮定すれば, $f(x_1), \cdots, f(x_n)$ 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 $f(x_1), \cdots, f(x_n), y$ 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)⇒(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 $f(x_1), \cdots, f(x_n)$ のうち $f(x_i)$ と $f(x_j)$ は等しいゆえ, 集合 $f(X) = \{ f(x_1), \cdots, f(x_n) \}$ 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. □

本論では「n 点からなる集合」の正確な定義には踏み込まずに素朴な立場で論じ, 構成する元の個数がちょうど n 個の集合を「n 点からなる集合」と定めよう[53]. この立場において次の命題は, 二つの集合の元の総数が一致することが全単射の存在によって特徴づけられることを述べている:

命題 19.6.2. (1) 集合 X, Y が n 点からなるとすると, 全単射 $f : X \to Y$ が存在する.

(2) X が n 点からなるとし, 全単射 $f : X \to Y$ があれば Y も n 点集合になる.

(3) Y が n 点からなるとし, 全単射 $g : X \to Y$ があれば X も n 点集合になる.

Proof. (1): X, Y がともに n 点からなるとすれば, 重複のない表示 $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ を用いて $X = \{ x_1, , \cdots, x_n \}$, $Y = \{ y_1, \cdots, y_n \}$ と書ける. このとき, $f : X \to Y$ を $f(x_i) := y_i$ $(i = 1, \ldots, n)$ と定めれば全単射である.

(2): X は n 点からなるゆえ重複のない表示 $x_1, \ldots, x_n$ を用いて $X = \{ x_1, \cdots, x_n \}$ と書ける. $f : X \to Y$ を全単射とし, $y_i := f(x_i)$ $(i = 1, \ldots, n)$ とおこう. f の全射性より $Y = f(X) = \{ f(x_1), \cdots, f(x_n) \} = \{ y_1, \cdots, y_n \}$ である. また, f の単射性は $y_1, \ldots, y_n$ の中に重複がないことを意味している. 実際, もし仮に重複があり $y_i = y_j$ $(i \neq j)$ となるならば, $f(x_i) = f(x_j)$ と単射性より $x_i = x_j$ となり, これは $x_1, \ldots, x_n$ に重複がないことに反する. 以上より, $Y = \{ y_1, \cdots, y_n \}$ は n 点からなる.

(3): $g^{-1} : Y \to X$ は全単射であり, $f = g^{-1}$ として (2) を適用すれば X は n 点集合である. □

そこで, 集合の元の総数が有限でない場合についても, 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 19.6.3. 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 19.6.4. 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 $f : X \to Y$ および $g : Y \to Z$ が存在する. このとき $g \circ f : X \to Z$ は命題 19.5.1(3) より全単射であり, したがって X と Z は対等である. □

有限集合の場合と異なり, 無限集合は自身の真部分集合[54]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 19.6.5. 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, $f : \mathbb{N} \to 2\mathbb{N}$ を $f(x) := 2x$ と定めればこれは全単射である.

一方で, 次に見るようにすべての無限集合が対等というわけではない. この命題の証明で用いた実数の連続性の詳細については解析学の本を参照せよ.

[53]そもそも n 点集合なる概念を集合論的な道具立てのみでいかに定義するかを考えると, 「集合 $X_n := \{ 1, \ldots, n \}$ との間に全単射がある集合」と定めるしかない. その意味において命題 19.6.2 は明らかであり, これはナンセンスな主張である. ただし, この厳密な定義を採用する場合, n 点集合と $n-1$ 点集合の間に全単射が存在しないことは明らかではなく, 別途証明する必要が生じる.

[54]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.

命題 19.6.6. $\mathbb{N}$ と $\mathbb{R}$ は対等ではない.

Proof. 全射 $f : \mathbb{N} \to \mathbb{R}$ が存在すると仮定すると矛盾が導けることを示そう. $f : \mathbb{N} \to \mathbb{R}$ を全射とする. まず $f(1)$ を含まない閉区間 $[a_1, b_1]$ を取る (ただし $a_1 < b_1$). 次に, $f(2)$ を含まない閉区間 $[a_2, b_2] \subset [a_1, b_1]$ を取る (ただし $a_2 < b_2$). 更に, $f(3)$ を含まない区間 $[a_3, b_3] \subset [a_2, b_2]$ を取る (ただし $a_3 < b_3$). この操作を順次繰り返していくと, 単調増加数列 $a_1 \leq a_2 \leq a_3 \leq \cdots$ および単調減少列 $b_1 \geq b_2 \geq b_2 \geq \cdots$ を得る. このとき閉区間の取り方から, 各 $n \in \mathbb{N}$ について $f(n) \notin [a_n, b_n]$ である. 上に有界な単調増加数列は収束する (実数の連続性) ゆえ数列 a_n は収束し, この極限を $a = \lim_{n\to\infty} a_n$ とする. さて, f の全射性より $f(n_0) = a$ を満たす $n_0 \in \mathbb{N}$ が存在する. このとき, 各 $n > n_0$ について $a_{n_0} \leq a_n < b_n \leq b_{n_0}$, つまり $a_{n_0} \leq a_n \leq b_{n_0}$ である. したがって, その極限においても $a_{n_0} \leq a \leq b_{n_0}$ が成り立つ. すなわち $f(n_0) \in [a_{n_0}, b_{n_0}]$ である. ところが, 閉区間 $[a_n, b_n]$ の取り方から $f(n_0) \notin [a_{n_0}, b_{n_0}]$ であり, 矛盾を得た. □

例 19.6.7. 次の事実の証明については, 集合論の入門的な本を参照されたい.

(1) $\mathbb{Z}$ や $\mathbb{Q}$ と $\mathbb{N}$ は対等である. 無理数全体や $\mathbb{C}$, 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, $X, \mathfrak{P}(X), \mathfrak{P}(\mathfrak{P}(X)), \mathfrak{P}(\mathfrak{P}(\mathfrak{P}(X))), \cdots$ は互いに対等ではない.

以上の事実から, 無限集合の元の総数[55]は無数の種類があることが分かる. とくに任意の無限集合 X が $X = \{ x_1, x_2, \cdots \}$ と表せるわけではない. もしこのような表示ができるならば写像 $f : \mathbb{N} \to X$ を $f(n) := x_n$ と定めれば f は全射となる. 更にあらかじめ部分列を取ることで $x_1, x_2, \cdots$ に重複がないようにしておけば f は全単射である. すなわち X と $\mathbb{N}$ は対等である. いまの議論から $\mathbb{R} = \{ x_1, x_2, \cdots \}$ と表せないことが分かる.

[55]無限集合も含めた文脈において, 集合 X を構成する元の総数のことを X の**濃度**という.

よりみち (無限集合の不思議).

$\mathbb{R}$ 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 $f : \mathbb{R} \to \mathbb{R}^n$ 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 項のコラムを参照). 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, $\mathbb{N}$ とも $\mathbb{R}$ とも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:

$$\int_a^b \int_c^d f(x,y)\,dxdy = \int_c^d \int_a^b f(x,y)\,dydx.$$

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.

20 線形写像

線形代数学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学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．本節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節以降に論じる．

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 $\boldsymbol{e}_i$ の間の積に関する次の性質

$$A\boldsymbol{e}_i = A \text{ の第 } i \text{ 列目}$$

を断りなく用いる (本節では例 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 V から線形空間 W への写像 $f: U \to V$ 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 $\boldsymbol{x}, \boldsymbol{y} \in U$ に対して, $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$,

(ii) すべての $\boldsymbol{x} \in U$ および $r \in \mathbb{R}$ について, $f(r\boldsymbol{x}) = rf(\boldsymbol{x})$.

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:

線形性 (iii) すべての $\boldsymbol{x}, \boldsymbol{y} \in V$ および $a, b \in \mathbb{R}$ について, $f(a\boldsymbol{x}+b\boldsymbol{y}) = af(\boldsymbol{x}) + bf(\boldsymbol{y})$.

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: $f(a\boldsymbol{x}+b\boldsymbol{y}) = f(a\boldsymbol{x}) + f(b\boldsymbol{y}) = af(\boldsymbol{x}) + bf(\boldsymbol{y})$. 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 $T_A: \mathbb{R}^n \to \mathbb{R}^m$ を $T_A(\boldsymbol{x}) := A\boldsymbol{x}$ 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．

$$T_A(a\boldsymbol{x}+b\boldsymbol{y}) = A(a\boldsymbol{x}+b\boldsymbol{y}) = A(a\boldsymbol{x}) + A(b\boldsymbol{y}) = a(A\boldsymbol{x}) + b(A\boldsymbol{y}) = aT_A(\boldsymbol{x}) + bT_A(\boldsymbol{y}).$$

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 $f: \mathbb{R} \to \mathbb{R}$ 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 $\boldsymbol{0}_V \in V$ にうつす定置写像 $f = \boldsymbol{0}_V: U \to V$ は線形写像である．実際, $f(a\boldsymbol{x}+b\boldsymbol{y}) = \boldsymbol{0}_V = a\boldsymbol{0}_V + b\boldsymbol{0}_V = af(\boldsymbol{x}) + bf(\boldsymbol{y})$. 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 $f: \mathbb{R} \to \mathbb{R}$ 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 $\det: M_n(\mathbb{R}) \to \mathbb{R}$ 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節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 $f: U \to V$ 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, $f(\mathbf{0}_U) = f(0 \cdot \mathbf{0}_U) = 0 \cdot f(\mathbf{0}_U) = \mathbf{0}_V$. □

線形性 (iii) は更に次のような有限和の形に一般化できる[56].

命題 20.1.6. $f: U \to V$ が線形性を満たすならば次も満たす:

線形性 (iii)' 各 $\boldsymbol{u}_k \in U$, $r_k \in \mathbb{R}$ $(k = 1, \cdots, \ell)$ について, $f\left(\sum_{k=1}^{\ell} r_k \boldsymbol{u}_k\right) = \sum_{k=1}^{\ell} r_k f(\boldsymbol{u}_k)$.

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.

$$
\begin{aligned}
f\left(\sum_{k=1}^{\ell+1} r_k \boldsymbol{u}_k\right) &= f\left(\left(\sum_{k=1}^{\ell} r_k \boldsymbol{u}_k\right) + r_{\ell+1}\boldsymbol{u}_{\ell+1}\right) \\
&= f\left(\sum_{k=1}^{\ell} r_k \boldsymbol{u}_k\right) + f\left(r_{\ell+1}\boldsymbol{u}_{\ell+1}\right) && \text{(線形性 (i))} \\
&= \sum_{k=1}^{\ell} r_k f(\boldsymbol{u}_k) + r_{\ell+1} f(\boldsymbol{u}_{\ell+1}) && \text{(帰納法の仮定と線形性 (ii))} \\
&= \sum_{k=1}^{\ell+1} r_k f(\boldsymbol{u}_k).
\end{aligned}
$$

□

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. $f: U \to V$ を線形写像とし, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in U$ とする.

(1) $\boldsymbol{x} \in U$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合で書けるならば, $f(\boldsymbol{x})$ は $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ の線形結合で書ける. (すなわち, $\boldsymbol{x} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle \Longrightarrow f(\boldsymbol{x}) \in \langle f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n) \rangle$.)

(2) $\sum_{i=1}^{n} a_i \boldsymbol{u}_i = \mathbf{0}_U \Longrightarrow \sum_{i=1}^{n} a_i f(\boldsymbol{u}_i) = \mathbf{0}_V$.

(3) $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ が線形独立ならば $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ も線形独立である.

(4) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が線形従属ならば $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ 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, $\boldsymbol{x} = \sum_{i=1}^{n} a_i \boldsymbol{u}_i$ と書けるならば $f(\boldsymbol{x}) = f(\sum_{i=1}^{n} a_i \boldsymbol{u}_i) = \sum_{i=1}^{n} a_i f(\boldsymbol{u}_i)$ であり, $f(\boldsymbol{x})$ は $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ の線形結合で書ける.

(2): $\boldsymbol{x} = \mathbf{0}_U$ 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 $\sum_{i=1}^{n} a_i \boldsymbol{u}_i = \mathbf{0}_U$ を仮定すれば (2) より $\sum_{i=1}^{n} a_i f(\boldsymbol{u}_i) = \mathbf{0}_V$ である. 組 $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ の線形独立性より $a_1 = \cdots = a_n = 0$. つまり $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は線形独立である.

(4): これは (3) の対偶にほかならない. □

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).

[56] 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.

例 20.1.8. $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ が定める線形写像 $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ について次が成り立つ:

(1) $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ とする．このとき, $f(\boldsymbol{x}) = \boldsymbol{e}_1$, $f(\boldsymbol{e}_1) = \boldsymbol{e}_1$ である．よって, $f(\boldsymbol{x})$ は $f(\boldsymbol{e}_1)$ の線形結合で書ける．しかし $\boldsymbol{x}$ を $\boldsymbol{e}_1$ の線形結合で書くことはできない.

(2) $0f(\boldsymbol{e}_1) + 1f(\boldsymbol{e}_2) = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$ である．一方で $0\boldsymbol{e}_1 + 1\boldsymbol{e}_2 = \boldsymbol{e}_2 \neq \boldsymbol{0}$ である.

(3) $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ は線形独立であるが, $f(\boldsymbol{e}_1)$, $f(\boldsymbol{e}_2)$ は線形独立でない.

補足. 線形写像 $f : U \to V$ 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 項における議論から $f^{-1}(\boldsymbol{0}_V)$ は 1 点集合ではない．したがって, $f(\boldsymbol{u}) = \boldsymbol{0}_V$ を満たす零でないベクトル $\boldsymbol{u} \in U$ が存在する．このとき, (1) $f(\boldsymbol{u}) = \boldsymbol{0}_V = f(\boldsymbol{0}_U)$ ゆえ $f(\boldsymbol{u})$ は $f(\boldsymbol{0}_U)$ の線形結合で書けるが, $\boldsymbol{u} \neq \boldsymbol{0}_U$ を $\boldsymbol{0}_U$ の線形結合で書くことはできない．(2) $1f(\boldsymbol{u}) + 1f(\boldsymbol{0}_U) = \boldsymbol{0}_V$ であるが, $1\boldsymbol{u} + 1\boldsymbol{0}_U = \boldsymbol{u} \neq \boldsymbol{0}_U$ である．(3) 1 つのベクトルからなる組 $\boldsymbol{u}$ は線形独立であるが, 組 $f(\boldsymbol{u})$ 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)**.** $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ と定めれば, $A = \{ \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3 \}$ は線形従属である．ところが, 例 20.1.8 で与えた $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ において, $f(\boldsymbol{a}_1) = f(\boldsymbol{a}_2) = f(\boldsymbol{a}_3) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 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 $f : U \to V$ が単射でない場合, 上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[57].

Proof. f が単射でない場合, 前例の補足で述べたように $f^{-1}(\boldsymbol{0}_V)$ は 1 点集合ではない．また, $f^{-1}(\boldsymbol{0}_V)$ 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 $\boldsymbol{v} \in f(\mathbb{R}^n)$ が存在し, 系 23.1.5 によれば $f^{-1}(\boldsymbol{v})$ の元の個数は $f^{-1}(\boldsymbol{0}_V)$ の元の個数と等しく, したがって $f^{-1}(\boldsymbol{v})$ も無限集合となる．そこで, 相異なる $n+1$ 個の元 $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n+1} \in f^{-1}(\boldsymbol{v}) \subset \mathbb{R}^n$ をとり, $A = \{ \boldsymbol{u}_1, \dots, \boldsymbol{u}_{n+1} \}$ とすれば $f(A) = \{ \boldsymbol{v} \}$ 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. $f : U \to V$ を線形写像とする．ベクトルの組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が U を生成し, 各 $i = 1, \cdots, n$ について $f(\boldsymbol{u}_i) = \boldsymbol{0}_V$ ならば, $f = \boldsymbol{0}_V$．すなわち, f は零ベクトルに値を取る定置写像である.

Proof. 各 $\boldsymbol{x} \in U$ について, 仮定より $\boldsymbol{x} = \sum_{i=1}^n a_i \boldsymbol{u}_i$ と書ける． ゆえに $f(\boldsymbol{x}) = f\big(\sum_{i=1}^n a_i \boldsymbol{u}_1\big) = \sum_{i=1}^n a_i f(\boldsymbol{u}_i) = \sum_{i=1}^n a_i \boldsymbol{0}_V = \boldsymbol{0}_V$. □

次の命題は, 線形写像の値は基底の行き先によって決定されることを述べている.

命題 20.1.11. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ を U の基底とし, $\boldsymbol{v}_1 \cdots, \boldsymbol{v}_n \in V$ とする ($\boldsymbol{v}_1 \cdots, \boldsymbol{v}_n$ 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, $f(\boldsymbol{u}_i) = \boldsymbol{v}_i$ を満たす線形写像 $f : U \to V$ が唯一つ存在する.

[57] 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.

Proof. U の各元 $\boldsymbol{x}$ は実数の組 $a_1, \cdots, a_n$ を用いて $\boldsymbol{x} = \sum_{i=1}^n a_i\boldsymbol{u}_i$ と一意的に書ける (命題 18.2.3). f が線形性を満たすには $f(\boldsymbol{x}) := \sum_{i=1}^n a_i\boldsymbol{v}_i$ 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 $\boldsymbol{x}, \boldsymbol{y} \in U$ を $\boldsymbol{x} = \sum_{i=1}^n a_i\boldsymbol{u}_i$, $\boldsymbol{y} = \sum_{i=1}^n b_i\boldsymbol{u}_i$ と書けば, $r\boldsymbol{x} + s\boldsymbol{y}$ は $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合として $r\boldsymbol{x} + s\boldsymbol{y} = \sum_{i=1}^n (ra_i + sb_i)\boldsymbol{u}_i$ と一意的に書けており, f の定義から $f(\boldsymbol{x}) = \sum_{i=1}^n a_i\boldsymbol{v}_i$, $\boldsymbol{y} = \sum_{i=1}^n b_i\boldsymbol{v}_i$, $f(r\boldsymbol{x} + s\boldsymbol{y}) = \sum_{i=1}^n (ra_i + sb_i)\boldsymbol{v}_i$ である. したがって,

$$f(r\boldsymbol{x} + s\boldsymbol{y}) = \sum_{i=1}^n (ra_i + sb_i)\boldsymbol{v}_i = \sum_{i=1}^n ra_i\boldsymbol{v}_i + \sum_{i=1}^n rb_i\boldsymbol{v}_i = r\left(\sum_{i=1}^n a_i\boldsymbol{v}_i\right) + s\left(\sum_{i=1}^n b_i\boldsymbol{v}_i\right) = rf(\boldsymbol{x}) + sf(\boldsymbol{y}).$$

すなわち, 線形性 (iii) を満たす f は線形写像である.

次に f の一意性を示すために, 線形写像 $g : U \to V$ も $g(\boldsymbol{u}_i) = \boldsymbol{v}_i$ を満たすと仮定し $f = g$ を導こう. そのためには各 $\boldsymbol{x} \in U$ について $f(\boldsymbol{x}) = g(\boldsymbol{x})$ を示せばよい. $\boldsymbol{x} = \sum_{i=1}^n a_i\boldsymbol{u}_i$ と書けており, f の定義と g の線形性 (iii)' を用いて変形すると,

$$f(\boldsymbol{x}) = \sum_{i=1}^n a_i\boldsymbol{v}_i = \sum_{i=1}^n a_i g(\boldsymbol{u}_i) = g\left(\sum_{i=1}^n a_i\boldsymbol{u}_i\right) = g(\boldsymbol{x}).$$

ゆえに $f = g$ である. □

上の証明において, 各 $\boldsymbol{x} \in U$ が $\boldsymbol{x} = \sum_{i=1}^n a_i\boldsymbol{u}_i$ 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, $\boldsymbol{x} = \sum_{i=1}^n a_i\boldsymbol{u}_i = \sum_{i=1}^n b_i\boldsymbol{u}_i$ と二通りに書けるとすると $f(\boldsymbol{x})$ を $\sum_{i=1}^n a_i\boldsymbol{v}_i$ と $\sum_{i=1}^n b_i\boldsymbol{v}_i$ のいずれにすべきか, あらかじめ指定する必要が生じる. 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. $f : U \to V$ および $g : V \to W$ を線形写像とし, これらの合成 $g \circ f : U \to W$ が線形性 (iii) を満たすことを示そう. $\boldsymbol{x}, \boldsymbol{y} \in U$ および $r, s \in \mathbb{R}$ に対して,

$$g\circ f(r\boldsymbol{x}+s\boldsymbol{y}) = g\Big(f(r\boldsymbol{x}+s\boldsymbol{y})\Big) = g\Big(rf(\boldsymbol{x})+sf(\boldsymbol{y})\Big) = rg\Big(f(\boldsymbol{x})\Big)+sg\Big(f(\boldsymbol{y})\Big) = r(g\circ f)(\boldsymbol{x})+s(g\circ f)(\boldsymbol{y}).$$

□

20.2 線形写像による像

写像 $f : X \to Y$ 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 $\operatorname{Im} f$ 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 $f : U \to V$ において, f による U の像 $f(U)$ のことを $\operatorname{Im} f$ と書く. すなわち, $\operatorname{Im} f := \{\, f(\boldsymbol{u}) \mid \boldsymbol{u} \in U \,\}$.

命題 20.2.2. 線形写像 $f : U \to V$ の像 $\operatorname{Im} f$ 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): $f(\boldsymbol{0}_U) = \boldsymbol{0}_V$ ゆえ $\boldsymbol{0}_V \in \operatorname{Im} f$ である.
(iv): $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \operatorname{Im} f$ 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, $f(\boldsymbol{u}_1) = \boldsymbol{v}_1$, $f(\boldsymbol{u}_2) = \boldsymbol{v}_2$ を満たす $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U$ が存在する. このとき, 各 $r, s \in \mathbb{R}$ について $r\boldsymbol{v}_1 + s\boldsymbol{v}_2 = rf(\boldsymbol{u}_1) + sf(\boldsymbol{u}_2) = f(r\boldsymbol{u}_1 + s\boldsymbol{u}_2)$ である. つまり, ベクトル $r\boldsymbol{v}_1 + s\boldsymbol{v}_2 \in V$ はベクトル $r\boldsymbol{u}_1 + s\boldsymbol{u}_2 \in U$ を f に代入することで得られる. ゆえに $r\boldsymbol{v}_1 + s\boldsymbol{v}_2 \in \operatorname{Im} f$. □

$f : U \to V$ 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 $f|_W : W \to V$ の像とも見なせる. また, $f|_W$ も線形写像であることから, $f(W) = \operatorname{Im} f|_W$ 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 $f : U \to V$ および組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in U$ において,

$$f(\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle) = \langle f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n) \rangle.$$

とくに $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が U を生成するとき, $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ は $\operatorname{Im} f$ を生成する.

Proof.

$$\begin{aligned}
f(\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle) &= \{\, f(\boldsymbol{x}) \mid \boldsymbol{x} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle \,\} \\
&= \left\{ f\left(\sum_{i=1}^{n} a_i \boldsymbol{u}_i \right) \;\middle|\; a_1, \ldots, a_n \in \mathbb{R} \right\} && (\text{この等号は後述する}) \\
&= \left\{ \sum_{i=1}^{n} a_i f(\boldsymbol{u}_i) \;\middle|\; a_1, \ldots, a_n \in \mathbb{R} \right\} && (f\ \text{の線形性 (iii)' を用いた}) \\
&= \langle f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n) \rangle.
\end{aligned}$$

上の変形で明らかでないのは $\{\, f(\boldsymbol{x}) \mid \boldsymbol{x} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle \,\} = \{\, f(\sum_{i=1}^{n} a_i \boldsymbol{u}_i) \mid a_1, \ldots, a_n \in \mathbb{R} \,\}$ のみゆえ, これを示そう. $A = \{\, f(\boldsymbol{x}) \mid \boldsymbol{x} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle \,\}$, $B = \{\, f(\sum_{i=1}^{n} a_i \boldsymbol{u}_i) \mid a_1, \ldots, a_n \in \mathbb{R} \,\}$ 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 $f(\boldsymbol{x}) \in A$ において, $\boldsymbol{x} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ より $a_1, \ldots, a_n$ を用いて $\boldsymbol{x} = \sum_{i=1}^{n} a_i \boldsymbol{u}_i$ とかける. つまり $f(\boldsymbol{x}) = f(\sum_{i=1}^{n} a_i \boldsymbol{u}_i) \in B$.

$(B \subset A)$: 各 $f(\sum_{i=1}^{n} a_i \boldsymbol{u}_i) \in B$ において, $\boldsymbol{x} = \sum_{i=1}^{n} a_i \boldsymbol{u}_i$ とおけば $\boldsymbol{x} \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ である. つまり $f(\sum_{i=1}^{n} a_i \boldsymbol{u}_i)$ は, $\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ の元 $\boldsymbol{x}$ を用いて $f(\boldsymbol{x})$ と書ける. したがって $f(\sum_{i=1}^{n} a_i \boldsymbol{u}_i) \in A$. □

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 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ による線形写像 $T_A : \mathbb{R}^n \to \mathbb{R}^m$ $(T_A(\boldsymbol{x}) := A\boldsymbol{x})$ において, 定義域 $\mathbb{R}^n$ の標準基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ を f に代入した値 $A\boldsymbol{e}_1, \cdots, A\boldsymbol{e}_n$ は A の各列ベクトル $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ 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, $\operatorname{Im} T_A = \langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \rangle$.

20.3 線形写像による逆像

写像 $f : X \to Y$ 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 $f : U \to V$ において, f による $\boldsymbol{0}_V$ の逆像 $f^{-1}(\{\boldsymbol{0}_V\})$ のことを $\operatorname{Ker} f$ と書き, これを f の核 (**kernel**) という. すなわち, $\operatorname{Ker} f := \{\, \boldsymbol{u} \in U \mid f(\boldsymbol{u}) = \boldsymbol{0}_V \,\}$.

例 20.3.2. (m, n)-行列 A による線形写像 $T_A : \mathbb{R}^n \to \mathbb{R}^m$ $(T_A(\boldsymbol{x}) := A\boldsymbol{x})$ の核とは, $A\boldsymbol{x} = \boldsymbol{0}$ を満たす点 $\boldsymbol{x} \in \mathbb{R}^n$ 全体からなる集合である. ゆえに $\operatorname{Ker} T_A$ は連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{0}$ の解空間 W_A に等しい.

命題 16.2.2 において斉次形連立 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 $f : U \to V$ 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): $f(\boldsymbol{0}_U) = \boldsymbol{0}_V \in W$ ゆえ $\boldsymbol{0}_U$ を f に代入すると W の元となる. すなわち, $\boldsymbol{0}_U \in f^{-1}(W)$ である.

(iv): $\boldsymbol{u}_1, \boldsymbol{u}_2 \in f^{-1}(W)$ とすれば, $f(\boldsymbol{u}_1), f(\boldsymbol{u}_2) \in W$ 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, $f(r\boldsymbol{u}_1 + s\boldsymbol{u}_2) = rf(\boldsymbol{u}_1) + sf(\boldsymbol{u}_2) \in W$ である. すなわち, $r\boldsymbol{u}_1 + s\boldsymbol{u}_2 \in f^{-1}(W)$. □

命題 20.3.4. 線形写像 $f : U \to V$ 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 $\boldsymbol{v} \in \operatorname{Im} f$ においても, その逆像 $f^{-1}(\boldsymbol{v})$ が 1 点からなる,

(2) $\operatorname{Ker} f = \{\mathbf{0}_U\}$.

Proof. (1)⇒(2): (2) は, (1) における $\boldsymbol{v}$ として $\mathbf{0}_V = f(\mathbf{0}_U) \in \operatorname{Im} f$ を考えた特別な場合である.

(2)⇒(1): $f(\boldsymbol{x}) = f(\boldsymbol{y})$ を仮定し, $\boldsymbol{x} = \boldsymbol{y}$ を示そう. ベクトル $\boldsymbol{x} - \boldsymbol{y}$ を f に代入すると $f(\boldsymbol{x} - \boldsymbol{y}) = f(\boldsymbol{x}) - f(\boldsymbol{y}) = \mathbf{0}_V$. つまり $\boldsymbol{x} - \boldsymbol{y} \in \operatorname{Ker} f$ であり, 仮定より $\operatorname{Ker} f$ は零ベクトル $\mathbf{0}_U$ のみからなるとしていたから $\boldsymbol{x} - \boldsymbol{y} = \mathbf{0}_U$ である. これを移項して $\boldsymbol{x} = \boldsymbol{y}$ を得る. □

例 20.3.5. (m, n)-行列 A において, T_A が単射であることと $A\boldsymbol{x} = \mathbf{0}$ 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. $f : U \to V$ を線形写像とし, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in U$ とする. f が単射であるとき次が成り立つ:

(1) $\boldsymbol{x} = \sum_{i=1}^n a_i \boldsymbol{u}_i \iff f(\boldsymbol{x}) = \sum_{i=1}^n a_i f(\boldsymbol{u}_i)$

(2) $\sum_{i=1}^n a_i \boldsymbol{u}_i = \mathbf{0}_V \iff \sum_{i=1}^n a_i f(\boldsymbol{u}_i) = \mathbf{0}_U$.

(3) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は線形独立である $\iff f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ は線形独立である.

(4) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は線形従属である $\iff f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ は線形従属である.

(5) $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ は V の基底である $\Longrightarrow \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は U の基底である.

Proof. まず (1) を示そう. (⇒) は命題 20.1.7 で得られているゆえ (⇐) のみ証明する. $f(\boldsymbol{x}) = \sum_{i=1}^n a_i f(\boldsymbol{u}_i)$ とする. このとき, $\boldsymbol{y} = \sum_{i=1}^n a_i \boldsymbol{u}_i$ とおき, $\boldsymbol{x} - \boldsymbol{y}$ を f に代入すると,

$$f(\boldsymbol{x} - \boldsymbol{y}) = f(\boldsymbol{x}) - f(\boldsymbol{y}) = f(\boldsymbol{x}) - f\left(\sum_{i=1}^n a_1 \boldsymbol{u}_i\right) = \sum_{i=1}^n a_i f(\boldsymbol{u}_i) - \sum_{i=1}^n a_i f(\boldsymbol{u}_i) = \mathbf{0}.$$

ゆえに $\boldsymbol{x} - \boldsymbol{y} \in \operatorname{Ker} f$ である. f は単射ゆえ $\operatorname{Ker} f$ は $\mathbf{0}_U$ のみからなる. ゆえに $\boldsymbol{x} - \boldsymbol{y} = \mathbf{0}_U$. すなわち $\boldsymbol{x} = \boldsymbol{y} = \sum_{i=1}^n a_i \boldsymbol{u}_i$.

(2) は, (1) において $\boldsymbol{x} = \mathbf{0}_U$ 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. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形独立性は (3) より分かっている. 各 $\boldsymbol{x} \in U$ に対して $f(\boldsymbol{x}) \in V$ であり, $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ が V を生成することから $f(\boldsymbol{x}) = \sum_{i=1}^n a_i f(\boldsymbol{u}_i)$ と書ける. (1) より $\boldsymbol{x} = \sum_{i=1}^n a_i \boldsymbol{u}_i$. すなわち $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ は U を生成する. □

命題 20.3.6(5) の逆が成り立つためには, $\operatorname{Im} f = V$ 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 $\boldsymbol{v}_1 \cdots, \boldsymbol{v}_n$ が線形独立であるとき, $f(\boldsymbol{u}_i) = \boldsymbol{v}_i$ を満たす線形写像 $f : U \to V$ は単射である. これを示せ.

解答例: $\operatorname{Ker} f = \{\mathbf{0}_U\}$ を示せばよい. そこで $\boldsymbol{x} \in \operatorname{Ker} f$ とする. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が U を生成することから $\boldsymbol{x} = \sum_{i=1}^n a_i \boldsymbol{u}_i$ と書けば,

$$\mathbf{0} = f(\boldsymbol{x}) = f\left(\sum_{i=1}^n a_i \boldsymbol{u}_i\right) = \sum_{i=1}^n a_i f(\boldsymbol{u}_i) = \sum_{i=1}^n a_i \boldsymbol{e}_i.$$

$\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ は線形独立ゆえ $a_1 = \cdots = a_n = 0$ であり, $\boldsymbol{x} = \sum_{i=1}^n 0\boldsymbol{u}_i = \mathbf{0}_U$ を得る. すなわち $\operatorname{Ker} f$ の元は $\mathbf{0}_U$ に限る. □

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 $p_n : \mathbb{R}^n \to \mathbb{R}$ ($f(x_1, \cdots, x_n) := x_i$) は第 i 座標への**射影** (**projection**) と呼ばれる．p_i は線形写像である．

例 20.4.2. n 次正方行列 $A = [a_{ij}]$ に対して A の対角成分の和 $\mathrm{tr}\, A := \sum_{i=1}^{n} a_{ii}$ を A の**跡** (**trace**) あるいは**トレース**, **対角和**などと呼ぶ．写像 $\mathrm{tr} : M_n(\mathbb{R}) \to \mathbb{R}$ は線形写像である．

命題 20.4.3. n 次正方行列 A, B について $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Proof. $A = [a_{ij}]_{i,j}$, $B = [b_{ij}]_{i,j}$ とすれば, $AB = \left[\sum_{k=1}^{n} a_{ik} b_{kj}\right]_{i,j}$ ゆえ $\mathrm{tr}(AB) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki}$. また, $BA = \left[\sum_{\ell=1}^{n} b_{p\ell}\, a_{\ell q}\right]_{p,q}$ ゆえ $\mathrm{tr}(BA) = \sum_{p=1}^{n} \sum_{\ell=1}^{n} b_{p\ell}\, a_{\ell p} = \sum_{\ell=1}^{n} \sum_{p=1}^{n} a_{\ell p}\, b_{p\ell}$. 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 $\mathrm{tr}(AB)$ と同じ式である． □

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 $S(x_1, x_2, x_3, \cdots) := (x_2, x_3, \cdots)$ で定められる写像 $S : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ 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 $\boldsymbol{y} = (y_1, y_2, y_3, \cdots)$ に対して, $\boldsymbol{x} = (0, y_1, y_2, \cdots)$ と定めれば $S(\boldsymbol{x}) = \boldsymbol{y}$ となる．また, $S(\boldsymbol{x}) = \boldsymbol{0}$ なる数列 $\boldsymbol{x}$ は, 初項を除いてすべて 0 なる数列であり, $\mathrm{Ker}\, S = \{\, (x, 0, 0, \cdots) \mid x \in \mathbb{R} \,\}$.

(3) 数列 $\boldsymbol{x} = (x_n)_{n \in \mathbb{N}}$ が線形漸化式 16.3.2:

$$x_{n+k} = a_{k-1} x_{n+k-1} + a_{k-2} x_{n+k-2} + \cdots + a_1 x_{n+1} + a_0 x_n$$

を満たすならば, 数列 $S(\boldsymbol{x}) = (x_{n+1})_{n \in \mathbb{N}}$ 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 $S : F \to F$ は全単射である．

Proof. 全射性: 各 $\boldsymbol{y} = (y_1, y_2, y_3, \cdots)$ に対して, $\boldsymbol{x} = (y_2 - y_1, y_1, y_2, y_3, \cdots)$ と定めれば $\boldsymbol{x} \in F$ であり $S(\boldsymbol{x}) = \boldsymbol{y}$ となる．なお, $\boldsymbol{x} = S(\boldsymbol{y}) - \boldsymbol{y}$ 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 $\boldsymbol{y}$ 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[58].

単射性: $\mathrm{Ker}\, S = \{\boldsymbol{0}\}$ を示せばよい．$\boldsymbol{x} = (x_1, x_2, x_3, \cdots) \in F$, $S(\boldsymbol{x}) = \boldsymbol{0}$ とすれば (2) より $\boldsymbol{x}$ 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 $\boldsymbol{x} = \boldsymbol{0}$ である． □

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' もまた何回も微分できる, すなわち $f' \in C^\infty(I)$. そこで, $D : C^\infty(I) \to C^\infty(I)$ 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 $\boldsymbol{0} \in C^\infty(I)$ になる関数は定数関数しかないことから[59], $\mathrm{Ker}\, D$ は定数関数全体のなす集合になる．

---

[58]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．

[59]実際, 対偶「f が定数関数でないならば $f' \neq \boldsymbol{0}$」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 $\boldsymbol{0}$ ではない．

(3) 関数 $y \in C^\infty(I)$ が線形微分方程式 16.3.4:

$$y^{(k)}(x) = a_{k-1}y^{(k-1)} + a_{k-2}y^{(k-2)} + \cdots + a_1 y^{(1)}(x) + a_0 y^{(0)}(x)$$

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 $D : W \to W$ 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: $\operatorname{Ker} D = \{\mathbf{0}\}$ 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. □

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 $\dfrac{d}{dx}$ で表す.

例 20.4.6 (発展)**.** $a \in \mathbb{R}$ を固定し, 写像 $T : C^\infty(\mathbb{R}) \to \mathbb{R}^{\mathbb{N}}$ を

$$T(f) := (f(a),\, f^{(1)}(a),\, f^{(2)}(a),\, \cdots)$$

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.

$$\begin{array}{cccccccc}
& C^\infty(\mathbb{R}) & \ni & f & \xmapsto{T} & (f(a),\, f^{(1)}(a),\, f^{(2)}(a),\, f^{(3)}(a),\, \cdots) & \in & \mathbb{R}^{\mathbb{N}} & \\
D & \downarrow & & \downarrow & & \downarrow & & \downarrow & S \\
& C^\infty(\mathbb{R}) & \ni & f^{(1)} & \xmapsto{T} & (f^{(1)}(a),\, f^{(2)}(a),\, f^{(3)}(a),\, f^{(4)}(a),\, \cdots) & \in & \mathbb{R}^{\mathbb{N}} &
\end{array}$$

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.

(1) $S_{a,b} : C(I) \to \mathbb{R}$ を $S_{a,b}(f) := \displaystyle\int_a^b f(x)dx$ と定めれば $S_{a,b}$ は線形写像である. $\operatorname{Ker} S_{a,b}$ の元は区間 $[a, b]$ における定積分が 0 となる関数である. $\operatorname{Im} S_{a,b}$ 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, $\operatorname{Im} S_{a,b} = \{0\}$. 言い換えると $\operatorname{Ker} S_{a,b} = C(I)$ 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, $\operatorname{Im} S_{a,b} = \mathbb{R}$.

(2) $C(I)$ の元 f に対して次で定義される $C(I)$ の元 $I_a(f) : I \to \mathbb{R}$ を対応させる写像 $I_a : C(I) \to C(I)$ 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. $\operatorname{Im} I_a$ は $F(a) = 0$ を満たす C^1-級関数[60] F たち全体に一致する.

[60] 導関数が連続となる関数を C^1-級関数という.

Proof. $V = \{ F \in C(I) \mid F(a) = 0$ かつ F は C^1-級$\}$ とし, $\operatorname{Im} I_a = V$ を示そう.

($\operatorname{Im} I_a \subset V$): $\operatorname{Im} I_a$ 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 $F \in V$.

($V \subset \operatorname{Im} I_a$): $F \in V$ 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[61], $F-G$ 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 $F = G = I_a(f) \in \operatorname{Im} I_a$. □

$I_a(f) = I_a(g)$ 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 $\operatorname{Ker} I_a = \{\mathbf{0}\}$.

(3) $\operatorname{Im} I_a$ 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 $D = \dfrac{d}{dx}$ に代入することができる. 微分積分学の基本定理より $D \circ I_a = \mathrm{id}_{C(I)}$ 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 $F : X \to Y$ を一つ与えると, 次のような写像 $T_F : C(Y) \to C(X)$ 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).

$$T_F(f) := f \circ F, \qquad (\text{ここで } f \in C(Y), \text{ つまり } f : Y \to \mathbb{R} \text{ は連続関数}).$$

(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 $f \circ F : X \to \mathbb{R}$ が連続であることを示さねばならないが, これは解析学の講義に譲り (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がら計算すると,

$$\begin{aligned}
T_F(rf + sg)(x) &= (rf + sg) \circ F(x) = (rf + sg)\Big(F(x)\Big) = (rf)\Big(F(x)\Big) + (sg)\Big(F(x)\Big) \\
&= r \cdot f\Big(F(x)\Big) + s \cdot g\Big(F(x)\Big). \\
\Big(rT_F(f) + sT_F(g)\Big)(x) &= (rT_F(f))(x) + (sT_F(g))(x) = \Big(r(f \circ F)\Big)(x) + \Big(s(g \circ F)\Big)(x) \\
&= r \cdot (f \circ F(x)) + s \cdot (g \circ F(x)) = r \cdot f\Big(F(x)\Big) + s \cdot g\Big(F(x)\Big).
\end{aligned}$$

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.

$$\begin{aligned}
T_F(fg)(x) &= (fg) \circ F(x) = (fg)\Big(F(x)\Big) = f\Big(F(x)\Big) \cdot g\Big(F(x)\Big). \\
\Big(T_F(f)T_F(g)\Big)(x) &= (T_F(f))(x) \cdot (T_F(g))(x) = (f \circ F)(x) \cdot (g \circ F)(x) = f\Big(F(x)\Big) \cdot g\Big(F(x)\Big).
\end{aligned}$$

□

[61]「$g' = \mathbf{0}$ 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.

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式変形をした経験はないだろうか:

$$(a+b)^2 = a^2 + b^2, \qquad \sqrt{a+b} = \sqrt{a} + \sqrt{b}.$$

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, 上述のミスを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): $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x})+f(\boldsymbol{y})$ 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, 線形性を要求するということは, 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 $\mathrm{sgn}(\sigma\tau) = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\tau)$ 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:

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right).$$

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.

21 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. 本節では, まずはじめに, 線形代数的な性質の相互翻訳が可能な 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. 本節の後半では, 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 $f : U \to V$ 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 $f : U \to V$ において, 終域を置き換えた写像 $f : U \to \operatorname{Im} f$ は線形同型である.

例 21.1.3. U の基底を $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ とする. このとき, 命題 20.1.11 より $f(\boldsymbol{u}_i) = \boldsymbol{e}_i$ を満たす線形写像 $f : U \to \mathbb{R}^n$ が存在する. この f は線形同型である.

Proof. 命題 20.2.3 より $\operatorname{Im} f = \langle f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)\rangle = \langle \boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\rangle = \mathbb{R}^n$. 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 19.4.6(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 $f : U \to V$ の逆写像 $f^{-1} : V \to U$ 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 $\boldsymbol{x}, \boldsymbol{y} \in V$ および $r, s \in \mathbb{R}$ を任意に取る. $\boldsymbol{a} = f^{-1}(\boldsymbol{x})$, $\boldsymbol{b} = f^{-1}(\boldsymbol{y})$ とおく. 逆写像の定義から $f(\boldsymbol{a}) = \boldsymbol{x}$, $f(\boldsymbol{b}) = \boldsymbol{y}$ である. また f の線形性より $f(r\boldsymbol{a} + s\boldsymbol{b}) = rf(\boldsymbol{a}) + sf(\boldsymbol{b}) = r\boldsymbol{x} + s\boldsymbol{y}$ である. この両辺をそれぞれ f^{-1} に代入することで

$$\begin{aligned} f^{-1}\Big(f(r\boldsymbol{a} + s\boldsymbol{b})\Big) &= f^{-1}(r\boldsymbol{x} + s\boldsymbol{y}) \\ r\boldsymbol{a} + s\boldsymbol{b} &= f^{-1}(r\boldsymbol{x} + s\boldsymbol{y}) \\ rf^{-1}(\boldsymbol{x}) + sf^{-1}(\boldsymbol{y}) &= f^{-1}(r\boldsymbol{x} + s\boldsymbol{y}). \end{aligned}$$

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. $f : U \to V$ が線形同型であるとき, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \in U$ について次が成り立つ.

$$\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \text{ は } U \text{ の基底である} \iff f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n) \text{ は } V \text{ の基底である}.$$

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. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ を U の基底とする. 命題 20.3.6(3) より $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ は線形独立である. また, f は全射ゆえ $\operatorname{Im} f = V$ であり, 命題 20.2.3 より $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)$ は V を生成する. □

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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$, $f : U \to V$ を線形写像とするとき, 次が成り立つ.

$$f \text{ は線形同型である} \iff f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n) \text{ は } V \text{ の基底である.}$$

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. □

命題 21.1.7. 線形同型の合成は線形同型である.

Proof. $f : U \to V$ および $g : V \to W$ 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 19.5.1(3) より全単射である. □

本項において, 前節で述べた一般の線形写像について片側の矢印: $\Longrightarrow$ あるいは $\Longleftarrow$ 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 $(\Longleftrightarrow)$ 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 $f : U \to V$ 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 $f : U \to V$ を $f(\boldsymbol{x}) := {}^t\boldsymbol{x}$ と定めればこれは線形同型であり, U において $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in U$ に関する線形空間的な性質 $P(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ が成立するとき, V において $P(f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_n))$ が成立することが示唆される[62]. 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.

$$(\text{反射律})\ \ U \simeq U, \quad (\text{対称律})\ \ U \simeq V \Longrightarrow V \simeq U, \quad (\text{推移律})\ \ U \simeq V, V \simeq W \Longrightarrow U \simeq W.$$

Proof. (反射律): 恒等写像 $\mathrm{id}_U : U \to U$ 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 $f : U \to V$ が存在し, その逆写像 $f^{-1} : V \to U$ 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 $f : U \to V$, $g : V \to W$ が存在し, これらの合成 $g \circ f : U \to W$ が線形同型であることから $U \simeq W$ を得る. □

[62] 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, 数理論理学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.

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に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∽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 19.4.6(i) に相当する. また, 推移律は命題 19.6.4 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$, V の基底を $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ とすれば, 命題 20.1.11 により $f(\boldsymbol{u}_i) = \boldsymbol{v}_i$ を満たす線形写像 $f : U \to V$ 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 15.2.7(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 15.2.7(2) で与えた次の対応:

$$f : U \to V, \qquad f(x_1, x_2, \cdots) := \text{“}x(n) := x_n \text{ で定義される写像 } x : \mathbb{N} \to \mathbb{R}\text{”}$$

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見なすことができ, 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.

補足 (発展): 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 $\boldsymbol{xy} = \boldsymbol{yx}$ 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 $X_m = \{1, \ldots, m\}$ 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 $f : \mathbb{R}^n \to \mathbb{R}^m$ が存在するとすれば, 定義域 $\mathbb{R}^n$ の標準ベクトル $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ について命題 21.1.5 (より根本的には命題 20.3.6(3)) より $f(\boldsymbol{e}_1), \cdots, f(\boldsymbol{e}_n)$ 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. □

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 $f : \mathbb{R}^n \to \mathbb{R}^m$ は存在しない．一方，逆向きの線形全射 $g : \mathbb{R}^m \to \mathbb{R}^n$ 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

1.7 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．1.7 項において $\mathbb{R}^2$ 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項で述べよう．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 $\mathrm{Hom}(U,V)$ と書く．とくに $U = V$ のとき，これを $\mathrm{End}(U)$ と書く．$\mathrm{Hom}(U,V)$ 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:

- $f, g \in \mathrm{Hom}(U,V)$ に対して $(f+g) : U \to V$ を次の写像として定める:
  各 $\boldsymbol{x} \in U$ について，$(f+g)(\boldsymbol{x}) := f(\boldsymbol{x}) + g(\boldsymbol{x})$.
- $f \in \mathrm{Hom}(U,V)$ および $r \in \mathbb{R}$ に対して $(rf) : U \to V$ を次の写像として定める:
  各 $\boldsymbol{x} \in U$ について，$(rf)(\boldsymbol{x}) := rf(\boldsymbol{x})$.

$\mathrm{Hom}(U,V)$ 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**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，$f, g \in \mathrm{Hom}(U,V)$, $r \in \mathbb{R}$ ならば $f+g, rf \in \mathrm{Hom}(U,V)$ 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:

$$\begin{aligned}(f+g)(s\boldsymbol{x}+t\boldsymbol{y}) &= f(s\boldsymbol{x}+t\boldsymbol{y}) + g(s\boldsymbol{x}+t\boldsymbol{y}) = \Big(sf(\boldsymbol{x})+tf(\boldsymbol{y})\Big) + \Big(sg(\boldsymbol{x})+tg(\boldsymbol{y})\Big) \\ &= s\Big(f(\boldsymbol{x})+g(\boldsymbol{x})\Big) + t\Big(f(\boldsymbol{y})+g(\boldsymbol{y})\Big) = s(f+g)(\boldsymbol{x}) + t(f+g)(\boldsymbol{y}).\end{aligned}$$

□

(2) $\mathrm{Hom}(U,V)$ 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 $T_A : \mathbb{R}^n \to \mathbb{R}^m$ を $T_A(\boldsymbol{x}) := A\boldsymbol{x}$ 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) 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．$T_A(\boldsymbol{e}_i)$ および $T_B(\boldsymbol{e}_i)$ はそれぞれ A, B の i 列目であり，仮定より $T_A(\boldsymbol{e}_i) = T_B(\boldsymbol{e}_i)$ である．ゆえに A, B の各列は等しい．

(2): $T_{rA+sB}(\boldsymbol{x}) = (rA + sB)(\boldsymbol{x}) = r(A\boldsymbol{x}) + s(A\boldsymbol{x}) = rT_A(\boldsymbol{x}) + sT_B(\boldsymbol{x})$.

(3): $T_{BA}(\boldsymbol{x}) = (BA)\boldsymbol{x} = B(A\boldsymbol{x}) = BT_A(\boldsymbol{x}) = T_B(T_A(\boldsymbol{x})) = T_B \circ T_A(\boldsymbol{x})$.

(4): A を可逆とすれば，(3) より $T_A \circ T_{A^{-1}} = T_E = \mathrm{id}_{\mathbb{R}^n}$, $T_{A^{-1}} \circ T_A = \mathrm{id}_{\mathbb{R}^n}$ である．したがって命題 19.4.4 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 $T_A : \mathbb{R}^n \to \mathbb{R}^m$ 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 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ について命題 20.3.6(3) より $T_A(\boldsymbol{e}_1), \ldots, T_A(\boldsymbol{e}_n)$ 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． □

定理 21.3.4. 任意の線形写像 $f : \mathbb{R}^n \to \mathbb{R}^m$ 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，$\mathcal{T} : M_{m,n}(\mathbb{R}) \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ を $\mathcal{T}(A) := T_A$ と定めれば，これは線形同型写像である．

Proof. $A_f := [f(\boldsymbol{e}_1), \cdots, f(\boldsymbol{e}_n)]$ と定めれば $f = T_{A_f}$ である．実際，$T_{A_f}(\boldsymbol{e}_i) =$ "A_f の i 列目" $= f(\boldsymbol{e}_i)$ 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 20.1.11 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． □

上の定理における $\mathcal{T}$ の逆写像 $\mathcal{S} : \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \to M_{m,n}(\mathbb{R})$ は，$f \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ に対して $A_f = [f(\boldsymbol{e}_1), \cdots, f(\boldsymbol{e}_n)]$ 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 $\mathrm{End}(\mathbb{R}^n)$ 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，本論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. $\mathrm{End}(U)$ 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，$E = \mathrm{id}_U$ 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，$\mathrm{End}(U)$ は多元環となる．

以降では，恒等写像 $\mathrm{id}_U : U \to U$ を $\mathrm{End}(U)$ 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. $F, G, H \in \mathrm{End}(U)$ 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 $\boldsymbol{u} \in U$ 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．

$$\begin{aligned}F(G + H)(\boldsymbol{u}) &= F \circ (G + H)(\boldsymbol{u}) = F\Big((G + H)(\boldsymbol{u})\Big) = F\Big(G(\boldsymbol{u}) + H(\boldsymbol{u})\Big)\\ &= F(G(\boldsymbol{u})) + F(H(\boldsymbol{u})) = F \circ G(\boldsymbol{u}) + F \circ H(\boldsymbol{u}) = (F \circ G + F \circ H)(\boldsymbol{u}) = (FG + FH)(\boldsymbol{u})\\ &\quad (\uparrow\ F \text{ の線形性を用いた})\end{aligned}$$

□

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, $\mathcal{T} : M_n(\mathbb{R}) \to \mathrm{End}(\mathbb{R}^n)$ 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 $\tau : \mathcal{M} \to \mathcal{L}$ 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[63]:

(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 $\mathcal{T} : M_n(\mathbb{R}) \to \mathrm{End}(\mathbb{R}^n)$ 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 $\mathrm{End}(\mathbb{R}^m)$ 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 $T_F : C(Y) \to C(X)$ 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

前項において $\mathrm{End}(U)$ 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 $F : U \to U$ 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 $F : U \to U$ 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. $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ を $D = \dfrac{d}{dx}$ 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.

- $y^{(k)}(x) + a_{k-1}y^{(k-1)}(x) + a_{k-2}y^{(k-2)}(x) + \cdots + a_1 y^{(1)}(x) + a_0 y^{(0)}(x) = 0,$
- $D^k(y) + a_{k-1}D^{k-1}(y) + a_{k-2}D^{k-2}(y) + \cdots + a_1 D(y) + a_0 I(y) = \mathbf{0},$
- $\Big(D^k + a_{k-1}D^{k-1} + a_{k-2}D^{k-2} + \cdots + a_1 D + a_0 I\Big)(y) = \mathbf{0},$
- $\Phi(D)(y) = \mathbf{0}$　　(ただし, $\Phi(t) := t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ 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.

[63] 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, 本論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.

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, $\Phi_A(t) = t^2 - (a+b)t + (ad-bc)$ とおくと $\Phi(A) = O$ となることを確かめよ.

補足. この公式を移行することで $A^2 = (a+b)A - (ad-bc)E$ を得る. 一般に, この左辺よりも右辺のほうが計算が難しくないことが多い. つまり, この公式は A の冪の計算において有用な場合がある. 一般の正方行列に関するケーリー・ハミルトンの定理は 28.4 項で論ずる.

多項式は, $\Phi(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ 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 $F: U \to U$ 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, $\Phi(t) = \lambda t^k \Psi(t) \Longrightarrow \Phi(F) = \lambda F^k \Psi(F)$.

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) $\Phi(t) = \sum_{k=1}^r \Psi_k(t) \Longrightarrow \Phi(F) = \sum_{k=0}^r \Psi_k(F)$.

解答例: $\Psi_k(t) = \sum_{i=0}^n a_{k,i} t^i \ (k = 1\cdots, r)$ 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) $\Phi(t) = \Phi_1(t)\Phi_2(t) \Longrightarrow \Phi(F) = \Phi_1(F)\Phi_2(F)$

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 $\Psi_i(t) = a_i t^i \Phi_2(t)$ $(i = 0, \cdots, n)$ 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, $\mathrm{End}(U)$ 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. $F : U \to U$ を線形変換とし, $\Phi(t)$ と $\Psi(t)$ を多項式とすれば,

$$\text{各 } \boldsymbol{u} \in U \text{ について,} \qquad \Phi(F)\circ\Psi(F)(\boldsymbol{u}) = \Psi(F)\circ\Phi(F)(\boldsymbol{u}).$$

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. $\mathcal{S} : \mathcal{M} \to \mathcal{L}$ 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. 同様にして $\mathcal{S}(F^n) = \mathcal{S}(F)^n$ が得られ, したがって

$$\begin{aligned}
\mathcal{S}(\Psi(A)) &= \mathcal{S}(a_n F^n + a_{n-1}F^{n-1} + \cdots + a_1 F + a_0 I_{\mathcal{M}}) \\
&= a_n\mathcal{S}(F^n) + a_{n-1}\mathcal{S}(F^{n-1}) + \cdots + a_1\mathcal{S}(F) + a_0 S(I_{\mathcal{M}}) \\
&= a_n\mathcal{S}(F)^n + a_{n-1}\mathcal{S}(F)^{n-1} + \cdots + a_1\mathcal{S}(F) + a_0 I_{\mathcal{L}} = \Psi(\mathcal{S}(F)).
\end{aligned}$$

□

例 21.5.8. $m = n$ として定理 21.3.4 で与えた $\mathcal{T} : M_n(\mathbb{R}) \to \mathrm{End}(\mathbb{R}^n)$ 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.

22　線形空間の次元

前節において, 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. 本節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 $f : U \to V$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ および $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ 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 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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 $\boldsymbol{v}_1 \ldots, \boldsymbol{v}_n$ 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**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**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 $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ について, 例 20.2.4 および 20.3.2 において $\operatorname{Im} T_A = \langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \rangle$, $\operatorname{Ker} T_A = W_A$ 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, $\dim \operatorname{Im} T_A = \dim\langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \rangle = \operatorname{rank} A$, $\dim \operatorname{Ker} T_A = \dim W_A = n - \operatorname{rank} A$ である. この事実から一般の線形写像 f においても $\dim \operatorname{Im} f$ のことを線形写像 f の**階数**と呼び $\operatorname{rank} f$ と書く. また, $\dim \operatorname{Ker} f$ にも特別な記号が与えられており, これを $\operatorname{null} f$ と書き, f の**退化次数**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, 本論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. この事実は次項で示す (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 $f: U \to V$ 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．$f: U \to V$ 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 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ が取れる (つまり $\dim U = n$)．命題 21.1.5 より $f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_n)$ 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 (ルベーグの被覆次元), 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, このような不変量は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 項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 $A\boldsymbol{x} = \boldsymbol{b}$ の解全体の集合 W は, 解が存在するとすれば斉次方程式 $A\boldsymbol{x} = \boldsymbol{0}$ 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 $A\boldsymbol{x} = \boldsymbol{0}$ の解空間 W_A が次のように表されているとする.

$$W_A = \{ c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_k\boldsymbol{a}_k \mid c_1, \ldots, c_k \in \mathbb{R} \} = \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \rangle.$$

4.7 項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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ の中に不要なベクトルがないということであり, これは $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ 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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ 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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ 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. $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ に水増しがなければこれは 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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ 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 $A\boldsymbol{x} = \boldsymbol{b}$ 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 $n - \operatorname{rank} A$ 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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in A$ が存在する.

(ii) いかなる $n+1$ 個の組 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n+1} \in A$ 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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in A$ について, $m \geq n+1 \implies \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ は線形従属.

(iv) 各 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in A$ 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に制限しても, あるいは全体まで広げても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 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \in A$ 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 $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n \in \langle A \rangle$ 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 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ の線形結合で書けることを示そう. 仮に $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ の線形結合で書けないベクトル $\boldsymbol{a} \in A$ があるとすれば $n+1$ 個のベクトルの組 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n, \boldsymbol{a}$ 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 $A \subset \langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \rangle$ である. V の部分空間 $\langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \rangle$ は線形結合で閉じていることから $\langle A \rangle \subset \langle \boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \rangle$ であり, 以上より $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ 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 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n \in B$ 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 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{\dim V}$ 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, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ 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 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n+1} \in V$ 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, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n+1}$ は線形従属である. また, n 個の組 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ 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): $A := \{ \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \}$ について, いま示した (1) と命題 22.3.2(2) より得る.

(3): 集合 $A = \{ \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \}$ における線形独立な最大個数を k とする. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の中から線形独立な組 $\boldsymbol{u}_{n_1}, \cdots, \boldsymbol{u}_{n_k}$ を取れば, 命題 22.3.2(1) より $\boldsymbol{u}_{n_1}, \cdots, \boldsymbol{u}_{n_k}$ は $\langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ の基底でとなる. 仮定より $V = \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n \rangle$ であり, ゆえに k 個の組 $\boldsymbol{u}_{n_1}, \cdots, \boldsymbol{u}_{n_k}$ 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 $\boldsymbol{u}_{n_1}, \cdots, \boldsymbol{u}_{n_k}$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n+1}$ 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 $\boldsymbol{u}_i$ を $\boldsymbol{u}_i$ 自身に写す線形写像 $f : \langle \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n+1} \rangle \to V$ 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 $\langle \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n+1} \rangle$ と $\mathbb{R}^{n+1}$ は同型である. これらを結ぶ同型写像を $g : V \to \mathbb{R}^n$, $h : \mathbb{R}^{n+1} \to \langle \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n+1} \rangle$ とすれば, 単射線形写像 $g \circ f \circ h : \mathbb{R}^{n+1} \to \mathbb{R}^n$ 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, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n+1}$ 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 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in V$ が存在する.)

Proof. (1)⇒(2): 各 $k \in \mathbb{N}$ において組 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$ が線形独立となるようにベクトルの無限列 $\boldsymbol{v}_n$ を見つけたい. このような $\boldsymbol{v}_n$ が帰納的に取れることを示そう. いま, $\boldsymbol{v}_k$ までが得られているとし, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$ 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, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$ の線形結合で表せない元 $\boldsymbol{v}_{k+1} \in V$ が存在する. このとき, $k+1$ 個の組 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{k+1}$ 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 $\boldsymbol{v}_n$ を集めた無限集合 $\{ \boldsymbol{v}_n \mid n \in \mathbb{N} \}$ は線形独立である.

(2)⇒(3): 明らか.

(3)⇒(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はできない. 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はできない. □

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 $\{ \boldsymbol{e}_n \mid n \in \mathbb{N} \}$ 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 $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m \in W$ を取ろう. $A = W$ について命題 22.3.2(2) を適用すれば $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ 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 $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ は V の基底となり, $W = V$ を得る. □

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 $\{\boldsymbol{0}\}$ 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 $\boldsymbol{u} \neq \boldsymbol{0}$ からなり, $W = \{ r\boldsymbol{u} \mid r \in \mathbb{R} \}$ は $\boldsymbol{u}$ と $\boldsymbol{0}$ を通る直線に一致する.
- $\dim W = 2$ なる部分空間は, 基底 $\boldsymbol{u}_1, \boldsymbol{u}_2 \in W$ を持つ. $W = \{ r\boldsymbol{u}_1 + s\boldsymbol{u}_2 \mid r, s \in \mathbb{R} \}$ であり, これは $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{0}$ 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) $t_1, \ldots, t_{n+1}$ を相異なる実数とする. 任意の実数 $r_1, \ldots, r_{n+1}$ に対して, $p(t_i) = r_i$ $(i = 1, \ldots, n+1)$ 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: $I := \{ t_1, \ldots, t_{n+1} \} \subset \mathbb{R}$ とおけば, 命題 15.2.5 より $\mathbb{R}[x]_n \subset C(I)$ とみなせる. 写像 $p : I \to \mathbb{R}$ を $p(t_i) := r_i$ $(i = 1, \ldots, n+1)$ 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[64]. 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, $p \in \mathbb{R}[x]_n$. □

[64] 具体的に次のような基底があることからも分かる: $f_i : I \to \mathbb{R}$ を $f_i(t_i) := 1$, $f_i(t_j) := 0$ (ただし $j \neq i$) と定めれば, $f_1, \ldots, f_{n+1}$ は $C(I)$ の基底である. 実際, $\sum_{i=1}^{n+1} a_i f_i(x) = \boldsymbol{0}$ 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.

(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 $f: U\to V$ が存在する.

(2) $\dim U \geq \dim V$ ならば線形全射 $f: U\to V$ が存在する.

Proof. (1) $\dim U = n$ とおき, $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n$ 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 $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n\in V$ が取れる．このとき, 例 21.2.5(1) より $f(\boldsymbol{u}_i)=\boldsymbol{v}_i$ を満たす線形同型 $f: U\to\langle \boldsymbol{v}_1,\ldots,\boldsymbol{v}_n\rangle$ 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 $f: U\to V$ 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, $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m$ を V の基底とし, $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_m,\boldsymbol{u}_{m+1},\ldots,\boldsymbol{u}_{m+k}$ を U の基底とする．命題 20.1.11 により

$$f(\boldsymbol{u}_i)=\boldsymbol{v}_i\ (i=1,\ldots,m),\quad f(\boldsymbol{u}_{m+j})=\boldsymbol{0}_V\ (j=1,\ldots,k)$$

を満たす線形写像 $f: U\to V$ を取れば, $\operatorname{Im} f=\langle f(\boldsymbol{u}_1),\ldots f(\boldsymbol{u}_{m+k})\rangle=\langle \boldsymbol{v}_1,\ldots \boldsymbol{v}_m,\boldsymbol{0}_V,\ldots,\boldsymbol{0}_V\rangle = V$ ゆえ f は全射である. □

命題 22.4.5. 有限次元線形空間の間の線形写像 $f: U\to V$ 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 $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n$ とすれば命題 20.3.6(3) より $f(\boldsymbol{u}_1),\ldots,f(\boldsymbol{u}_n)\in V$ 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 $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m$ とすれば, f の全射性より $f(\boldsymbol{u}_i)=\boldsymbol{v}_i$ を満たす $\boldsymbol{u}_i\in U$ $(i=1,\ldots,m)$ が存在する．命題 20.1.7(3) より $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_m$ 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 $\dim U\geq m=\dim V$. □

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 $f: U\to V$ は存在しない.

(2) $\dim U < \dim V$ ならば線形全射 $f: U\to V$ 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 $f: U\to V$ について, 次を示せ.

(1) f は単射である $\iff \dim U=\dim\operatorname{Im} f$.

解答例. ($\Rightarrow$): $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n$ を U の基底とすれば, $\operatorname{Im} f=\langle f(\boldsymbol{u}_1),\cdots,f(\boldsymbol{u}_n)\rangle$ である．f の単射性より $f(\boldsymbol{u}_1),\cdots,f(\boldsymbol{u}_n)$ 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 $\dim\operatorname{Im} f = n=\dim U$.

($\Leftarrow$): $f(\boldsymbol{u}_1),\cdots,f(\boldsymbol{u}_n)$ を $\operatorname{Im} f$ の基底とすれば, f の線形性より $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n$ 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, $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n$ 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 $f: U\to\operatorname{Im} f$ は全単射である．とくに f は単射である. □

(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 $f * X \to X$ について f の単射性と全射性は同値であった (命題 19.6.1). 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. U を有限次元線形空間とし, $f : U \to U$ を線形写像とすれば次の 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. $V = U$ として練習 22.4.7(2) を適用し, f の全射性を得る.

(2)$\Rightarrow$(1): f を全射とすれば練習 22.4.7(2) より $\dim U = \dim \operatorname{Im} f$ である. ゆえに練習 22.4.7(1) から 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).

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とで, 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節に進むことを勧める.

命題 22.5.1 (命題 20.1.7(1) 参照)**.** $f : U \to V$ を線形写像とし, $A \subset U$ とする. このとき $\boldsymbol{x} \in \langle A \rangle$ ならば $f(\boldsymbol{x}) \in \langle f(A) \rangle$ 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 $f : A \to V$ を与えておく. このとき, f の拡張である線形写像 $\widetilde{f} : U \to V$ が唯一つ存在する.

命題 22.5.3 (命題 20.2.3 参照)**.** 線形写像 $f : U \to V$ 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 $\operatorname{Im} f$ を生成する.

命題 22.5.4 (命題 20.3.6 参照)**.** $f : U \to V$ 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)**.** $f : U \to V$ 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 $A \subset U$, $f : U \to V$ を線形写像とするとき, 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論の枠を越えており, 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)⇒(2): $U \simeq V$ であると仮定し, $f : U \to V$ 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)⇒(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 $f : A \to B$ 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 $\widetilde{f} : U \to V$ が唯一つ存在する. $\widetilde{f}$ が線形同型であることは命題 22.5.6 より得る.

(2)⇒(3): 定理 22.5.7 より得られる.

(3)⇒(2): 明らか. □

次は 22.4 項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) $f : A \to B$ が単射ならば, その拡張である線形単射 $\widetilde{f} : U \to V$ が存在する.

(2) $f : A \to B$ が全射ならば, その拡張である線形全射 $\widetilde{f} : U \to V$ 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 $f : U \to V$ が存在するならば, 単射 $p : A \to B$ が存在する.

(2) 線形全射 $f : U \to V$ が存在するならば, 全射 $p : A \to B$ が存在する.

発展 ($\mathbb{R}[x]$ と $\mathbb{R}^{\mathbb{N}}$ は線形同型か).

本節において, 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 $\mathbb{R}^n$ の完備性を学ぶだろう). 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.

23　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間の関係, および線形常微分方程式とその斉次方程式の解の間の関係には類似性が認められていた (命題 6.3.2 および 6.3.3). 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節で述べる. 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並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, 与えられた平面と並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節の後半で見る.

ところで, 本節および次節において, $\boldsymbol{u}-\boldsymbol{v}$ と表記すればよいところをわざわざ読みにくい表記で $-\boldsymbol{v}+\boldsymbol{u}$ 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　空間の並行移動

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. ここで改めて正確な定義を述べておこう.

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 $\boldsymbol{u} \in U$ に対して, A を $\boldsymbol{u}$ 方向に平行移動した集合を $A+\boldsymbol{u}$ あるいは $\boldsymbol{u}+A$ とかく (図 4). すなわち,

$$A+\boldsymbol{u} := \{\boldsymbol{a}+\boldsymbol{u} \mid \boldsymbol{a} \in A\}, \qquad \boldsymbol{u}+A := \{\boldsymbol{u}+\boldsymbol{a} \mid \boldsymbol{a} \in A\}.$$

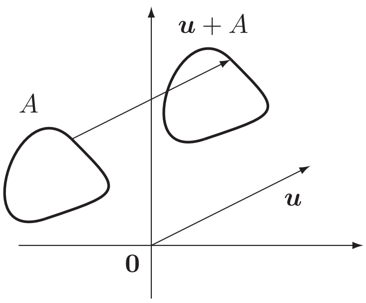


図 4: 図形 A の $\boldsymbol{u}$ 方向への平行移動

明らかに $A+\boldsymbol{u} = \boldsymbol{u}+A$ であり, この事実は和の演算の可換性 ($\boldsymbol{a}+\boldsymbol{u} = \boldsymbol{u}+\boldsymbol{a}$) に由来するものである. 以降では $\boldsymbol{u}+A$ という表記を主に用いることにする[65].

命題 23.1.2. A と $\boldsymbol{u}+A$ は対等である. すなわち, これらの元の個数は等しい.

Proof. 写像 $f : A \to U$ を $f(\boldsymbol{a}) := \boldsymbol{u}+\boldsymbol{a}$ で定めれば, $f(A) = \boldsymbol{u}+A$ 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. □

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 $\boldsymbol{u} \in W$ において, $\boldsymbol{u}+W = W$ 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, $\boldsymbol{u}+W \subset W$ を得る. また, 各 $\boldsymbol{v} \in W$ に対して $\boldsymbol{w} := -\boldsymbol{u}+\boldsymbol{v}$ とおけば $\boldsymbol{w} \in W$ であり, $\boldsymbol{v} = \boldsymbol{u}+\boldsymbol{w} \in \boldsymbol{u}+W$. つまり $W \subset \boldsymbol{u}+W$ である. □

[65] 実は, $A+\boldsymbol{u}$ のほうを採用していれば, 冒頭で述べた $-\boldsymbol{v}+\boldsymbol{u}$ 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 $\boldsymbol{u}+A$ と書くことにしたい.

次に述べる命題は, 命題 6.3.2 および 6.3.3 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[66]. 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. $f : U \to V$ を線形写像とし, $\boldsymbol{b} \in V$ とする. 更に $f(\boldsymbol{a}) = \boldsymbol{b}$ を満たす $\boldsymbol{a} \in U$ を一つ取って固定しよう (すなわち $\boldsymbol{a} \in f^{-1}(\boldsymbol{b})$). このとき次が成り立つ:

(1) 任意の $\boldsymbol{z} \in \operatorname{Ker} f$ に対し, $\boldsymbol{a} + \boldsymbol{z} \in f^{-1}(\boldsymbol{b})$ である.

(2) 任意の $\boldsymbol{y} \in f^{-1}(\boldsymbol{b})$ は, ある $\boldsymbol{z} \in \operatorname{Ker} f$ を用いて $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{z}$ と表せる.

(3) $f^{-1}(\boldsymbol{b}) = \boldsymbol{a} + \operatorname{Ker} f$

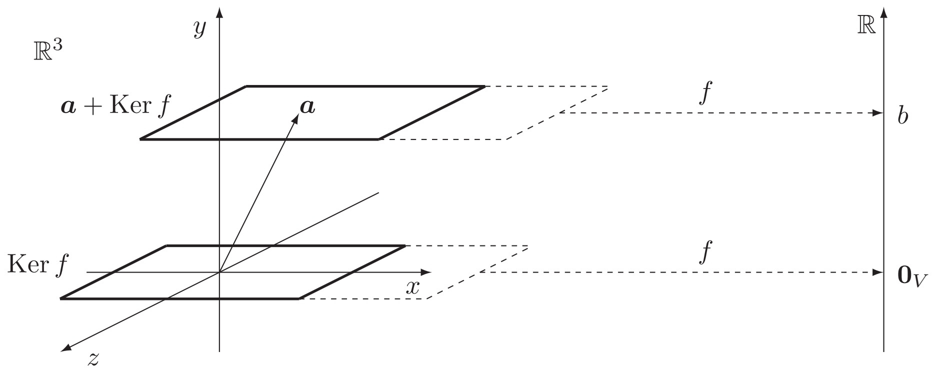


図 5: 線形写像の核とその並行移動

図 5 は線形写像 $f : \mathbb{R}^3 \to \mathbb{R}$ $(f(x, y, z) := y)$ について命題 23.1.4(3) を模式的に説明したものである. $\operatorname{Ker} f$ は x-z 平面に一致する. また, $\boldsymbol{a} := (x_1, b, z_1)$ とすれば $f(\boldsymbol{a}) = b$ である (つまり $\boldsymbol{a} \in f^{-1}(b)$). $\operatorname{Ker} f$ を $\boldsymbol{a}$ 方向にずらしたものが $\boldsymbol{a} + \operatorname{Ker} f$ 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 平面と並行な方向に無限に広がる平面である. つまり, $\operatorname{Ker} f$ を真上に持ち上げたもの ($(0, b, 0)$ 方向に並行移動させた平面) とも一致する. 命題 23.1.4(3) によれば $f^{-1}(b) = \boldsymbol{a} + \operatorname{Ker} f$ であり, $\boldsymbol{a} + \operatorname{Ker} f$ 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.

<u>命題 23.1.4 の証明.</u> (1) : $\boldsymbol{z} \in \operatorname{Ker} f$ とすると, $f(\boldsymbol{a} + \boldsymbol{z}) = f(\boldsymbol{a}) + f(\boldsymbol{z}) = \boldsymbol{b} + \boldsymbol{0} = \boldsymbol{b}$. ゆえに $\boldsymbol{a} + \boldsymbol{z} \in f^{-1}(\boldsymbol{b})$ である.

(2) : $\boldsymbol{y} \in f^{-1}(\boldsymbol{b})$ とする. $\boldsymbol{z} := -\boldsymbol{a} + \boldsymbol{y}$ とおこう. このとき $f(\boldsymbol{z}) = f(-\boldsymbol{a} + \boldsymbol{y}) = -f(\boldsymbol{a}) + f(\boldsymbol{y}) = -\boldsymbol{b} + \boldsymbol{b} = \boldsymbol{0}$ より $\boldsymbol{z} \in \operatorname{Ker} f$ である. また, $\boldsymbol{z}$ の定め方から $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{z}$ であり, 我々は主張を得た.

(3) : 両方の包含関係 $f^{-1}(\boldsymbol{b}) \subset \boldsymbol{a} + \operatorname{Ker} f$ および $f^{-1}(\boldsymbol{b}) \supset \boldsymbol{a} + \operatorname{Ker} f$ 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. □

命題 23.1.2 および 23.1.4(3) より, $f^{-1}(\boldsymbol{b}) \neq \emptyset$ であるとき $f^{-1}(\boldsymbol{b})$ と $\operatorname{Ker} f$ 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[67]:

系 23.1.5. 線形写像 $f : U \to V$ および各 $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \operatorname{Im} f$ について, $f^{-1}(\boldsymbol{b}_1)$ と $f^{-1}(\boldsymbol{b}_2)$ は対等である.

Proof. $f^{-1}(\boldsymbol{b}_1)$ と $f^{-1}(\boldsymbol{b}_2)$ はそれぞれ $\operatorname{Ker} f$ と対等であり, したがって $f^{-1}(\boldsymbol{b}_1)$ と $f^{-1}(\boldsymbol{b}_2)$ も対等である. □

---

[66] 実際, (m, n)-行列 A による線形写像 $T_A : \mathbb{R}^n \to \mathbb{R}^m$ に対して命題 23.1.4 を適用したものが命題 6.3.2 である. また, $T(f) := \sum_{k=0}^{n} \alpha_k \dfrac{d^k}{dx^k} f$ (ただし各 $\alpha_k \in C^\infty(\mathbb{R})$ はあらかじめ決めておいた関数) で定められる線形写像 $T : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ に対して命題 23.1.4 を適用したものが命題 6.3.3 である.

[67]「元の個数」といっても $\operatorname{Ker} f$ 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.

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. □

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 $f : U \to V$ において, $\boldsymbol{z}_1, \cdots, \boldsymbol{z}_n$ が $\mathrm{Ker}\, f$ の基底であり, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m \in U$ について $f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_m)$ が $\mathrm{Im}\, f$ の基底であるならば, $m+n$ 個の組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ 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 $\sum_{i=1}^m r_i \boldsymbol{u}_i + \sum_{j=1}^n s_j \boldsymbol{z}_j = \boldsymbol{0}_U$ を仮定する. これらに f をほどこすと

$$\begin{aligned} \sum_{i=1}^m r_i f(\boldsymbol{u}_i) + \sum_{j=1}^n s_j f(\boldsymbol{z}_j) &= f(\boldsymbol{0}_U) \\ \sum_{i=1}^m r_i f(\boldsymbol{u}_i) + \sum_{j=1}^n s_j \boldsymbol{0}_V &= \boldsymbol{0}_V \\ \sum_{i=1}^m r_i f(\boldsymbol{u}_i) &= \boldsymbol{0}_V. \end{aligned}$$

$f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_m)$ の線形独立性より $r_1 = \cdots = r_m = 0$ を得る. つまり $\sum_{j=1}^n s_j \boldsymbol{z}_j = \boldsymbol{0}_U$ であり, $\boldsymbol{z}_1, \cdots, \boldsymbol{z}_n$ の線形独立性より $s_1 = \cdots = s_n = 0$ を得る. 以上より $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ は線形独立である.

次に各 $\boldsymbol{y} \in U$ が $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ の線形結合で書けることを示そう. $\boldsymbol{b} := f(\boldsymbol{y})$ とおくと $\boldsymbol{b} \in \mathrm{Im}\, f$ ゆえ $\boldsymbol{b} = \sum_{i=1}^m r_i f(\boldsymbol{u}_i)$ と書ける. このとき $\boldsymbol{a} = \sum_{i=1}^m r_i \boldsymbol{u}_i$ とおくと, $f(\boldsymbol{a}) = f(\sum_{i=1}^m r_i \boldsymbol{u}_i) = \sum_{i=1}^m r_i f(\boldsymbol{u}_i) = \boldsymbol{b}$ より $\boldsymbol{a} \in f^{-1}(\boldsymbol{b})$ である. この $\boldsymbol{a}, \boldsymbol{b}$ に対して命題 23.1.4(2) を適用すれば, $\boldsymbol{y} \in f^{-1}(\boldsymbol{b})$ は $\boldsymbol{z} \in \mathrm{Ker}\, f$ をもちいて $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{z}$ と書ける. $\boldsymbol{z}_1, \cdots, \boldsymbol{z}_n$ は $\mathrm{Ker}\, f$ の基底であったから $\boldsymbol{z} = \sum_{j=1}^n s_j \boldsymbol{z}_j$ と書け, $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{z} = \sum_{i=1}^m r_i \boldsymbol{u}_i + \sum_{j=1}^n s_j \boldsymbol{z}_j$ を得る. □

いまの定理より直ちに次の公式を得る.

定理 23.2.3 (線形写像の次元公式)**.** 有限次元線形空間 U を定義域とする線形写像 $f : U \to V$ において,

$$\dim U = \dim \mathrm{Ker}\, f + \dim \mathrm{Im}\, f.$$

上の公式によれば, U および $\mathrm{Im}\, f$, $\mathrm{Ker}\, f$ の三つの空間のうち二つの次元が分かっていれば, 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れば, 残りの空間についてもある程度のことが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, $H := \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_m \rangle$ とおけば, $f|_H : H \to \mathrm{Im}\, f$ 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 $T_A : \mathbb{R}^n \to \mathbb{R}^m$ において, 例 22.1.3(2) より $\dim \mathrm{Im}\, T_A = \mathrm{rank}\, A$, $\dim \mathrm{Ker}\, f = n - \mathrm{rank}\, A$ である. ゆえに

$$\dim \mathrm{Ker}\, f + \dim \mathrm{Im}\, f = (n - \mathrm{rank}\, A) + \mathrm{rank}\, A = n = \dim \mathbb{R}^n$$

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, $\mathrm{Ker}\, f$ 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, $f : U \to V$ を線形写像とする．いま, $\boldsymbol{z}_1, \cdots, \boldsymbol{z}_n$ が $\mathrm{Ker}\, f$ 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 $\boldsymbol{z}_1, \cdots, \boldsymbol{z}_n, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ を U の基底とすることができる．このとき, $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_m)$ が $\mathrm{Im}\, f$ の基底となることを示せ.

解答例: 次元公式により $\dim \mathrm{Im}\, f = \dim U - \dim \mathrm{Ker}\, f = (n+m) - n = m$ である．命題 20.2.3 より

$$\mathrm{Im}\, f = \langle f(\boldsymbol{z}_1), \cdots, f(\boldsymbol{z}_n), f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_m)\rangle = \langle \boldsymbol{0}_V, \cdots, \boldsymbol{0}_V, f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_m)\rangle = \langle f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_m)\rangle.$$

ゆえに $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_m)$ は $\mathrm{Im}\, f$ を生成する．更に, 命題 22.3.3(3) より, $\dim \mathrm{Im}\, f$ 個の組 $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_m)$ は $\mathrm{Im}\, f$ の基底である. □

一方, $f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_m)$ が $\mathrm{Im}\, f$ の基底であるとき, 線形独立な組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ 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 $\mathrm{Ker}\, f$ 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 $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ において, $\mathrm{Im}\, T_A = \langle \boldsymbol{e}_1 \rangle = \langle T_A(\boldsymbol{e}_1)\rangle$ および $\mathrm{Ker}\, T_A = \langle \boldsymbol{e}_2 \rangle$ である．よって, $T_A(\boldsymbol{e}_1)$ は $\mathrm{Im}\, T_A$ の基底であり, $\boldsymbol{e}_1$ に $\boldsymbol{u} := \boldsymbol{e}_1 + \boldsymbol{e}_2$ を付け加えた $\boldsymbol{e}_1, \boldsymbol{u}$ は $\mathbb{R}^2$ の基底となる．このとき, $\boldsymbol{u} \notin \mathrm{Ker}\, T_A$ ゆえ $\boldsymbol{u}$ は $\mathrm{Ker}\, T_A$ の基底ではない.

23.3 商空間 (発展)

代入元 $\boldsymbol{u}$ を動かすことにより写像 $f : U \to V$ の値 $f(\boldsymbol{u})$ を変化させるにはどうすればよいか, あるいは逆に $f(\boldsymbol{u})$ を変化させないような代入元 $\boldsymbol{u}$ 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) $\boldsymbol{u}$ 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 $f : \mathbb{R}^n \to \mathbb{R}$ および点 $\boldsymbol{a} \in \mathbb{R}^n$ が与えられているとし, $\boldsymbol{x} = \boldsymbol{a}$ をわずかだけ動かして $f(\boldsymbol{x})$ の値を変化させることを考えよう．このとき, 勾配 $\mathrm{grad}\, f(\boldsymbol{a})$ 方向に動かすとき $f(\boldsymbol{x})$ の変化率[68]は最大となる．ここで,

$$\mathrm{grad}\, f(\boldsymbol{a}) = \left(\frac{\partial f}{\partial x_1}(\boldsymbol{a}), \frac{\partial f}{\partial x_2}(\boldsymbol{a}), \ldots, \frac{\partial f}{\partial x_n}(\boldsymbol{a}) \right)$$

である．また, $\mathrm{grad}\, f(\boldsymbol{a})$ に直行する方向に動かしたときの $f(\boldsymbol{x})$ の変化率は 0 である.

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 $f : U \to V$ および $\boldsymbol{u}, \boldsymbol{z} \in U$ において, $f(\boldsymbol{u}) = f(\boldsymbol{u} + \boldsymbol{z})$ となるための必要十分条件は $\boldsymbol{z} \in \mathrm{Ker}\, f$ となることである.

Proof. $\boldsymbol{b} := f(\boldsymbol{u})$ とおこう．$\boldsymbol{z} \in \mathrm{Ker}\, f$ を仮定すれば, 命題 23.1.4(1) より $f(\boldsymbol{u} + \boldsymbol{z}) = \boldsymbol{b}$ (すなわち $f(\boldsymbol{u}) = f(\boldsymbol{u}+\boldsymbol{z})$) である．逆に, $f(\boldsymbol{u}) = f(\boldsymbol{u}+\boldsymbol{z}) = \boldsymbol{b}$ ならば $f(\boldsymbol{z}) = f(-\boldsymbol{u}+\boldsymbol{u}+\boldsymbol{z}) = -f(\boldsymbol{u}) + f(\boldsymbol{u}+\boldsymbol{z}) = -\boldsymbol{b} + \boldsymbol{b} = \boldsymbol{0}$ より $\boldsymbol{z} \in \mathrm{Ker}\, f$ を得る. □

つまり, $\mathrm{Ker}\, f$ と並行な方向にベクトル $\boldsymbol{u}$ を移動させるか否かで $f(\boldsymbol{u})$ 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 $\mathrm{Ker}\, f$ 方向への移動は無意味である．そこで, U の中で $\mathrm{Ker}\, f$ 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 $\mathrm{Ker}\, f$ 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.

[68] ここでいう変化率とは方向微分のこと.

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 を並行移動させた集合 (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,

$$U/W := \{\, \boldsymbol{u} + W \mid \boldsymbol{u} \in U \,\}.$$

集合 $\boldsymbol{u} + W$ を商空間 U/W の元とみなすとき, これを略して $[\boldsymbol{u}]$ と書く.

$\boldsymbol{u} = \boldsymbol{u} + \boldsymbol{0} \in \boldsymbol{u} + W = [\boldsymbol{u}]$ より $\boldsymbol{u} \in [\boldsymbol{u}]$ であることに注意せよ．また, $[\boldsymbol{0}] = \boldsymbol{0} + W = W$ 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 と並行な平面によって構成される集合である．図 6 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.

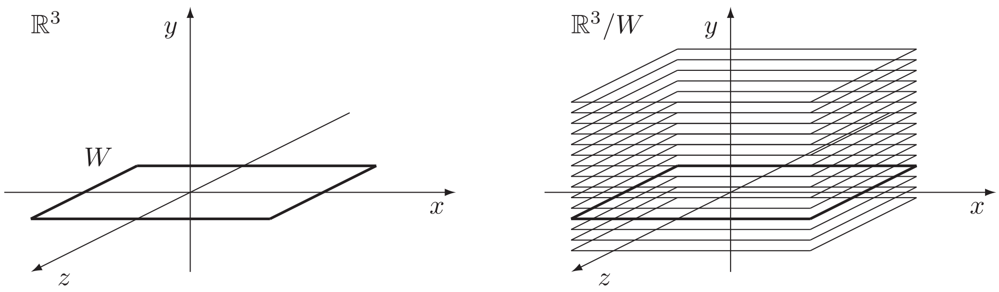


図 6: $\mathbb{R}^3$ 上の平面 W による商空間 $\mathbb{R}^3/W$

U の元 $\boldsymbol{u}$ の代わりに U/W の元 $[\boldsymbol{u}]$ 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 $\boldsymbol{u} \in U$ に対して, $[\boldsymbol{u}] \in U/W$ を対応させる写像 $q : U \to U/W$ を**商写像** (**quotient map**) と言う.

商写像 $q : U \to U/W$ は全射である．実際, U/W の各元 $\boldsymbol{u} + W$ に対して, $q(\boldsymbol{u}) = \boldsymbol{u} + W$ である.

次の命題は, $\boldsymbol{u}$ と $\boldsymbol{u} + \boldsymbol{w}$ が商空間 U/W において同じ元を表すこと (すなわち $[\boldsymbol{u}] = [\boldsymbol{u} + \boldsymbol{w}]$) と $\boldsymbol{w} \in W$ 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 $\boldsymbol{u}$ の代わりに $[\boldsymbol{u}]$ を考えるということは, W と並行な方向を捨象することを意味している.

命題 23.3.6. 各 $\boldsymbol{u}, \boldsymbol{v} \in U$ について, 次の条件はすべて同値である:

$$(1) \quad [\boldsymbol{u}] = [\boldsymbol{v}], \quad (2) \quad -\boldsymbol{v} + \boldsymbol{u} \in W, \quad (3) \quad \boldsymbol{v} - \boldsymbol{u} \in W, \quad (4) \quad \boldsymbol{v} \in [\boldsymbol{u}].$$

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)⇔(3) は明らかである.

(1)⇒(2): $[\boldsymbol{u}] = [\boldsymbol{v}]$ (すなわち $\boldsymbol{u} + W = \boldsymbol{v} + W$) を仮定しよう．すると $\boldsymbol{u} \in [\boldsymbol{u}] = \boldsymbol{v} + W$ ゆえ, ある $\boldsymbol{w} \in W$ を用いて $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}$ と書くことができる．この両辺に $-\boldsymbol{v}$ を加えて $-\boldsymbol{v} + \boldsymbol{u} = \boldsymbol{w} \in W$ を得る.

(2)⇐(1): $-\boldsymbol{v} + \boldsymbol{u} \in W$ を仮定しよう．$[\boldsymbol{u}] = [\boldsymbol{v}]$ を示すには $[\boldsymbol{u}] \subset [\boldsymbol{v}]$ および $[\boldsymbol{v}] \subset [\boldsymbol{u}]$ を示せばよい．まず $[\boldsymbol{u}] \subset [\boldsymbol{v}]$ を示すために $\boldsymbol{a} \in [\boldsymbol{u}]$ を任意に取る．すると, ある $\boldsymbol{w} \in W$ を用いて $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{w}$ と書くことができる．このとき, 別の $\boldsymbol{w}' \in W$ を用いて $\boldsymbol{a} = \boldsymbol{v} + \boldsymbol{w}'$ 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.

$$\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{w} = (\boldsymbol{v} - \boldsymbol{v}) + \boldsymbol{u} + \boldsymbol{w} = \boldsymbol{v} + (-\boldsymbol{v} + \boldsymbol{u}) + \boldsymbol{w}.$$

つまり, $\boldsymbol{w}' := (-\boldsymbol{v} + \boldsymbol{u}) + \boldsymbol{w}$ とすれば, $(-\boldsymbol{v} + \boldsymbol{u}), \boldsymbol{w} \in W$ および W が和の演算について閉じていることから $\boldsymbol{w}' \in W$ であり, $\boldsymbol{a} = \boldsymbol{v} + \boldsymbol{w}' \in \boldsymbol{v} + W$ を得る．いま $[\boldsymbol{u}] \subset [\boldsymbol{v}]$ であることが分かった．$\boldsymbol{u}$ と $\boldsymbol{v}$ の立場を入れ替えることで逆向きの包含関係 $[\boldsymbol{v}] \subset [\boldsymbol{u}]$ 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 $-\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} - \boldsymbol{u} \in W$ を用いればよい.

(1)⇒(4): $\boldsymbol{v} \in [\boldsymbol{v}] = [\boldsymbol{u}]$ より直ちに分かる.

(4)⇒(3): $\boldsymbol{v} \in [\boldsymbol{u}] = \boldsymbol{u} + W$ を仮定すれば, ある $\boldsymbol{w} \in W$ を用いて $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ と書ける．ゆえに $\boldsymbol{v} - \boldsymbol{u} = \boldsymbol{w} \in W$. □

系 23.3.7. 各 $\boldsymbol{w} \in U$ について, $[\boldsymbol{u}+\boldsymbol{w}] = [\boldsymbol{u}] \iff \boldsymbol{w} \in W$.

Proof. 前命題の $\boldsymbol{u}$ および $\boldsymbol{v}$ として, $\boldsymbol{u}+\boldsymbol{w}$ および $\boldsymbol{u}$ 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 と並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 $\mathbb{R}^2/X \to Y$ は全単射である.

(2) 上の例において $\mathbb{R}^2/X$ との間に 1 対 1 対応を与える直線を Y に限る必要はなく, 直線 X と並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 $\mathbb{R}^2/X \to L$ は全単射である.

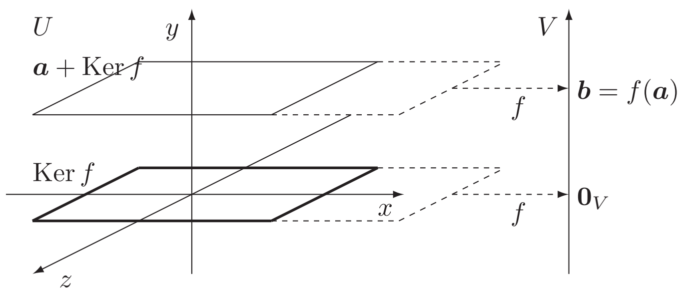
(3) $\mathbb{R}^2$ における直線 $y = \frac{1}{2}x$ 上の点全体を L とすれば, $\mathbb{R}^2/L$ と y 軸との間でも 1 対 1 対応がつく.

(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, 次節で詳しく述べよう.

線形写像 $f : U \to V$ に対して, f の値が動かないような代入元 $\boldsymbol{u}$ の移動とは, $\operatorname{Ker} f$ と並行な方向に移動させることに他ならなかった. そこで, $\operatorname{Ker} f$ 方向を捨象した $U/\operatorname{Ker} f$ を定義域とする f の代わりとなる写像 $F : U/\operatorname{Ker} f \to V$ を考えることができる.

例 23.4.2. $f : U \to V$ を線形写像とする. 各 $\boldsymbol{b} \in \operatorname{Im} f$ に対して $f(\boldsymbol{a}) = \boldsymbol{b}$ なる $\boldsymbol{a} \in U$ を取れば, 命題 23.1.4(3) より $f^{-1}(\boldsymbol{b}) = \boldsymbol{a} + \operatorname{Ker} f = [\boldsymbol{a}] \in U/\operatorname{Ker} f$ である.



そこで, 各 $f^{-1}(\boldsymbol{b}) = [\boldsymbol{a}] \in U/\operatorname{Ker} f$ に対して $\boldsymbol{b} \in V$ を対応させる写像を $F : U/\operatorname{Ker} f \to V$ 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 $U/\operatorname{Ker} f$ 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.

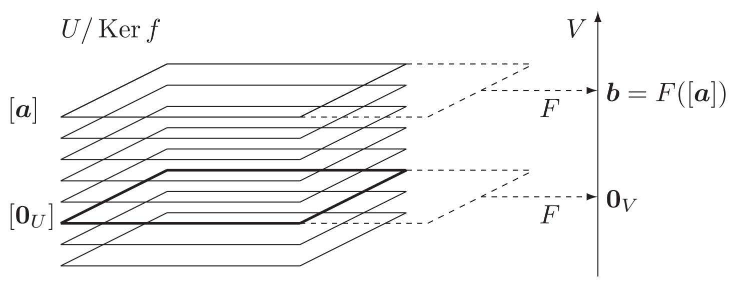


図 7: 線形写像 f が誘導する写像 $F : U/\operatorname{Ker} f \to V$

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(準同型定理) 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 $\boldsymbol{u} \in U$ に対して, 商空間 U/W の要素で $\boldsymbol{u}$ を含むものは $[\boldsymbol{u}]$ 唯一つのみである.

(2) 各 $L \in U/W$ および $\boldsymbol{u} \in L$ について $[\boldsymbol{u}] = L$. とくに $\boldsymbol{w} \in W$ について $[\boldsymbol{w}] = W = [\boldsymbol{0}]$.

(3) 各 $L \in U/W$ および $\boldsymbol{u} \in U$ について, $[\boldsymbol{u}] = L \iff \boldsymbol{u} \in L$.

(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[69].)

Proof. (1): U/W の二つの元 $[\boldsymbol{a}] = \boldsymbol{a} + W$ と $[\boldsymbol{b}] = \boldsymbol{b} + W$ がともに $\boldsymbol{u}$ を含むと仮定する. すると, ある $\boldsymbol{w}, \boldsymbol{w}' \in W$ を用いて $\boldsymbol{u} = \boldsymbol{a} + \boldsymbol{w} = \boldsymbol{b} + \boldsymbol{w}'$ と書くことができる. このとき, $-\boldsymbol{b} + \boldsymbol{a} = \boldsymbol{w}' - \boldsymbol{w} \in W$ であり, 命題 23.3.6 より $[\boldsymbol{a}] = [\boldsymbol{b}]$. すなわち, $\boldsymbol{u}$ を含む U/W の要素は一つしかない (注意: $[\boldsymbol{u}]$ は $\boldsymbol{u}$ を含む U/W の要素ゆえ $[\boldsymbol{u}] = [\boldsymbol{a}] = [\boldsymbol{b}]$ が成り立っている).

(2): $L \in U/W$ および $\boldsymbol{u} \in L$ について, $[\boldsymbol{u}]$ と L はともに $\boldsymbol{u}$ 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) より唯一つしかないことから, $[\boldsymbol{u}] = L$ である.

(3): ($\Leftarrow$) は (2) で示している. また, $[\boldsymbol{u}] = L$ とすれば $\boldsymbol{u} \in [\boldsymbol{u}] = L$ ゆえ $\boldsymbol{u} \in L$ である.

(4): $L, L' \in U/W$ および $L \cap L' \neq \emptyset$ とすれば $\boldsymbol{u} \in L \cap L'$ が取れる. このとき, (1) より $L = L' = [\boldsymbol{u}]$ 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. $\boldsymbol{u} \in U$ の属するグループが $[\boldsymbol{u}]$ 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[70] $\simeq$ 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.

$$[x] := \{\, y \in X \mid y \simeq x \,\}.$$

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.

---

[69] 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.

[70] 同値関係の定義は 21.2 項のコラムを見よ.

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 $[y] \subset [x]$ も同様にして示すことができる．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) と同様にして示される． □

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 $\boldsymbol{u}, \boldsymbol{v}$ に対して, 記号 $\boldsymbol{u} \simeq \boldsymbol{v}$ を「$-\boldsymbol{v}+\boldsymbol{u} \in W$ 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 $\boldsymbol{u} \in U$ について, いま定めた $\simeq$ に関する $\boldsymbol{u}$ の同値類と集合 $\boldsymbol{u}+W$ は一致する．

  Proof. 命題 23.3.6 では $\boldsymbol{u}+W$ のことを $[\boldsymbol{u}]$ と書いていたため, これと区別して同値関係 $\simeq$ に関する $\boldsymbol{u}$ の同値類を記号 $[[\boldsymbol{u}]]$ で表そう．すなわち, $[[\boldsymbol{u}]] := \{\, \boldsymbol{v} \in U \mid \boldsymbol{v} \simeq \boldsymbol{u} \,\}$ と定める．我々が示すべきことは $[[\boldsymbol{u}]] = [\boldsymbol{u}]$ であり, これは命題 23.3.6 より得られる:

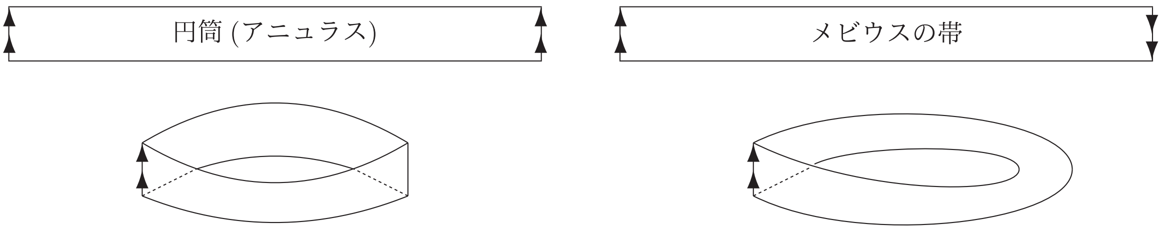
  $$\boldsymbol{v} \in [[\boldsymbol{u}]] \iff \boldsymbol{u} \simeq \boldsymbol{v} \iff -\boldsymbol{v}+\boldsymbol{u} \in W \underset{\uparrow \text{命題 } 23.3.6}{\iff} \boldsymbol{v} \in [\boldsymbol{u}].$$

  □

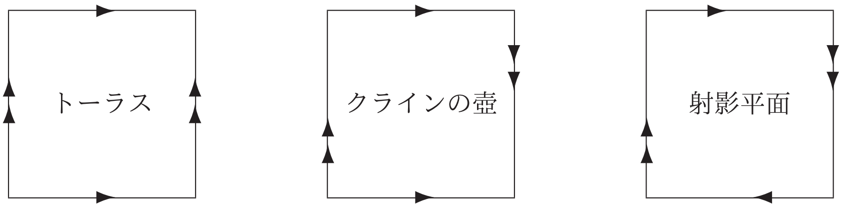
すなわち．この同値関係による商集合 $U/{\simeq}$ と本節で与えた商空間 U/W は一致する．

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る. 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だろうから, 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

$$A = \left\{\,(x,y,z) \in S^2 \,\middle|\, |z| \leq \frac{1}{10}\,\right\}$$

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対称な位置にある. 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 $F : S^2/\simeq \to P$ は全単射である. □

24 準同型定理と短完全列 (発展)

前節において, 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. 本節では, 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準同型定理を導く. 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) においても準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, 準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.

$$各\ [\boldsymbol{u}], [\boldsymbol{v}] \in U/W\ および\ r \in \mathbb{R}\ に対して,\ [\boldsymbol{u}] + [\boldsymbol{v}] := [\boldsymbol{u}+\boldsymbol{v}],\ r[\boldsymbol{u}] := [r\boldsymbol{u}].$$

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 $L = [\boldsymbol{u}] = [\boldsymbol{u}']$ としよう. このとき L のスカラー倍 rL は, $[r\boldsymbol{u}]$ と $[r\boldsymbol{u}']$ のいずれであると定めるべきだろうか. もし $[r\boldsymbol{u}] \neq [r\boldsymbol{u}']$ ならば L の代表として $\boldsymbol{u}$ と $\boldsymbol{u}'$ のいずれを選ぶのか (あるいは $\boldsymbol{u}, \boldsymbol{u}'$ 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 $[r\boldsymbol{u}] = [r\boldsymbol{u}']$ が成り立ち, L の中の

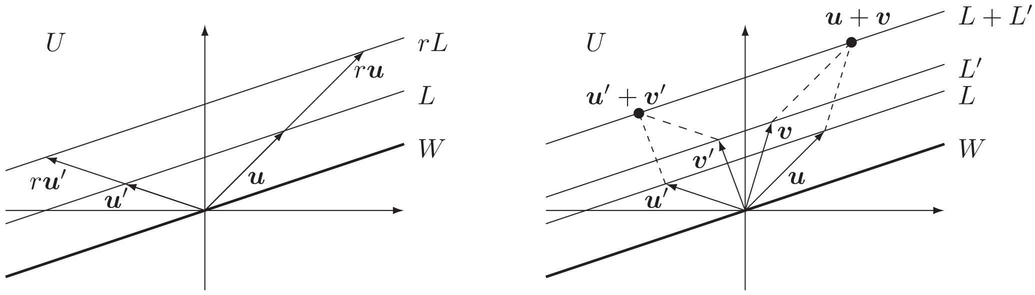


図 8: $L \in U/W$ のスカラー倍および $L' \in U/W$ との和

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 の代表元に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 8).

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:

$$(1)\ \ [\boldsymbol{u}] = [\boldsymbol{u}'] \Longrightarrow [r\boldsymbol{u}] = [r\boldsymbol{u}']. \qquad (2)\ \ [\boldsymbol{u}] = [\boldsymbol{u}']\ かつ\ [\boldsymbol{v}] = [\boldsymbol{v}'] \Longrightarrow [\boldsymbol{u}+\boldsymbol{v}] = [\boldsymbol{u}'+\boldsymbol{v}'].$$

(1): $[\boldsymbol{u}] = [\boldsymbol{u}']$ とすれば命題 23.3.6 より $-\boldsymbol{u}'+\boldsymbol{u} \in W$ であり, ゆえに $-r\boldsymbol{u}'+r\boldsymbol{u} = r(-\boldsymbol{u}'+\boldsymbol{u}) \in W$. 再び命題 23.3.6 より $[r\boldsymbol{u}] = [r\boldsymbol{u}']$ を得る.

(2): 仮定より $-\boldsymbol{u}'+\boldsymbol{u}, -\boldsymbol{v}'+\boldsymbol{v} \in W$ であり, ゆえに

$$-(\boldsymbol{u}'+\boldsymbol{v}') + (\boldsymbol{u}+\boldsymbol{v}) = (-\boldsymbol{u}'+\boldsymbol{u}) + (-\boldsymbol{v}'+\boldsymbol{v}) \in W$$

である[71]. したがって $[\boldsymbol{u}+\boldsymbol{v}] = [\boldsymbol{u}'+\boldsymbol{v}']$. □

[71] 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:

$$\begin{aligned} -(\boldsymbol{u}'+\boldsymbol{v}') + (\boldsymbol{u}+\boldsymbol{v}) &= (-\boldsymbol{v}'-\boldsymbol{u}') + (\boldsymbol{u}+\boldsymbol{v}) = -\boldsymbol{v}' + (-\boldsymbol{u}'+\boldsymbol{u}) + \boldsymbol{v} \\ &\in -\boldsymbol{v}' + \underline{(W+\boldsymbol{v})} = -\boldsymbol{v}' + \underline{(\boldsymbol{v}+W)} = (-\boldsymbol{v}'+\boldsymbol{v}) + W = [-\boldsymbol{v}'+\boldsymbol{v}] = W. \end{aligned}$$

非可換の世界では一般に $(W+\boldsymbol{v}) = (\boldsymbol{v}+W)$ は成立せず, 上の式変形を行うには条件 $(W+\boldsymbol{v}) = (\boldsymbol{v}+W)$ 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.

上の補題により, 任意の $\boldsymbol{u}, \boldsymbol{v} \in U$ 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.

$$[\boldsymbol{u}+\boldsymbol{v}] = [\boldsymbol{u}] + [\boldsymbol{v}], \qquad [r\boldsymbol{u}] = r[\boldsymbol{u}].$$

$[\mathbf{0}] = W \in U/W$ 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 $q : U \to U/W$ は線形全射であり, $\operatorname{Ker} q = W$.

Proof. 線形性 (iii) は直ちに確認できる:

$$q(r\boldsymbol{u} + s\boldsymbol{v}) = [r\boldsymbol{u} + s\boldsymbol{v}] = [r\boldsymbol{u}] + [s\boldsymbol{v}] = r[\boldsymbol{u}] + s[\boldsymbol{v}] = rq(\boldsymbol{u}) + sq(\boldsymbol{v}).$$

また, $[\mathbf{0}] = W$ に注意すると,

$$\boldsymbol{u} \in \operatorname{Ker} q \iff q(\boldsymbol{u}) = [\mathbf{0}] \iff [\boldsymbol{u}] = W \iff \boldsymbol{u} \in W \ (\text{命題 } 23.5.3(3)).$$

すなわち, $\operatorname{Ker} q = W$. □

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 $q : U \to U/W$ に対して次元公式を適用すればよい. q の全射性より $U/W = \operatorname{Im} q$ であること, また $\operatorname{Ker} q = W$ に注意すれば,

$$\begin{aligned} \dim \operatorname{Im} q + \dim \operatorname{Ker} q &= \dim U \\ \dim U/W + \dim W &= \dim U \\ \dim U/W &= \dim U - \dim W. \end{aligned}$$

□

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 $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n$ にいくつかのベクトルを付け加えて $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n, \boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ を U の基底とすれば $[\boldsymbol{u}_1], \cdots, [\boldsymbol{u}_m]$ 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

$$\operatorname{Im} q = \langle q(\boldsymbol{w}_1), \cdots, q(\boldsymbol{w}_n), q(\boldsymbol{u}_1), \cdots, q(\boldsymbol{u}_m)\rangle = \langle [\mathbf{0}], \cdots, [\mathbf{0}], [\boldsymbol{u}_1], \cdots, [\boldsymbol{u}_m]\rangle = \langle [\boldsymbol{u}_1], \cdots, [\boldsymbol{u}_m]\rangle.$$

ゆえに $[\boldsymbol{u}_1], \cdots, [\boldsymbol{u}_m]$ は $\operatorname{Im} q = U/W$ を生成する. 更に, 命題 22.3.3(3) より, $\dim U/W$ 個の組 $[\boldsymbol{u}_1], \cdots, [\boldsymbol{u}_m]$ は U/W の基底である. □

24.2 準同型定理

例 23.4.2 で述べた事実の詳細は次の通りである.

定理 24.2.1 (準同型定理)**.** 線形写像 $f : U \to V$ について, $U/\operatorname{Ker} f$ と $\operatorname{Im} f$ は線形同型である.

Proof. 写像 $F : U/\operatorname{Ker} f \to \operatorname{Im} f$ を $F([\boldsymbol{u}]) := f(\boldsymbol{u})$ 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「$[\boldsymbol{u}] = [\boldsymbol{v}]$ ならば $f(\boldsymbol{u}) = f(\boldsymbol{v})$」であることを確認しよう. 実際, $[\boldsymbol{u}] = [\boldsymbol{v}]$ ならば $-\boldsymbol{v}+\boldsymbol{u} \in \operatorname{Ker} f$ であり, これは $f(-\boldsymbol{v}+\boldsymbol{u}) = \mathbf{0}_V$ を意味する. つまり $-f(\boldsymbol{v})+f(\boldsymbol{u}) = f(-\boldsymbol{v}+\boldsymbol{u}) = \mathbf{0}_V$ であり, $f(\boldsymbol{u}) = f(\boldsymbol{v})$ 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:

$$F(r[\boldsymbol{u}] + s[\boldsymbol{v}]) = F([r\boldsymbol{u} + s\boldsymbol{v}]) = f(r\boldsymbol{u} + s\boldsymbol{v}) = rf(\boldsymbol{u}) + sf(\boldsymbol{v}) = rF([\boldsymbol{u}]) + sF([\boldsymbol{v}]).$$

また, 各 $f(\boldsymbol{u}) \in \operatorname{Im} f$ に対して, $F([\boldsymbol{u}]) = f(\boldsymbol{u})$ であるから F は全射である. 最後に単射性を示そう. $[\boldsymbol{u}] \in \operatorname{Ker} F$ とすれば $F([\boldsymbol{u}]) = \mathbf{0}_V$, すなわち $f(\boldsymbol{u}) = \mathbf{0}_V$ であり, したがって $\boldsymbol{u} \in \operatorname{Ker} f$ である. ゆえに命題 23.5.3(2) より $[\boldsymbol{u}] = \operatorname{Ker} f = [\mathbf{0}]$ であり, $\operatorname{Ker} F$ 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. □

練習 24.2.2. 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 $f : U \to V$ の準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 $\dim U/\operatorname{Ker} f = \dim U - \dim \operatorname{Ker} f = \dim \operatorname{Im} f$. したがって $U/\operatorname{Ker} f$ と $\operatorname{Im} f$ 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, 上のようにして準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, 準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 24.2.3 (不定積分の線形性)**.** 例 20.4.5(2) において, 微分作用素 $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ $(D(f) := f')$ の核 $\operatorname{Ker} D$ が定数関数全体に一致することを見た. つまり $C^\infty(\mathbb{R})/\operatorname{Ker} D$ 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,

$$[f(x)] = [f(x) + C] \in C^\infty(\mathbb{R})/\operatorname{Ker} D.$$

D は全射ゆえ $\operatorname{Im} D = C^\infty(\mathbb{R})$ であり, これに準同型定理を適用すれば線形同型 $\mathcal{D} : C^\infty(\mathbb{R})/\operatorname{Ker} D \to C^\infty(\mathbb{R})$ を得る. $\mathcal{D}$ の逆写像 $\mathcal{I} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})/\operatorname{Ker} D$ は, $f \in C^\infty(\mathbb{R})$ に対してその原始関数 F の同値類 $[F]$ を対応させる写像である:

$$\mathcal{I}(f) = [F] = F + \operatorname{Ker} D \in C^\infty(\mathbb{R})/\operatorname{Ker} D.$$

原始関数には定数関数の個数だけ任意性があるものの, $\mathcal{I}(f)$ は原始関数の選び方に依らずに決まる $C^\infty(\mathbb{R})/\operatorname{Ker} D$ の元である. $\mathcal{I}(f)$ 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:

$$\int f(x)dx.$$

こうして我々は, 不定積分が $C^\infty(\mathbb{R})/\operatorname{Ker} D$ 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 $I_a : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ は原始関数の一つを与える線形写像であったゆえ, $q : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})/\operatorname{Ker} D$ 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

$$\cdots \xrightarrow{f_{n-2}} U_{n-1} \xrightarrow{f_{n-1}} U_n \xrightarrow{f_n} U_{n+1} \xrightarrow{f_{n+1}} U_{n+2} \xrightarrow{f_{n+2}} \cdots \tag{24.3.1}$$

が各 $n \in \mathbb{Z}$ について $\operatorname{Im} f_n = \operatorname{Ker} f_{n+1}$ 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

$$\cdots \xrightarrow{\mathrm{id}} \{\mathbf{0}\} \xrightarrow{\mathrm{id}} \{\mathbf{0}\} \xrightarrow{f_0} W \xrightarrow{f_1} U \xrightarrow{f_2} V \xrightarrow{f_3} \{\mathbf{0}\} \xrightarrow{\mathrm{id}} \{\mathbf{0}\} \xrightarrow{\mathrm{id}} \cdots$$

を略して次のように書く[72]:

$$0 \xrightarrow{f_0} W \xrightarrow{f_1} U \xrightarrow{f_2} V \xrightarrow{f_3} 0 \tag{24.3.2}$$

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.

[72] 代数学では, 自明な空間 $\{\mathbf{0}\}$ を 0 と略すのが慣例となっている.

列 24.3.1 が完全系列であるとき, 各 $\boldsymbol{u} \in U_n$ について $f_n(\boldsymbol{u}) \in \operatorname{Im} f_n = \operatorname{Ker} f_{n+1}$ ゆえ $f_{n+1}(f_n(\boldsymbol{u})) = \mathbf{0}_{U_{n+2}}$ である. したがって合成写像 $f_{n+1} \circ f_n : U_n \to U_{n+2}$ 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) $\operatorname{Im} f_0 = \operatorname{Ker} f_1 \iff f_1 : W \to U$ は単射である (つまり W と $\operatorname{Im} f_1$ は同型).

(2) $\operatorname{Im} f_2 = \operatorname{Ker} f_3 \iff f_2 : U \to V$ は全射である.

Proof. (1): $\operatorname{Im} f_0 = \{\mathbf{0}_W\}$ ゆえ, $\operatorname{Im} f_0 = \operatorname{Ker} f_1 \iff \operatorname{Ker} f_1 = \{\mathbf{0}_W\} \iff f_1$ は単射.

(2): $\operatorname{Ker} f_3 = V$ ゆえ, $\operatorname{Im} f_2 = \operatorname{Ker} f_3 \iff \operatorname{Im} f_2 = V \iff f_2$ 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 $f : U \to V$ について, 次は短完全列である.

$$0 \longrightarrow \operatorname{Ker} f \xrightarrow{\mathrm{id}} U \xrightarrow{f} \operatorname{Im} f \longrightarrow 0$$

(2) U の部分空間 W および商写像 $q : U \to U/W$ について, 次は短完全列である.

$$0 \longrightarrow W \xrightarrow{\mathrm{id}} U \xrightarrow{q} U/W \longrightarrow 0$$

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: $\mathrm{id} : \operatorname{Ker} f \to U$ が単射であること, および $\mathrm{id}(\operatorname{Ker} f) = \operatorname{Ker} f$, $f : U \to \operatorname{Im} f$ 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 $\operatorname{Ker} q = W$ ゆえ (2) を得る. □

短完全列において準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 ($W \simeq \operatorname{Im} f_1 \subset U$). そこで W と $\operatorname{Im} f_1$ を同一視すれば, U/W と V は同型である.

Proof. $\operatorname{Im} f_1 = \operatorname{Ker} f_2$ および f_2 の全射性に注意すると, f_2 に関する準同型定理により次を得る:

$$V = \operatorname{Im} f_2 \simeq U/\operatorname{Ker} f_2 = U/\operatorname{Im} f_1 = U/W$$

. □

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 $f_2 : U \to V$ に関する次元公式により,

$$\begin{aligned} \dim U &= \dim \operatorname{Ker} f_2 + \dim \operatorname{Im} f_2 \\ &= \dim \operatorname{Im} f_1 + \dim \operatorname{Im} f_2 = \dim W + \dim V. \end{aligned}$$

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, $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n$ を W の基底, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ を V の基底とすれば, f_2 の全射性より $f(\boldsymbol{u}_i) = \boldsymbol{v}_i$ $(i = 1, \cdots, m)$ を満たす $\boldsymbol{u}_i \in U$ が取れる. このとき, 定理 23.2.2 により $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n, f_1(\boldsymbol{w}_1), \cdots, f_1(\boldsymbol{w}_n)$ 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.

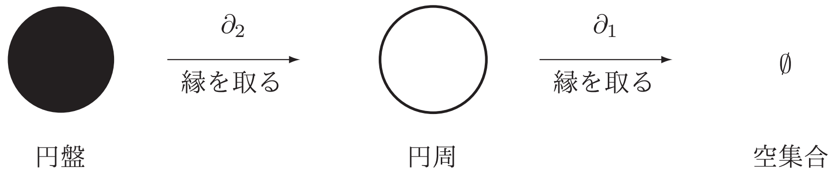
また, 例 24.3.3(2) および命題 24.3.4 によれば, 短完全列を与えることは商空間を与えることの言い換えに他ならない. 商空間を与えることは, 線形空間の各元を互いに交わらないグループに分類することを意味し, これは文字通りの空間の分解である. 短完全列から導かれる分解を荒っぽく述べれば, V の元の個数ぶんの W と合同な図形によって U は分解されている.

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 $\partial_n : C_n(X) \to C_{n-1}(X)$ 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 $\operatorname{Im} \partial_{n+1} \neq \operatorname{Ker} \partial_n$), $\operatorname{Im} \partial_{n+1} \subset \operatorname{Ker} \partial_n$ 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 $H_n(X) = \operatorname{Ker} \partial_n / \operatorname{Im} \partial_{n+1}$ 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．

$$\operatorname{rot} \circ \operatorname{grad} f = \mathbf{0}, \qquad \operatorname{div} \circ \operatorname{rot} F = \mathbf{0}. \tag{24.3.3}$$

ここで, 勾配ベクトル場 $\operatorname{grad} f$ およびベクトル場の回転 $\operatorname{rot} F$, ベクトル場の発散 $\operatorname{div} g$ は次で定められるのであった:

$$\begin{aligned}
\operatorname{grad} f &:= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right), \\
\operatorname{rot} F &:= \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right), \\
\operatorname{div} g &:= \frac{\partial g_1}{\partial x} + \frac{\partial g_2}{\partial y} + \frac{\partial g_3}{\partial z},
\end{aligned}$$

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ことになる．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 $\dfrac{\partial^2 f}{\partial y \partial x} = \dfrac{\partial^2 f}{\partial x \partial y}$) を用いよ．

25 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 $T : U \to \mathbb{R}^n$ を通して $\mathbb{R}^n$ の現象に書き換えられるはずである. 本節では, 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 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ について $[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n]$ 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:

$$[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n]\left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \dots & a_{1\ell} \\ a_{21} & a_{22} & \dots & a_{2\ell} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\ell} \end{array}\right] = \left[\sum_{i=1}^n a_{i1}\boldsymbol{u}_i \quad \sum_{i=1}^n a_{i2}\boldsymbol{u}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell}\boldsymbol{u}_i\right].$$

上の右辺には A の第 j 列成分を係数とする $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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[73] $[\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n]$ 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.

$$[\boldsymbol{v}_1, \boldsymbol{v}_2, \dots \boldsymbol{v}_n]\left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \dots & a_{1\ell} \\ a_{21} & a_{22} & \dots & a_{2\ell} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\ell} \end{array}\right] := \left[\sum_{i=1}^n a_{i1}\boldsymbol{v}_i \quad \sum_{i=1}^n a_{i2}\boldsymbol{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell}\boldsymbol{v}_i\right]. \qquad (25.1.1)$$

より抽象的に書けば,

$$[\boldsymbol{v}_h]_{h=1,\cdots,n}[a_{ij}]_{\substack{i=1,\cdots,n,\\ j=1,\cdots,\ell}} = \left[\sum_{i=1}^n a_{ij}\boldsymbol{v}_i\right]_{j=1,\cdots,\ell}.$$

式 (25.1.1) における $[\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n]$ 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[74]:

$$[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^n x_i\boldsymbol{u}_i.$$

例 25.1.1. (1) $[\boldsymbol{v}_1, \boldsymbol{v}_2]\begin{bmatrix} 3 & 2 & 1 \\ 1 & -1 & 4 \end{bmatrix} = \begin{bmatrix} 3\boldsymbol{v}_1 + \boldsymbol{v}_2 & 2\boldsymbol{v}_1 - \boldsymbol{v}_2 & \boldsymbol{v}_1 + 4\boldsymbol{v}_2 \end{bmatrix}$.

(2) $[\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n]E = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n]$.

ベクトルの組における和とスカラー倍も定めておこう:

- $[\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n] + [\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n] := [\boldsymbol{v}_1 + \boldsymbol{w}_1, \boldsymbol{v}_2 + \boldsymbol{w}_2, \cdots, \boldsymbol{v}_n + \boldsymbol{w}_n]$.
- $r[\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n] := [r\boldsymbol{v}_1, r\boldsymbol{v}_2, \dots, r\boldsymbol{v}_n]$.

[73] 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.

[74] $\sum_{i=1}^n \boldsymbol{u}_i x_i$ 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.

25.2 ベクトルの組と行列の演算の基本性質 (付録)

前項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 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, および (n, ℓ)-行列 A, $r \in \mathbb{R}$ に対して

$$[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n](rA) = (r[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n])A = r\big([\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]A\big).$$

Proof. $A = [a_{ij}]$ とおくと $rA = [ra_{ij}]$ であり,

$$\begin{aligned}
(\text{左辺}) &= \left[\textstyle\sum_{i=1}^n ra_{i1}\boldsymbol{v}_i \quad \sum_{i=1}^n ra_{i2}\boldsymbol{v}_i \quad \ldots, \quad \sum_{i=1}^n ra_{i\ell}\boldsymbol{v}_i\right] \\
&= \left[\textstyle\sum_{i=1}^n a_{i1}(r\boldsymbol{v}_i) \quad \sum_{i=1}^n a_{i2}(r\boldsymbol{v}_i) \quad \ldots, \quad \sum_{i=1}^n a_{i\ell}(r\boldsymbol{v}_i)\right] \quad (\text{この式は上の中央の式に相当する}) \\
&= \left[r\textstyle\sum_{i=1}^n a_{i1}\boldsymbol{v}_i \quad r\sum_{i=1}^n a_{i2}\boldsymbol{v}_i \quad \ldots, \quad r\sum_{i=1}^n a_{i\ell}\boldsymbol{v}_i\right] \\
&= r\left[\textstyle\sum_{i=1}^n a_{i1}\boldsymbol{v}_i \quad \sum_{i=1}^n a_{i2}\boldsymbol{v}_i \quad \ldots, \quad \sum_{i=1}^n a_{i\ell}\boldsymbol{v}_i\right] = (\text{右辺}).
\end{aligned}$$

□

命題 25.2.2. 線形空間 V のベクトルの組 $\boldsymbol{v}_1, \ldots \boldsymbol{v}_n$, および (n, ℓ)-行列 A, B に対して

$$[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]A + [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]B = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n](A + B).$$

Proof. $A = [a_{ij}]$, $B = [b_{ij}]$ とおくと,

$$\begin{aligned}
(\text{左辺}) &= \left[\textstyle\sum_{i=1}^n a_{i1}\boldsymbol{v}_i \quad \sum_{i=1}^n a_{i2}\boldsymbol{v}_i \quad \ldots \quad \sum_{i=1}^n a_{i\ell}\boldsymbol{v}_i\right] + \left[\textstyle\sum_{i=1}^n b_{i1}\boldsymbol{v}_i \quad \sum_{i=1}^n b_{i2}\boldsymbol{v}_i \quad \ldots \quad \sum_{i=1}^n b_{i\ell}\boldsymbol{v}_i\right] \\
&= \left[\textstyle\sum_{i=1}^n a_{i1}\boldsymbol{v}_i + \sum_{i=1}^n b_{i1}\boldsymbol{v}_i \quad \sum_{i=1}^n a_{i2}\boldsymbol{v}_i + \sum_{i=1}^n b_{i2}\boldsymbol{v}_i \quad \ldots \quad \sum_{i=1}^n a_{i\ell}\boldsymbol{v}_i + \sum_{i=1}^n b_{i\ell}\boldsymbol{v}_i\right] \\
&= \left[\textstyle\sum_{i=1}^n (a_{i1} + b_{i1})\boldsymbol{v}_i \quad \sum_{i=1}^n (a_{i2} + b_{i2})\boldsymbol{v}_i \quad \ldots \quad \sum_{i=1}^n (a_{i\ell} + b_{i\ell})\boldsymbol{v}_i\right] = (\text{右辺})
\end{aligned}$$

□

命題 25.2.3. 線形空間 V のベクトルの組 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, および (n, ℓ)-行列 A, (ℓ, r)-行列 B に対して

$$([\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]A)B = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n](AB).$$

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.

$$\begin{aligned}
(\text{左辺}) &= \left[\textstyle\sum_{i=1}^n a_{i1}\boldsymbol{v}_i \quad \ldots \quad \sum_{i=1}^n a_{i\ell}\boldsymbol{v}_i\right] B \\
&= \left[\boldsymbol{u}_1 \quad \ldots \quad \boldsymbol{u}_\ell\right] B \qquad \left(\boldsymbol{u}_j := \sum_{i=1}^n a_{ij}\boldsymbol{v}_i \ (j = 1, \cdots, \ell) \text{ とおいた}\right) \\
&= \left[\textstyle\sum_{j=1}^{\ell} b_{j1}\boldsymbol{u}_j \quad \ldots \quad \sum_{j=1}^{\ell} b_{ir}\boldsymbol{v}_i\right] = \left[\sum_{j=1}^{\ell} b_{jk}\boldsymbol{u}_j\right]_{k=1,\ldots,r} = \left[\sum_{j=1}^{\ell} b_{jk}\left(\sum_{i=1}^n a_{ij}\boldsymbol{v}_i\right)\right]_{k=1,\ldots,r} \\
&= \left[\sum_{j=1}^{\ell}\left(\sum_{i=1}^n a_{ij}b_{jk}\boldsymbol{v}_i\right)\right]_{k=1,\ldots,r} = \left[\sum_{i=1}^n\left(\sum_{j=1}^{\ell} a_{ij}b_{jk}\boldsymbol{v}_i\right)\right]_{k=1,\ldots,r} \qquad (\text{例 3.2.4 を用いた}) \\
&= \left[\sum_{i=1}^n\left(\sum_{j=1}^{\ell} a_{ij}b_{jk}\right)\boldsymbol{v}_i\right]_{k=1,\ldots,r} = \left[\sum_{i=1}^n z_{ik}\boldsymbol{v}_i\right]_{k=1,\ldots,r} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n][z_{ik}] = (\text{右辺}).
\end{aligned}$$

□

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 項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 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ が組 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell$ の線形結合で表され, また各 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell$ が組 $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ の線形結合で表されるならば, 各 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ は $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ の線形結合で表される.

Proof. 各 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ が $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell$ の線形結合で表されることから, (ℓ, r)-行列 B を用いて.$[\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r] = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell]B$ と書ける. また, 各 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell$ が $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ 個の線形結合で表されることから, (n, ℓ)-行列 A を用いて $[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell] = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n]A$ と書ける. ゆえに,

$$[\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r] = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell]B = ([\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n]A)B = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n](AB).$$

この式は各 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ が $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ 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, $\boldsymbol{u}_k$ $(k = 1, \cdots, r)$ を $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ 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. $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m \in U$ が線形独立であるとし, A, B を (m, n)-行列とする. このとき, $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]B$ ならば $A = B$ となることを示せ.

解答例: 移項して $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m](A - B) = [\boldsymbol{0}_V, \cdots, \boldsymbol{0}_V]$. $A - B = [x_{ij}]$ とおき, 各成分を比較すれば $j = 1, \cdots, n$ について $\sum_{i=1}^m x_{ij}\boldsymbol{v}_i = \boldsymbol{0}_V$ であり, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ 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, $f : U \to W$ を線形写像とする. $\boldsymbol{u}_j, \boldsymbol{v}_i \in U$ について次を示せ.

$$[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A \quad \Longrightarrow \quad [f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] = [f(\boldsymbol{v}_1), \cdots, f(\boldsymbol{v}_m)]A.$$

解答例: $A = [a_{ij}]$ とすれば, 仮定より,

$$[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n] = \begin{bmatrix} \sum_{i=1}^m a_{i1}\boldsymbol{v}_i & \ldots & \sum_{i=1}^m a_{in}\boldsymbol{v}_i \end{bmatrix} \qquad \text{であるから,}$$

$$\begin{aligned} [f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] &= \begin{bmatrix} f\left(\sum_{i=1}^m a_{i1}\boldsymbol{v}_i\right) & \ldots & f\left(\sum_{i=1}^m a_{in}\boldsymbol{v}_i\right) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^m a_{i1}f(\boldsymbol{v}_i) & \ldots & \sum_{i=1}^m a_{in}f(\boldsymbol{v}_i) \end{bmatrix} = [f(\boldsymbol{v}_1), \cdots, f(\boldsymbol{v}_m)]A. \end{aligned}$$

□

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. $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in \langle \boldsymbol{u}_1, \cdots, \boldsymbol{u}_m \rangle$ とする. このとき, $m < n$ ならば $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ は線形従属である.

Proof. 仮定より (m, n)-行列 A を用いて $[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n] = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]A$ と書ける. $n > m$ より連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{0}$ は自明でない解 ${}^t[c_1, \cdots, c_n] \neq \boldsymbol{0}$ を持つ. このとき,

$$\begin{aligned} c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} &= [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m]A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \\ &= [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m] \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = 0\boldsymbol{u}_1 + \cdots 0\boldsymbol{u}_n = \boldsymbol{0}. \end{aligned}$$

自明でない線形関係 $\sum_{j=1}^n c_j\boldsymbol{v}_j = \boldsymbol{0}$ があるゆえ $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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: $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ および $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ と $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ を U の基底とし, $F(\boldsymbol{u}_j) = \boldsymbol{e}_j$ なる線形同型 $F : U \to \mathbb{R}^m$ を取る.

$$\begin{array}{ccccccccc} & U & : & \boldsymbol{u}_1, & \boldsymbol{u}_2, & \cdots, & \boldsymbol{u}_m & & \boldsymbol{v} = \sum_{i=1}^m r_i \boldsymbol{u}_i \\ F & \downarrow & & \updownarrow & \updownarrow & & \updownarrow & & \updownarrow \\ & \mathbb{R}^m & : & \boldsymbol{e}_1, & \boldsymbol{e}_2, & \cdots, & \boldsymbol{e}_m & & \sum_{i=1}^m r_i \boldsymbol{e}_i = {}^t[r_1, \cdots, r_m] \end{array}$$

U において調べたいベクトルの組 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in U$ が与えられたとき, これに対応する $T(\boldsymbol{v}_1), \cdots, T(\boldsymbol{v}_n) \in \mathbb{R}^m$ を調べるというのが基本方針である. いま $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ は基底であるから各 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ は $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ の線形結合で書ける:

$$\boldsymbol{v}_1 = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m] \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \boldsymbol{v}_2 = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m] \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \ldots, \quad \boldsymbol{v}_n = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m] \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

これをまとめて書くと, (m, n)-行列 $A = [a_{ij}] = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ を用いて

$$[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n] = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m] \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \ldots & a_{1\ell} \\ a_{21} & a_{22} & \ldots & a_{2\ell} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{n\ell} \end{array}\right].$$

このとき, $[F(\boldsymbol{v}_1), \cdots, F(\boldsymbol{v}_n)] = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ である. 実際, 上式に F をほどこせば, 練習 25.3.2 により

$$[F(\boldsymbol{v}_1), \cdots, F(\boldsymbol{v}_n)] = [F(\boldsymbol{u}_1), \cdots, F(\boldsymbol{u}_m)]A = EA = A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n].$$

命題 20.3.6 より $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ の線形関係と $F(\boldsymbol{v}_1), \cdots, F(\boldsymbol{v}_n)$ (すなわち $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$) の線形関係は同等である. つまり, 各 $\boldsymbol{w} \in U$ および $s_1, \cdots, s_n \in \mathbb{R}$ について

$$\boldsymbol{w} = \sum_{j=1}^n s_j \boldsymbol{v}_j \iff F(\boldsymbol{w}) = \sum_{j=1}^n s_j \boldsymbol{a}_j. \tag{25.4.1}$$

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 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ の線形独立性 (したがって $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m$ を U の基底とする. ベクトルの組 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n \in U$ について $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n] = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_m]A$ とすれば次が成り立つ.

(1) $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$ とおけば, $\sum_{i=1}^n c_i \boldsymbol{v}_i = \boldsymbol{0}_U \iff \sum_{i=1}^n c_i \boldsymbol{a}_i = \boldsymbol{0}$.

(2) $n = m$ のとき次が成り立つ:

$\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ は線形独立 (したがって U の基底である) $\iff$ A は可逆.

補足: (2) の $\Leftarrow$ において $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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:

$[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n] = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]A$ の両辺に右から A^{-1} を掛ければ $[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]A^{-1}$. したがって, 基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の各ベクトルは $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ の線形結合で書けるゆえ, 命題 18.1.3 より $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ は U を生成する.

本項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. $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_5$ を U の基底とする. 次で与えるベクトルの組 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_5$ 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 $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_5 \rangle$ の基底を求めよ.

$$\begin{aligned} &\boldsymbol{v}_1 = \boldsymbol{u}_1 - \boldsymbol{u}_3 + 2\boldsymbol{u}_4 + \boldsymbol{u}_5, \quad \boldsymbol{v}_2 = \boldsymbol{u}_2 + \boldsymbol{u}_4 + \boldsymbol{u}_5, \quad \boldsymbol{v}_3 = -\boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{u}_3 - \boldsymbol{u}_4, \\ &\boldsymbol{v}_4 = \boldsymbol{u}_3, \quad \boldsymbol{v}_5 = -2\boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{u}_3 - 3\boldsymbol{u}_4 - \boldsymbol{u}_5. \end{aligned}$$

解答例: 上の線形結合を行列を用いた表示に書き直せば,

$$[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_5] = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_5] \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 & -3 \\ 1 & 1 & 0 & 0 & -1 \end{bmatrix}.$$

上式に現れる行列は例題 17.3.3 における $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_5]$ に一致している. $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_5$ 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 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_5$ は線形独立でない.

また, 例題 17.3.3 によれば $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_4$ は線形独立であり, これ以外のベクトルは $\boldsymbol{a}_3 = -\boldsymbol{a}_1 + \boldsymbol{a}_2$, $\boldsymbol{a}_5 = -2\boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{a}_4$ と書ける. したがって, 再び命題 25.4.1(1) より $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_4$ は線形独立であり, 更に式 (25.4.1) から $\boldsymbol{v}_3 = -\boldsymbol{v}_1 + \boldsymbol{v}_2$, $\boldsymbol{v}_5 = -2\boldsymbol{v}_1 + \boldsymbol{v}_2 - \boldsymbol{v}_4$ と書けることが分かる. 以上により, $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_4$ は $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_5 \rangle$ の基底である.

25.5 基底の変換行列

前項にて, U 上の各ベクトル $\boldsymbol{v}$ 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, 基底の取り変えにより $\boldsymbol{v}$ の数値化がど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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n$ が与えられているとする. このとき, 各 $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n \in U$ は基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の線形結合で書ける. これを式で書けば:

$$\begin{aligned} [\boldsymbol{u}'_1, \boldsymbol{u}'_2, \cdots, \boldsymbol{u}'_n] &= \left[\sum_{i=1}^n a_{i1}\boldsymbol{u}_i \quad \sum_{i=1}^n a_{i2}\boldsymbol{u}_i \quad \ldots, \quad \sum_{i=1}^n a_{in}\boldsymbol{u}_i \right] \\ &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n] \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \ldots, & a_{1n} \\ a_{21} & a_{22} & \ldots, & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{array} \right]. \end{aligned}$$

定義 25.5.1. 上式に現れる行列 $P = [a_{ij}]$ を基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による基底 $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n$ の**変換行列**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 $[\boldsymbol{u}'_1, \boldsymbol{u}'_2, \cdots, \boldsymbol{u}'_n] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n]P$ の両辺に右から P^{-1} をかけることで $[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n] = [\boldsymbol{u}'_1, \boldsymbol{u}'_2, \cdots, \boldsymbol{u}'_n]P^{-1}$ を得る. すなわち, 基底 $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n$ による基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の変換行列は P^{-1} である.

例 25.5.3. $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n \in \mathbb{R}^n$ を基底とする. 標準基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ による基底 $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n$ の変換行列は, $[\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n]$ である.

備考 25.5.4. 基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による基底 $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n$ の変換行列を P とする. $F_1 : U \to \mathbb{R}^n$ を $\boldsymbol{u}_i$ を $\boldsymbol{e}_i$ に写す線形同型とし, $F_2 : U \to \mathbb{R}^n$ を $\boldsymbol{u}'_i$ を $\boldsymbol{e}_i$ に写す線形同型とする.

$$
\begin{tikzcd}
 & U \arrow[ddl, "F_2"'] \arrow[ddr, bend left, "F_1"] & \\
 & \begin{array}{l} \quad\rotatebox{90}{$\in$} \\ \boldsymbol{v} = [\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n]\boldsymbol{x} \\ \quad = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]P\boldsymbol{x} \end{array} \arrow[dl, mapsto] \arrow[dr, mapsto] & \\
\mathbb{R}^n \ni \boldsymbol{x} \arrow[rr, mapsto] \arrow[rr, bend right, "T_P"'] & & P\boldsymbol{x} \in \mathbb{R}^n
\end{tikzcd}
$$

各ベクトル $\boldsymbol{v} = [\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n]\boldsymbol{x} \in U$ (ただし $\boldsymbol{x} \in \mathbb{R}^n$) に対応する $\mathbb{R}^n$ の元は, 同型 F_2 を用いれば $F_2(\boldsymbol{v}) = \boldsymbol{x}$ であり, 同型 F_1 を用いれば, $\boldsymbol{v} = [\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n]\boldsymbol{x} = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]P\boldsymbol{x}$ より $F_1(\boldsymbol{v}) = P\boldsymbol{x}$ である. このとき,

$$F_1 = T_P \circ F_2, \qquad T_P = F_1 \circ F_2^{-1}, \qquad F_2 = T_{P^{-1}} \circ F_1$$

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 $\boldsymbol{e}_j \in \mathbb{R}^n$ について $F_1 \circ F_2^{-1}(\boldsymbol{e}_j) = T_P(\boldsymbol{e}_j)$ であることを確かめればよい. $P = [\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n]$ と置けば, 変換行列の定義から $\boldsymbol{u}'_j = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]\boldsymbol{p}_j$ である. したがって,

$$
\begin{aligned}
F_1 \circ F_2^{-1}(\boldsymbol{e}_j) &= F_1(\boldsymbol{u}'_j) = F_1([\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]\boldsymbol{p}_j) = [F(\boldsymbol{u}_1), \cdots, F(\boldsymbol{u}_n)]\boldsymbol{p}_j \\
&= [\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n]\boldsymbol{p}_j = E\boldsymbol{p}_j = \boldsymbol{p}_j = P\boldsymbol{e}_j = T_P(\boldsymbol{e}_j).
\end{aligned}
$$

□

26 線形写像の表現行列

線形写像 $f : U \to V$ 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. 前節の対応と同様に, 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$, V の基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ をあらかじめ与えておく. 線形写像 $T : U \to V$ を (m,n)-行列 A による線形写像 $T_A : \mathbb{R}^n \to \mathbb{R}^m$ に翻訳する方法を考えよう. $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ が基底であることから, 各 $T(\boldsymbol{u}_1), \cdots, T(\boldsymbol{u}_n)$ は $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ の線形結合で書ける:

$$T(\boldsymbol{u}_1) = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m]\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T(\boldsymbol{u}_2) = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m]\begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \ldots, T(\boldsymbol{u}_n) = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m]\begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

つまり, (m,n)-行列 $A = [a_{ij}]$ を用いてまとめて書けば,

$$[T(\boldsymbol{u}_1), T(\boldsymbol{u}_2), \ldots, T(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m]\left[\begin{array}{c|c|c|c} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right]. \tag{26.1.1}$$

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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および V の基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ に関する $T : U \to V$ の**表現行列**または**行列表示**という. A の定義を次のように言い換えても良い:

$$A\text{ の第 } j \text{ 列} \quad = \quad \text{線形結合 } T(\boldsymbol{u}_j) = \sum_{i=1}^{m} a_{ij}\boldsymbol{v}_i \text{ に現れる係数を並べた列ベクトル}.$$

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 $T_B : \mathbb{R}^n \to \mathbb{R}^m$ $(T_B(\boldsymbol{x}) := B\boldsymbol{x})$ の表現行列は B に一致する.

例 26.1.3. (1) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ を U の基底とする. 定義域の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および終域の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ に関する恒等写像 $\mathrm{id}_U : U \to U$ の表現行列は単位行列 E である.

(2) $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ を U の基底とし, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ を V の基底とする. 各 $\boldsymbol{u}_i$ を $\boldsymbol{v}_i$ に対応させる線形同型 $f : U \to V$ について, U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および V の基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 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) $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ $(T_A(\boldsymbol{x}) := A\boldsymbol{x})$, ただし $A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$.

基底: 定義域と終域ともに $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ とせよ.

解答例:

$$\begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} = -2\boldsymbol{v}_1 + 0\boldsymbol{v}_2, \qquad \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 0\boldsymbol{v}_1 + 3\boldsymbol{v}_2.$$

ゆえに $[T_A(\boldsymbol{v}_1), T_A(\boldsymbol{v}_2)] = [\boldsymbol{v}_1, \boldsymbol{v}_2]\begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$ であり, T_A の表現行列は $\begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$.

(2) $D:\mathbb{R}[x]_3\to\mathbb{R}[x]_3$ $(D(p(x)):=\dfrac{d}{dx}p(x))$. 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 $\begin{bmatrix}0&1&0&0\\0&0&2&0\\0&0&0&3\\0&0&0&0\end{bmatrix}$.

(3) $I_{a,b}:\mathbb{R}[x]_3\to\mathbb{R}$ $(I_{a,b}(p(x)):=\displaystyle\int_a^b p(x)\,dx)$. 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:

$$I_{a,b}(x^3)=\left[\frac{1}{4}x^4\right]_a^b=\frac{1}{4}(b^4-a^4),\quad I_{a,b}(x^2)=\left[\frac{1}{3}x^3\right]_a^b=\frac{1}{3}(b^3-a^3),$$
$$I_{a,b}(x)=\left[\frac{1}{2}x^2\right]_a^b=\frac{1}{2}(b^2-a^2),\quad I_{a,b}(\mathbf{1})=\Big[x\Big]_a^b=(b-a)$$

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 $B=\{\,\boldsymbol{e}_n|n\in\mathbb{N}\,\}$ について, $V=\langle B\rangle$ と定める. 次の二つのシフト作用素

$$S_+:V\to V\qquad S_+(x_1,x_2,x_3,\cdots):=(0,x_1,x_2,x_3,x_4,\cdots)$$
$$S_-:V\to V\qquad S_-(x_1,x_2,x_3,\cdots):=(x_2,x_3,x_4,x_5,x_6\cdots)$$

の基底 B に関する表現行列は, 13.2 項のコラムで与えた A と B にそれぞれ等しい:

$$A=\begin{bmatrix}0&0&0&0&\cdots\\1&0&0&0&\cdots\\0&1&0&0&\cdots\\0&0&1&0&\cdots\\\vdots&\vdots&\vdots&\ddots&\end{bmatrix},\quad B=\begin{bmatrix}0&1&0&0&0&0&\cdots\\0&0&1&0&0&0&\cdots\\0&0&0&1&0&0&\cdots\\0&0&0&0&1&0&\cdots\\\vdots&\vdots&\vdots&\vdots&\vdots&\ddots&\end{bmatrix}.$$

$S_-\circ S_+=\mathrm{id}_V$ ゆえ BA は単位行列となる. 一方, $S_+\circ S_-(x_1,x_2,x_3,\cdots)=(0,x_2,x_3,\cdots)$ ゆえ $S_+\circ S_-\neq\mathrm{id}_V$. 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 19.6.1) 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 $S_+|_B : B \to B$ は単射であり, $S_-|_B : B \cup \{\mathbf{0}\} \to B \cup \{\mathbf{0}\}$ 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

$\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ を U の基底とし, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ を V の基底とする. また, これらの基底に関する線形写像 $T : U \to V$ の表現行列を A とする (つまり $[T(\boldsymbol{u}_1), \cdots, T(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A$). $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ とすれば, T はベクトル $[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]\boldsymbol{x} \in U$ をベクトル $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A\boldsymbol{x} \in V$ に写す写像である. 実際, 線形性 (iii)' より

$$T([\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]\boldsymbol{x}) = [T(\boldsymbol{u}_1), \cdots, T(\boldsymbol{u}_n)]\boldsymbol{x} = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A\boldsymbol{x}.$$

つまり, 次のような図式を得る:

$$\begin{array}{ccccccc} U & \ni & [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]\boldsymbol{x} & \xmapsto{T} & [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A\boldsymbol{x} & \in & V \\ F \downarrow & & \downarrow & & \downarrow & & \downarrow G \\ \mathbb{R}^n & \ni & \boldsymbol{x} & \xmapsto{T_A} & A\boldsymbol{x} & \in & \mathbb{R}^n \end{array}$$

ここで, $F : U \to \mathbb{R}^n$ は $F(\boldsymbol{u}_j) = \boldsymbol{e}_j$ を満たす線形同型写像, $G : V \to \mathbb{R}^m$ は $G(\boldsymbol{v}_i) = \boldsymbol{e}_i$ を満たす線形同型写像である. $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ が U の基底であることから, U の任意の元は $[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]\boldsymbol{x}$ 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:

$$\text{各 } \boldsymbol{u} \in U \text{ について,} \quad T_A \circ F(\boldsymbol{u}) = G \circ T(\boldsymbol{u}), \qquad \text{すなわち,} \quad T_A \circ F = G \circ T.$$

$T_A \circ F = G \circ T$ が成立するとき, これを次のような図式で表す.

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ {\scriptstyle F}\downarrow & & \downarrow{\scriptstyle G} \\ \mathbb{R}^n & \xrightarrow{T_A} & \mathbb{R}^m \end{array} \tag{26.1.2}$$

このような図式は**可換図式**と呼ばれる.

26.2 $\mathrm{Hom}(U, V)$ と $M_{m,n}(\mathbb{R})$

$\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ と $M_{m,n}(\mathbb{R})$ が同型であることを 21.3 項にて既に見たが, これらが $\mathrm{Hom}(U, V)$ (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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および V の基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ 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 $\mathcal{T} : \mathrm{Hom}(U, V) \to M_{m,n}(\mathbb{R})$ が線形同型を与えることを見よう.

各線形写像 $f : U \to V$ の表現行列を $\mathcal{T}(f)$ とする. つまり, $\mathcal{T}(f)$ は

$$[f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]\mathcal{T}(f)$$

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 $\mathcal{S}(A) : U \to V$ を,

$$[\mathcal{S}(A)(\boldsymbol{u}_1), \cdots, \mathcal{S}(A)(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A$$

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る:

- 各 $A \in M_{m,n}(\mathbb{R})$ について $\mathcal{T}(\mathcal{S}(A)) = A$ (つまり $\mathcal{T} \circ \mathcal{S} = \mathrm{id}_{M_{m,n}(\mathbb{R})}$).

  Proof. $\mathcal{S}(A) : U \to V$ は, $[\mathcal{S}(A)(\boldsymbol{u}_1), \cdots, \mathcal{S}(A)(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A$ を満たす線形写像であり, この写像 $\mathcal{S}(A)$ の表現行列は A である. すなわち, $\mathcal{T}(\mathcal{S}(A)) = A$. □

- 各 $f \in \mathrm{Hom}(U, V)$ について $\mathcal{S}(\mathcal{T}(f)) = f$ (つまり $\mathcal{S} \circ \mathcal{T} = \mathrm{id}_{\mathrm{Hom}(U,V)}$).

  Proof. 行列 $\mathcal{T}(f)$ とは, $[f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]\mathcal{T}(f)$ を満たす行列である. 一方, $\mathcal{S}(\mathcal{T}(f)) : U \to V$ は $[\mathcal{S}(\mathcal{T}(f))(\boldsymbol{u}_1), \cdots, \mathcal{S}(\mathcal{T}(f))(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]\mathcal{T}(f)$ を満たす線形写像であり,

  $$[f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] = [\mathcal{S}(\mathcal{T}(f))(\boldsymbol{u}_1), \cdots, \mathcal{S}(\mathcal{T}(f))(\boldsymbol{u}_n)]$$

  となる. $\mathcal{S}(\mathcal{T}(f))$ と f に U の各基底 $\boldsymbol{u}_j$ を代入した値はそれぞれ等しいゆえ, $\mathcal{S}(\mathcal{T}(f)) = f$. □

したがって, 命題 19.4.4 より $\mathcal{T}, \mathcal{S}$ は全単射かつ $\mathcal{T}^{-1} = \mathcal{S}$ 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 $\mathcal{T} : \mathrm{Hom}(U, V) \to M_{m,n}(\mathbb{R})$ および $f, g \in \mathrm{Hom}(U, V)$, $r, s \in \mathbb{R}$ について次を示せ.

$$\text{線形性 (iii):} \qquad \mathcal{T}(rf + sg) = r\mathcal{T}(f) + s\mathcal{T}(g).$$

解答例: $[f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]\mathcal{T}(f)$ および $[g(\boldsymbol{u}_1), \cdots, g(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]\mathcal{T}(g)$ であり, このとき

$$\begin{aligned}
[(rf + sg)(\boldsymbol{u}_1), \cdots, (rf + sg)(\boldsymbol{u}_n)] &= r[f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] + s[g(\boldsymbol{u}_1), \cdots, g(\boldsymbol{u}_n)] \\
&= r[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]\mathcal{T}(f) + s[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]\mathcal{T}(f) \\
&= [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m](r\mathcal{T}(f) + s\mathcal{T}(g)). \\
&\uparrow \text{ここで命題 25.2.1 および 25.2.2 を用いた.}
\end{aligned}$$

以上より, 線形写像 $rf + sg$ の表現行列は $r\mathcal{T}(f) + s\mathcal{T}(g)$ である. すなわち, $\mathcal{T}(rf + sg) = r\mathcal{T}(f) + s\mathcal{T}(g)$. □

系 26.2.2. (1) U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および V の基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ に関する $f, g : U \to V$ 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 $f : \mathbb{R}^n \to \mathbb{R}^m$ の表現行列が A ならば, $f = T_A$.

Proof. $\mathcal{T}$ の単射性から (1) は明らか. また, (2) の仮定において, A は標準基底に関する f および T_A の表現行列ゆえ, (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. $f : U \to V$, $g : V \to W$ を線形写像とし, U, V, W の基底をそれぞれ $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$, $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_\ell$ 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,

$$[f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]A, \qquad [g(\boldsymbol{v}_1), \cdots, g(\boldsymbol{v}_m)] = [\boldsymbol{w}_1, \cdots, \boldsymbol{w}_\ell]B$$

である. 練習 25.3.2 に注意すれば,

$$[g(f(\boldsymbol{u}_1)), \cdots, g(f(\boldsymbol{u}_n))] = [g(\boldsymbol{v}_1), \cdots, g(\boldsymbol{v}_m)]A = [\boldsymbol{w}_1, \cdots, \boldsymbol{w}_\ell]BA.$$

ゆえに $g \circ f$ の表現行列は BA である. □

系 26.2.4. U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および V の基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ に関する線形同型 $f : U \to V$ の表現行列を A とすれば, これらの基底に関する $f^{-1} : V \to U$ 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 $f^{-1} : V \to U$ の表現行列とすれば,

$$E = \text{“}\mathrm{id}_U : U \to U \text{ の表現行列”} = \text{“}f^{-1} \circ f : U \to U \text{ の表現行列”} = BA.$$

同様にして $AB = E$ 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:

<u>系 26.2.4 の別証明.</u> $[f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]A$ に逆写像 f^{-1} をほどこせば, 練習 25.3.2 より $[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n] = [f^{-1}(\boldsymbol{v}_1), \cdots, f^{-1}(\boldsymbol{v}_n)]A$ である. この両辺に右から A^{-1} をかけて $[f^{-1}(\boldsymbol{v}_1), \cdots, f^{-1}(\boldsymbol{v}_n)] = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]A^{-1}$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ を取るとき, 命題 26.2.3 は同型 $\mathcal{T} : \mathrm{End}(U) \to M_n(\mathbb{R})$ が積演算とも整合的であること, すなわち $\mathcal{T}(gf) = \mathcal{T}(g)\mathcal{T}(f)$ を意味している. 命題 21.5.7 より直ちに次が従う.

系 26.2.5. $\mathcal{T} : \mathrm{End}(U) \to M_n(\mathbb{R})$ を上で与えた同型とすれば, 任意の多項式 $\Psi(t)$ および $f \in \mathrm{End}(U)$ について $\mathcal{T}(\Psi(f)) = \Psi(\mathcal{T}(f))$. すなわち, U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ に関する f の表現行列を A とすれば, $\Psi(f) \in \mathrm{End}(U)$ 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 $T : U \to V$ 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.

$$\begin{array}{llcl} U \text{ の基底}: & \boldsymbol{u}_1, \cdots, \boldsymbol{u}_n & \text{および} & \boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n, \\ V \text{ の基底}: & \boldsymbol{v}_1, \cdots, \boldsymbol{v}_m & \text{および} & \boldsymbol{v}'_1, \cdots, \boldsymbol{v}'_m, \end{array}$$

とし, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ と $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ に関する T の表現行列を A, $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n$ と $\boldsymbol{v}'_1, \cdots, \boldsymbol{v}'_m$ に関する T の表現行列を B とする. また, 基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による基底 $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n$ の変換行列を P, 基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ による基底 $\boldsymbol{v}'_1, \cdots, \boldsymbol{v}'_m$ 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.

$$\begin{tikzcd} \mathbb{R}^n \arrow[d, "T_B"'] & U \arrow[l, "F_2"'] \arrow[r, "F_1"] \arrow[d, "T"'] & \mathbb{R}^n \arrow[d, "T_A"] \\ \mathbb{R}^m & V \arrow[l, "G_2"'] \arrow[r, "G_1"] & \mathbb{R}^m \end{tikzcd} \tag{26.3.1}$$

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:

$$\begin{array}{ll} F_1\colon F_1(\boldsymbol{u}_j) = \boldsymbol{e}_j \in \mathbb{R}^n \text{ を満たす写像}, & G_1\colon G_1(\boldsymbol{v}_i) = \boldsymbol{e}_i \in \mathbb{R}^m \text{ を満たす写像}, \\ F_2\colon F_2(\boldsymbol{u}'_j) = \boldsymbol{e}_j \in \mathbb{R}^n \text{ を満たす写像}, & G_2\colon G_2(\boldsymbol{v}'_i) = \boldsymbol{e}_i \in \mathbb{R}^m \text{ を満たす写像}. \end{array}$$

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.

<u>命題 26.3.1 の証明.</u>　$F_1 \circ F_2^{-1} = T_P$ 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:

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{T_P} & \mathbb{R}^n \\ {\scriptstyle T_B}\downarrow & & \downarrow{\scriptstyle T_A} \\ \mathbb{R}^m & \xrightarrow{T_Q} & \mathbb{R}^m \end{array}$$

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} 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

$$\begin{gathered} G_1 \circ G_1^{-1} \circ T_A \circ \underline{F_1 \circ F_2^{-1}} = \underline{G_1 \circ G_2^{-1}} \circ T_B \circ F_2 \circ F_2^{-1} \\ \mathrm{id}_{\mathbb{R}^m} \circ T_A \circ T_P = T_Q \circ T_B \circ \mathrm{id}_{\mathbb{R}^n} \\ T_A \circ T_P = T_Q \circ T_B. \end{gathered}$$

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.

<u>命題 26.3.1 の別証明.</u>　$[T(\boldsymbol{u}_1'), \cdots, T(\boldsymbol{u}_n')]$ を二通りの方法で計算する. $[\boldsymbol{u}_1', \cdots, \boldsymbol{u}_n'] = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]P$ に T をほどこすと練習 25.3.2 より

$$[T(\boldsymbol{u}_1'), \cdots, T(\boldsymbol{u}_n')] = [T(\boldsymbol{u}_1), \cdots, T(\boldsymbol{u}_n)]P = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]AP.$$

一方で,

$$[T(\boldsymbol{u}_1'), \cdots, T(\boldsymbol{u}_n')] = [\boldsymbol{v}_1', \cdots, \boldsymbol{v}_m']B = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]QB.$$

以上より $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]AP = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m]QB$ 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 $T : U \to U$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および終域の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ に関する T の表現行列のことを, 以下では単に基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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. $T : U \to U$ を線形変換とし, U の二組の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および $\boldsymbol{u}_1', \cdots, \boldsymbol{u}_n'$ が与えられているとする. このとき, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ に関する T の表現行列を A, $\boldsymbol{u}_1', \cdots, \boldsymbol{u}_n'$ に関する T の表現行列を B, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による $\boldsymbol{u}_1', \cdots, \boldsymbol{u}_n'$ 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. 次節以降では, 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, $\operatorname{tr} A = \operatorname{tr} B$.

Proof. 仮定より, $B = P^{-1}AP$ と書ける. 命題 20.4.3 を用いれば,

$$\operatorname{tr} B = \operatorname{tr}(P^{-1}AP) = \operatorname{tr}\big((P^{-1}A)P\big) = \operatorname{tr}\big(P(P^{-1}A)\big) = \operatorname{tr} A.$$

□

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) を調べる際の指標となることだろう (「指標」は代数学の専門用語でもある).

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節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.

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ {\scriptstyle F}\big\downarrow & & \big\downarrow{\scriptstyle G} \\ R_n & \xrightarrow{S} & R_m \end{array}$$

F および G が全単射であるとき次が成り立つ.

(1) $T = G^{-1} \circ S \circ F$,　$S = G \circ T \circ F^{-1}$.

(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, $T(X) = G^{-1}\Big(S\big(F(X)\big)\Big)$.

(4) $Y \subset V$ について, $T^{-1}(Y) = F^{-1}\Big(S^{-1}\big(G(Y)\big)\Big)$.

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, $T(X) = G^{-1} \circ S \circ F(X) = G^{-1}\Big(S\big(F(X)\big)\Big)$ 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 19.5.5 により

$$T^{-1}(Y) = (G^{-1} \circ S \circ F)^{-1}(Y) = F^{-1}\Big(S^{-1}\big((G^{-1})^{-1}(Y)\big)\Big) = F^{-1}\Big(S^{-1}\big(G(Y)\big)\Big).$$

□

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, $G^{-1}\big(S \circ F(X)\big) = G^{-1}\big(G(T(X))\big) = T(X)$.

(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 $x \in (S \circ F)^{-1}\big(G(Y)\big) = F^{-1}\big(S^{-1}(G(Y))\big)$. 一方, 各 $w \in F^{-1}\big(S^{-1}(G(Y))\big) = (S \circ F)^{-1}\big(G(Y)\big)$ 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 $T : U \to V$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ および V の基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ に関する $T : U \to V$ の表現行列を A とすれば, 次が成り立つ.

$$(1)\ \ \operatorname{Im} T \simeq \operatorname{Im} T_A, \qquad (2)\ \ \operatorname{Ker} T \simeq \operatorname{Ker} T_A$$

Proof. 可換図式 (26.1.2):

$$\begin{tikzcd} U \arrow[r, "T"] \arrow[d, "F"'] & V \arrow[d, "G"] \\ \mathbb{R}^n \arrow[r, "T_A"] & \mathbb{R}^m \end{tikzcd}$$

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

$$\operatorname{Im} T = T(U) = G^{-1}\Big(T_A\big(F(U)\big)\Big) = G^{-1}\Big(T_A(\mathbb{R}^n)\Big) = G^{-1}(\operatorname{Im} T_A).$$

(↑ ここで前命題 (3) を用いた)

すなわち, 線形同型写像 $G^{-1} : \operatorname{Im} T_A \to \operatorname{Im} T$ 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,

$$\operatorname{Ker} T = T^{-1}(\{\mathbf{0}_V\}) = F^{-1}\Big(T_A^{-1}\big(G(\{\mathbf{0}_V\})\big)\Big) = F^{-1}\Big(T_A^{-1}\big(\{\mathbf{0}\}\big)\Big) = F^{-1}\Big(\operatorname{Ker} T_A\Big).$$

(↑ ここで前命題 (4) を用いた)

線形同型写像 $F^{-1} : \operatorname{Ker} T_A \to \operatorname{Ker} T$ により, これらは同型である. □

$T_A : \mathbb{R}^n \to \mathbb{R}^m$ における次元定理は, 例 23.2.5 でみたように連立 1 次方程式の解法との関係からすぐに分かる. これと上の主張を合わせることで, 一般の線形写像 $f : U \to V$ における次元公式の別証明を得る.

27 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．しかし, 単純なベクトルと一言でいっても, 何を基準に単純なのかという疑問もあろう．そこで, ここでは次の二通りの方法を提案しよう．一つは, 標準的と考えられるベクトルへの分解であり, もう一つは与えられた線形写像と相性のよいベクトルへの分解である．

前者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かめることで, 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．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 27.0.3 (オイラーの公式)**.** $\quad e^{i\theta} = \cos\theta + i\sin\theta.$

このように, 関数の多項式による分解によって三角関数と指数関数の相互理解が深められる．一方で, 冪級数展開 $f(x) = \displaystyle\sum_{n=0}^{\infty} a_n x^n$ について次のような計算

$$\frac{d}{dx} f(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}, \qquad \int f(x)dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} + C \quad (C \text{ は積分定数})$$

も期待できるのであるが, この計算は x の冪がずれるゆえ少々計算が複雑になる．n 階微分になれば複雑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が考えられる．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)[75]:

$$f(x) = \sum_{n=-\infty}^{\infty} z_n e^{inx}, \quad \text{ここで, 各 } z_n := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt \text{ は複素数.}$$

このとき, $\dfrac{d}{dx} e^{ax} = a e^{ax}$ 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,

$$\frac{d}{dx} f(x) = \sum_{n=-\infty}^{\infty} in z_n e^{inx}, \quad \frac{d^n}{dx^n} f(x) = \sum_{n=-\infty}^{\infty} (in)^n z_n e^{inx}$$

となることが期待され, 指数関数による分解は多項式による分解よりも微分作用素と相性がよいことが示唆される．

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本節の主題となる．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, というもう一つ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 $\mathrm{End}(U)$ の元のみを考えることにする[76]．

[75] f が実数値関数の場合は, オイラーの公式により e^{inx} を三角関数に分解して表示するのが一般的である．実際, 定義から z_{-n} は z_n の共役な複素数であり, $a_n = z_n + z_{-n}$ および $b_n = (z_n - z_{-n})i$ は共に実数になる．このとき

$$z_n e^{inx} + z_{-n} e^{-inx} = a_n \cos(nx) + b_n \sin(nx)$$

であることは容易に確かめられ, 三角級数展開

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx) \big)$$

を得る．

[76] 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．

27.1 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 $f(\boldsymbol{u}) = a\boldsymbol{u}$ を満たすベクトル $\boldsymbol{u}$ 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 27.1.1. 線形変換 $f : U \to U$ において $f(\boldsymbol{v}) = \lambda\boldsymbol{u}$ を満たす $\lambda \in \mathbb{R}$ および零ベクトルでない $\boldsymbol{v} \in U$ が存在するとき, λ を f における**固有値** (**eigenvalue**) といい, $\boldsymbol{v}$ を固有値 λ に関する f の**固有ベクトル** (**eigenvector**) という. また, n 次正方行列 A による線形変換 $T_A : \mathbb{R}^n \to \mathbb{R}^n$ 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.

$\boldsymbol{u} \in U$ が f の固有値 λ_i に関する固有ベクトル $\boldsymbol{v}_i$ に分解されるとき, すなわち $\boldsymbol{u} = \sum_{i=1}^{n} a_i\boldsymbol{v}_i$ とかけるとき $f(\boldsymbol{u})$ の値は次のように簡単に計算できる:

$$
\begin{aligned}
f(\boldsymbol{u}) &= f\left(\sum_{i=1}^{n} a_i\boldsymbol{v}_i\right) = \sum_{i=1}^{n} a_i f(\boldsymbol{v}_i) = \sum_{i=1}^{n} a_i\lambda_i\boldsymbol{v}_i, \\
f^2(\boldsymbol{u}) &= f\left(\sum_{i=1}^{n} a_i\lambda_i\boldsymbol{v}_i\right) = \sum_{i=1}^{n} a_i\lambda_i f(\boldsymbol{v}_i) = \sum_{i=1}^{n} a_i{\lambda_i}^2\boldsymbol{v}_i, \\
f^3(\boldsymbol{u}) &= \sum_{i=1}^{n} a_i{\lambda_i}^3\boldsymbol{v}_i, \quad \cdots, \quad f^n(\boldsymbol{u}) = \sum_{i=1}^{n} a_i{\lambda_i}^n\boldsymbol{v}_i.
\end{aligned}
$$

例 27.1.2. (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 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ ついて, $B\boldsymbol{e}_i = \lambda_i\boldsymbol{e}_i$ である. ゆえに λ_i は B の固有値であり, $\boldsymbol{e}_i$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ について, 各 $\boldsymbol{u}_i$ が線形変換 $f \in \mathrm{End}(U)$ の固有値 λ_i に関する固有ベクトルであるとする. このとき, $f(\boldsymbol{u}_i) = \lambda_i\boldsymbol{u}_i$ であるから, 基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ に関する f の表現行列は対角行列 $\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ である.

例 27.1.3. ある自然数 n について $f^n = \boldsymbol{0}$ 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.

Proof. まず 0 が f の固有値となることを示そう. $\boldsymbol{x} \neq \boldsymbol{0}$ なるベクトルを一つとれば, $f^n(\boldsymbol{x}) = \boldsymbol{0}$ ゆえ, ある非負整数 m について $f^m(\boldsymbol{x}) \neq \boldsymbol{0}$ かつ $f^{m+1}(\boldsymbol{x}) = \boldsymbol{0}$ となる. このとき, $\boldsymbol{y} := f^m(\boldsymbol{x}) \neq \boldsymbol{0}$ とすれば, $f(\boldsymbol{y}) = \boldsymbol{0} = 0 \cdot \boldsymbol{y}$ ゆえ $\boldsymbol{y}$ は固有値 0 に関する固有ベクトルである.

一方, $\boldsymbol{x} \neq \boldsymbol{0}$ を固有値 λ に関する f の固有ベクトルとすれば, $\boldsymbol{0} = f^n(\boldsymbol{x}) = \lambda^n\boldsymbol{x}$ である. いま $\boldsymbol{x} \neq \boldsymbol{0}$ としていたから $\lambda^n = 0$ であり[77], ゆえに $\lambda = 0$. □

例 27.1.4. (1) シフト作用素 $S : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ $\left(S\Big((x_n)_{n\in\mathbb{N}}\Big)\right) := (x_{n+1})_{n\in\mathbb{N}}$ において, 初項 a 等比 λ なる等比数列 $\boldsymbol{x} = (a\lambda^{n-1})_{n\in\mathbb{N}}$ 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 $D : C^\infty(I) \to C^\infty(I)$ $(D(y) := y')$ 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.

---

[77] 仮に $\lambda^n \neq 0$ とすると, その逆数倍を $\lambda^n\boldsymbol{x} = \boldsymbol{0}$ の両辺にほどこせば $\boldsymbol{x} = \boldsymbol{0}$ となり矛盾してしまう.

例 27.1.5. (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 $\boldsymbol{x} \in \mathbb{R}^2$ (ただし $\boldsymbol{x} \neq \boldsymbol{0}$) は $A\boldsymbol{x}$ と平行でないゆえ $A\boldsymbol{x} \neq \lambda \boldsymbol{x}$ である. ゆえに体 $\mathbb{R}$ 上の線形空間としての線形写像 $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ 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 $T_A : \mathbb{C}^2 \to \mathbb{C}^2$ は固有ベクトルを持つ. 実際, $\boldsymbol{x} = \begin{bmatrix} i \\ 1 \end{bmatrix}$ 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 27.1.6. λ を正方行列 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.

解答例: $\Psi(t) = at^2 + bt + c$ の場合について示そう. 仮定より, 固有値 λ に関する A の固有ベクトル $\boldsymbol{v} \neq 0$ が存在する.

$$\begin{aligned} \Psi(A)\boldsymbol{v} &= (aA^2 + bA + cE)\boldsymbol{v} = aA^2\boldsymbol{v} + bA\boldsymbol{v} + cE\boldsymbol{v} \\ &= a\lambda^2\boldsymbol{v} + b\lambda\boldsymbol{v} + c\boldsymbol{v} = (a\lambda^2 + b\lambda + c)\boldsymbol{v} = \Psi(\lambda)\boldsymbol{v}. \end{aligned}$$

ゆえに $\boldsymbol{v}$ は $\Psi(A)$ の固有値 $\Psi(\lambda)$ に関する固有ベクトルである.

27.2 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,

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 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ に関する $T_A : \mathbb{R}^n \to \mathbb{R}^n$ 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 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n$ があるとすれば, 基底 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n$ に関する T_A の表現行列 B は例 27.1.2(2) により対角行列である. また, $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ による $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n$ の変換行列は $P = [\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n]$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ に関する f の表現行列を A とする. ある可逆行列 $P = [\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n]$ について $P^{-1}AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ となるならば, $[\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n] := [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]P$ で定められる[78]新たな

[78]すなわち, $\boldsymbol{q}_i := [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n]\boldsymbol{p}_i$ である.

基底[79]$\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n$ 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 $f(\boldsymbol{q}_i) = \lambda_i \boldsymbol{q}_i$ $(i = 1, \cdots, n)$ であり, ゆえに各 $\boldsymbol{q}_i$ 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:

$$A^n = (PBP^{-1})^n = PB\underline{P^{-1}\cdot P}B\underline{P^{-1}\cdot P}B\underline{P^{-1}\cdot}\cdots\cdot\underline{P}BP^{-1} = PB^nP^{-1}. \tag{27.2.1}$$

定義 27.2.1. 正方行列 A に対して, $P^{-1}AP$ が対角行列となるような可逆行列 P が存在するとき, A は**対角化可能である**といい, $P^{-1}AP$ を A の**対角化 (diagonalization)** という.

行列の対角化可能性については次節で詳しく解説する. 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.

27.3 特性多項式

線形空間 U における恒等写像 $\mathrm{id}_U \in \mathrm{End}(U)$ を I と書くのであった. 線形変換 $f : U \to U$ および $\lambda \in \mathbb{R}$ について, $\boldsymbol{x} \in U$ (ただし $\boldsymbol{x} \neq \boldsymbol{0}$) 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:

$$\begin{aligned} f(\boldsymbol{x}) = \lambda\boldsymbol{x} &\iff \lambda\boldsymbol{x} - f(\boldsymbol{x}) = \boldsymbol{0} \iff \lambda I(\boldsymbol{x}) - f(\boldsymbol{x}) = \boldsymbol{0} \\ &\iff (\lambda I - f)(\boldsymbol{x}) = \boldsymbol{0} \iff \boldsymbol{x} \in \mathrm{Ker}(\lambda I - f). \end{aligned}$$

したがって, 固有値 λ に関する固有ベクトルが存在することは, $\mathrm{Ker}(\lambda I - f) \neq \{\boldsymbol{0}\}$ であることと同値である:

命題 27.3.1. 線形空間 U および $f \in \mathrm{End}(U)$, $\lambda \in \mathbb{R}$ について次は同値である.

(1) λ は f の固有値である.

(2) $\mathrm{Ker}(\lambda I - f)$ は自明でない, すなわち $\mathrm{Ker}(\lambda I - f) \neq \{\boldsymbol{0}\}$.

(3) $\lambda I - f$ は単射でない.

とくに $U = \mathbb{R}^n$ および $f = T_A$ について, $(\lambda I - T_A)(\boldsymbol{x}) = (\lambda E - A)\boldsymbol{x}$ であり, $\mathrm{Ker}(\lambda I - T_A)$ が自明でないことは, 方程式 $(\lambda E - A)\boldsymbol{x}$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 27.3.2. $n \in \mathbb{N}$ 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: $\quad \lambda$ 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 $f \in \mathrm{End}(U)$ 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.

[79] 命題 25.4.1(2) により $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n$ は U の基底である.

定義 27.3.3. 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, 線形変換 $f \in \mathrm{End}(U)$ 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表す方程式 $\Phi_A(t) = 0$ 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 27.3.4. $f : U \to U$ を線形変換とし, 基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による f の表現行列を A, 基底 $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n$ による f の表現行列を B とすれば $\Phi_A = \Phi_B$.

Proof. 基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による基底 $\boldsymbol{u}'_1, \cdots, \boldsymbol{u}'_n$ 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

$$
\begin{aligned}
\Phi_B(t) &= |tE - B| = |tP^{-1}EP - P^{-1}AP| = |P^{-1}(tE)P - P^{-1}AP| = |P^{-1}((tE)P - AP)| \\
&= |P^{-1}(tE - A)P| = |P^{-1}| \cdot |tE - A| \cdot |P| = |tE - A| = \Phi_A(t).
\end{aligned}
$$

□

いまの証明から次の事実も導かれる:

系 27.3.5. 相似な正方行列の特性多項式は等しい.

例 27.3.6. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ について

$$
\begin{aligned}
\Phi_A(t) &= \begin{vmatrix} t-a & -b \\ -c & t-d \end{vmatrix} = (t-a)(t-d) - (-b)(-c) \\
&= t^2 - (a+d)t + (ad - bc) = t^2 - (\operatorname{tr} A)t + \det A.
\end{aligned}
$$

例 27.3.7. 上三角行列 $A = \begin{bmatrix} a_1 & & & \\ & a_2 & & * \\ & & \ddots & \\ O & & & a_n \end{bmatrix}$ について, 例 11.1.2 より

$$
\Phi_A(t) = \begin{vmatrix} t-a_1 & & & \\ & t-a_2 & & * \\ & & \ddots & \\ O & & & t-a_n \end{vmatrix} = (t-a_1)(t-a_2)\cdots(t-a_n).
$$

すなわち, A の対角成分は A の固有値である.

練習 27.3.8. n 次正方行列 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 $|B| = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)}$ の各項のうち, t の n 次多項式が現れるのは $\sigma = \mathrm{id}$ に関する項 $\mathrm{sgn}(\mathrm{id}) b_{11} b_{22} \cdots b_{nn} = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$ 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 27.3.9. A を m 次正方行列, B を n 次正方行列とする. $m+n$ 次正方行列 $X = \begin{bmatrix} A & * \\ O & B \end{bmatrix}$ について $\Phi_X(t) = \Phi_A(t)\Phi_B(t)$ を示せ.

練習 27.3.10. 正方行列 A について $\Phi_{{}^tA} = \Phi_A$ を示せ.

27.4 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．前項で論じたように, 線形変換 $f : U \to U$ および $\lambda \in \mathbb{R}$ について, $\mathrm{Ker}(\lambda I - f) \neq \{\mathbf{0}\}$ であるとき, $\mathrm{Ker}(\lambda I - f)$ 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 27.4.1. $\lambda \in \mathbb{R}$ を $f \in \mathrm{End}(U)$ の固有値とする．このとき, U の部分空間 $\mathrm{Ker}(\lambda I - f)$ を f の固有値 λ における**固有空間** (**eigenspace**) と呼び, 本論ではこれを $W(\lambda, f)$ で表す．また, n 次正方行列 A について, $T_A : \mathbb{R}^n \to \mathbb{R}^n$ 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 $(\lambda E - A)\boldsymbol{x} = \mathbf{0}$ の解空間 $W_{\lambda E - A}$ である．

例題 27.4.2. 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:

$$\Phi_A(t) = |tE - A| = \left| \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} - \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix} \right| = \begin{vmatrix} t-8 & 10 \\ -5 & t+7 \end{vmatrix}$$
$$= (t-8)(t+7) - (-5) \cdot 10 = t^2 - t - 56 + 50 = t^2 - t - 6 = (t-3)(t+2).$$

(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 27.3.2 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 $(\lambda E - A)\boldsymbol{x} = \mathbf{0}$ 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

$$\lambda E - A = \begin{bmatrix} \lambda-8 & 10 \\ -5 & \lambda+7 \end{bmatrix} \longrightarrow \begin{bmatrix} -5 & \lambda+7 \\ \lambda-8 & 10 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -\frac{\lambda+7}{5} \\ \lambda-8 & 10 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -\frac{\lambda+7}{5} \\ 0 & \frac{(\lambda+7)(\lambda-8)}{5} + 10 \end{bmatrix}$$

ここで, 最後の $(2,2)$-成分は

$$\frac{(\lambda+7)(\lambda-8)}{5} + 10 = \frac{1}{5}\Big((\lambda+7)(\lambda-8) + 50\Big) = \frac{1}{5}\Phi_A(\lambda) = 0.$$

したがって $\lambda = -2, 3$ における $(\lambda E - A)\boldsymbol{x} = \mathbf{0}$ の解はそれぞれ次のようになる:

$$W(-2, A) = \left\{ r \begin{bmatrix} 1 \\ 1 \end{bmatrix} \middle| r \in \mathbb{R} \right\}, \qquad W(3, A) = \left\{ r \begin{bmatrix} 2 \\ 1 \end{bmatrix} \middle| r \in \mathbb{R} \right\}.$$

上の例題における各固有空間の基底による組 $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ 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 28.1.1 にて示す．$\boldsymbol{v}_1, \boldsymbol{v}_2$ 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 27.1.2(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 $\boldsymbol{e}_1, \boldsymbol{e}_2$ による基底 $\boldsymbol{v}_1, \boldsymbol{v}_2$ に関する変換行列は $P = [\boldsymbol{v}_1, \boldsymbol{v}_2] = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ である．ゆえに系 26.3.2 より $B = P^{-1}AP$ を得る．

例題 27.4.3. $A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$ について, A^n を求めよ.

解答例: $B = P^{-1}AP$ とすれば, 式 (27.2.1) 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 27.4.4. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 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:

$$\lambda_i E - A = \begin{bmatrix} \lambda_i - a & -b \\ -c & \lambda_i - d \end{bmatrix} \longrightarrow \begin{bmatrix} -c & \lambda_i - d \\ \lambda_i - a & -b \end{bmatrix}$$
$$\longrightarrow \begin{bmatrix} 1 & -(\lambda_i - d)/c \\ \lambda_i - a & -b \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -(\lambda_i - d)/c \\ 1 & -b + (\lambda_i - a)(\lambda_i - d)/c \end{bmatrix}$$

$\Phi_A(t) = (t-a)(t-d) - bc$ 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, $W(\lambda_i, A) = \langle \boldsymbol{e}_i \rangle$ $(i = 1, 2)$ となる.
- $b \neq 0$ のとき, 固有値 a に関する固有空間は,

$$aE - A = \begin{bmatrix} 0 & -b \\ 0 & a - d \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 \\ 0 & a - d \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

より $W(a, A) = \langle \boldsymbol{e}_1 \rangle$ 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,

$$dE - A = \begin{bmatrix} d - a & -b \\ 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -b/(d-a) \\ 0 & 0 \end{bmatrix}$$

より $W(d, A) = \left\langle \begin{bmatrix} b/(d-a) \\ 1 \end{bmatrix} \right\rangle$.

27.5　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 27.3.2 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 27.5.1. U を n 次元線形空間とし $f : U \to U$ を線形変換とする. 更に, U の基底 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ による f の表現行列を A とすれば次が成り立つ.

(1) $\boldsymbol{x} \in \mathbb{R}^n$ が T_A の固有値 λ に関する T_A の固有ベクトルであることと, $\boldsymbol{y} = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\boldsymbol{x} \in U$ が固有値 λ に関する f の固有ベクトルであることは同値である.

(2) λ を f の固有値とする (したがって λ は A の固有値でもある). 列ベクトルの組 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k \in \mathbb{R}^n$ について, $\boldsymbol{y}_i = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\boldsymbol{x}_i$ $(i = 1, \cdots, k)$ とすれば, $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_k$ が $W(\lambda, f)$ の基底であることと $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ が $W(\lambda, A)$ の基底であることは同値である.

(3) $\dim W(\lambda, f) = \dim W(\lambda, A)$.

Proof. (1): 各 $\boldsymbol{v}_i \in U$ を $\boldsymbol{e}_i \in \mathbb{R}^n$ に対応させる線形同型を $F : U \to \mathbb{R}^n$ とする. このとき $g = \lambda I - f$ の表現行列は系 26.2.5 より $\lambda E - A$ である. また, 命題 26.4.2(2) の証明によれば $\mathrm{Ker}\, g = W(\lambda, f)$ と $\mathrm{Ker}(\lambda E - A) = W(\lambda, A)$ は, F によって 1 対 1 に対応する. 25.4 項で述べたことによれば, F は $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\boldsymbol{x}$ を $\boldsymbol{x}$ に対応させる線形同型写像であり, したがって, $\boldsymbol{x} \in W(\lambda, A)$ であることと $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\boldsymbol{x} \in W(\lambda, f)$ であることは同値である.

(2) および (3): 上で述べた対応 $F|_{W(\lambda,f)} : W(\lambda, f) \to W(\lambda, A)$ $([\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\boldsymbol{x} \mapsto \boldsymbol{x})$ が線形同型であることから明らか. □

補足. 上の (1) の証明が抽象的過ぎると感じる場合は, 次のような同値変形を行って確認してもよい:

<u>命題 27.5.1(1) の別証明.</u>　$\boldsymbol{x} = {}^t[x_1, \cdots, x_n]$ とおく. $\boldsymbol{y} = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\boldsymbol{x} = \sum_{i=1}^n x_i\boldsymbol{v}_i$ であること, および表現行列の定義より $[f(\boldsymbol{v}_1), \cdots, f(\boldsymbol{v}_n)] = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]A$ に注意すると,

$$\boldsymbol{y}\ \text{は}\ \lambda\ \text{に関する}\ f\ \text{の固有ベクトルである} \iff f\left(\sum_{i=1}^n x_i\boldsymbol{v}_i\right) = \lambda\boldsymbol{y}$$
$$\iff \sum_{i=1}^n x_i f(\boldsymbol{v}_i) = \lambda[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]\boldsymbol{x} \iff [f(\boldsymbol{v}_1), \cdots, f(\boldsymbol{v}_n)]\boldsymbol{x} = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n](\lambda\boldsymbol{x})$$
$$\iff [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n]A\boldsymbol{x} = [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n](\lambda\boldsymbol{x})$$
$$\iff A\boldsymbol{x} = \lambda\boldsymbol{x} \iff \boldsymbol{x}\ \text{は}\ \lambda\ \text{に関する}\ A\ \text{の固有ベクトルである.}$$

(↑ここで $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ の線形独立性および練習 25.3.1 を用いた.)

□

例 27.5.2. 線形写像 $D : \mathbb{R}[x]_3 \to \mathbb{R}[x]_3$ を $D(p(x)) := \dfrac{d}{dx}p(x)$ 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 $(0E - A)\boldsymbol{x} = \mathbf{0}$ を解いて

$$W(0, A) = \left\{ r\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \middle|\ r \in \mathbb{R} \right\}$$

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 27.5.1 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, $W(0, D) = \langle \mathbf{1} \rangle$.

28　固有空間分解と行列の対角化

前節において, 線形変換 $f: U \to U$ 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 $\boldsymbol{u}$ を固有ベクトルに分解すると $f(\boldsymbol{u})$ 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, 本節では, 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おり, 方程式が実数解を持たない場合にスカラー係数の範囲を複素数に広げたほうが都合がよい場合がある. しかしながら, スカラー係数を実数に限った話なのか複素数まで広げているのか議論が曖昧になっては困るだろう. そこで本節以降においては, 体 $\mathbb{K}$ を $\mathbb{R}$ または $\mathbb{C}$ のいずれかとし, 線形空間とは体 $\mathbb{K}$ 上の線形空間のこと, すなわちスカラー係数として $\mathbb{K}$ 上の元を取るとする. また, 行列とは $\mathbb{K}$ の元を各成分に持つ行列のこととする. もちろん, 線形写像あるいは正方行列の固有値は $\mathbb{K}$ の元であり, 多項式とは $\mathbb{K}$ の元を係数に持つ多項式のことである. また, 多項式 $\Phi(x)$ が因数分解されるとは, $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ を用いて $\Phi(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ と書けることと定める. 以上の設定のもとで, $\mathbb{K} = \mathbb{R}$ の場合として前節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.

28.1　固有ベクトルの線形独立性

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, 本節では 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 28.1.1. $f: U \to U$ を線形変換とし, その相異なる固有値を $\lambda_1, \cdots, \lambda_r$ とする. また $\dim W(\lambda, f) = n_\lambda$ $(\lambda = \lambda_1, \cdots, \lambda_r)$ とする. このとき, 各固有値 λ について固有空間 $W(\lambda, f)$ の基底 $\boldsymbol{u}_{\lambda,1}, \cdots, \boldsymbol{u}_{\lambda,n_\lambda}$ を一組えらべば, これらを集めたベクトルの組 $\mathcal{B} = \{\, \boldsymbol{u}_{\lambda,i} \mid \lambda = \lambda_1, \cdots, \lambda_r,\ i = 1, \cdots, n_\lambda \,\}$ は線形独立である.

Proof. $\mathcal{B}$ が線形従属であると仮定し, 矛盾を示そう. $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_\ell \in B$ を $\langle \mathcal{B} \rangle$ の基底とする. $\mathcal{B}$ 自身は線形従属ゆえ, $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_\ell$ のいずれでもない $\boldsymbol{b} \in B$ が取れる. ここで, 各 $\boldsymbol{b}_i$ $(i = 1, \cdots, \ell)$ は固有値 γ_i に関する固有ベクトルであるとし (γ_i には重複がある可能性もある), また $\boldsymbol{b}$ を固有値 γ に関する固有ベクトルとする. $\boldsymbol{b} = \sum_{i=1}^{\ell} c_i \boldsymbol{b}_i$ とし, この両辺に f をほどこすと,

$$\begin{aligned}
f(\boldsymbol{b}) &= c_1 f(\boldsymbol{b}_1) + \cdots + c_\ell f(\boldsymbol{b}_\ell) \\
\gamma \boldsymbol{b} &= c_1 \gamma_1 \boldsymbol{b}_1 + \cdots + c_\ell \gamma_\ell \boldsymbol{b}_\ell \\
\gamma\Big(c_1 \boldsymbol{b}_1 + \cdots + c_\ell \boldsymbol{b}_\ell\Big) &= c_1 \gamma_1 \boldsymbol{b}_1 + \cdots + c_\ell \gamma_\ell \boldsymbol{b}_\ell \\
\sum_{i=1}^{\ell} c_i (\gamma - \gamma_i) \boldsymbol{b}_i &= \boldsymbol{0}.
\end{aligned}$$

$\boldsymbol{b}_1, \cdots, \boldsymbol{b}_\ell$ の線形独立性より, $c_i(\gamma - \gamma_i) = 0$ $(i = 1, \cdots, \ell)$ であり, とくに γ_i と γ が異なる固有値であれば $\gamma - \gamma_i \neq 0$ ゆえ $c_i = 0$ を得る. したがって, 線形結合 $\boldsymbol{b} = \sum_{i=1}^{\ell} c_i \boldsymbol{b}_i$ に現れる係数 c_i のうち, 0 でないものは $\gamma_i = \gamma$ の場合, すなわち $\boldsymbol{b}_i \in W(\gamma, f)$ の場合に限る. このような $\boldsymbol{b}_i$ および $\boldsymbol{b}$ は, $\boldsymbol{u}_{\gamma,1}, \cdots, \boldsymbol{u}_{\gamma,n_\gamma}$ のいずれかのベクトルである. 命題 17.2.5 より $\boldsymbol{u}_{\gamma,1}, \cdots, \boldsymbol{u}_{\gamma,n_\gamma}$ は線形従属となり, これは $\boldsymbol{u}_{\gamma,1}, \cdots, \boldsymbol{u}_{\gamma,n_\gamma}$ を $W(\gamma, f)$ の基底として取ったことに矛盾する. □

系 28.1.2. A を n 次正方行列とし, その相異なる固有値を $\lambda_1, \cdots, \lambda_r$ とすれば, $\sum_{k=1}^{r} \dim W(\lambda_k, T_A) \leq n$.

備考 28.1.3. 実は, f の特性多項式が相異なる $\lambda_1, \cdots, \lambda_r \in \mathbb{K}$ について $\Phi_f(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_r)^{n_r}$ と因数分解されるとき, 後に学ぶ定理 32.1.1 によれば $\dim W(\lambda_k, f) \le n_k$ が成立する. 定理 32.1.1 に現れる部分空間 $\widetilde{W}(\lambda_k, f)$ は $W(\lambda_k, f)$ を部分空間として含み, したがって, $\dim W(\lambda_k, f) \le \dim \widetilde{W}(\lambda_k, f) = n_k$ を得る.

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, 命題 28.1.1 を背理法を用いずに示してみよう.

補題 28.1.4. f の相異なる固有値を $\lambda_1, \cdots, \lambda_r$ とする. $\boldsymbol{v}_k \in W(\lambda_k, f)$ $(k = 1, 2, \cdots, r)$ かつ $\boldsymbol{v}_1 + \boldsymbol{v}_2 + \cdots + \boldsymbol{v}_r = \boldsymbol{0}$ ならば, $\boldsymbol{v}_1 = \boldsymbol{v}_2 = \cdots = \boldsymbol{v}_r = \boldsymbol{0}$.

Proof. $\sum_{k=1}^{r} \boldsymbol{v}_k = \boldsymbol{0}$ の両辺に f をほどこすことで $\sum_{k=1}^{r} \lambda_k \boldsymbol{v}_k = \boldsymbol{0}$ を得る. この式に更に f をほどこせば, $\sum_{k=1}^{r} {\lambda_k}^2 \boldsymbol{v}_k = \boldsymbol{0}$ 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:

$$\begin{aligned}
\boldsymbol{v}_1 + \boldsymbol{v}_2 + \quad \cdots + \boldsymbol{v}_r &= \boldsymbol{0}, \\
\lambda_1 \boldsymbol{v}_1 + \lambda_2 \boldsymbol{v}_2 + \cdots + \lambda_r \boldsymbol{v}_r &= \boldsymbol{0}, \\
{\lambda_1}^2 \boldsymbol{v}_1 + {\lambda_2}^2 \boldsymbol{v}_2 + \cdots + {\lambda_r}^2 \boldsymbol{v}_r &= \boldsymbol{0}, \\
&\vdots \\
{\lambda_1}^{r-1} \boldsymbol{v}_1 + {\lambda_2}^{r-1} \boldsymbol{v}_2 + \cdots + {\lambda_r}^{r-1} \boldsymbol{v}_r &= \boldsymbol{0}.
\end{aligned}$$

$$\text{つまり,} \quad P = \begin{bmatrix} 1 & \lambda_1 & {\lambda_1}^2 & \cdots & {\lambda_1}^{r-1} \\ 1 & \lambda_2 & {\lambda_2}^2 & \cdots & {\lambda_2}^{r-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_r & {\lambda_r}^2 & \cdots & {\lambda_r}^{r-1} \end{bmatrix} \quad \text{とおけば} \quad [\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r]P = [\boldsymbol{0}, \cdots \boldsymbol{0}].$$

各 λ_k は相異なる固有値としていたゆえ, $1 \le p < q \le r$ ならば $\lambda_q - \lambda_p \ne 0$ 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 $|P| \ne 0$. つまり P は可逆であり, $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r]P = [\boldsymbol{0}, \cdots, \boldsymbol{0}]$ の両辺に右から P^{-1} をかけることで $[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_r] = [\boldsymbol{0}, \cdots, \boldsymbol{0}]$ を得る. □

上の証明では強力な公式 (ヴァンデルモンドの公式) に議論を帰着させている. この公式を用いない別証明を本節の後半で与えよう.

<u>命題 28.1.1 の別証明</u>

$$\sum_{\lambda=\lambda_1,\cdots,\lambda_r} \left(\sum_{i=1,\cdots,n_\lambda} c_{\lambda,i} \boldsymbol{u}_{\lambda,i} \right) = \boldsymbol{0}$$

と仮定し, $c_{\lambda,i} = 0$ を示そう. $\boldsymbol{v}_\lambda = \sum_{i=1,\cdots,n_\lambda} c_{\lambda,i} \boldsymbol{u}_{\lambda,i}$ とおけば $\sum_{\lambda=\lambda_1,\cdots,\lambda_r} \boldsymbol{v}_\lambda = \boldsymbol{0}$ である. 補題 28.1.4 より各 $\lambda = \lambda_1, \cdots, \lambda_r$ について $\boldsymbol{v}_\lambda = \boldsymbol{0}$ であり, ゆえに $\sum_{i=1,\cdots,n_\lambda} c_{\lambda,i} \boldsymbol{u}_{\lambda,i} = \boldsymbol{0}$. $\boldsymbol{u}_{\lambda,1}, \cdots, \boldsymbol{u}_{\lambda,n_\lambda}$ の線形独立性から $c_{\lambda,i} = 0$ を得る. □

補足. 実は, 補題 28.1.4 にある性質から命題 28.1.1 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ものである. 詳しくは 31.5 節を見よ.

28.2 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. 証明の一部は 27.2 項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 28.2.1. 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 $\lambda_1, \cdots, \lambda_r \in \mathbb{K}$ とすれば, $\sum_{k=1}^{r} \dim W(\lambda_k, A) = n$.
(注意: この条件は系 28.1.2 より $\sum_{k=1}^{r} \dim W(\lambda_k, A) \geq n$ と同値である.)

(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)⇒(2): 各 $\lambda = \lambda_1, \cdots, \lambda_r$ 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 $\mathcal{B} = \{ \boldsymbol{u}_{\lambda,i} \mid \lambda = \lambda_1, \cdots, \lambda_r,\ i = 1, \cdots, n_\lambda \}$ とすれば, 命題 28.1.1 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)⇒(3): 列ベクトルの組 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n$ を A の固有ベクトルからなる $\mathbb{K}^n$ の基底とし, $\boldsymbol{p}_i$ は固有値 γ_i に関する固有ベクトルとする (つまり $A\boldsymbol{p}_i = \gamma_i \boldsymbol{p}_i$). $P := [\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n]$ 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 27.2 項 (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. $P\boldsymbol{e}_i = \boldsymbol{p}_i$ および, この両辺に左から P^{-1} をかけて得られる $\boldsymbol{e}_i = P^{-1}\boldsymbol{p}_i$ に注意すれば

$$\text{“}P^{-1}AP \text{ の } i \text{ 列目”} = P^{-1}AP\boldsymbol{e}_i = P^{-1}A\boldsymbol{p}_i = P^{-1}\gamma_i\boldsymbol{p}_i = \gamma_i P^{-1}\boldsymbol{p}_i = \gamma_i \boldsymbol{e}_i.$$

ゆえに

$$P^{-1}AP = [\gamma_1\boldsymbol{e}_1, \cdots, \gamma_n\boldsymbol{e}_n] = \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix}.$$

(3)⇒(2):

$$P^{-1}AP = \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix} = B \qquad (\text{したがって } A = PBP^{-1})$$

であると仮定し, $P = [\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n]$ とおく. さらに, ベクトルの組 $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n \in \mathbb{K}^n$ を次で定める:

$$[\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n] := [\boldsymbol{e}_1, \cdots, \boldsymbol{e}_1]P = EP = P = [\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n].$$

このとき, $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n$ は $\mathbb{K}^n$ の基底であり, また $\boldsymbol{q}_i$ が固有値 γ_i の固有ベクトルとなることは, 27.2 項 (II) で述べた通りである. あるいは, $\boldsymbol{q}_i$ 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 ($P\boldsymbol{e}_i = \boldsymbol{q}_i$ に注意する):

$$\begin{aligned} A\boldsymbol{q}_i &= PBP^{-1}\boldsymbol{q}_i = PBP^{-1}(P\boldsymbol{e}_i) = PB\boldsymbol{e}_i \\ &= P\begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix}\boldsymbol{e}_i = P(\gamma_i\boldsymbol{e}_i) = \gamma_i P\boldsymbol{e}_i = \gamma_i\boldsymbol{q}_i. \end{aligned}$$

(2)⇒(1): A の相異なるすべての固有値を $\lambda_1, \cdots, \lambda_r$ とし, $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_n \in \mathbb{K}^n$ を A の固有ベクトルからなる基底とする.

$$\sum_{k=1}^{r} \dim W(\lambda_k, A) \geq \sum_{k=1}^{r} \Big(\text{“}\boldsymbol{q}_i \in W(\lambda_k, A) \text{ を満たす } i \text{ の個数”} \Big) \geq n.$$

上の式と系 28.1.2 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 31 および 32 節にて論じる.

系 28.2.2. n 次正方行列 A が n 個の相異なる固有値 $\lambda_1, \cdots, \lambda_n$ 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. □

備考 28.2.3. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 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 28.2.2 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 27.4.4 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) $\Phi_A(t) = 0$ の相異なる解 $\lambda_1, \cdots \lambda_k$ について, 各固有空間 $W(\lambda_k, A) = \mathrm{Ker}(\lambda_k I - T_A)$ の次元を求める. $W(\lambda_k, A)$ は連立 1 次方程式 $(\lambda_k E - A)\boldsymbol{x} = \boldsymbol{0}$ 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.

(3) 定理 28.2.1 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.

(4) 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 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n$ であり, 定理 28.2.1 における (2)⇒(3) の証明によれば $P = [\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n]$ について $B = P^{-1}AP$ は対角行列となる.

(5) このとき対角行列 B の (i, i)-成分は, $\boldsymbol{p}_i$ 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, 本論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 28.2.4. 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.

$$\Phi_A(t) = |tE - A| = \begin{vmatrix} t-5 & -6 & 0 \\ 1 & t & 0 \\ -1 & -2 & t-2 \end{vmatrix} = \begin{vmatrix} t-5 & -6 \\ 1 & t \end{vmatrix} (t-2) \qquad \text{(定理 12.5.1 を用いた)}$$

$$= \Big((t-5)t - (-6) \cdot 1\Big)(t-2) = (t^2 - 5t + 6)(t-2) = (t-2)^2(t-3).$$

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 $(3E-A)\boldsymbol{x}=\boldsymbol{0}$ を解くために $3E-A$ を簡約化すると

$$3E-A=\begin{bmatrix}-2&-6&0\\1&3&0\\-1&-2&1\end{bmatrix}\longrightarrow\begin{bmatrix}1&3&0\\-2&-6&0\\-1&-2&1\end{bmatrix}\longrightarrow\begin{bmatrix}1&3&0\\0&0&0\\0&1&1\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&-3\\0&1&1\\0&0&0\end{bmatrix}.$$

ゆえに $(3E-A)\boldsymbol{x}=\boldsymbol{0}$ 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 $W(2,A)$:

$$2E-A=\begin{bmatrix}-3&-6&0\\1&2&0\\-1&-2&0\end{bmatrix}\longrightarrow\begin{bmatrix}0&0&0\\1&2&0\\0&0&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&0\\0&0&0\\0&0&0\end{bmatrix}.$$

ゆえに $(2E-A)\boldsymbol{x}=\boldsymbol{0}$ 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 A の対角化が分かる. 途中の計算にミスがないか心配な場合は, 対角化 $B=P^{-1}AP$ の両辺に左から P をかけた等式 $PB=AP$ が成り立つかどうかを確認するとよい.

(2) $A=\begin{bmatrix}1&3&2\\0&-1&0\\1&2&0\end{bmatrix}.$

$$\Phi_A(t)=|tE-A|=\begin{vmatrix}t-1&-3&-2\\0&t+1&0\\-1&-2&t\end{vmatrix}=(-1)\begin{vmatrix}-1&t-1&-2\\t+1&0&0\\-2&-1&t\end{vmatrix}\qquad(1\,\text{列と}\,2\,\text{列の入れ替え})$$
$$=(-1)^2\begin{vmatrix}t+1&0&0\\-1&t-1&-2\\-2&-1&t\end{vmatrix}=(t+1)\begin{vmatrix}t-1&-2\\-1&t\end{vmatrix}$$
$$=(t+1)\Big((t-1)t-(-2)(-1)\Big)=(t+1)(t^2-t-2)=(t+1)^2(t-2).$$

- $\lambda=2$ の固有空間 $W(2,A)$:

$$2E{-}A=\begin{bmatrix}1&-3&-2\\0&3&0\\-1&-2&2\end{bmatrix}\longrightarrow\begin{bmatrix}1&-3&-2\\0&3&0\\0&-5&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&-3&-2\\0&1&0\\0&1&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&-2\\0&1&0\\0&0&0\end{bmatrix}.$$

ゆえに $(2E-A)\boldsymbol{x}=\boldsymbol{0}$ 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 $W(-1, A)$:

$$-E - A = \begin{bmatrix} -2 & -3 & -2 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに方程式 $(-E-A)\boldsymbol{x} = \boldsymbol{0}$ 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, 備考 28.1.3 において予告したことを認めると, $1 \leq \dim W(\lambda, A) \leq 1$ 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.

例 27.1.5 に現れた行列について改めて考えてみよう.

練習 28.2.5. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ 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[80]. ゆえに $P = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$ とすれば $P^{-1}AP = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$.

28.3 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 27.2 項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 28.3.1. n 次元線形空間 U 上の線形変換 $f : U \to U$ について次は同値である.

(i) f の相異なる固有値を $\lambda_1, \cdots, \lambda_r$ 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)⇔(ii): U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 28.2.1 における条件 (1), (2) と同値になる. 実際, (i)⇔(1) および (ii)⇔(2) は命題 27.5.1 による. 条件 (1), (2) の同値性より, 条件 (i), (ii) も同値である.

(i)⇒(iii): U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 27.5.1(3) より A は定理 28.2.1 における条件 (1) を満たし, したがって対角化できる.

[80] これらの計算は練習 27.4.4(1) から直ちに分かる.

(iii)⇒(i): U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 28.2.1 における条件 (1) を満たす. 命題 27.5.1(3) より, これは f が (i) を満たすことに他ならない.

(iii)⇒(iv): 明らか.

(iv)⇒(v): 27.2 項 (II) で述べた通り.

(v)⇒(iii): U の基底 $\boldsymbol{u}_1', \cdots, \boldsymbol{u}_n'$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による $\boldsymbol{u}_1', \cdots, \boldsymbol{u}_n'$ 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. □

系 28.2.2 に対応する主張は次の通りである:

系 28.3.2. n 次元線形空間 U 上の線形変換 $f : U \to U$ について, 特性方程式 $\Phi_f(t) = 0$ が n 個の相異なる解 $\lambda_1, \cdots, \lambda_n$ 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 28.2.2 より A は対角化可能であり, したがって定理 28.3.1 より f の固有ベクトルからなる基底が存在する. □

練習 28.3.3. 例 27.5.2 で定めた線形変換 $D : \mathbb{R}[x]_3 \to \mathbb{R}[x]_3$ $\left(D(p(x)) := \dfrac{d}{dx}p(x)\right)$ 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. □

基底 $\mathbf{1}, x, x^2, x^3$ に関する上の D の表現行列は, 例題 26.1.4(2) より $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ であり, $A^4 = O$ である (これは 3 次多項式を 4 回微分すると必ず $\mathbf{0}$ になることに相当している). このような行列は, 次のような一般論によって対角化できないことが分かる.

命題 28.3.4 (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 $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_n$ からなる $\mathbb{R}^n$ の基底がある. 例 27.1.3 より A の固有値は 0 のみであるゆえ, 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 $A = O$. □

28.4 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 28.4.1 (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 $f : U \to U$ について, $\Phi_f(f) = \mathbf{0}_U$.

練習 21.5.3 で扱ったように, 行列のサイズが小さければ成分表示した行列に対して定義通りに計算するだけで公式の成立が確かめられる. しかし, そのような方法は公式の正しさのみが分かるだけであり, また定理の成立を不思議に思う域を超えることはない. このような理解も数学の一つの捉え方ではあるものの, できれば定理の成立をあらかじめ想像できるに越したことはない. そこで本項では, 論理的には不要な議論となるが, A が対角化可能ならば固有ベクトルへの分解を通して必然性をもって定理が導かれることを解説しよう.

はじめに, 定理の (i) と (ii) がほぼ同じ主張であり, 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.

定理 28.4.1 における (i) と (ii) の同値性　(i)⇒(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)⇒(i): 行列 A から定まる自然な線形写像 $T_A : \mathbb{R}^n \to \mathbb{R}^n$ 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 $\Phi_A(T_A)$ 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 $\Phi_A(A)\boldsymbol{e}_j = T_{\Phi_A(A)}(\boldsymbol{e}_j) = \Phi_A(T_A)(\boldsymbol{e}_j) = \mathbf{0}_{\mathbb{R}^n}$. ゆえに, $\Phi_A(A)$ は零行列である. □

命題 28.4.2. n 次正方行列 A が対角化可能であるとき, $\Phi_A(A) = O$.

Proof. 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 $\boldsymbol{v}$ について $\Phi(A)\boldsymbol{v} = \mathbf{0}$ を示せば十分である (練習 18.1.4). $\boldsymbol{v}$ 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. $A^k\boldsymbol{v} = \lambda^k\boldsymbol{v}$ および $\Phi_A(\lambda) = 0$ に注意すれば

$$\Phi_A(A)\boldsymbol{v} = \left(\sum_{k=0}^{n} a_k A^k\right)\boldsymbol{v} = \sum_{k=0}^{n}(a_k A^k\boldsymbol{v}) = \sum_{k=0}^{n}(a_k\lambda^k\boldsymbol{v}) = \left(\sum_{k=0}^{n} a_k\lambda^k\right)\boldsymbol{v} = \Phi_A(\lambda)\boldsymbol{v} = \mathbf{0}.$$

□

いまの証明は簡明ではあるものの, 固有ベクトルが写像される様子がいま一つ見えてこない読者もいることだろう. そこで, 写像の合成の観点から上の命題の成立をもう少し詳しく調べてみたい. 例えば補題 28.1.4 は次のように見ることができる:

補題 28.1.4 (再掲)**.** f の相異なる固有値を $\lambda_1, \cdots, \lambda_r$ とする. $\boldsymbol{u}_k \in W(\lambda_k, f)$ $(k = 1, \cdots, r)$ かつ $\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r = \mathbf{0}$ ならば, $\boldsymbol{u}_1 = \cdots = \boldsymbol{u}_r = \mathbf{0}$.

Proof. $(f - \lambda_k I)(\boldsymbol{u}_k) = f(\boldsymbol{u}_k) - \lambda\boldsymbol{u}_k = \lambda_k\boldsymbol{u}_k - \lambda_k\boldsymbol{u}_k = \mathbf{0}$ に注意する. いまから, $\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r = \mathbf{0}$ の両辺に線形変換 $(f - \lambda_k I)$ $(k = 2, \cdots, r)$ を順次ほどこすことで $\boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ の項が消えて $\boldsymbol{u}_1$ 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 $k = 2, \cdots, r$ 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,

$$\begin{aligned}\Psi_1(f)(\boldsymbol{u}_k) &= \Big(\Theta_k(f)(f - \lambda_k I)\Big)(\boldsymbol{u}_k) = \Theta_k(f) \circ (f - \lambda_k I)(\boldsymbol{u}_k) \\ &= \Theta_k(t)\Big((f - \lambda_k I)(\boldsymbol{u}_k)\Big) = \Theta_k(t)(\mathbf{0}) = \mathbf{0}\end{aligned}$$

である. ゆえに $\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r = \mathbf{0}$ の両辺に線形変換 $\Psi_1(f)$ をほどこせば

$$\begin{aligned}\Psi_1(f)(\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r) &= \mathbf{0} \\ \sum_{k=1}^{r}\Psi_1(f)(\boldsymbol{u}_k) &= \mathbf{0} \\ \Psi_1(f)(\boldsymbol{u}_1) + \sum_{k=2}^{r}\Psi_1(f)(\boldsymbol{u}_k) &= \mathbf{0} \\ \Psi_1(f)(\boldsymbol{u}_1) &= \mathbf{0}.\end{aligned}$$

また, 各 $k = 2, \cdots, r$ について $(f - \lambda_k I)(\boldsymbol{u}_1) = \lambda_1\boldsymbol{u}_1 - \lambda_k\boldsymbol{u}_1 = (\lambda_1 - \lambda_k)\boldsymbol{u}_1$ である. これを順次繰り返

すことで

$$\begin{aligned}
\mathbf{0} = \Psi_1(f)(\boldsymbol{u}_1) &= \Big((f-\lambda_2 I)\cdots(f-\lambda_r I)\Big)(\boldsymbol{u}_1) \\
&= (f-\lambda_2 I)\circ\cdots\circ(f-\lambda_{r-1} I)\circ(f-\lambda_r I)\,(\boldsymbol{u}_1) \\
&= (f-\lambda_2 I)\circ\cdots\circ(f-\lambda_{r-1} I)\,\Big(\,(\lambda_1-\lambda_r)\boldsymbol{u}_1\,\Big) \\
&= (f-\lambda_2 I)\circ\cdots\circ(f-\lambda_{r-2} I)\,\Big(\,(\lambda_1-\lambda_{r-1})(\lambda_1-\lambda_r)\boldsymbol{u}_1\,\Big) \\
&= \cdots = (\lambda_1-\lambda_2)\cdots(\lambda_1-\lambda_r)\boldsymbol{u}_1.
\end{aligned}$$

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 $\boldsymbol{u}_1=\mathbf{0}$ である.

次に $\boldsymbol{u}_1+\cdots+\boldsymbol{u}_r=\mathbf{0}$ の両辺に $(f-\lambda_k I)$ $(k=1,3,4,\cdots,r)$ を順次ほどこせば, 同様の理由により $\boldsymbol{u}_2=\mathbf{0}$ を得る. これらと類似の議論を $\boldsymbol{u}_3,\cdots,\boldsymbol{u}_r$ についても行い, $\boldsymbol{u}_1=\cdots=\boldsymbol{u}_r=\mathbf{0}$ を得る. □

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 28.4.3. 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 $\lambda_1,\cdots,\lambda_r$ 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 $\boldsymbol{v}\in W(\lambda_k,A)$ について $\phi(A)\boldsymbol{v}=\mathbf{0}$ 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 $(A-\lambda_k E)\boldsymbol{v}=A\boldsymbol{v}-\lambda_k\boldsymbol{v}=\mathbf{0}$ から

$$\phi(A)\boldsymbol{v}=\Theta(A)(A-\lambda_k E)\boldsymbol{v}=\Theta(A)\mathbf{0}=\mathbf{0}.$$

以上より $\phi(A)=O$. □

本節の最後に付録として述べた因数定理 (定理 28.5.3) から, 命題 28.4.2 の別証明が導かれる:

<u>命題 28.4.2 の別証明.</u> A の相異なるすべての固有値を $\lambda_1,\cdots,\lambda_r$ 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 (系 28.5.4). ゆえに $\Phi_A(A)=\Psi(A)\phi(A)=\Psi(A)O=O$. □

上に述べた見方を発展させる形で一般の正方行列におけるケーリー・ハミルトンの定理を示すには, 一般固有空間分解について学ぶ必要がある (32.2 項をみよ). ここでは予備知識を仮定せずにできる定理の証明を与えておこう.

よりみち (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.

上のコラムにおける勘違いを反省し, また, ケーリー・ハミルトンの定理の別証明のための準備として, $\mathrm{End}(U)$ の元を成分とする行列を考えよう. $\mathrm{End}(U)$ の間には和と積が定義されていたゆえ, $\mathbb{K}$ を成分と

する行列の場合と同様にして $\mathrm{End}(U)$ 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 $\mathrm{End}(U)$ 成分の n 次正方行列とすれば, その行列式

$$\det X := \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma)(g_{1\sigma(1)} \circ g_{2\sigma(2)} \circ \cdots \circ g_{n\sigma(n)})$$

は $\mathrm{End}(U)$ 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, $\mathrm{End}(U)$ 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 $\mathrm{End}(U)$ 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 $g_{ij} \in \mathrm{End}(U)$ 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 $\widetilde{X}$ とすれば, 13.2 項で論じたことと同様にして

$$\widetilde{X}X = \begin{bmatrix} \det X & & \\ & \ddots & \\ & & \det X \end{bmatrix}$$

が成り立つ.

<u>定理 28.4.1 の証明.</u>　(ii) を示す. U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ による f の表現行列を $A = [a_{ij}]$ とすれば, $[f(\boldsymbol{u}_1), \cdots, f(\boldsymbol{u}_n)] = [\boldsymbol{u}_1, \cdots \boldsymbol{u}_n]A$ である. すなわち,

$$\begin{aligned} f(\boldsymbol{u}_1) &= a_{11}\boldsymbol{u}_1 + a_{21}\boldsymbol{u}_2 + \cdots + a_{n1}\boldsymbol{u}_n \\ f(\boldsymbol{u}_2) &= a_{12}\boldsymbol{u}_1 + a_{22}\boldsymbol{u}_2 + \cdots + a_{n2}\boldsymbol{u}_n \\ &\vdots \\ f(\boldsymbol{u}_n) &= a_{1n}\boldsymbol{u}_1 + a_{2n}\boldsymbol{u}_2 + \cdots + a_{nn}\boldsymbol{u}_n \end{aligned}$$

である. 上の連立式を移項し, また恒等写像 I をあえて用いれば次のように書ける:

$$\begin{aligned} (f - a_{11}I)(\boldsymbol{u}_1) + (-a_{21}I)(\boldsymbol{u}_2) + \cdots + (-a_{n1}I)(\boldsymbol{u}_n) &= \mathbf{0}_U \\ (-a_{12}I)(\boldsymbol{u}_1) + (f - a_{22}I)(\boldsymbol{u}_2) + \cdots + (-a_{n2}I)(\boldsymbol{u}_n) &= \mathbf{0}_U \\ &\vdots \\ (-a_{1n}I)(\boldsymbol{u}_1) + (-a_{2n}I)(\boldsymbol{u}_2) + \cdots + (f - a_{nn}I)(\boldsymbol{u}_n) &= \mathbf{0}_U \end{aligned}$$

ここで, $\mathrm{End}(U)$ の元 g と, U の元 $\boldsymbol{u}$ の積を $g\boldsymbol{u} := g(\boldsymbol{u})$ と定め, $\mathrm{End}(U)$ の元を成分とする行列と, ベクトルを成分とする列の間の積を導入すれば, 上式は次のように表せる:

$$\begin{bmatrix} f - a_{11}I & -a_{21}I & \cdots & -a_{n1}I \\ -a_{12}I & f - a_{22}I & \cdots & -a_{n2}I \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n}I & -a_{2n}I & \cdots & f - a_{nn}I \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_n \end{bmatrix} = \begin{bmatrix} \mathbf{0}_U \\ \mathbf{0}_U \\ \vdots \\ \mathbf{0}_U \end{bmatrix}.$$

上の左辺に現れる行列を X とすれば, X の各成分は互いに可換である (命題 21.5.6). また, $\mathrm{End}(U)$ 成分の行列 Y, Z, および U 成分の列ベクトル $\boldsymbol{x}$ の間の積について結合律 $Y(Z\boldsymbol{x}) = (YZ)\boldsymbol{x}$ が成り立つことが $\mathbb{K}$ 成分行列の場合と同様にして示され, したがって上の両辺に左から X の余因子行列 $\widetilde{X}$ をかけることで

$$\begin{bmatrix} \det X & & \\ & \ddots & \\ & & \det X \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_n \end{bmatrix} = \begin{bmatrix} \mathbf{0}_U \\ \vdots \\ \mathbf{0}_U \end{bmatrix} \tag{28.4.1}$$

を得る. 一方, 次の計算により $\det X = \Phi_f(f)$ である:

$$\begin{aligned}\det X &= \text{“多項式 } \Phi_{{}^tA}(t) = |tE - {}^tA| \text{ の展開式における } t^n \text{ を } f^n \text{ に置き換えた写像”} \\ &= \Phi_{{}^tA}(f) = \Phi_A(f) = \Phi_f(f).\end{aligned}$$

したがって式 (28.4.1) は, 各 $j = 1, \cdots, n$ について $\Phi_f(f)(\boldsymbol{u}_j) = \mathbf{0}_U$ 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. □

28.5 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 28.5.1. $\mathbb{K}$ 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 28.5.2. 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 $\Psi_1, \Psi_2, \Psi_3, \cdots$ 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 回目まで過程を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. □

定理 28.5.3 (因数定理)**.** $\mathbb{K}$ 係数多項式 $\Phi(x)$ について, $\lambda \in \mathbb{K}$ を $\Phi(x) = 0$ の解とすれば, $\Phi(x)$ は $(x - \lambda)$ を因子に持つ.

Proof. $\Phi(x)$ 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 $\Phi(x) = (x - \lambda)\Psi(x)$. □

系 28.5.4. $\Phi(x)$ を多項式とする. 相異なる数 $\lambda_1, \cdots, \lambda_r$ がそれぞれ方程式 $\Phi(x) = 0$ の解ならば, ある $n - r$ 次多項式 $\Psi(x)$ を用いて $\Phi(x) = (x - \lambda_1)\cdots(x - \lambda_r)\Psi(x)$ と書ける.

Proof. $\Phi(\lambda_1) = 0$ に対して因数定理を適用すれば, $n - 1$ 次多項式 $\Psi_1(x)$ を用いて, $\Phi(x) = (x - \lambda_1)\Psi_1(x)$ と書ける. この両辺に λ_2 を代入すると $(\lambda_2 - \lambda_1)\Psi_1(\lambda_2) = 0$ であり, 仮定より $\lambda_2 - \lambda_1 \neq 0$ ゆえ $\Psi_1(\lambda_2) = 0$. これに因数定理を適用すれば, $n - 2$ 次多項式 $\Psi_2(x)$ を用いて $\Psi_1(x) = (x - \lambda_2)\Psi_2(x)$ と書ける. したがって, $\Phi(x) = (x - \lambda_1)(x - \lambda_2)\Psi_2(x)$ である. これを順次繰り返し, $\Phi(x) = (x - \lambda_1)\cdots(x - \lambda_r)\Psi_r(x)$ を得る. □

次の定理は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論では証明を割愛する.

定理 28.5.5 (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, 任意の多項式が因数分解できることが分かる:

系 28.5.6. 任意の $\mathbb{C}$ 係数 n 次多項式 $\Phi(x)$ は, $\Phi(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ と因数分解できる.

Proof. n 次多項式 $\Phi(x)$ は代数学の基本定理により解 $\lambda_1 \in \mathbb{C}$ を持つ. ゆえに因数定理から $n-1$ 次多項式 $\Psi_1(x)$ を用いて $\Phi(x) = (x-\lambda_1)\Psi_1(x)$ と書ける. また, $n-1$ 次方程式 $\Psi_1(x) = 0$ は解 $\lambda_2 \in \mathbb{C}$ を持つゆえ $n-2$ 次多項式 $\Psi_2(x)$ を用いて $\Psi_1(x) = (x-\lambda_2)\Psi_2(x)$ と書けるゆえ, $\Phi(x) = (x-\lambda_1)(x-\lambda_2)\Psi_2(x)$ である. これを順次繰り返し, 因数分解 $\Phi(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$ を得る. □

複素数を成分とする行列について, 系 28.2.2 を次のように言い換えてもよい.

系 28.5.7. 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 28.5.6 より $\Phi_A(t) = (t-\lambda_1)\cdots(t-\lambda_n)$ と因数分解できる. 仮定より, $\lambda_1, \cdots, \lambda_n$ は相異なる複素数ゆえ系 28.2.2 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, $\overline{z} := a - bi$ を z の共役な複素数という. $z + \overline{z}$ および $z\overline{z}$ は共に実数である. また, $\overline{z+w} = \overline{z} + \overline{w}$ および $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$, $\overline{z^n} = \overline{z}^n$ が成り立つ. 実数 r について $\overline{r} = r$ である.

命題 28.5.8. $\mathbb{R}$ 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, その共役な複素数 $\overline{z}$ も $\Phi(x) = 0$ の解である.

Proof. $\Phi(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ とおく. $\Phi(z) = 0$ と仮定し, $\Phi(\overline{z}) = 0$ を導こう.

$$\begin{aligned}
\Phi(\overline{z}) &= \overline{z}^n + a_{n-1}\overline{z}^{n-1} + \cdots + a_1\overline{z} + a_0 = \overline{z^n} + \overline{a_{n-1}}\,\overline{z^{n-1}} + \cdots + \overline{a_1}\,\overline{z} + \overline{a_0} \\
&= \overline{z^n} + \overline{a_{n-1}z^{n-1}} + \cdots + \overline{a_1 z} + \overline{a_0} = \overline{z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0} \\
&= \overline{\Phi(z)} = \overline{0} = 0.
\end{aligned}$$

□

次の補題は 32 節において, 主定理の証明で用いる.

補題 28.5.9 (発展)**.** 相異なる $\lambda_1, \cdots, \lambda_r \in \mathbb{K}$ 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 $i = 2, 3, \cdots, r$ について, $\xi(x)$ は $(t-\lambda_i)$ を因子に持たない.

(ii) $\eta(x)$ は $(t-\lambda_1)$ 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 $(t-\lambda_1)$ を因子に持ち, ゆえに命題 28.5.2 により $(t-\lambda_1)$ 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:

$$(x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_r)^{n_r} = (x-\lambda_1)^m\zeta(x)\eta(x). \tag{28.5.1}$$

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 (28.5.1) 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 $\lambda_1,\cdots,\lambda_r$ 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 (28.5.1) 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, $\xi(x)=(x-\lambda_1)^{n_1}$. □

29　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, 本節および次節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. 本節と次節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ま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.

本節では初めに簡単な 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, 32 節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 32.3 項を見よ).

29.1　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

$$x_{n+2} + a_1 x_{n+1} + a_0 x_n = 0 \tag{29.1.1}$$

について論じよう. 漸化式 (29.1.1) 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, $S : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$ を $S\Big((x_n)_{n\in\mathbb{N}}\Big) := (x_{n+1})_{n\in\mathbb{N}}$ 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 29.1.1. $\dim F = 2$.

Proof. 初項が 1 で第 2 項が 0 の F の元を $\boldsymbol{u}_1$, 初項が 0 で第 1 項が 1 の F の元を $\boldsymbol{u}_2$ とすれば, 任意の F の元は $\boldsymbol{u}_1, \boldsymbol{u}_2$ の線形結合で書ける. 実際, $\boldsymbol{v} \in F$ の初項を p, 第 2 項を q とすれば, $\boldsymbol{v} = p\boldsymbol{u}_1 + q\boldsymbol{u}_2$ 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 (29.1.1) 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. □

S の固有ベクトルは次のように直ちに分かる:

命題 29.1.2. $S : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$ 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 $\boldsymbol{x} = (x_n)_{n\in\mathbb{N}}$ を S の固有値 λ に関する固有ベクトルとすれば, $S(\boldsymbol{x}) = \lambda\boldsymbol{x}$, 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 $\boldsymbol{x}$ は公比 λ の等比数列である. □

漸化式 (29.1.1) 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

$S|_F : F \to F$ 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 ${\lambda_1}^{n-1}$ および ${\lambda_2}^{n-1}$ 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 (29.1.1) を満たすとすれば,

$$\text{各 } n \in \mathbb{N} \text{ について,} \quad t^{n+1} + a_1 t^n + a_0 t^{n-1} = 0.$$

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 (29.1.1) 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**と言う[81]. 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 29.1.3. 漸化式 (29.1.1) 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 28.3.1 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 32.3.4).

例題 29.1.4. 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

$$\alpha = \frac{1+\sqrt{5}}{2}, \qquad \beta = \frac{1-\sqrt{5}}{2}.$$

固有値 α, β に関する固有ベクトルをそれぞれ

$$\boldsymbol{x}^\alpha = (1, \alpha, \alpha^2, \alpha^3, \alpha^4, \cdots), \qquad \boldsymbol{x}^\beta = (1, \beta, \beta^2, \beta^3, \beta^4, \cdots)$$

とおけば, 漸化式 $a_{n+2} = a_{n+1} + a_n$ を満たす数列 $\boldsymbol{x} := (x_n)_{n\in\mathbb{N}}$ 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

$$\boldsymbol{x} = p\boldsymbol{x}^\alpha + q\boldsymbol{x}^\beta, \qquad \text{とくに第 } n \text{ 項について } x_n = p\alpha^{n-1} + q\beta^{n-1}.$$

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

$$x_n = \frac{x_2 - x_1\beta}{\alpha - \beta}\alpha^{n-1} + \frac{x_1\alpha - x_2}{\alpha - \beta}\beta^{n-1}, \qquad \text{ただし } \alpha = \frac{1+\sqrt{5}}{2}, \beta = \frac{1-\sqrt{5}}{2}.$$

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 $0, 1, 1, 2, 3, 5, 8, 13, 21, \cdots$ の一般項は次で与えられる:

$$x_n = p\alpha^{n-1} + q\beta^{n-1} = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n-1}.$$

[81] 実は, Φ は $S|_F$ の特性多項式 $\Phi_{S|_F}$ に一致する. 詳しくは 29.3 項をみよ.

29.2 複素数列のなかの実数列

いま漸化式 (29.1.1) 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 $\beta = \overline{\alpha}$ が成り立つ (命題 28.5.8). このとき命題 29.1.3 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 29.2.1. p, q を複素数とし, $\alpha \in \mathbb{C}$ を実数でないとする. 数列 $x_n = p\alpha^{n-1} + q\overline{\alpha}^{n-1}$ が実数列となるための必要十分条件は $p = \overline{q}$.

Proof. 複素数 $p\alpha^{n-1}+q\overline{\alpha}^{n-1}$ 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

$$\alpha = re^{i\theta}, \quad \overline{\alpha} = re^{-i\theta}, \quad p = a_1 + b_1 i, \quad q = a_2 + b_2 i$$

と書ける. ただし $r > 0$ とし, 必要があれば α と $\overline{\alpha}$ の役割を入れ替えることで, $0 < \theta < \pi$ としてよい.

$$\begin{aligned}
x_n &= pr^{n-1}e^{i(n-1)\theta} + qr^{n-1}e^{i(n-1)\theta} \\
&= pr^{n-1}\big(\cos(n-1)\theta + i\sin(n-1)\theta\big) + qr^{n-1}\big(\cos(n-1)\theta - i\sin(n-1)\theta\big) \\
\frac{x_n}{r^{n-1}} &= (p+q)\cos(n-1)\theta + (p-q)i\sin(n-1)\theta \\
&= \big((a_1+a_2)+(b_1+b_2)i\big)\cos(n-1)\theta + \big((a_1-a_2)+(b_1-b_2)i\big)i\sin(n-1)\theta \\
&= \big((a_1+a_2)\cos(n-1)\theta - (b_1-b_2)\sin(n-1)\theta\big) \\
&\qquad + i\big((b_1+b_2)\cos(n-1)\theta + (a_1-a_2)\sin(n-1)\theta\big).
\end{aligned}$$

したがって, $p = \overline{q}$ (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. □

$p = \overline{q}$ のとき, 上の補題の証明によれば,

$$x_n = r^{n-1}\big(2a_1\cos(n-1)\theta - 2b_1\sin(n-1)\theta\big)$$

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 (29.1.1) を満たす実数列 x_n の一般項は次のように表せる:

$$x_n = r^{n-1}\big(C_1\cos(n-1)\theta + C_2\sin(n-1)\theta\big) \qquad (C_1, C_2 \text{ は任意定数}).$$

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.

29.3 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

$$x_{n+k} + a_{k-1}x_{n+k-1} + a_{k-2}x_{n+k-2} + \cdots + a_1x_{n+1} + a_0x_n = 0 \tag{29.3.1}$$

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.

$$x_{n+k} = -a_{k-1}x_{n+k-1} + -a_{k-2}x_{n+k-2} - \cdots - a_1x_{n+1} - a_0x_n \tag{29.3.2}$$

漸化式 (29.3.2) 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, 本項ではこれを $F = F_k$ 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 29.3.1. $\dim F = k$.

Proof. 次のような k 個の F の元からなる組 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_k$ を考えよう.

$$\begin{aligned}
\boldsymbol{u}_1 &= (\ 1,\ 0,\ \cdots,\ 0,\ -a_0,\ a_0a_{k-1},\ \cdots\ ),\\
\boldsymbol{u}_2 &= (\ 0,\ 1,\ \cdots,\ 0,\ -a_1,\ -a_0+a_1a_{k-1},\ \cdots\ ),\\
&\ \vdots\\
\boldsymbol{u}_k &= (\ 0,\ 0,\ \cdots,\ 1,\ -a_{k-1},\ a_{k-2}+a_{k-1}^2,\ \cdots\ ).
\end{aligned}$$

すなわち, 各 $\boldsymbol{u}_i$ 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 (29.3.2) 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, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_k$ の線形独立性はすぐに分かる. 次に, 各 $\boldsymbol{v} = (v_1, v_2, v_3, \cdots) \in F$ に対して,

$$\boldsymbol{v} = \sum_{i=1}^{k} v_i \boldsymbol{u}_i$$

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 (29.3.2) 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 29.3.2 として与えておこう. 以上により, $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_k$ は F の基底となる. □

数列 $\boldsymbol{x} = (x_n)_{n\in\mathbb{N}}$ に対して, $\boldsymbol{x}$ の第 k 項までに限った有限列を $\boldsymbol{x}|_k$ と書く. つまり $\boldsymbol{x}|_k = (x_1, \cdots, x_k)$ である.

補題 29.3.2. 各 $\boldsymbol{u}, \boldsymbol{v} \in F$ について, $\boldsymbol{u}|_k = \boldsymbol{v}|_k \Longrightarrow \boldsymbol{u} = \boldsymbol{v}$.

Proof. $\boldsymbol{u} = (u_1, u_2, \cdots)$, $\boldsymbol{v} = (v_1, v_2, \cdots)$ 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 $\boldsymbol{u}|_{k+n-1} = \boldsymbol{v}|_{k+n-1}$) と仮定すと, 漸化式 (29.3.2) より

$$\begin{aligned}
u_{n+k} &= a_{k-1}u_{n+k-1} + a_{k-2}u_{n+k-2} + \cdots + a_1u_{n+1} + a_0u_n\\
&= a_{k-1}v_{n+k-1} + a_{k-2}v_{n+k-2} + \cdots + a_1v_{n+1} + a_0v_n = v_{n+k}.
\end{aligned}$$

つまり $\boldsymbol{u}, \boldsymbol{v}$ の第 $k+n$ 項も一致する. □

29.1 項で述べたことと同様に, $S|_F$ の固有値を求めるには,

$$\Phi(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_1t + a_0 \tag{29.3.3}$$

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 29.3.1 で与えた F の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ について, それぞれを S に代入すると

$$\begin{aligned}
S(\boldsymbol{u}_1) &= (0, \cdots, 0, -a_0, a_0a_{k-1}, \cdots) = -a_0\boldsymbol{u}_k,\\
S(\boldsymbol{u}_2) &= (1, \cdots, 0, -a_1, -a_0+a_1a_{k-1}, \cdots) = \boldsymbol{u}_1 - a_1\boldsymbol{u}_k,\\
&\ \vdots\\
S(\boldsymbol{u}_i) &= \boldsymbol{u}_{i-1} - a_{i-1}\boldsymbol{u}_k,\\
&\ \vdots\\
S(\boldsymbol{u}_k) &= (0, \cdots, 1, -a_{k-1}, a_{k-2}+a_{k-1}^2, \cdots) = \boldsymbol{u}_{k-1} - a_{k-1}\boldsymbol{u}_k.
\end{aligned}$$

したがって基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ に関する S の表現行列は

$$A = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & \Large O & \\ & & \ddots & \ddots & \\ & \Large O & & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{k-2} & -a_{k-1} \end{bmatrix}. \tag{29.3.4}$$

$S|_F$ 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:

$$\Phi_{S|_F}(t) = |tE - A| = \begin{bmatrix} t & -1 & & & \\ & t & -1 & \Large O & \\ & & \ddots & \ddots & \\ & \Large O & & t & -1 \\ a_0 & a_1 & \cdots & a_{k-2} & t + a_{k-1} \end{bmatrix} = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0.$$

定義 29.3.3. 漸化式 (29.3.1) の左辺において, 各 x_{n+i} $(i = 0, \cdots, k)$ 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 (29.3.3) 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 28.1.1 により線形独立であり, したがって F の基底となる:

命題 29.3.4. 初項 1 公比 λ の等比数列を $\boldsymbol{x}^\lambda$ とする (つまり $\boldsymbol{x}^\lambda = (1, \lambda, \lambda^2, \lambda^3, \cdots)$). 漸化式 (29.3.1) の特性方程式が互いに異なる k 個の解 $\lambda_1, \cdots, \lambda_k$ を持つならば, この漸化式を満たす任意の数列 $\boldsymbol{x} = (x_n)_{n\in\mathbb{N}}$ は $\boldsymbol{x} = \sum_{i=1}^k r_i \boldsymbol{x}^{\lambda_i}$ と表せる. このとき, $\boldsymbol{x}$ の一般項の表示は $x_n = \sum_{i=1}^k r_i {\lambda_i}^{n-1}$ である.

30 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 項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節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. 本節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, それを超えるの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, これを一般の開区間 I としても同様に議論が進められるだろう.

30.1 線形常微分方程式と固有値

前節と同様に, 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

$$y''(x) + a_1 y'(x) + a_0 y(x) = 0 \tag{30.1.1}$$

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 (30.1.1) を満たす関数全体のなす $C^\infty(\mathbb{R})$ の部分空間を W とし, $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ 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 (30.1.1) 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 $y'(x)$ の値さえ決まれば, それ以降の n 次導関数 $y^{(2)}(x), y^{(3)}(x), y^{(4)}(x), \cdots$ 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 (29.1.1) 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 30.1.1. 微分作用素 $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ $(D(y) := y')$ 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,

$$\frac{dy}{dx} = \lambda y \tag{30.1.2}$$

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 (30.1.2) 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

$$\int \frac{1}{y}\frac{dy}{dx}\,dx = \int \lambda\,dx = \lambda x + C_1. \qquad (C_1 \text{ は積分定数})$$

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:

$$\int \frac{1}{y}\frac{dy}{dx}\,dx = \int \frac{1}{y}\,dy = \log|y| + C_2. \qquad (C_2 \text{ は積分定数})$$

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 30.1.2. (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 (30.1.2) 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 (30.1.2) 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

$D|_W : W \to W$ 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,

$$\begin{aligned} \left(e^{tx}\right)'' + a_1\left(e^{tx}\right)' + a_0 e^{tx} &= 0 \\ (t^2 + a_1 t + a_0)e^{tx} &= 0. \end{aligned}$$

$e^{tx} \neq 0$ 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 (30.1.1) 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 $\left(y^{(n-1)}(0)\right)_{n\in\mathbb{N}} = (t^{n-1})_{n\in\mathbb{N}}$ は漸化式 (29.1.1) を満たす. この事実からも $t^2+a_1t+a_0 = 0$ 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 30.1.3. 微分方程式 (30.1.1) 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 $y(x) = re^{\lambda_1 x} + se^{\lambda_2 x}$ と表される.

例題 30.1.4. 微分方程式 $\dfrac{d^2}{dx^2}y(x) = \dfrac{d}{dx}y(x) + y(x)$ 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

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \qquad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

ゆえに方程式の一般解は

$$y(x) = C_1 e^{\frac{1+\sqrt{5}}{2}x} + C_2 e^{\frac{1-\sqrt{5}}{2}x} \qquad (r_1, r_2 \text{ は任意定数}).$$

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 32.3.3).

30.2 特性多項式が複素解をもつ場合における実数解

29.2 項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 (30.1.1) の特性方程式が複素解 $\lambda, \overline{\lambda}$ 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 30.1.3 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 (30.1.1) の実数関数としての解を得る.

補題 30.2.1. c_1, c_2 および λ を複素数の定数とし, $y(x) := c_1 e^{\lambda x} + c_2 e^{\overline{\lambda} x}$ とする. 関数 $y : \mathbb{R} \to \mathbb{C}$ が実数値関数になるための必要十分条件は, $c_2 = \overline{c_1}$.

Proof. 実数 $a, b, \alpha_1, \beta_1, \alpha_2, \beta_2$ を用いて $\lambda, \overline{\lambda}, c_1, c_2$ を

$$\lambda = a + bi, \quad \overline{\lambda} = a - bi, \quad c_1 = \alpha_1 + \beta_1 i, \quad c_2 = \alpha_2 + \beta_2 i$$

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,

$$\begin{aligned}
y(x) &= c_1 e^{(a+bi)x} + c_2 e^{(a-bi)x} = c_1 e^{ax} e^{ibx} + c_2 e^{ax} e^{-ibx} \\
&= c_1 e^{ax}(\cos(bx) + i\sin(bx)) + c_2 e^{ax}(\cos(bx) - i \sin(bx)) \\
&= e^{ax}\Big((c_1 + c_2)\cos(bx) + (c_1 - c_2) i \sin(bx)\Big) \\
&= e^{ax}\Big(\big((\alpha_1 + \alpha_2) + (\beta_1 + \beta_2)i\big)\cos(bx) + \big((\alpha_1 - \alpha_2) + (\beta_1 - \beta_2)i\big) i \sin(bx)\Big) \\
&= e^{ax}\Big(\big((\alpha_1 + \alpha_2)\cos(bx) - (\beta_1 - \beta_2)\sin(bx)\big) + i\big((\beta_1 + \beta_2)\cos(bx) + (\alpha_1 - \alpha_2)\sin(bx)\big)\Big).
\end{aligned}$$

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

$$e^{ax}\Big((\beta_1 + \beta_2)\cos(bx) + (\alpha_1 - \alpha_2)\sin(bx)\Big)$$

が常に 0 であることと同値である.

いまの考察から, $c_2 = \overline{c_1}$ 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 $c_2 = \overline{c_1}$ 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[82]. つまり, $c_2 = \overline{c_1}$ である. また, このとき

$$y(x) = 2\alpha_1 e^{ax}\cos(bx) - 2\beta_1 e^{ax}\sin(bx) \tag{30.2.1}$$

□

式 (30.2.1) 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 30.2.2. 実数を係数とする微分方程式 (30.1.1) の特性方程式が実数でない複素数による解 $\lambda, \overline{\lambda}$ 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:

$$y(x) = r_1 e^{ax}\cos(bx) + r_2 e^{ax}\sin(bx) \qquad (r_1, r_2 \text{ は任意定数}).$$

練習 30.2.3. 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

$$y(x) = r_1 e^{-x}\cos(\sqrt{2}x) + r_2 e^{-x}\sin(\sqrt{2}x).$$

30.3 高次の線形常微分方程式 (発展)

高次の常微分方程式

$$y^{(k)}(x) + a_{k-1}y^{(k-1)}(x) + a_{k-2}y^{(k-2)}(x) + \cdots + a_1 y^{(1)}(x) + a_0 y^{(0)}(x) = 0. \tag{30.3.1}$$

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, 本項ではこれを $W = W_k$ 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.

[82] 例えば $x = 0$ を代入することで $\beta_1 + \beta_2 = 0$ を得る. $x = \dfrac{\pi}{2b}$ を代入することで $\alpha_1 - \alpha_2 = 0$ を得る.

定理 30.3.1. $a \in \mathbb{R}$ を一つ固定する．$z(x) \in W$ が $(z(a), z'(a), \cdots, z^{(k-1)}(x)) = \mathbf{0}$ 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 30.3.2. (1) $a \in \mathbb{R}$ および $\xi(x), \eta(x) \in W$ について,

$$(\xi(a), \xi'(a), \cdots, \xi^{(k-1)}(x)) = (\eta(a), \eta'(a), \cdots, \eta^{(k-1)}(a)) \implies \xi(x) = \eta(x).$$

(2) $\dim W \leq k$.

Proof. 線形写像 $G : W \to \mathbb{R}^k$ を $G(\xi) := (\xi(a), \xi'(a), \cdots, \xi^{(k-1)}(x))$ と定める．定理 30.3.1 によれば $\operatorname{Ker} G$ は自明であり, したがって命題 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 $\dim W \leq \dim \mathbb{R}^k = k$. □

練習 30.3.3. 上の結果を用いて命題 30.1.1 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 30.3.2(1) より $\xi(x) = \eta(x)$. □

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 32.3.3)．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, 29.3 項で与えた部分空間 $F \subset \mathbb{R}^{\mathbb{N}}$ のかりそめの基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_k$ 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 (30.3.1) の両辺を n 階微分することで漸化式 (29.3.1) 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

$$\boldsymbol{y} = (y(a), y'(a), y''(a), y'''(a), \cdots, y^{(n)}(a), \cdots)$$

とすれば, 数列 $\boldsymbol{y}$ は 29.3 項における 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 $u_1(x), \cdots, u_n(x)$ を考える:

$$\begin{gathered}
(u_1(a), u_1'(a), \cdots, {u_1}^{(k-1)}(x)) = (1, 0, \cdots, 0) \\
(u_2(a), u_2'(a), \cdots, {u_2}^{(k-1)}(x)) = (0, 1, \cdots, 0) \\
\vdots \\
(u_k(a), u_k'(a), \cdots, {u_k}^{(k-1)}(x)) = (0, 0, \cdots, 1)
\end{gathered}$$

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 $u_1(x), \cdots, u_k(x)$ の存在性は明らかではない．しかし天下り的に言えば, 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 $u_1(x), \cdots, u_k(x)$ 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 (29.3.1) 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 $(y(a), y'(a), y''(a), y'''(a), \cdots) = (b_1, b_2, b_3, b_4, \cdots)$ を得る．また, 項別微分可能性から, この $y(x)$ が式 (30.3.1) を満たすことも分かる．実際, b_n が式 (29.3.1) を満たすことから $b_{n+k} + a_{k-1}b_{n+k-1} + a_{k-2}b_{n+k-2} + \cdots + a_1 b_{n+1} + a_0 b_n = 0$ であり,

$$\begin{aligned}
&y^{(k)}(x) + a_{k-1}y^{(k-1)}(x) + a_{k-2}y^{(k-2)}(x) + \cdots + a_1 y^{(1)}(x) + a_0 y^{(0)}(x) \\
&\qquad = \sum_{n=0}^{\infty} \frac{b_{n+k} + a_{k-1}b_{n+k-1} + a_{k-2}b_{n+k-2} + \cdots + a_1 b_{n+1} + a_0 b_n}{n!}(x-a)^n = 0
\end{aligned}$$

次の二つの補題により, $u_1(x), \cdots, u_k(x)$ は W の基底となる．

補題 30.3.4. $u_1(x), \cdots, u_n(x)$ は線形独立である．

Proof. 系 30.3.2 の証明で与えた写像 G は $u_i(x)$ を $\boldsymbol{e}_i$ にうつす．$\boldsymbol{e}_i$ の線形独立性より $u_i(x)$ も線形独立である (命題 20.1.7). □

補題 30.3.5. 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 30.3.2(1) から, これらは同じ関数であることが分かる． □

基底 $u_1(x), \cdots, u_n(x)$ に関する $D : W \to W$ の表現行列を求めよう．そのためには, 各 $D(u_1), \cdots, D(u_n)$ を $u_1, \cdots, u_n$ の線形結合で表示した際の係数を見ればよい．

$$\begin{aligned}
&u_1 \text{ の点 } a \text{ における微分係数列：} && (\ 1,\ 0,\ \cdots,\ 0,\ -a_0,\ a_0 a_{k-1},\ \cdots\ ), \\
&u_2 \text{ の点 } a \text{ における微分係数列：} && (\ 0,\ 1,\ \cdots,\ 0,\ -a_1\ , -a_0 + a_1 a_{k-1},\ \cdots\ ), \\
& && \vdots \\
&u_k \text{ の点 } a \text{ における微分係数列：} && (\ 0,\ 0,\ \cdots,\ 1,\ -a_{k-1},\ a_{k-2} + a_{k-1}^2,\ \cdots\ ),
\end{aligned}$$

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ゆえ

$$\begin{aligned}
&D(u_1) \text{ の点 } a \text{ における微分係数列：} && (\ 0,\ \cdots,\ 0,\ -a_0,\ a_0 a_{k-1},\ \cdots\ ), \\
&D(u_2) \text{ の点 } a \text{ における微分係数列：} && (\ 1, \cdots, 0, -a_1, -a_0 + a_1 a_{k-1}, \cdots), \\
& && \vdots \\
&D(u_k) \text{ の点 } a \text{ における微分係数列：} && (\ 0,\ \cdots,\ 1,\ -a_{k-1},\ a_{k-2} + a_{k-1}^2,\ \cdots\ ).
\end{aligned}$$

したがって補題 30.3.5 より

$$D(u_1) = -a_0 u_k,\ D(u_2) = u_1 - a_1 u_k,\ \cdots,\ D(u_i) = u_{i-1} - a_{i-1}u_k,\ \cdots,\ D(u_k) = u_{k-1} - a_{k-1}u_k.$$

上式から, $D|_W$ の表現行列は 29.3 項で求めた $S|_F$ のそれと一致し, 式 (29.3.4) で与えらる行列 A となる．また, $D|_W$ の特性多項式は式 (29.3.3) で与えた $\Phi(t)$ に等しい．

定義 30.3.6. 微分方程式 (30.3.1) の左辺において, 各 $y^{(i)}(x)$ $(i = 0, \cdots, k)$ を t^i に置き換えた多項式 $\Phi(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0$ 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 29.3.4 と同様にして, 次が成り立つ:

命題 30.3.7. 微分方程式 (30.3.1) 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 $\lambda_1, \cdots, \lambda_k$ 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 (30.3.1) 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).

本項で与えた関数の組 $u_1, \cdots, u_k$ 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 30.3.4). 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 30.3.4 の証明にあるように, $k-1$ 階微分可能な関数の組 $f_1(x), f_2(x), \cdots, f_k(x)$ 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

$$\begin{bmatrix} f_1(a) \\ f_1'(a) \\ f_1''(a) \\ \vdots \\ f_1^{(k-1)}(a) \end{bmatrix}, \begin{bmatrix} f_2(a) \\ f_2'(a) \\ f_2''(a) \\ \vdots \\ f_2^{(k-1)}(a) \end{bmatrix}, \cdots, \begin{bmatrix} f_k(a) \\ f_k'(a) \\ f_k''(a) \\ \vdots \\ f_k^{(k-1)}(a) \end{bmatrix}$$

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:

$$W(f_1, f_2, \cdots, f_k)(a) = \begin{vmatrix} f_1(a) & f_2(a) & \cdots & f_k(a) \\ f_1'(a) & f_2'(a) & \cdots & f_k''(a) \\ \vdots & \vdots & \cdots & \vdots \\ f_1^{(k-1)}(a) & f_2^{(k-1)}(a) & \cdots & f_k^{(k-1)}(a) \end{vmatrix}.$$

行列式 $W(f_1, f_2, \cdots, f_k)(a)$ を関数の組 $f_1, f_2, \cdots, f_k$ の**ロンスキー行列式** (**Wronskian**) という.

補足. 組 $f_1, f_2, \cdots, f_k$ 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

$$W(f, g)(a) = \begin{vmatrix} a^2 & a|a| \\ 2a & |a| \end{vmatrix} = a^2|a| - a^2|a| = 0.$$

31 不変部分空間と冪零部分空間 (発展)

これまで, 線形変換 $f : U \to U$ 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 28.2.1(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 28.2.1(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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ 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.

$$A \text{ の第 } j \text{ 列} \quad = \quad \text{線形結合 } f(\boldsymbol{u}_j) = \sum_{i=1}^{n} a_{ij}\boldsymbol{u}_i \text{ に現れる係数を並べた列ベクトル.}$$

したがって, A をより簡単な行列にせよという課題は, $f(\boldsymbol{u}_j) = \sum_{i=1}^{n} a_{ij}\boldsymbol{u}_i$ 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. 本節では, 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.

31.1 不変部分空間

U の基底を $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ とする. このとき, j 番目のベクトル $\boldsymbol{u}_j$ として $f \in \mathrm{End}(U)$ の固有ベクトルを選ぶと, $f(\boldsymbol{u}_j) = \lambda \boldsymbol{u}_j$ ゆえ $f(\boldsymbol{u}_j)$ は $\boldsymbol{u}_j$ 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 $\lambda \boldsymbol{e}_j$ になる. ここで, $\boldsymbol{u}_j$ が固有ベクトルであるための条件が次のように書きかえられることに注意しよう.

$$f(\langle \boldsymbol{u}_j \rangle) \subset \langle \boldsymbol{u}_j \rangle.$$

そこで上の条件を一般化し, U の基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ の一部として次の条件を満たす組 $\boldsymbol{u}_\ell, \boldsymbol{u}_{\ell+1} \cdots, \boldsymbol{u}_{\ell+k}$ を考える:

$$f(\langle \boldsymbol{u}_\ell, \boldsymbol{u}_{\ell+1} \cdots, \boldsymbol{u}_{\ell+k} \rangle) \subset \langle \boldsymbol{u}_\ell, \boldsymbol{u}_{\ell+1} \cdots, \boldsymbol{u}_{\ell+k} \rangle.$$

すると, 各 $f(\boldsymbol{u}_{\ell+j})$ $(j = 0, \cdots, k)$ は $\boldsymbol{u}_\ell, \boldsymbol{u}_{\ell+1} \cdots, \boldsymbol{u}_{\ell+k}$ 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, $\ell, \ell+1, \cdots, \ell+k$ 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 31.1.1. 線形変換 $f : U \to U$ 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 31.1.2. U を有限次元線形空間とし, $f \in \mathrm{End}(U)$ とする. 各 $W_1, \cdots, W_r$ が U の f-不変部分空間であり, W_γ $(\gamma = 1, \cdots, r)$ の基底 $\boldsymbol{u}_{\gamma,1}, \cdots, \boldsymbol{u}_{\gamma,n_\gamma}$ 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 $\mathcal{B} = \{\, \boldsymbol{u}_{\gamma,i} \mid \gamma = 1, \cdots, r,\ i = 1, \cdots, n_\gamma \,\}$ 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 $(\gamma = 1, \cdots, r)$ 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 $j = 1, \cdots, n$ は次のように書ける:

$$j = n_0 + n_1 + n_2 + \cdots + n_{\gamma-1} + k \qquad (\text{ただし, } 1 \le \gamma \le r \text{ かつ } 1 \le k \le n_\gamma)$$

我々が示すべき事は A の第 j 列目の成分のうち第 $n_0 + \cdots + n_{\gamma-1} + 1$ 成分から第 $n_0 + \cdots + n_{\gamma-1} + n_\gamma$ 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 $j = n_0 + \cdots + n_{\gamma-1} + k$ 番目のベクトルは

$\boldsymbol{u}_{\gamma,k} \in W_\gamma$ であり, W_γ が f-不変なことから $f(\boldsymbol{u}_{\gamma,k}) \in W_\gamma$ である. つまり, 基底 $\mathcal{B}$ の線形結合で $f(\boldsymbol{u}_{\gamma,k})$ を表示した際に現れる係数のうち 0 でないものは $\boldsymbol{u}_{\gamma,1}, \cdots, \boldsymbol{u}_{\gamma,n_\gamma}$ 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 $\boldsymbol{u}_{\gamma,1}, \cdots, \boldsymbol{u}_{\gamma,n_\gamma}$ に対応する成分, すなわち第 $n_0+\cdots+n_{\gamma-1}+1$ 成分から第 $n_0+\cdots+n_{\gamma-1}+n_\gamma$ 成分に限られる. □

備考 31.1.3. 上の命題における各 A_γ は, 基底 $\boldsymbol{u}_{\gamma,1}, \cdots, \boldsymbol{u}_{\gamma,n_\gamma}$ に関する $f|_{W_\gamma} : W_\gamma \to W_\gamma$ の表現行列に等しい.

例 31.1.4. U が $f \in \mathrm{End}(U)$ の固有ベクトルからなる基底 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ を持つとき, 各 $W_j = \langle \boldsymbol{u}_j \rangle$ 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.

31.2 冪零部分空間

本節の始めに提示した問題を前項とは別の視点から論じよう. 前項では, 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. 本項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 $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$ を上手く取ることで $T(\boldsymbol{u}_j) = \sum_{i=1}^n a_{ij}\boldsymbol{u}_i$ 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, $\boldsymbol{u}_j$ が固有ベクトルであるとき, すなわち $f(\boldsymbol{u}_j)$ が $\boldsymbol{u}_j$ 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, 次に考えうる最も単純な形は, $\boldsymbol{u}_j$ 自身と, 基底をなす別のもう一つのベクトル $\boldsymbol{u}_i$ との線形結合によって $f(\boldsymbol{u}_j)$ が表せる場合であろう. ここで更に踏み込んで, 各 $f(\boldsymbol{u}_j)$ が, $\boldsymbol{u}_j$ と一つ隣のベクトル $\boldsymbol{u}_{j-1}$ の線形結合で書ける場合, すなわち

$$f(\boldsymbol{u}_j) = s_{j-1}\boldsymbol{u}_{j-1} + r_j\boldsymbol{u}_j \qquad (j = 2, \cdots, n)$$

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 $\boldsymbol{u}_1$ は固有ベクトルである (つまり $f(\boldsymbol{u}_1) = r_1\boldsymbol{u}_1$) 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 $\boldsymbol{u}_j$ は固有ベクトルである. 一方 $s_{j-1} \neq 0$ の場合は $\boldsymbol{u}_j$ の代わりに $\frac{1}{s_{j-1}}\boldsymbol{u}_j$ を基底として取れば,

$$f\left(\frac{1}{s_{j-1}}\boldsymbol{u}_j\right) = \frac{1}{s_{j-1}}f(\boldsymbol{u}_j) = \frac{1}{s_{j-1}}(s_{j-1}\boldsymbol{u}_{j-1} + r_j\boldsymbol{u}_j) = \boldsymbol{u}_{j-1} + r_j\left(\frac{1}{s_{j-1}}\boldsymbol{u}_j\right)$$

であるから $s_{j-1} = 1$ の場合が本質的である[83]. 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 27.3.7). そこで, 列 $r_1, \cdots, r_n$ が各固有値ごとに行儀よく並んで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 $\lambda_1, \cdots, \lambda_r$ とし, U の基底

[83] このとき, 次に並んでいるベクトル $\boldsymbol{u}_{j+1}$ に関する式 $f(\boldsymbol{u}_{j+1}) = s_j\boldsymbol{u}_j + r_{j+1}\boldsymbol{u}_{j+1}$ について, いま $\boldsymbol{u}_j$ 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 $\boldsymbol{u}_{j+1}$ 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 $s_{j-1}, s_j, \cdots, s_{n-1}$ を 0 または 1 に置き換えられる.

$\{\boldsymbol{u}_{\lambda_k,j} \mid k=1,\cdots,r,\ j=1,\cdots,n_k\}$ に関する表現行列 A が次のようになる場合である:

$$A=\begin{bmatrix} A_{\lambda_1} & & & \\ & A_{\lambda_2} & & \\ & & \ddots & \\ & & & A_{\lambda_r} \end{bmatrix}, \qquad A_{\lambda_k}=\begin{bmatrix} \lambda_k & s_{\lambda_k,1} & & & \\ & \lambda_k & s_{\lambda_k,2} & & \\ & & \ddots & \ddots & \\ & & & \lambda_k & s_{\lambda_k,n_k-1} \\ & & & & \lambda_k \end{bmatrix}. \tag{31.2.1}$$

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.
注意: 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 列目は固有ベクトルに対応する列であるとしてよく, ゆえに A_{λ_k} における $(1,1)$-成分の一つ上の成分は 0 であるとしてよい.
このとき, $\boldsymbol{u}_{\lambda_k,j}$ $(j=2,\cdots,n_k)$ が固有ベクトルでなければ $s_{\lambda_k,j-1}=1$ ゆえ $f(\boldsymbol{u}_{\lambda_k,j})=\boldsymbol{u}_{\lambda_k,j-1}+\lambda_k\boldsymbol{u}_{\lambda_k,j}$ である. ゆえに,

$$\boldsymbol{u}_{\lambda_k,j-1}=f(\boldsymbol{u}_{\lambda_k,j})-\lambda_k\boldsymbol{u}_{\lambda_k,j}=(f-\lambda_k I)(\boldsymbol{u}_{\lambda_k,j}).$$

つまり, $\boldsymbol{u}_{\lambda_k,j}$ に線形写像 $f-\lambda_k I$ を繰り返しほどこすことで $\boldsymbol{u}_{\lambda_k,j-1},\boldsymbol{u}_{\lambda_k,j-2},\cdots$ 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 $\boldsymbol{u}_{\lambda_k,j-\ell}$ を得る[84]. また,

$$(f-\lambda_k I)(\boldsymbol{u}_{\lambda_k,j-\ell})=f(\boldsymbol{u}_{\lambda_k,j-\ell})-\lambda_k\boldsymbol{u}_{\lambda_k,j-\ell}=\lambda_k\boldsymbol{u}_{\lambda_k,j-\ell}-\lambda_k\boldsymbol{u}_{\lambda_k,j-\ell}=\mathbf{0}$$

であるから

$$\boldsymbol{u}_{\lambda_k,j}\xmapsto{f-\lambda_k I}\boldsymbol{u}_{\lambda_k,j-1}\xmapsto{f-\lambda_k I}\boldsymbol{u}_{\lambda_k,j-2}\xmapsto{f-\lambda_k I}\ \cdots\ \xmapsto{f-\lambda_k I}\boldsymbol{u}_{\lambda_k,j-\ell}\xmapsto{f-\lambda_k I}\mathbf{0}.$$

すなわち,

$$\boldsymbol{u}_{\lambda_k,j}\in\operatorname{Ker}(f-\lambda_k I)^{\ell+1}. \tag{31.2.2}$$

こうして我々は冪零部分空間の概念に至る:

定義 31.2.1. 有限次元線形空間 U 上の線形変換 $g:U\to U$ に対して, 部分空間の増大列 $\operatorname{Ker} g\subset\operatorname{Ker} g^2\subset\operatorname{Ker} g^3\subset\cdots$ は下の補題により十分大きい N について, $\operatorname{Ker} g^N=\operatorname{Ker} g^{N+1}=\operatorname{Ker} g^{N+2}=\cdots$ を満たす. このとき, $\operatorname{Ker} g^N$ を g の**冪零部分空間**と呼ぼう[85].

定義 31.2.2. 線形変換 $f:U\to U$ の固有値 λ に対して, 線形変換 $(f-\lambda_k I):U\to U$ 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. 本論では, 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**と呼ぶ.

補足. $W(\lambda,f)=\operatorname{Ker}(\lambda I-f)=\operatorname{Ker}(f-\lambda I)$ 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.

以下では $g\in\operatorname{End}(U)$ 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 31.2.3. 有限次元線形空間 U 上の線形変換 $g:U\to U$ において次が成り立つ.

(1) $\operatorname{Ker} g\subset\operatorname{Ker} g^2\subset\operatorname{Ker} g^3\subset\cdots$.

(2) $\operatorname{Ker} g^n=\operatorname{Ker} g^{n+1}$ ならば, $\operatorname{Ker} g^n=\operatorname{Ker} g^{n+1}=\operatorname{Ker} g^{n+2}=\operatorname{Ker} g^{n+3}=\cdots$.

(3) $\dim U=N$ ならば, $\operatorname{Ker} g^N=\operatorname{Ker} g^{N+1}=\operatorname{Ker} g^{N+2}=\cdots$.

[84] 少なくとも $\boldsymbol{u}_{\lambda_k,1}$ 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.
[85] この名称は本論のみで通じるものである. $\operatorname{Ker} g^{\dim U}$ と記せばよいため, この空間に一般的な名称は与えられていない.

Proof. (1): $\boldsymbol{u} \in \operatorname{Ker} g^n$ とすれば, $g^n(\boldsymbol{u}) = \boldsymbol{0}$. よって, $g^{n+1}(\boldsymbol{u}) = g(g^n(\boldsymbol{u})) = g(\boldsymbol{0}) = \boldsymbol{0}$ ゆえ $\boldsymbol{u} \in \operatorname{Ker} g^{n+1}$.

(2): $\operatorname{Ker} g^{n+1} = \operatorname{Ker} g^{n+2}$ を示せば, あとは帰納的に各 $\operatorname{Ker} g^{n+m}$ がすべて一致することが分かる. $\operatorname{Ker} g^{n+1} \subset \operatorname{Ker} g^{n+2}$ は (1) で示したゆえ $\operatorname{Ker} g^{n+2} \subset \operatorname{Ker} g^{n+1}$ を示そう. 各 $\boldsymbol{u} \in \operatorname{Ker} g^{n+2}$ について, $\boldsymbol{0} = g^{n+2}(\boldsymbol{u}) = g^{n+1}(g(\boldsymbol{u}))$ より $g(\boldsymbol{u}) \in \operatorname{Ker} g^{n+1} = \operatorname{Ker} g^n$. よって, $g(\boldsymbol{u}) \in \operatorname{Ker} g^n$ ゆえ $g^n(g(\boldsymbol{u})) = \boldsymbol{0}$. つまり, $\boldsymbol{u} \in \operatorname{Ker} g^{n+1}$ である.

(3): g が単射ならば $\{\boldsymbol{0}\} = \operatorname{Ker} g = \operatorname{Ker} g^2 = \operatorname{Ker} g^3 = \cdots$ である. そこで, 単射でない場合を考えよう. このとき $\dim \operatorname{Ker} g \geq 1$ である. $\operatorname{Ker} g^{n+1}$ が $\operatorname{Ker} g^n$ よりも真に大きくなるとき, 命題 22.4.1 よりそれらの次元も真に大きくなる. 仮に, すべての $n = 1, \cdots, N$ について $\operatorname{Ker} g^{n+1}$ が $\operatorname{Ker} g^n$ よりも真に大きくなるとすれば,

$$\dim \operatorname{Ker} g^{N+1} \geq \dim \operatorname{Ker} g^N + 1 \geq (\dim \operatorname{Ker} g^{N-1} + 1) + 1 \geq \cdots \geq \dim \operatorname{Ker} g + N \geq 1 + N$$

となり $\dim U = N$ に矛盾してしまう. したがって, $n = 1, \cdots, N$ のいずれかにおいて $\operatorname{Ker} g^{n+1} = \operatorname{Ker} g^n$ となる必要があり, それ以降は (2) よりすべて一致する. □

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 (31.2.1) のような形の表現行列を得るためには, 式 (31.2.2) 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 32.1.1).

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 31.2.4. $f, g : U \to U$ 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, $W = \operatorname{Ker} g$ は f-不変部分空間である.

Proof. 各 $\boldsymbol{u} \in W$ について, $f(\boldsymbol{u}) \in W$ を示したい. そのためには $g(f(\boldsymbol{u})) = \boldsymbol{0}$ を言えばよい. g と f が可換になること, および $g(\boldsymbol{u}) = \boldsymbol{0}$ から,

$$g(f(\boldsymbol{u})) = g \circ f(\boldsymbol{u}) = f \circ g(\boldsymbol{u}) = f(g(\boldsymbol{u})) = f(\boldsymbol{0}) = \boldsymbol{0}.$$

□

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 31.2.5. $f, g : U \to U$ 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) $g : U \to U$ 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.

31.3 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 32.3 項を見よ).

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 31.3.1. 線形変換 $g : U \to U$ および $\boldsymbol{u} \in U$ について, $g^n(\boldsymbol{u}) \neq \boldsymbol{0}$ かつ $g^{n+1}(\boldsymbol{u}) = \boldsymbol{0}$ ならば, $n+1$ 個のベクトルからなる組 $\boldsymbol{u}, g(\boldsymbol{u}), g^2(\boldsymbol{u}), \cdots, g^n(\boldsymbol{u})$ は線形独立である.

Proof. 各 $k=0,\cdots,n$ について, $k+1$ 個の組 $g^n(\boldsymbol{u}), g^{n-1}(\boldsymbol{u}), \cdots g^{n-k}(\boldsymbol{u})$ 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 $g^n(\boldsymbol{u}), g^{n-1}(\boldsymbol{u}), \cdots g^{n-(k-1)}(\boldsymbol{u})$ が線形独立であると仮定し, $k+1$ 個の組 $g^n(\boldsymbol{u}), g^{n-1}(\boldsymbol{u}), \cdots g^{n-k}(\boldsymbol{u})$ の線形独立性を示す. 線形関係 $\sum_{i=0}^{k} r_i g^{n-i}(\boldsymbol{u}) = \boldsymbol{0}$ を仮定し, この両辺に g をほどこせば,

$$\begin{aligned} r_0 g^{n-0+1}(\boldsymbol{u}) + r_1 g^{n-1+1}(\boldsymbol{u}) + r_2 g^{n-2+1}(\boldsymbol{u}) + \cdots + r_{k-1} g^{n-k+1}(\boldsymbol{u}) &= \boldsymbol{0} \\ r_1 g^{n}(\boldsymbol{u}) + r_2 g^{n-1}(\boldsymbol{u}) + \cdots + r_{k-1} g^{n-(k-1)}(\boldsymbol{u}) &= \boldsymbol{0} \end{aligned}$$

$g^n(\boldsymbol{u}), g^{n-1}(\boldsymbol{u}), \cdots g^{n-(k-1)}(\boldsymbol{u})$ は線形独立であったゆえ $r_1 = r_2 = \cdots = r_{k-1} = 0$. ゆえに $r_0 g^n(\boldsymbol{u}) = \boldsymbol{0}$ であり, $g^n(\boldsymbol{u}) \neq \boldsymbol{0}$ より $r_0 = 0$. 以上より $k+1$ 個の組 $g^n(\boldsymbol{u}), g^{n-1}(\boldsymbol{u}), \cdots g^{n-k}(\boldsymbol{u})$ は線形独立である. □

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 31.3.2. $D : C^\infty(\mathbb{K}) \to C^\infty(\mathbb{K})$ 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 $x^n e^{\lambda x} \in \operatorname{Ker}(D-\lambda I)^{n+1}$ かつ $x^n e^{\lambda x} \notin \operatorname{Ker}(D-\lambda I)^n$.

(2) $e^{\lambda x},\ xe^{\lambda x},\ x^2 e^{\lambda x},\ \cdots,\ x^n e^{\lambda x}$ は線形独立である.

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 $e^{\lambda x} \in \operatorname{Ker}(D-\lambda I)$. 次に, $x^{n-1}e^{\lambda x} \in \operatorname{Ker}(D-\lambda I)^n$ かつ $x^{n-1}e^{\lambda x} \notin \operatorname{Ker}(D-\lambda I)^{n-1}$ を仮定して, $x^n e^{\lambda x} \in \operatorname{Ker}(D-\lambda I)^{n+1}$ および $x^n e^{\lambda x} \notin \operatorname{Ker}(D-\lambda I)^n$ を示そう.

$$\begin{aligned} (D-\lambda I)(x^n e^{\lambda x}) &= (x^n e^{\lambda x})' - \lambda x^n e^{\lambda x} = \Big(nx^{n-1}e^{\lambda x} + x^n \lambda e^{\lambda x}\Big) - \lambda x^n e^{\lambda x} \\ &= nx^{n-1}e^{\lambda x} \in \operatorname{Ker}(D-\lambda I)^n. \end{aligned}$$

ゆえに $x^n e^{\lambda x} \in \operatorname{Ker}(D-\lambda I)^{n+1}$ である. また, $nx^{n-1}e^{\lambda x} \notin \operatorname{Ker}(D-\lambda I)^{n-1}$ ゆえ $x^n e^{\lambda x} \notin \operatorname{Ker}(D-\lambda I)^n$ である[86].

(2): $g = D - \lambda I$ および $\boldsymbol{u} = x^n e^{\lambda x}$ について前補題を適用すると, $N+1$ 個の組

$$x^n e^{\lambda x},\ nx^{n-1}e^{\lambda x},\ n(n-1)x^{n-2}e^{\lambda x},\ \cdots,\ n!e^{\lambda x}$$

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 $x^n e^{\lambda x},\ x^{n-1}e^{\lambda x},\ x^{n-2}e^{\lambda x},\ \cdots,\ e^{\lambda x}$ も線形独立である. □

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も導かれる. 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 $i = 0, \cdots, k-1$ について $f^{(i)}(0) = 0$ であり, $f^{(k)}(0) = k!$ となる. したがって組 $e^{\lambda x},\ xe^{\lambda x},\ x^2e^{\lambda x},\ \cdots,\ x^n e^{\lambda x}$ 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. □

例 31.3.3. $S : \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$ をシフト作用素 $S\Big((x_n)_{n\in\mathbb{N}}\Big) := (x_{n+1})_{n\in\mathbb{N}}$ とする. $\lambda \neq 0$ および非負整数 N について, 数列 $\boldsymbol{x}_N^\lambda \in \mathbb{R}^{\mathbb{N}}$ を次で定める.

$$\boldsymbol{x}_N^\lambda := (1,\ 2^N\lambda,\ 3^N\lambda^2,\ 4^N\lambda^3,\ \cdots) \qquad (\text{つまり } \boldsymbol{x}_N^\lambda = (n^N \lambda^{n-1})_{n\in\mathbb{N}}).$$

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 $\boldsymbol{x}_N^0$ と定める.

(1) $\boldsymbol{x}_N^\lambda \in \operatorname{Ker}(S-\lambda I)^{N+1}$ かつ $\boldsymbol{x}_N^\lambda \notin \operatorname{Ker}(S-\lambda I)^N$.

(2) $\boldsymbol{x}_0^\lambda, \boldsymbol{x}_1^\lambda, \cdots, \boldsymbol{x}_N^\lambda$ は線形独立である.

---

[86] 何故なら, 仮に $\boldsymbol{u} = x^n e^{\lambda x} \in \operatorname{Ker}(D-\lambda I)^n$ と仮定すれば $(D-\lambda I)(\boldsymbol{u}) \in \operatorname{Ker}(D-\lambda I)^{n-1}$ となり, これは $(D-\lambda I)(\boldsymbol{u}) = nx^{n-1}e^{\lambda x} \notin \operatorname{Ker}(D-\lambda I)^{n-1}$ に矛盾してしまう.

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 $\boldsymbol{x}_N^\lambda$ を $\boldsymbol{x}_N$ と略そう.

(1): N に関する帰納法で示す. $N = 0$ の場合, $\boldsymbol{x}_0$ 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 $\boldsymbol{x}_0 \in \mathrm{Ker}(S - \lambda I)$ である. 次に各 $k = 0, \cdots, N-1$ について $\boldsymbol{x}_k \in \mathrm{Ker}(S - \lambda I)^{k+1}$ かつ $\boldsymbol{x}_k \notin \mathrm{Ker}(S - \lambda I)^k$ を仮定して, $\boldsymbol{x}_N \in \mathrm{Ker}(S - \lambda I)^{N+1}$ および $\boldsymbol{x}_N \notin \mathrm{Ker}(S - \lambda I)^N$ を示そう.

$$\begin{aligned}(S - \lambda I)(\boldsymbol{x}_N) &= (S - \lambda I)\Big((n^N \lambda^{n-1})_{n \in \mathbb{N}}\Big) = \Big((n+1)^N \lambda^n\Big)_{n \in \mathbb{N}} - \lambda\Big(n^N \lambda^{n-1}\Big)_{n \in \mathbb{N}} \\ &= \Big(((n+1)^N - n^N)\lambda^n\Big)_{n \in \mathbb{N}} = \lambda\Big(((n+1)^N - n^N)\lambda^{n-1}\Big)_{n \in \mathbb{N}}.\end{aligned}$$

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

$$\begin{aligned}\lambda\Big(((n+1)^N - n^N)\lambda^{n-1}\Big)_{n \in \mathbb{N}} &= \lambda\left(\left(\sum_{k=0}^{N-1} a_k n^k\right)\lambda^{n-1}\right)_{n \in \mathbb{N}} = \lambda\left(\sum_{k=0}^{N-1} a_k n^k \lambda^{n-1}\right)_{n \in \mathbb{N}} \\ &= \lambda \sum_{k=0}^{N-1}\left(a_k n^k \lambda^{n-1}\right)_{n \in \mathbb{N}} = \lambda \sum_{k=0}^{N-1} a_k \boldsymbol{x}_k \in \mathrm{Ker}(S - \lambda I)^N.\end{aligned}$$

以上より, $(S - \lambda I)(\boldsymbol{x}_N) \in \mathrm{Ker}(S - \lambda I)^N$ ゆえ $\boldsymbol{x}_N \in \mathrm{Ker}(S - \lambda I)^{N+1}$. また, 帰納法の仮定により各 $k = 0, \cdots, N-2$ について $(S - \lambda I)^{N-1}(\boldsymbol{x}_k) = \boldsymbol{0}$ および $(S - \lambda I)^{N-1}(\boldsymbol{x}_{N-1}) \neq \boldsymbol{0}$ であったから,

$$\begin{aligned}(S - \lambda I)^N(\boldsymbol{x}_N) &= (S - \lambda I)^{N-1}\Big((S - \lambda I)(\boldsymbol{x}_N)\Big) = (S - \lambda I)^{N-1}\left(\lambda \sum_{k=0}^{N-1} a_k \boldsymbol{x}_k\right) \\ &= \lambda \sum_{k=0}^{N-1} a_k (S - \lambda I)^{N-1}(\boldsymbol{x}_k) = \lambda a_{N-1}(S - \lambda I)^{N-1}(\boldsymbol{x}_{N-1}) \neq \boldsymbol{0}.\end{aligned}$$

ゆえに $\boldsymbol{x}_N \notin \mathrm{Ker}(S - \lambda I)^N$.

(2): $g = S - \lambda I$ について前補題を適用すれば, $\boldsymbol{x}_N, g(\boldsymbol{x}_N), \cdots, g^N(\boldsymbol{x}_N)$ は線形独立である. また, (1) の証明で行った計算によれば

$$\langle \boldsymbol{x}_N, g(\boldsymbol{x}_N), \cdots, g^N(\boldsymbol{x}_N) \rangle \subset \langle \boldsymbol{x}_N, \boldsymbol{x}_{N-1}, \cdots, \boldsymbol{x}_0 \rangle$$

である. したがって

$$N + 1 = \dim\langle \boldsymbol{x}_N, g(\boldsymbol{x}_N), \cdots, g^N(\boldsymbol{x}_N) \rangle \leq \dim\langle \boldsymbol{x}_N, \boldsymbol{x}_{N-1}, \cdots, \boldsymbol{x}_0 \rangle \leq N + 1$$

より $\dim\langle \boldsymbol{x}_N, \boldsymbol{x}_{N-1}, \cdots, \boldsymbol{x}_0 \rangle = N + 1$. これと命題 22.3.3(3) を合わせて, $\boldsymbol{x}_N, \boldsymbol{x}_{N-1}, \cdots, \boldsymbol{x}_0$ の線形独立性を得る. □

別解. $\boldsymbol{x}_0^\lambda, \boldsymbol{x}_1^\lambda, \cdots, \boldsymbol{x}_N^\lambda$ 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 $[\boldsymbol{x}_0^\lambda|_{N+1}, \boldsymbol{x}_1^\lambda|_{N+1}, \cdots, \boldsymbol{x}_N^\lambda|_{N+1}]$ 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 $\boldsymbol{x}_k^\lambda|_{N+1}$ 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

$$\det\Big(\boldsymbol{x}_0^\lambda|_{N+1}, \boldsymbol{x}_1^\lambda|_{N+1}, \cdots, \boldsymbol{x}_N^\lambda|_{N+1}\Big) = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \lambda & 2\lambda & 2^2\lambda & \cdots & 2^N\lambda \\ \lambda^2 & 3\lambda^2 & 3^2\lambda^2 & \cdots & 3^N\lambda^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \lambda^N & (N+1)\lambda^N & (N+1)^2\lambda^N & \cdots & (N+1)^N\lambda^N \end{vmatrix}$$

$$= \lambda^{1+2+\cdots+N} \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2^2 & \cdots & 2^N \\ 1 & 3 & 3^2 & \cdots & 3^N \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & (N+1) & (N+1)^2 & \cdots & (N+1)^N \end{vmatrix} = \lambda^{1+2+\cdots+N} \prod_{1 \leq i < j \leq N+1} (j - i) \neq 0.$$

□

31.4 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 31.4.1. 有限次元線形空間 U 上の線形変換 $g : U \to U$ に対して, 部分空間の減少列 $\operatorname{Im} g \supset \operatorname{Im} g^2 \supset \operatorname{Im} g^3 \supset \cdots$ は次の補題により十分大きい N について $\operatorname{Im} g^N = \operatorname{Im} g^{N+1} = \operatorname{Im} g^{N+2} = \cdots$ を満たす. この部分空間 $\operatorname{Im} g^N$ を g の**安定部分空間**と呼ぼう[87].

補題 31.4.2. 線形空間 U 上の線形変換 $g : U \to U$ において次が成り立つ.

(1) $\operatorname{Im} g \supset \operatorname{Im} g^2 \supset \operatorname{Im} g^3 \supset \cdots$.

(2) $\operatorname{Im} g^n = \operatorname{Im} g^{n+1}$ ならば, $\operatorname{Im} g^n = \operatorname{Im} g^{n+1} = \operatorname{Im} g^{n+2} = \cdots$.

(3) $\dim U = N$ ならば, $\operatorname{Im} g^N = \operatorname{Im} g^{N+1} = \operatorname{Im} g^{N+2} = \cdots$.

Proof. (1): 各 $g^{n+1}(\boldsymbol{u}) \in \operatorname{Im} g^{n+1}$ について, $g^{n+1}(\boldsymbol{u}) = g^n(g(\boldsymbol{u})) \in \operatorname{Im} g^n$. よって, $\operatorname{Im} g^{n+1} \subset \operatorname{Im} g^n$.

(2): $\operatorname{Im} g^{n+1} = \operatorname{Im} g^{n+2}$ さえ示せば, あとは帰納的に各 $\operatorname{Im} g^{n+m}$ (ただし $m \in \mathbb{N}$) がすべて一致することが分かる. $\operatorname{Im} g^{n+1} \supset \operatorname{Im} g^{n+2}$ は (1) で示したゆえ $\operatorname{Im} g^{n+1} \subset \operatorname{Im} g^{n+2}$ を示そう. 各 $g^{n+1}(\boldsymbol{u}) \in \operatorname{Im} g^{n+1}$ について, $g^{n+1}(\boldsymbol{u}) = g(g^n(\boldsymbol{u}))$ である. また, $g^n(\boldsymbol{u}) \in \operatorname{Im} g^n = \operatorname{Im} g^{n+1}$ ゆえ, ある $\boldsymbol{v} \in U$ を用いて $g^n(\boldsymbol{u}) = g^{n+1}(\boldsymbol{v})$ と書ける. よって, $g^{n+1}(\boldsymbol{u}) = g(g^n(\boldsymbol{u})) = g(g^{n+1}(\boldsymbol{v})) = g^{n+2}(\boldsymbol{v}) \in \operatorname{Im} g^{n+2}$.

(3): 補題 31.2.3(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 $U = \operatorname{Im} g = \operatorname{Im} g^2 = \operatorname{Im} g^3 = \cdots$ 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 $\dim U \geq \dim \operatorname{Im} g + 1$ である. $\operatorname{Im} g^{n+1}$ が $\operatorname{Im} g^n$ よりも真に小さくなるとき, 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 $n = 1, \cdots, N$ について $\operatorname{Im} g^{n+1}$ が $\operatorname{Im} g^n$ よりも真に小さくなるとすれば,

$$\dim U \geq \dim \operatorname{Im} g + 1 \geq (\dim \operatorname{Im} g^2 + 1) + 1 \geq \cdots \geq \dim \operatorname{Im} g^{N+1} + N + 1 \geq N + 1$$

となり $\dim U = N$ に矛盾してしまう. したがって, $n = 1, \cdots, N$ のいずれかにおいて $\operatorname{Im} g^{n+1} = \operatorname{Im} g^n$ となる必要があり, それ以降は (2) よりすべて一致する. □

補足. 上の (3) の証明, および補題 31.2.3(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, $g^0 = \mathrm{id}_U$ に注意し, すべての $n = 0, \cdots, N$ に対して部分空間の減少 (増大) 列が真に小さく (大きく) なるときに矛盾が生じることを示せばよい.

任意の線形変換 $g : U \to U$ 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 31.4.3. 有限次元線形空間上の線形変換 $g : U \to U$ 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) $g|_V : V \to V$ は線形同型である.

Proof. 自然数 N を十分大きく取り, $V = \operatorname{Im} g^N = \operatorname{Im} g^{N+1} = \cdots$ であるとする.
(1): $g(V) = g(\operatorname{Im} g^N) = \operatorname{Im} g^{N+1} = V$.
(2): (1) より $g|_V : V \to V$ は全射であり, ゆえに命題 22.4.8 より同型である. □

補題 31.2.4 および系 31.2.5 と類似の事実が安定部分空間においても成り立つ.

補題 31.4.4. $f, g : U \to U$ 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 $V = \operatorname{Im} g$ は f-不変部分空間である.

[87] これも本論でしか通じない名称である.

Proof. 各 $\boldsymbol{v} \in V$ について, $f(\boldsymbol{v}) \in V$ を示したい. $\boldsymbol{v} \in V = g(U)$ より, $g(\boldsymbol{u}) = \boldsymbol{v}$ を満たす $\boldsymbol{u} \in U$ が存在する. このとき, f と g の可換性より

$$f(\boldsymbol{v}) = f(g(\boldsymbol{u})) = f \circ g(\boldsymbol{u}) = g \circ f(\boldsymbol{u}) = g(f(\boldsymbol{u})) \in g(U) = V.$$

□

系 31.2.5 と同様の理由により, 次を得る:

系 31.4.5. 有限次元線形空間上の線形変換 $f : U \to U$ 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. 本論において次元定理が応用される最初の例となろう.

定理 31.4.6. 線形変換 $g : U \to U$ 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) $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_k$ を W の基底, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ 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, $W = \operatorname{Ker} g^N$ かつ $V = \operatorname{Im} g^N$ が満たされているとする.

(1): $\boldsymbol{u} \in W \cap V$ とすれば, $\boldsymbol{u} \in W$ より $g^N(\boldsymbol{u}) = \mathbf{0}$ である. 命題 31.4.3 より $g^N|_V : V \to V$ は単射であり, これと $\boldsymbol{u} \in \operatorname{Ker} g^N|_V$ を合わせれば $\boldsymbol{u} = \mathbf{0}$ を得る.

(2): まず線形独立性を示そう. $\sum_{i=1}^k c_i \boldsymbol{w}_i + \sum_{j=1}^m r_j \boldsymbol{v}_j = \mathbf{0}$ を仮定し, この式を次のように変形する:

$$W \ni \sum_{i=1}^k c_i \boldsymbol{w}_i = -\sum_{j=1}^m r_j \boldsymbol{v}_j \in V.$$

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 $\sum_{i=1}^k c_i \boldsymbol{w}_i = \mathbf{0}$ かつ $\sum_{j=1}^m r_j \boldsymbol{v}_j = \mathbf{0}$ であり, 組 $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_k$ および $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ 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 $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_k, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ 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. $g^N : U \to U$ に対して次元定理を適用すると,

$$\dim U = \dim \operatorname{Ker} g^N + \operatorname{Im} g^N = \dim W + \dim V = k + m.$$

□

31.5 直和分解 (付録)

先の定理 31.4.6 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 31.5.1. 線形空間 V の部分空間 $W_1, \cdots, W_n$ が次の条件を満たすとき, V は $W_1, \cdots, W_n$ たちによって**直和分解される**という.

(i) 和集合 $W_1 \cup W_2 \cup \cdots \cup W_n$ は V を生成する.

(ii) 各 $\boldsymbol{u}_i \in W_i$ について, $\quad \boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n = \mathbf{0} \implies \boldsymbol{u}_1 = \boldsymbol{u}_2 = \cdots = \boldsymbol{u}_n = \mathbf{0}$.

またこのとき,

$$V = \bigoplus_{i=1}^n W_i \quad \text{あるいは} \quad V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$$

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 31.5.2. 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 $\boldsymbol{v} \in V$ は, $\boldsymbol{w}_i \in W_i$ を用いて $\boldsymbol{v} = \boldsymbol{w}_1 + \cdots + \boldsymbol{w}_n$ 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 31.5.3. 線形空間 V の部分空間 $W_1, \cdots, W_r$ について, 次は同値である:

(1) 各 $\boldsymbol{u}_i \in W_i$ について, $\quad \boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r = \boldsymbol{0} \implies \boldsymbol{u}_1 = \cdots = \boldsymbol{u}_r = \boldsymbol{0}$.

(2) 任意の増大列 $1 \leq m_1 < m_2 < \cdots < m_\ell \leq r$ および任意の零ベクトルでない $\boldsymbol{u}_{m_i} \in W_{m_i}$ について, $\boldsymbol{u}_{m_1}, \cdots, \boldsymbol{u}_{m_\ell}$ は線形独立である.

(3) $\boldsymbol{v} \in V$ が各 $\boldsymbol{w}_i \in W_i$ 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) $\boldsymbol{u}_{i,1}, \boldsymbol{u}_{i,2}, \cdots, \boldsymbol{u}_{i,n_i}$ を W_i の基底とすれば, これらを全て集めた

$$\mathcal{B} = \{ \boldsymbol{u}_{i,k} \mid i = 1, \cdots, r,\ k = 1, \cdots, n_i \}$$

は線形独立なベクトルの組である.

Proof. (1)⇒(2): $\boldsymbol{u}_{m_i} \in W_{m_i}$ および $\boldsymbol{u}_{m_i} \neq \boldsymbol{0}$, $\sum_{j=1}^{\ell} c_j \boldsymbol{u}_{m_j} = \boldsymbol{0}$ を仮定する. (1) より $c_j \boldsymbol{u}_{m_j} = \boldsymbol{0}$ であり, これと $\boldsymbol{u}_{m_j} \neq \boldsymbol{0}$ から $c_j = 0$ を得る.

(2)⇒(3): $\sum_{i=1}^{r} \boldsymbol{u}_i = \sum_{i=1}^{r} \boldsymbol{w}_i$ (ただし $\boldsymbol{u}_i, \boldsymbol{w}_i \in W_i$) を仮定し, 各 $\boldsymbol{v}_i = \boldsymbol{u}_i - \boldsymbol{w}_i \in W_i$ が零ベクトルになることを示そう. いま $\sum_{i=1}^{r} \boldsymbol{v}_i = \boldsymbol{0}$ である. 仮に零ベクトルにならない $\boldsymbol{v}_i$ があるとし, それらをすべて列挙したものを $\boldsymbol{v}_{m_1}, \cdots, \boldsymbol{v}_{m_\ell}$ とすれば, (2) よりこれは線形独立である. 一方, $\sum_{j=1}^{\ell} \boldsymbol{v}_{m_j} = \boldsymbol{0}$ であり, これは非自明な線形関係ゆえ, $\boldsymbol{v}_{m_1}, \cdots, \boldsymbol{v}_{m_\ell}$ の線形独立性に反する.

(3)⇒(1): $\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_n = \boldsymbol{0}$ とする. 零ベクトルは $\boldsymbol{0} \in W_i$ をもちいて $\boldsymbol{0} = \sum_{i=1}^{r} \boldsymbol{0}$ と書ける. (3) より W_i の元の和として $\boldsymbol{0}$ を表す方法は一通りしかないことから, $\boldsymbol{u}_i = \boldsymbol{0}$ を得る.

(1)⇒(4): 28.1 項で述べた命題 28.1.1 の別証明と同様にして示される $\sum_{i=1}^{r} \sum_{k=1}^{n_i} c_{i,k} \boldsymbol{u}_{i,k} = \boldsymbol{0}$ とする. $\boldsymbol{w}_i := \sum_{k=1}^{n_i} c_{i,k} \boldsymbol{u}_{i,k}$ とおけば $\boldsymbol{w}_i \in W_i$ であり, $\sum_{i=1}^{r} \boldsymbol{w}_i = \boldsymbol{0}$. ゆえに (1) より $\boldsymbol{w}_i = \boldsymbol{0}$ となる. したがって $\sum_{k=1}^{n_i} c_{i,k} \boldsymbol{u}_{i,k} = \boldsymbol{0}$ であり, $\boldsymbol{u}_{i,1}, \boldsymbol{u}_{i,2}, \cdots, \boldsymbol{u}_{i,n_i}$ の線形独立性から $c_{i,k} = 0$.

(4)⇒(1): 各 $\boldsymbol{w}_i \in W_i$ について, $\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_r = \boldsymbol{0}$ を仮定しよう. このとき, (4) より $\boldsymbol{w}_i := \sum_{k=1}^{n_i} c_{i,k} \boldsymbol{u}_{i,k}$ と書ける. すなわち, $\sum_{i=1}^{r} \sum_{k=1}^{n_i} c_{i,k} \boldsymbol{u}_{i,k} = \boldsymbol{0}$ であり, $\mathcal{B}$ の線形独立性から $c_{i,k} = 0$. つまり $\boldsymbol{w}_i = \boldsymbol{0}$ 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. 本論では主に (4) を用いている.

系 31.5.4. 線形空間 V の部分空間 $W_1, \cdots, W_r$ について, 次は同値である:

(1) $V = \bigoplus_{i=1}^{r} W_i$.

(2) 各 $\boldsymbol{v} \in V$ が $\boldsymbol{w}_i \in W_i$ の和として表され, 任意の増大列 $1 \leq m_1 < m_2 < \cdots < m_\ell \leq r$ および任意の零ベクトルでない $\boldsymbol{u}_{m_i} \in W_{m_i}$ について, $\boldsymbol{u}_{m_1}, \cdots, \boldsymbol{u}_{m_k}$ は線形独立である.

(3) 各 $\boldsymbol{v} \in V$ が $\boldsymbol{w}_i \in W_i$ 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) $\boldsymbol{u}_{i,1}, \boldsymbol{u}_{i,2}, \cdots, \boldsymbol{u}_{i,n_i}$ を W_i の基底とすれば, これらを全て集めた

$$\mathcal{B} = \{\, \boldsymbol{u}_{i,k} \mid i = 1, \cdots, r,\ k = 1, \cdots, n_i \,\}$$

は V の基底である.

例 **31.5.5.** U を有限次元線形空間とする.

(1) 線形変換 $g : U \to U$ の冪零部分空間を W, 安定部分空間を V とすれば, $U = W \oplus W$ (定理 31.4.6).

(2) 線形変換 $f : U \to U$ 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 28.3.1).

32　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きる. 本節ではまず, 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.

32.1　一般固有空間分解

次の定理の証明は本節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 32.1.1 (一般固有空間分解)**.** 有限次元線形空間上の線形変換 $f : U \to U$ の相異なる固有値を $\lambda_1, \cdots, \lambda_r$ 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, $\dim \widetilde{W}(\lambda_k, f) = n_k$ $(k = 1, \cdots, r)$ が成り立つ. また, 各 $\widetilde{W}(\lambda_k, f)$ の基底 $\boldsymbol{u}_{\lambda_k,1}, \cdots, \boldsymbol{u}_{\lambda_k,n_k}$ 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

$$\mathcal{B} = \{\, \boldsymbol{u}_{\lambda_k,j} \mid k = 1, \cdots, r,\ j = 1, \cdots, n_k \,\}$$

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 28.5.6 により定理 32.1.1 の前提が必ず満たされ, したがって次が導かれる.

系 32.1.2. 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 $f : U \to U$ 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 32.1.3. $f : U \to U$ 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 28.3.1 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 $g = (f - \lambda_k I)|_{\widetilde{W}(\lambda_k, f)} : \widetilde{W}(\lambda_k, f) \to \widetilde{W}(\lambda_k, f)$ に対して $\dim \widetilde{W}(\lambda_k, f) = n_k$ および補題 31.2.3(3) を適用することで次を得る.

系 32.1.4. 定理 32.1.1 の前提のもとで, $\widetilde{W}(\lambda_k, f) = \mathrm{Ker}(f - \lambda_k I)^{n_k}$.

備考 32.1.5. (1) 定理 32.1.1 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:

$\boldsymbol{u}_{\lambda_k,1}, \cdots, \boldsymbol{u}_{\lambda_k,n_k}$ の選び方として, まず $\mathrm{Ker}(f-\lambda_k I)$ の基底を選び, それに $\mathrm{Ker}(f-\lambda_k I)^2$ の元を付け加えて $\mathrm{Ker}(f-\lambda_k I)^2$ の基底とし, 更に $\mathrm{Ker}(f-\lambda_k I)^3$ の元を付け加えて $\mathrm{Ker}(f-\lambda_k I)^3$ の基底とし…

という操作を繰り返して $\widetilde{W}(\lambda_k, f)$ の基底 $\boldsymbol{u}_{\lambda_k,1}, \cdots, \boldsymbol{u}_{\lambda_k,n_k}$ を得る．ここで $\boldsymbol{u}_{\lambda_k,j} \in \mathrm{Ker}(f-\lambda_k I)^{\ell}$ とすれば,

$$\begin{aligned}(f-\lambda_k I)(\boldsymbol{u}_{\lambda_k,j}) &\in \mathrm{Ker}(f-\lambda_k I)^{\ell-1}\\ f(\boldsymbol{u}_{\lambda_k,j}) - \lambda_k \boldsymbol{u}_{\lambda_k,j} &\in \mathrm{Ker}(f-\lambda_k I)^{\ell-1}\\ f(\boldsymbol{u}_{\lambda_k,j}) &\in \mathrm{Ker}(f-\lambda_k I)^{\ell-1} + \lambda_k \boldsymbol{u}_{\lambda_k,j}.\end{aligned}$$

上の最後の式は, $f(\boldsymbol{u}_{\lambda_k,j})$ が $\boldsymbol{u}_{\lambda_k,1}, \boldsymbol{u}_{\lambda_k,2}, \cdots, \boldsymbol{u}_{\lambda_k,j}$ 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 $f|_{\widetilde{W}(\lambda_k,f)} : \widetilde{W}(\lambda_k, f) \to \widetilde{W}(\lambda_k, f)$ の表現行列 A_k は上三角行列となる．これと命題 31.1.2 および備考 31.1.3 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 32.1.2 により $\mathbb{C}^n$ は $T_X : \mathbb{C}^n \to \mathbb{C}^n$ の一般固有ベクトルからなる基底を持つ．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, 内積空間の単元においては, 別の文脈から直交行列を用いた行列の上三角化が論じられる．

(3) 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 27.3.5). (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 27.3.7). したがって行列のトレースは, 重複を込めた意味での固有値の和に等しい．すなわち, 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,

$$\mathrm{tr}\, A = \sum_{k=1}^{r} \lambda_k n_k.$$

備考 32.1.6. 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 (31.2.1) 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**と言う．

次の定理は練習 27.1.6 の内容をさらに精査したものである．

定理 32.1.7 (フロベニウス)**.** 線形変換 $f : U \to U$ 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 $\gamma_1, \cdots, \gamma_n$ 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 32.1.5(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 27.3.7 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 $\gamma_1, \cdots, \gamma_n$ 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 ${\gamma_1}^k, \cdots, {\gamma_n}^k$ となる[88]．したがって $\Psi(A)$ も上三角であり, その対角成分は $\Psi(\gamma_1), \cdots, \Psi(\gamma_n)$ である．ゆえに $\Phi_{\Psi(f)}(t) = \Phi_{\Psi(A)}(t) = \prod_{i=1}^{n}(t-\Psi(\gamma_i))$. □

補足. 写像 $\Psi : \{\gamma_1, \cdots, \gamma_n\} \to \{\Psi(\gamma_1), \cdots, \Psi(\gamma_n)\}$ 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．

[88]定義 3.3.1 の直前にある計算を見よ．

32.2 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 28.4.3 と類似していることを確認してほしい．

定理 32.2.1. n 次正方行列 A について $\Phi_A(A) = O$.

Proof. 複素数の範囲において $\Phi_A(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_3}\cdots(t-\lambda_r)^{n_r}$ と因数分解しよう．ここで, $\lambda_1, \cdots, \lambda_r$ は相異なる A の固有値である．$T_A : \mathbb{C}^n \to \mathbb{C}^n$ に対して系 32.1.2 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 $\boldsymbol{v} \in \widetilde{W}(\lambda_k, A)$ について $\Phi_A(A)\boldsymbol{v} = \boldsymbol{0}$ を示せば十分である (練習 18.1.4)．系 32.1.4 より $\widetilde{W}(\lambda_k, A) = \mathrm{Ker}(T_A - \lambda_k I)^{n_k}$, つまり,

$$(A-\lambda_k E)^{n_k}\boldsymbol{v} = (T_A - \lambda_k I)^{n_k}(\boldsymbol{v}) = \boldsymbol{0}$$

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,

$$\Phi_A(A)\boldsymbol{v} = \Theta(A)(A-\lambda_k E)^{n_k}\boldsymbol{v} = \Theta(A)\boldsymbol{0} = \boldsymbol{0}.$$

以上より $\Phi_A(A) = O$. □

A が実数を成分とする正方行列であり, 実数の範囲において A の特性多項式が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．

32.3 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 (30.3.1) および線形漸化式 (29.3.1) の特性多項式は $\Phi(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ であった．関数 y が常微分方程式 (30.3.1) を満たすことは, 次の三つの同値な式に言い換えられる:

$$\begin{aligned} D^k(y) + a_{k-1}D^{k-1}(y) + a_{k-2}D^{k-2}(y) + \cdots + a_1 D(y) + a_0 I(y) &= \boldsymbol{0}, \\ \Big(D^k + a_{k-1}D^{k-1} + a_{k-2}D^{k-2} + \cdots + a_1 + a_0 I\Big)(y) &= \boldsymbol{0}, \\ \Phi(D)(y) &= \boldsymbol{0}. \end{aligned}$$

同様にして数列 $\boldsymbol{x} \in \mathbb{K}^{\mathbb{N}}$ が線形漸化式 (29.3.1) を満たすことは次を満たすことと同値である:

$$\Phi(S)(\boldsymbol{x}) = \boldsymbol{0}.$$

常微分方程式 (30.3.1) の解空間を W, 漸化式 (29.3.1) を満たす数列空間を F とする．

補題 32.3.1. 常微分方程式 (30.3.1) 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 $k = 1, \cdots, r$ について, 例 31.3.2 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 $e^{\lambda_k x}, xe^{\lambda_k x}, x^2 e^{\lambda_k x}, \cdots, x^{n_k-1}e^{\lambda_k x}$ は $D|_W : W \to W$ 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 32.1.1 より $\dim \widetilde{W}(\lambda_k, S) = n_k$ ゆえ, $e^{\lambda_k x}, xe^{\lambda_k x}, \cdots, x^{n_k-1}e^{\lambda_k x} \in W$ さえ示せばよい．つまり, 各 $y_i(x) = x^i e^{\lambda x}$ $(i = 0, \cdots, n_k - 1)$ について $\Phi(D)(y_i) = \boldsymbol{0}$ 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 31.3.2(1) より $y_i \in \mathrm{Ker}(D-\lambda_k I)^{n_k}$ ゆえ, $\Phi(D)(y_i) = \Theta(D)(D-\lambda_k I)^{n_k}(y_i) = \Theta(D)(\boldsymbol{0}) = \boldsymbol{0}$. □

いまと類似の議論で次を得る.

補題 32.3.2. 線形漸化式 (29.3.1) 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 $k = 1, \cdots, r$ について, 例 31.3.3 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 $\boldsymbol{x}_0^{\lambda_k}, \cdots, \boldsymbol{x}_{n_k-1}^{\lambda_k}$ は $S|_F : F \to F$ 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 32.1.1 の後半より次が従う:

定理 32.3.3. 常微分方程式 (30.3.1) の特性多項式が $\Phi(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_r)^{n_r}$ であるとする. ことのき, 次の集合 $\mathcal{B}$ は常微分方程式 (30.3.1) の解空間の基底となる:

$$\mathcal{B} = \{\, e^{\lambda_k x},\ xe^{\lambda_k x},\ x^2 e^{\lambda_k x},\ \cdots,\ x^{n_k-1}e^{\lambda_k x} \mid k = 1, \cdots, r \,\}.$$

つまり, 常微分方程式 (30.3.1) 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 32.3.4. 線形漸化式 (29.3.1) の特性多項式が $\Phi(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_r)^{n_r}$ 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:

$$\mathcal{B} = \{\, \boldsymbol{x}_0^{\lambda_k}, \cdots, \boldsymbol{x}_{n_k-1}^{\lambda_k} \mid k = 1, \cdots, r \,\}.$$

ここで, 各 $\boldsymbol{x}_N^{\lambda_k}$ は例 31.3.3 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.

32.4 定理 32.1.1 の証明

最後に主定理の証明を述べよう. 本項では定理 32.1.1 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 $f : U \to U$ の相異なる固有値を $\lambda_1, \cdots, \lambda_r$ 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.

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 31.2.5).

補題 32.4.1. 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, $(f-\gamma I)|_{\widetilde{W}(\lambda,f)} : \widetilde{W}(\lambda, f) \to \widetilde{W}(\lambda, f)$ は線形同型である.

(3) 多項式 $\Psi(t) = (t-\delta_1)\cdots(t-\delta_m)$ 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 $\boldsymbol{v} \neq \boldsymbol{0}$ について $\Psi(f)(\boldsymbol{v}) \neq \boldsymbol{0}$.

Proof. (1): $\widetilde{W}(\lambda, f) = \mathrm{Ker}(f-\lambda I)^N$ が成り立つよう十分大きな自然数 N を与えれば, 各 $\boldsymbol{v} \in \widetilde{W}(\lambda, f)$ について $(f-\lambda I)^N(\boldsymbol{v}) = \boldsymbol{0}$ が成り立つ. 仮にこの $\boldsymbol{v}$ が γ に関する固有ベクトルであるとすれば, $(f-\lambda I)(\boldsymbol{v}) = \gamma\boldsymbol{v} - \lambda\boldsymbol{v} = (\gamma-\lambda)\boldsymbol{v}$ であり, さらに $(f-\lambda I)$ をほどこすことで, $(f-\lambda I)^N(\boldsymbol{v}) = (\gamma-\lambda)^N\boldsymbol{v}$ を得る. ゆえに $(\gamma-\lambda)^N\boldsymbol{v} = \boldsymbol{0}$ となり, $\gamma-\lambda \neq 0$ ゆえ $\boldsymbol{v} = \boldsymbol{0}$. これは $\boldsymbol{v}$ が固有ベクトル (つまり $\boldsymbol{v} \neq \boldsymbol{0}$) であることに反する.

(2): 単射性を示せばよい. 各 $\boldsymbol{v} \in \widetilde{W}(\lambda, f)$ について, $(f-\gamma I)(\boldsymbol{v}) = \boldsymbol{0}$ とすれば $\boldsymbol{v}$ 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 $\boldsymbol{v} = \boldsymbol{0}$. 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. □

定理 32.1.1 の前半を示そう.

命題 32.4.2. $\dim \widetilde{W}(\lambda_k, f) = n_k$ $(k = 1, \cdots, r)$.

Proof. $W = \widetilde{W}(\lambda_k, f)$ とおく．W は線形写像 $(f - \lambda_k I) : U \to U$ の冪零部分空間である．W の基底を $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_d$ とする．W は f-不変であり, この基底に関する $f|_W : W \to W$ 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 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ とする．V も f-不変であり, この基底に関する $f|_V : V \to V$ の表現行列を B としよう．定理 31.4.6 より U は W と V に直和分解される．すなわち, $\boldsymbol{w}_1, \cdots, \boldsymbol{w}_d$ および $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ 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 $f : U \to U$ の表現行列を X とすれば, 命題 31.1.2 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:

$$\Phi_f(t) = \Phi_X(t) = \Phi_A(t)\Phi_B(t) = \Phi_{f|_W}(t)\Phi_{f|_V}(t).$$

$\Phi_{f|_W}(t) = (t - \lambda_k)^{n_k}$ を示すには, 補題 28.5.9 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 $\boldsymbol{v} \in W$ が存在する．しかしこれは補題 32.4.1(1) に矛盾してしまう． □

- $\Phi_{f|_V}(t)$ 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 $\boldsymbol{v} \neq \boldsymbol{0}$ を含む．このとき $(f - \lambda_k I)(\boldsymbol{v}) = \boldsymbol{0}$ であり, これは $(f - \lambda_k I) : V \to V$ の単射性 (命題 31.4.3(2)) に反する． □

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． □

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 28.1.4 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 32.4.3. $\boldsymbol{u}_k \in \widetilde{W}(\lambda_k, f)$ $(k = 1, \cdots, r)$ かつ $\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r = \boldsymbol{0}$ ならば, $\boldsymbol{u}_1 = \cdots = \boldsymbol{u}_r = \boldsymbol{0}$.

Proof. 系 32.1.4 より $\widetilde{W}(\lambda_k, f) = \operatorname{Ker}(f - \lambda_k I)^{n_k}$, つまり $(f - \lambda_k I)^{n_k}(\boldsymbol{u}_k) = \boldsymbol{0}$ である．いまから, $\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r = \boldsymbol{0}$ の両辺に $(f - \lambda_k I)^{n_k}$ $(k = 2, \cdots, r)$ を順次ほどこすことで $\boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ の項が消えて $\boldsymbol{u}_1$ に関する項のみが残ることを見よう．

多項式 $\Psi_1(t) = (t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$ は, 別の多項式 $\Theta_k(t)$ を用いて $\Psi_1(t) = \Theta_k(t)(t - \lambda_k)^{n_k}$ $(k = 2, \cdots, r)$ と書ける．つまり $\Psi_1(f) = \Theta_k(f)(f - \lambda_k I)^{n_k}$ である．$\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r = \boldsymbol{0}$ の両辺に線形変換 $\Psi_1(f)$ をほどこすと

$$\begin{aligned} \Psi_1(f)(\boldsymbol{u}_1, \cdots + \boldsymbol{u}_r) &= \boldsymbol{0} \\ \sum_{k=1}^{r} \Psi_1(f)(\boldsymbol{u}_k) &= \boldsymbol{0} \\ \Psi_1(f)(\boldsymbol{u}_1) + \sum_{k=2}^{r} \Theta_k(f)(f - \lambda_k I)^{n_k}(\boldsymbol{u}_k) &= \boldsymbol{0} \\ \Psi_1(f)(\boldsymbol{u}_1) &= \boldsymbol{0}. \qquad (\text{ここで } (f - \lambda_k I)^{n_k}(\boldsymbol{u}_k) = \boldsymbol{0} \text{ を用いた}) \end{aligned}$$

ゆえに補題 32.4.1(3) より $\boldsymbol{u}_1 = \boldsymbol{0}$ でなければならない．

次に $\boldsymbol{u}_1 + \cdots + \boldsymbol{u}_r = \boldsymbol{0}$ の両辺に $(f - \lambda_k I)^{n_k}$ $(k = 1, 3, 4, \cdots, r)$ を順次ほどこすことで $\boldsymbol{u}_2 = \boldsymbol{0}$ を得る．これらと類似の操作を順次繰り返し, $\boldsymbol{u}_1 = \cdots = \boldsymbol{u}_r = \boldsymbol{0}$ を得る． □

定理 32.1.1 の後半は次の通りである.

命題 32.4.4. 各 $\widetilde{W}(\lambda_k, f)$ の基底 $\boldsymbol{u}_{k,1}, \cdots, \boldsymbol{u}_{k,n_k}$ 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

$$\mathcal{B} = \{\, \boldsymbol{u}_{k,j} \mid k = 1, \cdots, r,\ j = 1, \cdots, n_k \,\}$$

を取れば, $\mathcal{B}$ は U の基底になる.

Proof. 命題 32.4.2 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 31.5.3 における (1) と (4) の同値性から導かれる. □