

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	6
1.1	大学で学ぶということ	6
1.2	数学教育の役割	6
1.3	この講義が目指すところ	7
1.4	数学概論	7
1.5	ユークリッド空間における和とスカラー倍	9
1.6	写像とその合成	10
1.7	そして線形代数学へ	11
2	行列の演算	15
2.1	行列の成分表示	15
2.2	行列の和とスカラー倍	16
2.3	行列の積	16
2.4	行列演算の性質	18
3	行列の表し方	21
3.1	クロネッカーのデルタ	21
3.2	\sum 記号の使い方	21
3.3	成分の空白と任意性	23
3.4	転置行列	24
3.5	行列の分割	26
4	連立1次方程式	28
4.1	導入	28
4.2	連立1次方程式と行列	28
4.3	逆行列を持つ場合	29
4.4	行列の行基本変形	29
4.5	簡約行列	30
4.6	連立1次方程式の解法	32
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	34
5	可逆行列	36
5.1	逆行列の性質	36
5.2	行基本変形再考	37
5.3	逆行列の求め方	39
5.4	基本行列と列基本変形 (発展)	40
6	行列の階数	41
6.1	簡約化の一意性	41
6.2	連立1次方程式と階数	43
6.3	同次形の方程式	43
7	行列式とは何か	46
7.1	\mathbb{R}^n 上の線形変換の面積拡大率	46
7.2	クラメルの公式	48
7.3	微積分学における行列式	49

8	置換	51
8.1	置換の定義	51
8.2	置換の表示	51
8.3	置換の積	52
8.4	巡回置換とその表示	53
8.5	置換の符号	55
9	置換の符号について	59
9.1	符号の正当性	59
9.2	文字列の並び替え (よりみち)	61
9.3	転倒数による符号の定義 (よりみち)	63
10	行列式の定義と性質	65
10.1	定義	65
10.2	カヴァリエリの原理	66
10.3	多重線形性	67
10.4	歪対称性	69
11	行列式の計算	71
11.1	サイズの小さい行列式との関係	71
11.2	計算例	72
11.3	体積との関係	74
12	行列式の性質 (証明)	76
12.1	$ A = {}^tA $ の証明	76
12.2	歪対称性の証明	77
12.3	多重線形性と歪対称性から導かれる性質	78
12.4	行列式の特徴づけ (証明)	79
12.5	命題 11.1.1 の証明	81
12.6	$ AB = A \cdot B $ の証明	82
13	余因子展開とクラメルの公式	83
13.1	余因子展開	83
13.2	余因子行列	85
13.3	クラメルの公式の証明	88
14	集合概念の基礎	90
14.1	集合の包含関係	90
14.2	集合の表し方	91
14.3	外延的か内包的か	93
14.4	集合論と逆理 (よりみち)	94
15	線形空間	97
15.1	ベクトル空間の公理	97
15.2	線形空間の例	98
15.3	体 K 上の線形空間 (発展)	100

16	いろいろな線形部分空間	103
16.1	定義	103
16.2	\mathbb{R}^n の部分空間	105
16.3	部分空間の様々な例	107
17	線形結合と線形独立性	109
17.1	線形結合	109
17.2	線形独立性	110
17.3	線形独立性の判定	112
18	基底	115
18.1	ベクトルの組が生成する部分空間	115
18.2	基底の定義と例	115
18.3	基底の探し方	118
18.4	一般の基底	120
19	写像概念の基礎	123
19.1	像と逆像	123
19.2	全射と単射	124
19.3	逆写像はいつ定まるか	126
19.4	逆写像とその性質	127
19.5	写像の合成	130
19.6	無限集合 (発展)	131
20	線形写像	135
20.1	線形写像の基本的性質	135
20.2	線形写像による像	138
20.3	線形写像による逆像	139
20.4	様々な線形写像の例	141
21	線形空間の同一視	145
21.1	線形同型写像	145
21.2	同型な線形空間	146
21.3	線形写像のなす空間	148
21.4	多元環とその準同型 (発展)	149
21.5	線形変換と多項式 (発展)	150
22	線形空間の次元	153
22.1	次元の定義	153
22.2	連立 1 次方程式の任意定数の個数	154
22.3	線形独立な最大個数	155
22.4	次元から分かること	157
22.5	無限次元の空間も含めた一般論 (発展)	159
23	次元公式と商空間	162
23.1	空間の並行移動	162
23.2	線形写像の次元公式	164
23.3	商空間 (発展)	165
23.4	商空間の例 (発展)	167
23.5	同値関係と商集合 (発展)	168

24 準同型定理と短完全列 (発展)	171
24.1 商空間の線形構造	171
24.2 準同型定理	172
24.3 完全系列と短完全列	173
25 線形結合の行列表示	177
25.1 線形結合の組と行列	177
25.2 ベクトルの組と行列の演算の基本性質 (付録)	178
25.3 線形結合再考	179
25.4 線形独立性の判定 (2)	180
25.5 基底の変換行列	181
26 線形写像の表現行列	183
26.1 表現行列	183
26.2 $\text{Hom}(U, V)$ と $M_{m,n}(\mathbb{R})$	185
26.3 基底の取りかたによる表現行列の違い	187
26.4 1対1の対応と可換図式	188
27 固有値と固有ベクトル	191
27.1 固有ベクトル	192
27.2 固有ベクトルからなる基底と行列の対角化	193
27.3 特性多項式	194
27.4 固有空間	195
27.5 一般の線形写像の固有空間	197
28 固有空間分解と行列の対角化	199
28.1 固有ベクトルの線形独立性	199
28.2 対角化可能条件	200
28.3 一般の線形変換の場合	204
28.4 対角化可能な行列に関するケーリー・ハミルトンの定理	205
28.5 多項式と方程式の解に関する基本的性質 (付録)	206
29 不変部分空間と冪零部分空間 (発展)	208
29.1 不変部分空間	208
29.2 冪零部分空間	209
29.3 $\text{Ker}(f - \lambda I)^n$ 上の元の例	211
29.4 広義固有空間分解	213
29.5 ケーリー・ハミルトンの定理	214
30 ジョルダン標準形 (発展)	217
30.1 冪零部分空間と安定像空間への分解	217
30.2 証明	218
30.3 ジョルダン標準系	220
31 線形漸化式と常微分方程式	221
31.1 線形漸化式と固有値	221
31.2 常微分方程式を解くための補題	223
31.3 定数係数常微分方程式の一般解	224
31.4 特性多項式が複素解をもつ場合における実数解	226
31.5 特性多項式が重解をもつ場合	226

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用いるのではなく、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、二つの異なる世界を一つの見方で繋げること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、それらの中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ことよい。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、大雑把に言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。このことは既知のことであるとして、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. $5 \in \mathbb{N}$, $-1 \notin \mathbb{N}$, $\frac{12}{13} \in \mathbb{Q}$, $\sqrt{2} \notin \mathbb{Q}$, $\text{ぶどう} \notin \{ \text{りんご, みかん, スイカ} \}$.

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、各 $n \in \mathbb{N}$ について、 n 個の実数の並び順を込めた意味での組 (x_1, x_2, \dots, x_n) たち全体からなる集合を \mathbb{R}^n と表し、これを n 次元ユークリッド空間 (n -dimensional Euclidean space) という。また、 \mathbb{R}^n の元のことをベクトル (vector) と呼ぶ。高校で学習したように、 \mathbb{R}^2 や \mathbb{R}^3 のベクトルの間には和とスカラー倍が定まっているのであった。同様のことが \mathbb{R}^n においても定義される。すなわち、 \mathbb{R}^n の二つのベクトル $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$ および実数 $r \in \mathbb{R}$ に対して、ベクトルの和 $(x_1, \dots, x_n) + (y_1, \dots, y_n)$ およびベクトルのスカラー倍 (scalar multiplication) $r(x_1, \dots, x_n)$ を次で定める:

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

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、 $\mathbf{x} = (x_1, \dots, x_n)$ 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。 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (ただし $n = 2, 3$) に対して、 $\mathbf{0}, \mathbf{x}, \mathbf{x} + \mathbf{y}, \mathbf{y}$ の4点を頂点とする四角形が平行四辺形になることは図示によって確認できる。もちろん $n \geq 4$ の場合についても同様のことが成り立つ。我々は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。

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。

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 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ことがある。

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

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

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

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

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

例 1.6.2. 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 $f : \mathbb{R} \rightarrow \mathbb{R}$ である。

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

定義 1.6.3. 二つの写像 $f : X \rightarrow Y$ および $g : Y \rightarrow Z$ が与えられているとする。このとき、各 $x \in X$ に対して、 Z の元 $g(f(x))$ を対応させる写像を f と g の合成 (**composition**) と呼び、記号 $g \circ f : X \rightarrow Z$ で表す。すなわち、 $(g \circ f)(x) := g(f(x))$ である。

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。ところで、 $g \circ f$ で一つの写像を表しているから、誤解が無い限り $(g \circ f)$ と括弧つきで書く必要はない。そこで、以下 $(g \circ f)(x)$ を $g \circ f(x)$ と書く。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 $f : \mathbb{R} \rightarrow \mathbb{R}$ を比例関数という⁸。

例 1.6.5. 関数 $g : \mathbb{R} \rightarrow \mathbb{R}$ および $f : \mathbb{R} \rightarrow \mathbb{R}$ を $g(x) := ax$, $f(x) := bx$ と定めれば $g \circ f$ も比例関数であり、 $g \circ f(x) = (ab)x$ である。実際、 $g \circ f(x) = g(f(x)) = g(bx) = a(bx) = (ab)x$ 。

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

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対応している。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”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 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ を線形写像 (**linear map**) という:

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

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

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 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ を $f \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$ と定める。このとき次に答えよ。

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

解答例:

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

(2) f が線形写像であることを示せ.

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

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

ゆえに線形写像の性質 (i) は成立する.

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

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

ゆえに線形写像の性質 (ii) は成立する. □

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

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

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

さて, 線形代数では四つの数を次のような形にならべたものを用いる:

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

これらを 2×2 行列と呼ぶことにしよう. すると上の議論は, 線形写像と 2×2 行列が 1 対 1 に対応することを示唆している. すなわち, 行列 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ に対して線形写像 $f\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$ を対応させれば, これは 1 対 1 対応である. ここで, 前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 $g, f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ に対応する行列をそれぞれ A, B とすれば, これらの和 $g + f$ および合成 $g \circ f$ もまた線形写像であり, $g + f$ および $g \circ f$ に対応する行列はそれぞれ $A + B, AB$ である.

この対応は 2 変数に限らず, \mathbb{R}^n から \mathbb{R}^m への線形写像全体と $m \times n$ 行列全体の間で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.

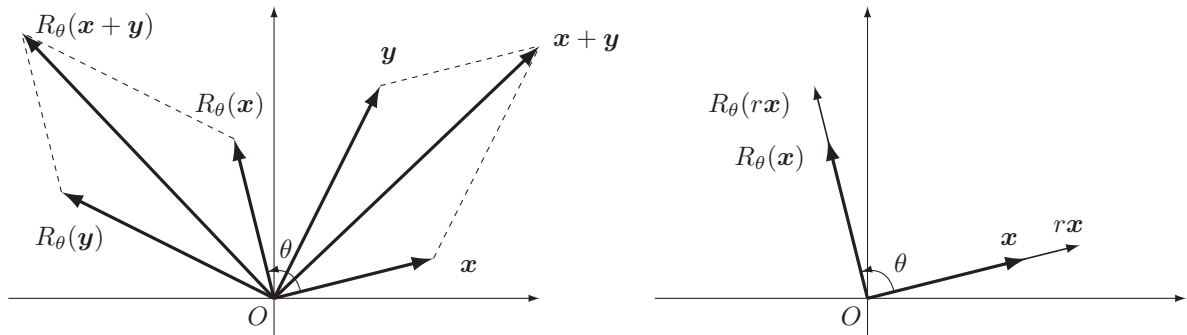
線形代数学で扱う行列とは線形写像をデータ化するために考えだされた概念である. そして, 抽象的な線形写像を行列によって数値化することで, 写像を視覚化するというのが行列理論のねらいである. 前期の講義では, 行列に関する計算を理解することを目標とする. また後期においては, 行列を駆使した線形写像の分析について学ぶ. 今後の講義において, しばらくのあいだ線形写像そのものは表立って出てはこないが, 行列理論の目的が線形写像の分析にあるということを念頭において学習してもらいたい.

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

よりみち (加法定理).

\mathbb{R}^2 の各ベクトル \mathbf{x} に対して, 原点 O を中心に \mathbf{x} を θ 回転させたベクトルを対応させる写像 $R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ を考える. これは明らかに線形写像である. 実際, 定義 1.7.1 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 $R_\theta(\mathbf{x} + \mathbf{y})$ は二つのベクトル $R_\theta(\mathbf{x}), R_\theta(\mathbf{y})$ で張られる平行四辺形の頂点となる. これは $R_\theta(\mathbf{x} + \mathbf{y}) = R_\theta(\mathbf{x}) + R_\theta(\mathbf{y})$ を意味している.
- (ii) 回転によってベクトルの長さが変化することはない. したがって \mathbf{x} と $r\mathbf{x}$ における長さの比と, これらを R_θ で写像した $R_\theta(\mathbf{x})$ と $R_\theta(r\mathbf{x})$ における長さの比は共に $1 : r$ である. 更に, \mathbf{x} と $r\mathbf{x}$ は平行ゆえ θ 回転後の $R_\theta(\mathbf{x})$ と $R_\theta(r\mathbf{x})$ も平行である. 以上のことから $R_\theta(r\mathbf{x}) = rR_\theta(\mathbf{x})$ が成り立つ.



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

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

さて, ベクトルを β 回転させた後に α 回転させることと, 一度に $\alpha + \beta$ 回転させることは同じ操作である. これは $R_\alpha \circ R_\beta = R_{\alpha+\beta}$ を意味する. $R_{\alpha+\beta}$ に対応する行列は $A_{\alpha+\beta} = \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}$ であり, $R_\alpha \circ R_\beta$ に対応する行列は $A_\alpha A_\beta$ ゆえ

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

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

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

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

を得る. 線形写像を知る者にとって加法定理は自明の理といえるだろう.

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:

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

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

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

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

一般の $m \times n$ 行列 A を成分表示すると次のようになる. A の (i, j) -成分を a_{ij} とするとき⁹,

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

毎回このような表示を用いては手間がかかるゆえ, 場合によっては, これを

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

と書く. 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:

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

A と B の和が定まるのは A, B のサイズが一致する場合のみである。

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

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

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

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

2.3 行列の積

何度も言うように、行列の演算とは線形写像のそれに対応するものである。 \mathbb{R}^2 を定義域とする線形写像のベクトル値は、 $ax + by$ という形の数をもつことを前節で見た。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 $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ の間の積 $A\mathbf{x}$ を、やや天下りの的では

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

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

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

例 2.3.1.

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

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

	大人	学生	幼児
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、

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

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

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ことと見やすく計算できる。左側の行列を行について分割し、右側の行列を列について分割している。

$$\begin{bmatrix} \hline 1 & 0 & 1 \\ 3 & 2 & 0 \\ \hline 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 3 & 1 \\ \hline 2 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 \cdot 0 + 0 \cdot 2 + 1 \cdot 0 & 1 \cdot 3 + 0 \cdot 0 + 1 \cdot 2 & 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 0 \\ 3 \cdot 0 + 2 \cdot 2 + 0 \cdot 0 & 3 \cdot 3 + 2 \cdot 0 + 0 \cdot 2 & 3 \cdot 1 + 2 \cdot 1 + 0 \cdot 0 \\ 1 \cdot 0 + 1 \cdot 2 + 2 \cdot 0 & 1 \cdot 3 + 1 \cdot 0 + 2 \cdot 2 & 1 \cdot 1 + 1 \cdot 1 + 2 \cdot 0 \end{bmatrix} \\ = \begin{bmatrix} 0 & 5 & 1 \\ 4 & 9 & 5 \\ 2 & 7 & 2 \end{bmatrix}.$$

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いよい(試験対策のためである)。

2.4 行列演算の性質

行列に関するいくつかの概念をここでまとめて定めておく。

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

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

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

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

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

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

□

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

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

□

性質 (5) および結合律 (18) の証明は, 次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.

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

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

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

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

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

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

発展 (代数構造).

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

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

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

他にも, まだまだ代数構造はたくさんある. 例えば, スカラー倍が環の中で定まっており, 命題 2.4.1 の性質 (9) から (15) が成立する (したがって命題 2.4.1 の性質すべてを満たす) 代数構造を多元環または代数 (**algebra**) という. 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, \dots, a_n たちの和 $a_1 + a_2 + \cdots + a_n$ を

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

と表す. 積 $a_1 a_2 \cdots a_n$ については

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

と表す. また, a_{ij} ($i = 1, \dots, m, j = 1, \dots, n$) たちの総和を $\sum_{\substack{i=1, \dots, m, \\ j=1, \dots, n}} a_{ij}$ と書く. こうした表記をより一般的な立場から眺めると, 次のような説明ができる.

P を, ある変数に関する条件とする. ここで変数は多変数としてもよい.

例 3.2.1. P の具体例に以下のようなものがある:

- $P(i)$: “ $i \in \mathbb{N}$ ”
- $P(i)$: “ $i = 1$ または $i = 2$ または \dots または $i = n$ ”
注: この条件文 $P(i)$ を略して, 我々は “ $i = 1, \dots, n$ ” と書いている.
- $P(i)$: “ $i \neq 1$ かつ $i \neq 2$ ”

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

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

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

- $P(i, j)$: “ $1 \leq i \leq j \leq 3$ ”
- $P(i, j)$: “ $1 \leq i \leq 5$ かつ $1 \leq j \leq 5$ ”

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

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

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

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

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

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

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

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

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

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

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

例 3.2.4.

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

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.

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

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

$(AB)C$ 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 $f: X \rightarrow Y, g: Y \rightarrow Z, h: Z \rightarrow W$ について, $(h \circ g) \circ f = h \circ (g \circ f)$.

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

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

次に $\psi = g \circ f$ とおくことで

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

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

線形写像の合成と行列の積が対応することの詳しい説明は, ユークリッド空間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.

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

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*と書く.

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

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

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

3.4 転置行列

(m, n) -行列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 $k = 1, \dots, n, h = 1, \dots, m$) を成分とする (n, m) -行列 $[b_{kh}]$ を A の転置行列 (**transposed matrix**) といい, 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,

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

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

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

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

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

転置行列の真の意味, すなわち線形写像としての意味をここで述べるのは難しい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 $A^t A$ で与えられる. 各列どうしの内積は, (n, n) -行列 ${}^t A A$ 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, ${}^t A$ の行と列を表す添え字 k, h を文字 i, j に置き換えれば, $[b_{ij}]_{\substack{i=1, \dots, n, \\ j=1, \dots, m}}$ となる. すなわち,

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

上の右辺と左辺では 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 $A = [a_{ij}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}}$ 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 $B = [a_{ji}]_{\substack{i=1, \dots, m, \\ j=1, \dots, n}}$ 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, $B = [a_{ji}]_{\substack{i=1, 2, \\ j=1, 2, 3}}$ と定めようとするれば, 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) ${}^t(A + B) = {}^t A + {}^t B$. (2) ${}^t(AB) = {}^t B {}^t A$.

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

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

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

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.

$${}^t B {}^t A = [b_{kj}]_{\substack{j=1, \dots, \ell, \\ k=1, \dots, n}} \cdot [a_{ji}]_{\substack{i=1, \dots, n, \\ j=1, \dots, m}} = [b_{\beta\alpha}]_{\substack{\alpha=1, \dots, \ell, \\ \beta=1, \dots, n}} \cdot [a_{\gamma\beta}]_{\substack{\beta=1, \dots, n, \\ \gamma=1, \dots, m}} \quad (\text{ここで添え字の文字を書き換えた})$$

$$= [y_{\alpha\beta}]_{\substack{\alpha=1, \dots, \ell, \\ \beta=1, \dots, n}} \cdot [x_{\beta\gamma}]_{\substack{\beta=1, \dots, n, \\ \gamma=1, \dots, m}} \quad (\text{ここで } y_{\alpha\beta} := b_{\beta\alpha}, x_{\beta\gamma} := a_{\gamma\beta} \text{ と置いた})$$

$$= \left[\sum_{\beta=1}^n y_{\alpha\beta} x_{\beta\gamma} \right]_{\substack{\alpha=1, \dots, \ell, \\ \gamma=1, \dots, m}} = \left[\sum_{\beta=1}^n b_{\beta\alpha} a_{\gamma\beta} \right]_{\substack{\alpha=1, \dots, \ell, \\ \gamma=1, \dots, m}} = \left[\sum_{\beta=1}^n a_{\gamma\beta} b_{\beta\alpha} \right]_{\substack{\alpha=1, \dots, \ell, \\ \gamma=1, \dots, m}}$$

$$= \left[\sum_{j=1}^n a_{kj} b_{ji} \right]_{\substack{i=1, \dots, \ell, \\ k=1, \dots, m}} \quad (\text{ここで再び添え字の文字を書き換えた}).$$

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.

一方で,

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

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

3.5 行列の分割

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

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

こうした分割による積の計算と元々の定義による積の計算が一致することは, 各成分がどのような成分の積たちの和になっているか調べれば分かる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.

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

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

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

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

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.

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

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

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

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。

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

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。

解の非存在証明。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 は任意の実数を取り得るため、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 $(x, y) = (1, 0)$ 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 $i = 1, \dots, m, j = 1, \dots, n$) 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, \dots, x_n に関する連立方程式

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

を解くことと、行列に関する次の関係式

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

を満たす n 次列ベクトル $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ を探すことは同じことである。式(4.2.1)あるいは式(4.2.2)の形

で与えられる方程式のことを連立1次方程式 (system of linear equations) という。

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

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

4.3 逆行列を持つ場合

まずは最も簡単な場合の解法を紹介する. それは A が正方行列であり, 更に $AB = BA = E$ を満たす正方行列 B が存在する場合である. このような B のことを A の逆行列 (**inverse**) という. このとき, 列ベクトル \mathbf{a} が方程式 $A\mathbf{x} = \mathbf{b}$ の解であると仮定すれば (すなわち $A\mathbf{a} = \mathbf{b}$ を満たす), この両辺に左から B をかけることで $BA\mathbf{a} = B\mathbf{b}$ を得る. 左辺を計算すると $BA\mathbf{a} = E\mathbf{a} = \mathbf{a}$ ゆえ $\mathbf{a} = B\mathbf{b}$ である. したがって方程式に解が存在するとすれば, その解は唯一解 $\mathbf{x} = B\mathbf{b}$ のみであることが分かった. また, 実際に $\mathbf{x} = B\mathbf{b}$ はこの方程式の解である. 何故なら, $A\mathbf{x} = AB\mathbf{b} = E\mathbf{b} = \mathbf{b}$ ゆえ $A\mathbf{x} = \mathbf{b}$ を満たすからである. 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 $A\mathbf{x} = \mathbf{b}$ は唯一解 $\mathbf{x} = B\mathbf{b}$ 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 $A\mathbf{x} = \mathbf{b}$ は, 変数の数が n 個, 式の数 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ることができるからである. この 3 つの操作に対応する行列の変形操作を行基本変形 (**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 \mathbf{b} を並べた, $(m, n + 1)$ -行列 $[A|\mathbf{b}]$ を考えよう.

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

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

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

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

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

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

- (I) 零ベクトルとなる行は零ベクトルでない行よりも下段にある,
- (II) 零ベクトルでない行の主成分は 1 である.
- (III) 第 i 行の主成分が第 j_i 列にあるとすれば, $j_1 < j_2 < j_3 < \cdots$ 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 を簡約化るといい, このとき B を A の簡約化と呼ぶ. 簡約化の方法は次の手順で得られる. ただし, これはあくまでも形式的なアルゴリズムであり, 実際に簡約化を行う際は, 各自の判断で手順を入れ替えたほうが効率が良い場合もある.

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

例 4.5.3. 次は簡約行列への行基本変形の一例である. 実際の変形においては, 2 目と 3 目の変形を, 4 目と 5 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 $[B|\mathbf{b}]$ における連立1次方程式 $B\mathbf{x} = \mathbf{b}$ の解法を具体例を通して説明する. $[B|\mathbf{b}]$ の主成分に関するある条件について, 二つのケースに分けて考える.

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

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

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

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

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

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

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

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

を移項することで次を満たさねばならないことが分かる.

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

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

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

逆に, 今度は実数 c_1, c_2, c_3 を勝手に選び, 上の式によって \mathbf{x} を定めてみよう. このとき \mathbf{x} が方程式 $B\mathbf{x} = \mathbf{b}$ を満たすことは明らかである. 何故なら, 上でもって \mathbf{x} を定めるということは, 式 $(\#2)$ が満たされるように定めたことを意味し, 式 $(\#2)$ を移項すれば元々の連立方程式 $(\#1)$ が得られるからである. これまでの議論により, 方程式 $B\mathbf{x} = \mathbf{b}$ の解は存在し, しかも解の形は次のもので出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も後々の議論において都合が良い。

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

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

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

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

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

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

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

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

たち全体のなす集合 H (これは \mathbb{R}^5 の原点を通る3次元の空間である) を、 $\begin{bmatrix} 0 \\ 4 \\ 0 \\ 5 \\ 0 \end{bmatrix}$ 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 $Ax = b$ の解法は次のとおりである. まず拡大係数行列 $[A|b]$ の簡約化 $[B|b']$ を求める. このとき, $Ax = b$ の解全体と $Bx = b'$ の解全体は一致する. $[B|b']$ の最後の列 b' に主成分がある場合は方程式の解は存在しない. そうでない場合は解は存在し, 方程式 $Bx = b'$ の解は, B における主成分を含まない列に対応する変数を任意定数とおくことで簡単に表せる. なお, B のすべての列に主成分がある場合, 任意定数の数は0で, 方程式の解は唯一つである. また, 方程式の解をベクトルの和に分解しておくこと, 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 $Ax = b$ の解法の一つ (はきだし法) を与えた. ここで, 式 (4.6.1) の形で与えられる $Ax = b$ の解の表し方にどれくらいの種類があるか検討しよう. 例えば, 次の方程式

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

の解として, 次の四つの形が挙げられる:

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

ただし, a, b, c, d, e はそれぞれ任意定数である. (1) は本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らない. それまでの間の当座の約束として、「連立 1 次方程式の解における任意定数の個数」とは本節で与えた解法 (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 26.2.3 にて述べる。

命題 5.1.2. A の逆行列が存在すれば、それは唯一つである。

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

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

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

例 5.1.3. 逆行列に関する簡単な性質をここでまとめておく。

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

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

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

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

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。

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

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

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

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

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

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

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

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

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

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

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

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

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

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

命題 5.2.3. 基本行列の逆行列は基本行列であり、

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

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

(2) および (3) も同様の考察から分かる。 □

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

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

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

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

以上を整理すると次のような主張になる。

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

(1) $PA = B$ を満たす m 次可逆行列 P が存在する。

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

(2) B を行基本変形することにより A に戻すこともできる。

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

5.3 逆行列の求め方

n 次正方行列 A の逆行列を求めるために、まずはその候補として $BA = E$ を満たす正方行列 B を探そう。ここで、行列 A が k 回の行基本変形によって E に変形できたと仮定しよう。すなわち、 A の簡約化は E であり、また前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 $\mathbf{0}$ に限る。 $\boldsymbol{x} = \mathbf{0}$ の場合、 $\mathbf{0}$ に基本変形をいくらほどこしても変化せず、 $\mathbf{0}$ の簡約化は $\mathbf{0}$ 自身以外にありえない。また、 $\boldsymbol{x} \neq \mathbf{0}$ の場合、 \boldsymbol{x} を $\mathbf{0}$ に行基本変形することはできない。何故なら、もし $\mathbf{0}$ が \boldsymbol{x} の簡約化であるならば、命題5.2.4により $\mathbf{0}$ を行基本変形して $\boldsymbol{x} \neq \mathbf{0}$ が得られるが、先の議論によりこれは不可能だからである。ゆえに \boldsymbol{x} の簡約化は \boldsymbol{e}_1 のみである。以上より、列ベクトルの簡約化は唯一つであることが分かった。

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

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

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

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

更に次の二つに場合分けをして考える。

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

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

(ii) (i) でない場合。すなわち、 $\boldsymbol{b}, \boldsymbol{c}$ 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。

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

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

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

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. $X[B|\mathbf{b}]$ を計算すると次のようになる.

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

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

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

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

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

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

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

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

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

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

(m, n) -行列 A について次は明らかである.

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

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

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

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

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

ゆえに簡約行列 B は C の簡約化である. □

6.2 連立1次方程式と階数

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

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

- (1) $Ax = b$ の解が存在するための必要十分条件は $\text{rank}[A|b] = \text{rank} A$ である。
- (2) $Ax = b$ の解が存在するとき、解の任意定数の個数は $n - \text{rank} A$ である。
- (3) $Ax = b$ の解が唯一つであるための必要十分条件は、 $\text{rank}[A|b] = \text{rank} A = n$ である。

とくに、 A が正方行列の場合は逆行列との関係を含めて次の主張を得る。

定理 6.2.2. n 次正方行列 A において次は同値である。

- (1) $\text{rank} A = n$, (2) A の簡約化は E である,
- (3) 任意の n 次列ベクトル b について方程式 $Ax = b$ は唯一つの解を持つ,
- (4) 方程式 $Ax = 0$ は唯一つの解 0 を持つ, (5) A は可逆である。

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

(1) \Rightarrow (2) : $\text{rank} A = n$ とすれば、 A の簡約化 B の主成分の個数は n である。 B は $n \times n$ 行列であるゆえ、 $B = E$ となる。

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

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

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

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

(5) \Rightarrow (4) : これは命題 4.3.1 の特別な場合に相当する。

$x = 0$ が方程式 $Ax = 0$ 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、 $Ax = 0$ の両辺に左から A^{-1} を掛けることで $x = 0$ を得る。

以上によって、すべての条件の同値性が示された。 □

6.3 同次形の方程式

(m, n) -行列 A において、 $Ax = 0$ 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 $x = 0$ 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 $Ax = 0$ の拡大係数行列 $[A|0]$ の簡約化は $[B|0]$ であるから $\text{rank}[A|0] = \text{rank}[B|0] = \text{rank} B = \text{rank} A$ 。確かに $\text{rank}[A|0] = \text{rank} A$ 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 $k = n - \text{rank } A$ である. 上の練習にもあるように拡大係数行列 $[A|\mathbf{0}]$ の簡約化は $[B|\mathbf{0}]$ なる形になり, ゆえに式 (4.6.1) において $\mathbf{a}_0 = \mathbf{0}$ となる. したがって解の表示は

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

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

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 $A\mathbf{x} = \mathbf{b}$ の解 \mathbf{a} を一つ取って固定しよう. このとき次が成り立つ.

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

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

(2) : \mathbf{y} を $A\mathbf{x} = \mathbf{b}$ の解とする. ここで $\mathbf{z} := \mathbf{y} - \mathbf{a}$ とおこう. このとき \mathbf{z} は $A\mathbf{x} = \mathbf{0}$ の解である. 何故なら, $A\mathbf{z} = A(\mathbf{y} - \mathbf{a}) = A\mathbf{y} - A\mathbf{a} = \mathbf{b} - \mathbf{b} = \mathbf{0}$ ゆえ \mathbf{z} は $A\mathbf{x} = \mathbf{0}$ を満たすからである. また, \mathbf{z} の定め方から $\mathbf{y} = \mathbf{a} + \mathbf{z}$ であり, 我々は主張を得た. \square

上の命題は, 同次形の方程式の解全体の集合を H とすれば, $A\mathbf{x} = \mathbf{b}$ の解全体の集合 W は H を \mathbf{a} 方向に平行移動した集合に一致することを言っている. この事実の特別な場合については 4.6 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 $A\mathbf{x} = \mathbf{0}$ の解法は, 方程式 $A\mathbf{x} = \mathbf{b}$ 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 $A\mathbf{x} = \mathbf{0}$ の解を求めておき, 更に, 何らかの方法で $A\mathbf{x} = \mathbf{b}$ の解 \mathbf{a} を一つでよいから見つけてくる. すると, $A\mathbf{x} = \mathbf{b}$ の解は, $A\mathbf{x} = \mathbf{0}$ の解全体を \mathbf{a} 方向へ平行移動することですべて得られる.

しかし, 既に連立 1 次方程式の解法を知っている我々にとっては, このような考え方は不要にも思える. そこで, 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.

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

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,

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

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,

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

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 の証明が平行であることに着目せよ. これはベクトル \mathbf{x} に対して $A\mathbf{x}$ 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.

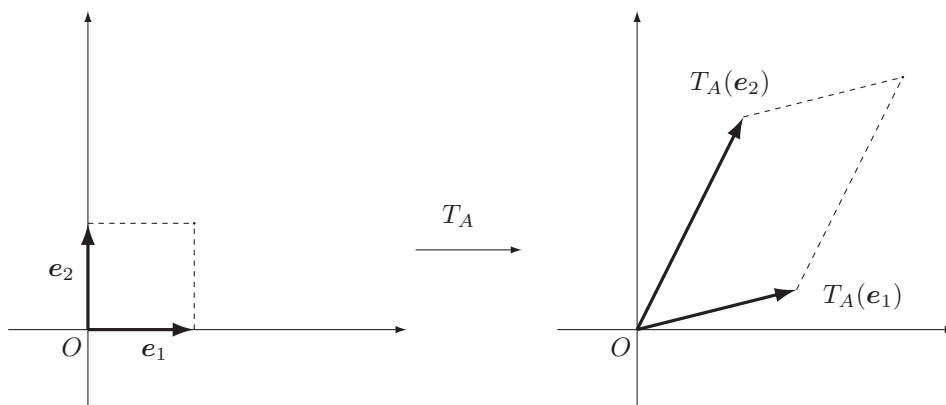
7 行列式とは何か

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 $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ を $T_A(\mathbf{x}) := A\mathbf{x}$ 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。



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

単位正方形とは $\mathbf{e}_1, \mathbf{e}_2$ で張られる平行四辺形のことであり、これを T_A で写像すると $T_A(\mathbf{e}_1)$ と $T_A(\mathbf{e}_2)$ で張られる平行四辺形になる。ここで、積分において負の面積を考えたように、平行四辺形についても向きを含めた面積を導入する。面積の符号は、半周未満の回転によってベクトル $T_A(\mathbf{e}_1)$ を $T_A(\mathbf{e}_2)$ に重ねるとき、その回転が反時計回りになるか時計回りになるかで判断すればよい。 $\mathbf{e}_1, \mathbf{e}_2$ について同じことを考えればこの回転は反時計回りであるから、 $T_A(\mathbf{e}_1), T_A(\mathbf{e}_2)$ についても反時計回りならば正、そうでなければ負の面積を対応させる。図 1 の平行四辺形の面積は、原点と $T_A(\mathbf{e}_1) + T_A(\mathbf{e}_2)$ を頂点に持つ長方形の面積からいくつかの台形や三角形の面積を引くことで得られるから、

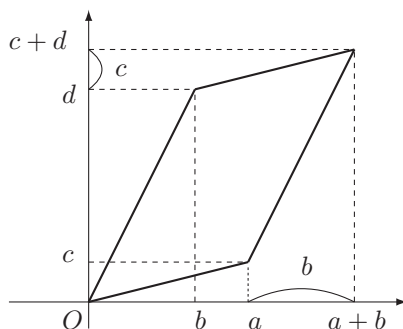


図 1: 平行四辺形の各頂点の座標表示

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

一般の n 次正方行列 A の行列式も同様に与えられる. すなわち, 線形変換 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ を $T_A(\mathbf{x}) := A\mathbf{x}$ で定め, T_A の体積拡大率を $\det A$ と定める. ここで, 1次元において長さ, 2次元において面積, 3次元において体積とそれぞれ異なる呼ばれ方をしていた各次元に関する量は, 4次元以上においてはすべて体積と呼ぶことにする. また n 次元空間において, 次元が n 未満の図形を面と呼ぶことにしよう. 3次元空間における2次元の図形が通常の意味での面であった. $n=3$ の場合 $\det A$ は, A の成分表示に現れる3つの列ベクトルで張られる平行六面体の体積に相当し, 図形の体積計算が得意な者ならば次式を得ることができるかもしれない¹⁹:

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

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

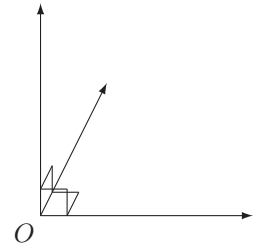
さて, 行列式が体積拡大率を意味するのであれば, T_E は元を動かさない恒等写像ゆえ $|E| = 1$ である. また, $|AB|$ は合成写像 $T_A \circ T_B$ の体積拡大率であるから, これは $|A| \cdot |B|$ に等しい. つまり $|AB| = |A||B|$ となる. 更に A が可逆であるとき, $|A||A^{-1}| = |AA^{-1}| = |E| = 1$ より $|A^{-1}| = |A|^{-1}$, とくに $|A| \neq 0$ である. 逆に, $|A| \neq 0$ ならば A が可逆であることも後に示され, したがって $|A|$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

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、

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

ここで、上式の最後の (1, 3)-成分は

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

以上より、行列式の言葉で方程式の解を表すと次のようになる:

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

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

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

方程式 $A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ 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 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ 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):

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

の値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

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

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

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

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

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

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

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

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

よりみち (行列と行列式).

行列は matrix の訳語であり行列式は determinant の訳語である。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な置換について述べる。置換とはその名の通り置き換え、あるいは入れ替え方を意味する。線形代数においては行列式の定義以外に置換が現れることは稀であるものの、置換は数学を記述する言葉として重要である。とくに、群を説明する道具として初等的役割を担う。

置換に関する話題は組み合わせ論的な色彩が強く、苦手意識を持つ読者も多いように思う。しかしながら本論において置換は、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 $X_n := \{1, 2, \dots, n\}$ とする。

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

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

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

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

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

8.2 置換の表示

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

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

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

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

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

置換の表示を扱いやすくするために、次の表示の仕方を許すとする。

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

以上二つの表示の仕方を認めると、次で表示された置換はすべて同じ写像を表す：

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

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

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

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

8.3 置換の積

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

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

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

となるから、 $\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ である。なお、 σ の上段が τ の下段と同じ並びになるようあらかじめ並び変えておくと、合成の表示が直ちに得られる。つまり $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix} =$

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

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

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

例 8.3.2. $\sigma = \begin{pmatrix} 2 & 4 & 9 \\ 4 & 9 & 2 \end{pmatrix}, \tau = \begin{pmatrix} 3 & 6 & 7 & 8 \\ 8 & 3 & 6 & 7 \end{pmatrix}$ のとき $\sigma\tau = \tau\sigma$ である。このように、二つの置換 σ, τ の表示において、同じ文字が一つも重複しないとき $\sigma\tau = \tau\sigma$ が成り立つ。

命題 3.2.5 の特別な場合として次が成り立つ:

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

次の二つの概念は, 本来は写像に関する一般論において定められるものである. ここでは置換の積演算と関連して, 次の呼称を与える (一般の写像については 19 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, $\sigma\sigma^{-1} = \text{id} = \sigma^{-1}\sigma$ もすぐに分かる.

同じ置換を繰り返し合成するときは冪を用いる:

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

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

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

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

8.4 巡回置換とその表示

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

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

例 8.4.2. • 次の巡回置換はすべて同じ置換を意味する:

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

- 上の置換に各元を代入した値は次の通りである:

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

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

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

答え. X_5 の各元を代入して確認する.

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

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

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

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

Proof. $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 1 & 6 & 2 & 7 & 5 & 3 \end{pmatrix}$ において確かめよう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,

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

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

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

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

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

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

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

命題 8.4.4 および 8.4.5 より直ちに次を得る:

系 8.4.6. 任意の置換は互換の積に分解される.

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

答え.

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

8.5 置換の符号

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

置換の偶奇性のことをパリティ (parity) とも呼ぶこともある.

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

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

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

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

したがって, σ が奇数個の互換の積で表せ, 更に偶数個の互換の積でも表せるとなると, σ の符号を定めることはできない. しかしながら, このようなことは起こらず, 置換を互換の積で表した際の互換の個数の偶奇は必ず一致することが知られている. この事実の証明にはいくつか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 $\text{sgn}(\sigma\tau) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau)$.

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

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

例 8.5.4.

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

練習 8.5.5. 次を示せ.

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

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

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

(\Leftarrow): 逆置換の定義より明らかである. □

- (2) 置換 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:

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

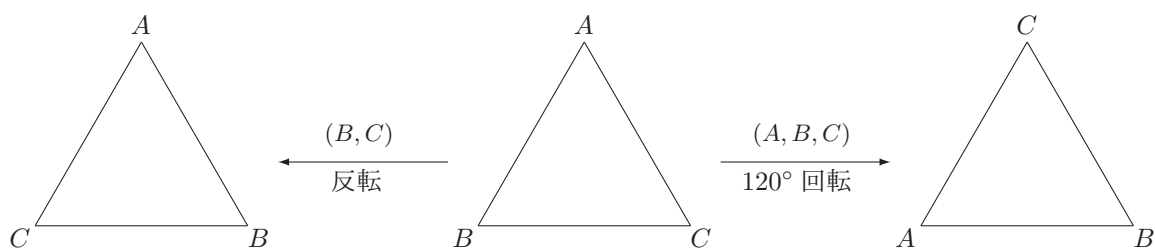
ゆえに (1) より $\tau = \sigma^{-1}$ である. □

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 倍によってお互いに入れ替わる。また、式 $x^2 + xy + y^2 + zw$ においては、式の意味が変わらないような変数の入れ替えを考えていた。ジャンケンにおいても同様である。命題「 A は B に勝ち、 B は C に勝ち、 C は A に勝つ」が成り立つような A, B, C にそれぞれジャンケンの手を対応させる。このとき、この命題が成立し続けるような A, B, C の入れ替えがジャンケンの対称性を記述する。なお、対称性とは入れ替えの総数のみを考えるのではなく、入れ替えの相互関係についても考慮する概念である。

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

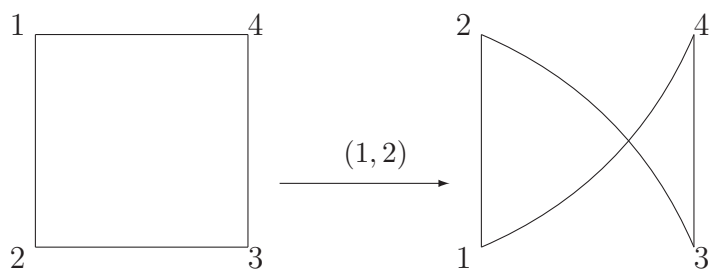


他に考えられる操作は 240° 回転に相当する巡回置換 (A, C, B) , 360° 回転 (結果的にこれは頂点を全く動かさない操作 id に等しい), B, C のいずれかを固定した反転に相当する互換 (A, C) および (A, B) であり、次の計 6 つである。

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

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

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



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

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

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

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

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

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

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

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読者が迷わずに済むための配慮である。もう一つは、数学を応用する立場から見ても文字列の並び替えは基本的な考え方・道具となりうるからである。

9.1 符号の正当性

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

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

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

$s(\sigma)$ が符号を意味することをこれから見ていく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.

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

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

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

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

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

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

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

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

↑ 変数を置き換えた

□

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

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

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

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

上の定理の証明から直ちに次を得る.

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

9.2 文字列の並び替え (よりみち)

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 (ii) における記号 $\xleftrightarrow{(i,j)}$ は i 列と j 列を入れ替える互換を指していることに注意する. これら各々の列の入れ替えの合成が σ であり, ゆえに次を得る:

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

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

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

9.3 転倒数による符号の定義 (よりみち)

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

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

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

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

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

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

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

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

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

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

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

(i) k_i について。

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

(ii) k_j について

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

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

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

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

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

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

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

□

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

最後に, 命題 8.5.3 に対応する性質は次のように示される.

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

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

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

□

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つてほしい。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. $\mathfrak{S}_1 = \{\operatorname{id}\}$ および $\operatorname{sgn}(\operatorname{id}) = 1$ より、

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

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

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

(3) $n = 3$ の場合. 例 8.5.4 より

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

である。したがって、

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

本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、 $\det {}^tA = \det A$ 。

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。

$$\begin{array}{ccc} a_{11} & a_{12} & \\ a_{21} & a_{22} & \end{array} \begin{array}{l} (-) \\ (+) \end{array} = a_{11}a_{22} - a_{12}a_{21}$$

3 次の場合はやや複雑になるが、次のように分けよう。

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

上の図式をもとに次の展開式を得る:

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

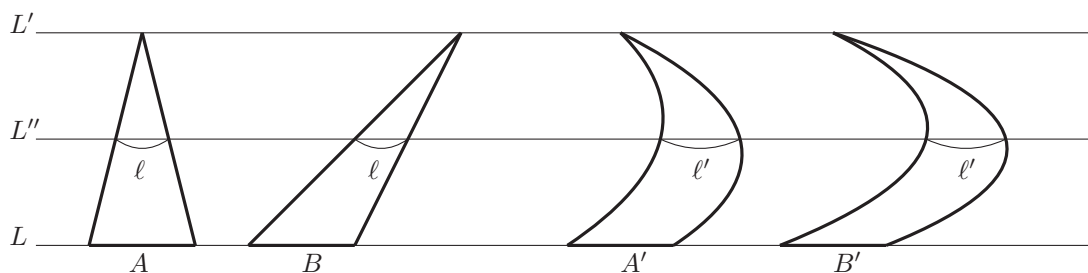
なお、4 次以上の行列式には、この様な方法は使えない。

10.2 カヴァリエリの原理

行列式の性質と図形の体積との関係を考察するにあたり、本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ふものの、この主張を正確に述べておこう。

カヴァリエリの原理.(Cavalieri's principle)

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



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

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

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

よりみち (体積とは何か)

いかなる図形に対してカヴァリエリの原理が成立するのだろうか. 読者の中にはこのような素朴な疑問を持つ者もいるかと思う. しかし残念なことに, 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 $f, g: [a, b] \rightarrow \mathbb{R}$ で挟まれた図形 A の面積は, 積分 $\int_a^b (f(x) - g(x)) dx$ によって得られる. すなわち, 図形 A の面積は, $h(x) := f(x) - g(x)$ のグラフと x 軸で挟まれた図形 B の面積に等しい. ここで, y 軸と平行な直線 $x = t$ (ただし $t \in [a, b]$) で A, B を切ってみよう. すると, それら線分の長さは共に $h(t)$ である. すなわち, 図形 A, B に対してカヴァリエリの原理が成立している. 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:

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

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

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

□

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

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

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

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

Proof. (1) は次より得る:

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

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

□

例 10.3.3.

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

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

正方行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ の各々の列ベクトルで張られる図形の符号付き体積に行列式 $|A|$ が一致することを認めたいので、多重線形性の幾何的な意味を $n = 2$ の場合について考察しよう。一つのベクトル

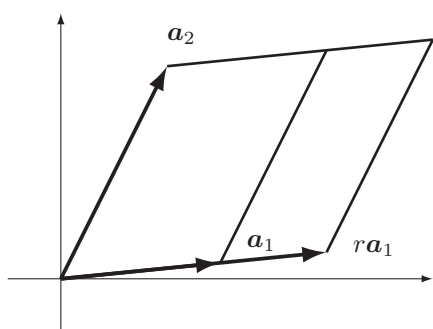


図 2: ベクトルのスカラー倍

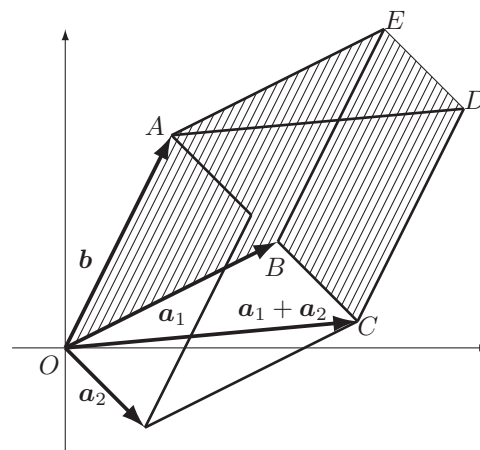


図 3: ベクトルの和

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. $\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}$ で張られる平行四辺形 AOC と六角形 $AOBCDE$ の面積は等しい. これは、 \mathbf{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 \mathbf{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 $\det(\mathbf{a}_1, \mathbf{b})$ と平行四辺形 $EBCD$ の面積の和に等しい. 平行四辺形 $EBCD$ は \mathbf{a}_2, \mathbf{b} で張られる平行四辺形を \mathbf{a}_1 方向に平行移動したものであるから、その面積は $\det(\mathbf{a}_2, \mathbf{b})$ である. 以上より、 $\det(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{b}) = \det(\mathbf{a}_1, \mathbf{b}) + \det(\mathbf{a}_2, \mathbf{b})$ 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 \mathbf{x}_k ($k = 1, \dots, n$) を n 次列ベクトルとすると、 $i < j$ について、

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

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

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

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

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

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

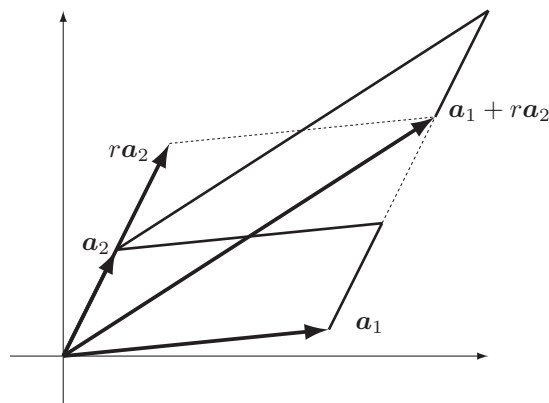
$$(2) \det {}^t[\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j + r\mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n] = \det {}^t[\mathbf{x}_1, \dots, \mathbf{x}_n].$$

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

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

行の場合についても同様の考察を行えばよい。□

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



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.

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

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

まずは特別な場合を考え,

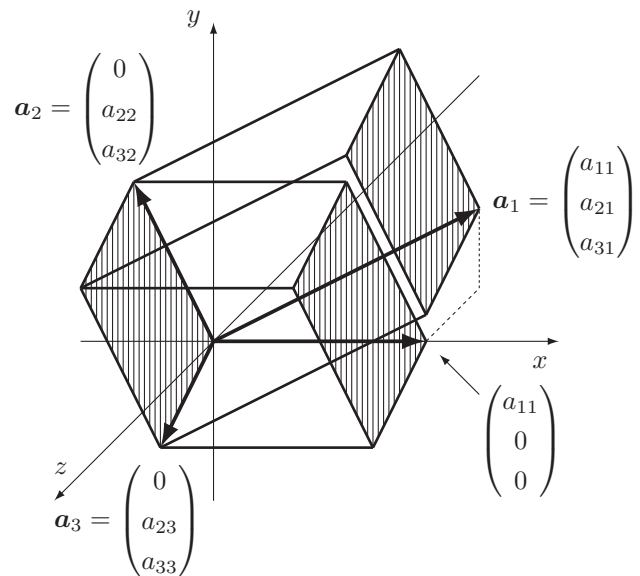
$$\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, \mathbf{a}_n = \begin{bmatrix} 0 \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} \quad \text{としよう.}$$

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

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

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

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



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:

$$\begin{vmatrix} a_1 & & & \\ & a_2 & * & \\ & & \ddots & \\ & 0 & & a_n \end{vmatrix} = a_1 \begin{vmatrix} a_2 & & * \\ & \ddots & \\ & & a_n \end{vmatrix} = \cdots = a_1 a_2 \cdots a_n.$$

とくに, $|E| = 1$.

11.2 計算例

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

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

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

(2)

$$\begin{aligned} \begin{vmatrix} 3 & 2 & 8 & 4 \\ 2 & 3 & 5 & 7 \\ 1 & 9 & 8 & 0 \\ 0 & 8 & 1 & 5 \end{vmatrix} &= \begin{vmatrix} 0 & -25 & -16 & 4 \\ 0 & -15 & -11 & 7 \\ 1 & 9 & 8 & 0 \\ 0 & 8 & 1 & 5 \end{vmatrix} = (-1)^2 \begin{vmatrix} 1 & 9 & 8 & 0 \\ 0 & -25 & -16 & 4 \\ 0 & -15 & -11 & 7 \\ 0 & 8 & 1 & 5 \end{vmatrix} = \begin{vmatrix} -25 & -16 & 4 \\ -15 & -11 & 7 \\ 8 & 1 & 5 \end{vmatrix} \\ &= (-1) \begin{vmatrix} -16 & -25 & 4 \\ -11 & -15 & 7 \\ 1 & 8 & 5 \end{vmatrix} = -(-1)^2 \begin{vmatrix} 1 & 8 & 5 \\ -16 & -25 & 4 \\ -11 & -15 & 7 \end{vmatrix} = - \begin{vmatrix} 1 & 8 & 5 \\ 0 & 103 & 84 \\ 0 & 73 & 62 \end{vmatrix} \\ &\uparrow 1 \text{ 列目と } 2 \text{ 列目を入れ替えた} \\ &= - \begin{vmatrix} 103 & 84 \\ 73 & 62 \end{vmatrix} = -2 \begin{vmatrix} 103 & 42 \\ 73 & 31 \end{vmatrix} = -2 \begin{vmatrix} 30 & 11 \\ 73 & 31 \end{vmatrix} = -2 \begin{vmatrix} 30 & 11 \\ 13 & 9 \end{vmatrix} = -2 \begin{vmatrix} 17 & 2 \\ 13 & 9 \end{vmatrix} \\ &= -2(17 \cdot 9 - 2 \cdot 13) = -2(153 - 26) = -2 \cdot 127 = -254. \end{aligned}$$

(3)

$$\begin{aligned}
& \begin{vmatrix} 3 & 1 & 4 & 1 & 5 \\ 9 & 2 & 6 & 5 & 3 \\ 5 & 8 & 9 & 7 & 9 \\ 3 & 2 & 3 & 8 & 4 \\ 6 & 2 & 6 & 4 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 3 & 2 & -2 & 3 & -7 \\ -19 & 8 & -23 & -1 & -31 \\ -3 & 2 & -5 & 6 & -6 \\ 0 & 2 & -2 & 2 & -7 \end{vmatrix} = (-1) \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 3 & -2 & 3 & -7 \\ 8 & -19 & -23 & -1 & -31 \\ 2 & -3 & -5 & 6 & -6 \\ 2 & 0 & -2 & 2 & -7 \end{vmatrix} \\
& = -(-1)^2 \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 3 & 2 & 3 & 7 \\ 8 & -19 & 23 & -1 & 31 \\ 2 & -3 & 5 & 6 & 6 \\ 2 & 0 & 2 & 2 & 7 \end{vmatrix} = - \begin{vmatrix} 3 & 2 & 3 & 7 \\ -19 & 23 & -1 & 31 \\ -3 & 5 & 6 & 6 \\ 0 & 2 & 2 & 7 \end{vmatrix} = - \begin{vmatrix} 3 & 2 & 3 & 7 \\ -1 & 35 & 17 & 73 \\ 0 & 7 & 9 & 13 \\ 0 & 2 & 2 & 7 \end{vmatrix} \\
& = - \begin{vmatrix} 0 & 107 & 54 & 226 \\ -1 & 35 & 17 & 73 \\ 0 & 7 & 9 & 13 \\ 0 & 2 & 2 & 7 \end{vmatrix} = -(-1) \begin{vmatrix} -1 & 35 & 17 & 73 \\ 0 & 107 & 54 & 226 \\ 0 & 7 & 9 & 13 \\ 0 & 2 & 2 & 7 \end{vmatrix} = (-1) \begin{vmatrix} 107 & 54 & 226 \\ 7 & 9 & 13 \\ 2 & 2 & 7 \end{vmatrix} \\
& = - \begin{vmatrix} 57 & 4 & 51 \\ 1 & 3 & -8 \\ 2 & 2 & 7 \end{vmatrix} = - \begin{vmatrix} 0 & -167 & 507 \\ 1 & 3 & -8 \\ 0 & -4 & 23 \end{vmatrix} = -(-1) \begin{vmatrix} 1 & 3 & -8 \\ 0 & -167 & 507 \\ 0 & -4 & 23 \end{vmatrix} \\
& = \begin{vmatrix} -167 & 507 \\ -4 & 23 \end{vmatrix} = (-1) \begin{vmatrix} 167 & 507 \\ 4 & 23 \end{vmatrix} = - \begin{vmatrix} 167 & 6 \\ 4 & 11 \end{vmatrix} \\
& = -(167 \cdot 11 - 6 \cdot 4) = -(1837 - 24) = -1813.
\end{aligned}$$

定理 11.2.2 (ヴァンデルモンドの行列式).

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

Proof. $|A| = |{}^t A|$ より一つ目の等式は明らか. 二つ目の等式を行列のサイズに関する帰納法で示す. 2次の場合は $\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix} = x_2 - x_1$ 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.

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

$$\begin{aligned}
&= \begin{bmatrix} x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ x_2(x_2 - x_1) & x_3(x_3 - x_1) & \cdots & x_n(x_n - x_1) \\ \vdots & \vdots & \vdots & \vdots \\ x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{bmatrix} \\
&= (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{bmatrix} \\
&= (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) \prod_{2 \leq i < j \leq n} (x_j - x_i) \quad (\text{ここで帰納法の仮定を用いた}) \\
&= \prod_{1 \leq i < j \leq n} (x_j - x_i).
\end{aligned}$$

□

11.3 体積との関係

前節と本節を通して、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 $1^n = 1$ であると考えよう。以上を満たす写像が行列式以外にないことを次の定理は主張している。すなわち、行列式とは、 n 個の列ベクトルで張られる図形の体積を符号付きで表す量のことである。

定理 11.3.1. 実数を成分に持つ n 次正方行列全体の集合を $M_n(\mathbb{R})$ とする²⁴。次の性質すべてを満たす写像 $D: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ を取れば、 D は行列式に一致する:

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

(1) 多重線形性:

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

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

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

(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:

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

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

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

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, $\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$.

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

補足. 上の補題は練習 8.5.5(2) から導くことができる.

補題 12.1.2. 写像 $I: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ を $I(\sigma) := \sigma^{-1}$ と定めれば, 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. \square

以上の補題を用いて定理 10.1.3 を証明しよう.

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

Proof. $A = [a_{ij}]$, ${}^tA = [b_{ij}]$ とおけば, 各 i, j について $b_{ij} = a_{ji}$ である. 行列式の定義および $\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$ (補題 12.1.1) から,

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

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

右辺の総和の各項に現れる積 $a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$ の並べ替えについて考えよう. まず $\sigma(1), \dots, \sigma(n)$ を小さい順に並べ替える:

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

つまり,

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

となり, ゆえに各 $i = 1, \dots, n$ について $\sigma(k_i) = i$ および $\sigma^{-1}(i) = k_i$ である. つまり $a_{\sigma(k_i)k_i} = a_{i\sigma^{-1}(i)}$ であり, 次のように並び替えができる:

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

写像 $I: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ を補題 12.1.2 で与えたものとし, 式 12.1.2 および 12.1.3 を合わせると,

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

σ が重複なく \mathfrak{S}_n 全体を動くとき, $\tau = I(\sigma)$ も重複なく \mathfrak{S}_n 全体を動く (補題 12.1.2). ゆえに, 上式の σ による総和は次の τ による総和に書き換えても良い:

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

□

12.2 歪対称性の証明

補題 12.2.1. あらかじめ置換 $\tau \in \mathfrak{S}_n$ を一つ取り, 固定しておく. このとき, 写像 $F: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ を $F(\sigma) := \sigma\tau$ と定めれば, F は 1 対 1 である. すなわち, σ が重複なく \mathfrak{S}_n 全体を動くとき, $F(\sigma)$ も重複なく \mathfrak{S}_n 全体を動く.

Proof. 補題 12.1.2 の証明と同様に, 単射性および全射性を示そう.

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

(単射性): 対偶である $F(\sigma_1) = F(\sigma_2) \Rightarrow \sigma_1 = \sigma_2$ を示す. $F(\sigma_1) = F(\sigma_2)$ とすれば F の定義より $\sigma_1\tau = \sigma_2\tau$ である. この両辺に右から τ^{-1} を掛けると

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

(全射性): 各 $\varphi \in \mathfrak{S}_n$ に対して, $\sigma := \varphi\tau^{-1}$ と定める. このとき, $F(\sigma) = F(\varphi\tau^{-1}) = \varphi\tau\tau^{-1} = \varphi \operatorname{id} = \varphi$. □

上の補題において置換 $\tau = (p, q)$ (ただし $p \neq q$) を考えると, $F: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ は偶置換を奇置換に写し, 奇置換を偶置換に写す写像となる. よって F により, 偶置換と奇置換の間に 1 対 1 の対応が与えられ, これらの総数が等しいことが分かる. すなわち:

命題 12.2.2. $n \geq 2$ について交代群 A_n の元の総数は $\frac{n!}{2}$ である.

行列式の歪対称性を証明しよう:

命題 10.4.1 (再掲). 列の入れ替え (または行の入れ替え) を行うと行列式は -1 倍される.

Proof. $p < q$ とし, p 行と q 行の入れ替えについて証明する. 正方行列 $A = [a_{ij}]$ の p 行と q 行を入れ替えた行列を $B = [b_{ij}]$ とすれば, 各 $j = 1, \dots, n$ について

$$(1) i \neq p, q \Rightarrow b_{ij} = a_{ij}, \quad (2) b_{pj} = a_{qj}, \quad (3) b_{qj} = a_{pj},$$

が成り立っている. $\tau = (p, q)$ とし, $F: \mathfrak{S}_n \rightarrow \mathfrak{S}_n$ を補題 12.2.1 で与えた写像とせよ. すなわち $F(\sigma) := \sigma\tau$ である. このとき,

$$(i) k \neq p, q \Rightarrow F(\sigma)(k) = \sigma(k), \quad (ii) F(\sigma)(p) = \sigma(q), \quad (iii) F(\sigma)(q) = \sigma(p).$$

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

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

↑ 各項において, 積の並び順を交換した (p 番目と q 番目を入れ替えた).

σ が重複なく \mathfrak{S}_n 全体を動くとき, $\varphi = F(\sigma)$ も重複なく \mathfrak{S}_n 全体を動く (補題 12.2.1). ゆえに, 上の最後の行にある σ による総和は次の φ による総和に書き換えても良い:

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

列の入れ替えについては, 定理 10.1.3 を用いて行の入れ替えに帰着させれば容易に示される. \square

12.3 多重線形性と歪対称性から導かれる性質

行列式の特徴づけ (定理 11.3.1) の証明を述べる前に, 多重線形性および歪対称性から導かれるいくつかの性質について調べておこう.

n 次正方形行列全体の集合を $M_n(\mathbb{R})$ と書くのであった. 本項および次項では写像 $F : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ について論じる. F の典型的な例は行列式である. 更に, 各 $X \in M_n(\mathbb{R})$ に, 列ベクトルを並べた表示 $X = [\mathbf{x}_1, \dots, \mathbf{x}_n]$ を与えておこう. このとき F は, 列ベクトルの組を代入する関数 $F(\mathbf{x}_1, \dots, \mathbf{x}_n)$ と見なすことができる. 写像 F における多重線形性および歪対称性を改めて述べておこう:

- 多重線形性: 各 $i = 1, \dots, n$ について,
 - (i) $F(\mathbf{x}_1, \dots, r\mathbf{x}_i, \dots, \mathbf{x}_n) = rF(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n)$,
 - (ii) $F(\mathbf{x}_1, \dots, \mathbf{y}_i + \mathbf{z}_i, \dots, \mathbf{x}_n) = F(\mathbf{x}_1, \dots, \mathbf{y}_i, \dots, \mathbf{x}_n) + F(\mathbf{x}_1, \dots, \mathbf{z}_i, \dots, \mathbf{x}_n)$.
- 歪対称性: 各 $i < j$ について,

$$F(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_j, \mathbf{x}_{i+1}, \dots, \mathbf{x}_{j-1}, \mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n) = -F(\mathbf{x}_1, \dots, \mathbf{x}_n).$$

次の主張は明らかと言えるが, ここでは形式的な証明を与えておく.

補題 12.3.1. F が多重線形ならば, 次の性質も満たす:

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

Proof. 和の個数 ℓ に関する帰納法で示す. $\ell = 1$ の場合は多重線形性の性質 (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。

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

□

次の補題は、命題 10.4.2 の証明と同様の論法により得られるゆえ、証明は省略しよう。

補題 12.3.2 (交代性). F が歪対称性を満たすとき、 $\mathbf{a}_i = \mathbf{a}_j$ (ただし $i < j$) ならば

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

歪対称性を用いて列の入れ替えを何度か行くと、それによる代入した値の変化は元の値の ± 1 倍である。このときの正負の符号は、入れ替えに対応する置換の符号に一致する。すなわち：

補題 12.3.3. F が歪対称性を満たすとき、任意の置換 $\sigma \in \mathfrak{S}_n$ に対して

$$F(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}) = \text{sgn}(\sigma) F(\mathbf{x}_1, \dots, \mathbf{x}_n).$$

Proof. σ を互換の積に分解した際の互換の個数 ℓ に関する帰納法により証明する。 σ 自身が互換 $\sigma = (i, j)$ の場合、 $F(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}) = \text{sgn}(\sigma) F(\mathbf{x}_1, \dots, \mathbf{x}_n)$ は歪対称性そのものである。

ℓ 個の互換の積で書ける任意の置換 $\tau \in \mathfrak{S}_n$ について $F(\mathbf{x}_{\tau(1)}, \dots, \mathbf{x}_{\tau(n)}) = \text{sgn}(\tau) F(\mathbf{x}_1, \dots, \mathbf{x}_n)$ が成り立つと仮定し、 $\sigma := \tau \cdot (i, j)$ (ただし $i < j$ とする) について $F(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}) = \text{sgn}(\sigma) F(\mathbf{x}_1, \dots, \mathbf{x}_n)$ が成り立つことを示そう。 i 列目と j 列目の添え字に注意して変形すると

$$\begin{aligned}
 F(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}) &= F(\mathbf{x}_{\tau(i,j)(1)}, \dots, \mathbf{x}_{\tau(i,j)(i)}, \dots, \mathbf{x}_{\tau(i,j)(j)}, \dots, \mathbf{x}_{\tau(i,j)(n)}) \\
 &= F(\mathbf{x}_{\tau(1)}, \dots, \mathbf{x}_{\tau(j)}, \dots, \mathbf{x}_{\tau(i)}, \dots, \mathbf{x}_{\tau(n)}) \\
 &= -F(\mathbf{x}_{\tau(1)}, \dots, \mathbf{x}_{\tau(i)}, \dots, \mathbf{x}_{\tau(j)}, \dots, \mathbf{x}_{\tau(n)}) \quad (\text{歪対称性を用いた}) \\
 &= -\text{sgn}(\tau) F(\mathbf{x}_1, \dots, \mathbf{x}_n) = \text{sgn}(\sigma) F(\mathbf{x}_1, \dots, \mathbf{x}_n). \quad (\text{帰納法の仮定を用いた})
 \end{aligned}$$

□

12.4 行列式の特徴づけ (証明)

n 次正方形行列 $A = [a_{ij}]$ の列ベクトル表示を $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ とする。これらの各列ベクトルを単純なベクトルの和に分解する方法を考えよう。そこで、長さ 1 の単位列ベクトルたち

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

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

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

である. 上の分解を用いて次の定理を示そう. 定理 11.3.1 は次の定理の特別な場合 ($F(E) = 1$ の場合) である.

定理 12.4.1. 写像 $F : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ が列ベクトルの組に関して多重線形性と歪対称性を満たすならば, $F(A) = \det A \cdot F(E)$ である.

Proof. 各 $A \in M_n(\mathbb{R})$ について $F(A) = |A| \cdot F(E)$ となることを示そう. そこで, $A = [a_{ij}] = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ とおく. 式 12.4.1 による分解, および補題 12.3.1 で示した多重線形性 (2) を各列に対して適用しながら展開すると,

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

いま得られた総和の各項において, $\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}$ の中に互いに同じものが一つでもあれば, 補題 12.3.2 より $F(\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}) = 0$ であるから, そのような項は無視して構わない. 一方, 無視できない各項において, $\mathbf{e}_{k_1}, \mathbf{e}_{k_2}, \dots, \mathbf{e}_{k_n}$ はいずれも異なる単位列ベクトルであり, とくに k_1, k_2, \dots, k_n に重複はない.

したがって, $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix}$ は置換である. このとき, 補題 12.3.3 より

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

以上より,

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

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

12.5 命題 11.1.1 の証明

次の定理における $r = 1$ の場合が命題 11.1.1 に相当する.

定理 12.5.1. A を r 次正方行列, X を s 次正方行列とすれば,

$$\begin{vmatrix} A & * \\ O_{s,r} & X \end{vmatrix} = |A| \cdot |X| = \begin{vmatrix} A & O_{r,s} \\ * & X \end{vmatrix}.$$

Proof. 一方の等式を示せば, それを転置することで他方の等式も得られる. ここでは右側の等式を示そう. 始めに, X が単位行列 E_s の場合について考える. そこで, $Y = \begin{bmatrix} A & O_{r,s} \\ * & E_s \end{bmatrix} = [a_{ij}]$ と置こう. Y は $n = r + s$ 次正方行列である. 行列式の定義によれば

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

である. ここで, 総和に現れる各項

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

について, この値が 0 でないための必要条件を検討しよう. まず, $O_{r,s}$ 成分に着目すると, $1 \leq i \leq r$ かつ $r+1 \leq j \leq n$ ならば $a_{ij} = 0$ である. よって $\{\sigma(1), \dots, \sigma(r)\}$ の中に r より大きな数 $\sigma(i)$ が一つでもあれば, $a_{i\sigma(i)} = 0$ ゆえこの置換 σ に関する式 (12.5.2) の値は 0 である. したがって, 式 (12.5.1) の総和における $\sigma \in \mathfrak{S}_n$ の動く範囲は,

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

を満たす部分のみを考えればよい. また, 上の条件を満たす σ においては

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

となっている. ここで Y の E_s 成分に着目すると, もし $\sigma(i) \neq i$ なる $i = r+1, \dots, n$ があれば Y の $(i, \sigma(i))$ -成分は E_s 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 $i = r+1, \dots, n$ について $\sigma(i) = i$ となる必要があり, したがって σ は次の形に限られる:

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

これは σ が \mathfrak{S}_r の元とみなせることに他ならない. 以上より, 式 (12.5.1) の総和における σ が動く範囲は \mathfrak{S}_r 上のみを考えればよいことから,

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

さて, 今度は X が一般の s 次正方行列の場合について考えよう. s 個の s 次列ベクトルの組 $X = [\mathbf{x}_1, \dots, \mathbf{x}_s]$ たち全体を定義域とする関数 $F: M_s(\mathbb{R}) \rightarrow \mathbb{R}$ を

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

と定めよう. すると, 行列式の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. \square

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

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

12.6 $|AB| = |A| \cdot |B|$ の証明

一般に, (m, n) -行列 A および (n, ℓ) -行列 $B = [\mathbf{b}_1, \dots, \mathbf{b}_\ell]$ に対して, $AB = [A\mathbf{b}_1, \dots, A\mathbf{b}_\ell]$ である (例 3.5.1). この事実注意到しながら次の定理を示そう.

定理 12.6.1. 積の行列式は行列式の積に等しい. すなわち, n 次正方行列 A, X について $|AX| = |A| \cdot |X|$.

Proof. n 次正方行列 $X = [\mathbf{x}_1, \dots, \mathbf{x}_n]$ たち全体を定義域とする写像 $F : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ を $F(\mathbf{x}_1, \dots, \mathbf{x}_n) := |AX|$ と定める. このとき,

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

であるから, F は多重線形性と歪対称性を満たす. 実際, 歪対称性は明らかであり, 多重線形性も次のように確認できる:

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

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

ゆえに定理 12.4.1 より $F(X) = |X| \cdot F(E) = |X| \cdot |AE| = |X| \cdot |A| = |A| \cdot |X|$ である. 以上により $|AX| = |A| \cdot |X|$ が示された. \square

次は 7.1 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. \square

上の命題の逆「 $|A| \neq 0$ ならば A は可逆である」もまた正しい. その証明は次節で与えよう (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、

$$A = \begin{pmatrix} \begin{matrix} A & \begin{matrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{nj} \end{matrix} & B \\ a_{i1} & \cdots & a_{in} \\ C & & D \end{matrix} \end{pmatrix} \quad \text{のとき} \quad A_{ij} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

例 13.1.2. $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ とすれば、 $A_{12} = \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix}$, $A_{22} = \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix}$ である。

列ベクトルによる成分表示を $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ とすれば、式 12.4.1 と同様にして、各 \mathbf{a}_j は次のように書ける:

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

したがって、補題 12.3.1 で示した行列式の多重線形性 (2) により、 $|A|$ は次の和に分解できる:

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

更に, 式 13.1.1 に現れる i 番目の項を計算すると

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

以上より, 式 13.1.1 は次のように書き下される:

$$|A| = a_{1j}(-1)^{1+j}|A_{1j}| + a_{2j}(-1)^{2+j}|A_{2j}| + \cdots + a_{nj}(-1)^{n+j}|A_{nj}| = \sum_{i=1}^n a_{ij}(-1)^{i+j}|A_{ij}|.$$

上の各項に現れる $(-1)^{i+j}|A_{ij}|$ を A の (i, j) -余因子 (cofactor) と呼び, この展開式のことを j 列に関する余因子展開 (cofactor expansion) という.

練習 13.1.3. j 列の代わりに i 行に関して同様の展開を行い, 次の等式を示せ.

$$|A| = a_{i1}(-1)^{i+1}|A_{i1}| + a_{i2}(-1)^{i+2}|A_{i2}| + \cdots + a_{in}(-1)^{i+n}|A_{in}| = \sum_{j=1}^n a_{ij}(-1)^{i+j}|A_{ij}|.$$

上式を i 行に関する余因子展開という.

成分に 0 が多く現れる行列式の値を求める際に余因子展開は有効である. 例えば命題 11.1.1 おける左の等号は 1 列目に関する余因子展開を, 右の等号は 1 行目に関する余因子展開を意味する.

例 13.1.4. 次の k 次正方行列について,

$$\begin{vmatrix} t & -1 & & & \\ & t & -1 & & O \\ & & \ddots & \ddots & \\ O & & & t & -1 \\ a_0 & a_1 & \cdots & a_{k-2} & t + a_{k-1} \end{vmatrix} = t^k + a_{k-1}t^{k-1} + \cdots + a_1t + a_0.$$

Proof. k に関する帰納法によって証明する. $k = 2$ の場合は次の計算で確かめられる:

$$\begin{vmatrix} t & -1 \\ a_0 & t + a_1 \end{vmatrix} = t(t + a_1) - a_0 \cdot (-1) = t^2 + a_1t + a_0.$$

サイズが $k-1$ のときに等式が成立すると仮定して、サイズが k の場合の等式を示そう。1 列目に関して左辺を余因子展開すると、第 1 項目に帰納法の仮定が適用できる:

$$\begin{aligned}
 (\text{左辺}) &= t \begin{vmatrix} t & -1 & & & \\ & \ddots & \ddots & & \\ & & t & -1 & \\ a_1 & \cdots & a_{k-2} & t+a_{k-1} & \end{vmatrix} + (-1)^{k+1} a_0 \begin{vmatrix} -1 & & & & \\ t & -1 & & & \\ & \ddots & \ddots & & \\ & & & t & -1 \end{vmatrix} \\
 &= t(t^{k-1} + a_{k-1}t^{k-2} + \cdots + a_2t + a_1) + (-1)^{k+1} a_0 \cdot (-1)^{k-1} = (\text{右辺}).
 \end{aligned}$$

なお、第 2 項目の行列式の値は、転置して例 11.1.2 を適用することで分かる。 \square

13.2 余因子行列

正方行列 A の逆行列の成分表示を与える方法を考えよう。 $C = A^{-1}$ ならば $CA = E$ を満たす。ゆえに C を行ベクトル分割し A を列ベクトル分割すれば次を満たすことになる。

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

すなわち、

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

上式を満たす行ベクトル \mathbf{c}_j を余因子展開を用いて導こう。

行列 $A = [a_{ij}]$ の第 j 列を別の列ベクトル \mathbf{b} に置き換えた行列 B について、 j 列に関する余因子展開を行えば、各 $i = 1, \dots, n$ について $B_{ij} = A_{ij}$ であるから、

$$\begin{vmatrix} a_{11} & \cdots & \overset{j}{\mathbf{b}_1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & \mathbf{b}_i & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & \mathbf{b}_n & \cdots & a_{nn} \end{vmatrix} = (-1)^{1+j} b_1 |A_{1j}| + (-1)^{2+j} b_2 |A_{2j}| + \cdots + (-1)^{n+j} b_n |A_{nj}|$$

この右辺は次の二つのベクトル

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

の内積と見ることができる。内積を行列の積として表現するには片方を転置する必要があり、ここでは

$\tilde{\mathbf{a}}_j$ を転置して

$$\begin{aligned} \left| \begin{array}{cccc} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & b_i & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{array} \right| &= \begin{bmatrix} (-1)^{1+j}|A_{1j}| & (-1)^{2+j}|A_{2j}| & \cdots & (-1)^{n+j}|A_{nj}| \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ &= {}^t\tilde{\mathbf{a}}_j \mathbf{b} \end{aligned} \quad (13.2.1)$$

を得る. とくに \mathbf{b} として A の各列 \mathbf{a}_i を取ると, 上式の左辺は $i = j$ の場合を除き i 列と j 列が等しい行列ゆえ 0 となり, $i = j$ の場合は $|A|$ である. すなわち,

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

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

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

ここで, $\tilde{A} = [a_{ij}^*]$ と成分表示すれば $a_{ij}^* = (-1)^{j+i}|A_{ji}|$ である (転置を取ったため添え字 i, j の位置が入れ替わっていることに注意せよ). すると式 (13.2.2) より $\tilde{A}A = |A|E$ を得る:

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

練習 13.2.1. 行に関する余因子展開に対して今と同様の議論を行うことで, $A\tilde{A} = |A|E$ を示せ.

定義 13.2.2. n 次正方行列 $A = [a_{ij}]$ に対して, 本項で定義した $\tilde{A} = [a_{ij}^*]$ (ただし $a_{ij}^* := (-1)^{j+i}|A_{ji}|$) を A の余因子行列 (**adjugate matrix**) という.

補足. A の (i, j) -余因子を (i, j) 成分とする行列, すなわち \tilde{A} の転置行列のことを **cofactor matrix** と言う. これを直訳すると「余因子行列」となるが, 邦語の文献では adjugate matrix のことを余因子行列と呼ぶのが慣例となっており, 本論もこれに従った.

これまでの議論により, $\tilde{A}A = A\tilde{A} = |A|E$ である. とくに $|A| \neq 0$ の場合は, 次の逆行列の公式を得る:

定理 13.2.3. 正方行列 A について, $|A| \neq 0$ ならば A は可逆であり $A^{-1} = \frac{1}{|A|}\tilde{A}$ である.

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

$$\frac{1}{|A|}\tilde{A} \cdot A = \frac{1}{|A|} \cdot |A|E = E, \quad A \cdot \frac{1}{|A|}\tilde{A} = \frac{1}{|A|}(A\tilde{A}) = \frac{1}{|A|} \cdot |A|E = E.$$

□

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. \square

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

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

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:

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

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:

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

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 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ (ただし各 \mathbf{a}_j は n 次列ベクトル) において, A が可逆 (すなわち $|A| \neq 0$) ならば, 連立 1 次方程式 $A\mathbf{x} = \mathbf{b}$ の唯一解 $\mathbf{x} = {}^t[x_1, \dots, x_n]$ は, $\mathbf{x} = A^{-1}\mathbf{b}$ である (命題 4.3.1). このとき, 各成分 x_i は次の式で与えられる:

$$x_j = \frac{|\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n|}{|A|} \quad (j = 1, \dots, n).$$

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

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

したがって、各成分 x_j は、

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

最後の等式は式 13.2.1 による. □

余因子行列を持ちださずともクラメルの公式を示すことはできる. 行列式への理解を深めるために, 余因子行列を用いない別証明を与えておこう:

クラメルの公式の別証明. $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ と置けば

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

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

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

ここで, 上式の第 i 項に現れる行列式は, $i = j$ の場合を除き i 列と j 列がともに \mathbf{a}_i であるから, その値は 0 である. また第 j 項は $x_j |A|$ である. 以上より,

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

これを移項して求める等式を得る. □