

行列と行列式

(c) 角田 保 (大東文化大学経済学部)

2018 年 12 月 25 日 ver.

- 目次に*がついている節は、 n 次正方行列の行列式への定義の準備なので、細部にこだわらず軽く読んでくれればよい.
- 目次に**がついている節はさらに高度なので、飛ばしても構わない.
- 参考文献リストは先に書いておいた.
- 索引はない. というのも PDF ファイルなので, Acrobat Reader ならば Ctrl と Shift を押しながら Enter すれば, 検索ができるためである. それを利用するのがとても楽である.

目次

0	まえがき (ここだけですます調)	4
	第 I 部 数ベクトルと行列計算の定義	6
1	数ベクトルと行列の定義, および四則演算	6
1.1	数ベクトルと行列の定義	6
1.2	ベクトルの演算の定義	7
1.3	2 項ベクトルの応用:平面ベクトル (1)	9
1.4	行列の四則演算の定義	13
2	いろいろな行列	15
	第 II 部 主に 2×2 の行列までの計算	20
3	2 次正方行列の行列式と逆行列	20
4	2 次正方行列のクラメールの公式	22
5	2 次正方行列の行列式の性質	23
6	2×2 までの行列の積を, 分割した行列で積で表す	26
7	2 次正方行列の固有値と固有ベクトル	27

8	実ベクトルの内積とベクトルの大きさ	30
9	内積までの 2 項ベクトルの応用:平面ベクトル (2)	33
9.1	法線ベクトルと点と直線の距離	33
9.2	平面全体を 2 つのベクトルの線形結合で表せる場合と、表せない場合	34
9.3	2 項列ベクトルの左から 2 次正方行列を掛けることと、行列式の関係	36
9.4	回転変換と三角関数の加法定理	38
9.5	線対称の点に移す変換を表す行列	40
10	2 次実正方行列の直交行列と三角化・対角化	41
11	2 次実正方行列の対角化の応用例	45
12	2 次実対称行列の性質	46
13	n 次実正方行列による、2 次形式の定義	48
14	2 次実正方行列と 2 次形式	49
15	第 II 部の Scilab 実習	51
15.1	[] を用いて行列とベクトルの作成. 和と積	51
15.2	実行列の転置行列を返す' と, sqrt 関数・norm 関数	52
15.3	行列式 det 関数と逆行列 inv 関数	53
15.4	固有値と固有ベクトルを返す spec 関数と 10 章の実習	54
15.5	行列の全成分を表す「:」の利用	55
15.6	spec 関数と 12 章の実習	55
15.7	よく使われる関数 (1): eye,zeros,ones	56
15.8	よく使われる関数 (2): sum,mean	56
15.9	成分ごとの掛け算・割り算・べき乗:「.*」「./」「.^」の利用	56
15.10	等差数列を返す行ベクトル:「:」の利用	57
第 III 部 主に 2×2 までの複素行列の計算		58
16	複素数入門	58
16.1	複素数の定義と四則	58
16.2	実数での数直線と複素数での複素平面	59
16.3	複素平面と複素数の絶対値の定義	60
16.4	* 複素数係数の 2 次方程式の解と係数の関係	61
17	複素ベクトルのエルミート内積: 第 I 部・第 II 部の拡張	62
18	** ユニタリー行列と 2×2 の複素行列の三角化・対角化	64

19	第 III 部の Scilab 実習	66
19.1	随伴行列 (実行列なら転置行列) を返す' と, sqrt 関数・norm 関数	67
19.2	複素数でも, 固有値と固有ベクトルを返す spec 関数	68
第 IV 部 3 × 3 の行列を中心に		69
20	* 3 × 3 の行列の行列式はどうあるべきか	69
21	* 3 × 3 の連立方程式のクラメールの公式	71
22	** 3 × 3 の逆行列の式	73
23	* 3 × 3 の行列と 2 次形式	74
第 V 部 一般の n 次正方行列を中心に		76
24	* 4 × 4 の行列の行列式はどうあるべきか	76
25	n 次正方行列の行列式の定義の準備	77
26	n 次正方行列の行列式	80
27	行列式の性質とクラメールの公式	85
28	n 次正方行列について, 積の行列式 = 行列式の積 の証明	88
29	正方行列の余因子展開	89
30	余因子展開から逆行列へ	92
31	分割された行列と, 行列式・逆行列の表記	94
32	第 V 部の Scilab 実習	98
第 VI 部 1 次独立と行列の階数		103
33	ベクトルの 1 次独立と 1 次従属および行列のランク (階数)	103
34	基本行列・基本変形と, 逆行列	104
35	n 元 1 次連立方程式を行変形のみで解く・逆行列の PC 計算	114
36	トレースの定義	117
37	第 VI 部の Scilab 実習	118

37.1	find 関数と sum 関数についての注意	118
37.2	基本変形・ランク・逆行列の復習:全て表示する関数	119
37.3	拡大係数行列による連立方程式の解や逆行列の復習	124
第 VII 部 固有値・対角化・三角化・2次形式		127
38	固有値・固有ベクトルと, 対角化・三角化	127
39	対称行列を表す「 λ 」と2次形式と定符号行列	133
40	2次形式までを n 変数実数値関数としてみる.	138
40.1	偏微分・ヤコビ行列・ヘッセ行列	138
40.2	2次形式と関数の凹凸と最大・最小	140
41	第 VII 部の Scilab 実習	142
第 VIII 部 計量経済学・確率論・統計学で出会うもの		146
42	回帰分析の回帰係数の導出	146
43	確率変数の期待値行列・共分散行列と線形変換 (正規分布)	151
44	同時密度関数・線形変換・同時正規分布の線形変換	154
45	2次形式・標準正規分布・カイ 2 乗分布・ t 分布・ F 分布	159
46	回帰分析での係数の t 検定や F 検定への導出	162
47	基本的な統計的推定や検定を回帰分析で説明する.	165
47.1	(47.10) 式の証明	177

0 まえがき (ここだけですます調)

この資料は,

- 大東文化大学経済学部 3 年生開講科目 (PC 経済分析 I,II) の参考資料.
- 大東文化大学経済学部角田ゼミ生新 3 年用の, 復習と予習.
- 将来, 本経済学部大学院生に対する簡単な数学プリント (math camp) への使用.

を全て兼ねたものです. 講義するなら 90 分× 5~6 回の分量を意識しています. ですから, 線形代数の浅いところ (具体的には, 行列と行列式の演算) に絞って, 説明します. それ以外の, 部分ベクトル空間, 底の変換, 次元, ジョルダン標準形などは, 参考文献を参考にしてください.

この資料の 1 つの目的は, 2×2 の行列までについて, 大がかりな理論なしに, 計算によっていろいろな結果を出すことです. そのため第 II 部の 2×2 の部分が結構な分量になっていますが, 2 次方程式までしっかり

勉強している人にとっては、第 II 部までは比較的簡単に読めることでしょう。特に第 II 部までの 4 割弱は、

1985 年に私が高 2 で学んだ「代数・幾何」の、行列や 1 次変換の部分

にあたり、その記憶を思い出しながら書きました。具体的には、

(1) 第 1 部での 2×2 までの場合。(2) 第 3 節・第 4 節・第 7 節。(3) 第 8 節を $n = 2$ で考えた場合。(4) 第 9 節。

よって、第 II 部までは、2015 年度以降の大学入学生が、自習するのにちょうどいい資料だと思います。
この資料は、参考文献にある齋藤正彦先生のあとがきになって、

日本の学生の数理能力を底上げすること

を目指して書きました。2015 年以降では、高校のカリキュラムの変更により、理系の線形代数の講義内容は、大幅な変更を余儀なくされます。個人的には、講義内容も学生の理解度も大幅に低下することは明白です。それを防ぐために、すでに免疫のある文系大学所属の私が、時間(と根性)を投入して、このような資料を作りました。線形代数で全く何やっているか分からない学生さんがいましたら、第 II 部まででも十分なので、一度見てほしいと思います。

さらに Scilab5.4.1 を使った実習もふんだんに盛り込みました。そのあたりは他とかなり違う特色があるでしょうし、Scilab の部分だけを自習してもかなり楽しめると思います。Scilab は、

www.scilab.org

から無料でダウンロードできますので、必ずダウンロードしてください。

参考文献

- 齋藤正彦 (1966) 「線形代数学」, 東京大学出版会
- 佐武一郎 (1974) 「線形代数学」, 裳華房
- 川久保勝夫 (2010) 「線形代数学 (新装版)」, 日本評論社
- 長谷川 浩司 (2004) 「線形代数」, 日本評論社
- 小山昭雄 (2010) 「[新装版] 経済数学教室 第 1 巻・第 2 巻」, 岩波書店

齋藤と佐武は、私の学生時代でも定番のものでした。2 つそろって、日本における線形代数学の教科書の父母のようなものです。が、現代の文系の学生が独学で読むのは難しいでしょう。川久保の第 1 版は、私が大学に勤める少し前に出たものです。新装版は誤植もずいぶん減って、かなりわかりやすく書かれていると思います。長谷川は、私が大学に勤めてから出たものです。簡素ですが、定評のあるものです。小山は旧版はもっと前に出たものです。経済学でよく使うラグランジアンに関する行列への対処など、経済学への応用を考えたものとなっています。

どれをとっても、このレジュメのレベルをはるかに超えるものですから、このレジュメを軽く読むことが、上記参考文献への橋渡しとなることでしょう。

第I部

数ベクトルと行列計算の定義

a, b が実数で, x に関する1次方程式

$$ax = b \quad (0.1)$$

は, $a \neq 0$ のとき $x = \frac{b}{a} = a^{-1}b$ という一意な解をもつことは, 中学校1年生で学んだ. また a^{-1} を a の逆数といった. そこで x, y に関する連立方程式

$$\begin{cases} ax + by = e \\ cx + dy = f \end{cases}$$

も,

$$Ax = b$$

と表すことができたりすると便利であり, さらに (0.1) 式の1次方程式の解での $a \neq 0$ のように, A に何らかの条件をつけるにしても,

$$x = A^{-1}b$$

と書けるととても便利であろう*1.

そのような計算ができるように, 数ベクトルや行列を定義し, 計算も定義する.

またこのレジュメでは, 文字

$$k, \ell, m, n$$

は, 全て正の整数を表すとする.

1 数ベクトルと行列の定義, および四則演算

1.1 数ベクトルと行列の定義

数を縦に並べたものを, 列ベクトルという. 数を横に並べたものを, 行ベクトルという*2.

$$(1). \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \quad (2). (1 \ 0),$$

で, (1) は, 3項列ベクトル・成分3個の列ベクトル・長さ3の列ベクトル, などという. 同様に, (2) は, 2項行ベクトル・成分2個の行ベクトル・長さ2の行ベクトル, などという. ベクトルの中身の数字を, 成分とか要素などという. 例えば,

$$b = (0 \ -3)$$

のとき, b の第1成分は0で, 第2成分は-3である.

*1 ところで今は太字 x, b を使っているので, 細字の x, b とは違うことを認識しておいてほしい.

*2 数と単に書いたのだが, この数とは, そして数ベクトルの理論とは複素数の範囲まで成り立つことであることがほとんどである. が, これは入門的なものなので, 固有値と固有ベクトルの章のみ複素数を扱うこととし, それ以外は全て, 実数を表すこととする.

次に行列について定義しよう。数を縦横に長方形になるように並べたものを行列という。

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

(3) は 2 行 3 列の行列といい, (4) は 2 行 2 列の行列である。行数と列数が同じものを正方行列という。よって上記 (4) のことを, 2 次正方行列ともいう。

数ベクトルや行列は \mathbf{a} や \mathbf{B} のように, 太字で表すことが一般的である (そうでない本もあるが)。またベクトルは小文字で, 行列は大文字になっていることも多い。

行列の成分の表示方法については, \mathbf{A} が m 行 n 列のとき, \mathbf{A} の i 行 j 列の値を, \mathbf{A} の (i, j) 成分といい, a_{ij} で表すことがほとんどである。あらわに書いてみると,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad (1.1)$$

である。 a_{11} で \mathbf{A} の「じゅういち」成分とは読まないようにすること。十一ではなくて, (1,1) 成分である。発音では略語的に「いちいち」成分とも読む人もいる。あくまでも a 行数列数 と, 行数と列数をそのまま並べたものである。

行列 \mathbf{A} について, その行と列をひっくり返したものを \mathbf{A} の転置行列といい, ${}^t\mathbf{A}, \mathbf{A}^T, \mathbf{A}'$ などで表す。つまり,

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \end{pmatrix} \implies {}^t\mathbf{A} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 0 \end{pmatrix}$$

である。日本の線形代数の本では ${}^t\mathbf{A}$ を使うことが多く, 米国の計量経済学の本などでは \mathbf{A}' などを使うことが多い。このレジュメでは ${}^t\mathbf{A}$ で表すことにする。

1.2 ベクトルの演算の定義

「数ベクトル」とわざわざベクトルというものに「数」という言葉を付けている。それは, ベクトルというものにはいろいろなものがあるが, 数ベクトルとは, その中の一部のものであることを意味している。そこで, ベクトル空間 V というものがどういうものであるかと, それに合わせて数ベクトルの計算の定義をここで行う。ベクトル空間の満たす条件は以下の 8 つである。その条件を示した後, それぞれについて, 数ベクトル \mathbf{a} についての計算が, その条件を満たしていることを確認してほしい。

(i). (交換法則) 任意の $\mathbf{u}, \mathbf{v} \in V$ について, $\mathbf{u} + \mathbf{v} \in V$ かつ $\mathbf{v} + \mathbf{u} \in V$ が成り立ち, かつ

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$$

- 数ベクトル \mathbf{a} と \mathbf{b} を考えよう。この 2 つが同じサイズのベクトルのとき, $\mathbf{a} = \mathbf{b}$ ということは, 各成分がそれぞれ等しいことを意味するように定義する。そして和 $\mathbf{a} + \mathbf{b}$ を計算可能とし, その計算方法は, 成分ごとの和と定義する。例えば, 2 つの n 項列ベクトルの場合

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

等号については,

$$\begin{cases} a_1 = b_1 \\ a_2 = b_2 \\ \vdots \\ a_n = b_n \end{cases}$$

和の計算については,

$$\mathbf{a} + \mathbf{b} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

である.

n 項行ベクトルの場合も同様に定義できる. つまり,

$$\mathbf{a} = (a_1 \ a_2 \ \cdots \ a_n), \quad \mathbf{b} = (b_1 \ b_2 \ \cdots \ b_n)$$

として, 等号については,

$$\mathbf{a} = \mathbf{b} \iff a_1 = b_1, \ a_2 = b_2, \ \cdots, \ a_n = b_n$$

であり, 和の計算については,

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1 \ a_2 + b_2 \ \cdots \ a_n + b_n)$$

である. 同じサイズの数ベクトルについては, このようにそれぞれの成分の和として定義するので, 交換法則 $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ も成り立っていることがわかる.

(ii). (結合法則) 任意の $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ について, 以下が成り立つ.

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$$

- 数ベクトルの場合は, (i) のように和の計算を定義したので, この交換法則が成り立つ.

(iii). (加法のゼロ元) 任意の $\mathbf{v} \in V$ について, 以下を満たす $\mathbf{0} \in V$ が唯一存在する.

$$\mathbf{v} + \mathbf{0} = \mathbf{v}$$

- 数ベクトルの場合は, 定義した数ベクトル \mathbf{a} と同じサイズで, 成分が全て 0 であるベクトルをゼロベクトル $\mathbf{0}$ と定義することにより, これが数ベクトルの場合の加法のゼロ元となる.

(iv). (加法の逆元) 任意の $\mathbf{u} \in V$ について, 以下を満たす $\mathbf{v} \in V$ が必ず唯一存在し, それを $-\mathbf{u}$ と表す.

$$\mathbf{u} + \mathbf{v} = \mathbf{0}$$

- 数ベクトルの場合は, \mathbf{a} の成分を全て -1 倍したものを $-\mathbf{a}$ としている. すると上記の逆元の条件を満たす.

(v). (スカラー倍) スカラーというものがあその集合を K とする, 任意の $\mathbf{v} \in V$, 任意のスカラー $c \in K$ について, 以下が成り立つ.

$$c\mathbf{v} \in V$$

- 数ベクトルの場合は、 K としてスカラーは実数の集合 \mathbf{R} や、複素数の集合 \mathbf{C} にすることが多い。
1.1 節で述べたように今は実数の集合 \mathbf{R} で定義しておく。数ベクトル \mathbf{a} について、 $c\mathbf{a}$ を以下のようにして、それぞれの成分に c を掛けることと定義する。

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \implies c\mathbf{a} = \begin{pmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_n \end{pmatrix}$$

これによって、スカラー倍の条件を満たすこととなる。

- (vi). (スカラーの交換法則) 任意の $\mathbf{v} \in V$ と、2つの任意のスカラー c, d について、

$$(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$$

- 数ベクトルの場合は、スカラー倍で定義した積の方法から、これが成り立つようになっている。

- (vii). (ベクトルの交換法則) 2つの任意の $\mathbf{u}, \mathbf{v} \in V$ と、任意のスカラー c について、

$$c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$$

- 数ベクトルの場合は、スカラー倍で定義した積の方法から、これが成り立つようになっている。

- (viii). (積の単位元) 任意の $\mathbf{v} \in V$ について、以下を満たすスカラー 1 が唯一存在する。

$$\mathbf{v} = 1\mathbf{v}$$

- 数ベクトルの場合は、実数の 1 を考えることによって、任意の数ベクトル \mathbf{a} について

$$\mathbf{a} = 1\mathbf{a}$$

が成り立つ。

以上から、ベクトル空間の満たすべき条件は8つあるが、現時点では特に覚えなくともよい。さらに言えば、数ベクトルの計算方法としては、

(i) の成分ごとの足し算と、(v) のスカラー倍の計算の2つをマスターすれば十分である。

それ以外は、(i)(v) の計算ができれば自由に得られる。とにかく、まずはこの2つをきちんと計算できるようにしておくことを目標にするとよい。そのために、中学や高校で学んだ xy 平面上の点を、ベクトルで表すことを紹介しよう。

1.3 2項ベクトルの応用:平面ベクトル (1)

xy 平面上の座標を (x, y) とおく。点 $A(x_1, y_1)$ から点 $B(x_2, y_2)$ を結ぶ直線の傾きを、 $x_1 \neq x_2$ のとき

$$\frac{y_2 - y_1}{x_2 - x_1} \tag{1.2}$$

とあらわしたのは、中学校のときである。そしてこれを利用して、2点 A, B を通る直線は、

$$\begin{cases} y = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1 & (x_2 \neq x_1 \text{ のとき}) \\ x = x_1 & (x_2 = x_1 \text{ のとき}) \end{cases} \tag{1.3}$$

とあらわしたことも思い出してほしい。そしてこれをもう少し考えよう。

傾きを表す (1.2) 式の意味は、

$$\frac{A \text{ から } B \text{ への } y \text{ 軸方向への増分}}{A \text{ から } B \text{ への } x \text{ 軸方向への増分}}$$

である。よって分数にしないで、

$$\overrightarrow{AB} = \begin{pmatrix} A \text{ から } B \text{ への } x \text{ 軸方向への増分} \\ A \text{ から } B \text{ への } y \text{ 軸方向への増分} \end{pmatrix}$$

という 2 項実列ベクトルで表すと便利であろう。というのも

$$\overrightarrow{OA} = \begin{pmatrix} O \text{ から } A \text{ への } x \text{ 軸方向への増分} \\ O \text{ から } A \text{ への } y \text{ 軸方向への増分} \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

と、座標を単純に列ベクトル化すればよいだけだからである。さらに言えば、数ベクトルの和を前節で定義したから、

$$\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB}$$

が成り立つ。実際、

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}$$

が成り立っていることがわかるであろう。いままで学んだことを簡単に定義すると以下である。

定義 1.1 xy 平面が与えられているとき、任意の点 $P(x_1, y_1)$ について、列ベクトル \overrightarrow{OP} を、座標を列ベクトル化したものとする。式で書くと、

$$\overrightarrow{OP} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

である。またこれを P の位置ベクトルと呼ぶ。 」

位置ベクトルは小文字の太字で表すことも多い。つまり、今の場合だと \overrightarrow{OP} を \mathbf{p} のように表す。さて座標を列ベクトルで表すと、ベクトルの和と定数倍という考え方から、任意の点 (x, y) について、

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

と表すことができる。右辺が x, y のままでは応用に欠けるので、実数 s, t にしてみよう。

命題 1.2 xy 平面上の原点を O とする。任意の点 $P(x, y)$ の位置ベクトルは、

$$\overrightarrow{OP} = s \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

によって、 s, t によって一意に定められる。 」

証明 一意性だけは証明しておこう。

$$\overrightarrow{OP} = \begin{cases} s \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ s' \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t' \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{cases}$$

と s, t, s', t' によってあらわされたと仮定しよう。これが成り立つならば、

$$\begin{aligned} s \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= s' \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t' \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \iff \begin{pmatrix} s \\ t \end{pmatrix} &= \begin{pmatrix} s' \\ t' \end{pmatrix} \end{aligned}$$

より、 $s' = s, t' = t$ が成り立つ。

□

命題 1.2 の形の s, t が実数上を自由に動くことによって、 P が平面全体を表すのである。すると、 s, t を制限すると、 P の動く範囲も制限されるのは明らかであろう。

例題 1.3 xy 平面上の任意の点 $P(x, y)$ の位置ベクトルについて、

$$\overrightarrow{OP} = s \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

とあらわしたとき、 s, t を以下に定めたときに、 P がどのような範囲になるかを考えよ。

- (i). $t = 0$
- (ii). $s = 0$
- (iii). $t = 0, s \geq 0$
- (iv). $s = 0, t \geq 0$
- (v). $0 \leq s \leq 1, 0 \leq t \leq 1$
- (vi). $0 \leq s \leq 1, t = 1 - s$

┘

(解答)

- (i). x 軸上の点である。式で書くと、 $y = 0$ 。
- (ii). y 軸上の点となる。式で書くと、 $x = 0$ 。
- (iii). x 軸上の点で、 O から x 軸の正の方向への半直線となる。式で書くと、 $y = 0, x \geq 0$ 。
- (iv). y 軸上の点で、 O から y 軸の正の方向への半直線となる。式で書くと、 $x = 0, y \geq 0$ 。
- (v). $A(1, 0), B(0, 1), C(1, 1)$ とすると、正方形 $OACB$ の内部と境界線を表す。
- (vi). (v) の A, B で、線分 AB 上の点を表す。というのも $t = 1 - s$ を代入すると $\overrightarrow{OP} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \overrightarrow{OB} + s \overrightarrow{BA}$ となる。 $\overrightarrow{OP} - \overrightarrow{OB} = \overrightarrow{BA}$ なので、 $\overrightarrow{BP} = s \overrightarrow{BA}$ となるため。

(解答終)

この (i) を考えてみよう。 s が任意の時、

$$\begin{pmatrix} x \\ y \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

が、直線 $y = 0$ を表すのである。 0 の代わりに a としよう。つまり

$$\begin{pmatrix} x \\ y \end{pmatrix} = s \begin{pmatrix} 1 \\ a \end{pmatrix} \tag{1.4}$$

である (s は任意の実数). $x = s, y = as$ だから, s を消去してなんと $y = ax$ である. つまり上の式は, 原点 O と点 $(1, a)$ を結ぶ直線となる. 傾き a で y 切片が 0 の直線の式 $y = ax$ が (1.4) のように表されるのである.

ではさらに応用して, $y = ax + b$ を考えてみよう. 上の式を y 軸方向に b だけ進めればよいので,

$$\begin{pmatrix} x \\ y \end{pmatrix} = s \begin{pmatrix} 1 \\ a \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix}$$

で表される. 和の順序を変えて, さらに傾き a だけでなく, 「 $x = \text{定数}$ 」という直線も含むような形で表すと, 以下がいえる.

命題 1.4 xy 平面上の異なる任意の 2 点 $A(x_1, y_1), B(x_2, y_2)$ について, 直線 AB 上の点 $P(x, y)$ は, 任意の実数 t によって

$$\begin{pmatrix} x \\ y \end{pmatrix} = \overrightarrow{OA} + t \overrightarrow{AB} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix} \quad (1.5)$$

と表すことができる. 」

s を t に変えたのは, 単に t を使う本が多いからである. このように表すことによって, (1.3) 式での x_1, x_2 に関する条件を外すことができるのである.

さてこの節の最後に, 直線の方法ベクトルを定義しておこう. (1.3) 式のように, 場合分けするのは面倒である. そこで以下のように, xy 平面上の直線は 1 つにまとめよう.

$$xy \text{ 平面上の直線は, } ax + by + c = 0 \quad (\text{ただし } a^2 + b^2 \neq 0) \quad (1.6)$$

実際 $b \neq 0$ ならば, $y = -\frac{a}{b}x - \frac{c}{b}$ であり, $b = 0$ ならば $x = -\frac{c}{a}$ で表される. この直線の表現を用いて,

定義 1.5 実数 a, b, c (ただし $a^2 + b^2 \neq 0$) として, xy 平面上の直線 $ax + by + c = 0$ が与えられたとき, この直線の方法ベクトルを,

$$s \begin{pmatrix} b \\ -a \end{pmatrix} \quad (\text{ただし } s \neq 0)$$

とする. 」

実際この直線 $ax + by + c = 0$ が点 $A(x_1, y_1)$ を通るとき, 直線上の点を $P(x, y)$ とすれば, 以下が成り立つ.

$$\begin{cases} ax + by + c = 0 \\ ax_1 + by_1 + c = 0 \end{cases}$$

この式を辺々引いて移項すると, $a(x - x_1) = -b(y - y_1)$ が成り立つ. よって任意の実数 t, s (ただし $s \neq 0$) によって,

$$\begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix} = ts \begin{pmatrix} b \\ -a \end{pmatrix}$$

がいえるから,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t \begin{pmatrix} sb \\ -sa \end{pmatrix}$$

と書ける.

(1.5) 式の形から方向ベクトルは、あるベクトルの 0 以外の定数倍である。直線 (1.6) 式が与えられたとき、この直線の簡単な方向ベクトルの例としては、

$$\begin{pmatrix} b \\ -a \end{pmatrix}, \begin{pmatrix} -b \\ a \end{pmatrix}$$

の 2 つがあげられる。これは単に、(1.5) 式に $s = 1, -1$ を代入したものである。

さて平面上に直線が存在すれば、それと直交する直線も必ず存在する。その直交を考えるのみベクトルが非常に役立つのである。直交については、9.1 節で詳しく述べる。

1.4 行列の四則演算の定義

行列が 1 行の場合は行ベクトルであり、行列が 1 列の場合は列ベクトルとなるので、それらの場合の計算も含むことを注意しておくこと。

- (i). (行列の加法) \mathbf{A} と \mathbf{B} が同じサイズ (行数・列数とも同じ場合) であるときのみ、各成分ごとに加えたものを $\mathbf{A} + \mathbf{B}$ で表す。もちろん $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ である。例えば

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 10 \end{pmatrix} + \begin{pmatrix} 0 & 1 & -1 \\ 2 & 1 & 20 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \\ 5 & 3 & 30 \end{pmatrix}$$

である。異なるサイズの行列については、加法は計算不能である。

- (ii). (スカラー倍) 行列のスカラー倍の定義も、数ベクトルの場合と同様である。つまり、スカラーを行列の成分全てに掛けたものと定義する。任意の実数 c について、

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \implies c\mathbf{A} = \mathbf{A}c = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}$$

- (iii). (行ベクトルと列ベクトルの積) n 項行ベクトル \mathbf{a} と、 n 項列ベクトル \mathbf{b} の積は、対応する成分の積の和と定義する。つまり、

$$\mathbf{a}\mathbf{b} = (a_1 \ a_2 \ \cdots \ a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n = \sum_{i=1}^n a_ib_i$$

ここで、「行ベクトルと列ベクトルの積」というのがポイントである。和の場合と異なり、交換法則は成り立たない。

- (iv). (行列の積) この (iii) の「行ベクトルと列ベクトルの積」の計算方法を応用して、行列の積を定義しよう。 \mathbf{A} は m 行 k 列、 \mathbf{B} は k 行 n 列の行列とする。このとき、 \mathbf{A} と \mathbf{B} の積は計算可能で、 \mathbf{AB} は m 行 n 列の行列となる。そしてその各成分の計算方法については、 \mathbf{AB} の (i, j) 成分を

$$\mathbf{AB} \text{ の } (i, j) \text{ 成分} = (\mathbf{A} \text{ の第 } i \text{ 行}) \text{ と } (\mathbf{B} \text{ の第 } j \text{ 列}) \text{ の積}$$

とする。つまり \mathbf{AB} の計算については、 \mathbf{A} は行ベクトルが縦に m 個並んだもので、 \mathbf{B} は列ベクトルが横に n 個並んだものと考え、(iii) での行ベクトルと列ベクトルの積の計算を繰り返すのである。

ではそのように \mathbf{A}, \mathbf{B} をベクトルの塊とみて, 計算をみてみよう. $\mathbf{a}_i (i = 1, 2, \dots, m)$ を k 項行ベクトル, $\mathbf{b}_j (j = 1, 2, \dots, n)$ を k 項列ベクトルと考えて,

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_m \end{pmatrix} (\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_j \ \cdots \ \mathbf{b}_n) = \begin{pmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \cdots & \mathbf{a}_1\mathbf{b}_j & \cdots & \mathbf{a}_1\mathbf{b}_n \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_j & \cdots & \mathbf{a}_2\mathbf{b}_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_i\mathbf{b}_1 & \mathbf{a}_i\mathbf{b}_2 & \cdots & \mathbf{a}_i\mathbf{b}_j & \cdots & \mathbf{a}_i\mathbf{b}_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \mathbf{a}_m\mathbf{b}_2 & \cdots & \mathbf{a}_m\mathbf{b}_j & \cdots & \mathbf{a}_m\mathbf{b}_n \end{pmatrix}$$

となるのである. また先に \mathbf{A} と \mathbf{B} の行数と列数についてそれぞれ, m 行 k 列と k 行 n 列と書いた. このように, \mathbf{A} の列数と \mathbf{B} の行数が一致するとき, そのときのみ \mathbf{AB} の計算が可能である.

上のように行列の積を一般形で書いたのだが, まずは 2 行 2 列までの計算に慣れることを目標としよう. ここで例示していこう.

(a) 行ベクトルと列ベクトルの積はもう定義した.

$$(a_1 \ a_2) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1b_1 + a_2b_2$$

である.

(b) では, \mathbf{A} が 2 行 2 列で, \mathbf{b} が 2 行 1 列の列ベクトルのときを考えよう. この場合 \mathbf{Ab} は計算可能で, 2 行 1 列の列ベクトルとなる. 成分で書くと,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + a_{12}b_2 \\ a_{21}b_1 + a_{22}b_2 \end{pmatrix}$$

である.

$$\begin{pmatrix} (\mathbf{A} \text{ の第 1 行目} \text{ と } \mathbf{b} \text{ の積}) \\ (\mathbf{A} \text{ の第 2 行目} \text{ と } \mathbf{b} \text{ の積}) \end{pmatrix}$$

という 2 つの「行ベクトルと列ベクトルの積」の形になっていることがわかるであろう.

(c) 次に \mathbf{a} が 2 行 1 列の行ベクトルで, \mathbf{B} が 2 行 2 列の行列のとき考えよう. この場合 \mathbf{aB} も計算可能で, 1 行 2 列の行ベクトルとなる. 成分で書くと,

$$(a_1 \ a_2) \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = (a_1b_{11} + a_2b_{21} \quad a_1b_{12} + a_2b_{22})$$

である.

$$(\mathbf{a} \text{ と } \mathbf{B} \text{ の 1 列目} \text{ の積} \quad \mathbf{a} \text{ と } \mathbf{B} \text{ の 2 列目} \text{ の積})$$

という 2 つの「行ベクトル * 列ベクトル」を計算していることがわかるであろう.

(d) \mathbf{A}, \mathbf{B} とも 2 次正方行列のときは \mathbf{AB} は計算可能で, やはり 2 次正方行列となる. 成分で表示すると,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

この 2 行 2 列の行列が,

$$\begin{pmatrix} \mathbf{A} \text{ の 1 行目と } \mathbf{B} \text{ の 1 列目の積} & \mathbf{A} \text{ の 1 行目と } \mathbf{B} \text{ の 2 列目の積} \\ \mathbf{A} \text{ の 2 行目と } \mathbf{B} \text{ の 1 列目の積} & \mathbf{A} \text{ の 2 行目と } \mathbf{B} \text{ の 2 列目の積} \end{pmatrix}$$

となっていることを, よく確認すること.

(e) 盲点になりやすいが, n 項列ベクトルと n 項行ベクトルの積も計算可能であり, n 次正方行列となる. $n = 2$ の場合で成分で書くと,

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (b_1 \ b_2) = \begin{pmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{pmatrix}$$

A, B がともに n 次正方行列であるならば, AB と BA の両方が計算可能である. このとき, 一般に,

$$AB \neq BA$$

であることに気を付けること. つまり一般に積の順序を入れ替えてはいけないのである. しかし次の場合のように, 順序を入れ替えなければ, 等号が成り立つ.

定理 1.6 (行列の積の結合法則) A は m 行 n 列, B は n 行 k 列, C は k 行 l 列とする. このとき以下が成り立つ.

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

┘

証明は簡単なのでしない. 行列の積の定義が, 面倒な形になっているのは, この定理が成り立つためでもある.

和と積を説明したので, それらを組み合わせた分配法則などを説明しよう.

命題 1.7 A, A_1, A_2 は m 行 k 列, B, B_1, B_2 は k 行 n 列の行列とする. このとき以下が成り立つ.

- (i). $A(B_1 + B_2) = AB_1 + AB_2$
- (ii). $(A_1 + A_2)B = A_1B + A_2B$
- (iii). $B_1 = B_2 \implies AB_1 = AB_2$
- (iv). $A_1 = A_2 \implies A_1B = A_2B$

┘

2 いろいろな行列

行列にもいろいろなものがあるので, それを紹介しよう. 以下の (i) から (viii) までの A は全て n 次正方行列とする.

- (i). 全ての成分が 0 の行列をゼロ行列といい, O で表すことが一般的である.
- (ii). A の (i, i) 成分 ($i = 1, 2, \dots, n$) のことを, 対角成分という. それ以外の成分を, 非対角成分という.
- (iii). 対角成分が全て 1 で, 非対角成分が全て 0 の行列を n 次単位行列という.

$$(a) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, (b) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

などがその例である. (a) は 2 次の単位行列, (b) は 3 次の単位行列である. 単位行列は I や E で表すことが多い. このレジュメでは I と表す. また次数を明記したいときには, (a) では I_2 , (b) では I_3 のように I_n と書くこともある.

(iv). ${}^t\mathbf{A} = \mathbf{A}$ となる行列 \mathbf{A} を, 対称行列という.

$$\begin{pmatrix} 0 & 3 \\ 3 & 10 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ 3 & 10 & -4 \end{pmatrix}$$

などがその一例である. 対角部分を境目にして線対称になっていることがわかるであろう.

(v). \mathbf{A} について, $i > j$ となる (i, j) 成分が全て 0 のとき, \mathbf{A} は上三角行列であるという.

$$\begin{pmatrix} 9 & 3 \\ 0 & 10 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 0 & 9 & 10 \\ 0 & 0 & 8 \end{pmatrix}$$

などがその一例である. 対角部分より下側の成分が全て 0 であることに注目すること.

(vi). \mathbf{A} について, $i < j$ となる (i, j) 成分が全て 0 のとき, \mathbf{A} は下三角行列であるという.

$$\begin{pmatrix} 9 & 0 \\ 1 & 10 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 3 & 9 & 0 \\ 1 & 9 & 8 \end{pmatrix}$$

などがその一例である. 対角部分の上側が全て 0 であることに注目すること.

(vii). $\mathbf{A}\mathbf{A}$ を \mathbf{A}^2 で表し, 一般に \mathbf{A} を n 回掛けたものを \mathbf{A}^n で表す. そして

$$\mathbf{A}^2 = \mathbf{A}$$

となる行列を, べき等行列という. \mathbf{A} がべき等行列ならば, 任意の正整数 k について, $\mathbf{A}^k = \mathbf{A}$ となる.

(viii). 各成分が全て 0 以上で, 各行の和がそれぞれ 1 になっているものを確率行列という.

$$\begin{pmatrix} 0.1 & 0.9 \\ 0.4 & 0.6 \end{pmatrix}, \begin{pmatrix} 0 & 0.8 & 0.2 \\ 0.1 & 0 & 0.9 \\ 1 & 0 & 0 \end{pmatrix}$$

などが一例である.

上三角行列では, 対角部分やその上方の数については, 何も条件は付けていないので, その中に 0 が含まれていても構わない. 下三角行列についても同様である.

練習 2.1 $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ が, 対称行列であり, 上三角行列であり, 下三角行列であり, べき等行列であり, 確率行列であることを確かめよ. 」

この練習問題から単位行列 \mathbf{I} はかなり重要であることが示唆される. 実際次の命題が重要である.

命題 2.2 \mathbf{I} が n 次単位行列であるとする. \mathbf{A} は m 行 n 列, \mathbf{B} は n 行 k 列, \mathbf{C} は n 次正方行列であるとする. このとき以下が成り立つ.

$$\begin{aligned} \mathbf{A}\mathbf{I} &= \mathbf{A} \\ \mathbf{I}\mathbf{B} &= \mathbf{B} \\ \mathbf{C}\mathbf{I} &= \mathbf{I}\mathbf{C} = \mathbf{C} \end{aligned}$$

」

証明は簡単なのでしない。単位行列は積に関して、実数における $a \cdot 1 = 1 \cdot a = a$ のように、あたかも実数の 1 のような働きをする。ところで、実数 $a \neq 0$ のとき逆数を a^{-1} で表したから、それと同様に、逆行列というものを定義しよう。

定義 2.3 A は n 次正方行列とし、 I は n 次単位行列とする。このとき、

$$AB = BA = I$$

を満たす n 次正方行列 B が存在するならば、それを A の逆行列といい、 A^{-1} で表す。 」

実数 a について逆数は存在しない場合 ($a = 0$) もあれば、存在する場合 ($a \neq 0$) もあったが、存在する場合は必ず 1 つだった。逆行列も同様である。

命題 2.4 n 次正方行列 A が逆行列を持つならば、それは 1 つのみである。 」

証明 A の逆行列が B と C の 2 つあったと仮定する。

$$B = BI = BAC = (BA)C = IC = C$$

が成り立つので、 $B = C$ となる。 □

また逆行列をもつ行列を、正則行列 (regular matrix) といったり、可逆 (invertible) であるといったりする。逆行列は、実数 x, y でいうところの、

$$xy = yx = 1 \implies y = x^{-1}$$

の類似である。一方で実数の性質で

$$「x = 0, \text{ または } y = 0」 \iff xy = 0$$

というものがあるが、この類似は、行列の積ではできない。

命題 2.5 $O_{m,n}$ は m 行 n 列のゼロ行列とする。 m 行 n 列の行列 A と、 n 行 k 列の行列 B について、

(i). $A = O_{m,n}$ または $B = O_{n,k} \implies AB = O_{m,k}$

(ii). $A = O_{m,n}$ または $B = O_{n,k} \not\Leftarrow AB = O_{m,k}$

証明 (i) は明らかなので、省略する。(ii) については \Leftarrow が成り立たない例を挙げればよい。

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

とすると、 $AB = O_{2,2}$ が成り立つ。 □

特にこの (ii) は最初に学んだときには、すぐ忘れやすいので要注意である。さて逆行列に戻ろう。2 つの正則な行列の積も、正則である。

命題 2.6 n 次正方行列 A, B がともに正則ならば, AB も正則であり,

$$(AB)^{-1} = A^{-1}B^{-1}$$

証明 仮定より

$$\begin{aligned} (AB)B^{-1}A^{-1} &= A(BB^{-1})A \\ &= A^{-1}A \\ &= I \end{aligned}$$

が成り立つ. 同様に, $ABB^{-1}A^{-1} = I$ も成り立つ. よって,

$$(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$$

なので, 逆行列の定義より, AB の逆行列は $B^{-1}A^{-1}$ である. つまり

$$(AB)^{-1} = B^{-1}A^{-1}$$

□

この式と似たようなものは, 行列の転置行列にもある. 単位行列以外であれば, 次の計算ができるようにしておいてほしい.

命題 2.7 A は m 行 n 列, B は n 行 k 列とする. このとき以下が成り立つ.

$${}^t(AB) = {}^tB {}^tA$$

転置行列と逆行列の関係では, 次のものが成り立つ.

命題 2.8 n 次正方行列 A が正則であるとき, 以下が成り立つ.

$$({}^tA)^{-1} = {}^t(A^{-1})$$

証明 逆行列の定義より,

$$AA^{-1} = A^{-1}A = I$$

これを全て転置させると, 前命題より

$${}^t(A^{-1}) {}^tA = {}^tA {}^t(A^{-1}) = I$$

この式は, tA の逆行列が ${}^t(A^{-1})$ であることを表している. つまり,

$$({}^tA)^{-1} = {}^t(A^{-1})$$

□

逆行列と転置行列を組み合わせたようなものは, 以下の直交行列である.

定義 2.9 n 次実正方行列 \mathbf{P} が以下を満たすとき, \mathbf{P} を直交行列という.

$${}^t\mathbf{P}\mathbf{P} = \mathbf{P} {}^t\mathbf{P} = \mathbf{I} \quad (2.1)$$

」

「実」とつけているので, \mathbf{P} の成分は実数のみを考えている. また逆関数の定義から, 直交行列の逆関数は必ず存在し, $\mathbf{P}^{-1} = {}^t\mathbf{P}$ である.

では逆行列その他について, 具体的な計算を 2×2 の場合について述べよう.

第 II 部

主に 2×2 の行列までの計算

第 1 部で、行列の積や逆行列について定義した。ここでは主に 2×2 の行列までについて、いろいろなことを説明する。ただし、8 章と 13 章では、一般の n 次正方行列について、いくつか定義している。

3 2 次正方行列の行列式と逆行列

まず中学校の学んだ x, y についての連立方程式

$$\begin{cases} ax + by = e \\ cx + dy = f \end{cases}$$

を考えてみよう。第 1 式を d 倍し、第 2 式を b 倍すると、

$$\begin{cases} adx + bdy = de \\ bcx + bdy = bf \end{cases}$$

第 1 式から第 2 式を辺々引くと、

$$(ad - bc)x = de - bf$$

が得られる。ここで、 $ad - bc \neq 0$ であれば、

$$x = \frac{de - bf}{ad - bc}$$

となり、これを第 1 式に代入して計算すると、

$$y = \frac{af - ce}{ad - bc}$$

が得られる。つまり $ad - bc \neq 0$ のときは、 x, y が解を 1 組のみ持つことと言えるのである。この $ad - bc$ に注目して、逆行列を考えることにしよう。

上で書いたことをまとめると、 $ad - bc \neq 0$ とするとき、

$$\begin{cases} ax + by = e \\ cx + dy = f \end{cases} \iff \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix} \quad (3.1)$$

$$\iff \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} de - bf \\ af - ce \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} e \\ f \end{pmatrix} \quad (3.2)$$

(3.1) 式での \iff は、 $ad - bc$ の値にかかわらず成り立つ。2 番目の、(3.1) 式と (3.2) 式を結ぶ \iff は、 $ad - bc \neq 0$ という条件があるため、成り立っている。

さて、(3.1) 式の右側と、(3.2) 式の最後の項を見れば、以下が簡単に言える。

命題 3.1 2 次正方行列 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ について、 $ad - bc \neq 0$ のとき逆行列 \mathbf{A}^{-1} が存在し、

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

で表される。 J

証明 以下の2つの式,

$$\mathbf{A} \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \mathbf{A}$$

をそれぞれ計算すると、ともに $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ となる。よって逆行列の定義から命題が成り立つ。 \square

(3.1) 式や (3.2) 式をみてまとめておこう。実数 a, b, c, d, e, f について, x, y に関する連立方程式

$$\begin{cases} ax + by = e \\ cx + dy = f \end{cases}$$

については,

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} e \\ f \end{pmatrix} \quad (3.3)$$

とすれば, 連立方程式を,

$$\mathbf{Ax} = \mathbf{b}$$

とおくことができ, さらに $ad - bc \neq 0$ のときは, この式の両辺に \mathbf{A}^{-1} を左から掛けることによって,

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

とできる。これで第1部の最初に書いたことが達成されるのである。

練習 3.2 次の連立方程式を, 逆行列を用いて求めよ。ただし a, b, θ は実数とする。

$$(1) \begin{cases} 2x + 3y = 10 \\ 4x + 5y = 18 \end{cases}, \quad (2) \begin{cases} x + 2y = 12 \\ 3x + 5y = 28 \end{cases}, \quad (3) \begin{cases} (a+1)x + by = 1 \\ bx - (a+1)y = 2 \end{cases}, \quad (4) \begin{cases} (\cos \theta)x - (\sin \theta)y = a \\ (\sin \theta)x + (\cos \theta)y = b \end{cases}$$

この練習問題でわかるように, 係数が実数ではなくて, 文字のときに, 逆行列のありがたみがわかるのである。

さて $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ のときに, $ad - bc \neq 0$ ならば, \mathbf{A}^{-1} を持つことを示したわけであるが, では $ad - bc = 0$ のときはどうなるのであろうか。それを示すのが以下の命題である。

命題 3.3 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ で, $ad - bc = 0$ の場合は, \mathbf{A} は逆行列を持たない。 \square

証明 (i) まず, ある実数 s, t (ただし $s^2 + t^2 \neq 0$) が存在して

$$s \begin{pmatrix} a \\ c \end{pmatrix} + t \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (3.4)$$

が成り立つことを, 以下(ア)(イ)(ウ)で場合分けして証明する。

(ア) $b \neq 0$ の場合: $s = -1, t = a/b$ とすると, (3.4) 式が成り立つ。

(イ) $b = 0$ かつ $a = 0$ の場合: $sc + td = 0$ を満たす s, t (ただし $s^2 + t^2 \neq 0$) が必ず存在する。このような s, t をとることによって, (3.4) 式が成り立つ。

(ウ) $b = 0$ かつ $d = 0$ の場合: $s = 0, t = 1$ をとると, (3.4) 式が成り立つ。

(ii) (3.4) 式を行列で表すと,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (3.5)$$

である. ここで仮に \mathbf{A} が逆行列を持つと仮定しよう. すると (3.5) 式の両辺に, \mathbf{A}^{-1} を左側から掛けることによって,

$$\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

となる. これは s, t が $s = t = 0$ 以外に存在するという (i) の結果に矛盾する. よって, \mathbf{A} は逆行列を持たない.

□

このように $ad - bc$ の値が非 0 か 0 かによって, 逆行列を持つか, 持たないかが判別される. ある意味判別式となっているのである. そこで, この $ad - bc$ の値を 2×2 の場合の行列式と定義する.

定義 3.4 2次正方行列 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ において, $ad - bc$ を \mathbf{A} の行列式といい,

$$\det(\mathbf{A}), \quad \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

などで表す.

□

\det は determinant の略である. 行列式の役割は, 逆行列を持つか否かの判別式であること以外にも存在する. どのようなものかは, いずれ説明することになるであろう.

4 2次正方行列のクラメールの公式

さて (3.1) 式と (3.2) 式を見てみよう. $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ で $\det(\mathbf{A}) \neq 0$ のとき, x, y は以下のような行列式の比で表すことができる.

$$x = \frac{\det \begin{pmatrix} e & b \\ f & d \end{pmatrix}}{\det(\mathbf{A})}, \quad y = \frac{\det \begin{pmatrix} a & e \\ c & f \end{pmatrix}}{\det(\mathbf{A})}$$

となっていることを確認してほしい. つまり連立方程式を行列化したもの

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

($\mathbf{A}, \mathbf{x}, \mathbf{b}$ は (3.3) 式と同じとする) について, $\det(\mathbf{A}) = ad - bc \neq 0$ ならば, $i = 1, 2$ について

$$\mathbf{x} \text{ の第 } i \text{ 成分} = \frac{\det(\mathbf{A} \text{ の } i \text{ 列目を } \mathbf{b} \text{ に変えた行列})}{\det(\mathbf{A})}$$

であらわされる. これを 2×2 の場合のクラメールの公式という. 3×3 や, $n \times n$ の場合も後に示す.

文字のある連立方程式や, あとは 0 や 1 が多い場合の連立方程式はこのクラメールの公式を使うと簡単に解くことができる.

練習 4.1 以下の x, y に関する連立方程式を解け. (1) の a, b は $a^2 + b^2 > 0$ とする. (2) の a, b は $a^2 \neq b^2$ とする.

$$(1). \begin{cases} ax + by = 1 \\ -b + ay = 0 \end{cases}, (2). \begin{cases} ax + by = 0 \\ x + ay = 1 \end{cases}$$

」

5 2 次正方行列の行列式の性質

積の行列式は, 行列式の積で表される.

定理 5.1 \mathbf{A}, \mathbf{B} ともに, 2 次正方行列とする. このとき以下が成り立つ.

- (i). $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$
- (ii). $\det({}^t \mathbf{A}) = \det(\mathbf{A})$

」

この証明は直接計算しても得られるのだが, 以下の定理 5.2 を性質を用いて, 後に証明する.

さて, この章でのこれからの性質は, 行列式の一般的な性質を示すための準備である. よって, この章の残りは軽く読み飛ばして次の章を読んでも構わない.

以下の性質については, 定理 5.1 よりももっと簡単に計算できるので, 証明はしない.

定理 5.2 (2 次正方行列の場合の行列式の基本的な性質) $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ とする. c はある数とする.

- (i). 列を入れ替えた行列式は, もとの行列の行列式の -1 倍となる. 行を入れ替えた行列式は, もとの行列の行列式の -1 倍となる. つまり,

$$\det \begin{pmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{pmatrix} = (-1) \det(\mathbf{A})$$

$$\det \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix} = (-1) \det(\mathbf{A})$$

- (ii). 1 列目と 2 列目が等しいとき, 行列式は 0 である. 1 行目と 2 行目が等しいとき, 行列式は 0 である.
- (iii). ある列を c 倍した行列の行列式は, 元の行列の行列式の c 倍となる. ある行を c 倍した行列の行列式は, 元の行列の行列式の c 倍となる. つまり,

$$\det \begin{pmatrix} ca_{11} & a_{12} \\ ca_{21} & a_{22} \end{pmatrix} = \det \begin{pmatrix} a_{11} & ca_{12} \\ a_{21} & ca_{22} \end{pmatrix} = c \det(\mathbf{A})$$

$$\det \begin{pmatrix} ca_{11} & ca_{12} \\ a_{21} & a_{22} \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} \\ ca_{21} & ca_{22} \end{pmatrix} = c \det(\mathbf{A})$$

- (iv). \mathbf{A} のある列が全て 0 ならば, $\det(\mathbf{A}) = 0$ である. 同様に, \mathbf{A} のある行が全て 0 ならば, $\det(\mathbf{A}) = 0$ である.
- (v). $\mathbf{a}_1, \mathbf{a}_2, \mathbf{b}, \mathbf{c}$ を 2 項列ベクトルとする. $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{pmatrix}$ と表した時,

$$\det(\mathbf{b} + \mathbf{c} \quad \mathbf{a}_2) = \det(\mathbf{b} \quad \mathbf{a}_2) + \det(\mathbf{c} \quad \mathbf{a}_2), \quad \det(\mathbf{a}_1 \quad \mathbf{b} + \mathbf{c}) = \det(\mathbf{a}_1 \quad \mathbf{b}) + \det(\mathbf{a}_1 \quad \mathbf{c})$$

$\mathbf{A} = \begin{pmatrix} {}^t\mathbf{a}_1 \\ {}^t\mathbf{a}_2 \end{pmatrix}$ と表した時,

$$\det \begin{pmatrix} {}^t\mathbf{b} + {}^t\mathbf{c} \\ {}^t\mathbf{a}_2 \end{pmatrix} = \det \begin{pmatrix} {}^t\mathbf{b} \\ {}^t\mathbf{a}_2 \end{pmatrix} + \det \begin{pmatrix} {}^t\mathbf{c} \\ {}^t\mathbf{a}_2 \end{pmatrix}, \quad \det \begin{pmatrix} {}^t\mathbf{a}_1 \\ {}^t\mathbf{b} + {}^t\mathbf{c} \end{pmatrix} = \det \begin{pmatrix} {}^t\mathbf{a}_1 \\ {}^t\mathbf{b} \end{pmatrix} + \det \begin{pmatrix} {}^t\mathbf{a}_1 \\ {}^t\mathbf{c} \end{pmatrix}$$

(vi). \mathbf{A} のある列に, 別の列を加えた行列の行列式は, $\det(\mathbf{A})$ のまま変わらない. 同様に, ある行に, 別を加えた行列の行列式は, もとの行列の行列式 $\det(\mathbf{A})$ とは変わらない. 式で書くと,

$$\det \begin{pmatrix} a_{11} + a_{12} & a_{12} \\ a_{21} + a_{22} & a_{22} \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} + a_{11} \\ a_{21} & a_{22} + a_{21} \end{pmatrix} = \det(\mathbf{A})$$

$$\det \begin{pmatrix} a_{11} + a_{21} & a_{12} + a_{22} \\ a_{21} & a_{22} \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} + a_{11} & a_{22} + a_{12} \end{pmatrix} = \det(\mathbf{A})$$

この定理 5.2 の (i)~(iv) より, または, 直接行列式を計算することによって, 以下は簡単に言える.

系 5.3 \mathbf{A} は 2 次正方行列とする.

- (i). \mathbf{A} のある行が, 別の行の定数倍ならば, $\det(\mathbf{A}) = 0$ である.
- (ii). \mathbf{A} のある列が, 別の列の定数倍ならば, $\det(\mathbf{A}) = 0$ である.

定理 5.2 の (v) をもう少し一般化しよう.

定理 5.4 \mathbf{A} は, 定理 5.2 の行列とする.

- (i). \mathbf{A} の i 行 ($i = 1, 2$) に, 行ベクトル $\mathbf{b} = (b_1 \ b_2)$ を加えたものの行列の行列式は, それぞれ以下のよ
うに表される.

$$\det \begin{pmatrix} a_{11} + b_1 & a_{12} + b_2 \\ a_{21} & a_{22} \end{pmatrix} = \det(\mathbf{A}) + \det \begin{pmatrix} b_1 & b_2 \\ a_{21} & a_{22} \end{pmatrix}$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} + b_1 & a_{22} + b_2 \end{pmatrix} = \det(\mathbf{A}) + \det \begin{pmatrix} a_{11} & a_{12} \\ b_1 & b_2 \end{pmatrix}$$

- (ii). \mathbf{A} の i 列 ($i = 1, 2$) に, ある列ベクトル $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ を加えたものの行列の行列式は, それぞれ以下のよ
うに表される.

$$\det \begin{pmatrix} a_{11} + b_1 & a_{12} \\ a_{21} + b_2 & a_{22} \end{pmatrix} = \det(\mathbf{A}) + \det \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}$$

$$\det \begin{pmatrix} a_{11} & a_{12} + b_1 \\ a_{21} & a_{22} + b_2 \end{pmatrix} = \det(\mathbf{A}) + \det \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}$$

証明は, 単に計算するだけで得られるので, 省略する. この定理 5.4 と系 5.3 から, 以下がいえる.

系 5.5 \mathbf{A} は 2 次正方行列とし, c はある数とする. このとき以下が成り立つ.

- (i). ある行に, 別の行を c 倍したものを加えたものの行列式は, もとの行列の行列式と変わらない.
(ii). ある列に, 別の列を c 倍したものを加えたものの行列式は, もとの行列の行列式と変わらない.

┘

証明 (i) について考える. 定理 5.2 の \mathbf{A} について, その 1 行目に, 2 行目の c 倍を加えた行列を考えると,

$$\begin{aligned} \det \begin{pmatrix} a_{11} + ca_{21} & a_{12} + ca_{22} \\ a_{21} & a_{22} \end{pmatrix} &= \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \det \begin{pmatrix} ca_{21} & ca_{22} \\ a_{21} & a_{22} \end{pmatrix} \\ &= \det(\mathbf{A}) + c \det \begin{pmatrix} a_{21} & a_{22} \\ a_{21} & a_{22} \end{pmatrix} \\ &= \det(\mathbf{A}) \end{aligned}$$

2 行目の場合も同様. (ii) の場合も同様に得られる. □

もちろん直接,

$$\begin{aligned} \det \begin{pmatrix} a_{11} + ca_{21} & a_{12} + ca_{22} \\ a_{21} & a_{22} \end{pmatrix} &= (a_{11} + ca_{21})a_{22} - (a_{12} + ca_{22})a_{21} \\ &= a_{11}a_{22} + ca_{21}a_{22} - a_{22}a_{21} - ca_{22}a_{21} \\ &= a_{11}a_{22} - a_{12}a_{21} \\ &= \det(\mathbf{A}) \end{aligned}$$

と計算してもよい. しかし 3 次以上の正方行列の場合を考えて, 直接計算しないですむ方法をとった.

いよいよこの性質を使って, 定理 5.1 を証明しよう. 以下の証明での各行の式変形において, 定理 5.4 のどの性質を使ったかを, 確認するとよい.

定理 5.1 の証明:

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

とすると,

$$\mathbf{AB} = (\mathbf{a}_1 \quad \mathbf{a}_2) \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = (b_{11}\mathbf{a}_1 + b_{21}\mathbf{a}_2 \quad b_{12}\mathbf{a}_1 + b_{22}\mathbf{a}_2)$$

と表すことができる (上の式の最後の等号が成り立つことは次章を参照). よって,

$$\begin{aligned} \det(\mathbf{AB}) &= \det(b_{11}\mathbf{a}_1 + b_{21}\mathbf{a}_2 \quad b_{12}\mathbf{a}_1 + b_{22}\mathbf{a}_2) \\ &= \det(b_{11}\mathbf{a}_1 \quad b_{12}\mathbf{a}_1 + b_{22}\mathbf{a}_2) + \det(b_{21}\mathbf{a}_2 \quad b_{12}\mathbf{a}_1 + b_{22}\mathbf{a}_2) \\ &= b_{11} \det(\mathbf{a}_1 \quad b_{12}\mathbf{a}_1 + b_{22}\mathbf{a}_2) + b_{21} \det(\mathbf{a}_2 \quad b_{12}\mathbf{a}_1 + b_{22}\mathbf{a}_2) \\ &= \{b_{11}b_{12} \det(\mathbf{a}_1 \quad \mathbf{a}_1) + b_{11}b_{22} \det(\mathbf{a}_1 \quad \mathbf{a}_2)\} + \{b_{21}b_{12} \det(\mathbf{a}_2 \quad \mathbf{a}_1) + b_{21}b_{22} \det(\mathbf{a}_2 \quad \mathbf{a}_2)\} \\ &= b_{11}b_{22} \det(\mathbf{a}_1 \quad \mathbf{a}_2) + b_{21}b_{12} \det(\mathbf{a}_2 \quad \mathbf{a}_1) \\ &= b_{11}b_{22} \det(\mathbf{a}_1 \quad \mathbf{a}_2) - b_{21}b_{12} \det(\mathbf{a}_1 \quad \mathbf{a}_2) \\ &= (b_{11}b_{22} - b_{21}b_{12}) \det(\mathbf{a}_1 \quad \mathbf{a}_2) \\ &= \det(\mathbf{B}) \cdot \det(\mathbf{A}) \end{aligned}$$

□

6 2×2 までの行列の積を，分割した行列で積で表す

前頁の定理 5.1 の証明では，行列 A を 2 つの列ベクトルに分割してして考えた．それを一般的によく考えるのがこの章である．ではまず 2 次正方行列の積

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \quad (6.1)$$

についてここで考えよう．この左辺を，

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix}$$

とする．すると，(6.1) 式は，

$$A (\mathbf{b}_1 \quad \mathbf{b}_2) \quad (6.2)$$

と表すことができる．この式はあたかも 1 行 1 列の行列と，1 行 2 列の行列の積のように見える．1 行 1 列 \times 1 行 2 列だと，1 行 2 列の行列ができるので，上のように行列を分割した場合，

$$A (\mathbf{b}_1 \quad \mathbf{b}_2) = (A\mathbf{b}_1 \quad A\mathbf{b}_2) \quad (6.3)$$

が成り立っていると便利である．そして実際に成り立っていることは，

$$A\mathbf{b}_1 = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} \\ a_{21}b_{11} + a_{22}b_{21} \end{pmatrix}, \quad A\mathbf{b}_2 = \begin{pmatrix} a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

を横に並べたものが，(6.1) 式の右辺と一致していることからわかる．

では次に (6.1) 式の行列の積を，それぞれ 2 つのベクトルで表してみよう．以下の 4 通りの積が考えられる．太字は 2 項列ベクトルである．

$$\begin{aligned} & \text{(i)} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix} (\mathbf{b}_1 \quad \mathbf{b}_2), \quad \text{(ii)} (\mathbf{a}_1 \quad \mathbf{a}_2) \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} \\ & \text{(iii)} (\mathbf{a}_1 \quad \mathbf{a}_2) (\mathbf{b}_1 \quad \mathbf{b}_2), \quad \text{(iv)} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} \end{aligned}$$

(i) は，2 行 1 列と 1 行 2 列の行列の積に見え，(ii) は，1 行 2 列と 2 行 1 列の行列の積に見え，それぞれ計算可能である．それぞれ 2 行 2 列と 1 行 1 列となる．それぞれ書気表すと，

$$\text{(i)} \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix} (\mathbf{b}_1 \quad \mathbf{b}_2) = \begin{pmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 \end{pmatrix}, \quad \text{(ii)} (\mathbf{a}_1 \quad \mathbf{a}_2) \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \mathbf{a}_1 \mathbf{b}_1 + \mathbf{a}_2 \mathbf{b}_2$$

となり，計算すれば，(6.1) 式の右辺と一致する．

しかし (iii) は 1 行 2 列 \times 1 行 2 列，(iv) は 2 行 1 列 \times 2 行 1 列のように見え，すると行列の積の計算の定義から計算不可能である．このように行列を分割するならば，分割した後に，積が計算できるように分割しなければならないことがわかる．

(6.1) 式の左辺の計算について，他の分割方法としては，定理 5.1 の証明でも書いたように，

$$\text{(v)} (\mathbf{a}_1 \quad \mathbf{a}_2) \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = (\mathbf{a}_1 b_{11} + \mathbf{a}_2 b_{21} \quad \mathbf{a}_1 b_{12} + \mathbf{a}_2 b_{22})$$

$$(vi) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \begin{pmatrix} a_{11} \mathbf{b}_1 + a_{12} \mathbf{b}_2 \\ a_{21} \mathbf{b}_1 + a_{22} \mathbf{b}_2 \end{pmatrix}$$

があげられる。一応説明すると、(v) は 1 行 2 列 × 2 行 2 列で、1 行 2 列のものが得られる。(vi) は 2 行 2 列 × 2 行 1 列なので、2 行 1 列のものができる。実際にそれぞれの右辺を計算すれば、(6.1) 式の右辺と一致するので、この表記が正しいことがわかる。

また (v) で $b_{11} = \lambda_1, b_{12} = b_{21} = 0, b_{22} = \lambda_2$ の場合で、左右逆にした表記

$$(\lambda_1 \mathbf{a}_1 \quad \lambda_2 \mathbf{a}_2) = (\mathbf{a}_1 \quad \mathbf{a}_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad (6.4)$$

が成り立つ。これは 10 章で多用する。

今までの 2 次正方行列の積の説明から、2 次正方行列と 2 行 1 列のベクトルの積は、

$$\begin{pmatrix} {}^t \mathbf{a}_1 \\ {}^t \mathbf{a}_2 \end{pmatrix} \mathbf{b} = \begin{pmatrix} {}^t \mathbf{a}_1 \mathbf{b} \\ {}^t \mathbf{a}_2 \mathbf{b} \end{pmatrix}, \quad (\mathbf{a}_1 \quad \mathbf{a}_2) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \mathbf{a}_1 b_1 + \mathbf{a}_2 b_2$$

などのように表せる。1 行 2 列と 2 次正方行列の積も、うまく分割して積が計算できる。

7 2 次正方行列の固有値と固有ベクトル

y, x, a, t を実数とする。小学校以来で、 xy 平面に $y = ax$ というグラフを何回も書いてきたことであろう。 x, y の間に $y = ax$ という関係があるとき、 x を t 倍すると、 y が at 倍となるのは明らかだろう。

さて行列も似たように、2 次正方行列 \mathbf{A} が与えられた時に、

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

となる、2 項列ベクトル \mathbf{x} (ただし $\mathbf{x} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$) と、実数 λ が存在すれば、 \mathbf{x} を t 倍したものは、 $\mathbf{A}\mathbf{x}$ を λt 倍したものと一致する。このような λ と \mathbf{x} を計算することによって、 \mathbf{A} の n 乗の計算など、いろいろな計算をすることができる。

ところで、2 項ゼロベクトルを $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ と書くのは面倒なので、この章では、 $\mathbf{0}$ と書くことにしよう。

定義 7.1 \mathbf{A} を 2 次正方行列とする。このとき、

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

を満たす実数 λ と 2 項列ベクトル $\mathbf{x} \neq \mathbf{0}$ が存在するとき、 λ を \mathbf{A} の固有値といい、 \mathbf{x} を \mathbf{A} の λ に関する固有ベクトルという。 ┘

\mathbf{A} の各成分が全て実数であるときでも、 λ が実数になるとは限らない。 λ の範囲を実数までにするのか、複素数 (複素数は第 III 部に簡単に解説してある) まで含むのかによって、固有値を持つかどうかの判断が分かれる場合が存在する。それは複素数を使った行列の章で詳述するので、今は先に進むことにしよう。

$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ とするとき、

$$\begin{aligned} \mathbf{A}\mathbf{x} = \lambda\mathbf{x} &\iff (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \\ &\iff \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned} \quad (7.1)$$

さて上の式を満たすような x_1, x_2 (ただし $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \mathbf{0}$) が存在することは

$$\det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} = 0 \quad (7.2)$$

と同値である。これを計算して λ の式にまとめた、

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0 \quad (7.3)$$

が重要なものである。

定義 7.2 $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ のとき、

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21})$$

を固有多項式といい、(7.3) 式を、 \mathbf{A} の固有方程式という。 ┘

注意: 固有値を求めたい場合は、 $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$ でも $\det(\lambda \mathbf{I} - \mathbf{A}) = 0$ でもどちらでもよい。計算自体は $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$ の方が、(7.2) 式のように、 \mathbf{A} の対角部分を λ で引くだけなので簡単に書けるし、手計算のミスも少ない。固有多項式の定義についても、2次の場合なら $\det(\lambda \mathbf{I} - \mathbf{A})$ でも $\det(\mathbf{A} - \lambda \mathbf{I})$ でも同じである。しかし2次正方行列ではなくて、 n 次正方行列の固有多項式を考えたとき、 λ^n の係数を正にしたいので、固有多項式を上のように $\det(\lambda \mathbf{I} - \mathbf{A})$ で定義する。

(注意終)

(7.3) 式の固有方程式は2次方程式だから、簡単に計算できるであろう。そこで λ が λ_1, λ_2 という実数解を持つとしよう。もし重解 (以下重複解と書く) をもつ場合は、つまり $(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21}) = 0$ の場合であるが、このときは $\lambda_1 = \lambda_2$ と考える。

次に各 λ_1, λ_2 について、固有ベクトルを求める方法を述べよう。 λ_i に関する固有ベクトルの求め方は、 λ_i をもとの (7.1) 式に代入した式

$$\begin{pmatrix} a_{11} - \lambda_i & a_{12} \\ a_{21} & a_{22} - \lambda_i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

を満たす x_1, x_2 (ただし $x_1 = x_2 = 0$ 以外のもの) を求めればよい。連立方程式として考えれば

$$\begin{cases} (a_{11} - \lambda_i)x_1 + a_{12}x_2 = 0 \\ a_{21}x_1 + (a_{22} - \lambda_i)x_2 = 0 \end{cases}$$

の2本の式がある。多くの場合には以下の (i) か (ii) の簡単な方によって、 λ_i の固有ベクトルを求めることができる。両方の条件に合わない場合は (iii) を行えばよい。

(i). $(a_{11} - \lambda_i)$ と a_{12} のどちらかが非0であれば、

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -a_{12} \\ a_{11} - \lambda_i \end{pmatrix} \text{ または } \begin{pmatrix} a_{12} \\ -a_{11} + \lambda_i \end{pmatrix}$$

が λ_i の固有ベクトル (の1つ) となる。

(ii). a_{21} と $(a_{22} - \lambda_i)$ のどちらかが非 0 であれば,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{22} - \lambda_i \\ -a_{21} \end{pmatrix} \text{ または } \begin{pmatrix} -a_{22} + \lambda_i \\ a_{21} \end{pmatrix}$$

が, λ_i の固有ベクトル (の 1 つ) となる.

(iii). (i)(ii) のどちらでもないということは, $\begin{pmatrix} a_{11} - \lambda_i & a_{12} \\ a_{21} & a_{22} - \lambda_i \end{pmatrix}$ がゼロ行列のときである. この場合は,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ や, } \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ が } \lambda_i \text{ の固有ベクトル (の 1 つ) となる.}$$

固有ベクトルの「1 つ」と書いたのは, 固有ベクトルが 1 つあれば, その c 倍 (ただし $c \neq 0$) も固有ベクトルとなるためである. Scilab をはじめとして, PC ソフトに, ある行列の固有値と固有ベクトルを求めさせると, 固有ベクトルの大きさ (ベクトルの大きさは $\sqrt{x_1^2 + x_2^2}$ のことで定義 8.1 参照のこと) を 1 にしたもの (これを「固有ベクトルの基準化」という) が返ってくる.

例題 7.3 $\mathbf{A} = \begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix}$ とするとき, \mathbf{A} の固有値と固有ベクトルを求めよ. 」

(解答) まず固有方程式を解く.

$$\begin{aligned} \lambda^2 - (2+4)\lambda + (2 \cdot 4 - 5 \cdot 3) &= 0 \\ \lambda^2 - 6\lambda - 7 &= 0 \\ (\lambda+1)(\lambda-7) &= 0 \\ \lambda &= -1, 7 \end{aligned}$$

(i). $\lambda = -1$ のときの固有ベクトルを求める.

$$\begin{pmatrix} 2 - (-1) & 3 \\ 5 & 4 - (-1) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (7.4)$$

を解くと,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$$

が 1 つの解である. よって $\begin{pmatrix} 3 \\ -3 \end{pmatrix}$ が固有ベクトルの 1 つである.

(ii). $\lambda = 7$ のときの固有ベクトルを求める.

$$\begin{pmatrix} 2 - 7 & 3 \\ 5 & 4 - 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (7.5)$$

を解くと,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

が解の 1 つなので, $\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ が固有ベクトルの 1 つである.

(解答終)

この解き方は前頁の説明の (i) 番に書いた通りに忠実に解いたわけである。しかし例えば, (7.4) 式などは,

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

と同値であるから, $\lambda = -1$ のときの固有ベクトルの 1 つを $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ としてもよい。ベクトルの大きさを 1 とするのであれば,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

としてもよい。

例題 7.4 この $A = \begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix}$ の固有値と固有ベクトルを, Scilab で求めたい。関数 spec を用いて求めよ。 」

(解答)

以下のように打ち込むことによって得られる。

```
A=[2 3; 5 4]
```

```
[P,Lam]=spec(A)
```

A,P,Lam は私が決めた変数名なので, 別に何に変えても構わない。Lam とつけたのは, 固有値行列 (ラムダ) の頭文字 3 つである。P とつけたのも, 一般に固有ベクトルを結合したものを P とすることが多いためである。

spec 関数によって, Lam は 2 行 2 列の対角行列であり, 2 つの対角要素が A の固有値となる。そして Lam の (1,1) 成分の固有値に対応する固有ベクトルの 1 つが, P の 1 列目の列ベクトルとなる。同様に Lam の (2,2) 成分の固有値に対応する固有ベクトルの 1 つが, P の 2 列目の列ベクトルとなる。

$1/\sqrt{2} = 0.7071068$ に気をつけて, Scilab で得られる $\lambda = -1$ のときの固有ベクトルが, $\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ になっているか, $\begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ になっていることを確かめよ。

(解答終)

8 実ベクトルの内積とベクトルの大きさ

題名に書いてある通り, ここではベクトルの成分を実数に限っている。ただし前章までの 2×2 ではなくて, n 項列ベクトルまでを扱う。

まず例として xy 平面上の点 $A = (a_1, a_2)$ と $B = (b_1, b_2)$ があるとする。原点 O と合わせて, 角 AOB が直角の直角三角形 OAB となる場合を考えよう。線分 AB , 線分 OA , 線分 OB の長さをそれぞれ $|AB|, |OA|, |OB|$ で表すと, 三平方の定理から,

$$|AB|^2 = |OA|^2 + |OB|^2$$

が成り立つ。この式を成分ごとに計算すると,

$$(a_1 - b_1)^2 + (a_2 - b_2)^2 = (a_1^2 + a_2^2) + (b_1^2 + b_2^2)$$

であり、これを整理すると、

$$a_1b_1 + a_2b_2 = 0$$

である。

今は座標で考えたが、位置ベクトルで同じよう考えることによって、ベクトルでも同じことが言える。今、2 項列ベクトル $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ と、 $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ を考える。そこで、 $x_1y_1 + x_2y_2 = 0$ のときこの列ベクトルは直交するという。さらにこれを n 項列ベクトルまで、拡張して以下のように定義しよう。

定義 8.1 2 つの n 項列ベクトル $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ について、

(i). 内積 $\langle \mathbf{x}, \mathbf{y} \rangle$ を、

$$\langle \mathbf{x}, \mathbf{y} \rangle = {}^t \mathbf{y} \mathbf{x} = \sum_{i=1}^n x_i y_i$$

で定義する。

(ii). ベクトル \mathbf{x} の大きさを

$$|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$$

で定義する。これを \mathbf{x} のユークリッドノルムともいう。

」

内積の値は実数である。よく (\mathbf{x}, \mathbf{y}) と小かっこを使う場合もあるが、初学の人には列ベクトルの結合と勘違いすることがあるので、「 \langle と \rangle 」を使った西洋式の表記を用いた。内積の定義によって、

$$\mathbf{x} \text{ と } \mathbf{y} \text{ が直交する} \iff \langle \mathbf{x}, \mathbf{y} \rangle = 0$$

を言いたいので、順にこれを示していく。まずは内積の性質からである。

定理 8.2 (実ベクトルの内積の性質) $\mathbf{x}, \mathbf{y}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2$ がともに n 項列ベクトルとする。 c は実数とする。このとき、

- (交換) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- (結合) $\begin{cases} \langle \mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2 \rangle = \langle \mathbf{x}, \mathbf{y}_1 \rangle + \langle \mathbf{x}, \mathbf{y}_2 \rangle \\ \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle \end{cases}$
- (定数倍) $c \langle \mathbf{x}, \mathbf{y} \rangle = \langle c\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, c\mathbf{y} \rangle$

」

ユークリッドノルムと内積の関係では、以下は簡単に得られる。

命題 8.3 \mathbf{x}, \mathbf{y} がともに n 項列ベクトルで, a, b が実数のとき,

$$|a\mathbf{x} + b\mathbf{y}|^2 = |\mathbf{x}|^2 a^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle ab + |\mathbf{y}|^2 b^2$$

」

ユークリッドノルムと内積の関係では, 三角不等式が重要な役割を果たすので, それを証明しよう.

定理 8.4 (三角不等式) \mathbf{x}, \mathbf{y} がともに n 項列ベクトルとする. このとき,

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq |\mathbf{x}| \cdot |\mathbf{y}| \quad (8.1)$$

が成り立つ. 等号が成り立つ必要十分条件は, $s\mathbf{y} = t\mathbf{x}$ となる実数 s, t (ただし $s^2 + t^2 \neq 0$) が存在する場合である.

」

証明 (i) $\mathbf{x} = \mathbf{0}$ の場合は, (8.1) 式が等号として成り立ち, $0\mathbf{y} = t\mathbf{x}$ となる $t \neq 0$ が存在することから, 命題が成り立つ.

(ii) $\mathbf{x} \neq \mathbf{0}$ の場合は, t を実数とすると以下がいえる.

$$|t\mathbf{x} - \mathbf{y}|^2 = |\mathbf{x}|^2 \cdot t^2 - 2 \cdot \langle \mathbf{x}, \mathbf{y} \rangle t + |\mathbf{y}|^2$$

ところで左辺は t にかかわらず 0 以上であるので, 右辺の t の 2 次式の判別式を D とすると $D/4 \leq 0$ が成り立つ.

$$D/4 = (\langle \mathbf{x}, \mathbf{y} \rangle)^2 - |\mathbf{x}|^2 \cdot |\mathbf{y}|^2 \leq 0$$

を変形すれば (8.1) 式が得られる.

(8.1) 式が等号で成り立つための必要十分条件は, $\mathbf{y} = t\mathbf{x}$ である. よって命題の等号条件が得られる.

□

三角不等式から以下のように 1 つの不等式と, 1 つの定義が導かれる. まず三角不等式の両辺を 2 乗すると, 以下が言える.

系 8.5 (シュワルツの不等式) n 個の実数 x_1, x_2, \dots, x_n と n 個の実数 y_1, y_2, \dots, y_n について,

$$(x_1 y_1 + x_2 y_2 + \dots + x_n y_n)^2 \leq (x_1^2 + \dots + x_n^2)(y_1^2 + \dots + y_n^2)$$

」

次にベクトルのなす角を定義しよう. (8.1) 式から, $\mathbf{x} \neq \mathbf{0}$ かつ $\mathbf{y} \neq \mathbf{0}$ ならば,

$$-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{|\mathbf{x}| \cdot |\mathbf{y}|} \leq 1$$

がいえる. 実数 θ を $0 \leq \theta \leq \pi$ とすると*3, $\cos \theta$ は -1 から 1 までの値をとり, θ と $\cos \theta$ の値は $1:1$ で対応している. よって, ベクトルのなす角を以下のように定義する.

定義 8.6 \mathbf{x}, \mathbf{y} がともに n 項列ベクトルとする.

*3 弧度法ではなくラジアンを用いている. x と y° の関係は, 「 $\frac{x}{\pi}$ ラジアン = $\frac{y}{180}$ 度」である.

(i). \mathbf{x}, \mathbf{y} ともにゼロベクトルではないときは,

$$\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{|\mathbf{x}| \cdot |\mathbf{y}|} \quad (8.2)$$

を満たす $\theta \in [0, \pi]$ を, \mathbf{x} と \mathbf{y} のなす角と定義する.

- (ii). \mathbf{x} または \mathbf{y} が 0 ベクトルのときは, \mathbf{x} と \mathbf{y} のなす角は $\pi/2$ と定義する.
 (iii). \mathbf{x} と \mathbf{y} のなす角が $\pi/2$ のとき, 2 つのベクトルは直交するといひ $\mathbf{x} \perp \mathbf{y}$ で表す.

」

9 内積までの 2 項ベクトルの応用: 平面ベクトル (2)

この章も, ベクトルや行列の成分を実数に限っている

9.1 法線ベクトルと点と直線の距離

1.3 節で xy 平面上の平面ベクトルについて学んだ. この節はその延長であり, 行列の積・行列式・内積についての平面ベクトルへの応用を考える.

まずは前節で直交ということ定義したので, 平面上の直線の法線ベクトルを定義しよう.

定義 9.1 (直線の法線ベクトルの定義) xy 平面上の直線 $ax + by + c = 0$ (ただし $a^2 + b^2 \neq 0$) が与えられたとき, この直線の法線ベクトルは, $s \neq 0$ の実数で,

$$s \begin{pmatrix} a \\ b \end{pmatrix}$$

と定義する.

」

定義 1.5 では方向ベクトルを定義した. その時は, 主な方向ベクトルとして $\begin{pmatrix} -b \\ a \end{pmatrix}, \begin{pmatrix} b \\ -a \end{pmatrix}$ のどちらにしてもよいと書いた. 今の法線ベクトルの場合も同様に

$$\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -a \\ -b \end{pmatrix}$$

が主なものである. しかし普通はマイナスを 2 つ使うことはしないので,

$$\begin{pmatrix} a \\ b \end{pmatrix}$$

を考えるとよい.

法線というのは, 直交する線という意味がある. 実際に定義 (9.1) での直線の方向ベクトルが $\begin{pmatrix} b \\ -a \end{pmatrix}$ であり, 法線ベクトル $\begin{pmatrix} a \\ b \end{pmatrix}$ との内積が 0 になることを確認してほしい.

次に, 高校の学習参考書に公式としては載ってはいるが, あまりきちんと証明されていないことを証明しよう.

命題 9.2 (点と直線の距離) xy 平面上で, 直線 $ax + by + c = 0$ と, 点 $P(x_0, y_0)$ の距離は,

$$\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$$

□

これは次の命題を先に示した後に簡単に言える.

命題 9.3 xy 平面上で, 点 $P(x_0, y_0)$ から直線 $ax + by + c = 0$ に下した垂線の脚を H とする. このとき以下が成り立つ.

$$\overrightarrow{PH} = -\frac{ax_0 + by_0 + c}{a^2 + b^2} \begin{pmatrix} a \\ b \end{pmatrix}$$

□

証明 \overrightarrow{PH} は $ax + by + c = 0$ と直交する. よって \overrightarrow{PH} は, この直線の法線ベクトルの定数倍で表される. つまり実数 t として,

$$\overrightarrow{PH} = t \begin{pmatrix} a \\ b \end{pmatrix} \quad (9.1)$$

と, 表すことができる. これを利用して,

$$\overrightarrow{OH} = \overrightarrow{OP} + \overrightarrow{PH} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + t \begin{pmatrix} a \\ b \end{pmatrix}$$

であり, H が直線 $ax + by + c = 0$ 上にあることから, この式に $x = x_0 + ta, y = y_0 + tb$ を代入して

$$\begin{aligned} a(x_0 + ta) + b(y_0 + tb) + c &= 0 \\ (a^2 + b^2)t &= -(ax_0 + by_0 + c) \\ t &= -\frac{ax_0 + by_0 + c}{a^2 + b^2} \end{aligned}$$

が成り立つ. これを (9.1) 式に代入して, 命題が成り立つ. □

命題 9.2 の証明 : 命題 9.3 の結果より,

$$\overrightarrow{PH} = -\frac{ax_0 + by_0 + c}{\sqrt{a^2 + b^2}} \cdot \frac{1}{\sqrt{a^2 + b^2}} \begin{pmatrix} a \\ b \end{pmatrix}$$

が成り立つ. 列ベクトル $\frac{1}{\sqrt{a^2 + b^2}} \begin{pmatrix} a \\ b \end{pmatrix}$ の大きさは 1 なので, 上の式の $\frac{1}{\sqrt{a^2 + b^2}} \begin{pmatrix} a \\ b \end{pmatrix}$ に掛かっている実数の絶対値, つまり,

$$\left| -\frac{ax_0 + by_0 + c}{\sqrt{a^2 + b^2}} \right| = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$$

が, P と直線 $ax + by + c = 0$ との距離である. □

9.2 平面全体を 2 つのベクトルの線形結合で表せる場合と, 表せない場合

命題 1.2 で, 平面全体を $s \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ によって一意に表した. まずはこれの応用である.

命題 9.4 xy 平面上の原点を O とする. 2 点 A, B の列ベクトルを

$$\overrightarrow{OA} = \begin{pmatrix} a \\ c \end{pmatrix}, \quad \overrightarrow{OB} = \begin{pmatrix} b \\ d \end{pmatrix}$$

とする. s, t は実数として, 以下の (i)(ii)(iii) は同値である.

(i). 任意の点 P の位置ベクトル $\overrightarrow{OP} = \begin{pmatrix} x \\ y \end{pmatrix}$ は, $s\overrightarrow{OA} + t\overrightarrow{OB}$ によって一意に表される. 成分で書くと以下である.

$$\begin{pmatrix} x \\ y \end{pmatrix} = s \begin{pmatrix} a \\ c \end{pmatrix} + t \begin{pmatrix} b \\ d \end{pmatrix} \quad (9.2)$$

(ii). $s\overrightarrow{OA} + t\overrightarrow{OB} = \mathbf{0}$ となる s, t が $s = t = 0$ のみである. 成分で書くと以下である.

$$s \begin{pmatrix} a \\ c \end{pmatrix} + t \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff s = t = 0 \quad (9.3)$$

(iii). $\det \begin{pmatrix} \overrightarrow{OA} & \overrightarrow{OB} \end{pmatrix} \neq 0$ である. 成分で書くと以下が成り立つ.

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0$$

┘

証明 (ii) \implies (iii) の証明: (9.3) 式は,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (9.4)$$

と表される. $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 0$ の場合は, (9.4) 式の \implies が成り立たない. よって, $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0$

(iii) \implies (i) の証明: 任意の x, y について, (ii) が成り立つので (iii) が成り立つことから, 任意の x, y について,

$$\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix}$$

と s, t を定めることができる. この s, t を用いて,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = s \begin{pmatrix} a \\ c \end{pmatrix} + t \begin{pmatrix} b \\ d \end{pmatrix}$$

と表すことができる. 次に一意性について証明しよう. 実数 s, t, u, v によって \overrightarrow{OP} が

$$\overrightarrow{OP} = s \begin{pmatrix} a \\ c \end{pmatrix} + t \begin{pmatrix} b \\ d \end{pmatrix} = u \begin{pmatrix} a \\ c \end{pmatrix} + v \begin{pmatrix} b \\ d \end{pmatrix}$$

のように 2 通りで示されたと仮定しよう. すると, この式の右側の等式より,

$$\begin{aligned} (s-u) \begin{pmatrix} a \\ c \end{pmatrix} + (t-v) \begin{pmatrix} b \\ d \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \iff \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} s-u \\ t-v \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

が得られる。(iii) の仮定より $\begin{pmatrix} s-u \\ t-t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ であり, これは $s = u, t = v$ を意味している。

(i) \implies (ii) の証明. 対偶で証明する。(ii) が成り立たないと仮定する。つまり, が成り立つような実数 $u, v (u^2 + v^2 \neq 0)$ が存在することを意味する。今 $u \neq 0$ としても一般性を失わないので, そのように仮定する。すると上の式は,

$$\begin{pmatrix} a \\ c \end{pmatrix} = -\frac{v}{u} \begin{pmatrix} b \\ d \end{pmatrix} \quad (9.5)$$

となるので, これを (9.2) 式に代入すると,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \left(-\frac{v}{u}s + t\right) \begin{pmatrix} b \\ d \end{pmatrix}$$

である。 $\left(-\frac{v}{u}s + t\right)$ は実数なので, s, t を任意に動かしても, 点 P は直線 OB 上にしか存在しない。つまり s, t によって xy 平面全体を表すことはできない。

□

従ってこの節の題名の答えの 1 つとして, 以下がいえる。

「2 つの非ゼロベクトル $\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}$ が,

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0$$

を満たすとき, そしてそのときのみ, 線形結合

$$s \begin{pmatrix} a \\ c \end{pmatrix} + t \begin{pmatrix} b \\ d \end{pmatrix}$$

によって, 平面全体を表すことができる。」

9.3 2 項列ベクトルの左から 2 次正方行列を掛けることと, 行列式の関係

行列 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ が与えられて, xy 上の点 $P(x, y)$ から, uv 平面上の点 $P'(u, v)$ へと,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \mathbf{A} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (9.6)$$

によって写されることを考えよう*4。行列 \mathbf{A} によって, xy 平面上の点は, uv 平面上のある 1 点に写るのだが, 前節の結果から, 以下がいえる。

命題 9.5 xy 平面から uv 平面への写りかた*5が, (9.6) 式で表されるとする。

*4 これを「 (u, v) は (x, y) の \mathbf{A} による写像である」というのだが, 写像の定義や, 線形写像と線形写像と行列の関係などの細かい話は, 参考文献を参照のこと。

*5 写像という数学用語がある。ある点について写す, とかある集合を写すという意味のイメージで写像という。正しい定義については, 詳しくは参考文献を参照のこと

- (i). \mathbf{A} が正則のときは, xy 平面上の点とそれに対応する uv 平面上の点は 1 対 1 に対応する.
(ii). $\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ のときは, xy 平面上の任意の点は, uv 平面上の原点に写される.
(iii). それ以外の場合は, xy 平面上の任意の点は, uv 上での原点を通るある直線上に写される.

証明 xy 平面上の任意の点 (x, y) は, ある実数 s, t によって,

$$\begin{pmatrix} x \\ y \end{pmatrix} = s \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

と一意にあらわされる. この式の両辺に左から \mathbf{A} を掛けると,

$$\begin{pmatrix} u \\ v \end{pmatrix} = s \begin{pmatrix} a \\ c \end{pmatrix} + t \begin{pmatrix} b \\ d \end{pmatrix}$$

となる. これに前節の命題 9.4 の結果を合わせると, 命題の (i)(ii)(iii) が全て成り立つ.

次にこの写像と, \mathbf{A} の行列式の関係を示そう. まず三角形の面積からである.

命題 9.6 原点を O とする座標平面上に, 2 点 A, B があり, 2 つの位置ベクトルをそれぞれ

$$\overrightarrow{OA} = \begin{pmatrix} a \\ c \end{pmatrix}, \quad \overrightarrow{OB} = \begin{pmatrix} b \\ d \end{pmatrix}$$

とする. このとき三角形 OAB の面積は

$$\frac{1}{2}|ad - bc|$$

であらわされる.

証明 \overrightarrow{OA} と \overrightarrow{OB} のなす角を θ とすると, 求める面積 S は,

$$\begin{aligned} S &= \frac{1}{2} |\overrightarrow{OA}| \cdot |\overrightarrow{OB}| \sin \theta = \frac{1}{2} \sqrt{a^2 + c^2} \sqrt{b^2 + d^2} \sin \theta \\ &= \frac{1}{2} \sqrt{(a^2 + c^2)(b^2 + d^2)(1 - \cos^2 \theta)} \\ &\quad \text{一方内積の定義 8.6 から,} \\ &\quad (a^2 + c^2)(b^2 + d^2) \cos^2 \theta = (ab + cd)^2 \text{ なので} \\ &= \frac{1}{2} \sqrt{(a^2 + c^2)(b^2 + d^2) - (ab + cd)^2} \\ &= \frac{1}{2} \sqrt{(ad - bc)^2} \\ &= \frac{1}{2} |ad - bc| \end{aligned}$$

この三角形の面積の結果によって, 以下がいえる.

命題 9.7 xy 平面から uv 平面への $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ による写像を

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

とする. xy 平面上の任意の 2 点 $P(x_1, y_1), Q(x_2, y_2)$ の \mathbf{A} による写像をそれぞれ P', Q' とする. このとき uv 平面上の三角形 $O'P'Q'$ の (O' は uv 平面上の原点とする) 面積は, xy 平面上の三角形 OPQ の, $|\det(\mathbf{A})|$ 倍となる. ┘

注意: $\det(\mathbf{A}) = 0$ の場合は, $O'P'Q'$ が一直線上に並ぶこととなるので, 三角形 $O'P'Q'$ の面積は 0 と考える.

(注意終)

証明 前命題より, 三角形 OPQ の面積は, $\frac{1}{2}|x_1y_2 - y_1x_2|$ である. uv 平面上での, P' の位置ベクトルと, Q' の位置ベクトルはそれぞれ

$$\begin{pmatrix} ax_1 + by_1 \\ cx_1 + dy_1 \end{pmatrix}, \quad \begin{pmatrix} ax_2 + by_2 \\ cx_2 + dy_2 \end{pmatrix}$$

である. よって, 三角形 $O'P'Q'$ の面積は,

$$\frac{1}{2}|(ax_1 + by_1)(cx_2 + dy_2) - (cx_1 + dy_1)(ax_2 + by_2)|$$

である. これを展開して整理すると, $\frac{1}{2}|x_1y_2 - y_1x_2|$ の $|ad - bc|$ 倍と一致する.

□

ということで, 写像された三角形の面積は, もとの三角形の面積の「行列式の絶対値」倍となるのである. 行列式には, 逆行列を持つか否かの判別式という意味以外にも, このような意味を持っているのである.

9.4 回転変換と三角関数の加法定理

前節では xy 平面から uv 平面への写像を考えた. この写像の中でも, 反時計回りに θ だけ回転する写像を考えよう. 写像の場合は, uv 平面に写したが, この節では同じ xy 平面上を動くことを考える. つまり, xy 平面上にある 2 点 $P(x, y)$ と点 $P'(x', y')$ があって,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

というように, 点 P から点 P' のように動くことを考えている. このような場合は写像というよりも, 変換という言葉がよく使われる.

では θ だけ反時計回りに回転する変換を考えてみよう. これは x 軸を θ だけ反時計回りに回し, y 軸を θ だけ反時計回りに回せばよいのである.

まず最初に 90° つまり $\frac{\pi}{2}$ ラジアンを反時計回りに動かす変換を考える. それは

- (i). x 軸の正の向きを, y 軸の正の向きに動かす.
- (ii). y 軸の正の向きを, x 軸の負の向きに動かす.

ということである。 xy 平面の任意の点の位置ベクトルは

$$s \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

であらわされるのだから、上の (i)(ii) をもう少し考えれば、

- (i). $(1, 0)$ を, $(0, 1)$ に移動する.
- (ii). $(0, 1)$ を, $(-1, 0)$ に移動する.

とすればよい。ということは、

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

とすればよいので、行列

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

が、反時計回りに $\pi/2$ ラジアン回転する変換を表す。

次に、反時計回りに (以下「反時計回り」を省略して書く) θ ラジアン回転する変換を考えよう。

xy 平面上の点で、

- (i). $(1, 0)$ を $(\cos \theta, \sin \theta)$ に移動させる.
- (ii). $(0, 1)$ を $(-\sin \theta, \cos \theta)$ に移動させる.

この (ii) の意味が分からなかったら (私も高校のとき最初はわからなかった), 次のように考えるとよい。

(i) によって元の x 軸が $(\cos \theta, \sin \theta)$ 方向に写った。そして元の y 軸は元の x 軸を反時計回りに $\pi/2$ ラジアン回したものである。ということは (ii) については、

$$(0, 1) \text{ を反時計回りに } \theta \text{ ラジアン回す} \iff (\cos \theta, \sin \theta) \text{ を, 反時計回りに } \pi/2 \text{ ラジアン回す}$$

と考えればよい。上の式の \iff の右側について、先の行列を用いて、 $(\cos \theta, \sin \theta)$ を、半時計回りに $\pi/2$ ラジアン回した座標は、

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

で得られる。計算して、点 $(-\sin \theta, \cos \theta)$ が得られるのである。

ということで、(i)(ii) を考え合わせれば、

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \tag{9.7}$$

という行列が、 θ ラジアン回転させる回転変換を表す行列となる。

この行列を $R(\theta)$ と表すことにしよう。ところで点 $(1, 0)$ を、 $(\alpha + \beta)$ だけ回転変換するということは、点 $(1, 0)$ を α 回転したあと、さらに β 回転することなので、

$$R(\alpha + \beta) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = R(\beta)R(\alpha) \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

が成り立つ*6. これをそれぞれ成分で書くと,

$$\begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

それぞれ計算して,

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

という $\sin(\alpha + \beta)$ と $\cos(\alpha + \beta)$ の加法定理が簡単に得られる. 加法定理からは特に倍角の公式

$$\cos 2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta = \cos^2\theta - \sin^2\theta \quad (9.8)$$

$$\sin 2\theta = 2\sin\theta\cos\theta \quad (9.9)$$

が簡単に得られる.

9.5 線対称の点に移す変換を表す行列

この節では xy 平面上の任意の点を, 原点を通る直線

$$\begin{pmatrix} x \\ y \end{pmatrix} = t \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix}$$

に対して, 線対称に移す変換を考える. この表現では直線がイメージにくいかもしれないが, 具体的には, $\cos \frac{\theta}{2} \neq 0$ ならば, 直線 $y = (\tan \frac{\theta}{2})x$ を表す. つまり, $\cos \frac{\theta}{2} \neq 0$ のときは, この直線は, 傾き $\tan \frac{\theta}{2}$ で原点を通る直線である. この線対称による変換によって,

- (i). 点 $(\cos \frac{\theta}{2}, \sin \frac{\theta}{2})$ は, そのまま変わらない.
- (ii). 点 $(-\sin \frac{\theta}{2}, \cos \frac{\theta}{2})$ は, 点 $(\sin \frac{\theta}{2}, -\cos \frac{\theta}{2})$ に写される.

となる. するとこの変換を表す行列 \mathbf{A} は,

$$\begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} = \mathbf{A} \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix}, \quad \begin{pmatrix} \sin \frac{\theta}{2} \\ -\cos \frac{\theta}{2} \end{pmatrix} = \mathbf{A} \begin{pmatrix} -\sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{pmatrix}$$

とあらわされる*7. まとめると,

$$\mathbf{A} \begin{pmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & -\cos \frac{\theta}{2} \end{pmatrix}$$

*6 この式の $R(\alpha + \beta) = R(\beta)R(\alpha)$ が成り立つことについて, 細かい部分を省略している. 気になる場合は, 例えば [川久保] や [長谷川] の該当部分を参照するとよい.

*7 この2点の座標の移動だけを考えれば十分であることも, 多少の説明や証明が必要なのだが, ここでは省略する. 興味のある人は参考文献を参照するとよい.

左辺の右側の逆行列を右から掛けて,

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & -\cos \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & -\cos \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \\ -\sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} & 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} \\ 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} & \sin^2 \frac{\theta}{2} - \cos^2 \frac{\theta}{2} \end{pmatrix} \\ &\text{倍角の公式 (9.8) 式} \cdot \text{(9.9) 式より} = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \end{aligned}$$

例えば, 直線 $y = x$ に関して線対称に移動することを表す行列を求めよう. $\theta/2 = \pi/4$ ラジアンなので, $\theta = \pi/2$ ラジアンである. (9.9) 式に代入して,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

である. つまり,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}$$

となる.

10 2次実正方行列の直交行列と三角化・対角化

8章で直交を学んだ. さらに行列まで考えることによって, 直交行列の意味を探り, さらにそれを用いて行列を変形していくのが, この章での目的である. この章も, ベクトルや行列の成分を実数に限っている.

命題 10.1 2次正方行列 \mathbf{P} の直交行列については, $p_1^2 + p_2^2 = 1$ を満たす p_1, p_2 によって,

$$\begin{pmatrix} p_1 & -p_2 \\ p_2 & p_1 \end{pmatrix}, \quad \begin{pmatrix} p_1 & p_2 \\ p_2 & -p_1 \end{pmatrix} \quad (10.1)$$

のどちらかの形で表される. $\sin \theta, \cos \theta$ の形で表すと,

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \quad (10.2)$$

注意: (10.2) 式の左は, 9.4 節の回転変換, 右は 9.5 節の線対称を表す変換の形である.

(注意終)

証明 2×2 の直交行列なので, $\mathbf{P} = \begin{pmatrix} p_1 & x_1 \\ p_2 & x_2 \end{pmatrix}$ として, p_1, p_2 を所与として

$$p_1^2 + p_2^2 = 1, x_1^2 + x_2^2 = 1, p_1 x_1 + p_2 x_2 = 0$$

を満たす x_1, x_2 を求めるとよい. $p_1 x_1 + p_2 x_2 = 0$ かつ $p_1^2 + p_2^2 \neq 0$ より, 実数 t によって,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = t \begin{pmatrix} p_2 \\ -p_1 \end{pmatrix} \quad (10.3)$$

とあらわされる. これを $x_1^2 + x_2^2 = 1$ に代入して計算すると, $(p_1^2 + p_2^2)t^2 = 1$. よって,

$$t = \pm 1$$

が得られる. これを (10.3) 式に代入して,

$$(x_1, x_2) = (-p_2, p_1), (p_2, -p_1)$$

の 2 通りのみの答えが出てくる.

□

(10.1) 式のどちらの行列についても, 1 列目の列ベクトルも 2 列目の列ベクトルも, 大きさは 1 であり, そして, その 2 つの列ベクトルは直交しているであることを確認してほしい. n 次直交行列についても同様のことがなりたつが, とりあえず 2 次の直交行列には, 以下のような性質がある.

命題 10.2 2 次正方行列 P が直交行列ならば, 以下が成り立つ.

- (i). $P^{-1} = {}^t P$
- (ii). $\det(P)$ は ± 1 のいずれかである.

┘

証明 (i) は直交行列の定義より明らかである. (ii) は, 直交行列の定義と定理 5.1 より得られる.

□

定義 10.3 A, P, Λ, U は 2 次正方行列とする. Λ は対角行列で, U は上三角行列とする. A が与えられたときに, 適当な正則行列 P が存在して

- (i). $P^{-1}AP = \Lambda$ となるとき, A は P で対角化可能である, という.
- (ii). $P^{-1}AP = U$ となるとき, A は P で三角化可能である, という.

┘

わざわざ三角化と対角化を分けているのは, 全ての 2 次正方行列が, 対角化できるわけではないためである. 三角化はその代理であり, より広いものである. まずは対角化できるものから求めてみよう.

定理 10.4 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ とする. A の固有方程式が異なる 2 つの実数解 λ_1, λ_2 を持つならば, A は, 適当な P を選ぶことによって,

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

のように対角化可能である.

┘

証明 固有方程式の異なる 2 つの実数解を λ_1, λ_2 とし, λ_1 の固有ベクトルの 1 つを x , λ_2 の固有ベクトルの 1 つを y とする. さて,

$$Ax = \lambda_1 x, Ay = \lambda_2 y,$$

が成り立つことから、 \mathbf{x} と \mathbf{y} を横に並べた行列 $\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}$ を用いて、

$$\mathbf{A} \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

がいえる。あとは、2 次正方行列 $\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}$ が正則であることを示せばよい。というのもこの行列を \mathbf{P} とすれば、 \mathbf{P}^{-1} を上の式の両辺の左側から掛けることによって、証明が完成するのである。そこで $\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}$ が正則であることを背理法で示そう。

$\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}$ が正則ではないと仮定する。それは

$$\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (10.4)$$

となる s, t (ただし $s = t = 0$ 以外) が存在する。上の式を簡単にすると、

$$s\mathbf{x} + t\mathbf{y} = \mathbf{0} \quad (10.5)$$

である。左から \mathbf{A} を掛けると、 \mathbf{x}, \mathbf{y} がそれぞれ \mathbf{A} の固有ベクトルであることから、

$$s\lambda_1\mathbf{x} + t\lambda_2\mathbf{y} = \mathbf{0}$$

がなりたつ。(10.5) 式の両辺を λ_1 倍して、上の式から引くと、

$$(\lambda_2 - \lambda_1)t\mathbf{y} = \mathbf{0}$$

$\lambda_1 \neq \lambda_2$ かつ $\mathbf{y} \neq \mathbf{0}$ なので、これより $t = 0$ である。 $t = 0$ を (10.5) 式に代入すると、

$$s\mathbf{x} = \mathbf{0}$$

であるが、 s, t の仮定より $s \neq 0$ であるから $\mathbf{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ となる。これは \mathbf{x} が固有ベクトルであることに矛盾する。よって、 \mathbf{x} と \mathbf{y} を横に並べた行列

$$\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}$$

は正則である。

□

例題 10.5 練習 7.3 の行列 \mathbf{A} について、その固有値と固有ベクトルの結果を用いることによって、 \mathbf{A} を対角化せよ。

(解答) 解答の結果から、固有値と固有ベクトルの関係から以下が成り立つ。

$$\mathbf{A} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (-1) \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{A} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = 7 \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

これを一つにまとめると、

$$\mathbf{A} \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 7 \end{pmatrix}$$

$\begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}$ は逆行列を持つので (行列式が非 0 となることを確かめよ), この逆行列を, 両辺の左側から掛けて

$$\begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}^{-1} \mathbf{A} \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 7 \end{pmatrix}$$

(解答終)

この対角化でのポイントを書いてみよう. \mathbf{A} が 2 次正方行列としたときに, λ_1, λ_2 (ただし $\lambda_1 \neq \lambda_2$) が実数の固有値で, \mathbf{x}, \mathbf{y} がそれぞれの固有ベクトルとする. このとき,

$$\begin{cases} \mathbf{Ax} = \lambda_1 \mathbf{x} \\ \mathbf{Ay} = \lambda_2 \mathbf{y} \end{cases} \iff \mathbf{A} \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

となることが, ポイントなのである*8.

なお重複解を持つ場合については, 対角化可能な場合も対角化不可能な場合もある. これ以上細かくは書かないが, 固有方程式が重複解を持つ場合には, 対角化可能か否かについて, よく考えなければならないということを押さえておくとよいだろう.

これに対して三角化は, 実数解で重複解があったとしても, 三角化できる. しかも直交行列によって三角化できる点が素晴らしいのである.

定理 10.6 \mathbf{A} は任意の 2 次正方行列とする. \mathbf{A} の固有方程式が実数解を持ちそれを λ_1, λ_2 とするとき (重複解の場合は $\lambda_1 = \lambda_2$ とする), ある直交行列 \mathbf{P} が存在して,

$$\mathbf{P}^{-1} \mathbf{AP} = \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix}$$

のように, \mathbf{A} を三角化することができる. 」

注意: 実数 u はどのような数字になってもよい. とにかく上三角行列になればよい. そこで, 証明の方針として, \mathbf{A} の固有値の一つを λ_1 とし, その固有ベクトルを $\begin{pmatrix} x \\ y \end{pmatrix}$ とする ($x^2 + y^2 = 1$ と基準化しておく). よって, この固有ベクトルを用いて, (10.1) 式の直交行列の形を参考にして,

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}^{-1} \mathbf{A} \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix}$$

が成り立つようにすればよい.

(注意終)

証明 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ とし, この固有値を λ_1, λ_2 とする. また λ_1 の固有ベクトルを $\begin{pmatrix} x \\ y \end{pmatrix}$ とし, すでに基準化しているものとする. 以上より,

$$ax + by = \lambda_1 x \tag{10.6}$$

$$cx + dy = \lambda_1 y \tag{10.7}$$

$$x^2 + y^2 = 1 \tag{10.8}$$

$$\lambda_1 + \lambda_2 = a + d \tag{10.9}$$

*8 以外と盲点であり, 慣れるためには手計算を何回も繰り返した覚えがある.

が成り立つ。(10.9) 式が成り立つのは, \mathbf{A} の固有方程式の解と係数の関係からである. さて,

$$\mathbf{P} = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \quad (10.10)$$

とすると, \mathbf{P} は明らかに直交行列である. この \mathbf{P} を用いて,

$$\begin{aligned} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} &= \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \\ (10.6) \text{ 式と } (10.7) \text{ 式より} &= \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \begin{pmatrix} \lambda_1 x & -ay + bx \\ \lambda_1 y & -cy + dx \end{pmatrix} \\ (10.8) \text{ 式より} &= \begin{pmatrix} \lambda_1 & bx^2 + (-a + d)xy - cy^2 \\ 0 & dx^2 + (-b - c)xy + ay^2 \end{pmatrix} \end{aligned} \quad (10.11)$$

ここで (10.6) 式の両辺に x を, (10.7) 式の両辺に y を, それぞれ掛けて移項することによって,

$$bxy = \lambda_1 x^2 - ax^2, \quad cxy = \lambda_1 y^2 - dy^2 \quad (10.12)$$

が成り立つので, (10.11) 式の (2, 2) 成分に代入して,

$$\begin{aligned} (10.11) \text{ 式の } (2, 2) \text{ 成分} &= dx^2 - bxy - cxy + ay^2 \\ &= dx^2 + (-\lambda_1 x^2 + ax^2) + (-\lambda_1 y^2 + dy^2) + ay^2 \\ (10.8) \text{ 式より} &= a + d - \lambda_1 \\ (10.9) \text{ 式より} &= \lambda_2 \end{aligned}$$

これと (10.11) 式から, 命題が成り立つ. □

なお証明には直接は関係ないので示さなかったが, (10.11) 式の (1, 2) 成分も $b - c$ となることが, (10.6) 式・(10.7) 式から, (10.12) 式と似たようなものを作ることによっていえる.

三角化は理論上では極めて重要であるが, 実際に数字が与えられて直交行列によって三角化することはあまりない.

11 2 次実正方行列の対角化の応用例

固有値と固有ベクトルと対角化については, いろいろと応用ができる. この章では \mathbf{A}^n について述べよう. この章も, ベクトルや行列の成分を実数に限っている.

命題 11.1 2 次正方行列 \mathbf{A} が \mathbf{P} によって, $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ と対角化できた場合 (λ_1, λ_2 は実数),

$$\mathbf{P}^{-1}\mathbf{A}^n\mathbf{P} = \mathbf{\Lambda}^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$$

」

注意: この P は直交行列である必要はない. A を対角化できればよい.

(注意終)

証明 右の等号は明らかなので, 左の等号について証明する. $\Lambda = (P^{-1}AP)$ の両辺を n 回掛けることにより,

$$\begin{aligned}\Lambda^n &= (P^{-1}AP)^n = (P^{-1}AP)(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)(P^{-1}AP) \\ &= P^{-1}A[(PP^{-1})A(PP^{-1})A\cdots(PP^{-1})A]P \\ &= P^{-1}(AA\cdots A)P \\ &= P^{-1}A^nP\end{aligned}$$

□

系 11.2 $A, P, \Lambda, \lambda_1, \lambda_2$ を命題 11.1 と同様としたとき, 任意の n について,

$$A^n = P \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} P^{-1}$$

」

微分積分の本などで, 実数列の収束について, $|a| < 1$ ならば $\lim_{n \rightarrow \infty} a^n = 0$ などを学んでいるだろう. それを用いると, 以下がいえる.

系 11.3 $A, P, \Lambda, \lambda_1, \lambda_2$ を命題 11.1 と同様としたとき,

(i). $|\lambda_1| < 1, |\lambda_2| < 1$ ならば,

$$\lim_{n \rightarrow \infty} A^n = O$$

(ii). $\lambda_1 = 1, |\lambda_2| < 1$ ならば,

$$\lim_{n \rightarrow \infty} A^n = P \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} P^{-1}$$

」

練習 11.4 (2 状態マルコフ過程の基礎) $0 < p_1 < 1, 0 < p_2 < 1$ として, $q_1 = 1 - p_1, q_2 = 1 - p_2$ とする. 全ての成分が正である確率行列 $A = \begin{pmatrix} p_1 & q_1 \\ q_2 & p_2 \end{pmatrix}$ を考える. このとき,

(i). A の固有値が 1 と $p_1 + p_2 - 1$ であることをし, $|p_1 + p_2 - 1| < 1$ を確かめよ.

(ii). それぞれの固有ベクトルを求めた後, 以下を示せ.

$$\lim_{n \rightarrow \infty} A^n = \frac{1}{q_1 + q_2} \begin{pmatrix} q_2 & q_1 \\ q_2 & q_1 \end{pmatrix}$$

」

12 2 次実対称行列の性質

この章も, ベクトルや行列の成分を実数に限っている.

実対称行列は, C^2 級の多変数の微分の場合にも出てくるので, 以下の 2 つは非常に重要な性質である.

定理 12.1 A が 2 次対称行列とする。このとき、

- (i). 固有値は全て実数である.
- (ii). A は直交行列で対角化できる.

」

注意: A が n 次対称行列でもなりたつ.

(注意終)

証明 (i) の証明: 対称行列なので, 実数 a, b, d で $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ とすると, 固有方程式は

$$\lambda^2 - (a + d)\lambda + (ad - b^2) = 0$$

この λ の 2 次方程式の判別式は

$$\begin{aligned} D &= (a + d)^2 - 4(ad - b^2) \\ &= (a^2 + 2ad + d^2) - 4ad + 4b^2 \\ &= (a^2 - 2ad + d^2) + 4b^2 \\ &= (a - d)^2 + 4b^2 \\ &\geq 0 \end{aligned}$$

より, λ は実数解のみをもつ.

(ii) の証明: 定理 10.6 と今の (i) から, A は, ある直交行列 P によって,

$$P^{-1}AP = U \tag{12.1}$$

と三角化できる (U は上三角行列). 直交行列なので $P^{-1} = {}^tP$ を代入して

$${}^tPAP = U \tag{12.2}$$

この左辺の転置行列は,

$${}^t({}^tPAP) = {}^tP^tA({}^tP) = {}^tPAP$$

であるので, (12.2) 式の両辺を転置すると,

$${}^tPAP = {}^tU \tag{12.3}$$

(12.2) 式と, (12.3) 式から,

$$U = {}^tU$$

が成り立つ. この左辺は上三角行列で, 右辺は下三角行列である. 上三角行列と下三角行列が等しいということは, それは非対角成分が 0 であることとなる. つまり U は対角行列である. よって (12.1) 式と合わせて, 命題がいえる.

□

13 n 次実正方行列による, 2 次形式の定義

この章も, ベクトルや行列の成分を実数に限っている. x についての 2 次式は, 実定数 a, b, c を用いて, 高校の時は $ax^2 + bx + c$ と書いた. この b の代わりに $2b$ と書いて

$$ax^2 + 2bx^2 + c \quad (13.1)$$

としても一般性は失わない. これをさらに一般化して x_1 と x_2 の 2 次式

$$ax_1^2 + 2bx_1x_2 + cx_2^2$$

を考えてみよう. この式を, x_1, x_2 の 2 元 2 次形式という. さらに対称行列を使って, 以下のように表すことができる.

$$ax_1^2 + 2bx_1x_2 + cx_2^2 = (x_1 \ x_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (13.2)$$

$\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ とすると, ${}^t\mathbf{x}\mathbf{A}\mathbf{x}$ で表される.

n 変数の場合でも同様なので, ここで x_1, x_2, \dots, x_n について定義しておこう.

定義 13.1 a_{ij} (i, j は $1 \leq i \leq j \leq n$) を実定数とする. n 個の実数 x_1, x_2, \dots, x_n の関数

$$Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_ix_j \quad (13.3)$$

を x_1, \dots, x_n に関する n 元 2 次形式, または単に 2 次形式と呼ぶ. このとき n 次対称行列 \mathbf{A} と n 項列ベクトル \mathbf{x} を

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

とすると, $Q(x_1, x_2, \dots, x_n) = {}^t\mathbf{x}\mathbf{A}\mathbf{x}$ と表される. よってこの対称行列 \mathbf{A} を, 2 次形式 $Q(x_1, \dots, x_n)$ の係数行列という. ┘

a, x を実数とすると, 中学や高校で学んだことから,

- (i). $a > 0$ とする. このとき任意の $x \neq 0$ について $ax^2 > 0$ が成り立つ.
- (ii). $a < 0$ とする. このとき任意の $x \neq 0$ について $ax^2 < 0$ が成り立つ.
- (iii). $a \geq 0$ とする. このとき任意の $x \neq 0$ について $ax^2 \geq 0$ が成り立つ.
- (iv). $a \leq 0$ とする. このとき任意の $x \neq 0$ について $ax^2 \leq 0$ が成り立つ.

は簡単にわかるであろう. 2 次形式についても, 同じようにして, 以下のように定義する.

定義 13.2 n 次対称行列 \mathbf{A} と n 項列ベクトル \mathbf{x} を考える.

- (i). 任意の $\mathbf{x} \neq \mathbf{0}$ について ${}^t\mathbf{x}\mathbf{A}\mathbf{x} > 0$ が成り立つとき, \mathbf{A} を正定定符号行列という. 正定定符号, または正定値ともいう.

- (ii). 任意の $\mathbf{x} \neq \mathbf{0}$ について ${}^t\mathbf{x}\mathbf{A}\mathbf{x} \geq 0$ が成り立つとき, \mathbf{A} を半正値定符号行列という. 半正値定符号, または半正定値ともいう.
- (iii). 任意の $\mathbf{x} \neq \mathbf{0}$ について ${}^t\mathbf{x}\mathbf{A}\mathbf{x} < 0$ が成り立つとき, \mathbf{A} を負値定符号行列という. 負値定符号, または負定値ともいう.
- (iv). 任意の $\mathbf{x} \neq \mathbf{0}$ について ${}^t\mathbf{x}\mathbf{A}\mathbf{x} \leq 0$ が成り立つとき, \mathbf{A} を半負値定符号行列という. 半負値定符号, または半負定値ともいう.

」

このように 2 次形式と定符号というものを定義したので, 次章では 2 次正方行列を中心に説明していく.

14 2 次実正方行列と 2 次形式

この章も, ベクトルや行列の成分を実数に限っている.

2 元 2 次形式についてこの章で学ぶことによって, C^2 級の 2 変数実数値関数 $f(x, y)$ についての極大・極小についての議論が, 十分に理解できるようになる.

さて (13.1) 式の正負については, $ax^2 + 2bx + c = 0$ の判別式を D として, $D/4 = b^2 - ac$ を用いて高校の時以下のように学んだことであろう. x は実数全体を考える.

- (i). $a > 0$ かつ $D/4 < 0$ ならば, 任意の x について $ax^2 + 2bx + c > 0$.
- (ii). $a > 0$ かつ $D/4 \leq 0$ ならば, 任意の x について $ax^2 + 2bx + c \geq 0$.
- (iii). $a < 0$ かつ $D/4 < 0$ ならば, 任意の x について $ax^2 + 2bx + c < 0$.
- (iv). $a < 0$ かつ $D/4 \leq 0$ ならば, 任意の x について $ax^2 + 2bx + c \leq 0$.
- (v). $a \neq 0$ かつ $D/4 > 0$ のときは, x の値によって, $ax^2 + 2bx + c$ は, 正, 負, 0 のいずれの値もとりうる.

2 次形式の場合も同様のことが言える. x_1, x_2 の 2 次形式

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \mathbf{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

として ($\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ である), 以下考えてみよう. 上で判別式 D を用いたが, それをよく見ると実は,

$D/4 = -\det(\mathbf{A})$ である. また, $x_1 = x_2 = 0$ のときは常に $ax_1^2 + 2bx_1x_2 + cx_2^2 = 0$ だから, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

の場合を考える. 以下は高校のときの平方完成の復習である.

- $a \neq 0$ ならば,

$$ax_1^2 + 2bx_1x_2 + cx_2^2 = a \left(x_1 + \frac{b}{a}x_2 \right)^2 + \frac{-b^2 + ac}{a}x_2^2 = a \left(x_1 + \frac{b}{a}x_2 \right)^2 + \frac{\det(\mathbf{A})}{a}x_2^2 \quad (14.1)$$

- $c \neq 0$ ならば,

$$ax_1^2 + 2bx_1x_2 + cx_2^2 = c \left(x_2 + \frac{b}{c}x_1 \right)^2 + \frac{-b^2 + ac}{c}x_1^2 = c \left(x_2 + \frac{b}{c}x_1 \right)^2 + \frac{\det(\mathbf{A})}{c}x_1^2$$

- $a = c = 0$ の場合だと,

$$ax_1^2 + 2bx_1x_2 + cx_2^2 = 2bx_1x_2$$

さてこれらと, $\det(\mathbf{A}) > 0 \iff b^2 < ac$ であることと合わせて, 以下がいえる.

命題 14.1 $\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ とし, \mathbf{x} は 2 項列ベクトルであるとする.

- (i). $a > 0$ かつ $\det(\mathbf{A}) > 0$ であることは, 任意の $\mathbf{x} \neq \mathbf{0}$ について ${}^t\mathbf{x}\mathbf{A}\mathbf{x} > 0$ であることと, 同値である.
- (ii). $a \geq 0$ かつ $c \geq 0$ かつ $\det(\mathbf{A}) \geq 0$ であることは, 任意の $\mathbf{x} \neq \mathbf{0}$ について ${}^t\mathbf{x}\mathbf{A}\mathbf{x} \geq 0$ であることと, 同値である.

証明 (i) の

$$\lceil a > 0 \text{ かつ } \det(\mathbf{A}) > 0 \implies \text{任意の } \mathbf{x} \neq \mathbf{0} \text{ について } {}^t\mathbf{x}\mathbf{A}\mathbf{x} > 0 \rceil$$

は (14.1) 式から簡単に言える. よって \Leftarrow を対偶で証明しよう.

(1) $a \leq 0$ を仮定する. このときある $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ が存在して, ${}^t\mathbf{x}\mathbf{A}\mathbf{x} = a \leq 0$ となる.

(2) $a > 0$ かつ $\det(\mathbf{A}) \leq 0$ を仮定する. (14.1) 式を見ると, ある $\mathbf{x} = \begin{pmatrix} -b \\ a \end{pmatrix}$ が存在して, ${}^t\mathbf{x}\mathbf{A}\mathbf{x} = \frac{\det(\mathbf{A})}{a}a^2 \leq 0$ がいえる.

(1)(2) より, $a \leq 0$ または $\det(\mathbf{A}) \leq 0$ ならば, ある $\mathbf{x} \neq \mathbf{0}$ があって, ${}^t\mathbf{x}\mathbf{A}\mathbf{x} \leq 0$ となる. よって対偶が証明された.

(ii) については自力で挑戦してみるとよい.

□

以下は簡単である.

命題 14.2 2 次対称行列 \mathbf{A} について,

- (i). $-\mathbf{A}$ が正値定符号であることは, \mathbf{A} が負値定符号であることと同値である.
- (ii). $-\mathbf{A}$ が半正値定符号であることは, \mathbf{A} が半負値定符号であることと同値である.

さて命題 14.1 から, 正値定符号と負値定符号についての, 必要十分条件を簡単に示すことができる.

定理 14.3 2 次対称行列 $\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ について,

- (i). $a > 0$ かつ $\det(\mathbf{A}) > 0$ であることは, \mathbf{A} が正値定符号であることと, 同値である.
- (ii). $a < 0$ かつ $\det(\mathbf{A}) > 0$ であることは, \mathbf{A} が負値定符号であることと, 同値である.

それに比べると, 半正値定符号と半負値定符号はやや面倒である.

定理 14.4 2 次対称行列 $\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ について,

- (i). $a \geq 0$ かつ $c \geq 0$ かつ $\det(\mathbf{A}) \geq 0$ であることは, \mathbf{A} が半正值定符号であることと, 同値である.
- (ii). $a \leq 0$ かつ $c \leq 0$ かつ $\det(\mathbf{A}) \geq 0$ であることは, \mathbf{A} が半負値定符号であることと, 同値である.

」

よって以下は簡単に言える.

系 14.5 2 次対称行列 $\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ について, $ac < 0$ ならば, \mathbf{A} は, 半正值定符号でもなければ, 半負値定符号でもない.

」

定理 14.4 のときに, c の条件を明示しているが, 定理 14.3 のときには明示していない. それは, 定理 14.3 の (i) の場合で考えると, $a > 0, \det(\mathbf{A}) = ac - b^2 > 0$ より, $c > 0$ が自動的にいえるためである. 定理 14.3 については, a の代わりに c を代入しても成り立つ.

練習 14.6 以下の行列を, 正值定符号・負値定符号・半正值定符号・半負値定符号・いずれでもない, の 5 通りで区分せよ.

$$(i) \begin{pmatrix} 1 & 3 \\ 3 & 4 \end{pmatrix}, (ii) \begin{pmatrix} 1 & 2 \\ 2 & -4 \end{pmatrix}, (iii) \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}, (iv) \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, (v) \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}$$

」

15 第 II 部の Scilab 実習

ここでは II 部の復習として Scilab の実習をまとめて行う. Scilab でいろいろなことができるが, このレジュメでは,

- (i). メイン画面に直接コマンドを打ち込むという, 対話形式.
- (ii). 自作関数 (ユーザー定義関数) を作っておき, .sce ファイルで保存した後, それを実行して関数を読み込ませる. その後でメイン画面で, 関数を用いる.
- (iii). .sce ファイルで保存して, 実行するバッチ形式.

の 3 つを主に行う. この章では (i) のみである.

15.1 [] を用いて行列とベクトルの作成. 和と積

例題 15.1 次のコマンドを, メイン画面に順に打ち込め. もちろん行末ごとに Enter キーを押すこと (Enter キーについては以下の実習でも同様).

```
a=[1 2]
b=[3;4]
c=10
A=[50 60;70 80]
```

```
B=[-1 -2;-3 -4]
```

```
c*a
```

```
a*b
```

```
A+B
```

```
A*B
```

```
A+c
```

```
a+b
```

```
D=[b 10*b]
```

ここでは、以下に気づいてほしい。

- (i). 変数の大文字と小文字は区別される。
- (ii). 行列やベクトルを作るときには [] でくくる。
- (iii). 行ベクトルの区切りは空白で、列ベクトルの区切りは:である。
- (iv). 2次正方行列 A, B を作るときには、行ベクトルを2つ並べて作るイメージである。
- (v). 答えは `ans` であらわされる。
- (vi). `A+c` では行列とスカラーの計算もしている。
- (vii). `a+b` ではエラーメッセージがでてくる。
- (viii). `D=[b 10*b]` のように、2行1列の行列を、横に結合して、新しい行列を作ることができる。

(v) の `ans` は、もちろん `answer` の略である。(vi) は、要注意である。行列 A とスカラー c の和は、行列 A の各成分すべてに c を加えることと、Scilab では定義している。ただし慣れると、この表現はとても楽である。(vii) は、行列のサイズが異なるので計算不能のためである。

15.2 実行列の転置行列を返す' と、sqrt 関数・norm 関数

例題 15.2 次のコマンドを打ち込め。clear

```
a=[1;2]
```

```
b=[1;%pi]
```

```
c=a'
```

```
d=b'
```

```
a'*a
```

```
b'*b
```

```
sqrt(a'*a)
```

```
norm(a)
```

ここでは、まず表題の転置行列について理解してほしい^{*9}。この「`'`」はよく使う。また「`'`」以外では、以下に気づいてほしい。

- (i). `clear` というものがある。
- (ii). `%pi` で、定数 π を表している。
- (iii). `sqrt` 関数と `norm` 関数というものがある。

`clear` は、Scilab 内にある変数や読み込んだ自作関数を、全て消去するというコマンドである。

`sqrt` 関数は、普通の PC ソフトでは平方根を表す。実数の場合も同様にして $a \geq 0$ のときに、 \sqrt{a} を返される。 a の代わりに行列を入れると、成分ごとに計算した値が返ってくる。

次に `norm` 関数を見てみよう。実数の列ベクトルの大きさを、ユークリッドノルムで定義した。 \mathbf{a} が実ベクトルのとき、 $\sqrt{\mathbf{a}'\mathbf{a}}$ が、 \mathbf{a} のユークリッドノルムである。そこで上の実習で `norm(a)` の値と、 $\mathbf{a}' * \mathbf{a}$ の平方根が等しいことから、`norm(a)` のノルムを返すことがわかるであろう。

例題 15.3 次のコマンドを順に打ち込み、`sqrt` 関数について理解せよ。

```
clear
A=[1 2;0 25]
sqrt(A)
```

」

15.3 行列式 `det` 関数と逆行列 `inv` 関数

Scilab では、行列式や逆行列を返す組み込み関数が用意されている。

例題 15.4 次のコマンドを打ち込め。

```
clear
A=[1 2;3 4];
B=[1 3;2 6];
C=[10^12 2;3 4]
det(A)
det(B)
det(C)
inv(A)
inv(B)
inv(C)
```

」

ここでは、まず表題の `det` 関数と `inv` 関数について理解してほしい。行列式は英語で *determinant* であり、逆行列は英語で *inverse* だから、それぞれその略語である。これ以外では、以下を気づいてほしい。

^{*9} 次の第 III 部で複素数まで習った場合には、「`'`」は随伴行列を返すコマンドであることがわかる。今は実数の場合であり、その場合だと転置行列を返すこととなる。

- (i). 行末に ; を付けると、画面で表示されない.
- (ii). 逆行列が存在しない場合には、エラーメッセージが表示される.
- (iii). 10^12 は、10 の 12 乗を表している. つまり「^」は累乗を表す.
- (iv). メイン画面で D+12 とか D-12 などが出ている.
- (v). C の逆行列は存在しているが、何か警告が発されている.

(iv) の D+12 の D は、10 を表す. そして、 $D+12 = 10^{12}$ を意味し、D-12 は 10^{-12} を表す. (v) 理論と数値計算との違いである. 行列式が 0 に近い時には、逆行列の数値計算は誤差が大きくなっていくことに気をつけること.

15.4 固有値と固有ベクトルを返す spec 関数と 10 章の実習

Scilab では、固有値と固有ベクトルを返す組み込み関数 spec が用意されている.

例題 15.5 次のコマンドを打ち込め.

```
clear
A=[2 3;5 4]
[Lvec,L]=spec(A)
inv(Lvec)*A*Lvec
```

この例題は、実習 7.3 の場合そのものである. さて spec 関数は値が 2 つ返ってくる. A が 2 次正方行列のときに、3 行目のように

$$[Lvec,L]=spec(A)$$

とすると、L は 2×2 の対角行列で、対角部分には固有値が返ってくる. Lvec も 2×2 の行列が返ってくる. Lvec の 1 列目は、A の固有値が L の (1,1) 成分であるときの、基準化された固有ベクトルである. Lvec の 2 列目は、A の固有値が L の (2,2) 成分のときの、基準化された固有ベクトルである.

よって、上の実習のように A が実行列でかつ、固有値も 2 つの異なる実数の場合だと、A は Lvec という行列によって、L に対角化される (定理 10.4 参照のこと). それを表したのが、上の 4 行目である. D-15 は 10 の -15 乗である. それは計算誤差であり、実際は 0 であるから、4 行目が L と同じになっていることが理解できるであろう.

では以下の実習で、A の固有値が、実数の重複解を持つ場合の固有ベクトルを求めてみよ.

練習 15.6 以下の 2 つの A について、その固有値と固有ベクトルを、Scilab の Spec 関数で求めよ.

$$(i). \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix}, \quad (ii). \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}$$

(i) の場合の 2 つの固有ベクトルを横に並べた行列は、正則であるが、(ii) の場合は非正則である. 非正則の場合は、上の例題のような対角化はできない. しかし直交行列による三角化はできる (定理 10.6 参照).

練習 15.7 上の練習の (ii) の A について、直交行列 P を、以下のように求めて、三角化せよ.

```
clear
A=[0 0;2 0];
[V,L]=spec(A)
P=[V(:,1) [-V(2,1);V(1,1)]]
inv(P)*A*P
```

$V(:,1)$ という表現については、次の節を参照すること。

15.5 行列の全成分を表す「:」の利用

Scilab で行列 $A = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 2 \\ 4 & 3 & 9 \end{pmatrix}$ としたときに、 A で、(2,3) 成分を b に代入したいときは、

$$b=A(2,3)$$

などと入力すればよい。また i (i は数字) 行目全体のベクトルを b としたいときは、

$$b=A(i,:)$$

でよい。同様に j (j は数字) 列目全体のベクトルを b としたいときは、

$$b=A(:,j)$$

とすればよい。似たように A の最初の 2 行 2 列の行列がほしいときは、

$$b=A(1:2,1:2)$$

とすれば得られる。

15.6 spec 関数と 12 章の実習

A が 2 次実対称行列であれば、直交行列 P で対角化されることを、数値例で、簡単に見てみよう。

例題 15.8 次のコマンドを打ち込め。

```
clear
A=[10 3;3 2];
[P,L]=spec(A)
P'-inv(P)
inv(P)*L*P-A
```

4 行目から P の転置行列 P' が、 P の逆行列と等しいことがわかり、5 行目から、 P によって A が、 L に対角化されることがわかるであろう。

15.7 よく使われる関数 (1): eye,zeros,ones

単位行列 I や, ゼロ O は今までの説明でもよく使った.

例題 15.9 次のコマンドを打ち込め.

```
clear
A=eye(4,4)
B=zeros(3,2)
C=ones(1,5)
```

このレジュメで, 単位行列を I で表した理由は, 実はこのコマンド `eye` を使いたいためだったのである. 先に書いたように日本の教科書では E を用いることが多い.

15.8 よく使われる関数 (2): sum,mean

行列の成分の和や平均を表すこの2つの関数は, 統計計算その他でよく使われるので, ここで紹介しておこう.

例題 15.10 次のコマンドを打ち込め.

```
clear
A=[1 2 3;4 5 0]
sum(A)
sum(A,1)
sum(A,2)
mean(A)
mean(A,1)
mean(A,2)
```

`sum(A)`, `mean(A)` のように第1引数のみであれば, 全体の和や平均を返す. 第2引数が1のときは, 列ごとの和や平均を返す. 第2引数が2のときは, 行ごとの和や平均を返す.

15.9 成分ごとの掛け算・割り算・べき乗: 「.*」「./」「.^」の利用

行列の掛け算は `*` であった. しかし, 行列の掛け算ではなく, 同じ成分で行列のごとに掛け算をしたい場合もある. そのときには題名のように, 「.`」` を利用する.

例題 15.11 次のコマンドを打ち込め.

```
clear
A=[1 0; 3 4]
```


$B = [1 \ 2; 10 \ 100]$

$A * B$

$A ./ B$

$A.^2$

$A.^B$

4 行目は $A * B$ と, 6 行目は A^2 と間違えないように.

15.10 等差数列を返す行ベクトル: 「:」の利用

Scilab で $a = (1 \ 2 \ 3 \ 4 \ 5 \ 6)$ がほしい時に,

$a = [1 \ 2 \ 3 \ 4 \ 5 \ 6]$

と打ち込むのは手間がかかる. 6 ぐらいならまだしも, 1 から 100 までなどは到底無理である. そこで表題のようなコマンドが用意されており, 上の a の式は

$a = [1:6]$

で簡単に書ける. また, 差分が 1 ではないときには,

最初の値:差分:最後の値

の形式で書けばよい. $b = (-2 \ 0 \ 2 \ 4 \ 6 \ 8 \ 10)$ がほしい時には,

$b = [-2:2:10]$

で得られる. 列ベクトルにしたいときは, 最後に 「'」 を付ければよい.

第 III 部

主に 2×2 までの複素行列の計算

16 複素数入門

16.1 複素数の定義と四則

a, b, c を実数とする以下の x の 2 次方程式 (ただし $a \neq 0$)

$$ax^2 + bx + c = 0 \quad (16.1)$$

については,

- (i). $b^2 - 4ac > 0$ ならば, x は 2 つの異なる実数解 $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ を持つ.
- (ii). $b^2 - 4ac = 0$ ならば, x は実数の重複解 $x = -\frac{b}{2a}$ を持つ.
- (iii). $b^2 - 4ac < 0$ ならば, x は実数解を持たない.

としてきたことであろう. もちろんこれら (i)(ii)(iii) は正しいことを言っている. が, (iii) の「実数解を持たない」ではなくて, 実数を少し拡張したいと考えるのは自然なことであろう.

数というものはかつては正の整数・正の有理数あたりまでしか定義されなかった. やがて 0, 負の数, 無理数という概念を獲得して, 少しずつ数というものが拡張されて来たのである. そこで, さらに実数を以下のように拡張することとなったのである.

まず上の 2 次方程式の基本的な形で, $a = 1, b = 0, c = -1$ にすると,

$$x^2 = -1 \quad (16.2)$$

である. これを満たす (実数ではない) 解の一方を

i

で表す^{*10*11}. その結果 (16.2) 式のもう一方の解は $-i$ である. そして, 数の集合

$$\{a + bi \mid a, b \text{ は実数}\} \quad (16.3)$$

を, 複素数と呼び, その集合を C で表すのである. 実数の集合は R と表す. R は (16.3) 式の $b = 0$ の場合だから,

$$R \subset C$$

つまり, 実数の集合は複素数の集合に含まれる.

四則演算などについては, 以下のように定義する.

定義 16.1 a, b, c, d は実数とする. 2 つの複素数 $z = a + bi, w = c + di$ について, 以下の計算が成り立つ.

$$(i). \quad z + w = w + z = (a + c) + (b + d)i$$

^{*10} 文字としての i を多用する場合は, $\sqrt{-1}$ と使う場合もある. 数論などの分野などでも $\sqrt{-5}$ のように使うことが多い.

^{*11} 行列での添字で i をよく用いている. そこで混乱を防ぐためにこのレジュメでは, $\sqrt{-1}$ を太字の \mathbf{i} で表している. i と \mathbf{i} が違うことに注意せよ.

$$(ii). zw = wz = (a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

(iii). $w \neq 0$ のとき,

$$\frac{z}{w} = \frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{(ac + bd) + (-ad + bc)i}{c^2 + d^2}$$

$$(iv). b > 0 \text{ のとき, } \sqrt{-b} = \sqrt{b} i$$

」

この計算の定義より, 複素数 z_1, z_2, z_3, z, w について, 以下のことがいえる.

(i). 以下の分配法則が成り立つ.

$$(z_1 + z_2)z_3 = z_3(z_1 + z_2) = z_1z_3 + z_2z_3$$

(ii). z が与えられたとき, $z + w = w + z = 0$ を満たす w が唯一存在する.

(iii). $1 \cdot z = z \cdot 1 = z$

(iv). $z \neq 0$ が与えられたとき, $zw = wz = 1$ を満たす w が唯一存在する. そのような w を z^{-1} で表す.

複素数 $a + bi$ で $b = 0$ のものを, 実数であることはすでに示した. 複素数の中で特別なものである. よって他にも,

- 複素数 $a + bi$ で $b \neq 0$ のものを, 虚数という.
- 複素数 $a + bi$ で $a = 0, b \neq 0$ のものを, 純虚数という.

と定義する. もちろん「純虚数の集合 \subset 虚数の集合 $\subset \mathbf{C}$ 」がいえる.

16.2 実数での数直線と複素数での複素平面

小学校以来, 実数では数直線を用いて, その 1 点と実数を対応させてきた. しかし複素数で $a + bi$ ($b \neq 0$) という形のは, その直線上にはもちろん存在しない. そこで普通は「直線がだめなら平面で考えればよい」と考えるだろうし, 実際そのように平面上の 1 点と複素数の値を対応させることをここで説明しよう.

複素数 $a + bi$ を考えるときには, 実数の部分 a を表すものを横軸 (これを実軸という) に, i の係数の実数 b を表すものを縦軸 (これを虚軸という) にとった複素平面というものを定義する. 数直線に右端と同様に, 横軸の右端は \rightarrow , 縦軸の上端は \uparrow のように, 方向の矢印は書いておく. しかし xy 平面で書くような, x や y といったものは書かない. 実際, 小学生の時に書いた数直線にも, 数字は書いたとしても文字などは何も書かなかったことであろう.

そうしてできた平面の座標で, 横軸が a で縦軸が b の点 (a, b) と, 複素数 $a + bi$ を対応させるのである. こうすることにより, 複素平面上の 1 点と, 1 つの複素数が対応する のである. また実数の数直線も, この平面の横軸と一致するので, 実数は複素数の一部として表現される.

数直線上の点と, 実数全体が 1 対 1 に対応していたのと同様に, 複素平面上の点と, 複素数全体とが 1 対 1 に対応するのである.

ところで xy 平面上で, $r > 0$ の定数としたとき, $x^2 + y^2 = r^2$ というのは, 半径 r の円を表した. また単位円 $x^2 + y^2 = 1$ 上の任意の 1 点は, $0 \leq \theta < 2\pi$ の範囲で, $(\cos \theta, \sin \theta)$ と表された.

この 2 つから, 0 以外の複素数は $a + bi$ という形以外の形式で書ける. 0 以外の複素数は, 複素平面上の 1

点と考えると, $r > 0$ と実数 θ で

$$a + bi = r \cos \theta + i \sin \theta \quad (16.4)$$

と表現できる. このように複素数を 2 つの実数 r と θ で表す形式を極形式という. θ も範囲を例えば $0 \leq \theta < 2\pi$ や, $-\pi < \theta \leq \pi$ のように決めておけば, 原点以外の複素平面上の点と r, θ の組は, 1 対 1 に対応する. 当然このとき, (a, b) と (r, θ) は 1 対 1 の関係である.

極形式は, 複素数の掛け算に有用である. 例えば $z = (a + bi)$ とするとき ($z \neq 0$ とする), 三角関数の倍角の公式より,

$$z^2 = (a + bi)^2 = r^2(\cos \theta + i \sin \theta)^2 \quad (16.5)$$

$$= r^2(\cos^2 \theta + 2i \cos \theta \sin \theta - \sin^2 \theta) \quad (16.6)$$

$$= r^2(\cos 2\theta + i \sin 2\theta) \quad (16.7)$$

が言える. これより $r = 1$ として, 数学的帰納法から以下が簡単に言える.

定理 16.2 n は整数, θ は実数とすると, 以下が成り立つ.

$$(\cos \theta + i \sin \theta)^n = \cos(n\theta) + i \sin(n\theta)$$

」

また $z = a + bi$ が与えられたとき, $a - bi$ を z の共役複素数 (きょうやくふくそすう) といい, \bar{z} で表す. よって, 複素数 z について

$$z = \bar{z} \iff z \text{ は実数}$$

である. 共役複素数については, 次のように, 和と積についての計算が成り立つ.

$$(i). \overline{z + w} = \bar{z} + \bar{w}$$

$$(ii). \overline{zw} = \bar{z}\bar{w}$$

実数と複素数の違いに関しては, 特に複素数の大小は定義されないことが大きな違いである. 実数については, 2 つの実数 a, b が存在すると,

$$a < b, a = b, a > b$$

の 3 つのうち 1 つのみが必ず成り立った. しかし複素数ではそのようなことはできないのである. それでは多少困ることがあるので, 次節で複素数 z の絶対値 $|z|$ を定義したい.

16.3 複素平面と複素数の絶対値の定義

複素平面で $z = a + bi$ を考えよう. z の表す点と原点の距離は $\sqrt{a^2 + b^2}$ となる. これを z の絶対値とするのは自然であろう. さらに実数 a があつた時, $|a| = \sqrt{a^2 + 0^2}$ となり, 実数の絶対値の場合も含まれるのである. ところで $z = a + bi$ ならば, $\bar{z} = a - bi$ であり,

$$a^2 + b^2 = (a + bi)(a - bi) = z\bar{z}$$

が成り立つから,

定義 16.3 複素数 z が与えられたとき, 絶対値 $|z|$ を

$$|z| = \sqrt{z\bar{z}}$$

と定義する. 実数 a, b によって, 複素数 $z = a + bi$ が与えられたとき, 以下が成り立つ.

$$|z| = \sqrt{a^2 + b^2}$$

$|z|$ は (0 以上の) 実数なので, 大小関係が計算される.

定理 16.4 複素数 z, w について, 以下が成り立つ.

- (i). (非負性) $|z| \geq 0$
- (ii). $|z - w| = 0 \iff z = w$
- (iii). (対称性) $|z - w| = |w - z|$
- (iv). (三角不等式) $|z + w| \leq |z| + |w|$

16.4 * 複素数係数の 2 次方程式の解と係数の関係

ここでは, 複素数係数の 2 次方程式の解と係数の関係が, 実数係数の場合とまったく同じになることを示す. (16.1) 式で, $b^2 - 4ac < 0$ の場合の一方の解は,

$$x = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{-b - \sqrt{4ac - b^2}i}{2a}$$

とあらわされ, もう一方の解は,

$$x = \frac{-b - \sqrt{b^2 - 4ac}}{2a} = \frac{-b + \sqrt{4ac - b^2}i}{2a}$$

で表される. つまり (16.1) 式で, $b^2 - 4ac < 0$ の場合は, x は 2 つの異なる虚数解

$$x = \frac{-b \pm \sqrt{4ac - b^2}i}{2a}$$

を持つといえる. 2 つの和と積はそれぞれ, $-b/a, c/a$ となるので, 実数の場合も含めると, 2 次方程式の解と係数の関係が, 複素数の場合も成り立つ

命題 16.5 (解と係数の関係) a, b, c が複素数である z の 2 次方程式 (ただし $a \neq 0$)

$$az^2 + bz + c = 0$$

の 2 解を α, β とすると,

$$\alpha + \beta = -b/a, \quad \alpha\beta = c/a$$

a, b, c が実数のみならず、複素数の場合でも解と係数の公式が成り立つのである。これは証明はしないが、一般には

定理 16.6 (代数学の基本定理) 複素数係数 c_0, c_1, \dots, c_n の複素数 z の n 次方程式 (ただし $c_0 \neq 0$)

$$c_0 z^n + c_1 z^{n-1} + \dots + c_{n-1} z + c_n = 0$$

の解の個数は、重複も含めて n 個である。 」

が成り立つことと、この $n = 2$ の場合から言える。

17 複素ベクトルのエルミート内積: 第 I 部・第 II 部の拡張

成分が複素数のベクトルを複素ベクトルといい、同様に、成分が複素数の行列を複素ベクトルという。復習もかねて、第 1 部・第 2 部で成り立ったことが、「複素ベクトル・複素行列・スカラーも複素数」、という場合についても成り立つかどうかを、考えてみよう。

- 1 章の計算については、平面座標の 1.3 章以外は、全く同様に計算できる。
- 2 章では、一番最後の直交行列の定義 2.9 以外は、全て同様に成り立つ。直交行列というものは、実行列についてのものである。複素行列の場合には、それを拡張したユニタリ行列というものを定義する(後述)。
- 3 章 (2 次正方行列の逆行列) から 6 章 (2×2 までの行列の積を、分割した行列で表す) までは、複素行列や複素数でも同様に成り立つ。
- 7 章 (2×2 の行列の固有値と固有ベクトル) についても計算は全く同様である。また定義 7.1 の下の説明については、2 次正方行列の固有値は、重複も含めて必ず 2 つ存在する。「必ず 2 つ存在する」という点は実行列とは異なる点である。
- 8 章 (実ベクトルの内積とベクトルの大きさ) は、題名を見た通り実数での話である。複素数ではこのままでは成り立たない。複素数のベクトルの場合では、この章で説明するエルミート内積を用いる。
- 9 章は 実数の平面のことなので、この章ではこれ以上のことはしない。
- 10 章 (2 次実正方行列の直交行列と三角化・対角化) から、12 章 (2 次実正方行列の対称行列の性質) も、実数で成り立つことである。複素数については次の 18 章で説明する。
- このレジュメでは、2 次形式は実数の場合のみを取扱うので、13 章 (n 次実正方行列による、2 次形式の定義) と、14 章 (2 次正方行列と 2 次形式) については、複素数の場合への拡張は考えない。

この章では 8 章の拡張を考える。つまり実ベクトルの内積を拡張した複素ベクトルの内積を定義する。まず転置行列と共役複素数を組み合わせた、随伴行列を定義する。

定義 17.1 成分が複素数の m 行 n 列の行列 A について、

- A の全ての成分を共役複素数にした行列を、 \bar{A} で表す。
- \bar{A} を転置した n 行 m 列の行列を A の随伴行列と定義し、 A^* で表す。式で書くと、

$$A^* = {}^t(\bar{A})$$

(i) から言えることで, m 行 n 列の行列 A について

$$A = \bar{A} \iff A \text{ の成分が全て実数}$$

は簡単であろう.

(ii) については, まず $*$ が A の右上についていることに注意してほしい. 右横ではない. また, 随伴行列は転置行列の計算と, 共役複素数の計算の両方をするものなので, 以下も簡単に理解できるであろう.

- (i). 行列の積 AB が存在するとき, $(AB)^* = B^*A^*$
- (ii). 複素数 c と行列 A について, $(cA)^* = \bar{c}A^*$
- (iii). $(A^*)^* = A$

さて, このバーをとって転置するという随伴行列を使うと, 内積が行列の計算によって, きれいに定義できる.

定義 17.2 複素数成分の, 2つの n 項列ベクトル $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ について,

(i). エルミート内積 $\langle \mathbf{x}, \mathbf{y} \rangle$ を,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{i=1}^n x_i \bar{y}_i$$

で定義する.

(ii). 複素ベクトル \mathbf{x} の大きさを

$$|\mathbf{x}| = \sqrt{\mathbf{x}^* \mathbf{x}} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{x_1 \bar{x}_1 + x_2 \bar{x}_2 + \cdots + x_n \bar{x}_n} = \sqrt{\sum_{i=1}^n x_i \bar{x}_i}$$

で定義する.

」

\mathbf{x} の成分が 1 のとき, つまり $\mathbf{x} = z$ という複素数のときは, (ii) から $|\mathbf{x}| = \sqrt{z\bar{z}} = |z|$ となっている. また, このエルミート内積について \mathbf{x} や \mathbf{y} が実数の場合を考えると, それは定義 8.1 で定義した実数の内積と一致することもいえる. つまり実数の内積を複素数に拡張したものであるといえる.

エルミート内積の主な性質は, 以下である.

定義 17.3 $\lambda, \lambda_1, \lambda_2$ は複素数とし, \mathbf{x}, \mathbf{y} は複素数成分の n 項列ベクトルとする. エルミート内積について, 以下が成り立つ.

- (i). $\langle \lambda \mathbf{x}, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle$
- (ii). $\langle \mathbf{x}, \lambda \mathbf{y} \rangle = \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle$
- (iii). $\langle (\lambda_1 + \lambda_2) \mathbf{x}, \mathbf{y} \rangle = \langle \lambda_1 \mathbf{x}, \mathbf{y} \rangle + \langle \lambda_2 \mathbf{x}, \mathbf{y} \rangle$
- (iv). $\langle \mathbf{x}, (\lambda_1 + \lambda_2) \mathbf{y} \rangle = \langle \mathbf{x}, \lambda_1 \mathbf{y} \rangle + \langle \mathbf{x}, \lambda_2 \mathbf{y} \rangle$

特に (ii) が重要である。この (ii) の性質が、このプリントの最後の方に出てくる、「実対称行列の固有値が実数のみである」という証明のキーとなる。

定義 17.4 (複素列ベクトルの基準化) n 項複素列ベクトル $\mathbf{x} \neq \mathbf{0}$ について、ベクトル

$$\frac{1}{|\mathbf{x}|} \mathbf{x} = \frac{1}{\sqrt{\mathbf{x}^* \mathbf{x}}} \mathbf{x}$$

は大きさ 1 のベクトルとなる。これを \mathbf{x} を基準化したベクトルという。

18 ** ユニタリー行列と 2×2 の複素行列の三角化・対角化

実行列の直交行列を拡張するものとして、ユニタリー行列が定義される。

定義 18.1 複素数成分の n 次正方行列 \mathbf{U} が以下をみたすとき、 \mathbf{U} をユニタリー行列という。

$$\mathbf{U}\mathbf{U}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I} \quad (18.1)$$

英語では unit は 1 単位という意味がある。行列だと単位といえは 1 ではなく単位行列 \mathbf{I} なので、「unitary matrix」だと、「単位行列になる行列」とか、「単位行列を作る行列」などのような意味あいがある。

逆行列の定義より、 \mathbf{U} がユニタリー行列ならば、 $\mathbf{U}^{-1} = \mathbf{U}^*$ であり、正則行列である。

対称行列の拡張として、エルミート行列などが定義されている。

定義 18.2 複素数成分の n 次正方行列 \mathbf{A} について、

- (i). $\mathbf{A}^* = \mathbf{A}$ を満たすとき、 \mathbf{A} をエルミート行列と呼ぶ。
- (ii). $\mathbf{A}^* = -\mathbf{A}$ を満たすとき、 \mathbf{A} を歪エルミート行列と呼ぶ。
- (iii). $\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$ を満たすとき、 \mathbf{A} を正規行列と呼ぶ。

エルミートは人名である。正規行列は、エルミート行列を含む。よって実直交行列も含む。また歪エルミート行列やユニタリー行列も含むので、後者からは直交行列も含んでいることがいえる。当然単位行列も含んでいる。わかりにくいだろうから、包含関係を \implies を用いたイメージで書くと、

- 単位行列 \implies 直交行列 \implies ユニタリー行列 \implies 正規行列
- 単位行列 \implies 実対角行列 \implies 実対称行列 \implies エルミート行列 \implies 正規行列

何回か述べているように、直交行列は、実行列であることに注意すること。このように、正規行列とは、今までにいろいろと学んできた特別な行列を含む行列である。重要な証明はこの章の最後に出てくる。

さて、この n 次までの定義を用いて、以下では具体的に 2×2 の行列の 10 章 (2×2 の行列の直交行列と三角化・対角化) から、12 章 (2×2 の行列での実対称行列の性質) までの、複素数への拡張を述べよう。次の証明は、 2×2 の直交行列の三角化の証明 (定理 10.6 の証明) とほとんど同じである。比較するとよい。

定理 18.3 任意の 2 次正方複素数行列 \mathbf{A} は, 2 次の適当なユニタリ行列 \mathbf{U} によって, 対角成分が固有値となる三角行列に三角化できる. 式で書くと,

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix}$$

であり, λ_1, λ_2 は \mathbf{A} の固有値である. u は適当な複素数である. □

証明 $a, b, c, d, \lambda_1, \lambda_2, x, y$ は複素数である. $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ とし, この固有値を λ_1, λ_2 とする. また λ_1 の固有ベクトルを $\begin{pmatrix} x \\ y \end{pmatrix}$ とし, すでに基準化しているものとする. 以上より,

$$ax + by = \lambda_1 x \tag{18.2}$$

$$cx + dy = \lambda_1 y \tag{18.3}$$

$$x\bar{x} + y\bar{y} = 1 \tag{18.4}$$

$$\lambda_1 + \lambda_2 = a + d \tag{18.5}$$

が成り立つ. ただし (18.5) 式が成り立つのは, \mathbf{A} の固有方程式の解と係数の関係からである.

さて, この x, y を用いて

$$\mathbf{U} = \begin{pmatrix} x & -\bar{y} \\ y & \bar{x} \end{pmatrix} \tag{18.6}$$

とする. この \mathbf{U} はユニタリ行列である (計算してみよ). この \mathbf{U} を用いて,

$$\begin{aligned} \mathbf{U}^{-1}\mathbf{A}\mathbf{U} &= \begin{pmatrix} \bar{x} & \bar{y} \\ -y & x \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & -\bar{y} \\ y & \bar{x} \end{pmatrix} \\ \text{(18.2) 式と (18.3) 式より} &= \begin{pmatrix} \bar{x} & \bar{y} \\ -y & x \end{pmatrix} \begin{pmatrix} \lambda_1 x & -a\bar{y} + b\bar{x} \\ \lambda_1 y & -c\bar{y} + d\bar{x} \end{pmatrix} \\ \text{(18.4) 式より} &= \begin{pmatrix} \lambda_1 & b\bar{x}\bar{x} + (-a+d)\bar{x}\bar{y} - c\bar{y}\bar{y} \\ 0 & dx\bar{x} - b\bar{x}y - c\bar{y}y + ay\bar{y} \end{pmatrix} \end{aligned} \tag{18.7}$$

ここで (18.2) 式の両辺に \bar{x} を, そして (18.3) 式の両辺に \bar{y} を, それぞれ掛けて移項することによって,

$$bx\bar{y} = \lambda_1 x\bar{x} - ax\bar{x}, \quad cx\bar{y} = \lambda_1 y\bar{y} - dy\bar{y}, \tag{18.8}$$

が成り立つので, これを (18.7) 式の (2,2) 成分に代入すると,

$$\begin{aligned} \text{(18.7) 式の (2,2) 成分} &= dx\bar{x} - b\bar{x}y - cx\bar{y} + ay\bar{y} \\ &= dx\bar{x} - (\lambda_1 x\bar{x} - ax\bar{x}) - (\lambda_1 y\bar{y} - dy\bar{y}) + ay\bar{y} \\ \text{(18.4) 式より} &= a + d - \lambda_1 \\ \text{(18.5) 式より} &= \lambda_2 \end{aligned}$$

これと (18.7) 式から, 命題が成り立つ. □

この章の最後として, 正規行列の対角化について証明しよう.

定理 18.4 2×2 の正規行列 A は, 2×2 の適当なユニタリ行列 U によって, 対角成分が A の固有値である対角行列に, 対角化される. 式で書くと,

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

であり, λ_1, λ_2 は A の固有値である. ┘

証明 定理 18.3 により, 2×2 の適当なユニタリ行列 U と, 適当な複素数 u で

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix} \quad (18.9)$$

と三角化される. ここから $u = 0$ となることを証明すればよい.

(18.9) 式の左から U を, 右から U^{-1} を掛け, $U^{-1} = U^*$ であることから, 以下がいえる.

$$A = U \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix} U^* \quad (18.10)$$

この両辺の随伴行列をとると,

$$A^* = U \begin{pmatrix} \bar{\lambda}_1 & 0 \\ \bar{u} & \bar{\lambda}_2 \end{pmatrix} U^* \quad (18.11)$$

正規行列の定義式の $AA^* = A^*A$ に, (18.10) 式と, (18.11) 式を代入すると,

$$U \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix} U^* U \begin{pmatrix} \bar{\lambda}_1 & 0 \\ \bar{u} & \bar{\lambda}_2 \end{pmatrix} U^* = U \begin{pmatrix} \bar{\lambda}_1 & 0 \\ \bar{u} & \bar{\lambda}_2 \end{pmatrix} U^* U \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix} U^*$$

上の式で $U^*U = I$ である. さらに, 両辺の左から U^* 掛けて, 右から U を掛けると,

$$\begin{aligned} \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \bar{\lambda}_1 & 0 \\ \bar{u} & \bar{\lambda}_2 \end{pmatrix} &= \begin{pmatrix} \bar{\lambda}_1 & 0 \\ \bar{u} & \bar{\lambda}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & u \\ 0 & \lambda_2 \end{pmatrix} \\ \begin{pmatrix} \lambda_1 \bar{\lambda}_1 + u \bar{u} & u \bar{\lambda}_2 \\ \lambda_2 \bar{u} & \lambda_2 \bar{\lambda}_2 \end{pmatrix} &= \begin{pmatrix} \bar{\lambda}_1 \lambda_1 & \bar{\lambda}_1 u \\ \bar{u} \lambda_1 & u \bar{u} + \bar{\lambda}_2 \lambda_2 \end{pmatrix} \end{aligned} \quad (18.12)$$

この両辺の成分を比較すると, $u = 0$ であり, これを (18.9) 式に代入して, 命題が成り立つ. □

この証明でもわかるように, 複素数 z では, 大小比較はできなくとも, $|z|^2 = z\bar{z} \geq 0$ であることや,

$$z\bar{z} = 0 \iff z = 0$$

のように, $z\bar{z}$ による比較ができるのである.

2 次の正規行列は, 適当なユニタリ行列によって対角化できることを, ここで証明した. n 次正規行列についても同様に適当なユニタリ行列で対角化できる. 証明は, 参考文献をみるとよい.

19 第 III 部の Scilab 実習

ここでは III 部の復習として Scilab の実習をまとめて行う.

19.1 随伴行列 (実行列なら転置行列) を返す' と, sqrt 関数・norm 関数

ここでは 15.2 節の複素数への応用である.

練習 19.1 次のコマンドを打ち込め.

```
clear
b=[1+%i;2-%i]
b'
b'*b
sqrt(b'*b)
norm(b)
sqrt(b(1,1))
sqrt(b)
```

ここでは、まず表題の随伴行列理解してほしい。3行目で b' が b の随伴行列となっていることを確認してほしい。それ以外では、

- (i). $\%i$ で、虚数単位 i を表している。
- (ii). `sqrt` 関数と `norm` 関数があり、引数が複素数の場合でもともに値を返している。

`clear` は、Scilab 内にある変数および、読み込んだ自作関数を、全て消去するというコマンドである。

`sqrt` 関数は、15.2 節でも説明したのだが、Scilab ではさらに拡張して、 a が複素数でも以下の値を返す仕様となっているようである (ヘルプにはきちんと書いてない)。

- $a = 0$ のときは、`sqrt(a)` は当然 0 を返す。
- 0 以外の複素数については $r > 0$, $-\pi < \theta \leq \pi$ として、 a を $a = r(\cos \theta + i \sin \theta)$ と極形式で表したときに、`sqrt(a)` は、 $\sqrt{r}(\cos(\theta/2) + i \sin(\theta/2))$ を返す。 $0 \leq \theta < 2\pi$ ではないことに気を付けること。

従って $a < 0$ のときに `sqrt(a)` と打ち込むと、 $\sqrt{|a|}i$ が返ってくる。 a の代わりに行列を入れると、成分ごとに計算した値が返ってくる。

$a \neq 0$ の複素数の場合 $x^2 = a$ を満たす複素数 x は 2 つあるのだが、Scilab の関数 `sqrt(a)` は、その片方のみを返すのである。最後の `sqrt(b)` からわかるように、 b がベクトルの場合は、それぞれの成分について計算した値が返ってくる。

次に `norm` 関数を見てみよう。実数の列ベクトルの大きさを表すのが、ユークリッドノルムと定義した。実数からさらに拡張して、 b が複素数の列ベクトルの場合は、エルミート内積を用いて、 $\sqrt{b'b}$ を b のノルムと定義した (定義 17.2 を参照のこと)。そこで上の実習で、`norm(b)` の値と、 $b'b$ の平方根が等しいことから、`norm(b)` は複素列ベクトルのノルムを返すことがわかるであろう。

練習 19.2 次のコマンドを順に打ち込み、`sqrt` 関数について理解せよ。

```
clear
```

```
A=[1 2;0 -25]
sqrt(A)
```

」

19.2 複素数でも、固有値と固有ベクトルを返す spec 関数

複素数の範囲でも、spec を使うことができる。

練習 19.3 次のコマンドを打ち込め。

```
clear
A=[0 -2;8 0]
[U,L]=spec(A)
inv(U)*A*U
```

」

この A は $A = \begin{pmatrix} 0 & -2 \\ 8 & 0 \end{pmatrix}$ と実 2 次正方行列であるが、固有値は $\pm 4i$ の 2 つである。15.4 節で説明したように、A の固有値が $L(1,1)$ のときの、(基準化された) 固有ベクトルが U の 1 列目である。A の固有値が $L(2,2)$ のときの、(基準化された) 固有ベクトルが U の 2 列目である。

今の場合固有値の値が異なるので、複素行列 U によって、A が L に対角化されることが 4 行目からわかるであろう。

第IV部

3 × 3 の行列を中心に

20 * 3 × 3 の行列の行列式はどうあるべきか

普通の工学系テキストでは、先を急がなければならないので行列式を先に定義してそこから論じているものが多い。が、我々は文系なので、ゆっくりと進めることにする。そういうわけで、行列式はどうあるべきかを考えた後に、行列式を定義することにしよう。この章の数は全て実数とする。

さて、2 × 2 の行列 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ の行列式は、 $\det(\mathbf{A}) = ad - bc$ と定義した。では

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (20.1)$$

という 3 × 3 の行列については、 \mathbf{A} の行列式 $\det(\mathbf{A})$ はどうあるべきか。やはり 2 × 2 の場合と同様に、逆行列を持つか否かの判別式となるべきであろう。そこでそのような判別式を探していこう。まず以下の x_1, x_2, x_3 についての方程式を行列にしたものを考える。

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (20.2)$$

2次正方行列の結果から、「この解が $x_1 = x_2 = x_3 = 0$ のみのとき、 \mathbf{A} は逆行列を持ち、そうでない場合は、 \mathbf{A} は逆行列を持たない」というように、したい。

では最初に \mathbf{A} の 2 列目までを見ておこう。

[手順 1] ある数 s, t (s, t の少なくとも一方は非 0) が存在して、

$$s \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = t \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} \quad (20.3)$$

が成り立つとき、

$$\mathbf{A} \begin{pmatrix} s \\ -t \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} s \\ -t \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

が成り立つので、 \mathbf{A} は逆行列を持たない。

[手順 2] では次に (20.3) 式が成り立たない場合を考えよう。2次正方行列の結果から、(20.3) 式が成り立つことは、

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = 0, \quad \det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix} = 0, \quad \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = 0 \quad (20.4)$$

が全て成り立つことと同値である。よって、(20.4) 式が成り立たない場合を考える。つまり 3 つの行列式のうちの、少なくとも 1 つは非 0 である場合を考えるのである。ところで、 \mathbf{A} の行の入れ替えについては、

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{pmatrix} \begin{pmatrix} x_2 \\ x_3 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

がいえるので, (20.4) の 3 つの行列式のうちの, 少なくとも 1 つは非 0 であるということは,

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \neq 0 \quad (20.5)$$

を仮定しても一般性は失わないことを意味する. よってこれを仮定しよう. (20.2) 式を書き直せば,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (20.6)$$

$$(a_{31} \ a_{32}) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + a_{33}x_3 = 0 \quad (20.7)$$

の 2 つの式である. (20.6) 式の左辺第 2 項を右辺に移項して, 左から $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ の逆行列を書けると,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = - \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3 \quad (20.8)$$

となる. これを, (20.7) 式に代入することによって,

$$\left[- (a_{31} \ a_{32}) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} + a_{33} \right] x_3 = 0$$

逆行列を成分で表示して, さらに両辺を $(a_{11}a_{22} - a_{12}a_{21})$ 倍すると,

$$\left[- (a_{31} \ a_{32}) \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} + a_{33}(a_{11}a_{22} - a_{12}a_{21}) \right] x_3 = 0 \quad (20.9)$$

この左辺の [] 内が非 0 ならば, この式を満たす解は $x_3 = 0$ のみであるが, そうでないならば, $x_3 = 0$ 以外の解をもつ. (20.8) 式と考え合わせれば, (20.9) 式の左辺の [] 内が非 0 であるとき, そのときのみ, ((20.5) 式の仮定の下で) (20.2) 式の解は $x_1 = x_2 = x_3 = 0$ のみとなる. つまり \mathbf{A} は逆行列を持つ.

よって, (20.9) 式の左辺の [] 内を 3×3 の行列式にすればよさそうである. 展開して,

$$a_{11}a_{12}a_{13} + a_{21}a_{32}a_{13} + a_{31}a_{23}a_{12} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21} \quad (20.10)$$

[手順 3] あとは, (20.3) 式が成り立つ場合, (20.10) 式が 0 になることを示さなければならないが, 代入して計算すれば, 0 になることは簡単に示すことができる.

よって, ようやく以下の定義と定理がいえる.

定義 20.1 3×3 の行列 \mathbf{A} を,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

とするとき, その行列式 $\det(\mathbf{A})$ を

$$\det(\mathbf{A}) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

とする. ┘

定理 20.2 3×3 の行列 \mathbf{A} について, 以下は同値である.

- (i). $\det(\mathbf{A}) \neq 0$
(ii). \mathbf{A} は逆行列を持つ.

┘

2×2 の行列のときの行列式の性質 (定理 5.2) が, 今の場合も同様に成り立つのだが, 1 つ 1 つ確かめるのは大変である.. ここでは 1 つだけ取り上げておこう.

定理 20.3 3×3 の行列 \mathbf{A} について,

- (i). ある行と別の行を入れ替えた行列の行列式は, $-\det(\mathbf{A})$ である.
(ii). ある列と別の列を入れ替えた行列の行列式は, $-\det(\mathbf{A})$ である.

┘

証明 1 行目と 2 行目を入れ替えてみよう. \mathbf{A} は定義 20.1 の通りとする.

$$\begin{aligned} \det \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{21}a_{12}a_{33} + a_{22}a_{13}a_{31} + a_{23}a_{11}a_{32} - a_{23}a_{12}a_{31} - a_{22}a_{11}a_{33} - a_{21}a_{13}a_{32} \\ &= -(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}) \\ &= -\det(\mathbf{A}) \end{aligned}$$

他の場合も同様である. □

21 * 3×3 の連立方程式のクラメールの公式

(20.1) 式が与えられたときに $\det(\mathbf{A}) \neq 0$ の場合の逆行列を求めたいのだが, それは次章に回す. まず次の連立 1 次方程式を考えて, その解の公式にあたるものを求めよう.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \quad (21.1)$$

x_1, x_2, x_3 の順序を適切に変更することにより, 前章と同様に $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \neq 0$ を仮定しても一般性は失わないので, そのように仮定する. (21.1) 式を解いていくことにしよう.

(20.6) 式・(20.7) 式と同様にして, (21.1) 式は以下の 2 式と同値となる.

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3 = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (21.2)$$

$$\begin{pmatrix} a_{31} & a_{32} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + a_{33}x_3 = b_3 \quad (21.3)$$

前章と同様に変形する. (21.2) 式を変形すると,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} - \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3 \quad (21.4)$$

であり, (21.3) 式に代入すれば, 以下が得られる.

$$(a_{31} \ a_{32}) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} - (a_{31} \ a_{32}) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3 + a_{33}x_3 = b_3$$

逆行列を計算したあと, 両辺に $a_{11}a_{22} - a_{12}a_{21}$ をかけて, 左辺第 1 項を右辺に移項すると,

$$\det(\mathbf{A})x_3 = (a_{11}a_{22} - a_{12}a_{21})b_3 - (a_{31} \ a_{32}) \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (21.5)$$

右辺を続けて計算していくと,

$$\begin{aligned} \det(\mathbf{A})x_3 &= (a_{11}a_{22} - a_{12}a_{21})b_3 + (-a_{31}a_{22} + a_{32}a_{21} \quad a_{31}a_{12} - a_{32}a_{11}) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ &= (a_{11}a_{22} - a_{12}a_{21})b_3 - a_{31}a_{22}b_1 + a_{32}a_{21}b_1 + a_{31}a_{12}b_2 - a_{32}a_{11}b_2 \\ &= \det \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix} \end{aligned}$$

よって,

$$x_3 = \frac{\det \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix}}{\det(\mathbf{A})} \quad (21.6)$$

(21.1) 式から, (21.6) 式が導かれた. よって, (21.1) 式を,

$$\begin{pmatrix} a_{13} & a_{12} & a_{11} \\ a_{23} & a_{22} & a_{21} \\ a_{33} & a_{32} & a_{31} \end{pmatrix} \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31} & a_{33} & a_{32} \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

と書き換えると, それぞれ,

$$x_1 = \frac{\det \begin{pmatrix} a_{13} & a_{12} & b_1 \\ a_{23} & a_{22} & b_2 \\ a_{33} & a_{32} & b_3 \end{pmatrix}}{\det \begin{pmatrix} a_{13} & a_{12} & a_{11} \\ a_{23} & a_{22} & a_{21} \\ a_{33} & a_{32} & a_{31} \end{pmatrix}}, \quad x_2 = \frac{\det \begin{pmatrix} a_{11} & a_{13} & b_1 \\ a_{21} & a_{23} & b_2 \\ a_{31} & a_{33} & b_3 \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{32} \\ a_{31} & a_{33} & a_{22} \end{pmatrix}}$$

が導かれる. x_1 の分子の行列式については, 定理 20.3 より, 1 列目と 3 列目を入れ替えた行列の行列式の -1 倍と等しくなる. つまり,

$$-\det \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}$$

と等しくなる. 分母の行列も同様の入れ替えをすると, 分母は $-\det(\mathbf{A})$ となる.

x_2 の分子の行列式については, 2 列目と 3 列目を入れ替えた行列式の -1 倍となる. 分母も同様にすると, $-\det(\mathbf{A})$ となる. 以上を (21.6) 式も含めてまとめると,

$$x_1 = \frac{\det \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}}{\det(\mathbf{A})}, \quad x_2 = \frac{\det \begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix}}{\det(\mathbf{A})}, \quad x_3 = \frac{\det \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix}}{\det(\mathbf{A})} \quad (21.7)$$

が得られる。2×2と同様に、3×3の場合のクラメールの公式が得られた。

2×2のときも説明したが、クラメールの公式は、 \mathbf{A} が文字の時や、0や1が多い時に役に立つ。また規則的な配列をしているときにも計算しやすい。

22 ** 3×3の逆行列の式

理論上以下のように計算できるということを示す。逆行列の第*i*行*j*列だけ知りたいというときには、手計算することもある。が、逆行列全体をこのように実際に手計算することはほぼない。とにかくまずは、

$$\det(\mathbf{A}) \neq 0 \text{ ならば逆行列が存在する.}$$

ということを確認してくればよい。

\mathbf{A}^{-1} が存在するならば、 $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ である。(21.1)式の形から、 $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ について計算するほうが楽なので、これを解く。つまり、 $\det(\mathbf{A}) \neq 0$ として、

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mathbf{A}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

となる3行3列の行列 \mathbf{A}^{-1} を求める。

まず \mathbf{A}^{-1} の1列目を $\begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix}$ とすると、

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

クラメールの公式では、(21.1)式の解が(21.7)式なのだから、(21.7)式に $b_1 = 1, b_2 = 0, b_3 = 0$ を代入して得られた x_1, x_2, x_3 が、それぞれ x_{11}, x_{21}, x_{31} である。よって、

$$x_{11} = \frac{\det \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix}}{\det(\mathbf{A})}, \quad x_{21} = \frac{\det \begin{pmatrix} a_{11} & 1 & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{pmatrix}}{\det(\mathbf{A})}, \quad x_{31} = \frac{\det \begin{pmatrix} a_{11} & a_{12} & 1 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix}}{\det(\mathbf{A})}$$

これらの分子を計算した後に、2×2の行列式の形で書くことができ、

$$x_{11} = \frac{\det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}}{\det(\mathbf{A})}, \quad x_{21} = \frac{-\det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}}{\det(\mathbf{A})}, \quad x_{31} = \frac{\det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}}{\det(\mathbf{A})},$$

と表される。同様に、 \mathbf{A}^{-1} の2列目と3列目については、それぞれ以下を解けばよい。

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_{13} \\ x_{23} \\ x_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

左の式の解は、(21.7)式に $b_1 = 0, b_2 = 1, b_3 = 0$ を代入すれば得られる。右の式の解は、(21.7)式に $b_1 = 0, b_2 = 0, b_3 = 1$ を代入すれば得られる。

計算して得られたものをまとめると,

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} & -\det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} & \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \\ -\det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} & \det \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} & -\det \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix} \\ \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} & -\det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix} & \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \end{pmatrix} \quad (22.1)$$

23 * 3 × 3 の行列と 2 次形式

この章によって, C^2 級の 3 変数実数値関数 $f(x, y, z)$ についての (狭義の) 極大・(狭義の) 極小についての議論が, 十分に理解できるようになる.

定理 23.1 3 次実対称行列 $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ を考える.

(i). 以下の 3 つの不等式が成り立つならば, \mathbf{A} は正値定符号である.

$$a_{11} > 0, \quad \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0, \quad \det(\mathbf{A}) > 0$$

(ii). 以下の 3 つの不等式が成り立つならば, \mathbf{A} は負値定符号である.

$$a_{11} < 0, \quad \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0, \quad \det(\mathbf{A}) < 0$$

」

証明 (i) の証明: x_1, x_2, x_3 の 3 元 2 次形式 $Q(x_1, x_2, x_3)$ を

$$Q(x_1, x_2, x_3) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

とする. 仮定より,

$$a_{11} > 0, a_{11}a_{22} - a_{12}^2 > 0, a_{11}a_{22}a_{33} + 2a_{13}a_{12}a_{23} - a_{13}^2a_{22} - a_{23}^2a_{11} - a_{12}^2a_{33} > 0$$

がいえる. $Q(x_1, x_2, x_3)$ を平方完成を駆使して変形していくと,

$$\begin{aligned} Q(x_1, x_2, x_3) &= a_{11}x_1^2 + 2(a_{12}x_2 + a_{13}x_3)x_1 + (a_{22}x_2^2 + 2a_{23}x_2x_3 + a_{33}x_3^2) \\ &= a_{11} \left(x_1 + \frac{a_{12}x_2 + a_{13}x_3}{a_{11}} \right)^2 - \frac{(a_{12}x_2 + a_{13}x_3)^2}{a_{11}} + (a_{22}x_2^2 + 2a_{23}x_2x_3 + a_{33}x_3^2) \\ &= a_{11} \left(x_1 + \frac{a_{12}x_2 + a_{13}x_3}{a_{11}} \right)^2 + \frac{(a_{11}a_{22} - a_{12}^2)x_2^2 + 2(a_{11}a_{23} - a_{12}a_{13})x_2x_3 + (a_{11}a_{33} - a_{13}^2)x_3^2}{a_{11}} \\ &= a_{11} \left(x_1 + \frac{a_{12}x_2 + a_{13}x_3}{a_{11}} \right)^2 \\ &\quad + \frac{1}{a_{11}} \left[(a_{11}a_{22} - a_{12}^2) \left\{ x_2 + \frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}a_{22} - a_{12}^2} x_3 \right\}^2 - \frac{(a_{11}a_{23} - a_{12}a_{13})^2}{a_{11}a_{22} - a_{12}^2} x_3^2 + (a_{11}a_{33} - a_{13}^2)x_3^2 \right] \end{aligned}$$

$$\begin{aligned}
&= a_{11} \left(x_1 + \frac{a_{12}x_2 + a_{13}x_3}{a_{11}} \right)^2 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11}} \left(x_2 + \frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}a_{22} - a_{12}^2} x_3 \right)^2 \\
&\quad + \frac{1}{a_{11}} \frac{-a_{11}^2 a_{23}^2 + 2a_{11}a_{23}a_{12}a_{13} - a_{12}^2 a_{13}^2 + a_{11}^2 a_{22}a_{33} - a_{11}a_{12}^2 a_{33} - a_{11}a_{22}a_{13}^2 + a_{12}^2 a_{13}^2}{a_{11}a_{22} - a_{12}^2} x_3^2 \\
&= a_{11} \left(x_1 + \frac{a_{12}x_2 + a_{13}x_3}{a_{11}} \right)^2 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11}} \left(x_2 + \frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}a_{22} - a_{12}^2} x_3 \right)^2 \\
&\quad + \frac{1}{a_{11}} \frac{-a_{11}^2 a_{23}^2 + 2a_{11}a_{23}a_{12}a_{13} + a_{11}^2 a_{22}a_{33} - a_{11}a_{12}^2 a_{33} - a_{11}a_{22}a_{13}^2}{a_{11}a_{22} - a_{12}^2} x_3^2 \\
&= a_{11} \left(x_1 + \frac{a_{12}x_2 + a_{13}x_3}{a_{11}} \right)^2 + \frac{a_{11}a_{22} - a_{12}^2}{a_{11}} \left(x_2 + \frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}a_{22} - a_{12}^2} x_3 \right)^2 \\
&\quad + \frac{-a_{11}a_{23}^2 + 2a_{23}a_{12}a_{13} + a_{11}a_{22}a_{33} - a_{12}^2 a_{33} - a_{22}a_{13}^2}{a_{11}a_{22} - a_{12}^2} x_3^2 \\
&\geq 0
\end{aligned}$$

ところで、この等号が成り立つ必要十分条件は、

$$x_3 = 0, \quad x_2 + \frac{a_{11}a_{23} - a_{12}a_{13}}{a_{11}a_{22} - a_{12}^2} x_3 = 0, \quad x_1 + \frac{a_{12}x_2 + a_{13}x_3}{a_{11}} = 0$$

であるのだが、これは、

$$x_3 = 0, x_2 = 0, x_1 = 0$$

と同値である。よって $x_1 = x_2 = x_3 = 0$ 以外ならば、 $Q(x_1, x_2, x_3) > 0$ であることがいえる。これは、 \mathbf{A} が正値定符号行列であることを意味する。

(ii) の証明: $-\mathbf{A}$ が正値定符号であることと、 \mathbf{A} が負値定符号であることは同値なので、(i) の結果から言える。

□

このように、高校1年次あたりで学んだ平方完成のみで証明できる。より少ない計算力による証明は、 n 次元の固有値・固有ベクトルを学んだ後、39章で説明する。

第V部

一般の n 次正方行列を中心に24 * 4×4 の行列の行列式はどうあるべきか

20章と同様にして、3次正方行列のクラメールの公式も用いて、4次正方行列の行列式はどうあるべきかを探る。この章では細かいことは軽く流し、項が24個出てくる点に注目してほしい。

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \quad (24.1)$$

とするとき、 x_1, x_2, x_3, x_4 についての連立方程式

$$\mathbf{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (24.2)$$

について、

$$\text{解が } x_1 = x_2 = x_3 = x_4 = 0 \text{ のみ} \iff \text{行列式が非0}$$

を満たすような行列式を定義したいというのが、この章の目的である。

[手順1]

20章の[手順1]と同様に、

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{pmatrix} s + \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ a_{42} \end{pmatrix} t + \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \\ a_{43} \end{pmatrix} u = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (24.3)$$

となるスカラー s, t, u が、 $s = t = u = 0$ 以外に存在する場合には、 $\det(\mathbf{A}) = 0$ になるように作りたい。ここではこれ以上計算はしない。

[手順2]

手順1の(24.3)式をみたす s, t, u が $s = t = u = 0$ のみの場合を考える。20章の[手順2]と同様に、

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \neq 0$$

を仮定しても一般性は失わない。(24.2)式を書き下すと、以下の2式となる。

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix} x_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (24.4)$$

$$a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 = 0 \quad (24.5)$$

(24.4)式の左辺第2項を右辺に移項して、(21.7)式のクラメールの公式によって、

$$x_1 = \frac{-x_4 \det \begin{pmatrix} a_{14} & a_{12} & a_{13} \\ a_{24} & a_{22} & a_{23} \\ a_{34} & a_{32} & a_{33} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}}, \quad x_2 = \frac{-x_4 \det \begin{pmatrix} a_{11} & a_{14} & a_{13} \\ a_{21} & a_{24} & a_{23} \\ a_{31} & a_{34} & a_{33} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}}, \quad x_3 = \frac{-x_4 \det \begin{pmatrix} a_{11} & a_{12} & a_{14} \\ a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}} \quad (24.6)$$

これを (24.5) 式に代入して、両辺 $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ 倍すると、

$$\left[\begin{array}{l} -a_{41} \det \begin{pmatrix} a_{14} & a_{12} & a_{13} \\ a_{24} & a_{22} & a_{23} \\ a_{34} & a_{32} & a_{33} \end{pmatrix} - a_{42} \det \begin{pmatrix} a_{11} & a_{14} & a_{13} \\ a_{21} & a_{24} & a_{23} \\ a_{31} & a_{34} & a_{33} \end{pmatrix} \\ -a_{43} \det \begin{pmatrix} a_{11} & a_{12} & a_{14} \\ a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \end{pmatrix} + a_{44} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \end{array} \right] x_4 = 0$$

この [] 内为非ゼロのとき, (24.2) 式の解は, $x_1 = x_2 = x_3 = x_4 = 0$ のみとなる. ゼロのときは, それ以外の解を持つ. よってこの式の [] 内を \mathbf{A} の行列式とするとよさそうである. そこでこれを展開すると, (3次正方行列の行列式の項数である)6項ずつ並べて

$$\begin{aligned} & -a_{14}a_{22}a_{33}a_{41} - a_{12}a_{23}a_{34}a_{41} - a_{13}a_{24}a_{32}a_{41} + a_{13}a_{22}a_{34}a_{41} + a_{14}a_{23}a_{32}a_{41} + a_{12}a_{24}a_{33}a_{41} \\ & -a_{11}a_{24}a_{33}a_{42} - a_{14}a_{23}a_{31}a_{42} - a_{13}a_{21}a_{34}a_{42} + a_{13}a_{24}a_{31}a_{42} + a_{14}a_{21}a_{33}a_{42} + a_{11}a_{23}a_{34}a_{42} \\ & -a_{11}a_{22}a_{34}a_{43} - a_{12}a_{24}a_{31}a_{43} - a_{14}a_{21}a_{32}a_{43} + a_{14}a_{22}a_{31}a_{43} + a_{12}a_{21}a_{34}a_{43} + a_{11}a_{24}a_{32}a_{43} \\ & + a_{11}a_{22}a_{33}a_{44} + a_{12}a_{23}a_{31}a_{44} + a_{13}a_{21}a_{32}a_{44} - a_{13}a_{22}a_{31}a_{44} - a_{12}a_{21}a_{33}a_{44} - a_{11}a_{23}a_{32}a_{44} \end{aligned} \quad (24.7)$$

である. これが4次正方行列 (24.1) 式の, 行列式の候補である.

[手順3] 本来ならば, (24.3) 式を満たす s, t, u が $s = t = u = 0$ 以外に存在する場合には, (24.7) 式が0になることを確かめなければならない. しかしそれは, 一般の形で行列式を定義した後, 一般の形で証明する. とにかく, 4次の場合の行列式 (の候補) には項が24個あって, このように定義するとよさそうであることを確認し, 次章を見てほしい.

25 n 次正方行列の行列式の定義の準備

2次の場合の行列式は,

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (25.1)$$

3次の場合の行列式は,

$$\begin{aligned} & \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ & = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \end{aligned} \quad (25.2)$$

そして, 4次の場合には, (24.7) 式が候補である.

これらから共通点を探ると,

- (i). 行列式の項数は、2 次の場合は項が $2! = 2$ 個、3 次の場合は項が $3! = 6$ 個、4 次の場合は項が $4! = 24$ 個。
- (ii). 各項は、2 次正方形行列のときは 2 つの数字の積、3 次の場合は 3 つの積、4 次の場合は、4 つの数字の積となっており、また係数は 1 か -1
- (iii). 行列内の対角成分の積の項つまり、2 次の場合は $a_{11}a_{22}$ 、3 次の場合は $a_{11}a_{22}a_{33}$ 、4 次の場合は $a_{11}a_{12}a_{33}a_{44}$ があり、それらの係数はすべて 1 である。
- (iv). 行列式内の各項を見ると、行数も列数にも重複がない。4 次の行列式の候補 (24.7) の最初の項 $-a_{14}a_{22}a_{33}a_{41}$ を見てみよう。 $a_{14}, a_{22}, a_{33}, a_{41}$ の値について、行番号をみると 1, 2, 3, 4 と重複なく 1 から 4 まで存在し、列番号を見ると、やはり 4, 2, 3, 1 と重複なく 1 から 4 まで存在している。これがどの項についてもいえる。

(i) と (iv) からは、順列が関係していることが示唆される。(ii) からは、順列によって符号を変えていることが示唆される。(iii) からは、対角成分の積を中心に考えることがよさそうである。

さてそれでは、1 から 3 の場合の順列をすべてあげよう。(25.2) 式の順番通りに、正負の表と順列について表にしよう。6 つの項について、行数が 1, 2, 3 のとき、列がいくつになっているかを見てみよう。つまり各項 $a_{1*}a_{2?}a_{3.}$ の形となっているので、* と ? と . の数字を順に並べたのである。正負については、係数が 1 のときは正、 -1 の時は負と書いてある。

	1 行	2 行	3 行	符号
第 1 項 $a_{11}a_{22}a_{33}$ の列番号	1	2	3	正
第 2 項 $a_{12}a_{23}a_{31}$ の列番号	2	3	1	正
第 3 項 $a_{13}a_{21}a_{32}$ の列番号	3	1	2	正
第 4 項 $a_{13}a_{22}a_{31}$ の列番号	3	2	1	負
第 5 項 $a_{12}a_{21}a_{33}$ の列番号	2	1	3	負
第 6 項 $a_{11}a_{23}a_{32}$ の列番号	1	3	2	負

6 つの列番号の数字が、1, 2, 3 の並べ方すべてをとっていることがわかる。次に 2 次正方形行列の行列式の定理 5.2 の (iii) を見てみよう。そこでは、行を交換したり、列を交換すると、行列式が -1 倍となった。そこで数字の交換を考えることにしよう。

第 2 項の列番号の順序は 2, 3, 1 となっている。1, 2, 3 の順序から何回数字を交換すると、2, 3, 1 という順序になるか考えてみよう。もちろん交換というのは 2 つの数字を入れ替えることである。

$$\begin{aligned}
 & 1, 2, 3 \\
 1 \text{ と } 2 \text{ を交換すると} & \rightarrow 2, 1, 3 \\
 3 \text{ と } 1 \text{ を交換すると} & \rightarrow 2, 3, 1
 \end{aligned}$$

という 2 回の数字の交換で得られる。

第 3 項の列番号の数字 3, 1, 2 は、

$$\begin{aligned}
 & 1, 2, 3 \\
 1 \text{ と } 3 \text{ を交換すると} & \rightarrow 3, 2, 1 \\
 1 \text{ と } 2 \text{ を交換すると} & \rightarrow 3, 1, 2
 \end{aligned}$$

という 2 回の数字の交換で得られる。第 4 項以降も同様に見てみると、

第 4 項では

$$1, 2, 3 \text{ で、} 1 \text{ と } 3 \text{ を交換すると} \rightarrow 3, 1, 2$$

という1回の数字の交換で得られる.

第5項では

$$1, 2, 3 \text{ で, } 1 \text{ と } 2 \text{ を交換すると } \rightarrow 2, 1, 3$$

という1回の数字の交換で得られる.

第6項では

$$1, 2, 3 \text{ で, } 2 \text{ と } 3 \text{ を交換すると } \rightarrow 1, 3, 2$$

という1回の数字の交換で得られる.

第1項では交換しなくともよいので, 0回の交換で得られる.

第1項から第3項では, 交換回数が偶数で, 符号が正である. 一方第4項から第6項までは, 交換回数が奇数で, 符号が負である. だから, 交換回数の偶奇によって, 符号を決めるとよさそうである.

ただし, ここで問題となるのは, 交換は上のように1通りではないことである. 例えば第6項の場合だと,

$$\begin{array}{l} 1, 2, 3 \\ 1 \text{ と } 2 \text{ を交換すると } \rightarrow 2, 1, 3 \\ 1 \text{ と } 3 \text{ を交換すると } \rightarrow 2, 3, 1 \\ 1 \text{ と } 2 \text{ を交換すると } \rightarrow 1, 3, 2 \end{array}$$

のように3回の交換によっても, 1, 3, 2の順序が得られる. とは言え3回というのは1回と同様に奇数である.

このようなことがあるので, 1から n までのある順序が与えられたとき, それは1から n まで順に並べたものから,

- 順序にかかわらず偶数回の交換で, ある順序が得られる.
- 順序にかかわらず奇数回の交換で, ある順序が得られる.

のいずれかが, 必ず成り立つということを証明しなければならないのである.

4次の場合も見てみよう. (24.7)についても見やすくするために各項の行数は1, 2, 3, 4の順に書いてある. よって列数の順に注目していこう. すでに書いたように交換方法はいろいろあるのだが, その最小の交換回数で得られるようにした.

項番号	並び順	1, 2, 3, 4 からの変形の例	交換回数
第1項	4, 2, 3, 1	1, 2, 3, 4 \rightarrow 4, 2, 3, 1	1回
第2項	2, 3, 4, 1	1, 2, 3, 4 \rightarrow 4, 2, 3, 1 \rightarrow 2, 4, 3, 1 \rightarrow 2, 3, 4, 1	3回
第3項	3, 4, 2, 1	1, 2, 3, 4 \rightarrow 3, 2, 1, 4 \rightarrow 3, 4, 1, 2 \rightarrow 3, 4, 2, 1	3回
第4項	3, 2, 4, 1	1, 2, 3, 4 \rightarrow 3, 2, 1, 4 \rightarrow 3, 2, 4, 1	2回
第5項	4, 3, 2, 1	1, 2, 3, 4 \rightarrow 4, 2, 3, 1 \rightarrow 4, 3, 2, 1	2回
第6項	2, 4, 3, 1	1, 2, 3, 4 \rightarrow 2, 1, 3, 4 \rightarrow 2, 4, 3, 1	2回
第7項	1, 4, 3, 2	1, 2, 3, 4 \rightarrow 1, 4, 3, 2	1回
第8項	4, 3, 1, 2	1, 2, 3, 4 \rightarrow 4, 2, 3, 1 \rightarrow 4, 3, 2, 1 \rightarrow 4, 3, 1, 2	3回
第9項	3, 1, 4, 2	1, 2, 3, 4 \rightarrow 3, 2, 1, 4 \rightarrow 3, 1, 2, 4 \rightarrow 3, 1, 4, 2	3回
第10項	3, 4, 1, 2	1, 2, 3, 4 \rightarrow 3, 2, 1, 4 \rightarrow 3, 4, 1, 2	2回
第11項	4, 1, 3, 2	1, 2, 3, 4 \rightarrow 4, 2, 3, 1 \rightarrow 4, 1, 3, 2	2回
第12項	1, 3, 4, 2	1, 2, 3, 4 \rightarrow 1, 3, 2, 4 \rightarrow 1, 3, 4, 2	2回
第13項	1, 2, 4, 3	1, 2, 3, 4 \rightarrow 1, 2, 4, 3	1回
第14項	2, 4, 1, 3	1, 2, 3, 4 \rightarrow 2, 1, 3, 4 \rightarrow 2, 4, 3, 1 \rightarrow 2, 4, 1, 3	3回
第15項	4, 1, 2, 3	1, 2, 3, 4 \rightarrow 4, 2, 3, 1 \rightarrow 4, 1, 3, 2 \rightarrow 4, 1, 2, 3	3回
第16項	4, 2, 1, 3	1, 2, 3, 4 \rightarrow 4, 2, 3, 1 \rightarrow 4, 2, 1, 3	2回
第17項	2, 1, 4, 3	1, 2, 3, 4 \rightarrow 2, 1, 3, 4 \rightarrow 2, 1, 4, 3	2回
第18項	1, 4, 2, 3	1, 2, 3, 4 \rightarrow 1, 4, 3, 2 \rightarrow 1, 4, 2, 3	2回
第19項	1, 2, 3, 4	1, 2, 3, 4	0回
第20項	2, 3, 1, 4	1, 2, 3, 4 \rightarrow 2, 1, 3, 4 \rightarrow 2, 3, 1, 4	2回
第21項	3, 1, 2, 4	1, 2, 3, 4 \rightarrow 3, 2, 1, 4 \rightarrow 3, 1, 2, 4	2回
第22項	3, 2, 1, 4	1, 2, 3, 4 \rightarrow 3, 2, 1, 4	1回
第23項	2, 1, 3, 4	1, 2, 3, 4 \rightarrow 2, 1, 3, 4	1回
第24項	1, 3, 2, 4	1, 2, 3, 4 \rightarrow 1, 3, 2, 4	1回

この表と (24.7) 式を照らし合わせれば、交換回数が偶数のときは、各項の符号は正になっており、交換回数が奇数のときは、各項の符号は負になっている。

この表を元にして、次章で置換と互換を定義し、行列式を定義しよう。

26 n 次正方行列の行列式

前章の数字の並べ替えを置換とよび、数学の関数で表すと次のような表現となる。

定義 26.1 正整数 n が 1 つ与えられているとする。

- (i). 1 から n までの整数が定義域で、値域も 1 から n までの整数である、1 対 1 の関数 σ を、置換と呼ぶ。
- (ii). n が与えられたとき、置換の集合を S_n とおく。 S_n は、 $n!$ 個の要素を持つ有限集合である。
- (iii). σ の表現方法として、1 行目に定義域の数 i 、2 行目に、その 1 行目の値に対応する $\sigma(i)$ を書いて、行

列の形で書くものとする。 S_2 の元を全てあげると、

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

の 2 通りである。 S_3 の元を全てあげると、

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

の 6 通りである。

」

1 行目は別に 1,2,3 の順に並んでいる必要はない。特に積の場合を考えると、並び替えて計算する必要がある。さて、実数の関数の合成関数で $g \circ f(x)$ と書かれていた場合、 x に f を施した後 g を施すという形になっている。これと同様に、置換の積を考えたとき、右から考えていくのがポイントである。

25 章であげた、24 項の表の第 4 項の置換を考えると、

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

と表すことができる。 行列の計算ではないことに十分気を付けてほしい。

さらに、同じ列で同じ数字の場合は動かさないわけだから、それは書かないことにするのが一般的である。それに従うと、上の場合は

$$\begin{pmatrix} 1 & 3 & 4 \\ 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \tag{26.1}$$

この右辺にあるものは、2 つの交換でありもっと簡単に表現することとしよう。

定義 26.2 置換の集合 S_n の要素について、

(i). 1 から n までの整数 $i, j (i \neq j)$ について、置換が

$$\begin{pmatrix} i & j \\ j & i \end{pmatrix}$$

となっているものを、互換 (交換) という。これを簡単に以下で表す。

$$(i \ j)$$

(ii). 文字が全く動かない置換、

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

が必ず 1 つ存在し、それを恒等置換と呼び、 1_n で表す。

(iii). $\sigma \in S_n$ が与えられたとき,

$$\tau\sigma = \sigma\tau = 1_n$$

となる置換 τ は, $\tau \in S_n$ であり,

$$\tau = \begin{pmatrix} \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ 1 & 2 & \cdots & n \end{pmatrix}$$

で必ず唯一存在する. この τ を σ の逆置換と呼び, σ^{-1} で表す.

」

互換の表現から (26.1) 式は,

$$\begin{pmatrix} 1 & 3 & 4 \\ 3 & 4 & 1 \end{pmatrix} = (1 \ 4)(1 \ 3)$$

であらわされる.

右辺をみると, 1, 2, 3, 4 の並びから, まず 1 と 3 を入れ替える. 次に 1 と 4 を入れ替える. その結果が左辺の置換と一致するのである. 先に書いたように, 右から置換を施していくことに気を付けること.

では次に, 前章で書いたように, 置換を互換の積で表すことができることを示そう.

命題 26.3 置換の集合 S_n の任意の元 σ は, 有限個の互換の積で表すことができる. より強く言えば, $n-1$ 個以下の互換の積で表すことができる.

」

証明 任意の $\sigma \in S_n$ について, 置換 $\tau_i (i = 1, 2, \dots, n-1)$ を

$$\tau_i = \begin{cases} \begin{pmatrix} i & \sigma(i) \\ 1_n \end{pmatrix} & (\text{if } i \neq \sigma(i)) \\ 1_n & (\text{if } i = \sigma(i)) \end{cases}$$

とする. すると,

$$\sigma = \tau_{n-1}\tau_{n-2}\tau_{n-3}\cdots\tau_2\tau_1$$

が成り立つ. $i = \sigma(i)$ の場合の恒等置換の場合には, 何も置換しないと考えると, この結果からは以下のことが言える. つまり, σ が与えられたとき, $i = 1, 2, \dots, n-1$ の中で $i \neq \sigma(i)$ の個数の互換の積によって, σ が表される. よって命題がなりたつ.

□

25 章であげた, 24 項の表の交換の順序はこの方式に従って行っている. 最初の 6 項ぐらいでもよいので, 余裕があれば確認するとよい.

上の定理では, σ を表す互換の積の個数の最小値が, $n-1$ 以下であることを示した. 最小値なので, もちろんそれよりも多い回数の互換の積によっても, σ を表すことができる. 例えば簡単な例ではあるが, ある 2 数を交換した後, また同じ 2 数を交換すれば, 元通りになる. では, 前章でも言ったように偶奇性について証明しよう.

定理 26.4 置換の集合 S_n の元の, 任意の置換 σ が与えられたとする. また τ_1, \dots, τ_k と ρ_1, \dots, ρ_ℓ は S_n における互換を表すとする. σ は, 2 通りの異なる互換の積によって, 以下のように

$$\sigma = \tau_1\tau_2\cdots\tau_k = \rho_1\rho_2\cdots\rho_\ell$$

と 2 通りの積で表されたとする. このとき,

- (i). k と l はともに偶数である.
- (ii). k と l はともに奇数である.

のいずれかである. ┘

注意: この証明は, 参考文献の [斉藤][佐竹][川久保] でも, 同じようなものなので, そちらをみてもよい.

証明 n 変数の多項式

$$\begin{aligned} \Delta(x_1, x_2, \dots, x_n) &= \prod_{1 \leq i < j \leq n} (x_i - x_j) \\ &= \begin{array}{ccccccc} (x_1 - x_2) & (x_1 - x_3) & \cdots & (x_1 - x_{n-1}) & (x_1 - x_n) \\ & (x_2 - x_3) & \cdots & (x_2 - x_{n-1}) & (x_2 - x_n) \\ & & \cdots & \cdots & \cdots \\ & & & (x_{n-2} - x_{n-1}) & (x_{n-2} - x_n) \\ & & & & (x_{n-1} - x_n) \end{array} \end{aligned}$$

を考える. この式を n 変数の差積と呼ぶ.

次に置換 σ のときの, n 変数の差積を

$$\Delta^\sigma(x_1, x_2, \dots, x_n) = \Delta(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})$$

と定義する. そうすると, 任意の互換 $\tau \in S_n$ については,

$$\Delta^\tau(x_1, x_2, \dots, x_n) = -\Delta(x_1, x_2, \dots, x_n) \quad (26.2)$$

が成り立つが, これは後で示す. 従って (26.2) 式と仮定より,

$$\Delta^\sigma(x_1, x_2, \dots, x_n) = \begin{cases} (-1)^k \Delta(x_1, x_2, \dots, x_n) \\ (-1)^\ell \Delta(x_1, x_2, \dots, x_n) \end{cases} \quad (26.3)$$

が成り立つ. この式から, $(-1)^k = (-1)^\ell$ が成り立つので, 命題が成り立つ.

任意の互換 $\tau = \begin{pmatrix} s & t \\ & \end{pmatrix}$ (s, t は 1 以上 n 以下の整数で $s \neq t$) について (26.2) 式が成り立つことの証明:

$s < t$ としても一般性は失わないので, そのように仮定する. 差積 $\Delta(x_1, x_2, \dots, x_n)$ から, 差積 $\Delta^\tau(x_1, x_2, \dots, x_n)$ へと変更するとき, $\Delta(x_1, x_2, \dots, x_n)$ の各 $(x_i - x_j)$ の中で, 互換 τ によって変わる部分は,

- (i). $(x_s - x_j)$ が $(x_t - x_j)$ へ変更 ($j = s + 1, \dots, t$)
- (ii). $(x_s - x_t)$ が $(x_t - x_s)$ へ変更
- (iii). $(x_s - x_j)$ が $(x_t - x_j)$ へ変更 ($j = t + 1, \dots, n$)
- (iv). $(x_t - x_j)$ が $(x_s - x_j)$ へ変更 ($j = t + 1, \dots, t$)
- (v). $(x_i - x_s)$ が $(x_i - x_t)$ へ変更 ($i = 1, \dots, s - 1$)
- (vi). $(x_i - x_t)$ が $(x_i - x_s)$ へ変更 ($i = 1, \dots, s - 1$)
- (vii). $(x_i - x_t)$ が $(x_i - x_s)$ へ変更 ($i = s + 1, \dots, t - 1$)

である. これを掛け合わせたものが差積 $\Delta(x_1, x_2, \dots, x_n)$ から, 差積 $\Delta^\tau(x_1, x_2, \dots, x_n)$ へと変更するとき, 変わる部分である. 掛けると, (i) と (vii), (iii) と (iv), (v) と (vi) がそれぞれ打ち消しあうため, 結局

(ii) のみを考えればよい。そこで (ii) をよく見ると、これは $(x_s - x_t)$ を、 $-(x_s - x_t)$ へ変更するということになるので、結局 (26.2) 式が成り立つ

□

この結果から $\sigma \in S_n$ において、符号 $\text{sign}(\sigma)$ がきちんと定義できる。

定義 26.5 $\sigma \in S_n$ において、

- (i). σ が偶数回数の互換で積で表されるものを偶置換といい、奇数回数の互換の積で表されるものを奇置換と定義する。
- (ii). σ の符号 $\text{sign}(\sigma)$ を、

$$\text{sign}(\sigma) = \begin{cases} 1 & (\sigma \text{ が偶置換のとき}) \\ -1 & (\sigma \text{ が奇置換のとき}) \end{cases}$$

」

sign は、符号の英語 sign の省略形である。さていよいよこれで、ようやく n 次正方行列の行列式を定義できる。

定義 26.6 n 次正方行列 $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ の行列式 $\det(\mathbf{A})$ を、

$$\det(\mathbf{A}) = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^n a_{i\sigma(i)} \quad (26.4)$$

と定義する。

」

$a_{i\sigma(i)}$ は、 \mathbf{A} の $(i, \sigma(i))$ 成分のことである。

もちろん、 2×2 の行列式の定義も、 3×3 の行列式の定義も、この定義に含まれている。また 4×4 の行列式の候補も、今や候補ではなく、この定義によって、 4×4 の行列式を書き下したものとなっている。

最後に、互換以外に重要な置換として、巡回置換を述べておくことにしよう。

定義 26.7 相異なる n 個の数 i_1, i_2, \dots, i_n について、置換

$$\sigma = \begin{pmatrix} i_1 & i_2 & \cdots & i_{n-1} & i_n \\ i_2 & i_3 & \cdots & i_n & i_1 \end{pmatrix}$$

を巡回置換と呼び、

$$\sigma = (i_1 \ i_2 \ \cdots \ i_{n-1} \ i_n)$$

で表す。

」

これは i_1 から i_{n-1} までの順番を変えずに、 i_n を i_1 の場所に持っていくことができる置換である。

定義 26.8 相異なる n 個の数 i_1, i_2, \dots, i_n の巡回置換は、 $n-1$ 個の互換の積で表すことができる。

」

証明

巡回置換

$$(1 \ 2 \ \cdots \ n-1 \ n)$$

を考えたも一般性は失わない。このとき、

$$(1 \ 2 \ \cdots \ n-1 \ n) = (1 \ n)(2 \ n) \cdots (n-2 \ n)(n-1 \ n)$$

が成り立つ。右辺は $n-1$ 個の互換の積である。

□

この巡回置換の性質は余因子展開のときに用いる。

27 行列式の性質とクラメールの公式

この章では、 \mathbf{A} は n 次正方行列

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad (27.1)$$

を表すものとする。

定理 27.1 (27.1) 式の \mathbf{A} の転置行列をとっても逆行列の値は変わらない。つまり、

$$\det({}^t\mathbf{A}) = \det(\mathbf{A}) \quad (27.2)$$

□

証明 $\sigma \in S_n$ の逆置換を σ^{-1} で表すとする。また $\mathbf{B} = {}^t\mathbf{A}$ とする。

$$\begin{aligned} \det(\mathbf{A}) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^n a_{i\sigma(i)} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^n a_{\sigma^{-1}(i)i} \end{aligned}$$

$$\text{明らかに } \text{sign}(\sigma^{-1}) = \text{sign}(\sigma) \text{ なので, } = \sum_{\sigma \in S_n} \text{sign}(\sigma^{-1}) \prod_{i=1}^n a_{\sigma^{-1}(i)i}$$

$\sum_{\sigma \in S_n}$ という形から σ は S_n 全てに渡っており、よって σ^{-1} も S_n 全てを渡っている。 σ^{-1} を τ で表せば

$$\begin{aligned} &= \sum_{\tau \in S_n} \text{sign}(\tau) \prod_{i=1}^n a_{\tau(i)i} \\ &= \sum_{\tau \in S_n} \text{sign}(\tau) \prod_{i=1}^n b_{i\tau(i)} \\ &= \det(\mathbf{B}) \end{aligned}$$

□

この証明から、以下のように行を $\sigma(i)$ で動かすことによっても、行列式が得られる。(26.6) の定義とよく比べるとよい。

系 27.2 (27.1) 式の \mathbf{A} について、

$$\det(\mathbf{A}) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^n a_{\sigma(i)i} \quad (27.3)$$

」

$a_{\sigma(i)i}$ は、 \mathbf{A} の $(\sigma(i), i)$ 成分のことである。

さて 2×2 の行列の場合で示した定理 5.2 が、一般の n 次正方行列についても当然成り立つことを示すために、まず記号を簡単にするために、以下の記号を定義したあとに、それを示す。

定義 27.3 $m \times n$ の行列 \mathbf{A} が与えられたとき、 \mathbf{A} の i 列のみを \mathbf{b} に変更した行列を、

$$\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{b})}$$

で表す。同様に \mathbf{A} の i 列と j 列を交換した行列を、

$$\mathbf{A}_{(\mathbf{a}_i \leftrightarrow \mathbf{a}_j)}$$

」

定理 27.4 (行列式の基本的な性質) (27.1) 式の n 次正方行列 \mathbf{A} について、

- (i). \mathbf{A} のある 2 つの列を入れ替えた行列の行列式は、 $-\det(\mathbf{A})$ である。つまり、

$$\det(\mathbf{A}_{(\mathbf{a}_i \leftrightarrow \mathbf{a}_j)}) = -\det(\mathbf{A}) \quad (\text{if } i \neq j)$$

同様にある 2 つの行を入れ替えた行列式は、 $-\det(\mathbf{A})$ である。

- (ii). \mathbf{A} において、ある列がすべて 0 ならば、 $\det(\mathbf{A}) = 0$ である。つまり、

$$\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{0})} = 0$$

同様に、ある行がすべて 0 ならば、 $\det(\mathbf{A}) = 0$ である。

- (iii). $i \neq j$ とする。 \mathbf{A} の第 i 列と、第 j 列が等しいとき、行列式は $\det(\mathbf{A}) = 0$ である。つまり、

$$\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{a}_j)} = 0 \quad (\text{if } i \neq j)$$

同様に \mathbf{A} の第 i 行と、第 j 行が等しいとき、 $\det(\mathbf{A}) = 0$ である。

- (iv). \mathbf{A} のある列を全て c 倍した行列の行列式は、 $c \cdot \det(\mathbf{A})$ である。つまり、

$$\mathbf{A}_{(\mathbf{a}_i \leftarrow c\mathbf{a}_i)} = c \det(\mathbf{A})$$

同様に、 \mathbf{A} のある行を全て c 倍した行列の行列式は、 $c \cdot \det(\mathbf{A})$ である。

- (v). \mathbf{b}, \mathbf{c} を n 項列ベクトルとする。 \mathbf{A} の第 i 列が $\mathbf{b} + \mathbf{c}$ で表されるとき、以下が成り立つ。

$$\det(\mathbf{A}) = \det(\mathbf{A} \text{ の } i \text{ 列を } \mathbf{b} \text{ に置き換えた行列}) + \det(\mathbf{A} \text{ の } i \text{ 列を } \mathbf{c} \text{ に置き換えた行列})$$

記号で書くと、

$$\det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{b} + \mathbf{c})}) = \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{b})}) + \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{c})})$$

また, \mathbf{A} の i 行目が ${}^t\mathbf{b} + {}^t\mathbf{c}$ で表されるとき, 以下が成り立つ.

$$\det(\mathbf{A}) = \det(\mathbf{A} \text{ の } i \text{ 行を } {}^t\mathbf{b} \text{ に置き換えた行列}) + \det(\mathbf{A} \text{ の } i \text{ 行を } {}^t\mathbf{c} \text{ に置き換えた行列})$$

(vi). \mathbf{A} のある列に, 別の列のスカラー倍を加えた行列の行列式は, $\det(\mathbf{A})$ のまま変わらない. つまり,

$$\det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{a}_i + c\mathbf{a}_j)}) = \det(\mathbf{A}) \quad (\text{if } i \neq j)$$

同様に, \mathbf{A} のある行に, 別の行のスカラー倍を加えた行列の行列式は, $\det(\mathbf{A})$ のまま変わらない.

」

注意: 全て列についてのみ証明する. というのも定理 27.1 より, 転置行列を考えれば, 行についても同様のことが言えるためである.

証明

(i) と (ii) と (iv) は, 行列式の定義より明らか.

(iii) の証明: \mathbf{A} の i 列目と j 列目を入れ替えた行列を \mathbf{B} とすると, (i) より,

$$\det(\mathbf{B}) = -\det(\mathbf{A})$$

ところで, 今 i 列目と j 列目の数字は等しいので, $\det(\mathbf{B}) = \det(\mathbf{A})$. この 2 式より $\det(\mathbf{A}) = 0$.

(v) の証明:

$$\begin{aligned} \det(\mathbf{A}_{\mathbf{a}_i \leftarrow \mathbf{b} + \mathbf{c}}) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \left[\left(\prod_{k=1}^{i-1} a_{\sigma(k)k} \right) (b_{\sigma(i)i} + c_{\sigma(i)i}) \left(\prod_{k=i+1}^n a_{\sigma(k)k} \right) \right] \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \left[\left(\prod_{k=1}^{i-1} a_{\sigma(k)k} \right) b_{\sigma(i)i} \left(\prod_{k=i+1}^n a_{\sigma(k)k} \right) \right] + \sum_{\sigma \in S_n} \text{sign}(\sigma) \left[\left(\prod_{k=1}^{i-1} a_{\sigma(k)k} \right) c_{\sigma(i)i} \left(\prod_{k=i+1}^n a_{\sigma(k)k} \right) \right] \\ &= \det(\mathbf{a}_1 \ \cdots \ \mathbf{a}_{i-1} \ \mathbf{b} \ \mathbf{a}_{i+1} \ \cdots \ \mathbf{a}_n) + \det(\mathbf{a}_1 \ \cdots \ \mathbf{a}_{i-1} \ \mathbf{c} \ \mathbf{a}_{i+1} \ \cdots \ \mathbf{a}_n) \end{aligned}$$

(vi) の証明: $i \neq j$ とする.

$$\begin{aligned} \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{a}_i + c\mathbf{a}_j)}) &= \det(\mathbf{A}) + \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow c\mathbf{a}_j)}) \\ &\quad ((\text{iv}) \text{ より}) = \det(\mathbf{A}) + c \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{a}_j)}) \\ &\quad ((\text{iii}) \text{ より}) = \det(\mathbf{A}) \end{aligned}$$

□

3 次正方行列のクラメールの公式 ((21.7) 式) の導出では, 成分ごとに計算した. 一般に $n \times n$ のときは, 成分ごとではなく, 上の行列式の性質から導出できる.

定理 27.5 (クラメールの公式) n 次正方行列 \mathbf{A} (ただし $\det \mathbf{A} \neq 0$) と, 2 つの n 項列ベクトル \mathbf{x}, \mathbf{b} について

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

が成り立つとする. \mathbf{x} の第 i 成分 x_i は,

$$x_i = \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{b})}) / \det(\mathbf{A})$$

」

証明 \mathbf{A} の第 j 列を \mathbf{a}_j で表す. 定義より, $\mathbf{b} = \mathbf{Ax} = \sum_{j=1}^n x_j \mathbf{a}_j$ である.

$$\begin{aligned} \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{b})}) &= \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \sum_{j=1}^n x_j \mathbf{a}_j)}) \\ \text{前定理 (v) より} &= \sum_{j=1}^n \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow x_j \mathbf{a}_j)}) \\ \text{前定理 (iv) より} &= \sum_{j=1}^n x_j \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{a}_j)}) \\ &= x_i \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{a}_i)}) + \sum_{j \neq i} x_j \det(\mathbf{A}_{(\mathbf{a}_i \leftarrow \mathbf{a}_j)}) \\ \text{前定理 (iii) より} &= x_i \det(\mathbf{A}) + 0 \end{aligned}$$

□

28 n 次正方行列について, 積の行列式 = 行列式の積 の証明

定理 5.1 では, 2 次正方行列において, $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$ を示した. n 次正方行列でも同様に成り立つ. 証明を読んでいて式がイメージしにくいときは, 定理 5.1 の証明 (5 章の一番最後) を参照するとよいだろう.

定理 28.1 \mathbf{A}, \mathbf{B} とともに n 次正方行列とする. このとき, 以下が成り立つ.

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$$

┘

証明 \mathbf{A} を, n 個の n 項列ベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ で以下のように表す.

$$\mathbf{A} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n)$$

\mathbf{B} は

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

とする. このとき,

$$\mathbf{AB} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n) \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix} = \left(\sum_{i=1}^n \mathbf{a}_i b_{i1} \quad \sum_{i=1}^n \mathbf{a}_i b_{i2} \quad \cdots \quad \sum_{i=1}^n \mathbf{a}_i b_{in} \right)$$

と表すことができる. 右辺の各列の添字を i ではなくて, それぞれの列に応じて i_1, i_2, \dots, i_n と変更した後, 両辺の行列式を求めると,

$$\begin{aligned}\det(\mathbf{AB}) &= \det\left(\sum_{i_1=1}^n \mathbf{a}_{i_1} b_{i_1 1} \quad \sum_{i_2=1}^n \mathbf{a}_{i_2} b_{i_2 2} \quad \cdots \quad \sum_{i_n=1}^n \mathbf{a}_{i_n} b_{i_n n}\right) \\ &= \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_n=1}^n (b_{i_1 1} b_{i_2 2} \cdots b_{i_n n}) \det(\mathbf{a}_{i_1} \quad \mathbf{a}_{i_2} \quad \cdots \quad \mathbf{a}_{i_n})\end{aligned}$$

これは n^n 個の行列式の和である。ところで行列式

$$\det(\mathbf{a}_{i_1} \quad \mathbf{a}_{i_2} \quad \cdots \quad \mathbf{a}_{i_n})$$

の添字 i_1, i_2, \dots, i_n は全て 1 から n までの整数のどれかの値である。行列式の性質 (定理 27.4 の (iii)) から n 個の数字の中に、1 組でも同じものが存在すると、この値は 0 となる。従って、 n 文字の置換の集合を S_n とすると、

$$\begin{aligned}\det(\mathbf{AB}) &= \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_n=1}^n (b_{i_1 1} b_{i_2 2} \cdots b_{i_n n}) \det(\mathbf{a}_{i_1} \quad \mathbf{a}_{i_2} \quad \cdots \quad \mathbf{a}_{i_n}) \\ &= \sum_{i_1, i_2, \dots, i_n \text{ は相異なる値で } 1 \leq i_1, i_2, \dots, i_n \leq n} [(b_{i_1 1} b_{i_2 2} \cdots b_{i_n n}) \det(\mathbf{a}_{i_1} \quad \mathbf{a}_{i_2} \quad \cdots \quad \mathbf{a}_{i_n})] \\ &\quad (\text{この和記号の条件は、置換 } \sigma \in S_n \text{ を用いて } i_1 = \sigma(1), i_2 = \sigma(2), \dots, i_n = \sigma(n) \text{ として、} \\ &\quad \sigma \text{ が } S_n \text{ 全体を渡ることに同じなので}) \\ &= \sum_{\sigma \in S_n} (b_{\sigma(1)1} b_{\sigma(2)2} \cdots b_{\sigma(n)n}) \det(\mathbf{a}_{\sigma(1)} \quad \mathbf{a}_{\sigma(2)} \quad \cdots \quad \mathbf{a}_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} (b_{\sigma(1)1} b_{\sigma(2)2} \cdots b_{\sigma(n)n}) \text{sign}(\sigma) \det(\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n) \\ &= \det(\mathbf{A}) \sum_{\sigma \in S_n} (b_{\sigma(1)1} b_{\sigma(2)2} \cdots b_{\sigma(n)n}) \text{sign}(\sigma) \\ &= \det(\mathbf{A}) \det(\mathbf{B})\end{aligned}$$

□

29 正方行列の余因子展開

3×3 以上の行列式は、そのまま計算するわけではない。定理 (27.4) の性質などを用いて、簡単にした後、計算するのである。また余因子展開という方法も用いるので、それを説明しよう。

定義 29.1 n 次正方行列 \mathbf{A} について、

- (i). i 行目と j 列目を除いてその部分を詰めた $n-1$ 次正方行列の行列式を、 \mathbf{A} の (i, j) 小行列式という。
- (ii). \mathbf{A} の (i, j) 小行列式に $(-1)^{i+j}$ を掛けたものを、 \mathbf{A} の (i, j) 余因子といい、 Δ_{ij} で表す。

┘

例題 29.2 次の \mathbf{A} について、 $(1, 1)$ 小行列式、 $(2, 1)$ 小行列式、 $(3, 1)$ 小行列式をそれぞれ求めよ。 $(1, 2)$ 余

因子, (2, 2) 余因子, (3, 2) 余因子をそれぞれ求めよ.

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad (29.1)$$

(解答)

(1, 1) 小行列式, (2, 1) 小行列式, (3, 1) 小行列式は, それぞれ

$$\det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, \quad \det \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}, \quad \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}$$

それぞれ計算して, 順に,

$$-3, \quad -6, \quad -3$$

である. (1, 2) 余因子, (2, 2) 余因子, (3, 2) 余因子は, それぞれ,

$$(-1)^{1+2} \det \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix}, \quad (-1)^{2+2} \det \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix}, \quad (-1)^{3+2} \det \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix}$$

である. 計算して順に,

$$6, \quad -11, \quad 6$$

(解答終)

余因子が重要である理由は, 行列式を展開できる点と, さらに, 逆行列を余因子を用いて求めることができる点にある. 後者は次章に回して, この章では行列式の実展開について述べよう.

定理 29.3 (余因子展開) n 次正方行列 $\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ を考える.

(i). (j 列の余因子展開) 任意の $j = 1, 2, \dots, n$ について, 以下が成り立つ.

$$\det(\mathbf{A}) = \sum_{i=1}^n a_{ij} \Delta_{ij}$$

(ii). (i 行の余因子展開) 任意の $i = 1, 2, \dots, n$ について, 以下が成り立つ.

$$\det(\mathbf{A}) = \sum_{j=1}^n a_{ij} \Delta_{ij}$$

証明 (ii) は (i) と同様にできるので, (i) のみ証明する. n 項列ベクトル \mathbf{e}_i (ただし $i = 1, 2, \dots, n$) は, 第 i 成分のみ 1 で, 残りの成分は 0 であるとする.

[手順1] \mathbf{A} の1列目が \mathbf{e}_1 の場合を考える. このとき1から n までの文字の置換の集合を S_n で表し, 2から n までの文字の置換の集合を S_{n-1} で表すとする. このとき,

$$\begin{aligned} \det(\mathbf{e}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \\ &= \sum_{\substack{\sigma \in S_n, \\ \sigma(1)=1}} \text{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_{n-1}} \text{sign}(\sigma) a_{\sigma(2)2} \cdots a_{\sigma(n)n} \\ &= \det \begin{pmatrix} a_{22} & \cdots & a_{n2} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix} \\ &= \Delta_{11} \end{aligned}$$

[手順2] \mathbf{A} の1列目が \mathbf{e}_i の場合を考える. \mathbf{A} の1行目から i 行目までに, 巡回置換 $\sigma = (1 \ 2 \ \cdots \ i)$ を, \mathbf{A} に施した行列式を考えると,

$$\begin{aligned} \det(\mathbf{e}_i \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n) &= \text{sign}(\sigma) \det \begin{pmatrix} 1 & a_{i2} & \cdots & a_{in} \\ 0 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{(i-1)2} & \cdots & a_{(i-1)n} \\ 0 & a_{(i+1)2} & \cdots & a_{(i+1)n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix} \\ &\text{\scriptsize } \sigma \text{ は } i-1 \text{ 回の互換の積で表されることと, 手順1から} \\ &= (-1)^{i-1} \cdot (\mathbf{A} \text{ の } (i, 1) \text{ 小行列式}) \\ &= (-1)^{i+1} \cdot (\mathbf{A} \text{ の } (i, 1) \text{ 小行列式}) \\ &= \Delta_{i1} \end{aligned}$$

[手順3] $j=1$ のとき命題が成り立つことを示す. \mathbf{A} の1列目は, $\sum_{i=1}^n a_{i1} \mathbf{e}_i$ で表すことができるので,

$$\begin{aligned} \det(\mathbf{A}) &= \det \left(\sum_{i=1}^n a_{i1} \mathbf{e}_i \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \right) \\ &= \sum_{i=1}^n a_{i1} \det(\mathbf{e}_i \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n) \\ \text{手順2より} \quad &= \sum_{i=1}^n a_{i1} \Delta_{i1} \end{aligned}$$

これで $j=1$ のとき命題が成り立つ.

[手順4] $j \geq 2$ のときは, 各列に巡回置換 $\sigma = (1 \ 2 \ \cdots \ j)$ を施すことによって,

$$\begin{aligned}
\det(\mathbf{A}) &= \text{sign}(\sigma) \det(\mathbf{a}_j \ \mathbf{a}_1 \ \cdots \ \mathbf{a}_{j-1} \ \mathbf{a}_{j+1} \ \cdots \ \mathbf{a}_n) \\
&\quad \sigma \text{ は } j-1 \text{ 回の互換の積で表されるので} \\
&= (-1)^{j-1} \det(\mathbf{a}_j \ \mathbf{a}_1 \ \cdots \ \mathbf{a}_{j-1} \ \mathbf{a}_{j+1} \ \cdots \ \mathbf{a}_n) \\
\text{手順 3 より} \quad &= (-1)^{j-1} \sum_{i=1}^n (-1)^{i+1} a_{ij} \det(\mathbf{e}_i \ \mathbf{a}_1 \ \cdots \ \mathbf{a}_{j-1} \ \mathbf{a}_{j+1} \ \cdots \ \mathbf{a}_n) \\
&= \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(\mathbf{e}_i \ \mathbf{a}_1 \ \cdots \ \mathbf{a}_{j-1} \ \mathbf{a}_{j+1} \ \cdots \ \mathbf{a}_n) \\
&= \sum_{i=1}^n a_{ij} \Delta_{ij}
\end{aligned}$$

□

30 余因子展開から逆行列へ

すでに 2 次正方行列と 3 次正方行列の場合に、逆行列やクラメールの公式を導いた。一般の n 次正方行列の場合は、余因子展開を用いることによって、それらが公式として得られる。まずは簡単な定義をしておこう。

定義 30.1 (クロネッカーのデルタ) i, j に関する実数 δ_{ij} を、以下のように定める。

$$\delta_{ij} = \begin{cases} 1 & (i = j \text{ のとき}) \\ 0 & (\text{その他のとき}) \end{cases}$$

┘

この δ_{ij} を用いて、余因子展開についての公式を表そう。

定理 30.2 (余因子展開) n 次正方行列 \mathbf{A} の (i, j) 余因子を Δ_{ij} とする。このとき、

$$\begin{aligned}
\text{(i).} \quad & \sum_{k=1}^n a_{ki} \Delta_{kj} = \delta_{ij} \cdot \det(\mathbf{A}) \\
\text{(ii).} \quad & \sum_{k=1}^n a_{ik} \Delta_{jk} = \delta_{ij} \cdot \det(\mathbf{A})
\end{aligned}$$

┘

証明 $i = j$ のときは、定理 29.3 より証明されている。よって $i \neq j$ のときを示す。(ii) は (i) と同様にできるので、(i) の場合のみ証明する。

i, j, k は 1 から n までの整数とし、まず任意の $i, j (i \neq j)$ を定める、 n 次正方行列 \mathbf{B} を、 i 列以外は \mathbf{A} と同じもので、 i 列目は、 \mathbf{A} の j 列目と同じものとする。すると \mathbf{B} の (k, i) 余因子は、 \mathbf{A} の (k, i) 余因子と一致

する. 明らかに $\det(\mathbf{B}) = 0$ である. そこで \mathbf{B} の i 列目を余因子展開すると,

$$\begin{aligned} 0 &= \det(\mathbf{B}) \\ &= \sum_{k=1}^n b_{ki} \Delta_{ki} \\ &= \sum_{k=1}^n a_{kj} \Delta_{ki} \end{aligned}$$

i, j は任意なので, (i) が証明された. □

定義 30.3 (余因子行列) \mathbf{A} と Δ_{ij} は前定理同様とする. このとき, (j, i) 成分が Δ_{ij} となる n 次正方行列 $\tilde{\mathbf{A}}$ を, \mathbf{A} の余因子行列と呼ぶ. 式で書くと以下である.

$$\tilde{\mathbf{A}} = \begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix} \quad (30.1)$$

$\tilde{\mathbf{A}}$ の (j, i) 成分が \mathbf{A} の (i, j) 余因子である. このように (i, j) でなく (j, i) としたのは, 定理 30.2 の系として, 次の式が成り立つからである.

系 30.4 n 次正方行列 \mathbf{A} とその余因子行列 $\tilde{\mathbf{A}}$ について,

$$\mathbf{A}\tilde{\mathbf{A}} = \tilde{\mathbf{A}}\mathbf{A} = \det(\mathbf{A}) \cdot \mathbf{I}$$

が成り立つ. ┘

よってこの式から逆行列は簡単に得られる.

定理 30.5 n 次正方行列 \mathbf{A} の余因子行列を $\tilde{\mathbf{A}}$ とする.

(i). \mathbf{A} が正則のとき, \mathbf{A}^{-1} は

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \tilde{\mathbf{A}}$$

(ii). \mathbf{A} は正則であることは, $\det(\mathbf{A}) \neq 0$ と同値である. ┘

注意: (i) の対偶から, 「 $\det(\mathbf{A}) = 0$ であることと, \mathbf{A} は正則ではないことと同値である。」ことも言える. (注意終)

証明 (i) の証明と (ii) の \Leftarrow の証明は系 30.4 から得られる.

(i) の \Rightarrow の証明: \mathbf{A} が正則なので,

$$\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A} = \mathbf{I}$$

を満たす \mathbf{B} が存在する. 両辺の行列式を取ると, $\det(\mathbf{A}\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$ より, $\det(\mathbf{A})\det(\mathbf{B}) = 1$ が成り立つ. よって $\det(\mathbf{A}) \neq 0$.

□

この定理の (i)(ii) はそれぞれ背反なので, (i)(ii) の逆も成り立つ.

上の式のように A の逆行列の定義は, $AB = BA = I$ となるような B のことであった. しかしこの式のように $AB = I, BA = I$ と, 2つの等式を示すのは面倒なので, 次の系が便利である.

系 30.6 n 次正方行列 A, B が $AB = I$ を満たすならば, A も B も正則で,

$$A^{-1} = B, \quad B^{-1} = A$$

┘

証明 $AB = I$ の両辺の行列式をとると, $\det(A)\det(B) = 1$. よって, $\det(A) \neq 0, \det(B) \neq 0$ が成り立つので, A, B ともに正則である.

A^{-1} を $AB = I$ の左から掛けると, $B = A^{-1}$ が得られる. B^{-1} を $AB = I$ の右から掛けると $A = B^{-1}$ が得られる. □

31 分割された行列と, 行列式・逆行列の表記

6章で説明したものを一般化し, さらに行列式や逆行列を求めるのが, この章の目的である. この章では, I_m, I_n はそれぞれ m 次と n 次の単位行列とする. O_{mn} は m 行 n 列のゼロ行列を表し, O_{nm} は n 行 m 列のゼロ行列を表す.

n 行 m 列の行列 A と, m 行 ℓ 列の行列 B で積 AB は計算可能である. このとき A を行ベクトル ${}^t a_i$ の集合 ($i = 1, 2, \dots, m$), B を列ベクトル b_j の集合 ($j = 1, 2, \dots, \ell$) として, 積 AB は

$$AB = \begin{pmatrix} {}^t a_1 \\ \vdots \\ {}^t a_m \end{pmatrix} (b_1 \quad \cdots \quad b_\ell) = \begin{pmatrix} {}^t a_1 b_1 & \cdots & {}^t a_1 b_\ell \\ \vdots & \ddots & \vdots \\ {}^t a_m b_1 & \cdots & {}^t a_m b_\ell \end{pmatrix}$$

と定義した. これをより一般化したものが以下となる. 証明は簡単なのではない.

命題 31.1 n 行 m 列の行列 A と, m 行 ℓ 列の行列 B を考える. I, J, K は整数で, それぞれ n 以下, k 以下, ℓ 以下とする. A_{ij} (i は 1 から I までの整数で j は 1 から J までの整数) は m_i 行 n_j 列の行列とし, A を IJ 個の行列の集合とみて,

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1J} \\ A_{21} & A_{22} & \cdots & A_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ A_{I1} & A_{I2} & \cdots & A_{IJ} \end{pmatrix}$$

のように分割するとする. ただし各 m_i, n_j の値は $\sum_{i=1}^I m_i = m, \sum_{j=1}^J n_j = n$ を満たす正整数とする.

次に B_{jk} (j と n_j は上の通りで k は 1 から K までの整数) は, n_j 行 ℓ_k 列の行列とし, B を JK 個の行列の集合とみて,

$$B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1K} \\ B_{21} & B_{22} & \cdots & B_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ B_{J1} & B_{J2} & \cdots & B_{JK} \end{pmatrix}$$

のように分割するとする. ただし各 l_k は $\sum_{k=1}^K l_k = \ell$ を満たす正整数とする. このとき積 \mathbf{AB} は, 以下のよう表すことができる.

$$\mathbf{AB} = \begin{pmatrix} \sum_{j=1}^J \mathbf{A}_{1j} \mathbf{B}_{j1} & \sum_{j=1}^J \mathbf{A}_{1j} \mathbf{B}_{j2} & \cdots & \sum_{j=1}^J \mathbf{A}_{1j} \mathbf{B}_{jK} \\ \sum_{j=1}^J \mathbf{A}_{2j} \mathbf{B}_{j1} & \sum_{j=1}^J \mathbf{A}_{2j} \mathbf{B}_{j2} & \cdots & \sum_{j=1}^J \mathbf{A}_{2j} \mathbf{B}_{jK} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^J \mathbf{A}_{Ij} \mathbf{B}_{j1} & \sum_{j=1}^J \mathbf{A}_{Ij} \mathbf{B}_{j2} & \cdots & \sum_{j=1}^J \mathbf{A}_{Ij} \mathbf{B}_{jK} \end{pmatrix} \quad (31.1)$$

┘

いくつか簡単な例を上げておこう. まず 2×2 の形である. $I = J = K = 2$ の場合を考えよう. n 行 m 列の行列 \mathbf{A} と, m 行 k 列の行列 \mathbf{B} について, 命題のように正整数 $m_1, m_2, n_1, n_2, l_1, l_2$ をとって,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}$$

と置くと, \mathbf{AB} は,

$$\mathbf{AB} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{11} \mathbf{B}_{11} + \mathbf{A}_{12} \mathbf{B}_{21} & \mathbf{A}_{11} \mathbf{B}_{12} + \mathbf{A}_{12} \mathbf{B}_{22} \\ \mathbf{A}_{21} \mathbf{B}_{11} + \mathbf{A}_{22} \mathbf{B}_{21} & \mathbf{A}_{21} \mathbf{B}_{12} + \mathbf{A}_{22} \mathbf{B}_{22} \end{pmatrix} \quad (31.2)$$

で表される. この 2×2 の形での積が証明でよく出てくる形なので, これを理解できればまずは OK である.

次に $I = 1, J = K = 2$ の場合の命題 31.1 でも, 正整数 n_1, n_2, l_1, l_2 をとって,

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{11} \mathbf{B}_{11} + \mathbf{A}_{12} \mathbf{B}_{21} & \mathbf{A}_{11} \mathbf{B}_{12} + \mathbf{A}_{12} \mathbf{B}_{22} \end{pmatrix}$$

として成り立つ.

$I = 1, J = 2, K = 1$ の場合だと, 正整数 n_1, n_2 をとって,

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{11} \\ \mathbf{B}_{21} \end{pmatrix} = \mathbf{A}_{11} \mathbf{B}_{11} + \mathbf{A}_{12} \mathbf{B}_{21}$$

が成り立つ.

また $I = 2, J = 1, K = 2$ の場合の命題 31.1 を考えると, 正整数 m_1, m_2, l_1, l_2 をとって,

$$\begin{pmatrix} \mathbf{A}_{11} \\ \mathbf{A}_{21} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{11} \mathbf{B}_{11} & \mathbf{A}_{11} \mathbf{B}_{12} \\ \mathbf{A}_{21} \mathbf{B}_{11} & \mathbf{A}_{21} \mathbf{B}_{12} \end{pmatrix} \quad (31.3)$$

もやはり成り立つ. この式で \mathbf{A}_{11} と \mathbf{A}_{21} の部分を \mathbf{A} に戻して,

$$\mathbf{A} \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \end{pmatrix} = \begin{pmatrix} \mathbf{AB}_{11} & \mathbf{AB}_{12} \end{pmatrix}$$

も成り立つし, (31.3) 式で, $\mathbf{B}_{11}, \mathbf{B}_{12}$ の部分を \mathbf{B} で表記すれば,

$$\begin{pmatrix} \mathbf{A}_{11} \\ \mathbf{A}_{21} \end{pmatrix} \mathbf{B} = \begin{pmatrix} \mathbf{A}_{11} \mathbf{B} \\ \mathbf{A}_{21} \mathbf{B} \end{pmatrix}$$

が成り立つ.

このように適切な長方形の行列をとることにより, 行列の積を簡単に表すことができる.

この 2×2 の形での積が証明でよく出てくる形であるといったが, その中でも正方行列を 2×2 で分割して, その対角部分も正方行列とするものが, よく出てくる. というのは, このように分割すると, 行列式の計算ができるのである.

定理 31.2 (分割された行列の行列式 1) $(m+n)$ 次正方行列 \mathbf{A} が,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{O}_{nm} & \mathbf{A}_{22} \end{pmatrix}$$

と分割されるとする. ただし, \mathbf{A}_{11} は m 次正方行列で, \mathbf{A}_{22} は n 次正方行列, \mathbf{A}_{12} は m 行 n 列である. このとき,

$$\det(\mathbf{A}) = \det(\mathbf{A}_{11}) \det(\mathbf{A}_{22})$$

」

注意: 上三角行列の行列式に類似していることに気が付くとよい.

(注意終)

証明

$$\begin{aligned} \det(\mathbf{A}) &= \sum_{\sigma \in S_{m+n}} \text{sign}(\sigma) \left\{ \prod_{k=1}^m a_{\sigma(k)k} \prod_{k=m+1}^{m+n} a_{\sigma(k)k} \right\} \\ &= \sum_{\sigma \in S_{m+n}, i \leq m \Rightarrow \sigma(i) \leq m} \text{sign}(\sigma) \prod_{k=1}^m a_{\sigma(k)k} \prod_{k=m+1}^{m+n} a_{\sigma(k)k} \end{aligned}$$

$i \leq m \Rightarrow \sigma(i) \leq m$ ということは, σ による 1 から m までの置換は, 1 から m までということである. $m+1$ 以上の値を考えないとは, $i > m$ のとき, $a_{\sigma(i)i} = 0$ となるためである. すると σ による, $m+1, \dots, n$ の置換 $\sigma(m+1), \dots, \sigma(n)$ も, やはり $m+1$ から n までの値をとる. よって $1, 2, \dots, m$ の m 個の置換の 1 つを τ_1 とする. $1, 2, \dots, n$ の n 個の置換の 1 つを τ_2 とする. このとき, $\sigma = \tau_1 \tau_2$ で表すことができ, τ_1 を用いる $\sigma \in S_{n+m}$ を考えると, σ による $m+1$ から $m+n$ までの置換は, S_n 通り存在する. よって, 2 つの置換 $\tau_1 \in S_m, \tau_2 \in S_n$ を用いて,

$$\begin{aligned} \det(\mathbf{A}) &= \sum_{\tau_1 \in S_m} \text{sign}(\tau_1 \tau_2) \left\{ \prod_{k=1}^m a_{\tau_1(k)k} \sum_{\tau_2 \in S_n} \prod_{j=1}^n a_{(m+\tau_2(j))(m+j)} \right\} \\ &= \sum_{\tau_1 \in S_m} \text{sign}(\tau_1) \prod_{k=1}^m a_{\tau_1(k)k} \left\{ \sum_{\tau_2 \in S_n} \text{sign}(\tau_2) \prod_{j=1}^n a_{(m+\tau_2(j))(m+j)} \right\} \\ &= \sum_{\tau_1 \in S_m} \text{sign}(\tau_1) \prod_{k=1}^m a_{\tau_1(k)k} \det(\mathbf{A}_{22}) \\ &= \det(\mathbf{A}_{22}) \sum_{\tau_1 \in S_m} \text{sign}(\tau_1) \prod_{k=1}^m a_{\tau_1(k)k} \\ &= \det(\mathbf{A}_{22}) \det(\mathbf{A}_{11}) \end{aligned}$$

□

この定理では \mathbf{A} を上三角行列の形で分割したが, 下三角行列の形

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{O}_{mn} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

に分割した場合でも、同様の結果

$$\det(\mathbf{A}) = \det(\mathbf{A}_{11}) \det(\mathbf{A}_{22})$$

が得られる。ではゼロ行列がない場合はどうなるかという、

定理 31.3 (分割された行列の行列式 2) $(m+n)$ 次正方行列 \mathbf{A} を、

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

と分割する。ただし、 \mathbf{A}_{11} は m 次正方行列、 \mathbf{A}_{22} は n 次正方行列、 \mathbf{A}_{12} は m 行 n 列、 \mathbf{A}_{21} は n 行 m 列である。このとき、

(i). \mathbf{A}_{11} が正則ならば、

$$\det(\mathbf{A}) = \det(\mathbf{A}_{11}) \det(\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12})$$

(ii). \mathbf{A}_{22} が正則ならば、

$$\det(\mathbf{A}) = \det(\mathbf{A}_{22}) \det(\mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21})$$

」

証明 同じようにできるので (i) のみ証明する。

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \\ \mathbf{O}_{nm} & \mathbf{I}_n \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{O}_{mn} \\ \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \end{pmatrix} \quad (31.4)$$

が成り立つので、この両辺の行列式をとると、

$$\det(\mathbf{A}) \det(\mathbf{A}_{11}^{-1}) = \det(\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12})$$

$\det(\mathbf{A}_{11}^{-1}) = [\det(\mathbf{A}_{11})]^{-1}$ なので、上の式の両辺に $\det(\mathbf{A}_{11})$ を掛けることによって、命題が得られる。

□

次は 2×2 の形に分割した行列の、逆行列である。

定理 31.4 \mathbf{A} を定理 31.3 のように分割する。このとき、

(i). \mathbf{A}_{11} が正則で、 $\mathbf{F} = \mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}$ も正則ならば、

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{F}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{F}^{-1} \\ -\mathbf{F}^{-1} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} & \mathbf{F}^{-1} \end{pmatrix} \quad (31.5)$$

(ii). \mathbf{A}_{22} が正則で、 $\mathbf{F} = \mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21}$ も正則ならば、

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{F}^{-1} & -\mathbf{F}^{-1} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{F}^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{F}^{-1} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \end{pmatrix} \quad (31.6)$$

」

注意: 仮定と定理 31.3 によって、 $\mathbf{A} \neq \mathbf{0}$ がいえる。

(注意終)

これらは, (31.5) 式と (31.6) 式の, それぞれの右辺の左から (もしくは右から), $\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$ を掛けると, \mathbf{I} となることを示せばよい (計算してみるとよい).

とはいえ (i)(ii) ともに天下りすぎるので, なぜこのような式が得られたかを (i) のみ示そう. それは (31.4) 式の右辺を \mathbf{F} で表した後に, さらに両辺の右から $\begin{pmatrix} \mathbf{I}_m & \mathbf{O}_{mn} \\ -\mathbf{F}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{F}^{-1} \end{pmatrix}$ を掛けて,

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{O}_{nm} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \mathbf{O}_{mn} \\ -\mathbf{F}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{F}^{-1} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{O}_{mn} \\ \mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{F} \end{pmatrix} \begin{pmatrix} \mathbf{I}_m & \mathbf{O}_{mn} \\ -\mathbf{F}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{F}^{-1} \end{pmatrix}$$

両辺計算すると,

$$\mathbf{A} \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{F}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{F}^{-1} \\ -\mathbf{F}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{F}^{-1} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{O}_{mn} \\ \mathbf{O}_{nm} & \mathbf{I}_n \end{pmatrix} \quad (31.7)$$

右辺は単位行列である. よって系 30.6 より, \mathbf{A}^{-1} が命題の式のように得られるためである.

32 第 V 部の Scilab 実習

ここでは V 部の復習として Scilab の実習をまとめて行う. メイン画面で行うのみでなく, 15 部の 51 ページの (ii) での,

自作関数 (ユーザー定義関数) を作っておき, .sce ファイルで保存した後, それを実行して関数を読み込ませる. その後でメイン画面で, 関数を用いる.

を主に行う. 必要なものは, 自分で作ればよいのである.

まず, 順列についての復習をする. 1 から整数 n までの順列は $n!$ 個あるが, それを全て出力する `perms` 関数が用意されている. $n = 4$ について確かめてみよう.

例題 32.1 次のコードをメイン画面に打ち込め.

```
perms([4 3 2 1])
```

24 行 4 列の数字が表示されたことであろう. 各行が, 1,2,3,4 の順列を表している. では次に, この `perms` を用いて, 整数 `numN` を引数としたときに, 1 から `numN` までの順列を出して, さらにその各行の最終列には, $1, 2, \dots, n$ からその順列を得るための交換回数を書くようなプログラム `myperms(numN)` を書いて, 実行しよう.

例題 32.2 まず Scilab のメインから, アプリケーション > Scinotes で Scinotes を開くこと.

(i). 次のコマンドを, Scinotes に打ち込んで (もしくは, PDF ファイルからコピーして). 適当な名前前で保存せよ. その後, Scinote での, 「実行する > ファイルを実行」によって, そのファイルを実行させよ.

```
function outF=myperms(numN)
    if length(numN)>1 then
        error('numN must be scalar');
```

```

elseif find([2:6]==numN)==[]
    error('numN must be 2,3,4,5,6');
end
eachTBL=perms([numN:-1:1]);
nTBL=size(eachTBL,1);
ansTBL=zeros(nTBL,1);

for irow=1:nTBL
    vecQ=eachTBL(irow,:);
    for piv=1:numN-1
        indx=find(vecQ==piv);
        if indx~=piv then
            vecQ([piv indx])=vecQ([indx piv]);
            ansTBL(irow)=ansTBL(irow)+1;
        end
    end
end

outF=[eachTBL ansTBL];
endfunction

```

(ii). 次にメイン画面から,

```
myperms(4)
```

と打ち込むこと.

ではこのプログラムについて説明しておこう.

(a) 実行すると 24 行 5 列の行列が出てくる. 最初の 1 行目と 2 行目を見てみよう 1 行目は,

```
1 2 3 4 0
```

であろう. これは 1,2,3,4 の順を表し, 最後の 0 は, 1, 2, 3, 4 からは 0 回の交換で得られることを示している. 2 行目は,

```
1 2 4 3 1
```

である. これは 1,2,3,4 の順を表し, 最後の 1 は, 並び順 1,2,3,4 から 1,2,4,3 へは, 1 回の交換で得られることを示している. 25 節の最後にある 24 項の表を見て, 交換回数が一致していることを確認すると面白いだろう. またプログラムでは最大 6 まで動くようにしておいた. よって myperms(6) でも動くが, 720 行も出てくるので気を付けること.

(b) 変数名に注意しておこう. myperms 内の function と endfunction の中にある変数は,

```
eachTBL, nTBL, ansTBL, irow, vecQ, piv, indx, outF
```

の 7 つである。最後の `outF` が、関数を返す値 (今回は行列) である。Scilab では変数名を宣言しなくてもよいので、

変数を作るときには、プログラムにコメント文を書かなくともわかるような、わかりやすい変数名にする
とよい。

- (c) さらに、変数の範囲の問題があるので、これを説明しておこう。上にあげた変数は、関数 `myperms` の中でのみ有効なローカル変数である。一方で Scilab 5.4.1 でのメイン画面で使う変数は、グローバル変数である。従って、

`function~endfunction` 内で使う変数名には、メイン画面で使わない変数名を用いること。

さもなければ、メイン画面で使った変数の影響を受けてしまい、組み込み関数が正しく動かなくなるおそれがある。このレジュメではメイン画面では、

`A,B,C,a,b,c,i,j,P,Q`

など 1 文字ものをよく使うことが多いので、それらは自作関数内では使わないようにしている。

- (d) `if-elseif-end` ブロックと `if-end` ブロックが 1 つずつ存在している。いわゆる if 文であり、Scilab では `endif` ではなくて、`end` で終わる。
- (e) 繰り返し文が、`for 整数変数=初期値:終値 end` ブロックとなっている。多くのプログラムでは、繰り返しの命令文の 1 つとして、`for` を用いられている。通称 For ループという。Scilab も同様である。ただし Basic 系の言語では `For-Next` ブロックなのであるが、Scilab では `Next` は使わず、`if-end` ブロック同様に `for-end` ブロックとなっている。
- (f) `~=` は、Scilab では「≠」を意味する。

では次に、行列式の積と積の行列式の復習を行おう。

例題 32.3 次のプログラムをメイン画面に打ち込んで、 $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ が成り立つことを確認せよ。

```
A=[1 2 3;4 5 6;7 8 0];
B=A-1
det(A)
det(B)
det(A*B)
```

次に、Scilab のメインから、アプリケーション > Scinotes で Scinotes という Scilab 専用テキストエディタを開くこと。

例題 32.4 正方行列 \mathbf{A} と正整数 i, j が与えられたときに、 \mathbf{A} の (i, j) 余因子行列式を返すような関数 `cof(matA, irow, jcol)` を作る。次のコマンドを、Scinotes に打ち込んで (もしくは、PDF ファイルからコピーして)。適当な名前でも保存せよ。その後、Scinote での、「実行する > ファイルを実行」によって、そのファイルを実行させよ。

```
function cofij=cof(matA,irow,jcol)
    [nA,colA]=size(matA);
    if nA ~=colA then
        error('matrix must be squared');
    end
    one2n=[1:nA];
    tmpRindx=find(one2n~=irow);
    tmpCindx=find(one2n~=jcol);
    if length(tmpRindx)==nA then
        error('2nd argument' +string(irow) +' is out of range.')
```

```
elseif length(tmpCindx)==nA then
        error('3rd argument' +string(jcol) +' is out of range.')
```

```
end
    matA_ij=matA(tmpRindx,tmpCindx);
    cofij=(-1)^(irow+jcol)*det(matA_ij);
endfunction
```

これを読み込んだ後に、1 行目についての余因子展開を Scilab のメイン画面で行おう。なお `endfunction` の 1 つ前の、行列式を求めるところは、簡単化のために Scilab の `det` 関数を用いている。

例題 32.5 メイン画面で、以下の 4 行を打ち込み、行列式と 1 行目の余因子展開が同じになることを確かめよ。

```
A=[1 2 3;4 5 6;7 8 0]
det(A)
i=1
A(i,1)*cof(A,i,1)+A(i,2)*cof(A,i,2)+A(i,3)*cof(A,i,3)
```

1 行目だけでなく、2 行目・3 行目についても余因子展開したい。そこで以下のようにすると、簡単に得られる。

(i). 次に `i=2` と打ち込んで Enter した後、メイン画面で ↑ キーを 2 回押すと、

$$A(i,1)*cof(A,i,1)+A(i,2)*cof(A,i,2)+A(i,3)*cof(A,i,3)$$

が出てくるから、Enter せよ。すると、2 行目に関する余因子展開が得られる。

(ii). 次に `i=3` と打ち込んで Enter した後、メイン画面で ↑ キーを 2 回押すと、やはり

$$A(i,1)*cof(A,i,1)+A(i,2)*cof(A,i,2)+A(i,3)*cof(A,i,3)$$

が出てくるから、Enter せよ。すると、3 行目に関する余因子展開が得られる。

全て 27 となることを確認せよ。また、上矢印や下矢印キー (もしくは PageUP と PageDown キー) によって前や後のコマンドが出てくることは、コマンド入力系の PC プログラムではよくあることである。

では、上の `cof` 関数を使って、行列 `matA` を引数として、その余因子行列を返す `cofmat` という関数を作ろう。

例題 32.6 例題 32.3 での `cof` を書いたファイルの一番下に、次のコードをコピーせよ。そして上書き保存した後、その後、Scinote での、「実行する>ファイルを実行」によって、そのファイルを実行させよ。

```
function outMat=cofmat(matA)
    outMat=zeros(matA);
    for rowi=1:size(matA,1)
        for colj=1:size(matA,2)
            outMat(colj,rowi)=cof(matA,rowi,colj);
        end
    end
endfunction
```

さてこれで関数 `cofmat` が読み込まれたので、Scilab のメイン画面で、次の 4 行を打ち込むこと。

```
A=[1 2 3;4 5 6;7 8 0];
cofmat(A)
cofmat(A)*A
A*cofmat(A)
```

以上により、系 30.4 の

$$A\tilde{A} = \det(A)I$$

がイメージできるであろう。

ここでは、余因子行列 \tilde{A} の (j, i) 成分が、 A の (i, j) 余因子となっていることを、プログラムのコードを見て復習してほしい。

第 VI 部

1 次独立と行列の階数

1 次独立と階数についてはいろいろなことがいえるが、ここでは最小限のことを述べる。次元 (*dim*) など、より詳しいことを知りたい場合は、参考文献を参照するとよい。

33 ベクトルの 1 次独立と 1 次従属および行列のランク (階数)

3 つのベクトル

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix},$$

とすると、どの 1 つのベクトルをとっても、残りの 2 つの線形結合で表すことはできない。一方

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \mathbf{b}_3 = \begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix},$$

の 3 つのベクトルについては、ある 1 つのベクトルは他の線形結合で表される。 $\mathbf{b}_3 = \mathbf{b}_1 + \mathbf{b}_2$ は簡単にわかるであろう。変形して、 $\mathbf{b}_1 + \mathbf{b}_2 - \mathbf{b}_3 = \mathbf{0}$ がある。この 2 つの考え方を踏まえて、次のように定義する。

定義 33.1 (数ベクトルの 1 次独立と 1 次従属) n 個の m 項列ベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ と n 個のスカラー c_1, c_2, \dots, c_n について

$$c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \dots + c_n \mathbf{a}_n = \mathbf{0} \quad (33.1)$$

を満たす c_1, \dots, c_n が $c_1 = c_2 = \dots = c_n = 0$ のみであるとき、 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ は、1 次独立であるという。1 次独立でない場合は、1 次従属であるという。

また、(33.1) 式が $c_1 = c_2 = \dots = c_n = 0$ 以外の解をもつとき、「(33.1) 式は非自明な解をもつ」という。」

$$\mathbf{A} = (\mathbf{a}_1 \quad \dots \quad \mathbf{a}_n), \mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \quad (33.2)$$

とおけば (33.1) 式を変えて以下の命題が成り立つ。

命題 33.2 (行列の列ベクトルの一次独立と一次従属) $m \times n$ の行列 \mathbf{A} の各列と n 項列ベクトル \mathbf{c} の各成分を (33.2) 式のように表す。このとき、

$$\mathbf{A}\mathbf{c} = \mathbf{0}$$

を満たす \mathbf{c} が、 $\mathbf{c} = \mathbf{0}$ のみのとき、 $\mathbf{a}_1, \dots, \mathbf{a}_n$ は一次独立である。そうでないならば、 $\mathbf{a}_1, \dots, \mathbf{a}_n$ は一次従属である。 ─

逆に、(33.2) のように $m \times n$ の行列 \mathbf{A} が与えられたしよう。例えば、

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3), \quad \mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad (33.3)$$

としたときに, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ は明らかに一次従属である. しかし, \mathbf{a}_1 と \mathbf{a}_2 の 2 つのベクトルは一次独立である. このように, 行列を列ベクトルの集まりと考えたときに, その中で何個のものを選ぶと一次独立になるかを数えたものが, 行列のランクである.

定義 33.3 (行列のランクの定義) $m \times n$ の行列 \mathbf{A} の各列の列ベクトルを $\mathbf{a}_1, \dots, \mathbf{a}_n$ とする. この n 個から適当に選んだときに, それらが 1 次独立となるような組み合わせの最大個数を \mathbf{A} のランクと定義し, $\text{rank}(\mathbf{A})$ で表す. ┘

従って上記の \mathbf{A} で $\mathbf{a}_1, \dots, \mathbf{a}_n$ が一次独立ならば, $\text{rank}(\mathbf{A}) = n$ であり, \mathbf{A} の成分が全て 0 ならば, $\text{rank}(\mathbf{A}) = 0$ となる. (33.3) 式の \mathbf{A} では, $\text{rank}(\mathbf{A}) = 2$ となる.

この \mathbf{A} はランクと列数が一致していない. そこで \mathbf{A} ランクと列数が一致するもの, つまり \mathbf{A} の列が全て一次独立なものについて考えよう.

命題 33.4 $m \times n$ の行列 \mathbf{A} のランクが n とする. \mathbf{A} の前から正則行列を掛けても, 後ろから正則行列を掛けてもランクは変わらない. つまり, m 次正則行列 \mathbf{P} と n 次正則行列 \mathbf{Q} について,

$$\text{rank}(\mathbf{PA}) = \text{rank}(\mathbf{A}) \quad (33.4)$$

$$\text{rank}(\mathbf{AQ}) = \text{rank}(\mathbf{A}) \quad (33.5)$$
┘

証明 \mathbf{A} の各列ベクトルが一次独立なので, n 項列ベクトル \mathbf{c} について,

$$\mathbf{Ac} = \mathbf{0} \iff \mathbf{c} = \mathbf{0}$$

\mathbf{P}, \mathbf{Q} は逆行列を持つので,

$$\begin{cases} \mathbf{Ac} = \mathbf{0} \iff \mathbf{PAc} = \mathbf{0} \\ \mathbf{Ac} = \mathbf{0} \iff \mathbf{AQQ}^{-1}\mathbf{c} = \mathbf{0} \end{cases}$$

\mathbf{Q} については, $\mathbf{Q}^{-1}\mathbf{c} = \mathbf{0} \iff \mathbf{c} = \mathbf{0}$ であるから, $\mathbf{d} = \mathbf{Q}^{-1}\mathbf{c}$ とすると,

$$\begin{cases} \mathbf{PAc} = \mathbf{0} \iff \mathbf{c} = \mathbf{0} \\ \mathbf{AQd} = \mathbf{0} \iff \mathbf{d} = \mathbf{0} \end{cases}$$

これは \mathbf{PA}, \mathbf{AQ} のランクがともに n であることを表す. □

この \mathbf{A} のように, 列数とランクが同じものは取扱が楽なので, 英語などでは特に full column rank などという.

\mathbf{A} が full column rank ではなくとも上の命題が成り立つことや, 実際に行列が与えられたときに, そのランクを求める方法などは, 次章で説明する.

34 基本行列・基本変形と, 逆行列

n 次正方行列が与えられたとき, 行列式の定義で, 行列式が計算されることや, 定理 30.5 で, 余因子行列を用いて \mathbf{A}^{-1} を示した. しかし実際に計算する場合には, これらの計算方法ではなく, 別の方法を用いる. 中でも一般的な, 基本変形と呼ばれる方法を紹介する.

その前に何度も出てくるが、 \mathbf{A} が与えられたとき、

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}$$

を満たす \mathbf{B} を、 \mathbf{A} の逆行列と定義した。これを忘れないように。

定義 34.1 (基本行列タイプ 1: 行や列の入れ替え) n 次正方行列で、 $i \neq j$ として、以下 2 つが成り立つものを、このレジュメでは) タイプ 1 の n 次基本行列といい、 $\mathbf{T}_1(i, j; n)$ で表す。

- (i, i) 成分と (j, j) 成分が 0 で、それ以外の対角成分が 1.
- (i, j) 成分と (j, i) 成分が 1. それ以外の非対角成分は 0.

」

\mathbf{A} が m 行 n 列として、その左側から $\mathbf{T}_1(i, j; m)$ を掛けると、 \mathbf{A} の i 行目と j 行目を入れ替えたものができる。 \mathbf{A} の右側から、 $\mathbf{T}_1(i, j; n)$ を掛けると、 \mathbf{A} の i 列目と j 列目を入れ替えたものができる。

定義 34.2 (基本行列タイプ 2: ある行や列のみを b 倍) $c \neq 0$ とする。 n 次正方行列で、 (i, i) 成分が b であり、その他の対角成分が全て 1 であるような対角行列をこのレジュメでは) タイプ 2 の n 次基本行列といい、 $\mathbf{T}_2(i, b; n)$ で表す。

」

\mathbf{A} が m 行 n 列として、その左側から $\mathbf{T}_2(i, b; m)$ を掛けると、 \mathbf{A} の i 行目を b 倍したものができる。 \mathbf{A} の右側から、 $\mathbf{T}_2(i, b; n)$ を掛けると、 \mathbf{A} の i 列目を b 倍したものができる。また $b = 1$ の場合は単位行列であり、それもタイプ 2 の基本行列に含まれている。

定義 34.3 (基本行列タイプ 3: ある行か列を c 倍して別のものに加える) n 次正方行列で、 $i \neq j$ として、以下 2 つが成り立つものを、このレジュメでは) タイプ 3 の n 次基本行列といい、 $\mathbf{T}_3(i, j, c; n)$ で表す。

- 対角成分が全て 1 である。
- (i, j) 成分が c である。

」

\mathbf{A} が m 行 n 列として、その左側から $\mathbf{T}_3(i, j, c; m)$ を掛けると、 \mathbf{A} の j 行目を c 倍したものを、 i 行目に加えたものができる。 \mathbf{A} の右側から、 $\mathbf{T}_3(i, j, c; n)$ を掛けると、 \mathbf{A} の i 列目を c 倍したものを、 j 列目に加えたものができる。タイプ 3 の場合は、左から掛ける場合と右から掛ける場合で、 i と j の役割が逆になることに気を付けること。 $c = 0$ の場合は単位行列であり、これもタイプ 3 に含まれる。

基本行列は全て正則行列であり、簡単な計算から逆行列と行列式は以下のように表される。

命題 34.4 $b \neq 0$ とする。3 種類の基本行列について以下が成り立つ。

基本行列	逆行列	行列式	転置行列
$\mathbf{T}_1(i, j; n)$	$\mathbf{T}_1(i, j; n)$	-1	$\mathbf{T}_1(i, j; n)$
$\mathbf{T}_2(i, b; n)$	$\mathbf{T}_2(i, 1/b; n)$	c	$\mathbf{T}_2(i, b; n)$
$\mathbf{T}_3(i, j, c; n)$	$\mathbf{T}_3(i, j, -c; n)$	1	$\mathbf{T}_3(j, i, c; n)$

以上より、基本行列の逆行列や転置行列も基本行列である。

」

$T_3(i, j, c; n)$ は上三角行列か下三角行列であることに注意すればよい。それ以外は簡単に求められる。
さて基本行列では、その有限個の積が重要なので、その集合を定義しておく。

定義 34.5 n 次正方行列で、基本行列の有限個の積で表される集合を $\mathcal{T}(n)$ とする。その中でも、タイプ 1 の基本行列の有限個の積のみで表される集合を $\mathcal{T}_1(n)$ とし、タイプ 3 の基本行列の有限個の積のみで表される集合を $\mathcal{T}_3(n)$ とする。 ┘

命題 34.4 に ${}^t(BC) = {}^tC {}^tB$ や $(BC)^{-1} = C^{-1}B^{-1}$ と合わせて、前定義を使って表すと、以下が言える。

系 34.6 $\mathcal{T}(n)$ は、逆行列・転置行列・積について閉じている。つまり記号で書くと、

$$\begin{aligned} P \in \mathcal{T}(n) &\implies P^{-1}, {}^tP \in \mathcal{T}(n) \\ P, Q \in \mathcal{T}(n) &\implies PQ \in \mathcal{T}(n) \end{aligned}$$

なので、 P, Q ともに $\mathcal{T}(n)$ に含まれるならば、 $(PQ)^{-1}, {}^t(P, Q)$ も含まれる。

命題 34.7 m 行 n 列の任意の行列は、基本行列を左右から掛けることによって、最初の r 個 (r は 0 以上 $\min(m, n)$ 以下の整数) の対角成分が 1 で、残りの対角成分が 0 である対角行列にすることができる。 ┘

注意: PC プログラミングに応用するために、 i, j といった変数名のみと同様に、 $\boxed{\text{piv}}$ という変数を用いる。意味は pivot(軸) の略であり、 $\boxed{\text{piv}}$ で一つの正整数を表している。 (注意終)

証明 A を m 行 n 列の行列とする。以下の第 1 プロセスと第 2 プロセスの順に行う。なお A の (i, j) 成分を $A(i, j)$ で表す。

第 1 プロセス: $\boxed{\text{piv}} = 1, \dots, \min(m, n)$ について、1 から $\min(m, n)$ まで順に以下のプロセスを繰り返す。

37 章も参照のこと。

(i). $i \geq \boxed{\text{piv}}, j \geq \boxed{\text{piv}}$ の $A(i, j)$ の中で、絶対値が最も大きいものを探す。

(a) その最も大きいものが 0 ならば、任意の i, j ($i \geq \boxed{\text{piv}}, j \geq \boxed{\text{piv}}$) について、 $A(i, j) = 0$ なので、第 1 プロセスを終了して、 $r = \boxed{\text{piv}} - 1$ として第 2 プロセスに進む。そうでなければ、絶対値が最大になるような i, j を記憶しておく。

(b) $i > \boxed{\text{piv}}$ のときは A の左から $T_1(\boxed{\text{piv}}, i; m)$ を掛けて行を入れ替える。そしてさらに $j > \boxed{\text{piv}}$ のときは A の右から $T_1(\boxed{\text{piv}}, j; n)$ を掛けて列を入れ替える。その結果 $A(\boxed{\text{piv}}, \boxed{\text{piv}})$ が非ゼロとなる。

(ii). A の左から $T_2(1, 1/A(\boxed{\text{piv}}, \boxed{\text{piv}}); m)$ を掛ける。その結果 $A(\boxed{\text{piv}}, \boxed{\text{piv}}) = 1$ となる。

(iii). $i = \boxed{\text{piv}} + 1, \dots, m$ について、順に左から $T_3(i, \boxed{\text{piv}}, -A(i, \boxed{\text{piv}}); m)$ を掛ける。その結果、 $\boxed{\text{piv}}$ 列の $\boxed{\text{piv}} + 1$ 行以下は全て 0 になる。

第 2 プロセス: 第 1 プロセスの (i)(b) で終了して第 2 プロセスに入った場合は、 r がすでに決めてある。そうではなく (i)(b) で終了しないで第 2 プロセスに入ったときは、 $r = \min(m, n)$ とする。この r により、第 1 プロセスの終了した A の形は、

$$\begin{pmatrix} U_{r,r} & *_{r,(n-r)} \\ O_{(m-r),r} & O_{(m-r),(n-r)} \end{pmatrix} \quad (34.1)$$

となっている。行列の添字は各行列の行数と列数を並べたものであり、添字の数が 0 になったときは、その行列は存在しない形となる。また $U_{r,r}$ は対角成分が 1 の上三角行列になっている。* 部分は不明ではあるが何らかの r 行 $(n-r)$ 列の行列である。

基本行変形 (基本行列のタイプ 3) を行う。 $j = 2, \dots, r$ の順について、「 $i = 1, \dots, j-1$ について順に、 $T_3(i, j, -\mathbf{A}(i, j); m)$ を \mathbf{A} の左側から掛ける。」というプロセスを繰り返す。

第 3 プロセス: 第 2 プロセス終了後には、

$$\begin{pmatrix} \mathbf{I}_{r,r} & *_{r,(n-r)} \\ \mathbf{O}_{(m-r),r} & \mathbf{O}_{(m-r),(n-r)} \end{pmatrix} \quad (34.2)$$

となっている。あとは $*_{r,(n-r)}$ 部分を、基本列変形 (基本行列のタイプ 3) によって全て 0 にすれば良い。

$i = 1, \dots, r$ の順について、「 $j = k+1, \dots, n$ について順に、 $T_3(i, j, -\mathbf{A}(i, j); m)$ を \mathbf{A} の右側から掛ける。」というプロセスを繰り返す。

□

定義 34.8 前命題より、 $m \times n$ の行列 \mathbf{A} について、ある $\mathbf{P} \in \mathcal{T}(m)$ と、ある $\mathbf{Q} \in \mathcal{T}(n)$ と、ある r 次単位行列 $\mathbf{I}_{r,r}$ が存在して、

$$\mathbf{PAQ} = \begin{pmatrix} \mathbf{I}_{r,r} & \mathbf{O}_{r,(n-r)} \\ \mathbf{O}_{(m-r),r} & \mathbf{O}_{(m-r),(n-r)} \end{pmatrix} \quad (34.3)$$

と分割された行列を作ることができる。この右辺を \mathbf{A} の標準形といい、 $\mathbf{F}(r)$ で表す。 」

この定義で上の式のように \mathbf{P}, \mathbf{Q} の存在は示したが、標準形を得るための \mathbf{P}, \mathbf{Q} は唯一ではないということに注意をしなければならない。

例えば、 $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ としたとき、

$$\begin{cases} \mathbf{P}_1 = \mathbf{T}_2(2, 1/3, 2)\mathbf{T}_2(1, 1/2, 2) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/3 \end{pmatrix} \\ \mathbf{Q}_1 = \mathbf{I}_2 \end{cases}$$

とすると、 $\mathbf{P}_1\mathbf{A}\mathbf{Q}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ である。一方、

$$\begin{cases} \mathbf{P}_2 = \mathbf{T}_2(1, 1/2, 2) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix} \\ \mathbf{Q}_2 = \mathbf{T}_2(2, 1/3, 2) = \begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix} \end{cases}$$

としても、 $\mathbf{P}_2\mathbf{A}\mathbf{Q}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ が成り立つのである。

このように \mathbf{A} から標準形を得るための \mathbf{P}, \mathbf{Q} の組み合わせは唯一ではない。

次の問題は \mathbf{P}, \mathbf{Q} の組み合わせごとに、 \mathbf{A} の標準形 $\mathbf{F}(r)$ について、その r の数は一定なのかどうかという点である。これについては一定であることが以下で示される。

定理 34.9 $m \times n$ の行列 A について, ある $P \in \mathcal{T}(m)$ とある $Q_1 \in \mathcal{T}_1(n)$ が存在して, 標準形

$$PAQ = F(r)$$

が得られたとする. このときの整数 r は, P, Q にかかわらず一定である. □

証明 2種類の形によって, 異なる $F(s)$ が得られたと仮定して矛盾を導く. 今 m 次の正則行列 P_1, P_2 と n 次の正則行列 Q_1, Q_2 によって, 2種類の基本形

$$P_1AQ_1 = F(r) \tag{34.4}$$

$$P_2AQ_2 = F(s) \tag{34.5}$$

が得られたと仮定する. $r \leq s$ として一般形は失わないので, そう仮定する. (34.4) 式の左から P_1^{-1} を掛け, 右から Q_1^{-1} を掛けると, $A = P_1^{-1}F(r)Q_1^{-1}$ となり, これを (34.5) 式に代入することによって,

$$PF(r)Q = F(s) \tag{34.6}$$

が成り立つ. ただし $P = P_2P_1^{-1}, Q = Q_1^{-1}Q_2$ であり, P, Q はそれぞれ m 次と n 次の, 正則行列である.

次に P, Q をそれぞれ 2×2 の形に,

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}, \quad Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$$

と分割する. P_{11} は r 行 r 列, P_{22} は $m-r$ 行 $m-r$ 列である. Q_{11} は r 行 r 列, Q_{22} は $n-r$ 行 $n-r$ 列である. これらを (34.6) 式に代入すると,

$$\begin{aligned} \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} I_r & O_{r(n-r)} \\ O_{(m-r)r} & O_{(m-r)(n-r)} \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} &= F(s) \\ \iff \begin{pmatrix} P_{11}Q_{11} & P_{11}Q_{12} \\ P_{21}Q_{11} & P_{21}Q_{12} \end{pmatrix} &= \begin{pmatrix} I_s & O_{s(n-s)} \\ O_{(m-s)s} & O_{(m-s)(n-s)} \end{pmatrix} \end{aligned} \tag{34.7}$$

ここで仮に $s > r$ としよう. このときこの分割された行列で, 左辺の (1, 1), (1, 2), (2, 1) 成分と右辺の比較からそれぞれ以下のように書ける.

$$P_{11}Q_{11} = I_r \tag{34.8}$$

$$P_{11}Q_{12} = O_{r(n-r)} \tag{34.9}$$

$$P_{21}Q_{11} = O_{(m-r)r} \tag{34.10}$$

(34.8) 式と系 30.6 から, $P_{11}Q_{11}$ とともに正則である. (34.9) 式の左から P_{11}^{-1} を掛け, (34.10) 式の右から Q_{11}^{-1} を掛けて,

$$Q_{12} = O_{r(n-r)}, \quad P_{21} = O_{(m-r)r}$$

を得る. $P_{21}Q_{12} = O$ となるので, これらを (34.7) 式の左辺に代入すると, $F(r)$ となる. (34.7) 式の右辺は $F(s)$ なので, $r = s$ が得られる. □

いよいよ標準形と前章のランクとの関係を述べる. ランクの定義から, ある行列について, 列を入れ替えた行列のランクは変わらない. 従って, 下の補題の (i) は簡単に言える.

補題 34.10 $m \times n$ の行列 A について、以下が言える。

- (i). $\text{rank}(A) = k$ であることは、 $m \times k$ の行列でランクが k の B_1 と $k \times (n-k)$ の行列 G と、 $Q_1 \in \mathcal{T}_1(n)$ が存在して、

$$A = (B_1 \ B_1 G) Q_1 \quad (34.11)$$

と表されることと同値である。

- (ii). $\text{rank}(A) = k$ ならば、 $m \times k$ の行列でランクが k の B_1 と $k \times (n-k)$ のゼロ行列 O と、 $Q \in \mathcal{T}(n)$ が存在して、

$$A = (B_1 \ O) Q \quad (34.12)$$

と表される。

」

注意: (ii) の逆が成り立つことは、次の補題で証明する。

証明 (34.11) 式が成り立つならば、(34.12) 式が成り立つことを言えば良い。(34.11) 式の B_1, G について、 $(B_1 \ O)$ を考える。 $k+1$ 列から n 列までのそれぞれの列についての列変形を行う。 $(B_1 \ O)$ の右側から、 $i = 1, \dots, k$ と $j = n+1, \dots, n$ について基本行列タイプ 3 の $T_3(j, i, G(i, j-k); n)$ を掛けることによって、 $(B_1 \ O)$ は、 $(B_1 \ B_1 G)$ に変形することができる。それはある $Q_3 \in \mathcal{T}_3(n)$ があって、

$$(B_1 \ B_1 G) = (B_1 \ O) Q_3$$

が存在することに他ならない。両辺の右から Q_1 を掛けて $Q = Q_3 Q_1 \in \mathcal{T}$ から、命題が成り立つ。

□

34.10 の (i) は、次の補題の証明で用いる。(ii) は、もう少し後の命題の証明で用いる。

補題 34.11 $m \times n$ の行列 A に、基本変形の行列を前からかけても後ろからかけても、行列のランクは変わらない。つまり、 P, Q とともに基本行列で、それぞれ m 次 n 次るとき、

$$\text{rank}(PA) = \text{rank}(A) \quad (34.13)$$

$$\text{rank}(AQ) = \text{rank}(A) \quad (34.14)$$

」

証明 $\text{rank}(A) = 0$ のときは明らかであり、 $\text{rank}(A) = n$ のときは、前章の命題 33.4 から言える。よって以下では $1 \leq \text{rank}(A) \leq n-1$ とする。

[1] まず (34.14) 式を示す。

基本行列のタイプ 1 の $T_1(i, j; n)$ を A の右から掛けても成り立つことは、補題 34.10 で示した。そこで、

$$A = (A_1 \ A_1 C) \quad (34.15)$$

として、 A_1 は $m \times k$ の行列でランク k であり、 C は $k \times (n-k)$ であるものとする。以下の証明では全てこの A について証明する。

基本行列のタイプ2とタイプ3を \mathbf{A} の右から掛けることについては、それらをまとめて計算するために、 n 次正方行列 \mathbf{Q} が、 k 次正方行列 \mathbf{Q}_{11} が正則なものとして、

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{O} & \mathbf{Q}_{22} \end{pmatrix}$$

と、分割された上三角行列で表される場合を考える。この \mathbf{Q} について、 $\mathbf{A}\mathbf{Q}$ を計算すると、

$$\begin{aligned} \mathbf{A}\mathbf{Q} &= (\mathbf{A}_1\mathbf{Q}_{11} \quad \mathbf{A}_1(\mathbf{Q}_{12} + \mathbf{B}\mathbf{Q}_{22})) \\ &= (\mathbf{A}_1\mathbf{Q}_{11} \quad \mathbf{A}_1\mathbf{Q}_{11}[\mathbf{Q}_{11}^{-1}(\mathbf{Q}_{12} + \mathbf{C}\mathbf{Q}_{22})]) \end{aligned}$$

命題 33.4 より、 $\mathbf{A}_1\mathbf{Q}_{11}$ は列数 k でランクも k なので、補題 34.10(i) から、 $\text{rank}(\mathbf{A}\mathbf{Q}) = k = \text{rank}(\mathbf{A})$ が得られる。

基本行列のタイプ2の $\mathbf{T}_2(i, b; n)$ は全て上の \mathbf{Q} の形をしている。基本行列のタイプ3の $\mathbf{T}_3(i, j, c; n)$ の中で、 \mathbf{Q} が上の形をしていないものは、 $i > k, j \leq k$ のときのものである。実際に上の \mathbf{Q} の形に合わせれば、

$$\mathbf{T}(i, j, c; n) = \begin{pmatrix} \mathbf{I}_{11} & \mathbf{O} \\ \mathbf{Q}_{21} & \mathbf{I}_{22} \end{pmatrix}$$

となっている。 \mathbf{I}_{11} と \mathbf{I}_{22} は、それぞれ k 次と $n - k$ 次の単位行列であり、 \mathbf{Q}_{21} は $(n - k) \times k$ の行列で、 $(i - k, j)$ 成分のみ定数 c で、残りの成分は0のものである。

$i > k, j \leq k$ を満たす任意の i, j についての証明は、 $i = n, j = k$ について固定した場合と同様にできるので、

$$\mathbf{A}\mathbf{T}_3(n, k, c; n) = \mathbf{A}_{\mathbf{a}_k \leftarrow \mathbf{a}_k + c\mathbf{a}_n} = (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_{k-1} \quad \mathbf{a}_k + c\mathbf{a}_n \quad \mathbf{a}_{k+1} \quad \cdots \quad \mathbf{a}_n)$$

のランクが k であることを示せば良い。また \mathbf{A} の最後の列 \mathbf{a}_n は $\mathbf{A}_1\mathbf{C}$ の最後の列なので、ある c_1, \dots, c_k があって、

$$\mathbf{a}_n = \sum_{j=1}^{k-1} c_j \mathbf{a}_j + c_k \mathbf{a}_k$$

が成り立っている。

(1) $c_k \neq 0$ の場合は、 $\mathbf{A}\mathbf{T}_3(n, k, c; n)$ の1列目から $k - 1$ 列目と n 列目の、計 k 個の列が一次独立となる。実際、 \mathbf{A}_1 の k 列目を c_k 倍して、その後 $j = 1, \dots, k - 1$ について、 \mathbf{a}_j の c_j 倍を、 \mathbf{A}_1 の k 列目に加えることによって、

$$\begin{aligned} &\mathbf{A}_1\mathbf{T}_2(k, c_k; k) [\mathbf{T}_3(k, 1, c_1; k)\mathbf{T}_3(k, 2, c_2; k) \cdots \mathbf{T}_3(k, k - 1, c_{k-1}; k)] \\ &= (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_{k-1} \quad c_k \mathbf{a}_k + \sum_{j=1}^{k-1} c_j \mathbf{a}_j) \\ &= (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_{k-1} \quad \mathbf{a}_n) \end{aligned}$$

命題 33.4 より、左辺のランクは \mathbf{A}_1 のランク k である。よって右辺の k 個の各列は一次独立である。

(2) $c_k = 0$ の場合は、 $\mathbf{A}\mathbf{T}_3(n, k, c; n)$ の1列目から k 列目までの、計 k 個の列が一次独立となる。実際 $\mathbf{a}_1, \dots, \mathbf{a}_{k-1}, \mathbf{a}_n = \sum_{j=1}^{k-1} c_j \mathbf{a}_j$ なので、 $j = 1, \dots, k - 1$ について、 \mathbf{a}_j の cc_j 倍を、 \mathbf{A}_1 の k 列目に加えることによって、

$$\begin{aligned} &\mathbf{A}_1 [\mathbf{T}_3(k, 1, cc_1; k)\mathbf{T}_3(k, 2, cc_2; k) \cdots \mathbf{T}_3(k, k - 1, cc_{k-1}; k)] \\ &= (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_{k-1} \quad \mathbf{a}_k + \sum_{j=1}^{k-1} cc_j \mathbf{a}_j) \\ &= (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_{k-1} \quad \mathbf{a}_k + c\mathbf{a}_n) \end{aligned}$$

(1) と同様にして、この右辺の各列が一次独立であることが言える。

\mathbf{A} の定義と (1)(2) および、 \mathbf{A} と $\mathbf{AT}_3(n, k, c; n)$ の違いは k 列目のみであることから、 $\mathbf{AT}_3(n, k, c; n)$ のランクは k である。

[2](34.13) 式を示す。

$$\begin{aligned} \mathbf{PA} &= \mathbf{P} \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1\mathbf{C} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{PA}_1 & \mathbf{PA}_1\mathbf{C} \end{pmatrix} \end{aligned}$$

\mathbf{P} が正則なので、命題 33.4 より、 $m \times k$ の行列 \mathbf{PA}_1 はランク k である。よって前補題より、右辺のランクは k である。つまり $\text{rank}(\mathbf{PA}) = \text{rank}(\mathbf{A})$ 。

□

定理 34.12 (行列のランクと標準形の r の一致) $m \times n$ の行列 \mathbf{A} の標準形が $\mathbf{F}(r)$ とする。このとき $\text{rank}(\mathbf{A}) = r$ ┘

証明 仮定より、ある $\mathbf{P} \in \mathcal{T}(m)$ と $\mathbf{Q} \in \mathcal{T}(n)$ があって、

$$\mathbf{PAQ} = \mathbf{F}(r)$$

が成り立つ。前補題より、 $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{PAQ}) = \text{rank}(\mathbf{F}(r))$ 。右辺は明らかに r である。

□

命題 34.7 で標準形の r は $r \leq \min(m, n)$ であった。よって、 $m \times n$ の行列 \mathbf{A} のランクについては、 $0 \leq \text{rank}(\mathbf{A}) \leq \min(m, n)$ が成り立つ。

ここまでは $m \times n$ の行列について述べてきた。次に正方行列の場合を考えて、ランクや標準形との関係を示したものが以下である。

定理 34.13 n 次正方行列 \mathbf{A} について、以下は同値である。

- (i). \mathbf{A} が正則である
- (ii). $\text{rank}(\mathbf{A}) = n$
- (iii). $\mathbf{A} \in \mathcal{T}(n)$.

┘

証明 \mathbf{A} のランクを r とする。つまり以下が言える。

$$\begin{aligned} \exists \mathbf{P} \in \mathcal{T}(n), \exists \mathbf{Q} \in \mathcal{T}(n); \\ \mathbf{PAQ} = \mathbf{F}(r) \end{aligned} \tag{34.16}$$

(i) \implies (ii) の証明: (34.16) 式の両辺の行列式をとると、定理 28.1 より、

$$\det(\mathbf{P}) \det(\mathbf{A}) \det(\mathbf{Q}) = \det(\mathbf{F}(r))$$

が言える。仮定より左辺は非ゼロなので、 $\det(\mathbf{F}(r)) \neq 0$ 。 $\mathbf{F}(r)$ は対角成分が 1 か 0 の対角行列なので、右辺については $r < n$ のとき 0 であり、 $r = n$ のとき 1 である。よって (ii) が成り立つ。

(ii) \implies (iii) の証明:仮定より, (34.16) 式の右辺 $F(r)$ は n 次単位行列 I となる. (34.16) 式の左から P^{-1} を掛けて, 右から Q^{-1} を掛けると, $A = P^{-1}Q^{-1}$. $\mathcal{T}(n)$ は積と逆行列について閉じているので, $P^{-1}Q^{-1} \in \mathcal{T}(n)$. よって (iii) が成り立つ.

(iii) \implies (i) は, 明らか.

□

(i) と (iii) の $A \in \mathcal{T}(n)$ の同値性は一見かなりの驚きである. 正則行列は, 基本行列の有限個の積なのである. $\mathcal{T}(n)$ は逆行列についても閉じているから,

さらにこれから, 補題 34.11 が拡張できる.

定理 34.14 $m \times n$ の行列 A に, 正則行列を前からかけても後ろからかけても, 行列のランクは変わらない. つまり, P, Q ともに正則行列で, それぞれ m 次, n 次するとき,

$$\begin{aligned} \text{rank}(PA) &= \text{rank}(A) \\ \text{rank}(AQ) &= \text{rank}(A) \end{aligned}$$

」

定理 34.13 と $\mathcal{T}(n)$ の性質から, A が正則ならば, A^{-1} も有限個の基本行列の積で表される. さらに, 正則の定義 $A^{-1}A = AA^{-1} = I$ から以下が言える.

系 34.15 n 次正方行列が正則ならば, 基本行列を左側から掛けることのみで, 単位行列 I_n に変形することができる. 同様に, 基本行列を右から掛けることのみで, 単位行列に変形することができる.

」

他にランクの性質としては, 以下があげられる.

命題 34.16 $m \times n$ の行列 A と, $n \times k$ の行列 B について,

- (i). $\text{rank}({}^tA) = \text{rank}(A)$
- (ii). $\text{rank}(AB) \leq \text{rank}(A) \cdot \text{rank}(B)$

」

証明 (i) では, $P \in \mathcal{T}(m), Q \in \mathcal{T}(n)$ で $PAQ = F_{m,n}(r)$ と標準形を作ったとしよう. $F_{m,n}$ と添字をつけたのは, これが m 行 n 列であることを明示している. 両辺の転置行列をとると,

${}^tQ {}^tA {}^tP = F_{n,m}(r)$ である. 右辺は n 行 m 列となるが, r の値は変化しない. 系 34.6 より, ${}^tQ \in \mathcal{T}(n), {}^tP \in \mathcal{T}(m)$. 従って標準形の定義より, tA の標準形が $F_{n,m}(r)$ である. よって, $\text{rank}({}^tA) = r$

(ii) でまず, $a = \text{rank}(A), b = \text{rank}(B)$ とする. $P_1, P_2 \in \mathcal{T}(m)$ と, $Q_1, Q_2 \in \mathcal{T}(n)$ によって, A, B がそれぞれ

$$\begin{cases} P_1 A Q_1 &= F(a) \\ P_2 B Q_2 &= F(b) \end{cases}$$

で基本変形されたとする. 逆行列を掛けて変形して,

$$\begin{aligned} \text{rank}(AB) &= \text{rank}(P_1^{-1}F(a)Q_1^{-1}P_2^{-1}F(b)Q_2^{-1}) \\ &(\text{定理 34.14 より}) = \text{rank}(F(a)Q_1^{-1}P_2^{-1}F(b)) \end{aligned}$$

標準形から,

$$\mathbf{F}(a) = \begin{pmatrix} \mathbf{I}_a & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}, \mathbf{F}(b) = \begin{pmatrix} \mathbf{I}_b & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$$

と \mathbf{O} の添字は省略して書き下す. $k = \min(a, b)$ として, さらに計算して,

$$\mathbf{F}(a)\mathbf{Q}_1^{-1}\mathbf{P}_2^{-1}\mathbf{F}(b) = \begin{pmatrix} *_{k,k} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$$

が得られる. ただし $*_{k,k}$ は k 次正方行列であり, $\mathbf{Q}_1^{-1}\mathbf{P}_2^{-1}$ の 1 行目から $\min(a, b)$ 行目までの 1 列目から $\min(a, b)$ 列目を表す. 従って,

$$\begin{aligned} \text{rank}(\mathbf{AB}) &= \text{rank}(*_{k,k}) \\ &\leq k = \min(a, b) \end{aligned}$$

□

ここまでは成分が複素数の行列でも成り立つことである. 実数の場合だと以下が言える.

命題 34.17 $m \times n$ の実行列 \mathbf{A} について, \mathbf{A} のランクと ${}^t\mathbf{AA}$ のランクは一致する. 」

注意: よって \mathbf{A} が full column rank ならば, ${}^t\mathbf{AA}$ は非特異である.

(注意終)

証明 [1] $\text{rank}(\mathbf{A}) = n \iff \text{rank}({}^t\mathbf{AA}) = n$ を証明する. これは対偶を取って

$$\text{rank}(\mathbf{A}) < n \iff \text{rank}({}^t\mathbf{AA}) < n$$

を示すほうが簡単なので, これを示す.

(1) まず $\text{rank}(\mathbf{A}) < n \implies \text{rank}({}^t\mathbf{AA}) < n$ を示す. 仮定より, ある n 項列ベクトル $\mathbf{x} \neq \mathbf{0}_n$ が存在して, $\mathbf{Ax} = \mathbf{0}_m$. これより両辺に ${}^t\mathbf{A}$ を左から掛けた

$${}^t\mathbf{AAx} = \mathbf{0}_n$$

を満たす $\mathbf{x} \neq \mathbf{0}$ が存在するので, ${}^t\mathbf{AA}$ は非正則である. よって $\text{rank}({}^t\mathbf{AA}) < n$ である.

(2) 次に $\text{rank}(\mathbf{A}) < n \iff \text{rank}({}^t\mathbf{AA}) < n$ を示す. 仮定より, ある n 項列ベクトル $\mathbf{x} \neq \mathbf{0}_n$ が存在して,

$${}^t\mathbf{AAx} = \mathbf{0}_n$$

となる. これより ${}^t\mathbf{x}{}^t\mathbf{AAx} = 0$ が成り立つが, 変形して ${}^t(\mathbf{Ax})\mathbf{Ax} = 0$ であり, この左辺は, m 項列ベクトル \mathbf{Ax} の各項の 2 乗和である. 仮定より各成分は実数なので, 2 乗和が 0 ということは, 各成分が全て 0 であることと同値である. 以上より, ${}^t\mathbf{x}{}^t\mathbf{AAx} = 0$ は $\mathbf{Ax} = \mathbf{0}_m$ と同値である.

すなわち, ある $\mathbf{x} \neq \mathbf{0}_n$ が存在して, $\mathbf{Ax} = \mathbf{0}_m$ が成り立つので, \mathbf{A} の各列は一次従属である. 従って $\text{rank}(\mathbf{A}) < n$.

[2] $\text{rank}(\mathbf{A}) < n$ のときに, $\text{rank}(\mathbf{A}) = \text{rank}({}^t\mathbf{AA})$ を示す. $\text{rank}(\mathbf{A}) = r$ とする. 補題 34.10 の (ii) より, $m \times r$ でランク r の行列 \mathbf{A}_1 と $r \times (n - r)$ のあるゼロ行列と, ある $\mathbf{Q} \in \mathcal{T}(n)$ によって,

$$\mathbf{A} = (\mathbf{A}_1 \quad \mathbf{O})\mathbf{Q}$$

とおける.

$$\begin{aligned} \text{rank}({}^t\mathbf{A}\mathbf{A}) &= \text{rank}({}^t\mathbf{Q} {}^t[(\mathbf{A}_1 \ \mathbf{O})] (\mathbf{A}_1 \ \mathbf{O}) \mathbf{Q}) \\ \text{補題 34.11 より} &= \text{rank}\left[\begin{pmatrix} {}^t\mathbf{A}_1 \\ \mathbf{O} \end{pmatrix} (\mathbf{A}_1 \ \mathbf{O})\right] \\ &= \text{rank}\begin{pmatrix} {}^t\mathbf{A}_1\mathbf{A}_1 & \mathbf{O}_{r,n-r} \\ \mathbf{O}_{n-r,r} & \mathbf{O}_{n-r,n-r} \end{pmatrix} \end{aligned}$$

[1] より ${}^t\mathbf{A}_1\mathbf{A}_1$ のランクは r なので, 右辺は r である.

□

この命題の周辺の話は, 2 次形式で詳しく議論する.

この章の最後に, 標準形の作り方として, 命題 34.7 を紹介した理由を述べよう. 数学の教科書では, これよりももう少し簡単に表現されている. にもかかわらずこの方法で紹介したのは, PC プログラムのためである. 標準形ではなく, 元の行列のランクだけを知りたいときや, n 次正方行列での行列式を知りたいときなどは, 第 1 プロセスだけで事が済むからである. それに合わせて標準形について述べたものが以下である.

命題 34.18 $m \times n$ の行列 \mathbf{A} について, ある $\mathbf{P} \in \mathcal{T}(m)$ と, ある $\mathbf{Q}_1 \in \mathcal{T}_1(n)$ と, 対角成分が全て 1 である r 次正方上三角行列 \mathbf{U}_{rr} によって

$$\mathbf{P}\mathbf{A}\mathbf{Q}_1 = \begin{pmatrix} \mathbf{U}_{r,r} & \mathbf{O}_{r,(n-r)} \\ \mathbf{O}_{(m-r),r} & \mathbf{O}_{(m-r),(n-r)} \end{pmatrix}$$

(ただし添字は行列のサイズ) と分割した行列を作ることができる. このとき, $\text{rank}(\mathbf{A}) = r$ である. ◻

証明 この後, $\mathbf{P}\mathbf{A}\mathbf{Q}_1$ に, 命題 34.7 の第 2 プロセスを施せば標準形 $\mathbf{F}(r)$ が得られることと, 補題 34.11 から命題が成り立つ.

□

35 n 元 1 次連立方程式を行変形のみで解く・逆行列の PC 計算

前章の系 34.15 から, a_{ij} と b_j を定数として, $n \times n$ の連立方程式

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \quad (35.1)$$

について, 解が存在するならば, それは行列の行変形のみで解くことができるということが言える. まず拡大係数行列を定義しよう.

定義 35.1 定数 $a_{ij} (1 \leq i, j \leq n)$ と定数 $b_i (i = 1, 2, \dots, n)$ についての n 元 1 次連立方程式 (35.1) を解くとき, $n \times (n+1)$ の行列

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{pmatrix}$$

を拡大係数行列という。 ┘

もとの (35.1) 式を行列で表せば,

$$Ax = b$$

とできるので, 拡大係数行列を C とすれば,

$$C = (A \ b)$$

である. $\det(A) \neq 0$ のときは C の行についての以下のような基本変形のプロセスで, 方程式の解を求めることができる. 命題 34.7 の証明のように, $\boxed{\text{piv}}$ を使って表現すると, 以下である.

- (i). $\boxed{\text{piv}} = 1$ とする.
- (ii). 任意の $i \geq \boxed{\text{piv}}$ について, $C(i, \boxed{\text{piv}})$ の絶対値が最も大きくなるものの行数を i^* とする. $i^* > \boxed{\text{piv}}$ ならば, C の i^* 行目と $\boxed{\text{piv}}$ 行目を入れ替える.
- (iii). C の $\boxed{\text{piv}}$ 行目全体を $C(\boxed{\text{piv}}, \boxed{\text{piv}})$ で割る. $C(\boxed{\text{piv}}, \boxed{\text{piv}}) = 1$ となる.
- (iv). $i = 1, \dots, \boxed{\text{piv}} - 1$ について, C の i 行目から, C の $\boxed{\text{piv}}$ 行目の $C(i, \boxed{\text{piv}})$ 倍を引く.
- (v). $i = \boxed{\text{piv}} + 1, \dots, n$ について, C の i 行目から, C の $\boxed{\text{piv}}$ 行目の $C(i, \boxed{\text{piv}})$ 倍を引く.
- (vi). $\boxed{\text{piv}} < n$ のときは, $\boxed{\text{piv}}$ に 1 を加えて (ii) へ. $\boxed{\text{piv}} = n$ のときは, プロセス終了.

プロセスが終了すると, C の最初の n 行 n 列は, I_n となり, 最後の列が方程式の解となる.

もし $\det(A) = 0$ のときは, このプロセス (ii) で, ある $\boxed{\text{piv}}$ があって, 任意の $i = \boxed{\text{piv}}, \dots, n$ について, $a_{ik} = 0$ となってしまい, プロセスが途中で止まってしまうのである.

例題 35.2 3元連立方程式

$$\begin{cases} x + 2y + 3z = 14 \\ 5x + 6y + 7z = 38 \\ 3x + 5y + 2z = 19 \end{cases}$$

を, 上のプロセスによって解く Scilab プログラムを書くこと。 ┘

(解答) 例としては以下のプログラムを参考にするとよい. 最後に逆行列を求める関数 `inv` を用いて, C の最終列と比較している. `dai=` の行と `disp` の行は, C の変形の途中経過を示すために, 入れたものである.

```
clear
A=[1 2 3;5 6 7;3 5 2];
b=[14 38 19]';
n=size(A,1);
C=[A b];
disp(C,'initial C=');
for k=1:n
    disp('k='+string(k));
    mxabs=max(abs(C(k:n,k)));
    indexes=find(abs(C(k:n,k))==mxabs);
    istar=indexes(1)+k-1;
```

```

if istar>k
    C([k istar],:)=C([istar k],:);
    dai='row '+string(istar)+'& row ' + string(k) + ' was exchanged';
    disp(C,dai);
end

C(k,:)=C(k,+)/C(k,k);
dai='row ' +string(k)+ ' was divided by C(' +string(k)+','+string(k)+')';
disp(C,dai);

for i=1:k-1
    C(i,:)=C(i,)-C(i,k)*C(k,);
end
for i=k+1:n
    C(i,:)=C(i,)-C(i,k)*C(k,);
end
dai='column '+string(k)+' was swept out.' ;
disp(C,dai);
end
disp(inv(A)*b,'check with inv(A)*b');

```

逆行列がほしい場合は、拡大係数行列の拡張として、 n 行 $2n$ 列の

$$C = \begin{pmatrix} A & I_n \end{pmatrix}$$

をとって、上のプロセスを同様に行えばよい。すると、 A が正則ならば、 C の $n+1$ 列目以降が、 A の逆行列となる。

練習 35.3 上の Scilab プログラムの 5 行目を

$$C=[A \text{ eye}(n,n)];$$

に変える。そして最後の行を 2 か所の

$$*b$$

を削除して、別名で名前を付けて保存をしてから、プログラムを実行せよ。最後の C の $n+1$ 列目以降と、 $\text{inv}(C)$ が一致していることを確かめよ。

A が正則であるということがわかっている場合のプログラムが上のものであり、正則かどうか分かっていない場合には、エラー処理などを付け加えなければならない。しかし、このプログラムが理解できれば、行列の行変形にかなり強くなることであろう。

練習 35.4 前の例題のプログラムを Scinote で開き、

- (i). A を別の正方行列に変えて、上書き保存をする。

(ii). プログラムを実行せよ.

36 トレースの定義

トレースは正方行列に対して定義されるもので,

定義 36.1 n 次正方行列 \mathbf{A} の対角成分の和をトレースといい $tr(\mathbf{A})$ であらわす. 成分で表示すると,

$$tr(\mathbf{A}) = \sum_{i=1}^n a_{ii}$$

命題 36.2 トレースについて, 以下が成り立つ.

- (i). n 次正方行列 \mathbf{A}, \mathbf{B} について, $tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$
(ii). n 行 m 列の行列 \mathbf{A} と m 行 n 列の行列 \mathbf{B} について, $tr(\mathbf{AB}) = tr(\mathbf{BA})$

証明 (i) の和のトレースは簡単である. (ii) の積のトレースについては, n 項行ベクトルと n 項列ベクトルについて成り立つことを示せば, 行列の積の定義より十分である. 行ベクトル \mathbf{a} と列ベクトル \mathbf{b} を

$$\mathbf{a} = (a_1 \quad \cdots \quad a_n), \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

とすると, 明らかに $\mathbf{ab} = \sum_{i=1}^n a_i b_i$ が成り立ち, スカラーなので, トレースも当然 $tr(\mathbf{ab}) = \sum_{i=1}^n a_i b_i$ である. 一方,

$$\begin{aligned} tr(\mathbf{ba}) &= tr \left(\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} (a_1 \quad \cdots \quad a_n) \right) \\ &= tr \begin{pmatrix} b_1 a_1 & \cdots & b_1 a_n \\ \vdots & \ddots & \vdots \\ b_n a_1 & \cdots & b_n a_n \end{pmatrix} \\ &= \sum_{i=1}^n b_i a_i \\ &= tr(\mathbf{ab}) \end{aligned}$$

□

ただの対角成分の足し算に何の意味があるか疑問に思うかもしれない. 一つは, 正方行列 \mathbf{A} と固有値の関係でトレースが関係してくる. 2 次正方行列 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ が固有値 λ_1, λ_2 を持つとき, (10.9) 式で解と係数の関係から

$$a + d = \lambda_1 + \lambda_2$$

が成り立つことを示したが、これはよく見ると、

$$\text{tr}(\mathbf{A}) = \lambda_1 + \lambda_2$$

となっている。

もう一つは、このレジュメの最後の方のべき等対称行列において、トレースを使った 1 例があるので、それを見るとよいであろう。

37 第 VI 部の Scilab 実習

ここでは VI 部の復習として Scilab の実習をまとめて行う。メイン画面で行うのみでなく、15 部の 51 ページの (iii) での、

.sce ファイルで保存して、実行するバッチ形式。

を主に行う。

37.1 find 関数と sum 関数についての注意

まずは行列志向プログラミングとして重要な、find 関数について説明する。他の C 言語・Visual Basic その他と違って、Scilab のこの関数を使いこなすことによって、行列計算が簡単になる。

find(条件式) という形で、条件式にあてはまるインデックスを返してくれる。条件式が、行ベクトルや列ベクトルの場合だと簡単である。

例題 37.1 以下のコードをメイン画面に打ち込め。

```
A=[2:2:10]
B=[4 8 9]'
find(A==6)
find(B>6)
find(B<0)
```

3 行目は 6, 4 行目は $\begin{pmatrix} 2 & 3 \end{pmatrix}$, そして 5 行目は答えがないので, [] が返ってくるであろう。

ところが条件式の部分が行列の場合は、行列をあたかも 1 つの列ベクトルとして、インデックスがかえってくる。

例題 37.2 以下のコードをメイン画面に打ち込め。

```
clear
A=[2:4;11:13]
find(A>11)
```

find(A>11) の結果としては、 $\begin{pmatrix} 6 & 8 \end{pmatrix}$ がかえってくる。これは A について、その 4 つの列ベクトルを左

から順に並べていった

$$A = \begin{pmatrix} 2 \\ 11 \\ 3 \\ 12 \\ 4 \\ 13 \end{pmatrix}$$

ものとして、そのあと find で、成分が 11 より大きいインデックスを返しているのである。find 関数は便利であり、うまく使えば for ループを使うよりも相当早いのだが、行列に find 関数を用いる場合には、このように注意しておかなければならない。

またベクトルや行列を使った条件式は、T(True を意味する) と F(False を意味する) のどちらかを返すのだが、数字でいうと 1 と 0 となる。よって、条件を満たす個数が、条件式と sum 関数を重ねたもので得られる。

例題 37.3 以下のコードをメイン画面に打ち込め。

```
clear
A=[1:3;4:6]
B=[2 2 3 4]
sum(A>=5)
sum(B<2)
```

」

37.2 基本変形・ランク・逆行列の復習:全て表示する関数

まず Scilab のメインから、アプリケーション> Scinotes で Scinotes を開くこと。

例題 37.4 Scilab 上で、matA という行列があったときに、その基本変形を行うことによって、ランク、各列の中で一次独立なものの列番号の組み合わせの一例、行列式 (matA が正則な場合)、逆行列 (matA が正則な場合) の 4 つを全て返すプログラムが、以下の Ftrans である。

- (i). 次のコマンドを、Scinotes に打ち込んで (もしくは、PDF ファイルからコピーして)、適当な名前前で保存せよ。
- (ii). その後、Scinote での、「実行する>ファイルを実行」によって、そのファイルを実行させよ。
- (iii). //以降の部分は Scilab 6.0 ではコメント文であっても日本語があると、エラーがでることがあるので、コメントも全て英語にしてある。

```
function [r,Qseq,detA,invA]=Ftrans(matA,varargin)
//r is rank of inA. Qseq is column numbers of independent column vector of in A
//deA and B is determinant and inverse of inA if inA is regular
//process are shown if argument 2 exists. For example, main3(matA,1)

FlagPrintA= length(varargin) >=1
```

```
A=matA;myeps = 10 ^ (-14);m = size(A, 1);n = size(A, 2);minMN = min(m,n)
FlagFullRowRank=1
Qseq=[1:n]
if m==n then
    detA=1
    B=eye(n,n) // if PAQ=I then PIQ=A^{-1}
else
    detA=[]
    B=[]
end
if FlagPrintA then; disp(A,"A="); end

//1st. process:
for piv=1:minMN
    //Max_{i>=piv,j>=piv} abs(A(i,j))
    //mxabsval is optimun value. istar & jstar is optimal argument.
    mxabsval=max(abs(A(piv:m,piv:n)))
    vecindex=min(find(abs(A(piv:m,piv:n))==mxabsval))
    pivMatrow=m-piv+1
    quot=floor((vecindex-1)/pivMatrow)
    jstar=quot+piv
    istar=vecindex-1-quot*pivMatrow+piv

    if mxabsval < myeps then
        r = piv - 1
        FlagFullRowRank=0
        break
    end

    if istar > piv then
        A([piv istar], :) = A([istar piv], :)
        detA=-detA
        B([piv istar], :) = B([istar piv], :)
        if FlagPrintA then; disp(A,"type1:row"+string(piv)+" <-> row "+string(istar)); end
    end

    if jstar > piv then
        A(:, [piv jstar]) = A(:, [jstar piv])
        detA=-detA
        B(:, [piv jstar]) = B(:, [jstar piv])
    end
end
```



```

        Qseq([piv jstar]) = Qseq([jstar piv])
        if FlagPrintA then; disp(A,"type1:col"+string(piv)+" <-> col"+string(jstar)); end
    end

    divVal=A(piv,piv)
    A(piv,:) = A(piv,+)/divVal
    B(piv,:) = B(piv,+)/divVal
    detA=divVal*detA
    if FlagPrintA then; disp(A,"type2:row"+string(piv) +" divided by " +string(divVal)); end

    for i= piv + 1: m
        multVal=-A(i,piv)
        A(i,:)=A(i,)+multVal*A(piv,:)
        B(i,:)=B(i,)+multVal*B(piv,:)
        if FlagPrintA then; disp(A,"type3:row"+ string(i) ..
            + "=row"+string(i)+"+ row"+ string(piv) +"*( " +string(multVal)+")" ); end
    end
end

if FlagFullRowRank==1 then
    r = min(M,N)
end

//2nd process(1): if A is regular, U_rr to I_rr by type 3 row transformation

if m==n & r==m then
    for j = 2:r
        for i=1:j-1
            multVal=-A(i,j)
            A(i,:)=A(i,)+multVal*A(j,:)
            B(i,:)=B(i,)+multVal*B(j,:)
            if FlagPrintA then; disp(A,"type3:row"+ string(i) ..
                + "=row"+string(i)+"+ row"+ string(piv) +"*( " +string(multVal)+")" ); end
        end
    end
end

//forall i and j, invA(i,j)=B(Qseq(i),Qseq(j))
invA=zeros(n,n)
invA(:,Qseq)=B
invA(Qseq,:)=invA
//Instead this codes, below 6 codes works as well.

```

```
//invA=zeros(n,n)
//for i=1:n
//    for j=1:n
//        invA(Qseq(i),Qseq(j))=B(i,j)
//    end
//end

elseif m==n then
    detA=0
    invA=[]
else
    invA=[]
end
Qseq=Qseq(1:r) // shows only independent column numbers

//We solve rank,seqQ,det,invA, so far.
//below codes shows fundamental transformation makes standard form F(r)

//2nd process(2): if A is not regular, U_rr to I_rr by type 3 rows transformation
if m<>n | r<m then
    for j = 2:r
        for i=1:j-1
            multVal=-A(i,j)
            A(i,:)=A(i,:)+multVal*A(j,:)
            if FlagPrintA then; disp(A,"type3:row"+ string(i) ..
                + "row"+string(i)+" row"+ string(piv) +"*( " +string(multVal)+")" ); end
        end
    end
end

//3rd process: this makes F(r) by type 3 columns transformation
for i=1:r
    for j = r + 1: n
        A(:,j)=A(:,j)-A(i, j)*A(:,i)
        if FlagPrintA then; disp(A,"type3:col"+ string(j) ..
            + "col"+string(j)+" row"+ string(i) +"*( " +string(multVal)+")" ); end
    end
end
endfunction
```

命題 34.7 のプロセスと、全く同じプロセスで計算を行っている。

ただしメイン画面では、上に入っているものと同じ変数名を使わないように。では

$$C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, D = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 3 & 4 \end{pmatrix}$$

の 2 つの行列について、Ftrans を実行させてみよう。メイン画面で、以下の順に打ち込むと良い。出てきた結果を変数とともに吟味すると良いだろう。

```
C=[1 2;3 4]
D=[[0;0] C]
[rankC,indpC,detC,invC]=Ftrans(C)
[rankD,indpD,detD,invD]=Ftrans(D)
```

Scilab などは、行列の変更が簡単にできる点が、利点である。というのも

```
A([piv istar], :) = A([istar piv], :)
```

の部分は、A の piv 行と istar 行に、A の istar 行と piv 行をそれぞれ代入するという意味である。このプログラム 1 行だけで、行の変更ができるのである。

さてこのプログラムでは 1st process: が終了した時点で、命題 34.18 の状態となっている。この後、ランク・逆行列・行列式を求めるにあたって、 Q_1 は全く変更しない。 Q_1 は行や列変換の行列の積であるから、行列になっていなくても、A のどの列がもともとどの列から変換されたかを記録しておけばいいので、1 行の行ベクトルの形にして記録しておけばよい。それがプログラム内の行ベクトルの Qseq なのである。

上の D=[[0;0] C] についてプログラムを動かした結果 indpD=[3,2] と出てくることであろう。これは、プログラム内で、seqQ=seqQ(1:r) によって切り取った結果であり、1st process: が終了した時点での seqQ は $\begin{pmatrix} 3 & 2 & 1 \end{pmatrix}$ となっている。この意味は、列交換を繰り返した結果、 $i = 1, 2, 3$ について、この時点の A の i 列目は、元の seqQ(i) 列目なのである。実際コードで

```
[rankD,indpD,detD,invD]=Ftrans(D,1)
```

と第 2 引数に適当なものを入れると、変形の途中が表示されるので、それが確認されるであろう。

次に、コードの中で逆行列を求めるときに、

```
invA=zeros(n,n)
invA(:,Qseq)=B
invA(Qseq,:)=invA
```

となっているので、これを説明しよう。A は正則で、ある $P \in T(n)$ とある $Q \in T_1(n)$ があって、このコードの前段階で、

$$PAQ = I \quad (37.1)$$

$$PIQ = B \quad (37.2)$$

と変形されている。I は単位行列で、A, B は、プログラム内の matA, B である。

よって (37.2) 式の両辺の右から Q^{-1} を掛けた $P = BQ^{-1}$ を, (37.1) 式に代入すると,

$$BQ^{-1}AQ = I$$

両辺の右から Q^{-1} を掛けた後に, 両辺の Q を掛けると, $QBQ^{-1}A = I$ なので,

$$A^{-1} = QBQ^{-1}$$

と言える. さらに命題 34.4 の, $T_1(i, j; n)^{-1} = {}^tT_1(i, j; n) = T_1(i, j; n)$ から, $Q^{-1} = {}^tQ$ が成り立つので, 結局

$$A^{-1} = QB {}^tQ$$

が成り立つ. $Q = T_1(i, j; n)$ の場合には, $QB {}^tQ$ は B の i 列と j 列を交換したあとに, i 行と j 行を交換する (行と列をどちらを先にしてもよい).

$Q = T_2(s, t; n)T_1(i, j; n)$ のときは,

$$A^{-1} = QB {}^tQ = T_2(s, t; n)T_1(i, j; n)BT_1(i, j; n)T_2(s, t; n)$$

この列変更部分 (B の右側) を見ると B の i 列と j 列を交換した後に s 列と t 列を交換している. 行変更部分を見ても, まず i 行と j 行を交換した後に, s 行と t 行を交換している.

有限個の積でも同様に, 行と列の交換について, 交換する順序が, 行と列で同じなのである. プログラム内の変数 `Qseq` には交換した結果が記録されているので, その順序に従って, B の列を変更して行を変更すれば, A^{-1} が得られるのである.

37.3 拡大係数行列による連立方程式の解や逆行列の復習

n 次正方行列 C が正則で, n 項列ベクトル b が与えられていて,

$$Ax = b$$

を満たす x を知りたいときには, 前節のプログラムで A の逆行列 A^{-1} を求めて, その後 $C^{-1}b$ を計算するという方法が 1 つ考えられる. もう一つはこのプログラムを使うが, 35 章の拡大係数行列の考え方を使い, 前節のプログラムを少し帰ることで, 答えが得られる.

まず前節の中プログラムでは, `matA` に $C = \begin{pmatrix} A & b \end{pmatrix}$ を入れて n 行 $n+1$ 列の行列にする. そして `process1` で基本行変形のみを行い (つまり絶対値の最大値を探す範囲を同じ列のみにする), その後 `process2` が終われば, この右端の列が求める答えとなる.

例題 37.5 行列 `matA` が与えられたときに, その逆関数を, 基本変形によって求める関数 `myinv(matA)` を作る.

- (i). 前節のプログラムのファイルを違う名前にして保存する.
- (ii). 関数名 `function [r,Qseq,detA,B]=Ftrans(matA,varargin)` を,
`function [A1 A2]=FtransR(matA,varargin)`

に変更する. 目的は m 行 n 列の `matA` が与えられたときに, 行のみの基本変形で前節のプロセス 2 までを行ったときに, 最初の m 列までの行列を `A1` で, 残りの $m+1$ 列目以降の行列を `A2` で返すものを作ることである.

(iii). 以下の部分を,

```
//Max_{i>=piv,j>=piv} abs(A(i,j))
//mxabsval is optimun value. istar & jstar is optimal argument.
mxabsval=max(abs(A(piv:m,piv:n)))
vecindex=min(find(abs(A(piv:m,piv:n))==mxabsval))
pivMatrow=m-piv+1
quot=floor((vecindex-1)/pivMatrow)
jstar=quot+piv
istar=vecindex-1-quot*pivMatrow+piv
```

次のように変更する.

```
//Max_{i>=piv} abs(A(i,piv))
//mxabsval is optimun value. istar is optimal argument.
mxabsval=max(abs(A(piv:m,piv)))
vecindex=min(find(abs(A(piv:m,piv))==mxabsval))
istar=vecindex-1+piv
```

(iv). 次の部分は, 第1プロセスの列変形なので, 削除する.

```
if jstar > piv then
    A(:, [piv jstar]) = A(:, [jstar piv])
    detA=-detA
    B(:, [piv jstar]) = B(:, [jstar piv])
    Qseq([piv jstar]) = Qseq([jstar piv])
    if FlagPrintA then; disp(A,"type1:col"+string(piv)+" <-> col"+string(jstar)); end
end
```

(v). //3rd process: this makes F(r) by type 3 columns transformation の行から,

endfunction の手前の end

まで全て削除する. 第3プロセスは不要なためである. そして以下のコードを書き入れる.

```
A1=A(:,1:m)
if m<n then
    A2=A(:,m+1:n)
else
    A2=[]
end
```

その後, Scinote での, 「実行する>ファイルを実行」によって, そのファイルを実行させよ.

」

例えば, $\mathbf{G} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 10 & 4 \\ 1 & 0 & 4 \end{pmatrix}$ で $\mathbf{q} = \begin{pmatrix} 3 \\ 10 \\ -2 \end{pmatrix}$ で, $\mathbf{G}\mathbf{x} = \mathbf{q}$ となる \mathbf{x} を求めたいとしよう. Scilab のメイン画面で, 組み込みの逆行列を使う場合には

```
G=[1 2 3;3 10 4;1 0 4]
```

```
q=[3; 10; -2]
```

```
inv(G)
```

```
inv(G)*q
```

で, 逆行列 \mathbf{G}^{-1} や, \mathbf{x} の答 $\mathbf{G}^{-1}\mathbf{q}$ が得られる. これを, 上のプログラムの基本行変形のみによる計算によって得られることを示そう. まず \mathbf{G} の右に \mathbf{q} を結合した行列を基本行変形することにより,

```
[G1,G2]=FtransR([G q])
```

を計算せよ. そして $\mathbf{G2}$ が $\mathbf{G}^{-1}\mathbf{q}$ と一致することを確認せよ. また \mathbf{G} の逆行列についてもメイン画面で, \mathbf{G}

の右に $\mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ を結合した行列によって,

```
[G3,G4]=FtransR([G eye(3,3)])
```

を計算せよ. そしてこの $\mathbf{G4}$ が先に計算した $\text{inv}(\mathbf{G})$ の値と一致することを確認すると良い. さらに前節のプログラム `Ftrans` が組み込まれているならば,

```
[rankG,seqQ,devG,G5]=Ftrans(G)
```

としてこの $\mathbf{G5}$ も, 当然それらと一致することを確認するとよい.

第 VII 部

固有値・対角化・三角化・2次形式

38 固有値・固有ベクトルと、対角化・三角化

固有値と固有ベクトルについては、2次正方行列の場合と全く同様に定義できる。

実ベクトル・実行列という場合には、ベクトルや行列は実数である。スカラーもそうである。そう書いていない場合は、ベクトルの成分・行列の成分・スカラーを、複素数の範囲まで考えている。この章の定義・定理・命題について、実数までの話なのか、複素数まで拡張可能なのかを具体的に書くと、

- (i). 定義 38.1・定義 38.2・定理 38.3 は、複素数の場合まで拡張可能である。
- (ii). 命題 38.7 は、実行列で成り立つことを言っているが、証明ではエルミート内積を用いて、複素数の範囲まで考えている。
- (iii). それ以外は、全て実数のみの場合である。

ではまず固有値・固有ベクトル・固有多項式の定義からである。2次正方行列の場合と全く同様である。

定義 38.1 n 次正方行列 \mathbf{A} と単位行列 \mathbf{I} について、

- (i). λ に関する n 次多項式

$$\det(\lambda\mathbf{I} - \mathbf{A}) \tag{38.1}$$

を固有多項式と呼び $\phi_{\mathbf{A}}(\lambda)$ であらわす。

- (ii). λ に関する n 次方程式

$$\det(\lambda\mathbf{I} - \mathbf{A}) = 0 \tag{38.2}$$

を固有方程式と呼ぶ。

- (iii). 固有方程式の解を固有値という。また、固有値 λ_0 に関して、

$$(\lambda_0\mathbf{I} - \mathbf{A})\mathbf{x}_0 = \mathbf{0} \tag{38.3}$$

をみたすベクトル $\mathbf{x}_0 \neq \mathbf{0}$ を、 λ_0 の固有ベクトルという。

」

ある λ_0 について、 \mathbf{x}_0 が固有ベクトルの1つであれば、任意のスカラー $t \neq 0$ について、 $t\mathbf{x}_0$ も固有ベクトルであることも、2次正方行列の場合と同様である。

対角化と三角化についても、2次正方行列の場合と全く同様に定義できる。

定義 38.2 n 次正方行列 \mathbf{A} と n 次単位行列 \mathbf{I} について、

- (i). 適当な正則 n 次行列 \mathbf{P} が存在して、

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{\Lambda} \tag{38.4}$$

と、ある n 次対角行列 $\mathbf{\Lambda}$ で表されるとき、 \mathbf{A} は \mathbf{P} によって $\mathbf{\Lambda}$ に対角化される、という。

- (ii). 適当な正則 n 次行列 \mathbf{P} が存在して、

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{U} \tag{38.5}$$

と、ある n 次上三角行列 U で表されるとき、 A は P によって U に三角化される、という。

A が P によって三角化された場合、 A の固有多項式と、三角行列 $P^{-1}AP$ の固有多項式と一致する。

定理 38.3 n 次正方行列 A, P について、 P は正則であるとする。このとき A の固有多項式 $\phi_A(\lambda)$ と $P^{-1}AP$ の固有多項式 $\phi_{P^{-1}AP}(\lambda)$ は等しい。

証明

$$\begin{aligned}\phi_{P^{-1}AP}(\lambda) &= \det(\lambda I - P^{-1}AP) = \det(\lambda P^{-1}IP - P^{-1}AP) \\ &= \det(P^{-1}(\lambda I - A)P) \\ &= \det(P^{-1}) \det(\lambda I - A) \det(P) \\ &= \det(\lambda I - A) \\ &= \phi_A(\lambda)\end{aligned}$$

□

A は対角化はできるとは限らない。しかし三角化は、固有方程式が n 個の解をもつならば、必ずできる。さらに A が実行列でかつ、固有方程式が n 個の解をもつならば、直交行列によって三角化できる。これらは 2 次正方行列の場合と同様である。 n 次正方行列の場合で証明する前にまずは、 2×2 の場合を思い出してほしい。1 つの基準化された固有ベクトル $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ をとったとき、 $\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ というベクトルを使って、

$$\begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix}$$

が直交行列になることを証明で用いた。ここでも直交行列をうまく作りたいため、そのための定理を先に示す。

定理 38.4 (グラム・シュミットの直交化の利用) n 個の n 次実列ベクトル $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ が、1 次独立であるとする。このとき、 \mathbf{a}_1 を基準化した列ベクトル $\mathbf{a}_1/\|\mathbf{a}_1\|$ をある列に含む直交行列 P を作ることができる。

証明 まず、 \mathbf{g}_1 を $\mathbf{a}_1/\|\mathbf{a}_1\|$ とする。 $\|\mathbf{g}_1\| = 1$ なので、 $\langle \mathbf{g}_1, \mathbf{g}_1 \rangle = 1$ である。

次に実数 c_1 を $c_1 = \langle \mathbf{a}_2, \mathbf{g}_1 \rangle$ として、 \mathbf{f}_2 を $\mathbf{f}_2 = \mathbf{a}_2 - c_1 \mathbf{g}_1$ とする。 \mathbf{g}_1 は \mathbf{a}_1 の定数倍なので、 \mathbf{a}_2 と \mathbf{a}_1 の 1 次独立性より、 $\mathbf{f}_2 \neq \mathbf{0}$ である。また計算すると、 $\langle \mathbf{f}_2, \mathbf{g}_1 \rangle = 0$ が成り立つ。そこで $\mathbf{g}_2 = \mathbf{f}_2/\|\mathbf{f}_2\|$ とする。 $\|\mathbf{g}_2\| = 1$ かつ、 $\langle \mathbf{g}_1, \mathbf{g}_2 \rangle = 0$ がいえる。

以下 $i = 3, \dots, n$ の順に、以下のプロセスを繰り返す。

- (i). $i-1$ 個の実数 c_1, c_2, \dots, c_{i-1} を、 $c_k = \langle \mathbf{a}_i, \mathbf{g}_k \rangle$ ($k = 1, 2, \dots, i-1$) とする。
- (ii). 次に \mathbf{f}_i を、 $\mathbf{f}_i = \mathbf{a}_i - \sum_{k=1}^{i-1} c_k \mathbf{g}_k$ とする。 $\sum_{k=1}^{i-1} c_k \mathbf{g}_k$ は $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{i-1}$ の線形結合なので、1 次独立性より、 $\mathbf{f}_i \neq \mathbf{0}$ である。式より、 $k = 1, 2, \dots, i-1$ について、 $\langle \mathbf{f}_i, \mathbf{g}_k \rangle = 0$ が成り立つ。
- (iii). $\mathbf{g}_i = \mathbf{f}_i/\|\mathbf{f}_i\|$ とする。 $\|\mathbf{g}_i\| = 1$ であり、(ii) より $k = 1, 2, \dots, i-1$ について、 $\langle \mathbf{g}_i, \mathbf{g}_k \rangle = 0$ が成り立つ。

このように得られた $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$ を横に並べた行列

$$\mathbf{P} = (\mathbf{g}_1 \quad \mathbf{g}_2 \quad \cdots \quad \mathbf{g}_n)$$

は (上のは一例であり, 並び順は任意でよい), \mathbf{g}_i の作り方から, ${}^t\mathbf{P}\mathbf{P} = \mathbf{I}$ を満たす.

□

系 38.5 任意の n 次実列ベクトル $\mathbf{q} \neq \mathbf{0}$ について, \mathbf{q} を基準化した列ベクトル $\mathbf{q}/\|\mathbf{q}\|$ を任意の列に含む n 次直交行列 \mathbf{P} を作ることができる. 」

証明 少なくとも 1 つの i があって, \mathbf{q} の第 i 成分非 0 である. この i について, 単位行列 \mathbf{I} の i 列目を \mathbf{q} に変更した行列を作った後, 1 列目と i 列目を変更する. この行列は正則であり, 前定理より 1 列目が $\mathbf{q}/\|\mathbf{q}\|$ となるような直交行列 \mathbf{P} を作るができる. さらに直交行列の性質から, 任意の $k \geq 2$ について, 1 列目と k 列目を入れ替えても, 直交行列であるので, 命題が成り立つ.

□

いよいよ目的の三角化である.

定理 38.6 n 次正方実行列 \mathbf{A} の固有多項式が重複を含めて n 個の実数解 $\lambda_1, \lambda_2, \dots, \lambda_n$ をもつ場合, ある適当な直交行列 \mathbf{P} が存在して,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & u_{12} & \cdots & u_{1n} \\ 0 & \lambda_2 & \cdots & u_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad (38.6)$$

と, 対角成分が固有値となるような上三角行列に三角化できる. 」

証明 数学的帰納法で証明する. $n = 1$ のとき成り立つのは明らか. よって $n \geq 2$ として $n - 1$ 次正方行列の場合に, 命題が成り立つと仮定する. n 次正方行列 \mathbf{A} の固有値の 1 つを λ_1 とし, その固有ベクトルを \mathbf{q} (ただし $\|\mathbf{q}\| = 1$ と基準化している) とすると,

$$\mathbf{A}\mathbf{q} = \lambda_1\mathbf{q} \quad (38.7)$$

また, 系 38.5 より, n 行 $n - 1$ 列の行列 \mathbf{Q}_2 を用いて

$$\mathbf{Q} = (\mathbf{q} \quad \mathbf{Q}_2)$$

という形の直交行列 \mathbf{Q} が存在する. 従って,

$${}^t\mathbf{Q}_2\mathbf{q} = {}^t\mathbf{0}, \quad {}^t\mathbf{Q}_2\mathbf{Q}_2 = \mathbf{I}_{n-1}$$

である. 以下 \mathbf{I}_{n-1} は $(n - 1)$ 次単位行列とする.

この \mathbf{Q} を用いて $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ を計算すると,

$$\begin{aligned}
Q^{-1}AQ &= {}^tQAQ = \begin{pmatrix} {}^tq \\ {}^tQ_2 \end{pmatrix} A (q \ Q_2) \\
&= \begin{pmatrix} {}^tqAq & {}^tqAQ_2 \\ {}^tQ_2Aq & {}^tQ_2AQ_2 \end{pmatrix} \\
&= \begin{pmatrix} {}^tq\lambda_1q & {}^tqAQ_2 \\ {}^tQ_2\lambda_1q & {}^tQ_2AQ_2 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & {}^tqAQ_2 \\ \mathbf{0} & Q_2^{-1}AQ_2 \end{pmatrix} \tag{38.8}
\end{aligned}$$

両辺の固有多項式を求めると,

$$\phi_{Q^{-1}AQ}(\lambda) = \det \begin{pmatrix} \lambda - \lambda_1 & {}^tqAQ_2 \\ \mathbf{0} & \lambda I_{n-1} - Q_2^{-1}AQ_2 \end{pmatrix}$$

左辺は定理 38.3 より, 右辺は分割された行列式より, それぞれ計算して

$$\phi_A(\lambda) = (\lambda - \lambda_1) \det(\lambda I_{n-1} - Q_2^{-1}AQ_2)$$

左辺は A の固有方程式に関する仮定より, 右辺も固有多項式に直して,

$$\prod_{i=1}^n (\lambda - \lambda_i) = (\lambda - \lambda_1) \phi_{Q_2^{-1}AQ_2}(\lambda)$$

この式は λ について恒等的に成り立つので,

$$\phi_{Q_2^{-1}AQ_2}(\lambda) = \prod_{i=2}^n (\lambda - \lambda_i)$$

が成り立つ. この式において, $Q_2^{-1}AQ_2$ は $n-1$ 次正方行列なので, 帰納法の仮定より, 適当な $n-1$ 次直交行列 P_2 が存在して,

$$P_2^{-1}(Q_2^{-1}AQ_2)P_2 = \begin{pmatrix} \lambda_2 & * & * \\ & \ddots & * \\ \mathbf{0} & & \lambda_n \end{pmatrix}$$

と三角化される ($i < j$ となる (i, j) 成分には適当な数が入る). この右辺を U_2 として, 左と右からそれぞれ P_2, P_2^{-1} を両辺から掛けると,

$$Q_2^{-1}AQ_2 = P_2 U_2 P_2^{-1}$$

であり, これを (38.8) 式に代入すると,

$$\begin{aligned}
Q^{-1}AQ &= \begin{pmatrix} \lambda_1 & {}^tqAQ_2 \\ \mathbf{0} & Q_2^{-1}AQ_2 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & {}^tqAQ_2 \\ \mathbf{0} & P_2 U_2 P_2^{-1} \end{pmatrix} \\
&= \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & {}^tqAQ_2 P_2 \\ \mathbf{0} & U_2 \end{pmatrix} \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2^{-1} \end{pmatrix}
\end{aligned}$$

この式で P_2 が直交行列であることから, $\begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix}$ も直交行列である. そこで, 両辺の左側から $\begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix}^{-1}$ を掛けて, 両辺の右側から $\begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix}$ を掛けて

$$\begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix}^{-1} Q^{-1} A Q \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & {}^t q A Q_2 P_2 \\ \mathbf{0} & U_2 \end{pmatrix}$$

が成り立つ. この式は, A が, 直交行列 $Q \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix}$ によって, 上三角行列に三角化されていることを表している. よって $n-1$ の時に命題が成り立つならば, n のとき命題が成り立つ.

以上から数学的帰納法より, 命題が成り立つ.

□

次に実対称行列について考えよう. 2 次の場合と同様に, まずは固有値が全て実数であることを示そう.

命題 38.7 n 次実対称行列の固有値は, 重複も含めて, n 個全て実数である. 」

証明 A が実対称行列であるとする. 代数学の基本定理より, A の固有方程式の解は, 複素数の範囲で, 重複も含めて n 個存在する. この複素数解の 1 つを λ_0 とし, この固有値 λ_0 の固有ベクトルを \mathbf{x}_0 とする.

エルミート内積 (定義 17.2) を用いて,

$$\mathbf{x}_0^* A \mathbf{x}_0 = \langle A \mathbf{x}_0, \mathbf{x}_0 \rangle = \langle \lambda_0 \mathbf{x}_0, \mathbf{x}_0 \rangle = \lambda_0 \langle \mathbf{x}_0, \mathbf{x}_0 \rangle$$

同様に,

$$\begin{aligned} \mathbf{x}_0^* A \mathbf{x}_0 &= \langle \mathbf{x}_0, A^* \mathbf{x}_0 \rangle \\ (A \text{ は対称行列なので}) &= \langle \mathbf{x}_0, A \mathbf{x}_0 \rangle \\ &= \langle \mathbf{x}_0, \lambda_0 \mathbf{x}_0 \rangle \\ &= \bar{\lambda}_0 \langle \mathbf{x}_0, \mathbf{x}_0 \rangle \end{aligned}$$

従って, この 2 つの式から,

$$\lambda_0 \langle \mathbf{x}_0, \mathbf{x}_0 \rangle = \bar{\lambda}_0 \langle \mathbf{x}_0, \mathbf{x}_0 \rangle$$

\mathbf{x}_0 は固有ベクトルなので $\mathbf{x}_0 \neq \mathbf{0}$, よって $\langle \mathbf{x}_0, \mathbf{x}_0 \rangle \neq 0$. だから,

$$\lambda_0 = \bar{\lambda}_0$$

従って, λ_0 は実数である. λ_0 は任意の解なので, 命題が成り立つ.

□

さて以下が, 微分積分でも重要な働きをする定理である.

定理 38.8 n 次実対称行列 A は, ある適当な直交行列 P によって, 対角成分が固有値である対角行列 Λ に対角化できる. つまり,

$$P^{-1} A P = \Lambda$$

」

証明 定理 38.6 と命題 38.7 から, n 次実対称行列 A は, ある適当な直交行列 P によって, 固有値を対角成分に持つ上三角行列 U に対角化される. つまり,

$$P^{-1}AP = U$$

P は直交行列なので,

$${}^tPAP = U \quad (38.9)$$

両辺の転置行列をとると,

$${}^tP^tA({}^tP) = {}^tU$$

A は対称行列なので,

$${}^tPAP = {}^tU \quad (38.10)$$

(38.9) 式と (38.10) 式から,

$$U = {}^tU$$

左辺は上三角行列なので, 左辺の, $i > j$ のときの (i, j) 成分は 0 である. 右辺は下三角行列なので, 左辺の, $i < j$ のときの (i, j) 成分は 0 である. よって U は対角行列である. U の対角成分は固有値なので, 命題が成り立つ.

□

定理 38.9 n 次実対称行列 A がべき等行列 $A^2 = A$ のとき, つまり A がべき等対称行列のとき,

- (i). 固有値は 0 または 1 のみである.
- (ii). A のランクは, トレース $tr(A)$ と等しい.

┘

証明 (i) A は実対称行列なので, 重複を含めて n 個の固有値は全て実数である. n 個の固有値の任意の 1 つを λ とし, そのときの固有ベクトルを x とする. 定義より, $Ax = \lambda x$ が成り立つ. また,

$$Ax = A^2x = A\lambda x = \lambda^2x$$

が成り立つので,

$$\lambda x = \lambda^2x$$

がいえる. $x \neq 0$ なので, これは $\lambda^2 = \lambda$ と同値である. これを解けば

$$\lambda = 0, 1$$

(ii) (i) の結果と前定理から, A はある適当な直交行列 P によって,

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad (38.11)$$

と対角化される. $\lambda_1, \lambda_2, \dots, \lambda_n$ は 0 か 1 のどちらかであり, 適当な P によって,

$$1 \geq \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$$

を仮定してもよいので、そう仮定する。 P, P^{-1} が正則であることから、右辺は A の標準形である。したがって、右辺のトレースは、 A のランクを表す。一方 (38.11) 式の左辺のトレースは

$$\text{tr}(P^{-1}AP) = \text{tr}(PP^{-1}A) = \text{tr}(A)$$

である。以上より、命題は成り立つ。

□

39 対称行列を表す「'」と 2 次形式と定符号行列

14 章で説明した 2 次形式の、 n 次元バージョンである。この章では、 A は n 次の実対称行列であり、 x は n 項実列ベクトルとする。

そしてこの章以降実行列しか出てこないで、転置行列を ' X ' ではなく、 X' で表すこととする。というのは、米国の教科書でもこのように書いてあることが多いためである。

また PC プログラミング言語の、Scilab, Matlab, Octava, Gauss では、実行列の場合ならば、 ' X ' の代わりに X' を使えば同じ効果が得られる。上記ソフトウェアでは、行列 X について X' は、転置したものの共役複素数を返す。仮に $a = \begin{pmatrix} 1 & 2+i \end{pmatrix}$ であれば、 a' と書くと、 $\begin{pmatrix} 1 \\ 2-i \end{pmatrix}$ を返す。従って成分が全て実数ならば、' も転置行列を表すこととなる。

さて 2 次形式の定義については、定義 13.1 で済ませてあり、正值定符号・負値定符号・半正值定符号・半負値定符号の定義も、定義 13.2 で示した通りである。

この章では一般の n 次正方行列の場合での、2 次形式の性質を述べる。まず、前章で示した実対称行列の直交行列による対角化の結果から、以下が簡単に示される。

定理 39.1 n 次実対称行列 A について、

- (i). A が正值定符号であることは、 A の固有値が全て正であることと同値である。
- (ii). A が負値定符号であることは、 A の固有値が全て負であることと同値である。
- (iii). A が半正值定符号であることは、 A の固有値が全て 0 以上であることと同値である。
- (iv). A が半負値定符号であることは、 A の固有値が全て 0 以下であることと同値である。

┘

証明 (i) のみ証明しよう。 A は対称行列なので、ある直交行列 P と、対角成分が固有値である実対角行列 Λ によって、 $A = P\Lambda P'$ と表される。 Λ の (i, i) 成分を λ_i 、ある x について、 $q = P'x$ とすると、

$$x'Ax = (x'P)\Lambda(P'x) = q'\Lambda q = \lambda_1 q_1^2 + \lambda_2 q_2^2 + \cdots + \lambda_n q_n^2 \quad (39.1)$$

が成り立つ。

⇐ の証明: P は正則なので、 $q \neq 0$ と $x \neq 0$ は同値である。よって、(39.1) 式から、任意の $x \neq 0$ について $x'Ax > 0$ がいえるから、 A は正值定符号。

⇒ の証明: 背理法で証明する。少なくとも 1 つの固有値が 0 以下であると仮定する。 $\lambda_1 \leq 0$ と仮定しても一般性は失わないのでそう仮定する。このとき、 q が $q_1 = 1, q_2 = \cdots = q_n = 0$ となるような $x \neq 0$ は存

在する (実際計算すると \mathbf{P} の 1 列目である). このとき, (39.1) 式から, $\mathbf{x}'\mathbf{A}\mathbf{x} = \lambda_1 \leq 0$ となる. これは \mathbf{A} が正値定符号であることに矛盾する. よって全ての固有値は正である.

□

\mathbf{A} が正値定符号ならば, 任意の $\mathbf{x} \neq \mathbf{0}$ について $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ であったから, \mathbf{x} の成分に適当に 0 を代入することによって, \mathbf{A} から成分をいくつか抜き出した行列も, やはり正値定符号であることが言える. 3 次正方行列を例にして説明しよう.

例題 39.2 $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{33} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ が正値定符号であると仮定する. このとき,

(i). $a_{11} > 0, a_{22} > 0, a_{33} > 0$ であることを確かめよ.

(ii). \mathbf{A} から, $1 \leq i < j \leq 3$ について取り出した 2 行 2 列の行列 $\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix}$ は全部で 3 種類考えられる. そのどの行列も正値定符号であることを示せ.

┘

(解答) (i) の証明: 単位ベクトル

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

とすると, \mathbf{A} が正値定符号であることから, どの $i = 1, 2, 3$ についても, $\mathbf{e}_i'\mathbf{A}\mathbf{e}_i > 0$ である. $\mathbf{e}_i'\mathbf{A}\mathbf{e}_i = a_{ii}$ であるから, $a_{11} > 0, a_{22} > 0, a_{33} > 0$ が言える.

(ii) の証明: 単位ベクトルを用いて $\mathbf{x} = y_i\mathbf{e}_i + y_j\mathbf{e}_j$ とすると, \mathbf{A} が正値定符号であることから, $\mathbf{y} = \begin{pmatrix} y_i \\ y_j \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ のとき, $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ であるから,

$$0 < \mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{y}' \begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix} \mathbf{y}$$

である. これは, 対称行列 $\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix}$ が正値定符号であることを表す.

(解答終)

よって, 任意の n 次正方行列についても, 同様の証明によって, 以下が言える.

命題 39.3 n 次実対称行列 \mathbf{A} について,

(i). \mathbf{A} が正値定符号ならば, 任意の $k (1 \leq k \leq n)$ をとり, 任意の k 個の整数 $i_1, i_2, \dots, i_k (1 \leq i_1 < i_2 < \dots < i_k \leq n)$ について,

$$\begin{pmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_k} \\ a_{i_2 i_1} & a_{i_2 i_2} & \cdots & a_{i_2 i_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k i_1} & a_{i_k i_2} & \cdots & a_{i_k i_k} \end{pmatrix}$$

が正値定符号である.

- (ii). (i) の 2 つの「正値定符号」の部分で「負値定符号」とした場合も、同様に成り立つ。
 (iii). (i) の 2 つの「正値定符号」の部分で「半正値定符号」とした場合も、同様に成り立つ。
 (iv). (i) の 2 つの「正値定符号」の部分で「半負値定符号」とした場合も、同様に成り立つ。

」

ところで上の命題の (i) で書いた行列は全部で $2^n - 1$ 個あるわけだが、全てを調べるにはそれは多すぎる。実は n 個の行列を調べることで、正値定符号性と負値定符号性がいえる場合がある。そのためにまずは、任意の行列 \mathbf{A} について、その左上からの正方行列を定義しよう。

定義 39.4 m 行 n 列の行列 \mathbf{A} について、 $1 \leq k \leq \min(m, n)$ となる k をとって、

$$\begin{pmatrix} a_{11} & \cdots & a_{k1} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

となる正方行列をとったとき、これを \mathbf{A} の k 次首座小行列という。さらにこれの行列式、

$$\det \begin{pmatrix} a_{11} & \cdots & a_{k1} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

を、 \mathbf{A} の k 次首座小行列式という。

」

これを用いて、

定理 39.5 n 次実対称行列 \mathbf{A} を考える。

- (i). \mathbf{A} の任意の首座小行列式が正であるとき、つまり、任意の $k = 1, 2, \dots, n$ について

$$\det \begin{pmatrix} a_{11} & \cdots & a_{k1} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0$$

が成り立つならば、 \mathbf{A} は正値定符号である。

- (ii). $k = 1, 2, \dots, n$ の全ての k 次首座小行列式について、

$$(-1)^k \det \begin{pmatrix} a_{11} & \cdots & a_{k1} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0 \tag{39.2}$$

が成り立つならば、 \mathbf{A} は負値定符号である。

」

注意: 2 次対称行列の場合は、定理 14.3 で示している。

(注意終)

証明 \mathbf{A} が負値定符号であることは $-\mathbf{A}$ が正値定符号と同値なので、(i) のみ数学的帰納法で証明する。
 $n = 1$ のときは明らかに成り立つので、 $n \geq 2$ とする。

$n-1$ 次正方行列のときは命題が成り立つとする \mathbf{A} を $(n-1)$ 次正方行列 \mathbf{A}_1 と $(n-1)$ 次列ベクトル \mathbf{a}_2 とスカラー a によって

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{a}_2 \\ \mathbf{a}'_2 & a \end{pmatrix}$$

と分割する. このとき, 分割行列の行列式より, $\det(\mathbf{A}) = \det(\mathbf{A}_1) \det(a - \mathbf{a}'_2 \mathbf{A}_1^{-1} \mathbf{a}_2)$ が成り立つ. また帰納法の仮定より \mathbf{A}_1 は正値定符号である.

仮定より $\det(\mathbf{A}) > 0, \det(\mathbf{A}_1) > 0$ であり, $a - \mathbf{a}'_2 \mathbf{A}_1^{-1} \mathbf{a}_2$ が実数なので,

$$a > \mathbf{a}'_2 \mathbf{A}_1^{-1} \mathbf{a}_2 \quad (39.3)$$

がいえる. n 次列ベクトル \mathbf{x} を, $n-1$ 次列ベクトル \mathbf{u} とスカラー v で

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ v \end{pmatrix}$$

と分割する. このとき \mathbf{A} の 2 次形式は,

$$\begin{aligned} \mathbf{x}' \mathbf{A} \mathbf{x} &= (\mathbf{u}' \quad v) \begin{pmatrix} \mathbf{A}_1 & \mathbf{a}_2 \\ \mathbf{a}'_2 & a \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ v \end{pmatrix} \\ &= \mathbf{u}' \mathbf{A}_1 \mathbf{u} + \mathbf{u}' \mathbf{a}_2 v + v \mathbf{a}'_2 \mathbf{u} + av^2 \\ &= \mathbf{u}' \mathbf{A}_1 \mathbf{u} + 2(\mathbf{a}'_2 \mathbf{u})v + av^2 \\ (39.3) \text{ 式より, } &> \mathbf{u}' \mathbf{A}_1 \mathbf{u} + 2(\mathbf{a}'_2 \mathbf{u})v + (\mathbf{a}'_2 \mathbf{A}_1^{-1} \mathbf{a}_2)v^2 \\ &= (\mathbf{u}' + v \mathbf{a}'_2 \mathbf{A}_1^{-1}) \mathbf{A}_1 (\mathbf{u} + \mathbf{A}_1^{-1} \mathbf{a}_2 v) \\ &= (\mathbf{u} + \mathbf{A}_1^{-1} \mathbf{a}_2 v)' \mathbf{A}_1 (\mathbf{u} + \mathbf{A}_1^{-1} \mathbf{a}_2 v) \end{aligned} \quad (39.4)$$

$$\mathbf{A}_1 \text{ は正値定符号なので, } \geq 0 \quad (39.5)$$

以上より, 任意の $\mathbf{x} \neq \mathbf{0}$ について $\mathbf{x}' \mathbf{A} \mathbf{x} > 0$ となるので, n 次正方行列のときも命題が成り立つ.

以上より, 数学的帰納法から, 任意の正整数 n について命題が成り立つ.

□

ここからは正値定符号の性質を中心に説明しよう. まず正値定符号は逆行列について閉じている.

命題 39.6 \mathbf{A} が正値定符号ならば, \mathbf{A}^{-1} も正値定符号である. このとき \mathbf{A}^{-1} の固有値全ては, \mathbf{A} の固有値全ての逆数である. ┘

証明 n 次正方行列 \mathbf{A} について, はある直交行列 \mathbf{P} があって, 対角成分が全て \mathbf{A} の正の固有値である対角行列 $\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}$ があって, $\mathbf{P}' \mathbf{A} \mathbf{P} = \mathbf{\Lambda}$ と対角化できる. $\mathbf{P}' = \mathbf{P}^{-1}$ に気をつけて両辺の逆行列をとると,

$$\mathbf{P}' \mathbf{A}^{-1} \mathbf{P} = \begin{pmatrix} 1/\lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1/\lambda_n \end{pmatrix}$$

従って命題が成り立つ.

□

一般に正方行列には $1/2$ 乗や $-1/2$ 乗が存在しないが, 以下の行列のみ定義する.

定義 39.7 対角成分が全て 0 以上の対角行列 Λ について, その対角成分を全て平方根にした行列を $\Lambda^{1/2}$ で表す. さらに Λ の対角成分が全て正のときは, $\Lambda^{1/2}$ の逆行列を $\Lambda^{-1/2}$ で表す. 」

明らかに $\Lambda^{-1/2}$ は, $\Lambda^{1/2}$ の対角成分を逆数にした対角行列である. 2次正方行列を例にして示そう. ならば $a > 0, b > 0$ として,

$$\Lambda = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \implies \Lambda^{1/2} = \begin{pmatrix} \sqrt{a} & 0 \\ 0 & \sqrt{b} \end{pmatrix}, \Lambda^{-1/2} = \begin{pmatrix} 1/\sqrt{a} & 0 \\ 0 & 1/\sqrt{b} \end{pmatrix}$$

命題 34.17 で実行列 A のランクと, $A'A$ のランクの一致を示した. このような $A'A$ と表す行列は正値定符号と関係が深く, また統計などでもよく出てくる. 上の命題と合わせて, 以下が簡単に言える.

定理 39.8 $n \times k$ の行列 X について,

- (i). $\text{rank}(X) = k$ のとき, つまり X が Full Column rank のとき, $X'X, (X'X)^{-1}$ はともに正値定符号である.
- (ii). $\text{rank}(X) < k$ のとき, $X'X$ は半正値定符号であるが, 正値定符号ではない. 従って $(X'X)^{-1}$ は存在しない. 」

ある程度逆の形も言える.

命題 39.9 A が n 次の正値定符号行列ならば, ある n 次正方行列 X によって

$$A = X'X$$

が成り立つ. 」

証明 正値定符号のときに命題 39.6 で P, Λ について, $A = P\Lambda P'$ を示しているので, $X = \Lambda^{1/2}P'$ とすればよい. □

この証明から以下も言える.

系 39.10 A が n 次の半正値定符号行列ならば, ある n 次直交行列 Q と, 対角成分が 0 以上の対角行列 Λ があって,

$$A = (\Lambda^{1/2}Q)'(\Lambda^{1/2}Q)$$

が成り立つ. 」

$X'X = X'IX$ であるが, 定理 39.8 で, X' と X の間に I の代わりに正値定符号行列を挟んでも, X が Full Column Rank であれば, 以下のように正値定符号性が伝播することが言える.

系 39.11 A を n 次の正値定符号とする. X が $n \times k$ で $\text{rank}(X) = k$ であれば,

$$X'AX$$

も正値定符号であり, 当然ランク k である. 」

証明 前命題の P, Λ について,

$$X'AX = X'P\Lambda^{1/2}\Lambda^{1/2}P'X$$

より, $Y = \Lambda^{1/2}P'X$ とすると, $\Lambda^{1/2}P'$ は正則なので, 定理 34.14 より, $\text{rank}(Y) = \text{rank}(X) = k$ であり, かつ

$$X'AX = Y'Y$$

Y は $n \times k$ の行列なので, 定理 39.8 より $Y'Y$ は正値定符号である. よって命題はなりたつ.

□

40 2次形式までを n 変数実数値関数としてみる.

40.1 偏微分・ヤコビ行列・ヘッセ行列

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「 n 変数実数値関数の定義から 1 階微分までの重要事項」章や「 n 変数 m 実数値関数の定義と偏導関数」章から, 定数 a_1, \dots, a_n について, x_1, \dots, x_n の n 変数実数値関数

$$f(x_1, \dots, x_n) = a_1x_1 + \dots + a_nx_n$$

の偏導関数ベクトル (ヤコビ行列) は,

$$\frac{\partial}{\partial \mathbf{x}}(x_1, \dots, x_n) = \left(\frac{\partial f}{\partial x_1} \quad \dots \quad \frac{\partial f}{\partial x_n} \right) = (a_1 \quad \dots \quad a_n)$$

で表される. よって以下が成り立つ.

命題 40.1 \mathbf{a}, \mathbf{x} を n 次列ベクトルとする. このとき,

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{a}'\mathbf{x}) = \mathbf{a}'$$

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{a}) = \mathbf{a}'$$

」

この命題の \mathbf{a} の代わりに, $m \times n$ の行列 \mathbf{A} を考えると, $\mathbf{A}\mathbf{x}$ は x_1, \dots, x_n についての, n 変数 m 項列ベクトル値関数となる. 同様にヤコビ行列を考えると, 以下となる.

命題 40.2 \mathbf{A} は $m \times n$ の行列, \mathbf{x} を n 次列ベクトルとする. このとき, n 変数 m 項列ベクトル値関数 $\mathbf{A}\mathbf{x}$ のヤコビ行列は,

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{A}\mathbf{x}) = \mathbf{A}$$

」

次に, \mathbf{x} はそのまま, \mathbf{A} が n 次正方行列とする. このとき $\mathbf{x}'\mathbf{A}\mathbf{x}$ は実数なので, \mathbf{x} による n 変数実数値関数と考えられる. この偏導関数ベクトルは, $\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{A}\mathbf{x})$ は, 上の 2 つの命題と偏微分の性質から, $p = \mathbf{A}\mathbf{x}, \mathbf{q}' = \mathbf{x}'\mathbf{A}$ として,

$$\begin{aligned}\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{A}\mathbf{x}) &= \frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'p) + \frac{\partial}{\partial \mathbf{x}}(\mathbf{q}'\mathbf{x}) \\ &= p' + q' \\ &= \mathbf{x}'\mathbf{A}' + \mathbf{x}'\mathbf{A} \\ &= \mathbf{x}'(\mathbf{A} + \mathbf{A}')\end{aligned}$$

これは n 項行ベクトルである. ヘッセ行列はこれを列ベクトルにした

$${}^t \left[\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{A}\mathbf{x}) \right] = (\mathbf{A}' + \mathbf{A})\mathbf{x}$$

を, n 変数 n 項列ベクトル値関数と見て, \mathbf{x} で偏微分したものなので, 前命題より,

$$\frac{\partial}{\partial \mathbf{x}} {}^t \left[\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}'\mathbf{A}\mathbf{x}) \right] = \mathbf{A}' + \mathbf{A}$$

この左辺の表記が面倒なので,

定義 40.3 開集合で定義された \mathbf{x} の n 変数実数値関数 $f(\mathbf{x})$ が C^1 級のとき, 偏導関数ベクトル

$$\frac{\partial}{\partial \mathbf{x}} f(\mathbf{x})$$

を列ベクトル化したものを

$$\frac{\partial}{\partial \mathbf{x}'} f(\mathbf{x}), \frac{\partial}{\partial {}^t \mathbf{x}} f(\mathbf{x})$$

などで表す. ┘

こう定義すれば,

命題 40.4 \mathbf{A} は $m \times n$ 次行列で, \mathbf{B} は n 次正方行列とする. \mathbf{x} は n 次列ベクトルとする. \mathbf{x} の n 変数 m 項ベクトル値関数 $\mathbf{A}\mathbf{x}$ と \mathbf{x} の n 変数実数値関数 $\mathbf{x}'\mathbf{B}\mathbf{x}$ について, 以下が成り立つ.

$$\frac{\partial}{\partial \mathbf{x}'} (\mathbf{A}\mathbf{x}) = \mathbf{A}' \tag{40.1}$$

$$\frac{\partial}{\partial \mathbf{x}'} (\mathbf{x}'\mathbf{A}) = \mathbf{A} \tag{40.2}$$

$$\frac{\partial}{\partial \mathbf{x}'} (\mathbf{x}'\mathbf{B}\mathbf{x}) = (\mathbf{B} + \mathbf{B}')\mathbf{x} \tag{40.3}$$

$$\frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{x}'} ({}^t \mathbf{x}\mathbf{B}\mathbf{x}) = \mathbf{B} + \mathbf{B}' \tag{40.4}$$

2 次形式では \mathbf{A} が対称行列だったので, その場合は以下となる. ┘

系 40.5 上の命題で \mathbf{A} は n 次対称行列のとき

$$\frac{\partial}{\partial \mathbf{x}'}(\mathbf{A}\mathbf{x}) = \mathbf{A} \quad (40.5)$$

$$\frac{\partial}{\partial \mathbf{x}'}(\mathbf{x}'\mathbf{A}) = \mathbf{A} \quad (40.6)$$

$$\frac{\partial}{\partial \mathbf{x}'}(\mathbf{x}'\mathbf{A}\mathbf{x}) = 2\mathbf{A}\mathbf{x} \quad (40.7)$$

$$\frac{\partial^2}{\partial \mathbf{x}\partial \mathbf{x}'}(\mathbf{x}'\mathbf{A}\mathbf{x}) = 2\mathbf{A} \quad (40.8)$$

」

40.2 2 次形式と関数の凹凸と最大・最小

引き続き、定数の n 次正方対称行列 \mathbf{A} について、 n 変数実数値関数 $\mathbf{x}'\mathbf{A}\mathbf{x}$ を考える。

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「 n 変数実数値関数の凹凸性と偏導関数：次章・次々章のための準備」の命題 7.4 で、凹関数や凸関数が和について閉じていることを示した。また、 ax^2 が $a > 0$ のときは x の狭義凸関数で $a < 0$ のときは x の凹関数であることと、定理 39.1 の証明で各 q_i をある種の合成関数と見れば、比較的簡単に以下が言える。以下では合成関数を考えずに、示すとする。

命題 40.6 \mathbf{A} は n 次正方行列で、それぞれの成分は定数とする。 \mathbf{x} は n 次列ベクトルとする。 \mathbf{x} の n 変数実数値関数 $\mathbf{x}'\mathbf{A}\mathbf{x}$ について、以下が成り立つ。

- (i). \mathbf{A} が正値定符号ならば、 $\mathbf{x}'\mathbf{A}\mathbf{x}$ は狭義凸関数である。
- (ii). \mathbf{A} が負値定符号ならば、 $\mathbf{x}'\mathbf{A}\mathbf{x}$ は狭義凹関数である。
- (iii). \mathbf{A} が半正値定符号ならば、 $\mathbf{x}'\mathbf{A}\mathbf{x}$ は凸関数である。
- (iv). \mathbf{A} が半負値定符号ならば、 $\mathbf{x}'\mathbf{A}\mathbf{x}$ は凹関数である。

」

証明 同様にできるので (i) のみ証明する。 \mathbf{x} の関数 $f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ について、任意の 2 つの異なる n 項列ベクトル \mathbf{u}, \mathbf{v} と、任意の $\lambda \in (0, 1)$ について、 $f((1-\lambda)\mathbf{u} + \lambda\mathbf{v}) < (1-\lambda)f(\mathbf{u}) + \lambda f(\mathbf{v})$ を示せば良い。

$$\begin{aligned} & f((1-\lambda)\mathbf{u} + \lambda\mathbf{v}) - [(1-\lambda)f(\mathbf{u}) + \lambda f(\mathbf{v})] \\ &= ((1-\lambda)\mathbf{u} + \lambda\mathbf{v})'\mathbf{A}((1-\lambda)\mathbf{u} + \lambda\mathbf{v}) - (1-\lambda)\mathbf{u}'\mathbf{A}\mathbf{u} - \lambda\mathbf{v}'\mathbf{A}\mathbf{v} \\ &= (1-\lambda)^2\mathbf{u}'\mathbf{A}\mathbf{u} + \lambda(1-\lambda)\mathbf{v}'\mathbf{A}\mathbf{u} + \lambda(1-\lambda)\mathbf{u}'\mathbf{A}\mathbf{v} + \lambda^2\mathbf{v}'\mathbf{A}\mathbf{v} - (1-\lambda)\mathbf{u}'\mathbf{A}\mathbf{u} - \lambda\mathbf{v}'\mathbf{A}\mathbf{v} \\ &= -\lambda(1-\lambda)\mathbf{u}'\mathbf{A}\mathbf{u} + \lambda(1-\lambda)\mathbf{v}'\mathbf{A}\mathbf{u} + \lambda(1-\lambda)\mathbf{u}'\mathbf{A}\mathbf{v} - \lambda(1-\lambda)\mathbf{v}'\mathbf{A}\mathbf{v} \\ &= -\lambda(1-\lambda)[\mathbf{u}'\mathbf{A}\mathbf{u} - \mathbf{v}'\mathbf{A}\mathbf{u} + -\mathbf{u}'\mathbf{A}\mathbf{v} + \mathbf{v}'\mathbf{A}\mathbf{v} - (1-\lambda)\mathbf{u}'\mathbf{A}\mathbf{u} - \lambda\mathbf{v}'\mathbf{A}\mathbf{v}] \\ &= -\lambda(1-\lambda) \cdot (\mathbf{u} - \mathbf{v})'\mathbf{A}(\mathbf{u} - \mathbf{v}) \end{aligned}$$

\mathbf{A} は正値定符号なので < 0

□

線形関数 $-p_1x_1 - \cdots - p_nx_n$ は、凹関数でもあり凸関数でもあるので、 $\mathbf{x}'\mathbf{A}\mathbf{x}$ に加えても、上の結論は変わらない。

系 40.7 \mathbf{A} は n 次正方行列で、 \mathbf{p}' は n 次行ベクトル、 \mathbf{q} は n 次列ベクトルで、それぞれの成分は定数とする。 c も定数とする。 \mathbf{x} は n 次列ベクトルとする。 \mathbf{x} の n 変数実数値関数

$$g(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} - \mathbf{p}'\mathbf{x} - \mathbf{x}'\mathbf{q} + c$$

について、以下が成り立つ。

- (i). \mathbf{A} が正値定符号ならば、 $g(\mathbf{x})$ は \mathbf{x} について狭義凸関数である。
- (ii). \mathbf{A} が負値定符号ならば、 $g(\mathbf{x})$ は \mathbf{x} について狭義凹関数である。
- (iii). \mathbf{A} が半正値定符号ならば、 $g(\mathbf{x})$ は \mathbf{x} について凸関数である。
- (iv). \mathbf{A} が半負値定符号ならば、 $g(\mathbf{x})$ は \mathbf{x} について凹関数である。

」

この $g(\mathbf{x})$ は \mathbf{R}^n で定義される \mathbf{x} について関数であり、定義域で C^1 級である。偏導関数ベクトルを列ベクトル化すると、命題 40.4 や系 40.5 より、

$$g(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} - \mathbf{p}'\mathbf{x} - \mathbf{x}'\mathbf{q} + c \implies \frac{\partial}{\partial \mathbf{x}} g(\mathbf{x}) = 2\mathbf{A}\mathbf{x} - \mathbf{p} - \mathbf{q} \quad (40.9)$$

となる。

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「 n 変数実数値凹凸関数の制約なしの最大化・最小化」章の定理 8.1 から、上の系の (i) や (ii) の場合ならば「微分 = 0」で最大値・最小値が得られる。(i)(ii) の場合ならば \mathbf{A}^{-1} が存在し、上の (40.9) 式を解いて、最適点という意味で \mathbf{x} の代わりに \mathbf{x}^* とすると、

$$\mathbf{x}^* = \mathbf{A}^{-1}(\mathbf{p} + \mathbf{q})/2 \quad (40.10)$$

が成り立つ。 \mathbf{x}^* は n 項列ベクトルである。以上より、

定理 40.8 \mathbf{A} は n 次正方行列で、 \mathbf{p}' は n 次行ベクトル、 \mathbf{q} は n 次列ベクトルで、それぞれの成分は定数とする。実数 c も定数とする。 \mathbf{x} は n 次列ベクトルとして、 \mathbf{x} の n 変数実数値関数

$$g(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} - \mathbf{p}'\mathbf{x} - \mathbf{x}'\mathbf{q} + c \quad (40.11)$$

について、以下が成り立つ。

- (i). \mathbf{A} が正値定符号のとき、 $g(\mathbf{x})$ は (40.10) 式の \mathbf{x}^* で、唯一の最小値を持つ。
- (ii). \mathbf{A} が負値定符号のとき、 $g(\mathbf{x})$ は (40.10) 式の \mathbf{x}^* で、唯一の最大値を持つ。

」

\mathbf{A} が半正値定符号や半負値定符号の場合では、 \mathbf{A}^{-1} が存在しない場合があるので、上のようには書けない。

41 第 VII 部の Scilab 実習

ここでは VII 部の復習として Scilab の実習をまとめて行う。

まずはグラムシュミットの直交化についての復習である (定理 38.4 参照)。

まず Scilab のメインから、アプリケーション > Scinotes で Scinotes を開くこと。

例題 41.1 n 項列ベクトル $\text{vecV} (\neq \mathbf{0})$ が与えられたときに、それを基準化したベクトルを 1 列目とするような、 n 次直交行列を返す関数 $\text{myortho}(\text{vecV})$ を作る。

- (i). 次のコマンドを、Scinotes に打ち込んで (もしくは、PDF ファイルからコピーして)。適当な名前前で保存せよ。その後、Scinote での、「実行する > ファイルを実行」によって、そのファイルを実行させよ。

```
function myG=myortho(vecV)
    myeps=10^(-14);
    [rowV,colV]=size(vecV);
    if colV>=2 | rowV==1 then
        error('argument must be a column vector');
    end

    mxabs=max(abs(vecV));
    if mxabs<myeps then
        error('argument must be a non zero vector');
    end
    indexes=find(abs(vecV)==mxabs);
    istar=indexes(1);
    vecQ=vecV;
    if istar~=1 then
        vecQ([1 istar])=vecQ([istar 1]);
    end
    myG=eye(rowV,rowV);
    myG(:,1)=vecQ/sqrt(vecQ'*vecQ);

    for piv=2:rowV
        vecAi=myG(:,piv);
        Cs=myG(:,1:piv-1)'*vecAi;
        sumCG=[];
        for colj=1:piv-1
            sumCG=sumCG+Cs(colj)*myG(:,colj);
        end
        Fi=vecAi-sumCG;
```

```

    myG(:,piv)=Fi/sqrt(Fi'*Fi);
end

if istar~=1 then
    myG([1 istar],:)=myG([istar 1],:);
end
endfunction

```

- (ii). その後以下のようにメイン画面に打ち込んで、
- (a) ベクトル V を基準化したものが、 P の 1 列目と一致していること
 - (b) P が直交行列であること
- を確認せよ。

```

V=[1:4]'
V1=V/sqrt(V'*V)
P=myortho(V)
P'*P

```

」

この直交行列を返すプログラムをよく見ると、「 \uparrow 」は随伴行列を返すので、実は引数のベクトル $\text{vec}V$ が複素列ベクトルの場合には、ユニタリー行列を返すこととなる。

例題 41.2 前例題に続けて、以下のようにメイン画面に打ち込んで、

- (i). ベクトル V を基準化したものが、 P の 1 列目と一致していること
- (ii). P が複素行列であること。かつユニタリー行列であること

を確認せよ。

```

V=[1+%i;2;3]
P=myortho(V)
P'*P

```

」

では次に三角化についてみてみよう。定理 38.6 についての復習である。 n 次実正方行列 $\text{mat}A$ が与えられたときに、それを三角化できるような直交行列 $\text{out}Q$ を返すものとする。プログラムについては、Scilab の `spec` 関数を利用し、(38.8) 式をどんどん繰り返していくというものとする。従って無駄が多いプログラムなのではあるが、定理 38.6 の証明のよい復習となるであろう。

例題 41.3 上のような値を返す関数 `mytriangle(matA)` というプログラムを作る。

- (i). 先の例題 41.1 で作った `myortho(vecV)` を書いたファイルを、Scinotes で開き、その `endfunction` の下側に、以下のコードを打ち込み (もしくは、PDF ファイルからコピーして)。適当な名前前で保存せよ。その後、Scinote での、「実行する > ファイルを実行」によって、そのファイルを実行させよ。

```

function outQ=mytriangle(matA)
myeps=10^(-14);
[nA,colA]=size(matA);
if colA~=nA | nA==1 then
    error('argument must be squared matrix');
end

outQ=eye(nA,nA);
nowA=matA;
for piv=1:nA-1
    [specvec,Lambda]=spec(nowA);
    vecq=specvec(:,1);
    myP=myortho(vecq);
    if piv==1 then
        outQ=myP;
    else
        outQ=outQ*[eye(piv-1,piv-1) zeros(piv-1,nA-piv+1);zeros(nA-piv+1,piv-1) myP];
    end
    nowA=myP(:,2:$)'+nowA*myP(:,2:$);
end
endfunction

```

- (ii). その後 2 つの行列について、以下のようにメイン画面に打ち込んで、行列 A が直交行列 P によって、上三角行列に三角化されている様子を見ること。

```

A=[1:4; 5:8; 9:12;13:16]
P=mytriangle(A)
P'*P
inv(P)*A*P

```

$inv(P)*A*P$ の部分は数値計算誤差の関係上、下三角部分が「0」という数字にはなかなかならない。しかしよく見ると D-17 などが出ているので、「0」と考えても差し支えはない。前例題と同様に、もし行列 A が複素行列であれば、`mytriangle` で返される行列は、ユニタリー行列である。

例題 41.4 前例題に引き続いて、以下のようにメイン画面に打ち込んで、 P が複素行列であることを。かつユニタリー行列であることを確認し、 B が P によって、三角化されることを確認せよ。

```

B=rand(3,3)+%i*rand(3,3)
P=mytriangle(B)
P'*P
inv(P)*B*P

```


証明はしていないが、任意の n 次正方複素行列は、適当なユニタリ行列で三角化可能であることを示している。 ┘

また `rand` 関数は、第 3 引数に何も書かない場合は、 $(0, 1)$ の一様分布に従う乱数 (疑似乱数) が得られる。従って `B=rand(3,3)+%i*rand(3,3)` と打ち込むたびに、 B は毎回異なる値が出てくる。しかしどの場合の B でも、それに応じた適当な P があって、 $\text{inv}(P)*B*P$ は上三角行列となることを確かめるとよい。

第 VIII 部

計量経済学・確率論・統計学で出会うもの

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の第 IV 部を読んだ後の続きとなる。

簡単化のため確率変数はすべて、連続形確率変数としている。

42 回帰分析の回帰係数の導出

n, k はともに正整数で $k \leq n$ である。 n サンプル k 変数回帰分析では、 $i = 1, \dots, n$ について、 $(y_i, x_{1i}, x_{2i}, \dots, x_{ki})$ が既知の値で、これについて、

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \epsilon_i, \quad (i = 1, 2, \dots, n)$$

とモデル化する。 β_1, \dots, β_k は未知で、 ϵ_i も未知として誤差項と呼ぶ。このとき β_1, \dots, β_k の推定量について、誤差項の 2 乗和 $Q = \sum_{i=1}^n \epsilon_i^2$ を最小化するようなものを選ぶ方法を最小 2 乗法という。

実際のデータが与えられたときには、PC で簡単に値が得られるが、この最小 2 乗法による β_1, \dots, β_k の推定量の求め方を、ここで行列を用いて説明する。

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad (42.1)$$

とする。OLS モデルは、

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (\text{ただし } \mathbf{X} \text{ のランクは } k) \quad (42.2)$$

と書ける。ランクの条件があるのは、そうでなければ β_1, \dots, β_k は一意に定まらないためである。

式より $\boldsymbol{\epsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ である。誤差の 2 乗和 Q は $\boldsymbol{\epsilon}'\boldsymbol{\epsilon}$ と書けるので、変形して、

$$\begin{aligned} Q &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= (\mathbf{y}' - \boldsymbol{\beta}'\mathbf{X}')(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \end{aligned} \quad (42.3)$$

$$\begin{aligned} &= \boldsymbol{\beta}'(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \mathbf{y}'\mathbf{y} \\ & \quad (\text{第 3 項は実数の第 2 項を転置したものなので第 2 項と同じ}) \\ &= \boldsymbol{\beta}'(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \mathbf{y}'\mathbf{y} \end{aligned} \quad (42.4)$$

これを最小化するような $\boldsymbol{\beta}$ を求めるのが最小 2 乗法なのである。

n 行 k 列の行列 \mathbf{X} のランクが k である (よって当然 ($k \leq n$)) ことから、定理 39.8 より $(\mathbf{X}'\mathbf{X})$ は正値定符号である。定理 40.8 より、

$$\boldsymbol{\beta}^* = (\mathbf{X}'\mathbf{X})^{-1}(2\mathbf{X}'\mathbf{y})/2 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

が、誤差の 2 乗和を最小化する β_1, \dots, β_k である。これを回帰ベクトル $\boldsymbol{\beta}$ の最小 2 乗推定量といい、一般に $\hat{\boldsymbol{\beta}}$ で表す。

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (42.5)$$

上で、定理 40.8 から $\hat{\beta}$ を一気に導出したが、これを順に書いていくことにしよう。まず (42.4) 式を β の偏導関数ベクトルが $\mathbf{0}$ となる β を、 $\hat{\beta}$ とした。命題 40.4 と系 40.5 から、

$$\frac{\partial Q}{\partial \beta'} = 2(\mathbf{X}'\mathbf{X})\hat{\beta} - 2\mathbf{X}'\mathbf{y} = \mathbf{0} \quad (42.6)$$

上の等式を 2 で割って、

$$(\mathbf{X}'\mathbf{X})\hat{\beta} - \mathbf{X}'\mathbf{y} = \mathbf{0} \quad (42.7)$$

左辺第 2 項を右辺に移項して、両辺 $(\mathbf{X}'\mathbf{X})^{-1}$ を左から掛けて、(42.5) 式が導出されたのである。

以下で、簡単な実習しよう。

例題 42.1 以下のようなデータセットがあるとする。5 サンプル 3 変数モデル

i	y	x1	x2	x3
1	4	1	3	1
2	5	1	4	2
3	2	1	2	3
4	4	1	1	4
5	1	1	-1	10

で、 y を x_1, x_2, x_3 で回帰したときの回帰係数を Scilab で求めたいので、以下のコードを Scinote に書いて、実行すると良い。 y と X を表示して、`bhat` という名前で、回帰係数が出てくる。上から順に $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ である。 ┘

(解答)

```
y=[4 5 2 4 1]';
X0=[3 1;4 2;2 3;1 4;-1 10];
X=[ones(5,1) X0];
disp([y X], "y & X=")
b=inv(X'*X)*X'*y;
disp(b, "bhat=")
```

結果として推定式は以下ようになる。

$$\hat{y}_i = 2.7x_1 + 0.52x_2 - 0.099x_3$$

(解答終)

$\hat{\beta}$ を推定した後は、 \mathbf{y} の推定ベクトル

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$$

と計算して、次に誤差 ϵ の推定ベクトルを残差ベクトルと定義し、

$$\hat{\epsilon} = \mathbf{y} - \hat{\mathbf{y}}$$

とする。

命題 42.2 上の説明変数行列 \mathbf{X} 、 \mathbf{y} の推定ベクトル $\hat{\mathbf{y}}$ 、残差ベクトル $\hat{\epsilon}$ について、以下が成り立つ。

- (i). $\hat{\boldsymbol{\epsilon}}$ と, \mathbf{X} の各列は直交する. すなわち, $\mathbf{X}'\hat{\boldsymbol{\epsilon}} = \mathbf{0}$
(ii). $\hat{\boldsymbol{\epsilon}}$ と $\hat{\mathbf{y}}$ は直交する. すなわち, $\hat{\mathbf{y}}'\hat{\boldsymbol{\epsilon}} = 0$

証明 (i) の証明

$$\begin{aligned}\mathbf{X}'\hat{\boldsymbol{\epsilon}} &= \mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \\ & \text{((42.7) 式より)} = \mathbf{0}\end{aligned}$$

(ii) の証明 $\hat{\mathbf{y}}'\hat{\boldsymbol{\epsilon}} = \hat{\boldsymbol{\beta}}'\mathbf{X}'\hat{\boldsymbol{\epsilon}}$ と (i) より, 命題が成り立つ.

□

この命題の (i) は, 「残差ベクトルは説明変数行列と直交する」といい, (ii) は「残差ベクトルは \mathbf{y} の推定ベクトル (\mathbf{y} の理論値ベクトル) と直交する」という.

\mathbf{X} のある列が i のときには, 「総変動=回帰変動+残差変動」が成り立つ. その意味は \mathbf{y} の平均を \bar{y} で表したときに, 成分で書くと,

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n \hat{\epsilon}_i^2$$

である. なお一般に, 統計学の用語で変動というのは, 平均からの偏差の 2 乗を表す. 上の式を行列で書いたものが以下の (ii) になる.

命題 42.3 OLS モデルの (42.2) の \mathbf{X} について, ある列が $a\mathbf{i}$ の場合 (ただし $a \neq 0$), 以下が成り立つ.

- (i). $\hat{\mathbf{y}}$ の平均は \mathbf{y} の平均と等しい. つまり $\mathbf{i}'\hat{\mathbf{y}}/n = \mathbf{i}'\mathbf{y}/n$
(ii). $(\mathbf{y} - \mathbf{i}\bar{y})'(\mathbf{y} - \mathbf{i}\bar{y}) = (\hat{\mathbf{y}} - \mathbf{i}\bar{y})'(\hat{\mathbf{y}} - \mathbf{i}\bar{y}) - \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}$

□

証明 前命題 (ii) より, $\mathbf{i}'\hat{\boldsymbol{\epsilon}} = 0$ が成り立つ. また $a = 1$ とおいても一般性は失わないので, そう仮定する.

(i) の証明: $\mathbf{i}'\mathbf{y}/n = \mathbf{i}'(\hat{\mathbf{y}} + \hat{\boldsymbol{\epsilon}})/n = \mathbf{i}'\hat{\mathbf{y}}/n$

(ii) の証明:

$$\begin{aligned}(\mathbf{y} - \mathbf{i}\bar{y})'(\mathbf{y} - \mathbf{i}\bar{y}) &= (\hat{\mathbf{y}} + \hat{\boldsymbol{\epsilon}} - \mathbf{i}\bar{y})'(\hat{\mathbf{y}} + \hat{\boldsymbol{\epsilon}} - \mathbf{i}\bar{y}) \\ &= (\hat{\mathbf{y}} - \mathbf{i}\bar{y})'(\hat{\mathbf{y}} - \mathbf{i}\bar{y}) - (\hat{\mathbf{y}} - \mathbf{i}\bar{y})'\hat{\boldsymbol{\epsilon}} - \hat{\boldsymbol{\epsilon}}'(\hat{\mathbf{y}} - \mathbf{i}\bar{y}) + \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}} \\ \text{前命題 (ii) と } \mathbf{i}'\hat{\boldsymbol{\epsilon}} = 0 \text{ より} &= (\hat{\mathbf{y}} - \mathbf{i}\bar{y})'(\hat{\mathbf{y}} - \mathbf{i}\bar{y}) - 0 - 0 + \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}\end{aligned}$$

□

この (ii) の性質があるので, 回帰分析ではできるだけ定数項を含める推計が望ましい.

次は制約条件付きの, 最小二乗法について説明する. (42.1) 式と (42.2) 式の OLS モデルと同じく n サンプル k 変数モデルの回帰分析について考える. 例えば今 $k = 5$ としてみよう.

係数に関する検定などで,

$$\beta_3 + \beta_4 = 1, \quad \beta_5 = 0$$

のように, β_1, \dots, β_5 の推定を, 線形の制約式を付けたときの, 最小 2 乗推定値として求めることがある. 上の場合, 制約式は

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

となる.

$$\mathbf{A}' = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{q} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

として, $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$, とすると, 制約式は

$$\mathbf{A}'\boldsymbol{\beta} = \mathbf{q} \quad (42.8)$$

と書ける. r は $r \leq k$ を満たす正整数として, \mathbf{A} が k 行 r 列で $\text{rank}(\mathbf{A}) = r$ の Full Column Rank で, \mathbf{q} が r 行 1 列として, この行列式が線形制約を表す.

そこで, (42.1)・(42.2) の OLS モデルについて, k 行 r 列の行列 \mathbf{A} (Full Column rank つまりランクは r) と r 項列ベクトル \mathbf{q} で, (42.2) を制約として, 誤差項の 2 乗和を最小化する $\boldsymbol{\beta}$ を求めたい. これは

$$\min_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \quad \text{s.t. } \mathbf{A}'\boldsymbol{\beta} - \mathbf{q} = \mathbf{0}$$

の解である. 目的関数を展開すると,

$$\boldsymbol{\beta}'(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \mathbf{y}'\mathbf{y}$$

であり, 系 40.7 よりこれは $\boldsymbol{\beta}$ について凸関数. $\mathbf{q} - \mathbf{A}\boldsymbol{\beta}$ の r 個の各成分は, $\boldsymbol{\beta}$ に関する線形関数である.

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「 n 変数実数値凹凸関数の制約付き最大化: ラグランジュ乗数法」章を参考にして, $2\boldsymbol{\lambda}$ は, ラグランジュ定数を表す r 項列ベクトル変数として, ラグランジュ関数を作ると,

$$L(\boldsymbol{\beta}, \boldsymbol{\lambda}) = \boldsymbol{\beta}'(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} - 2\boldsymbol{y}'\mathbf{X}\boldsymbol{\beta} + \mathbf{y}'\mathbf{y} + 2\boldsymbol{\lambda}'(\mathbf{A}'\boldsymbol{\beta} - \mathbf{q})$$

である. 目的関数も制約式も凸関数なので, 一般的なラグランジュ条件

$$(i). \frac{\partial L}{\partial \boldsymbol{\beta}'} = \begin{pmatrix} 0 & \dots & 0 \end{pmatrix}'$$

$$(ii). \mathbf{A}'\boldsymbol{\beta} = \mathbf{q}$$

(iii). $2\boldsymbol{\lambda}$ の成分が全て 0 以上

である $\boldsymbol{\beta}$ が存在すれば, それが最小解となるのだが, さらに今の等式制約式の場合だと (iii) が成り立たなくとも良いことが知られているので, (iii) は無視する.

命題 40.4 と系 40.4 より,

$$\frac{\partial L}{\partial \boldsymbol{\beta}'}(\boldsymbol{\beta}, \boldsymbol{\lambda}) = 2(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} - 2\mathbf{X}'\mathbf{y} + 2\mathbf{A}\boldsymbol{\lambda} = \mathbf{0}$$

2で割って、左辺第2項と第3項を右辺に移項して

$$(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} = \mathbf{X}'\mathbf{y} - \mathbf{A}\boldsymbol{\lambda}$$

$(\mathbf{X}'\mathbf{X})$ は正値定符号で正則なので、両辺の左から $(\mathbf{X}'\mathbf{X})^{-1}$ を掛けて、

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}\boldsymbol{\lambda} \quad (42.9)$$

両辺に \mathbf{A}' を左から掛けると、

$$\mathbf{q} = \mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - \mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}\boldsymbol{\lambda}$$

\mathbf{X} の仮定より定理 39.8 から $(\mathbf{X}'\mathbf{X})^{-1}$ が正値定符号である。 \mathbf{A} が Full Column Rank であることから、定理 39.11 より $\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}$ も正値定符号であり、正則である。そこで右辺第2項を左辺に移項し、左辺を移項した後に、左から $[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}$ を掛けて、

$$\boldsymbol{\lambda} = [\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - \mathbf{q}) \quad (42.10)$$

この式を (42.9) 式に代入して、

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - \mathbf{q}) \quad (42.11)$$

この $\boldsymbol{\beta}$ が最適解である。よってこれを $\tilde{\boldsymbol{\beta}}$ で表し、また OLS 推定量の $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ を $\hat{\boldsymbol{\beta}}$ で表すと、

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})$$

OLS の残差ベクトルを $\hat{\boldsymbol{\epsilon}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ で表したのと同様に、制約条件付き OLS の残差ベクトルを $\tilde{\boldsymbol{\epsilon}} = \mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}$ とする。後者の2乗和と前者の2乗和の差をとることが多いので、これを計算しておく。計算の簡単化のため上の式で、 $\mathbf{B} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}$ とする。つまり、

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} - \mathbf{B}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})$$

となる。よって、

$$\tilde{\boldsymbol{\epsilon}} = \mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}\mathbf{B}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q}) = \hat{\boldsymbol{\epsilon}} + \mathbf{X}\mathbf{B}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})$$

$\tilde{\boldsymbol{\epsilon}}$ の各項の2乗和は、

$$\begin{aligned} \tilde{\boldsymbol{\epsilon}}'\tilde{\boldsymbol{\epsilon}} &= [\hat{\boldsymbol{\epsilon}} + \mathbf{X}\mathbf{B}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})]'[\hat{\boldsymbol{\epsilon}} + \mathbf{X}\mathbf{B}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})] \\ &= \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}} + 2\hat{\boldsymbol{\epsilon}}'\mathbf{X}\mathbf{B}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q}) + (\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})'\mathbf{B}'\mathbf{X}'\mathbf{X}\mathbf{B}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q}) \\ &\quad (\text{命題 42.2(i) から第2項は0で第3項は } \mathbf{X}'\mathbf{X}\mathbf{B} = \mathbf{A}[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1} \text{ より}) \\ &= \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}} + (\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})'\mathbf{B}'\mathbf{A}[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q}) \\ &\quad (\mathbf{A}'\mathbf{B} = \mathbf{I} \text{ より}) = \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}} + (\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})'[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q}) \end{aligned}$$

従って

$$\tilde{\boldsymbol{\epsilon}}'\tilde{\boldsymbol{\epsilon}} - \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}} = (\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q})'[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}'\hat{\boldsymbol{\beta}} - \mathbf{q}) \quad (42.12)$$

ここまでの、誤算の2乗和を最小化するという計算である。実は誤差項ベクトル $\boldsymbol{\epsilon}$ の各項が確率変数を仮定しているのだが、これについては次章以降で、やはり行列を元にして説明する。

43 確率変数の期待値行列・共分散行列と線形変換 (正規分布)

行列 \mathbf{A} の各成分が確率変数のとき, \mathbf{A} は確率変数行列という. 行列が 1 列の場合は, 確率変数列ベクトル, もしくは確率変数ベクトルという.

定義 43.1 $m \times n$ の確率変数行列 \mathbf{X} について, それぞれの成分の期待値を取った行列を,

$$\begin{pmatrix} E[X_{11}] & \cdots & E[X_{mn}] \\ \vdots & \ddots & \vdots \\ E[X_{m1}] & \cdots & E[X_{mn}] \end{pmatrix} \quad (43.1)$$

\mathbf{X} の期待値行列といい, $E[\mathbf{X}]$ や $E(\mathbf{X})$ などで表す. ┘

(43.1) 式の行列の各成分が全て実数で存在するときに, $E[\mathbf{X}]$ は存在するという.

X_1, X_2 が確率変数のとき, 定数 a, b について $E[aX_1 + bX_2] = aE[X_1] + bE[X_2]$ が成り立つことなどから, 以下が言える.

命題 43.2 $m \times n$ の, 2 つの確率変数行列 \mathbf{X}, \mathbf{Y} があり, $E[\mathbf{X}], E[\mathbf{Y}]$ が存在するとき,

(i). $E[\mathbf{X} + \mathbf{Y}] = E[\mathbf{X}] + E[\mathbf{Y}]$

(ii). $k \times m$ の定数行列 \mathbf{A} と, $n \times \ell$ の定数行列 \mathbf{B} について, \mathbf{AXB} は $k \times \ell$ の確率変数行列であり,

$$E[\mathbf{AXB}] = \mathbf{A}E[\mathbf{X}]\mathbf{B}$$
┘

次に 2 変数の共分散の定義 $E[(X - E[X])(Y - E[Y])]$ を拡張して, 以下を定義する.

定義 43.3 \mathbf{X} が m 項列ベクトル, \mathbf{Y} が n 項列ベクトルとともに, 確率変数ベクトルとする. このとき行列

$$\begin{aligned} E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{Y} - E[\mathbf{Y}])'] &= E \left[\begin{pmatrix} X_1 - E[X] \\ \vdots \\ X_m - E[X_m] \end{pmatrix} (Y_1 - E[Y] \quad \cdots \quad Y_n - E[Y_n]) \right] \\ &= \begin{pmatrix} E[(X_1 - E[\mathbf{X}])(Y_1 - E[\mathbf{Y}])] & \cdots & E[(X_1 - E[\mathbf{X}])(Y_n - E[\mathbf{Y}])] \\ \vdots & \ddots & \vdots \\ E[(X_m - E[\mathbf{X}])(Y_1 - E[\mathbf{Y}])] & \cdots & E[(X_m - E[\mathbf{X}])(Y_n - E[\mathbf{Y}])] \end{pmatrix} \end{aligned}$$

を, \mathbf{X}, \mathbf{Y} の共分散行列といい, $Cov(\mathbf{X}, \mathbf{Y})$ で表す. ┘

共分散公式から, 上の \mathbf{X}, \mathbf{Y} については

$$Cov(\mathbf{X}, \mathbf{Y}) = E[\mathbf{XY}'] - E[\mathbf{X}]E[\mathbf{Y}']$$

が成り立つ. この場合, 特に $\mathbf{Y} = \mathbf{X}$ については, 以下のように定義する.

定義 43.4 \mathbf{X} が n 項列ベクトルの確率変数ベクトルとする. このとき,

$$E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])'] = \begin{pmatrix} E[(x_1 - E[\mathbf{X}])(y_1 - E[\mathbf{X}])] & \cdots & E[(x_1 - E[\mathbf{X}])(x_n - E[\mathbf{X}])] \\ \vdots & \ddots & \vdots \\ E[(x_n - E[\mathbf{X}])(y_1 - E[\mathbf{X}])] & \cdots & E[(x_n - E[\mathbf{X}])(x_n - E[\mathbf{X}])] \end{pmatrix}$$

を \mathbf{X} の分散・共分散行列といい, $Var[\mathbf{X}]$ や $V[\mathbf{X}]$ などです。 」

共分散公式からやはり, この \mathbf{X} については, 以下が成り立つ。

$$Var[\mathbf{X}] = E[\mathbf{X}\mathbf{X}'] - E[\mathbf{X}]E[\mathbf{X}']$$

さて, 期待値ベクトルと分散・共分散行列が存在するときには, 以下のように記号で簡単に書く。

定義 43.5 \mathbf{X} が n 項列ベクトルの確率変数ベクトルとする。 $\boldsymbol{\mu}$ は n 項列ベクトルの定数ベクトルで, \mathbf{G} は n 次半正値定符号行列の定数行列とする。 $E[\mathbf{X}] = \boldsymbol{\mu}, Var[\mathbf{X}] = \mathbf{G}$ が成り立つことを,

$$\mathbf{X} \sim (\boldsymbol{\mu}, \mathbf{G})$$

で表す。 」

統計学では \mathbf{X} の線形変換がよく出てくる。

命題 43.6 \mathbf{X} は n 項確率変数ベクトルとする。 $\boldsymbol{\mu}, \mathbf{G}, \mathbf{A}, \mathbf{q}$ の各成分は定数で, それぞれ n 項列ベクトル, n 次半正値定符号行列, $n \times m$ 行列, m 項列ベクトルとする。 m 項確率変数ベクトル $\mathbf{A}'\mathbf{x} - \mathbf{q}$ について,

$$\mathbf{X} \sim (\boldsymbol{\mu}, \mathbf{G}) \implies \mathbf{A}'\mathbf{X} - \mathbf{q} \sim (\mathbf{A}'\boldsymbol{\mu} - \mathbf{q}, \mathbf{A}'\mathbf{G}\mathbf{A})$$

注意: 特に \mathbf{A} が列ベクトルで \mathbf{q} がスカラーのときは, それぞれを \mathbf{a}, q で表して,

$$\mathbf{a}'\mathbf{X} - q \sim (\mathbf{a}'\boldsymbol{\mu} - q, \mathbf{a}'\mathbf{G}\mathbf{a})$$

と書く。 (注意終)

証明 $E[\mathbf{A}'\mathbf{X} - \mathbf{q}] = \mathbf{A}'E[\mathbf{X}] - \mathbf{q}$ となることは簡単なので, 分散・共分散行列のみ示す。

$$\begin{aligned} Var(\mathbf{A}'\mathbf{X} - \mathbf{q}) &= E[(\mathbf{A}'\mathbf{X} - \mathbf{q} - E[\mathbf{A}'\mathbf{X} - \mathbf{q}])(\mathbf{A}'\mathbf{X} - \mathbf{q} - E[\mathbf{A}'\mathbf{X} - \mathbf{q}])'] \\ &= E[(\mathbf{A}'(\mathbf{X} - E[\mathbf{X}])(\mathbf{A}'(\mathbf{X} - E[\mathbf{X}]))'] \\ &= E[\mathbf{A}'(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])'\mathbf{A}] \\ &= \mathbf{A}'E[\mathbf{A}(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])'\mathbf{A}] \\ &= \mathbf{A}'\mathbf{G}\mathbf{A} \end{aligned}$$

□

この命題では, 確率変数ベクトルの分布には言及していない。期待値と分散がわかっているならば, 必ずこの様になるということである。

例えば X_1 が標準正規分布に従い, X_2 が自由度 4 のカイ 2 乗分布に従い, X_1, X_2 が無相関とする。このとき

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \mathbf{x} \sim \left(\begin{pmatrix} 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 8 \end{pmatrix} \right)$$

である。命題より a_1, a_2 が定数のとき,

$$a_1X_1 + a_2X_2 \sim (4a_2, a_1^2 + 8a_2^2)$$

となることがわかる。ただし、 $a_1X_1 + a_2X_2$ がどのような分布になるかは、簡単にはわからないのである。
ただし、ともに正規分布ならば簡単にわかる。

まず、定義 43.5 と同様に正規分布について以下のように定義する。記号で簡単に書きたいので、以下の記号を決めておく。

定義 43.7 n 項確率変数列ベクトル \mathbf{X} の成分 X_i がそれぞれ正規分布に従うとする。このとき、以下のよう
に表す。

$$\mathbf{X} \sim N(E[\mathbf{X}], \text{Var}[\mathbf{X}])$$

\mathbf{X} が 1 変数のときと同様の書き方である。共分散行列は $\text{Var}[\mathbf{X}]$ は、必ず半正値定符号である。 $\text{Var}[\mathbf{X}]$ が正値定符号であれば、 \mathbf{X} の同時密度関数が定義され、 \mathbf{X} は非退化の正規分布の分布であるという。正値定符号ではない場合は、 \mathbf{X} は退化した正規分布という。退化していても、定義 43.7 のように書くことができる。

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「確率・統計・計量経済学で出てくる積分 (2)」章で、正規分布は線形結合に閉じていることを示した。これより、以下が簡単に言える。

補題 43.8 $\sigma_1 > 0, \sigma_2 > 0, |\rho| \leq 1$ とする。 X_1 が平均 μ_1 分散 σ_1^2 の正規分布、 X_2 が平均 μ_2 分散 σ_2^2 の正規分布にそれぞれ従い、共分散は $\text{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2$ とする。このとき、 $a_1X_1 + a_2X_2$ は、平均 $a_1\mu_1 + a_2\mu_2$ 、分散 $a_1^2\sigma_1^2 + 2\rho a_1a_2\sigma_1\sigma_2 + a_2^2\sigma_2^2$ の正規分布に従う。

証明

$|\rho| < 1$ の場合は、

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「確率・統計・計量経済学で出てくる積分 (2)」章の一番最後で示してある。

$|\rho| = 1$ の場合を示す。2つの確率変数が相関係数の絶対値が 1 となることは、その 2つの確率変数が完全な線形関係にあることと同値である。よって X_1 と X_2 については、ある b_1, b_2, c があって、

$$b_1X_1 - b_2X_2 = c$$

が成り立つ。 b_1, b_2 の一方が 0 の場合は X_1, X_2 の一方が定数となってしまう不適当なので b_1, b_2 ともに非 0 である。

期待値を取ると $b_1\mu_1 - b_2\mu_2 = c$ 。従って

$$b_1(X_1 - \mu_1) - b_2(X_2 - \mu_2) = 0$$

次に上の両辺に $X_1 - \mu_1$ を掛けて期待値を取ると、

$$b_1\sigma_1^2 - b_2\rho\sigma_1\sigma_2 = 0$$

$b = b_1/b_2$ として、 $b = \rho\sigma_2/\sigma_1$ が成り立つ。

$b_1(X_1 - \mu_1) - b_2(X_2 - \mu_2) = 0$ より, $X_2 = b(X_1 - \mu_1) + \mu_2$ なので,

$$\begin{aligned} a_1X_1 + a_2X_2 &= a_1X_1 + a_2b(X_1 - \mu_1) + a_2\mu_2 \\ &= (a_1 + a_2b)X_1 - a_2b\mu_1 + a_2\mu_2 \end{aligned}$$

右辺は $X_1 \sim N(\mu, \sigma^2)$ のアフィン変換なので, 平均 $a_1\mu_1 + a_2\mu_2$, 分散 $(a_1 + a_2b)^2\sigma_1^2$ の正規分布に従う. この分散の b に $b = \rho\sigma_2/\sigma_1$ を代入して計算すると

$$\begin{aligned} (a_1 + a_2b)^2\sigma_1^2 &= (a_1 + a_2\rho\sigma_2/\sigma_1)^2\sigma_1^2 \\ &= a_1\sigma_1^2 + 2\rho a_1a_2\sigma_1\sigma_2 + a_2\rho^2\sigma_2^2 \end{aligned}$$

なので, $a_1X_1 + a_2X_2 \sim N(a_1\mu_1 + a_2\mu_2, a_1\sigma_1^2 + 2\rho a_1a_2\sigma_1\sigma_2 + a_2\rho^2\sigma_2^2)$ が言える.

□

この補題を定義 43.7 で書くと,

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

としたとき,

$$\mathbf{X} \sim N(\mathbf{E}[\mathbf{X}], \text{Var}[\mathbf{X}]) \implies \mathbf{a}'\mathbf{X} \sim N(\mathbf{a}'\mathbf{E}[\mathbf{X}], \mathbf{a}'\text{Var}[\mathbf{X}]\mathbf{a})$$

である. 前補題を繰り返すことにより, 上の式が \mathbf{X} が n 次列ベクトルの場合でも成り立つ. \mathbf{a}' の代わりに定数行列 \mathbf{A} を考えて定数を加えると, 命題 43.6 から, 以下が言える.

定理 43.9 \mathbf{X} は n 次確率変数ベクトルとし, \mathbf{A} は $n \times m$ の定数行列, \mathbf{q} は m 次定数列ベクトルとする. このとき,

$$\mathbf{X} \sim N(\mathbf{E}[\mathbf{X}], \text{Var}[\mathbf{X}]) \implies \mathbf{A}'\mathbf{X} - \mathbf{q} \sim N(\mathbf{A}'\mathbf{E}[\mathbf{X}] - \mathbf{q}, \mathbf{A}'\text{Var}[\mathbf{X}]\mathbf{A})$$

□

この定理と, 命題 43.6 を比較すると良い. 命題では平均や, 分散・共分散行列の値は分かるが, 分布はわからないと説明した. しかし, 正規分布の場合であれば, アフィン変換 (線形変換に定数項をつけた変換) でも正規分布となっているのである. つまり, 正規分布はアフィン変換について閉じているのである.

系 43.8 の証明から $\rho < 1$ と $\rho = \pm 1$ の間に差があることが分かるであろう. これについて, 次章で述べる.

44 同時密度関数・線形変換・同時正規分布の線形変換

さて, n 変数同時密度関数の定義を済ませておこう. 2 変数の場合は,

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「確率・統計・計量経済学で出てくる積分 (2)」章の最初に示してある.

定義 44.1 (n 変数連続形確率変数の定義)

- (i). X_1, \dots, X_n が n 変数連続形確率変数であるとは, \mathbf{R}^n で定義された $f(x_1, \dots, x_n) \geq 0$ があって, 任意の領域 $A \subset \mathbf{R}^n$ について, (X_1, \dots, X_n) が A の範囲の値をとる確率が, 重積分で以下のように

$$\Pr((X_1, \dots, X_n) \in A) = \int \cdots \int_A f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

と表されるときである.

- (ii). (i) での $f(x_1, \dots, x_n)$ を X_1, \dots, X_n の同時密度関数という.
 (iii). (i) での $f(x_1, \dots, x_n)$ から導出される

$$g_{X_i}(x) = \int \cdots \int_{\mathbf{R}^{n-1} \cap x_i=x} f(x_1, \dots, x_n) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n,$$

を X_i の周辺密度関数と呼ぶ.

- (iv). 任意の $i, j (i \neq j)$ で, 任意の実数 a, b について,

$$\Pr(X_i \leq a \cap X_j \leq b) = \Pr(X_i \leq a) \Pr(X_j \leq b)$$

が成り立つ時, X_1, \dots, X_n は各々独立であるという.

┘

以下は簡単に言える.

命題 44.2 連続形確率変数 $X_i (i = 1, \dots, n)$ について,

- (i). 各 X_i がそれぞれ密度関数 $g_{X_i}(x_i)$ を持ち, かつ X_i が各々独立のとき, X_1, \dots, X_n の同時密度関数は,

$$\prod_{i=1}^n g_{X_i}(x_i)$$

である.

- (ii). X_1, \dots, X_n の同時密度関数 $f(x_1, \dots, x_n)$ で, 各 X_i の周辺密度関数が $g_{X_i}(x_i)$ であり,

$$f(x_1, \dots, x_n) = \prod_{i=1}^n g_{X_i}(x_i)$$

が任意の (x_1, \dots, x_n) で成り立つならば, 各 X_i は各々独立である.

┘

確率変数を線形変換した場合の同時密度関数については, 2 変数実数値連続形確率変数の変数変換と, n 変数の微分積分の変数変換から以下が言える.

定理 44.3 X_1, \dots, X_n の同時密度関数を, $f(\mathbf{x})$ とする. ただし $\mathbf{x} = {}^t(x_1, \dots, x_n)$ である. n 次定数正平方行列 \mathbf{A} が正則で, かつ

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \mathbf{A} \begin{pmatrix} V_1 \\ \vdots \\ V_n \end{pmatrix}$$

が成り立つとき, V_1, \dots, V_n の同時密度関数は,

$$f(\mathbf{A}\mathbf{v}) |\det(\mathbf{A})|$$

で表される。ただし $\mathbf{v} = {}^t(v_1, \dots, v_n)$ 」

この定理と、前章の正規分布の定理 43.9 から、非退化の正規分布についての同時密度関数の存在を示す。

補題 44.4 $i = 1, \dots, n$ について、確率変数 Z_i は、それぞれ独立に標準正規分布に従う。このとき、 Z_1, \dots, Z_n の同時密度関数は $f_{\mathbf{Z}}(z_1, \dots, z_n)$ は

$$f_{\mathbf{Z}}(z_1, \dots, z_n) = \frac{1}{(2\pi)^{n/2}} \exp \left[-\frac{z_1^2 + \dots + z_n^2}{2} \right]$$
」

前章の分散・共分散行列から、以下が言える。

命題 44.5 $\boldsymbol{\mu}$ は n 次定数列ベクトル、 $\boldsymbol{\Sigma}$ は n 次正方定数行列とする。 n 次確率変数ベクトル \mathbf{X} が $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ を満たし、 $\boldsymbol{\Sigma}$ が正値定符号とする。このとき、 X_1, \dots, X_n の同時密度関数は、

$$\frac{1}{(2\pi)^{n/2} |\det(\boldsymbol{\Sigma})|^{1/2}} \exp \left[-\frac{(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}{2} \right] \quad (44.1)$$

ただし、 $\mathbf{x} = {}^t(x_1, \dots, x_n)$ とする。 」

証明

$\boldsymbol{\Sigma}$ は対称行列なので、固有値 $\lambda_1, \dots, \lambda_n$ を対角成分とする対角行列 $\boldsymbol{\Lambda}$ と、ある直交行列 \mathbf{P} によって

$$\boldsymbol{\Sigma} = \mathbf{P} \boldsymbol{\Lambda} \mathbf{P}'$$

と表すことができる。さらに $\boldsymbol{\Sigma}$ が正値定符号なので、 $\boldsymbol{\Lambda}$ の対角成分は全て正である。従って対角成分が $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ で全て正の対角行列 $\boldsymbol{\Lambda}^{1/2}$ を持ち、

$$\boldsymbol{\Sigma} = (\mathbf{P} \boldsymbol{\Lambda}^{1/2}) (\mathbf{P} \boldsymbol{\Lambda}^{1/2})'$$

で表される。

Z_1, \dots, Z_n が補題 44.4 の通りとする。 $\mathbf{Z} = {}^t(Z_1 \ \dots \ Z_n)$ とすると、

$$\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I})$$

で表される。 $(\mathbf{0}$ は成分全て 0 の n 次列ベクトルで \mathbf{I} は単位行列) このとき、

$$\mathbf{X} = (\mathbf{P} \boldsymbol{\Lambda}^{1/2}) \mathbf{Z} + \boldsymbol{\mu} \quad (44.2)$$

とすると、定理 43.9 より、 $E[\mathbf{X}] = \boldsymbol{\mu}$ である。分散・共分散行列は、

$$\text{Var}[\mathbf{X}] = (\mathbf{P} \boldsymbol{\Lambda}^{1/2}) (\mathbf{P} \boldsymbol{\Lambda}^{1/2})' = \boldsymbol{\Sigma}$$

なので、 $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ である。一方、 Z_1, \dots, Z_n の同時密度関数と、定理 44.3 から、 $\mathbf{x} = {}^t(x_1, \dots, x_n)$ としたときに、(44.2) 式から

$$\mathbf{z} = \boldsymbol{\Lambda}^{-1/2} \mathbf{P}' (\mathbf{x} - \boldsymbol{\mu})$$

(ただし $\Lambda^{-1/2}$ は $\lambda^{1/2}$ の対角成分の逆数を取ったもの) で表されるので, X_1, \dots, X_n の同時密度関数 $f_{\mathbf{X}}(x_1, \dots, x_n)$ は存在して,

$$\begin{aligned} f_{\mathbf{X}}(x_1, \dots, x_n) &= \frac{1}{(2\pi)^{n/2}} \exp \left[-\frac{[(\Lambda^{-1/2})' \mathbf{P}'(\mathbf{x} - \mathbf{u})][\Lambda^{-1/2} \mathbf{P}'(\mathbf{x} - \mathbf{u})]}{2} \right] \left| \det \left(\frac{\partial \mathbf{z}}{\partial \mathbf{x}} \right) \right| \\ &= \frac{1}{(2\pi)^{n/2}} \exp \left[-\frac{(\mathbf{x} - \mathbf{u})' \mathbf{P} \Lambda^{-1} \mathbf{P}' (\mathbf{x} - \mathbf{u})}{2} \right] \left| \det(\Lambda^{-1/2} \mathbf{P}') \right| \end{aligned}$$

この式について, まず明らかに $\mathbf{P} \Lambda^{-1} \mathbf{P}' = \Sigma^{-1}$ である.

また, $|\det(\Lambda^{-1/2} \mathbf{P}')| = |\det(\Sigma)|^{-1/2}$ が成り立つことは, $(\mathbf{P} \Sigma^{-1/2})(\Sigma^{-1/2} \mathbf{P}') = \Sigma^{-1}$ と, 転置行列の行列式が等しいこと ($\det(\mathbf{P} \Sigma^{-1/2}) = \det(\Sigma^{-1/2} \mathbf{P}')$), 積の行列式の性質 ($\det(\mathbf{P} \Sigma^{-1/2} \Sigma^{-1/2} \mathbf{P}') = \det(\mathbf{P} \Sigma^{-1/2}) \det(\Sigma^{-1/2} \mathbf{P}')$) と, 対角成分が全て正の対角行列の性質 ($\det(\Sigma^{-1/2}) = (\det(\Sigma))^{-1/2}$) から言える.

以上より, X_1, \dots, X_n の密度関数は (44.1) 式となる. □

定義 44.6 X_1, \dots, X_n の同時密度関数が (44.1) 式となる時, そのときのみ, 確率変数ベクトル列ベクトル $\mathbf{X} = {}^t(X_1 \cdots X_n)$ は, 平均ベクトル \mathbf{u} ・ 分散・共分散行列 Σ の同時正規分布に従うという. ┘

同時正規分布は非退化の正規分布であり, この定義から明らかのように Σ が正値定符号のときのみ, 同時正規分布であり, 同時密度関数が存在する.

この定義と, 積の行列のランクから以下が簡単に言える.

命題 44.7 \mathbf{Y} が n 次確率変数列ベクトルで非退化の正規分布に従い, $\mathbf{Y} \sim N(\mathbf{E}[\mathbf{Y}], \text{Var}[\mathbf{Y}])$ とする.

\mathbf{A}, \mathbf{c} は定数の行列で, それぞれ $n \times m$ の行列・ m 次列ベクトルとする. このとき, 定理 43.9 より

$$\mathbf{Y} \sim N(\mathbf{E}[\mathbf{Y}], \text{Var}[\mathbf{Y}]) \implies \mathbf{A}'\mathbf{Y} + \mathbf{c} \sim N(\mathbf{A}'\mathbf{E}[\mathbf{Y}] + \mathbf{c}, \mathbf{A}'\text{Var}[\mathbf{Y}]\mathbf{A})$$

であるが, さらに以下が言える

- (i). $\text{rank}(\mathbf{A}) = m$ つまり \mathbf{A} が Full Column Rank のとき, $\mathbf{A}'\mathbf{Y} + \mathbf{c}$ は非退化の正規分布である. そうでなければ, $\mathbf{A}'\mathbf{Y} + \mathbf{c}$ は退化した正規分布である.
 - (ii). $[\text{Var}[\mathbf{Y}]]^{-1/2}(\mathbf{Y} - \mathbf{E}[\mathbf{Y}]) \sim N(\mathbf{0}_n, \mathbf{I}_n)$
- ┘

証明 (i) は $\mathbf{A}'\text{Var}[\mathbf{Y}]\mathbf{A}$ が正値定符号か否かで判定できる. (ii) は仮定より $\text{Var}[\mathbf{Y}]$ が正値定符号なので, $[\text{Var}[\mathbf{Y}]]^{-1/2}$ が存在することと, (i) から言える. □

(i) から $m > n$ の場合は必ず退化した正規分布となる. では退化したものについて簡単に説明しよう. $Z \sim N(0, 1)$ とする. 非 0 の定数 a, b について, $Y_1 = aZ, Y_2 = bZ$ とすると, 1 変数の変数変換より,

$$Y_1 \sim N(0, a^2), \quad Y_2 \sim N(0, b^2) \tag{44.3}$$

が言える。一方上の命題より,

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} Z \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} a & \\ & b \end{pmatrix}^{-1} \begin{pmatrix} a & b \end{pmatrix}\right)$$

が成り立つ。分散・共分散行列を計算すると,

$$\begin{pmatrix} a^2 & ab \\ ba & b^2 \end{pmatrix}$$

対角成分は a^2 と b^2 であり, それは (44.3) 式の分散部分と一致する。が相関係数を計算すると, 共分散を分散の積の平方根で割ったものだから

$$\frac{ab}{\sqrt{a^2b^2}} = 1$$

相関係数が 1 のときに同時正規分布は作れない。このように, 列ベクトル確率変数 \mathbf{Y} が退化した正規分布であるということは, 各成分は 1 変数の正規分布に従うのであるが, Y_1, \dots, Y_n についての n 変数同時密度関数をもたない正規分布なのである。

この章の最後に, 正規分布における, 独立と共分散 0 の関係について述べる。

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「確率・統計・計量経済学で出てくる積分 (2)」章で 2 つの正規分布について, 共分散が 0 であることと独立であるは同値であることを述べた。行列を用いて表すと, 以下が言える。

定理 44.8 (i). m 項確率変数ベクトル \mathbf{X} と n 項確率変数ベクトル \mathbf{Y} がともに正規分布,

$$\mathbf{X} \sim N(E[\mathbf{X}], \text{Var}[\mathbf{X}]), \quad \mathbf{Y} \sim N(E[\mathbf{Y}], \text{Var}[\mathbf{Y}])$$

に従っているとする。このとき共分散行列がゼロ行列のとき, つまり

$$E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{Y} - E[\mathbf{Y}])'] = \mathbf{O}_{mn}$$

が成り立つときは, \mathbf{X} の任意の 1 つの成分と, \mathbf{Y} の任意の 1 つの成分は, 独立である。このとき, \mathbf{X} と \mathbf{Y} は独立であるという。

(ii). n 項確率変数ベクトル \mathbf{Y} が以下のように, 各々独立で, 分散一定の正規分布に従うとする。

$$\mathbf{Y} \sim N(E[\mathbf{Y}], \sigma^2 \mathbf{I})$$

また行列 \mathbf{A}, \mathbf{B} がともに行数が n の定数行列とする。2 つの確率変数行列 $\mathbf{A}'\mathbf{Y}$ と $\mathbf{B}'\mathbf{Y}$ について,

$$\mathbf{A}'\mathbf{B} = \mathbf{O} \implies \mathbf{A}'\mathbf{Y} \text{ の任意の成分と, } \mathbf{B}'\mathbf{Y} \text{ の任意の成分は独立}$$

証明 (i) は正規分布の性質から明らかである。(ii) については,

$$\mathbf{A}'\mathbf{Y} \sim N(\mathbf{A}'E[\mathbf{Y}], \sigma^2 \mathbf{A}'\mathbf{A}), \quad \mathbf{B}'\mathbf{Y} \sim N(\mathbf{B}'E[\mathbf{Y}], \sigma^2 \mathbf{B}'\mathbf{B})$$

であり, $A'Y, B'Y$ の各成分ともにそれぞれ正規分布に従う. $A'Y, B'Y$ の共分散行列は,

$$\begin{aligned} \text{Cov}(A'Y, B'Y) &= E[(A'Y - A'E[Y])(B'Y - B'E[Y])'] \\ &= A'E[(Y - E[Y])(Y - E[Y])']B \\ &= A' \cdot \sigma^2 IB \\ &= O \end{aligned}$$

(i) より命題が成り立つ. □

この命題の (ii) では A, B の列数に制限はつけていない. つまりこの命題は非退化の正規分布の場合でも成り立つのである.

45 2次形式・標準正規分布・カイ2乗分布・t分布・F分布

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「確率・統計・計量経済学で出てくる積分 (2)」章で「 $(n$ 個の正規分布を標準化したものの 2 乗和) $i = 1, \dots, n$ について $X_i \sim N(\mu, \sigma^2)$ とし, 各 X_i は各々独立とする. このとき,

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n)$$

」を示した. これを標準正規分布の行列の書き方で書きなおすと, 以下となる.

系 45.1 n 項確率変数列ベクトル $Z \sim N(0, I_n)$ について,

$$Z'Z \sim \chi^2(n)$$

さらに上の Z' と Z の間にべき等対称行列 A を挟んだ $Z'AZ$ の分布について述べたものが以下である.

命題 45.2 n 項確率変数列ベクトル $Z \sim N(0, I)$ と, n 次べき等対称行列 A について,

$$Z'AZ \sim \chi^2(\text{tr}(A))$$

証明 $A = A^2 = A'A$ より, $Z'AZ = Z'A'AZ$. $A'A$ は対称行列なので, ある直交行列 Q と対角行列 Λ があって,

$$Q'(A'A)Q = \Lambda \tag{45.1}$$

と対角化できる. Λ の対角成分は, $A'A$ の固有値であるが $A'A = A$ より, それらは A の固有値である. 定理 38.9 より, 固有値は 1 か 0 のみであり, かつ Λ の対角成分の 1 の数は $\text{rank}(A)$ 個である. $r = \text{rank}(A)$ として, Λ の対角成分について

$$\Lambda_{i,i} = \begin{cases} 1 & (\text{if } i \leq r) \\ 0 & (\text{if } i > r) \end{cases}$$

としても一般性は失わないので、こう仮定する。(45.1) 式に $\mathbf{A}'\mathbf{A} = \mathbf{A}$ を代入して、 \mathbf{Q} を左から、 \mathbf{Q}' を右から掛けると、

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'$$

そこで $\mathbf{Y} = \mathbf{Q}'\mathbf{Z}$ とすると、 $E[\mathbf{Y}] = \mathbf{0}$ と $Var[\mathbf{Y}] = \mathbf{Q}'\mathbf{I}\mathbf{Q} = \mathbf{I}$ より、 $\mathbf{Y} \sim N(\mathbf{0}, \mathbf{I})$ が言える。よって

$$\begin{aligned} \mathbf{Z}'\mathbf{A}\mathbf{Z} &= \mathbf{Z}'\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}'\mathbf{Z} \\ &= \mathbf{Y}'\mathbf{\Lambda}\mathbf{Y} \\ &= \sum_{i=1}^n \Lambda_{i,i} Y_i^2 = \sum_{i=1}^r Y_i^2 \\ &= (Y_1 \ \cdots \ Y_r) \begin{pmatrix} Y_1 \\ \vdots \\ Y_r \end{pmatrix} \\ &\sim \chi^2(r) \end{aligned}$$

この式の最後の部分は、系 45.1 と $\begin{pmatrix} Y_1 \\ \vdots \\ Y_r \end{pmatrix} \sim N(\mathbf{0}, \mathbf{I}_r)$ から言える。 $r = rank(\mathbf{A}) = tr(\mathbf{A})$ なので、命題が成り立つ。 □

<http://www.ic.daito.ac.jp/~tkadoda/math/index.htm>

の [2] 経済学で出会う数学 + α (作成中) の「確率・統計・計量経済学で出てくる積分 (2)」章で

- (標準正規分布と t 分布の関係) n は正整数とする。 $Z \sim N(0, 1), Y \sim \chi^2(n)$ で独立とする。このとき、

$$\frac{Z}{\sqrt{Y/n}} \sim t(n)$$

- (カイ 2 乗分布と F 分布の関係) m, n は正整数とする。 $X \sim \chi^2(m), Y \sim \chi^2(n)$ で独立の時、 $\frac{X/m}{Y/n} \sim F(m, n)$

の 2 つを示した。これらと上の命題から以下が言える。

系 45.3 n 項確率変数列ベクトル \mathbf{Z} が各々独立な標準正規分布に従うとする。つまり、 $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I})$ 。このとき、

- (i). n 項列ベクトル \mathbf{a} と n 次べき等対称行列 \mathbf{B} について、 $\mathbf{a}'\mathbf{B} = 0$ のとき、

$$\frac{\mathbf{a}'\mathbf{Z}/\sqrt{\mathbf{a}'\mathbf{a}}}{\sqrt{\mathbf{Z}'\mathbf{B}\mathbf{Z}/tr(\mathbf{B})}} \sim t(tr(\mathbf{B}))$$

- (ii). 2 つの n 次べき等対称行列 \mathbf{A}, \mathbf{B} が、 $\mathbf{A}\mathbf{B} = \mathbf{O}$ を満たすとき、

$$\frac{\mathbf{Z}'\mathbf{A}\mathbf{Z}/tr(\mathbf{A})}{\mathbf{Z}'\mathbf{B}\mathbf{Z}/tr(\mathbf{B})} \sim F(tr(\mathbf{A}), tr(\mathbf{B}))$$

」

注意: $\mathbf{a}'\mathbf{Z}/\sqrt{\mathbf{a}'\mathbf{a}} \sim N(0,1)$ であること, つまり標準正規分布に従うことに注意すること. (注意終)

命題 45.4 n 項確率変数列ベクトルが非退化の同時正規分布 $\mathbf{X} \sim N(\boldsymbol{\mu}, \text{Var}[\mathbf{X}])$ に従うとする. このとき,

$$(\mathbf{X} - \boldsymbol{\mu})'[\text{Var}[\mathbf{X}]]^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(n)$$

」

証明 命題 44.7 より, $[\text{Var}[\mathbf{X}]]^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \mathbf{I}_n)$ である. n 個の独立な標準正規分布の 2 乗和は, 自由度 n のカイ 2 乗分布に従うので,

$$\{[\text{Var}[\mathbf{X}]]^{-1/2}(\mathbf{X} - \boldsymbol{\mu})\}'[\text{Var}[\mathbf{X}]]^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(n)$$

この転置部分を計算して, $\mathbf{X}^{-1/2}$ が対称行列であることから結局上の式は,

$$(\mathbf{X} - \boldsymbol{\mu})'[\text{Var}[\mathbf{X}]]^{-1}(\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(n)$$

となる.

□

補題 45.5 \mathbf{A} が n 次べき等対称行列とし, $\mathbf{B} = \mathbf{I} - \mathbf{A}$ とする. このとき, \mathbf{B} もべき等対称行列であり, さらに $\mathbf{AB} = \mathbf{BA} = \mathbf{O}$ である. 」

証明 \mathbf{B} が対称行列なのはすぐわかる. べき等であることは,

$$\begin{aligned} \mathbf{B}^2 &= (\mathbf{I} - \mathbf{A})(\mathbf{I} - \mathbf{A}) \\ &= \mathbf{I} - \mathbf{A} - \mathbf{A} + \mathbf{A}^2 \\ &= \mathbf{I} - \mathbf{A} - \mathbf{A} + \mathbf{A} \\ &= \mathbf{B} \end{aligned}$$

$\mathbf{A} + \mathbf{B} = \mathbf{I}$ の両辺の左から \mathbf{B} を掛けると,

$$\mathbf{AB} + \mathbf{B}^2 = \mathbf{B}$$

\mathbf{B} がべき等なので, $\mathbf{AB} = \mathbf{O}$ が言える. また同様に $\mathbf{A} + \mathbf{B} = \mathbf{I}$ の両辺の右側から \mathbf{A} を掛けて $\mathbf{BA} = \mathbf{O}$ が言える. □

統計学・計量経済学でよく出てくるのが, 以下である.

定理 45.6 $n \times k$ の行列 \mathbf{X} のランクが k とする (Full Column Rank). このとき, 2 つの n 次正方行列 $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ と $\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ はともにべき等対称行列であり, ランクはそれぞれ k と $n - k$ である. 」

証明 前補題より, $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ がべき等対称行列であることのみ示せばよい. 対称性については以下から言える.

$$(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')' = (\mathbf{X}')'[(\mathbf{X}'\mathbf{X})^{-1}]'\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

べき等性については以下から言える.

$$[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']^2 = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

べき等対称行列のランクは、そのトレースに等しいので、

$$\begin{aligned} \text{rank}(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') &= \text{tr}(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \\ &\text{命題 36.2 より } = \text{tr}(\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}) \\ &= \text{tr}(\mathbf{I}_k) \\ &= k \end{aligned}$$

□

46 回帰分析での係数の t 検定や F 検定への導出

OLS モデルで誤差項に正規性を仮定したモデルは、 n 項列ベクトル確率変数 \mathbf{y} が、 $n \times k$ の定数行列 \mathbf{X} (ランクは k) と、各項が未知定数である、 k 項列ベクトル $\boldsymbol{\beta}$ と n 項誤差ベクトルによって

$$\begin{cases} \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I}) \end{cases} \quad (46.1)$$

と表されるモデルである。OLS 推定量は

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (46.2)$$

であることはすでに示した。これを利用して、 t 検定や F 検定のための理論を示す。(46.1) 式を代入して、

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\epsilon} \quad (46.3)$$

なので、 $E[\hat{\boldsymbol{\beta}}] = \boldsymbol{\beta}$ であり、

$$\begin{aligned} \text{Var}[\hat{\boldsymbol{\beta}}] &= [(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\epsilon}] \sigma^2 \mathbf{I} [(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\epsilon}]' \\ &= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{X} (\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

以上より、

命題 46.1 (46.1) 式の OLS モデルでの、(46.3) 式の $\hat{\boldsymbol{\beta}}$ について、

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}) \quad (46.4)$$

┘

命題 46.2 (46.1) の OLS モデルにおいて、(46.3) 式の OLS 推定量ベクトル $\hat{\boldsymbol{\beta}}$ によって、 \mathbf{y} の推定ベクトル $\hat{\mathbf{y}}$ と、 $\boldsymbol{\epsilon}$ の推定ベクトル (残差ベクトル) $\hat{\boldsymbol{\epsilon}}$ が

$$\begin{cases} \hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\epsilon} \\ \hat{\boldsymbol{\epsilon}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{y} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\boldsymbol{\epsilon} \end{cases}$$

と表すことができる。

┘

$[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']$ がランク $n - k$ のべき等対称行列であることは、前章の 45.6 で示してある。残差は回帰分析での推定や検定において、大きな役割を持つ。

命題 46.3 (46.1) 式の OLS モデルの残差ベクトル $\hat{\epsilon}$ について,

- (i). $\hat{\epsilon} \sim N(\mathbf{0}, \sigma^2[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'])$
- (ii). $\frac{\hat{\epsilon}'\hat{\epsilon}}{\sigma^2} = \left(\frac{\epsilon}{\sigma}\right)' [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \left(\frac{\epsilon}{\sigma}\right) \sim \chi^2(n-k)$
- (iii). 残差ベクトル $\hat{\epsilon}$ の任意の 1 つの成分と, OLS 推定量ベクトル $\hat{\beta}$ の任意の 1 つの成分は独立である. これを, 残差ベクトルと, 回帰ベクトルは独立であるという.
- (iv). 残差ベクトル $\hat{\epsilon}$ の任意の 1 つの成分と, \mathbf{y} の推定量 $\hat{\mathbf{y}}$ の任意の 1 つの成分は独立である. これを, 残差ベクトルと, \mathbf{y} の推定量ベクトルは独立であるという.

□

証明 記号の簡単化のため $\mathbf{B} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ と置く. ランク $n-k$ のべき等対称行列である.

(i) の証明: 定理 43.9 より,

$$\hat{\epsilon} = \mathbf{B}\epsilon \sim N(\mathbf{0}, \mathbf{B}\sigma^2\mathbf{B}')$$

が成り立つ. \mathbf{B} はべき等対称行列なので, $\mathbf{B}\sigma^2\mathbf{B}' = \sigma^2\mathbf{B}$ である.

(ii) の証明: $\epsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ から $\frac{\epsilon}{\sigma} \sim N(\mathbf{0}, \mathbf{I})$ が成り立つことと, $\frac{\hat{\epsilon}}{\sigma} = \mathbf{B}\frac{\epsilon}{\sigma}$ かつ, \mathbf{B} がべき等対称行列でランクが $n-k$ であることから, 命題 45.2 より, $\left(\frac{\epsilon}{\sigma}\right)' \mathbf{B} \left(\frac{\epsilon}{\sigma}\right) \sim \chi^2(n-k)$ である. \mathbf{B} はべき等対称行列なので,

$$\begin{aligned} \left(\frac{\epsilon}{\sigma}\right)' \mathbf{B} \left(\frac{\epsilon}{\sigma}\right) &= \left(\mathbf{B}\frac{\epsilon}{\sigma}\right)' \left(\mathbf{B}\frac{\epsilon}{\sigma}\right) \\ &= \frac{\hat{\epsilon}'\hat{\epsilon}}{\sigma^2} \end{aligned}$$

(iii) の証明: (i) と (46.4) 式から $\hat{\beta}$ の各成分は正規分布に従い, (i) より残差の各成分も正規分布に従う. 定理 44.8 なので, $\hat{\epsilon}$ と $\hat{\beta}$ の共分散行列がゼロ行列であることを示せば良い.

$$\epsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I}), \quad \hat{\beta} = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon, \quad \hat{\epsilon} = \mathbf{B}\epsilon$$

と, $E[\hat{\beta}] = \beta, E[\hat{\epsilon}] = \mathbf{0}$ から, $\hat{\beta}$ と $\hat{\epsilon}$ の共分散行列は,

$$\begin{aligned} E[(\hat{\beta} - \beta)\hat{\epsilon}'] &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon[\mathbf{B}\epsilon]'] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon\epsilon'\mathbf{B}'] \\ \mathbf{B} \text{ を戻して} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}[\mathbf{X}' - \mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \\ &= \mathbf{0} \end{aligned}$$

(iv) の証明: $\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$ と (iii) から成り立つ.

□

(46.4) 式より, n 項定数列ベクトル \mathbf{a} と定数 q について,

$$\mathbf{a}'\hat{\beta} - q \sim N(\mathbf{a}'\beta - q, \sigma^2\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a})$$

上の左辺はスカラーである. $\mathbf{a}'\beta - q = 0$ が成り立つならば, 以下が言える.

$$\frac{\mathbf{a}'\hat{\beta} - q}{\sqrt{\sigma^2\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}}} \sim N(0, 1) \quad (46.5)$$

σ^2 は未知なので, 少し変更することによって以下が成り立つ.

定理 46.4 (46.1) 式の OLS モデルにおいて, $\mathbf{a}'\boldsymbol{\beta} - q = 0$ が真のとき, $s^2 = \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}/(n-k)$ すると,

$$\frac{\mathbf{a}'\hat{\boldsymbol{\beta}} - q}{\sqrt{s^2 \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}}} \sim t(n-k) \quad (46.6)$$

証明 (46.5) 式と命題 46.3(ii)(iii), と p160 の (標準正規分布と t 分布の関係) によって,

$$\frac{\mathbf{a}'\hat{\boldsymbol{\beta}} - q}{\sqrt{\sigma^2 \mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a}}} \sim t(n-k)$$

$$\frac{\mathbf{a}'\hat{\boldsymbol{\beta}} - q}{\sqrt{\frac{\hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}}{\sigma^2}/(n-k)}} \sim t(n-k)$$

σ^2 を約分して, s^2 を代入すると命題が成り立つ. □

よって, この s^2 や s については特に以下のように定義する.

定義 46.5 (46.1) 式の OLS モデルにおいて, σ^2 の推定量を $s^2 = \hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}/(n-k)$ であらわし, 誤差項の不偏分散という. その正の平方根 s を, 「回帰式の標準誤差」または, 「誤差項の標準誤差」という. また, $\hat{\boldsymbol{\beta}}$ の分散・共分散行列 $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$ の推定量 $\hat{V}[\hat{\boldsymbol{\beta}}]$ を,

$$\hat{V}[\hat{\boldsymbol{\beta}}] = s^2(\mathbf{X}'\mathbf{X})^{-1}$$

で表す. ┘

データで OLS 分析をした場合, 多くの PC ソフトでは, この $\hat{V}[\hat{\boldsymbol{\beta}}]$ の値を明示することはないが, OLS 推定後に, なんらかのコマンドを入力したり, マウスで選択することによって, 表示させることは可能である.

この定義から, 前定理と合わせて簡単に以下が言える. また推定や検定にもすぐに使うことができる.

系 46.6 (46.1) 式の OLS モデルにおいて,

$$\mathbf{a}'\boldsymbol{\beta} - q = 0 \implies \frac{\mathbf{a}'\hat{\boldsymbol{\beta}} - q}{\sqrt{\mathbf{a}'\hat{V}[\hat{\boldsymbol{\beta}}]\mathbf{a}}} \sim t(n-k) \quad (46.7)$$

サンプルの大きさ n と回帰係数の個数 k は, それぞれ \mathbf{X} の行数と列数に一致するので, $n-k$ についてはそれで覚えると良い.

さて上の式で, 特に \mathbf{a} の第 i 成分のみ 1 で, その他が 0 として $q = 0$ とおいた場合は,

$$\beta_i = 0 \implies \frac{\hat{\beta}_i}{\sqrt{\hat{V}[\hat{\boldsymbol{\beta}}]_{i,i}}} \sim t(n-k) \quad (46.8)$$

が成り立つ. $\hat{V}[\hat{\boldsymbol{\beta}}]_{i,i}$ は, 分散・共分散行列の推定値の第 (i, i) 成分であり, $\text{Var}[\hat{\beta}_i]$ の推定値である. その平方根 $\sqrt{\hat{V}[\hat{\boldsymbol{\beta}}]_{i,i}}$ は $\hat{\beta}_i$ の標準誤差である. 標準的な回帰分析表では, $i = 1, \dots, k$ について, それぞれの回帰係数の横に, それぞれの標準誤差が書かれている.

またデータから (46.8) 式を計算した値は, $\beta_i = 0$ が真のときの t 値といい, 標準的な回帰分析表では, $i = 1, \dots, k$ について, それぞれの標準誤差の右横にこの t 値が書かれている.

今の場合 β に関する制約が $\mathbf{a}'\beta - q = 0$ という方程式 1 本であった。例えば $k = 4$ の場合で、 $\beta_3 + \beta_4 = 2, \beta_2 = 0$ という 2 本ある場合は、

$$\mathbf{A}' = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{q} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

とにおいて、 $\mathbf{A}'\beta - \mathbf{q} = \mathbf{0}$ と制約式を表すことができる。これを利用して、 F 分布を使った検定を行うことができる。

定理 46.7 この章の今までの OLS モデルにおいて、 k 行 r 列の定数行列 \mathbf{A} (ただしランクは r) と、 r 項定数ベクトル \mathbf{q} について、

$$\mathbf{A}'\beta - \mathbf{q} = \mathbf{0} \implies \frac{(\mathbf{A}'\hat{\beta} - \mathbf{q})'[\mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}'\hat{\beta} - \mathbf{q})/r}{s^2} \sim F(r, n - k) \quad (46.9)$$

証明 (46.4) 式より、

$$\mathbf{A}'\beta - \mathbf{q} \sim N(\mathbf{0}, \sigma^2 \mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A})$$

$(\mathbf{X}'\mathbf{X})^{-1}$ は正則であり、 \mathbf{A} は Full Column Rank なので、系 39.11 より、 $\sigma^2 \mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}$ も正則であり、ランク r である。命題 45.4 より

$$[\mathbf{A}'\beta - \mathbf{q}]'[\sigma^2 \mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}[\mathbf{A}'\beta - \mathbf{q}] \sim \chi^2(r)$$

が成り立つ。さらに命題 46.3(ii)(iii) と、p160 の (カイ 2 乗分布と F 分布の関係) から、

$$\frac{[\mathbf{A}'\beta - \mathbf{q}]'[\sigma^2 \mathbf{A}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}]^{-1}[\mathbf{A}'\beta - \mathbf{q}]/r}{\frac{\hat{\epsilon}'\hat{\epsilon}/(n - k)}{\sigma^2}} \sim F(r, n - k)$$

が成り立ち、 σ^2 を約分して、 s^2 を代入すると、命題が成り立つ。

□

この式の分子は、(42.12) 式で示した形をしている。つまり制約付きの残差ベクトル $\tilde{\epsilon}$ を用いて以下のように書くことができる。

系 46.8 前定理の制約条件を満たしたとき OLS モデルでの残差ベクトルを $\tilde{\epsilon}$ とする。このとき、

$$\frac{(\tilde{\epsilon}'\tilde{\epsilon} - \hat{\epsilon}'\hat{\epsilon})/r}{\hat{\epsilon}'\hat{\epsilon}/(n - k)} \sim F(r, n - k)$$

□

多くの統計ソフトでは、回帰分析の後で制約式を指定すると、すぐに上記の値が計算されて表示される。

47 基本的な統計的推定や検定を回帰分析で説明する。

前章での回帰分析を応用して、初等的な統計学で学ぶものを、回帰分析によって再解釈して説明する。前節の OLS モデルで、 $\mathbf{X}, \mathbf{y}, \epsilon, \beta$ は同じものとし、推定量 $\hat{\beta}, \hat{\mathbf{y}}$ と残差 $\hat{\epsilon}$ や誤差項の標準誤差 s も同様とする。 \mathbf{X} や \mathbf{y} を工夫することによって、初等的な統計学での学びを再確認する。

まず最初の 1 群の平均値の検定や推定である。この仮定は、 $i = 1, \dots, n$ について、未知の数 μ, σ として、 Y_i が各々平均 μ 、分散 σ^2 の正規分布に従うというものであった。

例題 47.1 平均 μ 分散 σ^2 の正規母集団からの大きさ n の無作為標本 Y_1, \dots, Y_n について, σ が未知のときの, 真の平均 μ に関する, 区間推定や t 検定について, 自由度 $n-1$ の t 分布が関係していることを述べよ。」

(解答) (46.1) 式の OLS モデルの \mathbf{y} を, Y_1, \dots, Y_n を列ベクトル化したものとし, $\mathbf{X} = \mathbf{i}$ とする. この結果

$$\begin{aligned}\mathbf{y} &= \mathbf{i}\mu + \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} &\sim N(0, \sigma^2)\end{aligned}$$

とモデル化できる. $(\mathbf{i}'\mathbf{i})^{-1} = 1/n$ なので, μ の OLS 推定量 $\hat{\mu}$ は

$$\hat{\mu} = (\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'\mathbf{y} = (1/n)(Y_1 + \dots + Y_n) = \bar{Y}$$

と \mathbf{y} の標本平均となる. 残差は

$$\hat{\boldsymbol{\epsilon}} = \mathbf{y} - \mathbf{i}\hat{\mu} = \begin{pmatrix} Y_1 - \bar{Y} \\ \vdots \\ Y_n - \bar{Y} \end{pmatrix}$$

であり, $s^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2 / (n-1)$ が言える. ある実数 μ_0 をとって, $\mu = \mu_0$ が真のとき, 定理 46.4 の $\alpha = 1, q = \mu_0, \beta = \mu, \hat{\beta} = \hat{\mu}$ とおく. また定理 46.4 の $(\mathbf{X}'\mathbf{X})^{-1}$ は, 今の場合 $(\mathbf{i}'\mathbf{i}) = 1/n$ なので,

$$\frac{\hat{\mu} - \mu_0}{\sqrt{s^2(\mathbf{X}'\mathbf{X})^{-1}}} = \frac{\bar{Y} - \mu_0}{\sqrt{s^2/n}} \sim t(n-1)$$

がいえる.

(解答終)

これをもとにして, 真の値 μ の 95% 信頼区間を求めよう. それは自由度 $n-1$ の t 分布の上側 2.5% 点を t^* とし, $\hat{\mu}$ に \bar{Y} を代入して, 信頼区間

$$\left[\bar{Y} - t_* \frac{s}{\sqrt{n}}, \bar{Y} + t_* \frac{s}{\sqrt{n}} \right]$$

が得られる. これが初等的な統計学で学んだ通りの結果となっていることを確認すると良い.

次は 2 群の平均値についてである.

例題 47.2 μ, ν, σ は未知とする. $i = 1, \dots, m$ について, U_i が各々平均 μ , 分散 σ^2 の正規分布に従うとする. $j = 1, \dots, n$ について, V_j が各々平均 ν , 分散 σ^2 の正規分布に従うとする. 当然任意の U_i と V_j についても独立とする. ある実数 q_0 で, $\mu - \nu = q_0$ についての区間推定や t 検定について, 自由度 $m+n-2$ の t 分布が関係していることを述べよ。」

(解答) U_1, \dots, U_m と V_1, \dots, V_n のそれぞれの標本平均を \bar{U}, \bar{V} とする. 上のモデルを回帰分析で書くことにしよう. \mathbf{u}, \mathbf{v} を, U_1, \dots, U_m と V_1, \dots, V_n をそれぞれ列ベクトル化したものとする. \mathbf{y} はそれを縦に並べた $\mathbf{y} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$ とし, $\mathbf{X} = \begin{pmatrix} \mathbf{i}_m & \mathbf{0}_m \\ \mathbf{0}_n & \mathbf{i}_n \end{pmatrix}$ とする. 誤差項の仮定は $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{m+n})$ である. これで (46.1) 式の OLS モデルのモデル化ができた.

$$\begin{aligned}(\mathbf{X}'\mathbf{X})^{-1} &= \left[\begin{pmatrix} \mathbf{i}'_m & \mathbf{0}'_n \\ \mathbf{0}'_m & \mathbf{i}'_n \end{pmatrix} \begin{pmatrix} \mathbf{i}_m & \mathbf{0}_m \\ \mathbf{0}_n & \mathbf{i}_n \end{pmatrix} \right]^{-1} \\ &= \begin{pmatrix} m & 0 \\ 0 & n \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1/m & 0 \\ 0 & 1/n \end{pmatrix}\end{aligned}\tag{47.1}$$

このとき, $\hat{\beta}$ は

$$\begin{aligned}\hat{\beta} &= \begin{pmatrix} 1/m & 0 \\ 0 & 1/n \end{pmatrix} \begin{pmatrix} \mathbf{i}'_m & \mathbf{0}'_n \\ \mathbf{0}'_m & \mathbf{i}'_n \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \\ &= \begin{pmatrix} \bar{U} \\ \bar{V} \end{pmatrix}\end{aligned}$$

と, それぞれの標本平均が β の OLS 推定量となる. 残差は

$$\begin{aligned}\hat{\epsilon} &= \mathbf{y} - \mathbf{X}\hat{\beta} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \begin{pmatrix} \mathbf{i}_m \bar{U} \\ \mathbf{i}_n \bar{V} \end{pmatrix} \\ &= \begin{pmatrix} (U_1 - \bar{U}) \\ \vdots \\ (U_m - \bar{U}) \\ (V_1 - \bar{V}) \\ \vdots \\ (V_n - \bar{V}) \end{pmatrix}\end{aligned}$$

であり, \mathbf{u}, \mathbf{v} からそれぞれの標本平均を引いたものとなる. \mathbf{X} の行数が $m+n$ で列数が 2 なので, 誤差項の不偏分散は $s^2 = [\sum_{i=1}^m (U_i - \bar{U})^2 + \sum_{i=1}^n (V_i - \bar{V})^2] / (m+n-2)$ である. ある実数 q_0 をとって, $\nu - \mu = q_0$ が真のとき, 定理 46.4 の $\mathbf{a}' = \begin{pmatrix} -1 & 1 \end{pmatrix}$ で $q = q_0$ において, (47.1) 式と合わせて,

$$\frac{\bar{V} - \bar{U} - q_0}{\sqrt{s^2 \begin{pmatrix} -1 & 1 \end{pmatrix} (\mathbf{X}'\mathbf{X})^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix}}} = \frac{\bar{V} - \bar{U} - q_0}{\sqrt{s^2(1/m + 1/n)}} \sim t(m+n-2)$$

がいえる.

(解答終)

このように \mathbf{X} をとると, $(\mathbf{X}'\mathbf{X})^{-1}$ の計算が簡単である. しかし統計学などでは, 定数項以外の全ての変数が 0 の場合の検定を考えることが多い. よって \mathbf{X} に, 全てが 1 の列 \mathbf{i} を入れることが一般的である. 実際すでに, 命題 42.3 の後でも説明したように, やはり \mathbf{i} を入れたほうがよい. \mathbf{i} の列は, PC ソフトによっては最初の列だったり最後の列だったりするが, 教科書では最初の列に入れていることが多い. ここでも最初の列に入れた上で, この例題の別解を書いておこう.

(解答) \mathbf{u}, \mathbf{v} は前の解答と同様とする. $\mathbf{X} = \begin{pmatrix} \mathbf{i}_m & \mathbf{0}_m \\ \mathbf{i}_n & \mathbf{i}_n \end{pmatrix}$ とする. \mathbf{X} の第 1 列が定数項で, 第 2 列が「2 群ダミー」を表す変数となる. 回帰モデル $\mathbf{y} = \mathbf{X}\beta + \epsilon$ から, y_i が 1 群の場合は, $y_i = \beta_1 + \epsilon_i$ となり, 2 群の場合は, $y_i = \beta_1 + \beta_2 + \epsilon_i$ と表される. すなわち $\beta_1 = \mu, \beta_2 = \nu - \mu$ である.

$$\begin{aligned}(\mathbf{X}'\mathbf{X})^{-1} &= \left[\begin{pmatrix} \mathbf{i}'_m & \mathbf{i}'_n \\ \mathbf{0}'_m & \mathbf{i}'_n \end{pmatrix} \begin{pmatrix} \mathbf{i}_m & \mathbf{0}_m \\ \mathbf{i}_n & \mathbf{i}_n \end{pmatrix} \right]^{-1} \\ &= \begin{pmatrix} m+n & n \\ n & n \end{pmatrix}^{-1} \\ &= \frac{1}{mn} \begin{pmatrix} n & -n \\ -n & m+n \end{pmatrix}\end{aligned}\tag{47.2}$$

となるので, $\hat{\beta}$ は

$$\begin{aligned}\hat{\beta} &= \frac{1}{mn} \begin{pmatrix} n & -n \\ -n & m+n \end{pmatrix} \begin{pmatrix} \mathbf{i}'_m & \mathbf{i}'_n \\ \mathbf{0}'_m & \mathbf{i}'_n \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \\ &= \frac{1}{mn} \begin{pmatrix} n & -n \\ -n & m+n \end{pmatrix} \begin{pmatrix} m\bar{U} + n\bar{V} \\ n\bar{V} \end{pmatrix} \\ &= \begin{pmatrix} \bar{U} \\ \bar{V} - \bar{U} \end{pmatrix}\end{aligned}$$

となる. 残差は

$$\begin{aligned}\hat{\epsilon} &= \mathbf{y} - \mathbf{X}\hat{\beta} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \begin{pmatrix} i_m\bar{U} \\ i_n\bar{U} + i_n(\bar{V} - \bar{U}) \end{pmatrix} \\ &= \begin{pmatrix} (U_1 - \bar{U}) \\ \vdots \\ (U_m - \bar{U}) \\ (V_1 - \bar{V}) \\ \vdots \\ (V_n - \bar{V}) \end{pmatrix}\end{aligned}$$

である. この残差は上の解答と全く同じ形となるので, 誤差分散 s^2 も同じとなる. ある実数 q_0 をとって, $\nu - \mu = q_0$ が真のときとは, $\beta_2 = \nu - \mu$ であるから, 今の場合 $\beta_2 = q_0$ が真のときとなる. 定理 46.4 の $\mathbf{a}' = (0 \ 1)$ で $q = q_0$ とおいて, (47.2) 式と合わせて,

$$\frac{\bar{V} - \bar{U} - q_0}{\sqrt{s^2 (0 \ 1) (\mathbf{X}'\mathbf{X})^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}}} \sim t(m+n-2)$$

$(0 \ 1) (\mathbf{X}'\mathbf{X})^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ は, $(\mathbf{X}'\mathbf{X})^{-1}$ の (2,2) 成分なので, $(m+n)/mn$ つまり $(1/m + 1/n)$. よって,

$$\frac{\bar{V} - \bar{U} - q_0}{\sqrt{s^2(1/m + 1/n)}} \sim t(m+n-2)$$

と, 前解答と全く同じことがこの解答からもいえる.

(解答終)

\mathbf{X} が前の解答とは異なるので, 回帰係数も異なるのだが, 残差は同じものとなる.

このダミー変数の考え方で, 次は m 群の平均値について検定する. ただし, \mathbf{X} の形が込み入ってくるので, 表記を簡単にするために以下の記号を導入する.

定義 47.3 (i). $m \times n$ の行列 \mathbf{A} の各 i 列を \mathbf{a}_i について, これを順に縦に並べた列ベクトル

$$\begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

を \mathbf{A} のベクトル化といい $\text{vec}(\mathbf{A})$ で表す.

- (ii). m_a 行 n_a 列の行列 \mathbf{A} と, m_b 行 n_b 列の行列 \mathbf{B} について, \mathbf{A} の各成分 a_{ij} に \mathbf{B} を掛けて, それを分割された行列に並べた $m_a \times m_b$ 行, $n_a \times n_b$ 列の行列

$$\begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n_a}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n_a}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m_a1}\mathbf{B} & a_{m_a2}\mathbf{B} & \cdots & a_{m_an_a}\mathbf{B} \end{pmatrix}$$

を $\text{kron}(\mathbf{A}, \mathbf{B})$ で表し, \mathbf{A} と \mathbf{B} のクロネッカー積という.

- (iii). n 個の行列 $\mathbf{A}_1, \dots, \mathbf{A}_n$ について, これらを分割された行列の対角成分にとった行列

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_2 & \ddots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \cdots & \mathbf{A}_n \end{pmatrix}$$

を $\text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_n)$ で表す.

」

イメージしやすいように,

$$\mathbf{U} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix}$$

として上の計算を行ってみよう. $\text{vec}(\mathbf{U}) = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 4 \end{pmatrix}$ は明らかであろう. それ以外は以下となるので, 確認する

とよい.

$$\text{kron}(\mathbf{U}, \mathbf{v}) = \begin{pmatrix} 1 \begin{pmatrix} a \\ b \end{pmatrix} & 2 \begin{pmatrix} a \\ b \end{pmatrix} \\ 3 \begin{pmatrix} a \\ b \end{pmatrix} & 4 \begin{pmatrix} a \\ b \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a & 2a \\ b & 2b \\ 3a & 4a \\ 3b & 4b \end{pmatrix}$$

$$\text{kron}(\mathbf{v}, \mathbf{U}) = \begin{pmatrix} a \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ b \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a & 2a \\ 3a & 4a \\ b & 2b \\ 3b & 4b \end{pmatrix}$$

$$\text{diag}(\mathbf{U}, \mathbf{v}) = \begin{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} & \mathbf{O} \\ \mathbf{O} & \begin{pmatrix} a \\ b \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & a \\ 0 & 0 & b \end{pmatrix}$$

$$\text{diag}(\mathbf{v}, \mathbf{U}) = \begin{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} & \mathbf{O} \\ \mathbf{O} & \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ b & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 3 & 4 \end{pmatrix}$$

では m 群の正規母集団の平均値の検定を, 回帰分析で確かめよう.

例題 47.4 m 群の異なる正規母集団を考える. $j = 1, \dots, m$ について, それぞれの群の母集団は, 平均 μ_j , 分散 σ^2 の正規分布に従っているとす. 各群は独立とする. $j = 1, \dots, m$ について, それぞれ n_1, \dots, n_m の大きさの無作為標本を取り出す. 各群の無作為標本を縦に並べた列ベクトルを $\mathbf{y}_1, \dots, \mathbf{y}_m$ とする. この合計 $n_1 + \dots + n_m$ 個のサンプルから, 全ての群の真の平均が等しいという帰無仮説 $H_0: \mu_1 = \mu_2 = \dots = \mu_m$ が真の場合, F 分布を用いた検定を行うのだが, そのときの F 分布の自由度や, またその F 統計量を計算せよ.

注意: ダミー変数を使う方法で考える.

(解答) $\mathbf{y}_1, \dots, \mathbf{y}_m$ を順に縦に並べた列ベクトルを \mathbf{y} とする. 長さ $\sum_{i=1}^m n_i$ なので, これを n とする. 各 j についての \mathbf{y}_j の成分の平均を \bar{y}_j で表し, \mathbf{y} に含まれる全ての平均を \bar{y} で表す. 従って, \mathbf{y}_j の成分の和は $n_j \bar{y}_j$ で, \mathbf{y} の成分の和は $n \bar{y}$ で表される. 列ごとの総和が, 全体の総和が等しいことから $n_1 \bar{y}_1 + \dots + n_m \bar{y}_m = n \bar{y}$ がいえる.

$$\mathbf{X} = \left(\mathbf{i}_n \left(\text{diag}(\mathbf{0}_{n_1, m-1}, \mathbf{i}_{n_2}, \mathbf{i}_{n_3}, \dots, \mathbf{i}_{n_m}) \right) \right) \quad (47.3)$$

とする. \mathbf{X} の 1 列目が定数項の列であり, $j = 2, \dots, m$ について第 j 列が, 「 j 群ダミー変数」を表す. 誤差項の仮定は $\epsilon \sim N(0, \sigma^2 \mathbf{I}_n)$ である. これより, 回帰モデル $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ によって, 第 1 群に入っている y_k については, $y_k = \beta_1 + \epsilon_k$ となり, $j \geq 2$ についての第 j 群に入っている y_k については, $y_k = \beta_1 + \beta_j + \epsilon_k$ となる. よって β_1 の真の値は μ_1 , $j \geq 2$ についての β_j の真の値は $\mu_j - \mu_1$ となる. よって帰無仮説 H_0 が真であるとは, $\beta_2 = \beta_3 = \dots = \beta_m = 0$ を意味する.

$$\begin{aligned} (\mathbf{X}'\mathbf{X})^{-1} &= \left[\left(\begin{pmatrix} \mathbf{i}'_n & \mathbf{0}'_{n_1, m-1} \\ \text{diag}(\mathbf{i}'_{n_2}, \mathbf{i}'_{n_3}, \dots, \mathbf{i}'_{n_m}) \end{pmatrix} \begin{pmatrix} \mathbf{i}_n & \text{diag}(\mathbf{0}_{n_1, m-1}, \mathbf{i}_{n_2}, \mathbf{i}_{n_3}, \dots, \mathbf{i}_{n_m}) \end{pmatrix} \right) \right]^{-1} \\ &= \begin{pmatrix} n & (n_2 \ n_3 \ \dots \ n_m) \\ (n_2 \ n_3 \ \dots \ n_m)' & \text{diag}(n_2, n_3, \dots, n_m) \end{pmatrix}^{-1} \end{aligned} \quad (47.4)$$

(注意: $F = n - (n_2 \ n_3 \ \dots \ n_m) [\text{diag}(n_2, n_3, \dots, n_m)]^{-1} (n_2 \ n_3 \ \dots \ n_m)' = n_1 \neq 0$ なので分割された行列の逆行列の公式より, $\mathbf{B} = [\text{diag}(n_2, n_3, \dots, n_m)]^{-1}$ として)

$$\begin{aligned} &= \begin{pmatrix} n_1^{-1} & -n_1^{-1} (n_2 \ n_3 \ \dots \ n_m) [\text{diag}(n_2, n_3, \dots, n_m)]^{-1} \\ \text{右上の成分の転置} & \mathbf{B} + \mathbf{B} \begin{pmatrix} n_2 & n_3 & \dots & n_m \end{pmatrix} n_1^{-1} \begin{pmatrix} n_2 & n_3 & \dots & n_m \end{pmatrix} \mathbf{B} \end{pmatrix} \\ &= \begin{pmatrix} 1/n_1 & (-1/n_1) \mathbf{i}'_{m-1} \\ (-1/n_1) \mathbf{i}_{m-1} & \mathbf{B} + (1/n_1) \mathbf{i}_{m-1} \mathbf{i}'_{m-1} \end{pmatrix} \end{aligned} \quad (47.5)$$

このとき, $\hat{\beta}$ は

$$\begin{aligned}
 \hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\
 &= \begin{pmatrix} 1/n_1 & (-1/n_1)\mathbf{i}'_{m-1} \\ (-1/n_1)\mathbf{i}_{m-1} & \mathbf{B} + (1/n_1)\mathbf{i}_{m-1}\mathbf{i}'_{m-1} \end{pmatrix} \begin{pmatrix} \mathbf{O}_{m-1, n_1} & \text{diag}(\mathbf{i}'_{n_2}, \mathbf{i}'_{n_3}, \dots, \mathbf{i}'_{n_m}) \end{pmatrix} \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} \\
 &= \frac{1}{n_1} \begin{pmatrix} 1 & -\mathbf{i}'_{m-1} \\ -\mathbf{i}_{m-1} & n_1\mathbf{B} + \mathbf{i}_{m-1}\mathbf{i}'_{m-1} \end{pmatrix} \begin{pmatrix} n\bar{y} \\ n_2\bar{y}_2 \\ \vdots \\ n_m\bar{y}_m \end{pmatrix} \\
 &= \frac{1}{n_1} \begin{pmatrix} n\bar{y} - (n_2\bar{y}_2 + \dots + n_m\bar{y}_m) \\ -n\mathbf{i}_{m-1}\bar{y} + n_1 \begin{pmatrix} \bar{y}_2 \\ \vdots \\ \bar{y}_m \end{pmatrix} + \mathbf{i}_{m-1} \sum_{j=2}^m n_j\bar{y}_j \end{pmatrix} \\
 &= \frac{1}{n_1} \begin{pmatrix} n_1\bar{y}_1 \\ n_1 \begin{pmatrix} \bar{y}_2 \\ \vdots \\ \bar{y}_m \end{pmatrix} - \mathbf{i}_{m-1}n_1\bar{y}_1 \end{pmatrix} \\
 &= \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 - \bar{y}_1 \\ \vdots \\ \bar{y}_m - \bar{y}_1 \end{pmatrix}
 \end{aligned}$$

が言える. 前例題と同様に, β_1 の推定量は \bar{y}_1 で, 第 j ダミー変数の係数を表す β_j の推定量は $\bar{y}_j - \bar{y}_1$ で表される. 前例題同様に残差は, もとの値から各群の平均値を引いたものとなり,

$$\begin{aligned}
 \hat{\epsilon} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\beta} &= \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} - \begin{pmatrix} \mathbf{i}_{n_1}\bar{y}_1 \\ \mathbf{i}_{n_2}\bar{y}_2 \\ \vdots \\ \mathbf{i}_{n_m}\bar{y}_m \end{pmatrix} \\
 &= \begin{pmatrix} \mathbf{y}_1 - \mathbf{i}_{n_1}\bar{y}_1 \\ \mathbf{y}_2 - \mathbf{i}_{n_2}\bar{y}_2 \\ \vdots \\ \mathbf{y}_m - \mathbf{i}_{n_m}\bar{y}_m \end{pmatrix}
 \end{aligned}$$

である. \mathbf{y}_j の第 k 成分を $y_j(k)$ で表すと, 残差 2 乗和は $\hat{\epsilon}'\hat{\epsilon} = \sum_{j=1}^m \sum_{k=1}^{n_j} (y_j(k) - \bar{y}_j)^2$ と表され, \mathbf{X} の行数が n で列数が $m-1$ なので, 誤差分散は s^2 は, この残差 2 乗和を $(n-m+1)$ で割ったものとなる.

次に H_0 の制約付きの OLS の推定量 $\tilde{\beta}$ と, そのときの残差 $\tilde{\epsilon}$ を求める. $\tilde{\beta}$ を (42.11) 式でも求めることができるが, それよりももっと簡単に得られる. 今の制約がある場合は, 先の \mathbf{X}, \mathbf{y} に対して,

$$\beta = \begin{pmatrix} \beta_1 \\ \mathbf{0}_{m-1} \end{pmatrix}$$

となる。もとの $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ に代入すると、

$$\mathbf{y} = \mathbf{X} \begin{pmatrix} \beta_1 \\ \mathbf{0}_{m-1} \end{pmatrix} = \mathbf{i}\beta_1 + \boldsymbol{\epsilon}$$

となり、制約付きの OLS は、このときの β_1 を OLS で推計すれば良い。例題 47.1 と同じ計算から、 β_1 の OLS 推定量は \mathbf{y} の標本平均つまり \bar{y} である。以上より制約付きの OLS 推定量は

$$\tilde{\boldsymbol{\beta}} = \begin{pmatrix} \bar{y} \\ \mathbf{0}_{m-1} \end{pmatrix}$$

と書ける。この、制約条件付き OLS の残差 $\tilde{\boldsymbol{\epsilon}}$ は、

$$\tilde{\boldsymbol{\epsilon}} = \mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}} = \mathbf{y} - \mathbf{i}_n\bar{y}$$

よってこのときの残差 2 乗和は

$$\tilde{\boldsymbol{\epsilon}}' \tilde{\boldsymbol{\epsilon}} = \sum_{j=1}^m \sum_{k=1}^{n_j} (y_j(k) - \bar{y})^2$$

が成り立ち、平均からの偏差の 2 乗和なので総変動である。 \mathbf{X} の行数が n で列数が m なので、前章の系 46.8 から、以下が成り立つ。

$$\frac{[\sum_{j=1}^m \sum_{k=1}^{n_j} (y_j(k) - \bar{y})^2 - \sum_{j=1}^m \sum_{k=1}^{n_j} (y_j(k) - \bar{y}_j)^2]/(m-1)}{\sum_{j=1}^m \sum_{k=1}^{n_j} (y_j(k) - \bar{y}_j)^2/(n-m)} \sim F(m-1, n-m)$$

さらにこの分子の [] 内は、総変動 - 残差変動となっているので、回帰変動となる。計算すると、以下のように変形できる。

$$\frac{\sum_{j=1}^m n_j (\bar{y}_j - \bar{y})^2 / (m-1)}{\sum_{j=1}^m \sum_{k=1}^{n_j} (y_j(k) - \bar{y}_j)^2 / (n-m)} \sim F(m-1, n-m) \quad (47.6)$$

(解答終)

初等的な統計学の講義では、(47.6) 式の結果のみが与えられることが多いが、回帰分析で解釈すると以上のように得られる。この証明から、回帰分析の理論面では残差が重要であることがわかるであろう。

この章の最後に、繰り返しのない、二元配置モデルを回帰モデルで解釈する。個体数 n のものを、 m 回実験を行い、その m 群の平均値の検定を、個体差をコントロールした上で行うというものである。

例えば $n = 5$ で、 $m = 3$ のときを考えてみよう。ある 5 人が、テストを 3 回受けた結果表が以下とする。

	1 回目	2 回目	3 回目
個人 1	49	60	68
個人 2	42	40	44
個人 3	20	27	28
個人 4	14	12	16
個人 5	80	91	99

(47.7)

このとき、個人の能力をコントロールした上で、1 回目のテスト・2 回目のテスト・3 回目のテストについて、平均について差があるかどうかを検定したいわけである。

モデルとしては、各回の点数を表す μ_1, μ_2, μ_3 と、個人の能力差を表す v_1, v_2, v_3, v_4, v_5 があって、8 つとも未知とする。個人 i の j 回目のテストが平均 $\mu_j + v_i$ で分散 σ^2 の、各々独立な正規分布に従っている仮定して、 $\mu_1 = \mu_2 = \mu_3$ について検定するのである。一般形で述べたものが以下となる。

例題 47.5 上で述べた n 個体, m 群に関する二元配置について, (47.7) 式のように縦に $i = 1, \dots, n$, 横に $j = 1, \dots, m$ として, i 個体の, j 群の確率変数を並べた $n \times m$ の行列を \mathbf{Y} とする. i 個体での j 群での確率変数を $Y_{i,j}$ とし, ある未知の m 個の実数 μ_1, \dots, μ_m とある未知の n 個の実数 v_1, \dots, v_n で,

$$Y_{i,j} \sim N(\mu_i + \nu_j, \sigma^2)$$

と仮定し, 各々独立と仮定する. この mn のサンプルから, 全ての群の真の平均が等しいという帰無仮説 $H_0: \mu_1 = \mu_2 = \dots = \mu_m$ が真の場合, F 分布を用いた検定を行うのだが, そのときの F 分布の自由度や, またその F 統計量を計算せよ. ┘

(解答) まず各 j の標本平均を $\bar{y}_{.j}$ で表す. さらに各個体 i の, m 回の平均 $(Y_{i,1} + Y_{i,2} + \dots + Y_{i,m})/m$ を $\bar{h}_{i.}$ で表す. 全体の平均は \bar{y} で表す. なので, 全体の合計は $mn\bar{y}$ で表される.

次に OLS モデルを考える. $\mathbf{y} = \text{vec}(\mathbf{Y})$ とする. 説明変数 \mathbf{X} についてはまず, 一元配置の式 (47.3) で, $n_1 = \dots = n_m = n$ のときの,

$$\left(\mathbf{i}_{mn} \quad \begin{pmatrix} \mathbf{O}_{n,m-1} \\ \text{kron}(\mathbf{I}_{m-1}, \mathbf{i}_n) \end{pmatrix} \right)$$

が, 定数項の列と, 2 群ダミーから m 群ダミーを含んだ列となる.

さらにコントロールするダミー変数として, $i = 2, \dots, n$ ダミー変数を表す説明変数行列が,

$$\text{kron} \left(\mathbf{i}_m, \begin{pmatrix} \mathbf{0}'_{n-1} \\ \mathbf{I}_{n-1} \end{pmatrix} \right)$$

である. この 2 つを並べて説明変数行列

$$\mathbf{X} = \left(\mathbf{i}_{mn} \quad \begin{pmatrix} \mathbf{O}_{n,m-1} \\ \text{kron}(\mathbf{I}_{m-1}, \mathbf{i}_n) \end{pmatrix} \quad \text{kron} \left(\mathbf{i}_m, \begin{pmatrix} \mathbf{0}'_{n-1} \\ \mathbf{I}_{n-1} \end{pmatrix} \right) \right) \quad (47.8)$$

を作る. これで OLS モデル $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ とする. 二元配置モデルで $\mu_1, \mu_2, \mu_3, v_1, v_2, v_3, v_4, v_5$ で計 8 つの数を未知としたが, \mathbf{X} の列数はそれより 1 つ少ない 7 である. 一元配置の場合は, bmX の列数と未知母数の数はあっていたが, 二元配置の場合は, そうではないのがポイントである.

例えば (47.7) 式の場合だと, 以下となる.

$$\mathbf{y} = \begin{pmatrix} 49 \\ 42 \\ 20 \\ 14 \\ 80 \\ 60 \\ 40 \\ 27 \\ 12 \\ 91 \\ 68 \\ 44 \\ 28 \\ 16 \\ 99 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (47.9)$$

一般の形に戻って, $\mathbf{X}'\mathbf{X}$ を計算する. これは簡単に, 以下が得られる.

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} mn & n\mathbf{i}'_{m-1} & m\mathbf{i}'_{n-1} \\ n\mathbf{i}_{m-1} & n\mathbf{I}_{m-1} & \mathbf{i}_{m-1}\mathbf{i}'_{n-1} \\ m\mathbf{i}_{n-1} & \mathbf{i}_{n-1}\mathbf{i}'_{m-1} & m\mathbf{I}_{n-1} \end{pmatrix}$$

これより,

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} \frac{m+n-1}{mn} & -\frac{1}{n}\mathbf{i}'_{m-1} & -\frac{1}{m}\mathbf{i}'_{n-1} \\ -\frac{1}{n}\mathbf{i}_{m-1} & \frac{1}{n}[\mathbf{I}_{m-1} + \mathbf{i}_{m-1}\mathbf{i}'_{m-1}] & \mathbf{O} \\ -\frac{1}{m}\mathbf{i}_{n-1} & \mathbf{O} & \frac{1}{m}[\mathbf{I}_{n-1} + \mathbf{i}_{n-1}\mathbf{i}'_{n-1}] \end{pmatrix} \quad (47.10)$$

が成り立つ. 上の式の証明は長いので, この章末 (47.1 節) で述べるとする. 次に $\mathbf{X}'\mathbf{y}$ は,

$$\mathbf{X}'\mathbf{y} = \begin{pmatrix} mn\bar{y} \\ (n\bar{y}_{.,2}) \\ \vdots \\ (n\bar{y}_{.,m}) \\ (m\bar{h}_{2,.}) \\ \vdots \\ (m\bar{h}_{n,.}) \end{pmatrix}$$

以上より OLS 推定量 $\hat{\beta}$ は,

$$\begin{aligned} \hat{\beta} &= \begin{pmatrix} \frac{m+n-1}{mn} & -\frac{1}{n}\mathbf{i}'_{m-1} & -\frac{1}{m}\mathbf{i}'_{n-1} \\ -\frac{1}{n}\mathbf{i}_{m-1} & \frac{1}{n}[\mathbf{I}_{m-1} + \mathbf{i}_{m-1}\mathbf{i}'_{m-1}] & \mathbf{O} \\ -\frac{1}{m}\mathbf{i}_{n-1} & \mathbf{O} & \frac{1}{m}[\mathbf{I}_{n-1} + \mathbf{i}_{n-1}\mathbf{i}'_{n-1}] \end{pmatrix} \begin{pmatrix} mn\bar{y} \\ (n\bar{y}_{.,2}) \\ \vdots \\ (n\bar{y}_{.,m}) \\ (m\bar{h}_{2,.}) \\ \vdots \\ (m\bar{h}_{n,.}) \end{pmatrix} \\ &= \begin{pmatrix} (m+n-1)\bar{y} - \sum_{j=2}^n \bar{y}_{.,j} - \sum_{i=2}^n \bar{h}_{i,.} \\ -m\bar{y}\mathbf{i}_{m-1} + \begin{pmatrix} \bar{y}_{.,2} \\ \vdots \\ \bar{y}_{.,m} \end{pmatrix} + \mathbf{i}_{m-1} \sum_{j=2}^m \bar{y}_{.,j} \\ -n\bar{y}\mathbf{i}_{n-1} + \begin{pmatrix} \bar{h}_{2,.} \\ \vdots \\ \bar{h}_{n,.} \end{pmatrix} + \mathbf{i}_{m-1} \sum_{i=2}^n \bar{h}_{i,.} \end{pmatrix} \\ &= \begin{pmatrix} (\bar{y}_{1,.} + \bar{h}_{.,1} - \bar{y}) \\ (\bar{y}_{.,2} - \bar{y}_{.,1}) \\ \vdots \\ (\bar{y}_{.,m} - \bar{y}_{.,1}) \\ (\bar{h}_{2,.} - \bar{h}_{1,.}) \\ \vdots \\ (\bar{y}_{n,.} - \bar{h}_{1,.}) \end{pmatrix} \end{aligned}$$

定数項が, 1 群の平均と個体 1 の平均から, 全体の平均 \bar{y} を引いたものとなる. 2 群ダミーから, m 群ダミーについては, それぞれの j 群の平均 $\bar{y}_{.,j}$ から, 1 群の平均 $\bar{y}_{.,1}$ を引いたものとなっている.

個体 2 ダミーから, 個体 n ダミーについては, それぞれの個体 i の平均 $\bar{h}_{i,.}$ から, 個体 1 の平均 $\bar{h}_{1,.}$ を引いたものとなっている.

\mathbf{y} の推定量 $\hat{\mathbf{y}}$ は, 行列で書くよりも, 各成分で書いたほうが分かりやすい. $Y_{i,j}$ つまり, 個体 i の j 群で確率変数について, その推定量は, $j \geq 2, i \geq 2$ のときは, 定数項 $\bar{y}_{1,.} + \bar{h}_{1,.} - \bar{y}$ と, 第 j 群ダミー $\bar{y}_{.,j} - \bar{y}_{.,1}$

と、個体第 i ダミー $\overline{h_{i..}} - \overline{h_{1..}}$ の和であり、計算して、

$$\overline{y_{.j}} + \overline{h_{i..}} - \bar{y}$$

である。そしてこれは $j = 1$ や $i = 1$ の場合でも成り立つ。すなわち二元配置モデルの $\mu_j + v_i$ の推定量が

$$\widehat{\mu_j + v_i} = \overline{y_{.j}} + \overline{h_{i..}} - \bar{y}$$

と表されるわけである。

計算例として (47.7) 式の場合を考えよう。平均の推定値は

$$(\overline{y_{.1}} \quad \overline{y_{.2}} \quad \overline{y_{.3}}) = (41 \quad 46 \quad 51), \quad \begin{pmatrix} \overline{h_{.1}} \\ \overline{h_{.2}} \\ \overline{h_{.3}} \\ \overline{h_{.4}} \\ \overline{h_{.5}} \end{pmatrix} = \begin{pmatrix} 59 \\ 42 \\ 25 \\ 14 \\ 90 \end{pmatrix}, \quad \bar{y} = 46$$

となるので、 $Y_{i,j}$ の推定値 $\hat{Y}_{i,j} = \widehat{\mu_j + v_i} = \overline{y_{.j}} + \overline{h_{i..}} - \bar{y}$ は、以下のように表される。

	1回目	2回目	3回目
個人1	54	59	64
個人2	37	42	47
個人3	20	25	30
個人4	9	14	19
個人5	85	90	95

さて一般形に戻る。残差の各成分は、 $Y_{i,j}$ からその推定量 $\overline{y_{.j}} + \overline{h_{i..}} - \bar{y}$ を引いたものであるから、残差 2 乗和は、

$$\sum_{i=1}^n \sum_{j=1}^m (Y_{i,j} - \overline{y_{.j}} + \overline{h_{i..}} - \bar{y})^2 \quad (47.11)$$

である。OLS の説明変数行列 \mathbf{X} の行数が mn で、列数が $1 + (m-1) + (n-1) = m+n-1$ なので、その差は $(mn - m - n + 1) = (m-1)(n-1)$ 。誤差項の不偏分散 s^2 は、上の式を $(m-1)(n-1)$ で割ったものである。

ではすべての群の平均が等しい場合についての検定について考える。全ての群の平均が等しいとは、2群ダミーから m 群ダミーの係数が全て 0 の場合である。制約の本数は、前例題同様に、 $m-1$ 本となる。このときの制約付きの OLS のを考えてその残差 2 乗和を求めたい。この制約の場合は、(47.8) 式の真ん中の列に対する係数が全て 0 である制約なので、2 列目を取り去った、

$$\tilde{\mathbf{X}} = \left(\mathbf{i}_{mn} \quad \text{kron} \left(\mathbf{i}, \begin{pmatrix} \mathbf{0}'_{n-1} \\ \mathbf{I}_{n-1} \end{pmatrix} \right) \right)$$

を説明変数として、 \mathbf{y} を回帰してその残差を求めれば良い。するとこの説明変数は、定数項と個体 2 から個体 n までのダミー変数のみなので、前例題の一元配置の考え方から、その OLS 推定量 $(\tilde{\mathbf{X}}' \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}' \mathbf{y}$ は、

$$\begin{pmatrix} \overline{h_{1..}} \\ \left(\overline{h_{2..}} - \overline{h_{1..}} \right) \\ \vdots \\ \left(\overline{h_{n..}} - \overline{h_{1..}} \right) \end{pmatrix}$$

である。このときの y の推定量も、行列でなく各成分で書いたほうが分かりやすい。 $Y_{i,j}$ つまり、個体 i の j 群で確率変数について、その推定量は、 $j \geq 2$ のときは、定数項 $\overline{h_{1..}}$ と、

個体第 i ダミー $\overline{h_{i..}} - \overline{h_{1..}}$ の和であり、計算して、 $\overline{h_{i..}}$ である。これは $j = 1$ でも成り立つ。 j にかかわらず、 $Y_{i,j}$ の推定量がこのようになる。(47.7) 式の場合だと、この制約付きの場合の $Y_{i,j}$ の推定値は、

	1 回目	2 回目	3 回目
個人 1	59	59	59
個人 2	42	42	42
個人 3	25	25	25
個人 4	14	14	14
個人 5	90	90	90

である。このように j について変化しないのである。

さて一般形に戻ると、制約付きの OLS の残差 2 乗和は、制約なしの場合と同様に成分ごとに書くと、

$$\sum_{i=1}^n \sum_{j=1}^m (Y_{i,j} - \overline{h_{i..}})^2$$

で表される。

従って前例題と同様に、制約付きの残差 2 乗和を制約なしの残差 2 乗和で引いて、制約本数 $m - 1$ で割ったものを分子に、分母は制約なしの誤差項の不偏分散にして、

$$\frac{[\sum_{i=1}^n \sum_{j=1}^m (Y_{i,j} - \overline{h_{i..}})^2 - \sum_{i=1}^n \sum_{j=1}^m (Y_{i,j} - \overline{y_{..j}} + \overline{h_{i..}} - \overline{y})^2]/(m-1)}{s^2} \sim F(m-1, (m-1)(n-1))$$

が成り立つ。分子を計算し、

$$\frac{[\sum_{j=1}^m n(\overline{y_{..j}} - \overline{y})^2]/(m-1)}{s^2} \sim F(m-1, (m-1)(n-1)) \quad (47.12)$$

(解答終)

(47.7) 式の場合で、(47.12) 式の値を計算すると、

$$\frac{(250 - 168)/2}{168/8} = 125/21 = 5.95$$

自由度 (2, 8) の F 分布の上側 2.6% 点であり、

$$\begin{cases} H_0: \text{全ての群の平均が等しい.} \\ H_1: H_0 \text{ ではない.} \end{cases}$$

という仮説検定については、有意水準 5% で H_0 が棄却される結果となる。このあたりは統計で学ぶことであろう。ここでは、行列を使った回帰分析を用いて、 F 分布を使った検定をできることや、分子と分母の自由度が、それぞれ $m - 1, (m - 1)(n - 1)$ となる理由を理解すれば良い。

また、各個体差が存在しないという検定も同様に可能である。

$$\begin{cases} H_0: \text{個体差が存在しない.} \\ H_1: H_0 \text{ ではない.} \end{cases}$$

については、 H_0 が真のときには、前例題と全く同様にして、

$$\frac{[\sum_{i=1}^n m(\overline{h_{i.}} - \bar{y})^2]/(n-1)}{s^2} \sim F(n-1, (m-1)(n-1)) \quad (47.13)$$

が言える。(47.12) 式の分子の [] 内を二元配置における列変動といい、(47.13) 式の分子の [] 内を二元配置における行変動という。残差変動と総変動との関係が以下である。

定理 47.6 前例題の二元配置モデルで、

$$\sum_{i=1}^n \sum_{j=1}^m (Y_{i,j} - \bar{y}) = \sum_{i=1}^n m(\overline{h_{i.}} - \bar{y})^2 + \sum_{j=1}^m n(\overline{y_{.j}} - \bar{y})^2 + \sum_{i=1}^n \sum_{j=1}^m (Y_{i,j} - \overline{y_{.j}} + \overline{h_{i.}} - \bar{y})^2$$

が成り立つ。つまり、

$$\text{総変動} = \text{行変動} + \text{列変動} + \text{残差変動}$$

┘

証明 説明変数行列に、定数項の列 \mathbf{i}_{mn} が含まれているので、命題 42.3 より総変動 = 回帰変動 + 残差変動が言えるので、

$$\begin{aligned} \text{総変動} &= \sum_{i=1}^n \sum_{j=1}^m (\widehat{Y}_{i,j} - \bar{y})^2 + \text{残差変動} \\ &= \sum_{i=1}^n \sum_{j=1}^m (\overline{y_{.j}} + \overline{h_{i.}} - \bar{y} - \bar{y})^2 + \text{残差変動} \\ &= \sum_{i=1}^n \sum_{j=1}^m [(\overline{y_{.j}} - \bar{y}) + (\overline{h_{i.}} - \bar{y})]^2 + \text{残差変動} \\ &= \sum_{i=1}^n \sum_{j=1}^m (\overline{y_{.j}} - \bar{y})^2 + 2 \sum_{i=1}^n \sum_{j=1}^m (\overline{y_{.j}} - \bar{y})(\overline{h_{i.}} - \bar{y}) + \sum_{i=1}^n \sum_{j=1}^m (\overline{h_{i.}} - \bar{y})^2 + \text{残差変動} \\ &\quad \left(\sum_{j=1}^m (\overline{y_{.j}} - \bar{y}) = \sum_{j=1}^m \overline{y_{.j}} - m\bar{y} = 0 \text{ なので} \right) \\ &= \sum_{j=1}^m n(\overline{y_{.j}} - \bar{y})^2 + 0 + \sum_{i=1}^n m(\overline{h_{i.}} - \bar{y})^2 + \text{残差変動} \\ &= \text{列変動} + \text{行変動} + \text{残差変動} \end{aligned}$$

□

この証明自体に行列は含まれないが、説明変数行列に、定数項 \mathbf{i} が含まれているという意味で、行列の概念が入っている。

47.1 (47.10) 式の証明

煩雑なので、ゼロ行列 \mathbf{O} やゼロベクトル $\mathbf{0}$ の添字は省略して書く。

証明その1: 適当な m, n を用いて Scilab を使って, 推論から得る方法.

$m = 3, n = 5$ の場合の (47.9) 式の \mathbf{X} を作り, $(\mathbf{X}'\mathbf{X})^{-1}$ を計算させる. Scilab コードの一例は以下である.

```
X=[ones(15,1) zeros(15,6)];
X(6:10,2)=1;
X(11:15,3)=1;
X(1:5,4:7)=[zeros(1,4);eye(4,4)];
X(6:10,4:7)=[zeros(1,4);eye(4,4)];
X(11:15,4:7)=[zeros(1,4);eye(4,4)];
X'*X
inv(X'*X)
```

これで最後の2行によって表示されるのが $\mathbf{X}'\mathbf{X}, (\mathbf{X}'\mathbf{X})^{-1}$ であり,

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 15 & 5 & 5 & 3 & 3 & 3 & 3 \\ 5 & 5 & 0 & 1 & 1 & 1 & 1 \\ 5 & 0 & 5 & 1 & 1 & 1 & 1 \\ 3 & 1 & 1 & 3 & 0 & 0 & 0 \\ 3 & 1 & 1 & 0 & 3 & 0 & 0 \\ 3 & 1 & 1 & 0 & 0 & 3 & 0 \\ 3 & 1 & 1 & 0 & 0 & 0 & 3 \end{pmatrix}$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 0.46667 & -0.2 & -0.2 & -0.333 & -0.333 & -0.333 & -0.333 \\ -0.2 & 0.4 & 0.2 & 0 & 0 & 0 & 0 \\ -0.2 & 0.2 & 0.4 & 0 & 0 & 0 & 0 \\ -0.333 & 0 & 0 & 0.667 & 0.333 & 0.333 & 0.333 \\ -0.333 & 0 & 0 & 0.333 & 0.667 & 0.333 & 0.333 \\ -0.333 & 0 & 0 & 0.333 & 0.333 & 0.667 & 0.333 \\ -0.333 & 0 & 0 & 0.333 & 0.333 & 0.333 & 0.667 \end{pmatrix}$$

が得られる. $1/m$ が約 0.3333 で, $1/n = 0.2$ だから, 一般的な m, n のときに

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} \text{ア} & -\frac{1}{n}\mathbf{i}'_{m-1} & -\frac{1}{m}\mathbf{i}'_{n-1} \\ -\frac{1}{n}\mathbf{i}_{m-1} & \text{イ} & \mathbf{O} \\ -\frac{1}{m}\mathbf{i}_{n-1} & \mathbf{O} & \text{ウ} \end{pmatrix}$$

となるだろうと予測できる. アは何かの実数なので a とする. イとウは不明だが, $\mathbf{X}'\mathbf{X}$ には $\mathbf{I}, \mathbf{i}, \mathbf{i}', \mathbf{i}\mathbf{i}'$ の定数倍しか含まれていないので, 上の数字から, 実数 b, c, d, e で

$$\text{イ} = b\mathbf{I}_{m-1} + c\mathbf{i}_{m-1}\mathbf{i}'_{m-1}, \quad \text{ウ} = d\mathbf{I}_{n-1} + e\mathbf{i}_{n-1}\mathbf{i}'_{n-1}$$

となっていることが予想される. この予想から,

$$\begin{pmatrix} mn & n\mathbf{i}'_{m-1} & m\mathbf{i}'_{n-1} \\ n\mathbf{i}_{m-1} & n\mathbf{I}_{m-1} & \mathbf{i}_{m-1}\mathbf{i}'_{n-1} \\ m\mathbf{i}_{n-1} & \mathbf{i}_{m-1} & m\mathbf{I}_{n-1} \end{pmatrix} \begin{pmatrix} a & -\frac{1}{n}\mathbf{i}'_{m-1} & -\frac{1}{m}\mathbf{i}'_{n-1} \\ -\frac{1}{n}\mathbf{i}_{m-1} & b\mathbf{I}_{m-1} + c\mathbf{i}_{m-1}\mathbf{i}'_{m-1} & \mathbf{O} \\ -\frac{1}{m}\mathbf{i}_{n-1} & \mathbf{O} & d\mathbf{I}_{n-1} + e\mathbf{i}_{n-1}\mathbf{i}'_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{O}' & \mathbf{O}' \\ \mathbf{O} & \mathbf{I}_{m-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{I}_{n-1} \end{pmatrix}$$

を満たす a, b, c, d, e を探す. 左辺の分割された行列の積で, 1行1列, 2行2列, 3行3列を計算すると,

$$\begin{cases} mna + n\mathbf{i}'_{m-1}\frac{-1}{n}\mathbf{i}_{m-1} + m\mathbf{i}'_{n-1}\frac{-1}{m}\mathbf{i}_{n-1} & = 1 \\ n\mathbf{i}_{m-1}\frac{-1}{n}\mathbf{i}'_{n-1} + nb\mathbf{I}_{m-1} + nci_{m-1}\mathbf{i}'_{m-1} & = \mathbf{I}_{m-1} \\ m\mathbf{i}_{n-1}\frac{-1}{m}\mathbf{i}'_{n-1} + md\mathbf{I}_{n-1} + me\mathbf{i}_{n-1}\mathbf{i}'_{n-1} & = \mathbf{I}_{n-1} \end{cases}$$

簡単化して

$$\begin{cases} mna - (m-1) - (n-1) & = 1 \\ nb\mathbf{I}_{m-1} + (nc-1)\mathbf{i}_{m-1}\mathbf{i}'_{m-1} & = \mathbf{I}_{m-1} \\ md\mathbf{I}_{n-1} + (me-1)\mathbf{i}_{n-1}\mathbf{i}'_{n-1} & = \mathbf{I}_{n-1} \end{cases}$$

これより, $a = \frac{m+n-1}{mn}, b = c = \frac{1}{n}, d = e = \frac{1}{m}$ のとき上の式が成り立つ. $\mathbf{X}'\mathbf{X}$ が対称行列なので, 計算していない $(1,2), (1,3), (2,1), (2,3), (3,1), (3,2)$ の残りの6成分のうち, 上三角側の $(1,2), (1,3), (2,3)$ 成分がこの a, b, c, d, e のときに, 全てゼロ行列になることを確かめれば良いし, 実際にゼロ行列になる.

よって (47.10) 式が成り立つ.

Scilab のコードは初心者でも理解できるように \mathbf{X} についてわかりやすく書いた. クロネッカー積を用いて \mathbf{X} の定義どおりに以下のように書けば, 一発で \mathbf{X} が得られる.

```
X=[ones(m*n,1) [zeros(n,m-1);kron(eye(m-1,m-1),ones(n,1))] kron(ones(m,1),[zeros(1,n-1);eye(n-1,n-1)])]
```

なおコードが長すぎる場合には, Scilab のコマンドラインで「...」のあと Enter キーを打って次の行に続きを書くことができる.

PC をこのように推論の1つとして利用することは一般的なもので, Scilab に限らず, ある程度使えるようになることが望ましい.

証明その2: $\mathbf{X}'\mathbf{X}$ の拡大係数行列から基本行変形を用いて作る方法を示す. $(\mathbf{X}'\mathbf{X})$ が \mathbf{I}, \mathbf{i} などを用いた比較的簡単な形をしているので, 比較的簡単に, 基本行変形で得られる. 拡大係数行列を以下のように書く.

$$\left(\begin{array}{ccc|ccc} mn & n\mathbf{i}'_{m-1} & m\mathbf{i}'_{n-1} & 1 & \mathbf{0}' & \mathbf{0}' \\ n\mathbf{i}_{m-1} & n\mathbf{I}_{m-1} & \mathbf{i}_{m-1}\mathbf{i}'_{n-1} & \mathbf{0} & \mathbf{I}_{m-1} & \mathbf{O} \\ m\mathbf{i}_{n-1} & \mathbf{i}_{n-1}\mathbf{i}'_{m-1} & m\mathbf{I}_{n-1} & \mathbf{0} & \mathbf{O} & \mathbf{I}_{n-1} \end{array} \right) \quad (47.14)$$

2段目の計 $(m-1)$ 行を全て加えたものは,

$$\left(n(m-1) \quad n\mathbf{i}'_{m-1} \quad (m-1)\mathbf{i}'_{n-1} \mid 0 \quad \mathbf{i}'_{m-1} \quad \mathbf{0}' \right)$$

これを, n で割った行ベクトルは, 以下となる.

$$\left(m-1 \quad \mathbf{i}'_{m-1} \quad \frac{m-1}{n}\mathbf{i}'_{n-1} \mid 0 \quad \frac{1}{n}\mathbf{i}'_{m-1} \quad \mathbf{0}' \right)$$

前者を (47.14) 式1段目から引き, 後者を, (47.14) 式の3段目の行全てから全て引くと,

$$\left(\begin{array}{ccc|ccc} n & \mathbf{0}' & \mathbf{i}'_{n-1} & 1 & -\mathbf{i}'_{m-1} & \mathbf{0}' \\ n\mathbf{i}_{m-1} & n\mathbf{I}_{m-1} & \mathbf{i}_{m-1}\mathbf{i}'_{n-1} & \mathbf{0} & \mathbf{I}_{m-1} & \mathbf{O} \\ \mathbf{i}_{n-1} & \mathbf{O} & m\mathbf{I}_{n-1} - \frac{m-1}{n}\mathbf{i}_{n-1}\mathbf{i}'_{n-1} & \mathbf{0} & -\frac{1}{n}\mathbf{i}_{n-1}\mathbf{i}'_{m-1} & \mathbf{I}_{m-1} \end{array} \right) \quad (47.15)$$

これで, 2列目の非対角部分が \mathbf{O} となった. 次は1列目である. 1行目の $\left(n \quad \mathbf{0}' \quad \mathbf{i}'_{n-1} \mid 1 \quad -\mathbf{i}'_{m-1} \quad \mathbf{0}' \right)$ を2段目の行のそれぞれから引くと,

$$\left(\begin{array}{ccc|ccc} n & \mathbf{0}' & \mathbf{i}'_{n-1} & 1 & -\mathbf{i}'_{m-1} & \mathbf{0}' \\ \mathbf{0} & n\mathbf{I}_{m-1} & \mathbf{O} & -\mathbf{i}_{m-1} & \mathbf{I}_{m-1} + \mathbf{i}_{m-1}\mathbf{i}'_{m-1} & \mathbf{O} \\ \mathbf{i}_{n-1} & \mathbf{O} & m\mathbf{I}_{n-1} - \frac{m-1}{n}\mathbf{i}_{n-1}\mathbf{i}'_{n-1} & \mathbf{0} & -\frac{1}{n}\mathbf{i}_{n-1}\mathbf{i}'_{n-1} & \mathbf{I}_{m-1} \end{array} \right) \quad (47.16)$$

1 行目を n で割って、それを 3 段目の各行から引くと、

$$\left(\begin{array}{ccc|ccc} 1 & \mathbf{0}' & & \frac{1}{n}\mathbf{i}'_{n-1} & & \\ \mathbf{0} & n\mathbf{I}_{m-1} & & \mathbf{O} & -\frac{1}{n}\mathbf{i}'_{m-1} & \\ \mathbf{0} & \mathbf{O} & m\mathbf{I}_{n-1} - \frac{m}{n}\mathbf{i}_{n-1}\mathbf{i}'_{n-1} & & \mathbf{I}_{m-1} + \mathbf{i}_{m-1}\mathbf{i}'_{m-1} & \mathbf{O}' \\ & & & -\frac{1}{n}\mathbf{i}_{n-1} & \mathbf{O} & \mathbf{I}_{n-1} \end{array} \right) \quad (47.17)$$

これで 1 列目の非対角部分が \mathbf{O} となった。残りは 3 列目である。3 段目の各行を全て加えた行ベクトルは、

$$\left(\mathbf{0} \quad \mathbf{0}' \quad \frac{m}{n}\mathbf{i}'_{n-1} \mid -\frac{n-1}{n} \quad \mathbf{0}' \quad \mathbf{i}'_{n-1} \right)$$

この $\frac{1}{m}$ 倍を 1 行目から引き、さらに上のこの行ベクトルを 3 段目の各行に加えると、

$$\left(\begin{array}{ccc|ccc} 1 & \mathbf{0}' & \mathbf{0}' & \frac{1}{n} + \frac{n-1}{nm} & -\frac{1}{n}\mathbf{i}'_{m-1} & -\frac{1}{m}\mathbf{i}'_{n-1} \\ \mathbf{0} & n\mathbf{I}_{m-1} & \mathbf{O} & -\mathbf{i}_{m-1} & \mathbf{I}_{m-1} + \mathbf{i}_{m-1}\mathbf{i}'_{m-1} & \mathbf{O} \\ \mathbf{0} & \mathbf{O} & m\mathbf{I}_{n-1} & -\mathbf{i}_{n-1} & \mathbf{O} & \mathbf{I}_{n-1} + \mathbf{i}_{n-1}\mathbf{i}'_{n-1} \end{array} \right) \quad (47.18)$$

2 段目の各行を n で割って、3 段目の各行を m で割ると、以下のように左側が単位行列となる。

$$\left(\begin{array}{ccc|ccc} 1 & \mathbf{0}' & \mathbf{0}' & \frac{m+n-1}{nm} & -\frac{1}{n}\mathbf{i}'_{m-1} & -\frac{1}{m}\mathbf{i}'_{n-1} \\ \mathbf{0} & \mathbf{I}_{m-1} & \mathbf{O} & -\frac{1}{n}\mathbf{i}_{m-1} & \frac{1}{n}\mathbf{I}_{m-1} + \frac{1}{n}\mathbf{i}_{m-1}\mathbf{i}'_{m-1} & \mathbf{O} \\ \mathbf{0} & \mathbf{O} & \mathbf{I}_{n-1} & -\frac{1}{m}\mathbf{i}_{n-1} & \mathbf{O} & \frac{1}{m}\mathbf{I}_{n-1} + \frac{1}{m}\mathbf{i}_{n-1}\mathbf{i}'_{n-1} \end{array} \right) \quad (47.19)$$

この右側が $(\mathbf{X}'\mathbf{X})^{-1}$ となるので、(47.10) 式が成り立つ。