

線形代数学演習—行列と行列式*

計算問題の略解付き

黒木 玄 (東北大学大学院理学研究科数学専攻)

最終更新 2004 年 2 月 9 日 (作成 2004 年 2 月 9 日)

目 次

1	行列に関する雑多な問題	4
1.1	回転行列	4
1.2	対角行列	4
1.3	行列単位	5
1.4	置換行列	6
1.5	確率行列	7
1.6	Neumann 級数	9
1.7	Neumann 級数 (続き)	10
1.8	三角行列	10
1.9	行列の指数関数	11
1.10	複素数の実 2 次正平方行列による表現	13
1.11	コラム: 線形代数と量子力学	13
2	行列式	14
2.1	行列式の定義と基本性質	14
2.2	行列式の余因子展開と Cramér の公式	20
2.3	Cayley-Hamilton の定理	25
2.4	Gauss 分解	29
2.5	行列式と面積や体積の関係	32
2.6	小行列式と外積代数	37
3	色々な行列式	43
3.1	Vandermonde の行列式 (ヴァンデルモンドの行列式)	43
3.2	Cauchy の行列式と Hilbert の行列式	44
3.3	Gram 行列式	46
3.4	Sylvester の行列式と判別式	47

*この演習問題集は 2003 年度における東北大学理学部数学科 1 年生前期の代数学序論 A のために作成された。

3.5	Jacobian	48
3.6	特性多項式の係数	49
3.7	行列式の微分と行列の指数関数の行列式	49
3.8	Pfaffian (パフィアン)	49
3.9	雑多な問題	50
4	一次方程式の理論	52
4.1	一次方程式とその解空間	52
4.2	Cramér の公式について	55
4.3	行列の基本操作と基本変形の導入	55
4.4	行列の rank の定義	58
4.5	行列の基本変形による行列の簡単化	62
4.6	K^n の部分空間とその次元	65
4.7	rank と image と kernel の関係	70
4.8	行列の rank の計算の仕方	72
4.9	斉次な一次方程式の解法	75
4.10	非斉次な一次方程式の解法	77
5	体上のベクトル空間の理論	81
5.1	可換環上の加群と体上のベクトル空間の定義	81
5.2	加群の準同型とベクトル空間の線形写像の定義	85
5.3	一次独立性と基底	89
5.4	同型写像	89
5.5	線形写像の行列表示	90
5.6	商ベクトル空間	99
5.7	双対空間	104
6	計量ベクトル空間	107
6.1	計量ベクトル空間の定義	107
6.2	Cauchy-Schwarz の不等式と三角不等式	108
6.3	正規直交基底	110

記号法 (Notation)

以下のような記号を用いる:

$$\mathbb{Z} = \{ \text{整数全体} \} = \{ \dots, -2, -1, 0, 1, 2, \dots \},$$

$$\mathbb{R} = \{ \text{実数全体} \} \ni \sqrt{2}, \pi, e,$$

$$\mathbb{C} = \{ \text{複素数全体} \} \ni i = \sqrt{-1},$$

$$M_{m,n}(\mathbb{R}) = \{ \text{実 } m \times n \text{ 行列全体} \},$$

$$M_{m,n}(\mathbb{C}) = \{ \text{複素 } m \times n \text{ 行列全体} \},$$

$$M_n(\mathbb{R}) = M_{n,n}(\mathbb{R}) = \{ \text{実 } n \text{ 次正方行列全体} \},$$

$$M_n(\mathbb{C}) = M_{n,n}(\mathbb{C}) = \{ \text{複素 } n \text{ 次正方行列全体} \},$$

$$\mathbb{R}^n = \{ \text{実 } n \text{ 次元縦ベクトル全体} \} = M_{n,1}(\mathbb{R}),$$

$$\mathbb{C}^n = \{ \text{複素 } n \text{ 次元縦ベクトル全体} \} = M_{n,1}(\mathbb{C}),$$

$${}^t A = (\text{行列 } A \text{ の転置}),$$

$$\text{tr } A = (\text{正方行列 } A \text{ のトレース}) = (A \text{ の対角成分の和}),$$

$$\det A = |A| = (\text{正方行列 } A \text{ の行列式}),$$

$$\delta_{i,j} = (\text{Kronecker のデルタ}) = (i = j \text{ のとき } 1 \text{ でそうでないとき } 0).$$

面倒なのでベクトルを太字で印刷したりしない.

1 行列に関する雑多な問題

1.1 回転行列

実数 θ に対して次の行列 $R(\theta)$ を回転行列と呼ぶことにする:

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

$R(\theta)$ は xy 平面の一次変換

$$v = \begin{bmatrix} x \\ y \end{bmatrix} \mapsto R(\theta)v = \begin{bmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{bmatrix}$$

を定める.

[1] $R(\theta)$ が定める xy 平面の一次変換が反時計回りに角度 θ の回転になっていることを図を描いて説明せよ. \square

[2] 三角関数の加法公式¹ を用いて $R(\theta + \phi) = R(\theta)R(\phi)$ ($\theta, \phi \in \mathbb{R}$) が成立することを証明せよ. 逆に $R(\theta + \phi) = R(\theta)R(\phi)$ ($\theta, \phi \in \mathbb{R}$) が成立することを仮定して, 三角関数の加法公式を証明せよ. さらに $\det R(\theta) = 1$ も示せ. \square

参考 1.1 $R(\theta + \phi) = R(\theta)R(\phi)$ という公式は xy 平面をまず角度 ϕ だけ回転させてから, 角度 θ の回転を行なうことは一度に角度 $\theta + \phi$ の回転を行なうことに等しいという意味を持っている. これは直観的には当然のことである. 上の問題の結果より, 三角関数の加法公式はこの直観から導かれることがわかる. \square

1.2 対角行列

A が実 n 次正方行列であり, その $(1, 1), \dots, (n, n)$ 成分以外がすべて 0 であるとき, A は実 n 次対角行列 (**diagonal matrix**) であるという. 複素対角行列も同様に定義する. n 次の実対角行列と複素対角行列全体の集合をそれぞれ次のように表わすことにする:

$$\begin{aligned} D_n(\mathbb{R}) &= \{ \text{実 } n \text{ 次対角行列全体} \} \subset M_n(\mathbb{R}), \\ D_n(\mathbb{C}) &= \{ \text{複素 } n \text{ 次対角行列全体} \} \subset M_n(\mathbb{C}). \end{aligned}$$

さらに対角成分が順に a_1, a_2, \dots, a_n であるような n 次対角行列を次のように表わす:

$$\text{diag}(a_1, \dots, a_n) = \begin{bmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_n \end{bmatrix}.$$

[3] $K = \mathbb{R}, \mathbb{C}$ の両方に対して以下が成立する²:

¹ $\cos(\theta + \phi), \sin(\theta + \phi)$ を $\cos \theta, \cos \phi, \sin \theta, \sin \phi$ で表わす公式のこと. 高校で習う.

² \mathbb{R}, \mathbb{C} は体をなす. 体は英語では field と呼ばれ, ドイツ語では Körper と呼ばれる. $K = \mathbb{R}, \mathbb{C}$ の K は kelper の頭文字の K である.

1. $D_n(K)$ はともに行列の和と差と積について閉じている.
2. 任意の $A, B \in D_n(K)$ に対して $AB = BA$.
3. $A = \text{diag}(a_1, \dots, a_n) \in D_n(K)$, $X \in M_{n,m}(K)$, $Y \in M_{m,n}(K)$ のとき, A の X への左からの積は X の第 i 行を a_i 倍し, A の X への右からの積は X の第 j 列を a_j 倍する.
4. $A = \text{diag}(-1, 1, e^{\pi i}) \in D_3(\mathbb{C})$ であるとき A と可換な³複素 3 次正方行列全体の集合は次に一致する:

$$\left\{ \begin{bmatrix} x_{11} & 0 & x_{13} \\ 0 & x_{22} & 0 \\ x_{31} & 0 & x_{33} \end{bmatrix} \mid x_{11}, x_{13}, x_{22}, x_{31}, x_{33} \in \mathbb{C} \right\}. \quad \square$$

[4] 対角行列 $A = \text{diag}(a_1, \dots, a_n) \in D_n(\mathbb{C})$ に対して, 以下の 2 つの条件は互いに同値である:

- (a) a_1, \dots, a_n は互いに異なる. すなわち⁴ $i \neq j$ ならば $a_i \neq a_j$ である.
- (b) A と可換な任意の $X \in M_n(\mathbb{C})$ は対角行列になる. \square

1.3 行列単位

(i, j) 成分だけが 1 で他の成分がすべて 0 であるような n 次正方行列 E_{ij} を行列単位 (matrix unit) と呼ぶ⁵:

$$E_{ij} = [(i, j) \text{ 成分だけが } 1 \text{ で他はすべて } 0 \text{ の } n \text{ 次正方行列}].$$

行列単位はほとんどの成分が 0 の行列を扱うときに便利である.

第 i 成分だけが 1 で他の成分がすべて 0 であるような n 次元縦ベクトルを e_i と書くことにする:

$$e_i = [\text{第 } i \text{ 成分だけが } 1 \text{ で他はすべて } 0 \text{ の } n \text{ 次元縦ベクトル}].$$

[5] 行列単位 E_{ij} と e_i に関して次が成立している:

$$E_{ij}e_k = \delta_{jk}e_i, \quad E_{ij}E_{kl} = \delta_{jk}E_{il}. \quad \square$$

[6] n 本の m 次元縦ベクトル $a_j = [a_{ij}]_{i=1}^m$ ($j = 1, \dots, n$) を並べてできる $m \times n$ 行列 A を次のように表わす:

$$A = [a_1, \dots, a_n] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad a_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}.$$

このとき次が成立する:

$$Ae_j = a_j \quad (j = 1, \dots, n).$$

³行列 A, B が $AB = BA$ を満たしているとき A と B は可換 (commutative) であるという.

⁴「すなわち」は直前に述べたことを同値な主張を述べ直すときに用いる接続詞である。「つまり」は直前に述べたことをより「詰めて」要約する場合に用いる。「つまり」は「要するに」と同じ意味である.

⁵単位行列 (unit matrix) E と区別せよ.

すなわち行列 A への e_j の右からの積は A の第 j 列を取り出す操作になっている. さらに任意の $l \times m$ 行列 B に対して

$$BA = [Ba_1, \dots, Ba_n].$$

すなわち B の A への左からの積は A の各々の列ベクトルに B を左からかけたものに等しい. \square

1.4 置換行列

集合 $\{1, 2, \dots, n\}$ からそれ自身への全単射を $1, 2, \dots, n$ の置換 (permutation) と呼び, それら全体の集合を S_n と表わす. 置換 $\sigma \in S_n$ を次のよう表わすことがある:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

[7] $\sigma, \tau \in S_n$ に対して写像の合成 $\sigma \circ \tau$ を $\sigma\tau$ と表わすことにする. 集合 $\{1, 2, \dots, n\}$ からそれ自身への恒等写像を e と表わすことにする. 恒等写像は全単射なので $e \in S_n$ である. $\sigma \in S_n$ の逆写像を σ^{-1} と表わすことにする. 全単射の逆写像もまた全単射なので $\sigma^{-1} \in S_n$ である. このとき以下が成立する.

1. 任意の $\sigma, \tau, \rho \in S_n$ に対して $(\sigma\tau)\rho = \sigma(\tau\rho)$.
2. 任意の $\sigma \in S_n$ に対して $e\sigma = \sigma e = \sigma$.
3. 任意の $\sigma \in S_n$ に対して $\sigma\sigma^{-1} = \sigma^{-1}\sigma = e$. \square

上の問題の結果は置換全体の集合 S_n が自然に群 (group) をなすことを意味している. S_n は置換群 (permutation group) もしくは対称群 (symmetric group) と呼ばれる. 置換群 S_n には行列式 (determinant) を学んだときにすでに出会っているはずである.

[8] 互換 (transposition), 偶置換 (even permutation), 奇置換 (odd permutation), 行列式 (determinant) の定義を説明せよ. \square

置換 $\sigma \in S_n$ に対して次の行列 $P(\sigma)$ を置換行列と呼ぶ:

$$P(\sigma) = \sum_{i=1}^n E_{\sigma(i), i} = [e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}].$$

[9] すべての $\sigma \in S_3$ に対して置換行列 $P(\sigma)$ を具体的に書き下し, それぞれの行列式を計算せよ. \square

ヒント. たとえば巡回置換 $\sigma = (1, 2, 3) \in S_3$ に対して,

$$P(\sigma) = [e_2, e_3, e_1] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

よって $\det P(\sigma) = 1$ である. \square

[10] 置換行列について以下が成立している:

1. $P(\sigma)e_i = e_{\sigma(i)} \quad (\sigma \in S_n, i = 1, \dots, n).$
2. $P(\sigma\tau) = P(\sigma)P(\tau) \quad (\sigma, \tau \in S_n).$
3. $\det P(\sigma) = \text{sign } \sigma \quad (\sigma \in S_n).$ すなわち σ が偶置換, 奇置換であるかに応じて $\det P(\sigma) = 1, \det P(\sigma) = -1$ となる. \square

[11] 写像 $f: \{1, \dots, m\} \rightarrow \{1, \dots, n\}$ に対して $n \times m$ 行列 $X(f)$ を次のように定める:

$$X(f) = [e_{f(1)}, e_{f(2)}, \dots, e_{f(m)}].$$

この行列は $X(f)e_j = e_{f(j)}$ を満たしている. \square

参考 1.2 このように有限集合間の写像に対して有限サイズの行列を自然に対応させることができる. この意味で有限サイズの行列は有限集合間の写像の一般化であるとみなせる. より一般に集合, 空間, 図形, などなどあいだの写像に対して行列 (もしくは線形写像) を対応させてその性質を調べることは数学の常套手段になっている. 特に種々のコホモロジー論 (cohomology theory) は重要である. \square

1.5 確率行列

行列は確率論でも役に立つ.

実行列 A のすべての成分が 0 以上であるとき A は非負行列 (non-negative matrix) であるという. 同様に実縦ベクトル v のすべての成分が 0 以上であるとき v は非負ベクトル (non-negative vector) であるという. 非負行列と非負行列 (もしくは非負ベクトル) の積と和もまた非負行列 (もしくは非負ベクトル) である.

[12] 実 n 次正方行列 $A \in M_n(\mathbb{R})$ に関して次の 2 条件は互いに同値である:

- (a) A は可逆⁶でかつ A^{-1} は非負行列である.
- (b) 任意の $v \in \mathbb{R}^n$ に対して Av が非負ベクトルならば v 自身も非負ベクトルである. \square

ヒント. (a) \implies (b) の証明は $A^{-1}Av = v$ より易しい. (b) \implies (a) の証明は, A が可逆でないとき (b) が成立しないことおよび A が可逆で A^{-1} が非負行列でないとき (b) が成立しないことを証明すれば良い. \square

解答. (a) \implies (b): 非負行列と非負ベクトルの積もまた非負ベクトルなので Av が非負であることから $A^{-1}(Av) = v$ も非負であることが出る.

A が可逆でないとき (b) が成立しないこと: ある $u \in \mathbb{R}^n$ で $u \neq 0$ かつ $Au = 0$ を満たすものが存在する. $-u$ も同じ性質を満たしている. $\pm u$ のどちらかは非負ベクトルではない. 非負でない方を v と書くと $Av = 0$ は非負なのに v は非負ではない.

A が可逆で A^{-1} が非負行列でないとき (b) が成立しないこと: A^{-1} が非負でなければそのある列ベクトルが非負でないものである j で $v = A^{-1}e_j$ は非負でないものが存在する. しかし $Av = e_j$ は非負である. \square

⁶ A が逆行列を持つとき A は可逆 (invertible) もしくは正則 (regular) であるという.

$P = [p_{ij}] \in M_n(\mathbb{R})$ が非負行列でかつ

$$\sum_{i=1}^n p_{ij} = 1 \quad (j = 1, \dots, n)$$

を満たしているとき, P は確率行列 (stochastic matrix) であるという⁷. $x = [x_i] \in \mathbb{R}^n$ が非負ベクトルでかつ

$$\sum_{i=1}^n x_i = 1$$

を満たしているとき, x は確率ベクトルであるという.

[13] 確率行列と確率ベクトルについて以下が成立する:

1. 単位行列 E は確率行列である.
2. P が確率行列でかつ v が確率ベクトルならば Pv も確率ベクトルである.
3. P, Q が確率行列ならば PQ も確率行列である.
4. P が確率行列ならば $k = 0, 1, 2, \dots$ に対する P^k も確率行列である. \square

[14] A, B, C 地点の3ヶ所を移動する嫌な奴がいる. その嫌な奴は10分ごとに以下のような確率で今までいた場所に止まったり, 他の場所に移動したりする:

- A 地点にいるとき, 10分後に A, B, C 地点にいる確率はそれぞれ $1/3, 1/2, 1/6$.
- B 地点にいるとき, 10分後に A, B, C 地点にいる確率はそれぞれ $1/3, 1/3, 1/3$.
- C 地点にいるとき, 10分後に A, B, C 地点にいる確率はそれぞれ $1/6, 1/6, 2/3$.

現在その嫌な奴は B 地点にいる. 30分後にその嫌な奴が A, B, C 地点にそれぞれにいる確率を求めよ. 30分後にその嫌な奴と出会う確率を最小にするためには A, B, C 地点のどこに行けば良いか? \square

ヒント. 確率行列 P を次のように定める:

$$P = \begin{bmatrix} 1/3 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 1/6 & 1/3 & 2/3 \end{bmatrix}.$$

嫌な奴がある時点で A, B, C 地点にいる確率がそれぞれ a, b, c ならば次の10分後にそいつが A, B, C 地点にいる確率 a', b', c' は

$$\begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} = P \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

で計算される. よって P^3 の第2列の成分が求める確率になっている. \square

⁷ P の転置のことを確率行列と呼ぶ流儀も存在する. もしかしたらそちらの方が普通かもしれないが以下ではこの定義を採用する.

解答. P, P^2, P^3 の計算結果は次のようになる:

$$P = \frac{1}{6} \begin{bmatrix} 2 & 2 & 1 \\ 3 & 2 & 1 \\ 1 & 2 & 4 \end{bmatrix}, \quad P^2 = \frac{1}{36} \begin{bmatrix} 11 & 10 & 8 \\ 13 & 12 & 9 \\ 12 & 14 & 19 \end{bmatrix}, \quad P^3 = \frac{1}{216} \begin{bmatrix} 60 & 58 & 53 \\ 71 & 68 & 61 \\ 85 & 90 & 102 \end{bmatrix}.$$

P^3 の第 2 列の成分は上から $58/216, 68/216, 90/216$ である. それらがそれぞれ 30 分後に嫌な奴が A, B, C 地点にいる確率である. 30 分後に嫌な奴がいる確率が最も小さいのは A 地点である. \square

参考 1.3 一般に確率行列は固有値 1 を持ち, 他の固有値の絶対値は 1 以下である. しかも固有値 1 に属する固有ベクトルとして確率ベクトルを取ることができる.

たとえば上の問題の P は固有値 $1, (1 + \sqrt{2})/6, (1 - \sqrt{2})/2$ を固有値に持ち, 固有値 1 に属する固有ベクトルとして,

$$v = \begin{bmatrix} 6/23 \\ 7/23 \\ 10/23 \end{bmatrix} = \begin{bmatrix} 0.2608695652 \dots \\ 0.3043478261 \dots \\ 0.4347826087 \dots \end{bmatrix}$$

が取れる. P の 1 以外の固有値の絶対値が 1 より小さいことから

$$\lim_{k \rightarrow \infty} P^k = [v, v, v]$$

が導かれる. この結果は長時間平均で嫌な奴が A, B, C 地点にいる確率がそれぞれ 26%, 30%, 43% 程度であることを意味している. \square

1.6 Neumann 級数

複素正方行列 A に対して

$$E + A + A^2 + A^3 + \dots$$

を A の Neumann 級数 (ノイマン級数, Neumann series) と呼ぶ. ここで E は単位行列である.

[15] $A \in M_n(\mathbb{C})$ の Neumann 級数が収束するならば $E - A$ は可逆になり

$$(E - A)^{-1} = E + A + A^2 + A^3 + \dots$$

が成立する. \square

[16] $A = [a_{ij}] \in M_n(\mathbb{C})$ は $i \geq j$ ならば $a_{ij} = 0$ を満たしているとする. すなわち A は次のような形をしているとする:

$$A = \begin{bmatrix} 0 & a_{12} & \cdots & a_{n1} \\ & 0 & \ddots & \vdots \\ & & \ddots & a_{n-1,n} \\ 0 & & & 0 \end{bmatrix}.$$

このとき $A^n = 0$ が成立する. よって A の Neumann 級数は第 n 項で切れ, 有限和になる. このとき $E - A$ は可逆になり,

$$(E - A)^{-1} = E + A + A^2 + \dots + A^{n-1}.$$

が成立する. \square

1.7 Neumann 級数 (続き)

[17] 次の形の行列の逆行列を Neumann 級数を用いて求めよ:

$$A = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & x & 0 & 0 \\ 0 & 1 & y & 0 \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad \square$$

ヒント. $X = A, B, C$ に対して $N = X - E$ と置くと, $X = E + N$ であり,

$$X^{-1} = (E + N)^{-1} = E - N + N^2 - N^3 + N^4 - \dots.$$

この形の級数をも Neumann 級数と呼ぶことにする. \square

解答. $A^{-1} = \begin{bmatrix} 1 & -x \\ 0 & 1 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 1 & -x & xy - z \\ 0 & 1 & -y \\ 0 & 0 & 1 \end{bmatrix}, \quad C^{-1} = \begin{bmatrix} 1 & -x & xy & -xyz \\ 0 & 1 & -y & yz \\ 0 & 0 & 1 & -z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad \square$

[18] 以下の行列の逆行列を Neumann 級数のアイデアを用いて求めよ:

$$A = \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 2 \end{bmatrix}. \quad \square$$

ヒント. $A = D^{-1}A', D = \text{diag}(1, 2, 3, 2)$ と置くと, $A^{-1} = A'^{-1}D$. A' に Neumann 級数の方法を適用せよ. \square

解答. $A'^{-1} = \begin{bmatrix} 1 & 3 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -3 & 2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} 1 & -3/2 & 2/3 & 1/2 \\ 0 & 1/2 & -1/3 & 0 \\ 0 & 0 & 1/3 & -1 \\ 0 & 0 & 0 & 1/2 \end{bmatrix}. \quad \square$

1.8 三角行列

[19] (三角行列) 次の形の複素正方行列を (上) 三角行列 ((upper) triangular matrix) と呼ぶ:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \ddots & \vdots \\ & & \ddots & a_{n-1,n} \\ 0 & & & a_{nn} \end{bmatrix}.$$

A が可逆⁸であるための必要十分条件は $a_{ii} \neq 0$ ($i = 1, \dots, n$) の成立であることを行列式を使わずに証明せよ. \square

[20] 問題 [19] の三角行列 A の逆行列を $n = 2, 3, 4$ について求めよ. \square

参考 1.4 一般の行列の逆行列を求めるのはかなり面倒だが, 三角行列の逆行列を求めるのはずっと易しい. \square

⁸invertible. 逆行列を持つという意味.

1.9 行列の指数関数

複素 n 次正方行列 A の指数関数 $\exp A = e^A$ を次のように定義する:

$$\exp A = e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = E + A + \frac{1}{2} A^2 + \frac{1}{3!} A^3 + \frac{1}{4!} A^4 + \cdots.$$

ここで E は単位行列である⁹.

次の問題はできるだけ多くの人に解いてもらいたい.

[21] (定数係数線形常微分方程式の解) A は複素 n 次正方行列であるとし, $u_0 \in \mathbb{C}^n$ であるとする. このとき \mathbb{R} 上の \mathbb{C}^n に値を持つ関数 $u = u(t)$ に関する方程式

$$\frac{d}{dt}u = Au, \quad u(0) = u_0 \quad (*)$$

は $u(t) = e^{At}u_0$ を解に持つことを示せ. この事実を利用して $n = 2$ で

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad u_0 = e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

のとき方程式 $(*)$ を解け¹⁰. \square

ヒント. 問題 [22], [23] の結果を使う. \square

参考 1.5 実は $u(t) = e^{At}u_0$ は方程式 $(*)$ の一意的な解である. そのことは後で常微分方程式の解の存在と一意性を習ったときに証明されるだろう. \square

略解. $P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, $D = \text{diag}(-1, 3)$ と置くと $A = PDP^{-1}$, ${}^tP = P^{-1}$ なので

$$e^{At} = Pe^{Dt}P^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{3t} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{-t} + e^{3t} & -e^{-t} + e^{3t} \\ -e^{-t} + e^{3t} & e^{-t} + e^{3t} \end{bmatrix}.$$

よって $u(t) = e^{At}u_0 = e^{At}e_1 = \frac{1}{2} \begin{bmatrix} e^{-t} + e^{3t} \\ -e^{-t} + e^{3t} \end{bmatrix}$. \square

[22] (行列の指数関数の導関数) A は複素正方行列であるとする. このとき, 複素数 t の行列値関数 e^{At} は次を満たしている¹¹:

$$\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A, \quad e^{A0} = E. \quad \square$$

⁹この演習では, この定義の無限級数が複素正方行列 A に関して広義一様絶対収束するという事実や A の成分に関する (偏) 微分を項別微分によって計算できるという事実などを証明抜きで自由に用いて良い. 無限級数の収束性などについては気にせずに形式的な計算を自由に行なって良い.

¹⁰線形常微分方程式の初期値問題の解の存在と一意性を自由に用いて良い. この演習では特別に断わらない限り, 解析学で将来習うことを自由に形式的に用いて良いことにする.

¹¹複素変数 t に関する巾級数の t による形式的微分を自由に行なってかまわない.

[23] (行列の相似変換との指数函数の可換性) A, P が複素正方行列で P が可逆ならば

$$e^{PAP^{-1}} = Pe^AP^{-1}. \quad \square$$

[24] (可換な行列の和の指数函数) 2つの複素 n 次正方行列 A, B が互いに可換¹²ならば,

$$e^{A+B} = e^A e^B = e^B e^A. \quad \square$$

ヒント. $AB = BA$ であれば次の二項定理を利用できる:

$$(A+B)^k = \sum_{i=0}^k \binom{k}{i} A^i B^{k-i}.$$

ただし二項係数 $\binom{k}{i}$ は次のように定義する:

$$\binom{k}{i} = \frac{k!}{i!(k-i)!}. \quad (k, i \in \mathbb{Z}, 0 \leq i \leq k) \quad \square$$

注意 1.6 可換性の仮定は本質的である. その条件を外すとこの問題の結論は一般に成立しなくなる. たとえば次の問題を見よ. \square

[25] $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ と $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ に対して $e^{At+Bs}, e^{At}e^{Bs}, e^{Bs}e^{At}$ は互いに異なる. \square

ヒント. $e^{At} = \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix}, e^{Bs} = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}, e^{At+Bs} = \begin{bmatrix} e^t & st^{-1} \sinh t \\ 0 & e^{-t} \end{bmatrix}.$ \square

[26] $\alpha, \beta \in \mathbb{C}$ であるとし, 複素正方行列 A, B, C を

$$A = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}, \quad B = \begin{bmatrix} \alpha & 1 \\ 0 & \alpha \end{bmatrix}, \quad C = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

と定義する.

1. e^{At}, e^{Bt}, e^{Ct} を計算せよ.
2. その計算結果を用いて e^{At}, e^{Bt}, e^{Ct} の t による導函数を直接求めて, この場合に [22] の結論が確かに成立していることをチェックせよ.
3. $e^{C(t+s)} = e^{Ct}e^{Cs}$ から三角函数の加法公式を導け. \square

[27] A は複素 m 次正方行列であり, B は複素 n 次正方行列であるとし, $m+n$ 次正方行列 X を $X = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ と定める. このとき, $e^X = \begin{bmatrix} e^A & 0 \\ 0 & e^B \end{bmatrix}.$ \square

[28] 複素数 α に対して n 次正方行列 $J = J_n(\alpha)$ を次のように定める:

$$J = J_n(\alpha) = \begin{bmatrix} \alpha & 1 & & 0 \\ & \alpha & \ddots & \\ & & \ddots & 1 \\ 0 & & & \alpha \end{bmatrix} \quad (n \text{ 次正方行列}).$$

この形の行列を **Jordan ブロック (Jordan block)** と呼ぶ. e^{Jt} を計算せよ. \square

¹² A と B が可換 (commutative) であるとは $AB = BA$ が成立することである.

ヒント. 対角成分の一つ右上に 1 が並び他の成分が 0 の n 次正方行列を $N = N_n$ と書くと, $J = \alpha E + N$ である. αE と N は互いに可換なので, [24] より,

$$e^{Jt} = e^{\alpha t E} e^{tN} = e^{\alpha t} e^{tN}.$$

よって, e^{tN} を計算すればよい. たとえば

$$e^{tN_4} = \begin{bmatrix} 1 & t & \frac{1}{2}t^2 & \frac{1}{6}t^3 \\ 0 & 1 & t & \frac{1}{2}t^2 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

この結果を見れば n が一般の場合にどうなるかが推測できるだろう. \square

1.10 複素数の実 2 次正方行列による表現

[29] 複素数 $z = x + iy \in \mathbb{C}$ ($x, y \in \mathbb{R}$) に対して実 2 次正方行列 $A(z) = A(x + iy)$ を次のように定める:

$$A(z) = A(x + iy) := \begin{bmatrix} x & -y \\ y & x \end{bmatrix}.$$

このとき $z, w \in \mathbb{C}$ に対して次が成立する:

$$\begin{aligned} A(z + w) &= A(z) + A(w), & A(zw) &= A(z)A(w), & A(1) &= 1; \\ \det A(z) &= |z|^2, & \operatorname{tr} A(z) &= 2 \operatorname{Re} z, & e^{A(z)} &= A(e^z). \end{aligned} \quad \square$$

1.11 コラム: 線形代数と量子力学

拡張された意味での「線形代数」と我々の宇宙の基礎的物理法則である「量子力学」の関係について知りたい方はディラック [D] を読めば良いだろう. 量子力学をめぐる話はどれも非常に面白いので知っておいて損はない. 物理学における量子力学の発見は, 「量子化」および「古典極限」という考え方が数学的にも普遍的かつ基本的であるという事実の発見でもあった (図 1.1 を見よ).

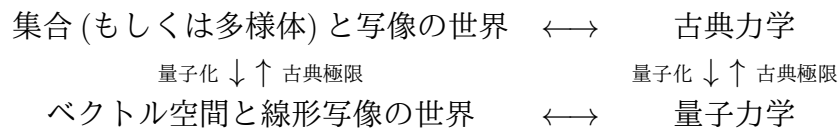


図 1.1: 古典力学的数学と量子力学的数学の対応

集合と写像の言葉で表現される数学のあらゆる対象の「量子化とは何か」すなわち「ベクトル空間と写像の世界における対応物とは何か」について考えてみるのは結構楽しい. たとえば n 次の置換群 S_n の量子化は可逆な n 次正方行列全体のなす群 GL_n であると考えることができる¹³. もちろんそれだけが正しい解答ではない.

¹³ GL_n は一般線形群 (general linear group) と呼ばれる. より正確には行列の成分が何であるかを指定してはじめて本物の群になる. たとえば $GL_n(\mathbb{R})$ は実数を成分に持つ可逆な n 次正方行列全体のなす群であり, $GL_n(\mathbb{C})$ は複素数を成分に持つ可逆な n 次正方行列全体のなす群である.

このような一見怪しげに見えることを正しく考え続けるためには論理を数学的に素早く自由にあやつる能力を身に付けることが必要である。数学において論理と直観は常に相補的であり、論理抜きの直観は役に立たないし、逆に直観抜きの論理も役に立たない。

2 行列式

なにはともあれ、様々な行列式の計算をしてみたいという方は第3節を見よ。理論的な基礎に興味のある人は第2.1節から順番に読んで行けば良い。

2.1 行列式の定義と基本性質

行列式の定義を理解したと自信を持って言えるようになるためにはこの subsection のすべての問題を何の準備もなしにいつでも解けるように自分の状態を持って行く必要がある¹⁴。そのためには行列式に関する自前のノートをきちんと作っておかなければいけない。この subsection の内容をそのために役に立てて欲しい。

定義 2.1 (行列式) n 次正方行列 $A = [a_{ij}]$ の行列式 (**determinant**) $\det A = |A|$ は次のように定義される:

$$\det A = |A| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

ここで S_n は n 次の置換群であり、 $\text{sign}(\sigma)$ は置換 σ の符号数 (signature) である¹⁵。置換の符号数はその置換が偶置換であれば 1、奇置換であれば -1 と定義される。□

[30] (2×2 , 3×3 の行列式の「たすきがけ」の公式) 上の行列式の定義に基いて次の公式を示せ:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc, \quad \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = aek + bfg + cdh - afh - bdk - ceg.$$

これらの公式では、 $+$ の項は左上から右下の方向に進む積で表わされており、 $-$ の項は右上から左下に進む積になっている。□

注意 2.2 上の問題の結果のような「たすきがけ」型の公式は 3 次以下の行列式特有のものであり、4 次以上では通用しない。次の問題が「たすきがけ」が通用しない場合になっている。□

[31] 上の行列式の定義に基いて次の公式を示せ:

$$\begin{vmatrix} x & -1 & 0 & 0 \\ 0 & x & -1 & 0 \\ 0 & 0 & x & -1 \\ d & c & b & x+a \end{vmatrix} = x^4 + ax^3 + bx^2 + cx + d. \quad \square$$

¹⁴「参考」の欄での解説は様々なことに興味を持ってもらうための「おまけ」の解説なので必ずしも理解しなくてもよい。

¹⁵ $\text{sign}(\sigma)$ の代わりに $\text{sign } \sigma$, $\text{sgn}(\sigma)$, $\text{sgn } \sigma$ と書いてもよい。

ヒント. 行列式を定義通りに計算すると生き残るのは次の4項しかないことがわかる:

$$\begin{aligned} & \text{sign} \begin{pmatrix} 1234 \\ 1234 \end{pmatrix} xxx(x+a) + \text{sign} \begin{pmatrix} 1234 \\ 1243 \end{pmatrix} xxb(-1) \\ & + \text{sign} \begin{pmatrix} 1234 \\ 1423 \end{pmatrix} xc(-1)(-1) + \text{sign} \begin{pmatrix} 1234 \\ 4123 \end{pmatrix} d(-1)(-1)(-1). \quad \square \end{aligned}$$

参考 2.3 (コンパニオン行列) 次の形の n 次正方行列のを コンパニオン行列 (同伴行列, companion matrix) と呼ぶ:

$$C(a_0, \dots, a_{n-1}) = \begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & \ddots & & \\ & & \ddots & 1 & \\ 0 & & & 0 & 1 \\ -a_{n-1} & -a_{n-2} & \cdots & -a_1 & -a_0 \end{bmatrix}.$$

コンパニオン行列 $C = C(a_0, \dots, a_{n-1})$ の特性多項式¹⁶は

$$p_C(\lambda) = \det(\lambda E - C(a_0, \dots, a_{n-1})) = \lambda^n + a_0\lambda^{n-1} + a_1\lambda^{n-2} + \cdots + a_{n-2}\lambda + a_{n-1}$$

となる. 上の問題の結果はこの公式の $n = 4$ の場合になっている.

コンパニオン行列の最小多項式は特性多項式に等しく, しかもその固有値 α に属する Jordan 細胞は唯一になることが知られている¹⁷. \square

[32] (転置行列の行列式) 行列式の定義に基いて $|{}^t A| = |A|$ を証明せよ. すなわち転置行列の行列式ともとの行列の行列式が一致することを示せ. \square

注意 2.4 この問題の結果を使えば行列式の列 (もしくは行) に関する結果から行 (もしくは列) に関する結果を導くことができる. \square

ヒント. 置換 $\sigma \in S_n$ に対して, 積の順序の並び換えによって

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} = a_{\sigma^{-1}(1)1}a_{\sigma^{-1}(2)2}\cdots a_{\sigma^{-1}(n)n}$$

が成立することがわかる¹⁸. 置換 $\sigma \in S_n$ の逆置換 σ^{-1} 全体は S_n 自身に一致するので, $\sigma \in S_n$ の全体にわたる和を $\sigma \in S_n$ に対する σ^{-1} 全体にわたる和に書き直すことができる. さらに σ の偶奇と σ^{-1} の偶奇は等しい. 以上の3つの事実を使えば容易に証明できる.

上の3つの事実が成立する理由を友人にどのように説明すれば良いか? 一般の場合についてどのように説明して良いかがわからない場合はまず $n = 3$ の場合に挑戦してみよ. \square

¹⁶一般に n 次正方行列 A の特性多項式 (characteristic polynomial) $p_A(\lambda)$ は $p_A(\lambda) = \det(\lambda E - A)$ と定義される. ここで E は n 次の単位行列である.

¹⁷「最小多項式」や「Jordan 細胞」などの用語の意味は後で **Jordan 標準形 (Jordan normal form, Jordan canonical form)** の理論を習うときに教わることになるだろう. もちろん各自が自由に自習して構わない. 数学の得意な人の特徴は学校の授業の先の勉強を勝手にやってしまうことである.

¹⁸ $i = \sigma^{-1}(j)$ と置くと $a_{i\sigma(i)} = a_{\sigma^{-1}(j)j}$.

[33] (列や行の置換) $\tau \in S_n$ で A の列を置換した行列式は A 自身の行列式の $\text{sign}(\tau)$ 倍になる:

$$\begin{vmatrix} a_{1\tau(1)} & a_{1\tau(2)} & \cdots & a_{1\tau(n)} \\ a_{2\tau(1)} & a_{2\tau(2)} & \cdots & a_{2\tau(n)} \\ \vdots & \vdots & & \vdots \\ a_{n\tau(1)} & a_{n\tau(2)} & \cdots & a_{n\tau(n)} \end{vmatrix} = \text{sign}(\tau) \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

行の置換についても同様の結果が成立する. \square

ヒント. 問題 [32] と同様の方針で証明できる. 置換 $\tau \in S_n$ を固定する. 置換 $\sigma \in S_n$ に対して, 積の順序の並び換えによって

$$a_{\sigma(1)\tau(1)} a_{\sigma(2)\tau(2)} \cdots a_{\sigma(n)\tau(n)} = a_{\sigma(\tau^{-1}(1))1} a_{\sigma(\tau^{-1}(2))2} \cdots a_{\sigma(\tau^{-1}(n))n}$$

が成立することがわかる. 置換 $\sigma \in S_n$ に対する $\sigma\tau^{-1}$ の全体は S_n 自身に一致するので, $\sigma \in S_n$ の全体にわたる和を $\sigma \in S_n$ に対する $\sigma\tau^{-1}$ 全体にわたる和に書き直すことができる. さらに $\text{sign}(\sigma\tau^{-1}) = \text{sign}(\sigma)\text{sign}(\tau)$ が成立する. 以上の3つの事実を使えばよい. \square

参考 2.5 (行列式の反対称性) 任意の置換が互換の積で表わされることに注意すれば問題 [33] の結果は「行列式は2つの列を交換すると -1 倍になり, 行についても同様である」と述べることができる.

一般に (x_1, \dots, x_n) の関数 f が

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n) \quad (\sigma \in S_n)$$

を満たしているとき, f は対称 (symmetric) であるといい,

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \text{sign}(\sigma)f(x_1, \dots, x_n) \quad (\sigma \in S_n)$$

を満たしているとき, f は反対称 (anti-symmetric) もしくは交代的 (alternative) であるという. この用語法を用いれば問題 [33] の結果を「行列式は列と行の置換について反対称である」と述べ直すことができる. \square

[34] (2つの列や行が一致する場合) 2つの列が一致する正方行列の行列式は0になる. すなわち $1 \leq k < l \leq n$ でかつ n 次正方行列 $A = [a_{ij}]$ において $a_{ik} = a_{il}$ ($i = 1, \dots, n$) が成立しているならば $|A| = 0$ となる. 行についても同様の結果が成立する. \square

ヒント 1. 問題 [33] の結果を2つの一致する列の互換に適用し, 数 d に対して $d = -d$ ならば $d = 0$ であることを使う. \square

参考 2.6 ヒント 1 では $d = -d$ ならば $2d = 0$ であるから $d = 0$ となることを使った. しかし, もしも $2 = 0$ ならば2で割ることができなくなる. 複素数や実数の世界では $2 \neq 0$ であり, 2で割ることができるので問題が生じないが, 標数が2の体の世界¹⁹で問題が生じてしまう. たとえば0と1だけからなり, 加法と乗法が

$$0 + 0 = 0, \quad 0 + 1 = 1, \quad 1 + 0 = 1, \quad 1 + 1 = 0,$$

¹⁹もしも体 K の中で p 個の1の和が0になり, p 個未満の1の和が0にならないならば, K の標数は p であるという. そのとき p は必ず素数になる. そのような p が存在しないとき K の標数は0であるという.

$$0 \cdot 0 = 0, \quad 0 \cdot 1 = 1, \quad 1 \cdot 0 = 1, \quad 1 \cdot 1 = 0$$

と定義された 2 元体 $\mathbb{F}_2 = \{0, 1\}$ の中では $2 = 1 + 1 = 0$ なので 2 で割ることができない. 標数が 2 の体はデジタル・コンピューターとの相性の良さから, 実用的にも重要であり, 標数が 2 の体にも適用できる証明をしておくことが好ましい. そこで次のヒント 2 には $2 = 0$ であっても通用する証明法を書いておく²⁰. \square

ヒント 2. $A = [a_{ij}]$ は n 次正方行列であり, $1 \leq k < l \leq n$ であり, A の第 k 列と第 l 列が等しいと仮定する. すなわち $a_{ik} = a_{il}$ ($i = 1, \dots, n$) が成立していると仮定する. このとき A の行列式の定義式の $\sigma \in S_n$ に対応する項は次のように変形される:

$$\begin{aligned} & \text{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(k)k} \cdots a_{\sigma(l)l} \cdots a_{\sigma(n)n} \\ &= \text{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(k)l} \cdots a_{\sigma(l)k} \cdots a_{\sigma(n)n} \\ &= \text{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(l)k} \cdots a_{\sigma(k)l} \cdots a_{\sigma(n)n} \\ &= -\text{sign}(\sigma\tau) a_{\sigma\tau(1)1} \cdots a_{\sigma\tau(k)k} \cdots a_{\sigma\tau(l)l} \cdots a_{\sigma\tau(n)n}. \end{aligned}$$

最後の等号で k と l を交換する互換を τ と書き, $j \neq k, l$ のとき $\sigma\tau(j) = \sigma(j)$ であることおよび $\text{sign}(\sigma\tau) = -\text{sign}(\sigma)$ であることを用いた. 最後の結果は A の行列式の定義中の $\sigma\tau$ に対応する項の -1 倍に等しい. よって行列式の定義式の中の σ に対応する項と $\sigma\tau$ に対応する項は互いにキャンセルして消えることになる.

感じがつかめなければ $n = 3$ の場合に確かめてみよ. \square

[35] (多重線形性) $A = [a_{ij}]$ は n 次正方行列であり, その k 列が

$$a_{ik} = \beta b_i + \gamma c_i$$

と表わされている²¹ならば A の行列式は次のように計算される:

$$\begin{vmatrix} a_{11} & \cdots & \beta b_1 + \gamma c_1 & \cdots & a_{1n} \\ a_{21} & \cdots & \beta b_2 + \gamma c_2 & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \beta b_n + \gamma c_n & \cdots & a_{nn} \end{vmatrix} = \beta \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix} + \gamma \begin{vmatrix} a_{11} & \cdots & c_1 & \cdots & a_{1n} \\ a_{21} & \cdots & c_2 & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & c_n & \cdots & a_{nn} \end{vmatrix}.$$

行に関しても同様のことが成立する. \square

ヒント. 次の式を見ればほとんど明らかであろう:

$$a_{\sigma(1)} \cdots (\beta b_{\sigma(k)} + \gamma c_{\sigma(k)}) \cdots a_{\sigma(n)n} = \beta a_{\sigma(1)} \cdots b_{\sigma(k)} \cdots a_{\sigma(n)n} + \gamma a_{\sigma(1)} \cdots c_{\sigma(k)} \cdots a_{\sigma(n)n}. \quad \square$$

参考 2.7 上の問題の結果は以下のように言い直すことができる. A の第 j 列を $a_j = {}^t[a_{1j}, \dots, a_{nj}]$ と書き, A を $A = [a_1, \dots, a_n]$ と表わし, 第 k 列 a_k が

$$a_k = \beta b + \gamma c, \quad b = {}^t[b_1, \dots, b_n], \quad c = {}^t[c_1, \dots, c_n]$$

²⁰標数が 2 でない場合だけに通用する簡明な証明法も知っておくことは大事なことである. ヒント 1 もヒント 2 も両方重要である. 一般的な証明法だけが重要なわけではない.

²¹ β, γ, b_i, c_i は a_{ij} が含まれている体の元であるとする.

と表わされているならば A の行列式は次のように計算される:

$$\det[a_1, \dots, \beta b + \gamma c, \dots, a_n] = \beta \det[a_1, \dots, b, \dots, a_n] + \gamma \det[a_1, \dots, c, \dots, a_n].$$

一般にベクトル $a = a_k$ の関数 f が

$$f(\beta b + \gamma c) = \beta f(b) + \gamma f(c)$$

を満たしているとき, f は線形 (**linear**) であるという. 上の問題の結果は $\det[a_1, \dots, a_n]$ が各列ベクトル a_j に関して線形であることを意味している. $(a_1, \dots, a_n) \mapsto \det[a_1, \dots, a_n]$ のように複数のベクトルの組の関数が各 a_j ごとに線形であるとき, その関数は多重線形 (**multilinear**) であるという²². この述語を用いれば上の問題の結果を「行列式は多重線形である」と一言で述べることができる. \square

[36] (ある列のスカラー倍を別の列に加えた場合) $A = [a_{ij}]$ は n 次正方行列であり, α は任意の数であり, $1 \leq k, l \leq n$, $k \neq l$ であるとし, A の第 k 列の α 倍を第 l 列に加えてできる行列の行列式と A の行列式は等しい:

$$\begin{vmatrix} a_{11} & \cdots & a_{1l} + \alpha a_{1k} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2l} + \alpha a_{2k} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nl} + \alpha a_{nk} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1l} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2l} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nl} & \cdots & a_{nn} \end{vmatrix} \quad (k \neq l).$$

すなわち, ある列のスカラー倍²³を別の列に加えても行列式は変化しない. 行についても同様のことが成立する. \square

ヒント. 問題 [35] の結果を示すべき等式の左辺の第 l 列に適用し, 問題 [34] の結果を用いる. \square

多重線形性 [35] と反対称性 [33] は行列式の最も基本的な性質である. 実はその二つの性質だけで行列式はほとんど特徴付けられてしまう. 次の問題を見よ.

[37] (行列式の特徴付け) f は n 個の n 次元縦ベクトルの組 $(a_1, \dots, a_n) \in (\mathbb{C}^n)^n$ の複素数値関数であり, 以下の条件を満たしていると仮定する:

(a) 多重線形性. 任意に $k = 1, \dots, n$ を選ぶとき, a_k が

$$a_k = \beta b + \gamma c, \quad \beta, \gamma \in \mathbb{C}, \quad b, c \in \mathbb{C}^n$$

と表わされているならば

$$f(a_1, \dots, a_k, \dots, a_n) = \beta f(a_1, \dots, b, \dots, a_n) + \gamma f(a_1, \dots, c, \dots, a_n).$$

(b) 反対称性. 任意の置換 $\sigma \in S_n$ に対して,

$$f(a_{\sigma(1)}, \dots, a_{\sigma(n)}) = \text{sign}(\sigma) f(a_1, \dots, a_n).$$

²²特に $n = 2$ ならば双線形 (**bilinear**) であるという.

²³基礎になっている体の元による積をスカラー倍という.

以上の仮定のもとで次が成立する:

$$f(a_1, \dots, a_n) = f(e_1, \dots, e_n) \det[a_1, \dots, a_n] \quad (a_1, \dots, a_n \in \mathbb{C}^n).$$

ここで e_j は第 j 成分だけが 1 で他の成分が 0 であるような n 次元縦ベクトルである. 特にもしも $f(e_1, \dots, e_n) = 1$ が成立しているならば f は行列式に等しくなる. \square

ヒント. まず, 問題 [34] のヒント 1 の方法で次を示せ²⁴:

(c) 互いに異なるある k, l で $a_k = a_l$ を満たすものが存在すれば $f(a_1, \dots, a_n) = 0$ である.

各 a_j を $a_j = \sum_{i=1}^n a_{ij} e_i$ と表わし, $f(a_1, \dots, a_n)$ に多重線形性 (a) を適用すると,

$$f(a_1, \dots, a_n) = \sum_{i_1, \dots, i_n=1}^n a_{i_1 1} \cdots a_{i_n n} f(e_{i_1}, \dots, e_{i_n}).$$

(c) より $\sigma(j) = i_j$ と定められた $\{1, 2, \dots, n\}$ からそれ自身への写像 σ が置換になっている場合以外は $f(e_{i_1}, \dots, e_{i_n}) = 0$ となる. あとは反対称性 (b) を使えば示したい結果が得られる. \square

[38] A, B が n 次の正方行列であるとき $|AB| = |A||B|$ かつ $|E| = 1$ であることを証明せよ. ここで E は n 次の単位行列である. そのことを用いてもしも A が可逆であれば $|A| \neq 0$ であつ $|A^{-1}| = |A|^{-1}$ であることを示せ²⁵. \square

ヒント. この問題の解答は [37] と非常に似ている²⁶. A の第 j 列を a_j と表わし, B の (i, j) 成分を b_{ij} と表わし, AB の第 j 列を c_j と表わすことにすると,

$$c_j = \sum_{i=1}^n a_i b_{ij} \quad (j = 1, \dots, n).$$

行列式の多重線形性 [35] より

$$|AB| = \det[c_1, \dots, c_n] = \sum_{i_1, \dots, i_n=1}^n \det[a_{i_1}, \dots, a_{i_n}] b_{i_1 1} \cdots b_{i_n n}.$$

行列式の反対称性 [33], [34] より

$$|AB| = \sum_{\sigma \in S_n} \det[a_1, \dots, a_n] \operatorname{sign}(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n} = |A||B|. \quad \square$$

²⁴問題 [34] のヒント 2 に書いてあるように標数 2 の場合にはその証明は通用しない. 標数 2 の場合に問題 [37] と同様の結果を得るためには最初から, 互いに異なるある k, l で $a_k = a_l$ を満たすものが存在すれば $f(a_1, \dots, a_n) = 0$ であることを仮定しておけばよい.

²⁵ $|A| \neq 0$ ならば A が可逆であることを示すためには余因子の概念が必要になる.

²⁶これは偶然ではない. それはなぜか?

2.2 行列式の余因子展開と Cramér の公式

定義 2.8 (余因子) n 次正方行列 $A = [a_{ij}]$ の (i, j) 余因子 (cofactor²⁷) とは A から第 i 行と第 j 列を取り除いてできる行列の行列式の $(-1)^{i+j}$ 倍のことである. すなわち A の (i, j) 余因子を Δ_{ij} と書くと

$$\Delta_{ij} = (-1)^{i+j} \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}. \quad \square$$

次の問題の結果は空気のごとく自由に使われる.

[39] (余因子の別の表現) n 次正方行列 $A = [a_{ij}]$ の (i, j) 余因子 Δ_{ij} を次のように表わすこともできる:

$$\Delta_{ij} = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{in} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}, \quad (1)$$

$$\Delta_{ij} = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,i} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}. \quad (2)$$

すなわち A の (i, j) 余因子は A の第 j 列を ${}^t[0, \dots, 1, \dots, 0]$ (第 i 成分だけが 1) で置き換えた行列の行列式および A の第 i 行を $[0, \dots, 1, \dots, 0]$ (第 j 成分だけが 1) で置き換えた行列の行列式に等しい. \square

[40] (行列式の余因子展開) n 次正方行列 $A = [a_{ij}]$ の (i, j) 余因子を Δ_{ij} と書くと

$$\sum_{k=1}^n \Delta_{ki} a_{kj} = |A| \delta_{ij}, \quad (1)$$

$$\sum_{k=1}^n a_{ik} \Delta_{jk} = |A| \delta_{ij}. \quad (2)$$

²⁷ “co” は「余」と訳され, “factor” は「因子」と訳されることが多い. したがって「余因子」は “cofactor” のほとんど直訳になっている. 一方, “determinant” は「決定するもの」という意味だが「行列式」と訳されてしまった.

すなわち Δ_{ij} を (i, j) 成分に持つ行列を Δ と書くと

$${}^t\Delta A = A {}^t\Delta = |A|E.$$

ここで E は n 次の単位行列である. したがって, もしも $|A| \neq 0$ ならば A は可逆であり,

$$A^{-1} = |A|^{-1} {}^t\Delta$$

が成立する. 以下において Δ を A の余因子行列と呼ぶことにする. \square

注意 2.9 上の問題 [40] の結果は極めて有用である!

たとえば Cayley-Hamilton の定理 (第 2.3 節) を証明するためにも役に立つ. 他にも様々な応用がある. \square

ヒント. (1) を証明するためには, Δ_{ij} の表示として問題 [39] の (1) を採用し, 行列式の列に関する多重線形性 [35] を右辺を左辺に変形する形で使い, 最後に [34] を使う. (2) の証明も問題 [39] の (2) を使えば同様である. \square

注意 2.10 問題 [40] と問題 [38] の結果を合わせると正方行列 A が可逆であることと $|A| \neq 0$ であることは同値であることがわかる²⁸.

[41] (2×2 , 3×3 の場合) A, B はそれぞれ次のように表わされた 2 次および 3 次の正方行列であるとする:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}.$$

$|A| \neq 0$, $|B| \neq 0$ のとき, 問題 [40] の結果を A, B に適用することによって次が成立することを確かめよ:

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

$$B^{-1} = \frac{1}{aek + bfg + cdh - afh - bdk - ceg} \begin{bmatrix} ek - fh & ch - bk & bf - ce \\ fg - dk & ak - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{bmatrix}.$$

さらに実際にこれらが A, B の逆行列になっていることを直接の計算によって確かめよ. \square

[42] n 次正方行列 A の余因子行列を Δ と書くと $|\Delta| = |A|^{n-1}$. \square

ヒント. $A {}^t\Delta = |A|E$ の両辺の行列式を取ってみよ. \square

上の問題のヒントの論法を少し一般化すると次の問題の結果が得られる.

²⁸ここでは体の元を成分に持つ行列を考えているのでこうなる. もしもより一般的に可換環の元を成分に持つ行列を考えた場合には, A が可逆であることと $|A|$ が可逆であることが同値になる. ((可換) 体とは 0 でない元がすべて可逆であるような可換環のことである.)

[43] (Jacobi の公式) n 次正方行列 $A = [a_{ij}]$ において次の公式が成立している:

$$\begin{vmatrix} \Delta_{k+1,k+1} & \cdots & \Delta_{n,k+1} \\ \vdots & & \vdots \\ \Delta_{k+1,n} & \cdots & \Delta_{n,n} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} |A|^{n-k-1}$$

ここで Δ_{ij} は A の (i, j) 余因子である. $k = 0$ のとき右辺に表われる空な行列式は 1 に等しいと約束しておく. 特に $k = n - 2$ のとき,

$$\Delta_{n-1,n-1}\Delta_{n,n} - \Delta_{n,n-1}\Delta_{n-1,n} = \begin{vmatrix} a_{11} & \cdots & a_{1,n-2} \\ \vdots & & \vdots \\ a_{n-2,1} & \cdots & a_{n-2,n-2} \end{vmatrix} |A|. \quad \square$$

ヒント. $\sum_{k=1}^n a_{ik}\Delta_{jk} = |A|\delta_{ij}$ より,

$$A \begin{bmatrix} 1 & 0 & \Delta_{k+1,1} & \cdots & \Delta_{n,1} \\ & \ddots & \vdots & & \vdots \\ 0 & 1 & \Delta_{k+1,k} & \cdots & \Delta_{n,k} \\ 0 & \cdots & 0 & \Delta_{k+1,k+1} & \cdots & \Delta_{n,k+1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \Delta_{k+1,n} & \cdots & \Delta_{n,n} \end{bmatrix} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,k} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,k} & 0 & \cdots & 0 \\ a_{k+1,1} & \cdots & a_{k+1,k} & |A| & & 0 \\ \vdots & & \vdots & & \ddots & \\ a_{n,1} & \cdots & a_{n,k} & 0 & & |A| \end{bmatrix}$$

この等式の両辺の行列式を取ってみよ. \square

注意 2.11 この問題のヒントの方針では最後に「両辺を $|A|$ で割る」操作が必要になる. そのとき厳密には $|A| \neq 0$ を仮定しなければいけない. しかし, 上の問題の場合には $|A| \neq 0$ を仮定して証明された公式が $|A| = 0$ の場合にも成立することがわかる.

その理由は示すべき公式の両辺が行列 A の成分 a_{ij} たちの多項式になっているからだ. 実際, 余因子 Δ_{ij} は a_{ij} たちの $n - 1$ 次の多項式であるから, それらの $n - k$ 次の行列式 $|\Delta_{ij}|_{i,j=k+1,\dots,n}$ は a_{ij} たちの $(n - 1)(n - k)$ 次の多項式であり, k 次の行列式 $|a_{ij}|_{i,j=1,\dots,k}$ は a_{ij} たちの k 次の多項式であり, $|A|^{n-k-1}$ は a_{ij} たちの $n(n - k - 1)$ 次の多項式である. $((n - 1)(n - k) = k + n(n - k - 1))$ なので両辺の次数が等しいことに注意せよ.)

厳密には [43] の公式は次のような方針で証明される.

Step 1. まず, 行列 A の成分 a_{ij} を数ではなく文字だとみなし, a_{ij} たちの有理函数²⁹全体のなす体の中ですべての計算を行なうことにする. そのとき $|A|$ は文字式として 0 でないので $|A|$ で割るという操作を自由に行なって構わない. したがって, 問題 [43] のヒントの方針によって行列 A の成分を数ではなく文字とみなした場合に関する公式が証明される.

Step 2. 次に, Step 1 で証明された公式の両辺は A の成分の多項式 (分母がない!) であることに注目する. 一般に有理函数の中の文字には自由に数を代入できない. なぜならば有理函数の分母が数の代入によって 0 になってしまうかもしれないからだ³⁰. しかし, 多項式の中の文字には任意の数を代入できる. したがって, Step 1 で証明された公式の両辺の a_{ij} に任意の数を代入して得られた等式も成立することがわかる.

一般に, 両辺が多項式であるような等式の証明では, 「0 で割ってはいけないという理由で生じる場合分け」を回避できることが多い. \square

²⁹多項式分の多項式を有理函数と呼ぶ.

³⁰たとえば x, y の有理函数 $f(x, y) = (x + y)/(x - y)$ の (x, y) には $(1, 1)$ を代入できない.

参考 2.12 上の問題の Jacobi の公式の $k = n - 2$ の場合は実対称行列 (同じことだが実対称形式) の符号数を主小行列式を用いて決定する方法の基礎付けに使われる. 高木 [Tkg1] 第 9 章第 61 節の定理 9.5 (295 頁) を見よ. \square

[44] (**Cramér の公式**) $A = [a_{ij}]$ は可逆な n 次正方行列であるとし, $b = [b_i]$ は n 次元縦ベクトルであるとする. このとき以下が成立する.

1. x_1, \dots, x_n に関する連立一次方程式

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

の解は次のように表わされる:

$$x_j = \frac{\begin{vmatrix} a_{11} & \cdots & \overset{j}{b_1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}} \quad (j = 1, \dots, n).$$

ここで右辺の分子は A の第 j 列を縦ベクトル b で置き換えてできる行列の行列式であり, 分母は A 自身の行列式である.

2. x_1, \dots, x_n に関する連立一次方程式

$$\begin{aligned} x_1a_{11} + x_2a_{21} + \cdots + x_na_{n1} &= b_1, \\ x_1a_{12} + x_2a_{22} + \cdots + x_na_{n2} &= b_2, \\ &\dots\dots\dots \\ x_1a_{1n} + x_2a_{2n} + \cdots + x_na_{nn} &= b_n, \end{aligned}$$

の解は次のように表わされる:

$$x_i = \frac{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ {}^i)b_1 & \cdots & b_n \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}} \quad (i = 1, \dots, n).$$

ここで右辺の分子は A の第 i 行を縦ベクトル tb で置き換えてできる行列の行列式であり, 分母は A 自身の行列式である.

以上の解の公式を **Cramér** の公式 (クラメールの公式) と呼ぶ. \square

ヒント. $x = {}^t[x_1, \dots, x_n]$ と置くと, 1 の一次方程式は行列を用いて次のように表わされる:

$$Ax = b.$$

同様に 2 の一次方程式は次のように表わされる:

$${}^t_x A = {}^t_b.$$

よって転置行列を考えることによって前者に関する結果から後者に関する結果を導ける.

A は可逆であると仮定してあるので, 1 の一次方程式の解は $x = A^{-1}b$ と表わされる. 問題 [40] の結果より, $A^{-1} = |A|^{-1} {}^t\Delta$ であるから $x = |A|^{-1} {}^t\Delta b$ である. よって $|A|x_j$ は以下のように計算される:

$$|A|x_j = \sum_{i=1}^n \Delta_{ij} b_i = \sum_{i=1}^n \begin{vmatrix} a_{11} & \cdots & \overset{j}{0} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & \cdots & b_i & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & \overset{j}{b_1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}.$$

二番目の等号は問題 [39] の (1) の公式を用いた. これで 1 が示された. \square

[45] (2×2 , 3×3 の場合) A, B はそれぞれ次のように表わされた 2 次および 3 次の正方行列であるとする:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}.$$

$|A| \neq 0, |B| \neq 0$ のとき, 問題 [44] の結果を A, B に適用すると, x, y もしくは x, y, z に関する連立一次方程式たち

$$\begin{cases} ax + by = p, \\ cx + dy = q; \end{cases} \quad \begin{cases} ax + by + cz = p, \\ dx + ey + fz = q, \\ gx + hy + kz = r \end{cases}$$

の解がそれぞれ次のように表わされることがわかる:

$$x = \frac{\begin{vmatrix} p & b \\ q & d \end{vmatrix}}{|A|}, \quad y = \frac{\begin{vmatrix} a & p \\ c & q \end{vmatrix}}{|A|};$$

$$x = \frac{\begin{vmatrix} p & b & c \\ q & e & f \\ r & h & k \end{vmatrix}}{|B|}, \quad y = \frac{\begin{vmatrix} a & p & c \\ d & q & f \\ g & r & k \end{vmatrix}}{|B|}, \quad z = \frac{\begin{vmatrix} a & b & p \\ d & e & q \\ g & h & r \end{vmatrix}}{|B|}.$$

これらが実際に解になっていることを直接の計算によって確かめよ. ただしできるだけ工夫して要領よくチェックせよ. \square

参考 2.13 Cramér の公式は連立一次方程式の解を行列式を用いてコンパクトに表示する公式である. 行列式による表示が可能であるという事実は理論的に重要になることが多い. その意味で Cramér の公式は有用な公式である.

たとえば第 2.4 節にあるように行列の Gauss 分解を行列式を用いて表示できるという事実は Cramér の公式の直接的な帰結である.

しかし, 行列式の数値計算には大変な手間がかかることが知られている. したがって連立一次方程式の解を数値的に求めるために Cramér の公式を用いるのは賢い考え方ではない. 素朴な掃き出し法のような他のアルゴリズムを用いた方が圧倒的に効率が良いことが知られている. \square

2.3 Cayley-Hamilton の定理

[46] (2×2 行列に関する Cayley-Hamilton の定理) 2×2 行列 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ に対して λ の多項式 $p_A(\lambda)$ を

$$p_A(\lambda) = \det(\lambda E - A) = \begin{vmatrix} \lambda - a & -b \\ -c & \lambda - d \end{vmatrix}.$$

と定める. ここで E は 2 次の単位行列である. $p_A(\lambda)$ を A の特性多項式 (characteristic polynomial) と呼ぶ. このとき以下が成立することを直接的な計算で確かめよ:

1. $p_A(\lambda) = \lambda^2 - \text{tr}(A)\lambda + |A|$.
2. $p_A(A) = A^2 - \text{tr}(A)A + |A|E = 0$.

$p_A(\lambda)$ の λ に A を代入するとき定数項 $|A|$ には単位行列 E をかけることを忘れないように注意せよ. \square

[47] 2×2 行列の Cayley-Hamilton の定理を用いて以下の行列の k 乗 ($k = 0, 1, 2, 3, \dots$) を計算せよ:

$$A = \begin{bmatrix} x & ax \\ y & ay \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}.$$

C^k, D^k については k が偶数と奇数の場合に分けて結果を書け. \square

略解. $|A| = 0$ なので $r = x + ay$ と置くと $A^2 = \text{tr}(A)A = rA$ である. よって $k = 0, 1, 2, \dots$ に対して

$$A^k = r^{k-1}A = \begin{bmatrix} x(x+ay)^{k-1} & ax(x+ay)^{k-1} \\ y(x+ay)^{k-1} & ay(x+ay)^{k-1} \end{bmatrix}.$$

B^k は $x = 1, y = -2, a = -2$ の場合の A^k に等しい. よって

$$B^k = \begin{bmatrix} 5^{k-1} & -2 \cdot 5^{k-1} \\ -2 \cdot 5^{k-1} & 4 \cdot 5^{k-1} \end{bmatrix}.$$

$\text{tr}(C) = 0$ なので $s = a^2 + bc$ と置くと $C^2 = -|C|E = sE$ である. よって $l = 0, 1, 2, \dots$ に対して

$$C^{2l} = s^l E = \begin{bmatrix} (a^2 + bc)^l & 0 \\ 0 & (a^2 + bc)^l \end{bmatrix}, \quad C^{2l+1} = s^l C = \begin{bmatrix} a(a^2 + bc)^l & b(a^2 + bc)^l \\ c(a^2 + bc)^l & -a(a^2 + bc)^l \end{bmatrix}.$$

D^k は $a = 1, b = 1, c = 2$ の場合の C^k に等しい. よって

$$D^{2l} = \begin{bmatrix} 3^l & 0 \\ 0 & 3^l \end{bmatrix}, \quad D^{2l+1} = \begin{bmatrix} 3^l & 3^l \\ 2 \cdot 3^l & -3^l \end{bmatrix}. \quad \square$$

[48] (3×3 行列に関する **Cayley-Hamilton** の定理) 3 次正方行列 $A = [a_{ij}]_{i,j=1}^3$ に対して λ の多項式 $p_A(\lambda)$ を

$$p_A(\lambda) = \det(\lambda E - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix}$$

と定める. ここで E は 3 次の単位行列である. $p_A(\lambda)$ を A の特性多項式 (**characteristic polynomial**) と呼ぶ. このとき以下が成立することを直接的な計算で確かめよ:

1. a, b, c を

$$a = \text{tr}(A) = a_{11} + a_{22} + a_{33},$$

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

$$c = |A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}$$

と定めると $p_A(\lambda) = \lambda^3 - a\lambda^2 + b\lambda - c$.

2. $p_A(A) = A^3 - aA^2 + bA - cE = 0$. \square

参考 2.14 (一般の **Cayley-Hamilton** の定理) 一般に任意の n 次正方行列 $A = [a_{ij}]_{i,j=1}^n$ に対してその特性多項式 $p_A(\lambda)$ を

$$p_A(\lambda) = \det(\lambda E - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & \lambda - a_{nn} \end{vmatrix}$$

と定める. このとき $p_A(A) = 0$ が成立する. この結果を **Cayley-Hamilton** の定理と呼ぶ. (**Hamilton-Cayley** の定理と呼ぶ場合もある.)

Cayley-Hamilton の定理の証明として次は誤りであることに注意せよ:

$$p_A(A) = \det(AE - A) = \det(A - A) = \det 0 = 0 \quad (\text{これは誤り}).$$

どこが誤りであるかを理解するためには記号に騙されないように注意しなければならない. $p_A(A)$ は行列である. $AE - A$ も行列である. しかし $\det(AE - A)$ は数である. $p_A(A) = \det(AE - A)$ という計算は左辺が行列で右辺が数なのでナンセンスである.

ついでに述べておけば, 「 $\det 0$ 」の 0 は行列のゼロであるが, その次の「 $= 0$ 」の 0 は数のゼロである. この2つの「 0 」は同じ記号で書かれているが意味が違うことに注意せよ. 違うものを同じ記号で表わすことは数学において結構あるので文脈には気を付けなければいけない. この演習では行列のゼロもベクトルのゼロも単に「 0 」と書くことが多い. \square

一般の **Cayley-Hamilton** の定理の直接的証明. $A = [a_{ij}]$ は n 次正方行列であるとし, その特性多項式を $p_A(\lambda) = \det(\lambda E - A)$ と表わす. $\lambda E - A$ の (i, j) 余因子を $f_{ij}(\lambda)$ と書くと,

$$p_A(\lambda)\delta_{ik} = \sum_{j=1}^n f_{ij}(\lambda)(\delta_{kj}\lambda - a_{kj}).$$

この等式の両辺は λ の多項式なので λ に A を代入できる:

$$p_A(A)\delta_{ik} = \sum_{j=1}^n f_{ij}(A)(\delta_{kj}A - a_{kj}E).$$

さらにこの等式の両辺を e_k に³¹左から作用させて $k = 1, \dots, n$ について和を取ると,

$$p_A(A)e_i = \sum_{j=1}^n f_{ij}(A)\left(Ae_j - \sum_{k=1}^n a_{kj}e_k\right) = 0.$$

最後の等号は $Ae_j = \sum_{k=1}^n e_k a_{kj}$ から出る. よって $p_A(A) = 0$ である. \square

[49] 上の証明の細部を埋め, 黒板を用いて詳しく説明せよ. \square

参考 2.15 上に示した Cayley-Hamilton の定理の直接的証明の利点は行列式の余因子による展開以外に何も用いていないことである. そのおかげで上の証明法は任意の可換環上でも通用する.

そのおかげで Cayley-Hamilton の定理は可換環論における**中山の補題 (Nakayama-Azumaya-Krull (NAK) の補題**と呼ばれる場合もある) の証明に役に立つ. たとえば松村 [M] pp.9–12, 堀田 [H3] pp.37–39, リード [R] pp.46–49 を見よ³².

Cayley-Hamilton の定理の証明法には上に示したものの以外に少なくとも二種類の方法がある. 一つ目は行列係数の多項式の剰余定理を用いる方法であり, たとえば杉浦 [Sg] の 65–66 頁にある. 二つ目は以下に示すように行列の三角化可能性を用いる方法である. \square

[50] 複素 n 次正方行列 A と複素 n 次可逆行列 P と λ の複素係数多項式 $f(\lambda)$ に対して以下が成立する³³:

1. $p_{PAP^{-1}}(\lambda) = p_A(\lambda)$.
2. $f(PAP^{-1}) = Pf(A)P^{-1}$. \square

ヒント. 1 は $|AB| = |A||B|$ を使う. 2 は $f(\lambda)$ を $f(\lambda) = a_N\lambda^N + a_{N-1}\lambda^{N-1} + \dots + a_1\lambda + a_0$ ($a_i \in \mathbb{C}$) と表わして λ に A を代入して, $(PAP^{-1})^k = PA^kP^{-1}$ を用いればよい. \square

³¹ e_k は第 k 成分だけが 1 で他の成分が 0 であるような n 次元縦ベクトル.

³²可換とか限らない環における中山の補題を Cayley-Hamilton の定理を使わずに証明する流儀もある. たとえば服部 [Ht] pp.93–97 を参照せよ.

³³ここでは複素行列と複素係数多項式に関して結果を述べているが, 実際には任意の体上の行列と多項式についても同様の結果が成立する.

[51] (複素正方行列の三角化可能性) A は複素 n 次正方行列であるとする³⁴. A の特性多項式 $p_A(\lambda) = \det(\lambda E - A)$ の互いに異なる根の全体が $\alpha_1, \dots, \alpha_r$ であり, $p_A(\lambda)$ が次のように表わされているとする:

$$p_A(\lambda) = \det(\lambda E - A) = (\lambda - \alpha_1)^{n_1} \cdots (\lambda - \alpha_r)^{n_r}.$$

このとき, ある可逆な複素 n 次正方行列 P で $P^{-1}AP$ が上三角行列になり, しかも $P^{-1}AP$ の対角部分が特性多項式の根を重複を含めてすべて並べた $\text{diag}(\overbrace{\alpha_1, \dots, \alpha_1}^{n_1}, \dots, \overbrace{\alpha_r, \dots, \alpha_r}^{n_r})$ (各 α_i が n_i 個ずつ順番に並ぶ) に等しくなるものが存在する. \square

ヒント. n に関する数学的帰納法. $\alpha = \alpha_1$ のとき固有値 α を持つ A の固有ベクトル $v \neq 0$ が存在する. v を含む基底 $p_1 = v, p_2, \dots, p_n$ が取れる³⁵. このとき, $P = [p_1, \dots, p_n]$ と置くと, $P^{-1}AP$ は次の形になる:

$$P^{-1}AP = \begin{bmatrix} \alpha & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{bmatrix}.$$

行列 $B = [b_{ij}]_{2 \leq i, j \leq n}$ に帰納法の仮定を用いよ. \square

[52] 問題 [50], [51], [16] の結果を用いて複素正方行列に関する Cayley-Hamilton の定理を証明せよ³⁶. \square

ヒント. 問題 [50] より $p_{P^{-1}AP}(\lambda) = p_A(\lambda)$ であり, $p_A(P^{-1}AP) = P^{-1}p_A(A)P$ であるので, A は問題 [51] における $P^{-1}AP$ の形をしていると仮定してよい. すなわち, A は上三角行列でかつ A の対角線上には対角成分がすべて α_i であるような n_i 次の上三角行列たちが並んでいると仮定してよい. そのとき $(A - \alpha_j E)^{n_j}$ の対角線上には対角成分がすべて $(\alpha_i - \alpha_j)^{n_i}$ であるような n_i 次上三角行列たちが並ぶ. 問題 [16] の結果より, 対角線上の i 番目のブロックの n_i 乗は 0 になる. 実はこのことだけから $(A - \alpha_j E)^{n_j}$ を $j = 1, \dots, r$ についてかけ合わせた結果が零行列になることを示せる. たとえば $r = 4$ の場合は

$$\begin{aligned} & \begin{bmatrix} 0 & * & * & * \\ & * & * & * \\ & & * & * \\ & & & * \end{bmatrix} \begin{bmatrix} * & * & * & * \\ & 0 & * & * \\ & & * & * \\ & & & * \end{bmatrix} \begin{bmatrix} * & * & * & * \\ & * & * & * \\ & & 0 & * \\ & & & * \end{bmatrix} \begin{bmatrix} * & * & * & * \\ & * & * & * \\ & & * & * \\ & & & 0 \end{bmatrix} \\ = & \begin{bmatrix} 0 & 0 & * & * \\ & 0 & * & * \\ & & * & * \\ & & & * \end{bmatrix} \begin{bmatrix} * & * & * & * \\ & * & * & * \\ & & 0 & * \\ & & & * \end{bmatrix} \begin{bmatrix} * & * & * & * \\ & * & * & * \\ & & * & * \\ & & & 0 \end{bmatrix} \end{aligned}$$

³⁴ より一般に代数閉体の元を成分に持つ行列を考えても良い.

³⁵ 実はより強く v として単位ベクトルを取り, $p_1 = v, p_2, \dots, p_n$ として正規直交基底を取れる. このことを使えば P としてユニタリ行列を取れることもわかる.

³⁶ 複素数体を任意の代数閉体で置き換えてもこの証明法は通用する. 任意の体は代数閉体に埋め込めるので結果的にこの証明法は任意の体上で通用することになる.

$$\begin{aligned}
&= \begin{bmatrix} 0 & 0 & 0 & * \\ & 0 & 0 & * \\ & & 0 & * \\ & & & * \end{bmatrix} \begin{bmatrix} * & * & * & * \\ & * & * & * \\ & & * & * \\ & & & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix} \cdot \square
\end{aligned}$$

2.4 Gauss 分解

可逆行列の Gauss 分解は Cramér の公式の良い応用先になっている。以下そのことを説明しよう。

定義 2.16 (Gauss 分解) 可逆な n 次正方行列 $A = [a_{ij}]$ に対して, ある可逆な n 次正方行列 $W = [w_{ij}]$, $Z = [z_{ij}]$ で

$$w_{ii} = 1, \quad w_{ij} = 0 \quad (i < j), \quad z_{ij} = 0 \quad (i > j)$$

および $A = W^{-1}Z$ を満たす行列が存在するとき³⁷, A は **Gauss 分解可能**であると言い, (W, Z) は A の **Gauss 分解**であるという。□

参考 2.17 (Gauss 分解と可積分系) 古典および量子可積分系および Painlevé 方程式の理論やその一般化としてのモノドロミー保存変形の理論では行列を下三角と上三角の積で表現する Gauss 分解が基本的な役目を果たしている。興味のある方は高崎 [Tk] や野海 [Nm] を参照せよ。量子群における普遍 R 行列は実は Gauss 分解の量子化を与えている。量子群に関しては神保 [J] や谷崎 [Tn] およびその参考文献を参照せよ。Gauss 分解のような素朴な対象であっても研究すべきことはまだ残っているように思われる。□

[53] (Gauss 分解の一意性) Gauss 分解は (もしも存在するならば) 一意である。□

ヒント. (W, Z) と (W', Z') は共に A の Gauss 分解であるとする。このとき, $A = W^{-1}Z = W'^{-1}Z'$ であるから $W'W^{-1} = Z'Z^{-1}$ である。左辺の $W'W^{-1}$ は対角成分がすべて 1 の下三角行列になり, 右辺の $Z'Z^{-1}$ は可逆な上三角行列になる。したがって $W'W^{-1} = Z'Z^{-1}$ は単位行列になる。□

次の問題の結果は Cramér の公式から得られる。

[54] (Gauss 分解の行列式表示) A が Gauss 分解可能であるための必要十分条件は

$$\begin{vmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{vmatrix} \neq 0 \quad (i = 1, 2, \dots, n) \quad (*)$$

³⁷ これは A が対角成分がすべて 1 の下三角行列 W と対角成分がどれも 0 でないような上三角行列 Z によって $A = W^{-1}Z$ と表わされるということである。

が成立することであり³⁸, そのとき $(W, Z) = ([w_{ij}], [z_{ij}])$ を A の Gauss 分解とすると W, Z の非自明な成分は以下のように表わされる:

$$w_{ij} = (-1)^{i+j} \frac{\begin{vmatrix} a_{11} & \cdots & a_{1,i-1} \\ \vdots & \ddots & \vdots \\ \text{第 } j \text{ 行を削る} \\ \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{i,i-1} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1,i-1} \\ \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,i-1} \end{vmatrix}} \quad (i = 2, 3, \dots, n, j = 1, 2, \dots, i-1), \quad (\text{a})$$

$$z_{ij} = \frac{\begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & a_{1j} \\ \vdots & & \vdots & \vdots \\ a_{i1} & \cdots & a_{i,i-1} & a_{ij} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1,i-1} \\ \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,i-1} \end{vmatrix}} \quad (i = 1, 2, \dots, n, j = i, i+1, \dots, n). \quad (\text{b})$$

ただし $i = 1$ のとき z_{ij} の表示における分母の行列式は 1 であると約束しておく. \square

ヒント. 条件(*)の必要性の証明. A が $A = W^{-1}Z$ と Gauss 分解可能であると仮定する. そのとき W^{-1} は対角成分がすべて 1 の下三角行列であり, $Z = [z_{ij}]$ は対角成分がどれも 0 でないような上三角行列である. よって $i = 1, 2, \dots, n$

$$\begin{bmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ * & & 1 \end{bmatrix} \begin{bmatrix} z_{11} & & * \\ & \ddots & \\ 0 & & z_{ii} \end{bmatrix}$$

が成立することがわかる. 右辺の行列式は $z_{11} \cdots z_{ii} \neq 0$ に等しいので左辺の行列式も 0 にならないことがわかる.

条件(*)の十分性の証明. 条件(*)を仮定する. $W = [w_{ij}]$ は対角成分がすべて 1 であるような下三角行列であり, $Z = [z_{ij}]$ は上三角行列であると仮定する. $A = W^{-1}Z$ は $WA = Z$ と同値であり, $WA = Z$ は W と Z の成分に関する連立一次方程式とみなせる. 実際 $WA = Z$ の第 i 行目部分は

$$[\overbrace{w_{i1}, \dots, w_{i,i-1}}^{i-1}, 1, \overbrace{0, \dots, 0}^{n-i}]A = [\overbrace{0, \dots, 0}^{i-1}, \overbrace{z_{ii}, z_{i,i+1}, \dots, z_{in}}^{n-i+1}]$$

であり, さらにこの方程式を第 $1, \dots, i-1$ 成分に関する部分と第 $i, i+1, \dots, n$ 成分に関する部分に分割するとそれぞれ以下と同値になる:

$$[w_{i1}, \dots, w_{i,i-1}] \begin{bmatrix} a_{11} & \cdots & a_{1,i-1} \\ \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,i-1} \end{bmatrix} = -[a_{i1}, \dots, a_{i,i-1}], \quad (1)$$

³⁸ $|A| \neq 0$ は A が可逆であるという仮定より常に成立している.

$$[w_{i1}, \dots, w_{i,i-1}, 1] \begin{bmatrix} a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{ii} & \cdots & a_{in} \end{bmatrix} = [z_{ii}, \dots, z_{in}]. \quad (2)$$

前者の (1) は w_{ij} ($i > j$) に関する連立一次方程式の形をしており, 後者の (2) は w_{ij} たちで z_{ij} たちを表わす式になっている.

条件 (*) より, (1) に Cramér の公式 (問題 [44] の 2 の公式) を適用することができる. よって $i > j$ のとき

$$w_{ij} = - \frac{\begin{vmatrix} a_{11} & \cdots & a_{1,i-1} \\ \vdots & & \vdots \\ j) a_{i1} & \cdots & a_{i,i-1} \\ \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,i-1} \end{vmatrix}}{(A \text{ の左上の } (i-1) \times (i-1) \text{ 部分の行列式})}.$$

分子の行列式における $[a_{i1}, \dots, a_{i,i-1}]$ の行を最下段まで落とせば $(-1)^{(i-1)-j}$ の因子が出る. それとある -1 の因子をかけ合わせると $(-1)^{i+j}$ に等しい. これで w_{ij} の表示 (a) が得られた.

公式 (a) より, w_{ik} の表示における $(-1)^{i+k}$ と分子の行列式の積は行列 A の左上の $i \times i$ 部分の (k, i) 余因子に等しく, w_{ik} の表示における分母は (i, i) 余因子に等しい. よって $i \leq j$ のとき (2) より,

$$z_{ij} = \sum_{k=1}^{i-1} w_{ik} a_{kj} + a_{ij} = \frac{\sum_{k=1}^i (A \text{ の左上の } i \times i \text{ 部分の } (k, i) \text{ 余因子}) \cdot a_{kj}}{(A \text{ の左上の } (i-1) \times (i-1) \text{ 部分の行列式})}.$$

この式の分子は A の左上の $i \times (i-1)$ 部分の右に列 ${}^t[a_{1j}, \dots, a_{ij}]$ を追加して得られる行列の行列式に等しい. これで z_{ij} の表示 (b) が得られた.

公式 (b) に条件 (*) を適用すれば $z_{ii} \neq 0$ であることがわかる. これで A の Gauss 分解の存在が示されたことになる. \square

参考 2.18 問題 [54] の結果は条件 (*) が成立していればある下三角行列 W を A の左側からかけて $Z = WA$ が上三角行列になるようにできることを意味している. よって, そのとき $b = {}^t[b_1, \dots, b_n]$ が与えられたならば, $x = {}^t[x_1, \dots, x_n]$ に関する一次方程式 $Ax = b$ は三角行列 Z が係数行列であるような一次方程式 $Zx = Wb$ と同値になる. 三角行列が係数行列であるような一次方程式は簡単に解ける. たとえば $z_{ii} \neq 0$ のとき次の方程式を解いてみよ:

$$\begin{aligned} z_{11}x_1 + z_{12}x_2 + z_{13}x_3 &= c_1, \\ z_{22}x_2 + z_{23}x_3 &= c_2, \\ z_{33}x_3 &= c_3. \quad \square \end{aligned}$$

[55] (2×2 , 3×3 の場合) 次の行列を Gauss 分解せよ:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}. \quad \square$$

[56] A は $m+n$ 次正方行列であるとし, W, Z は次のような形の可逆な $m+n$ 次正方行列であるとする:

$$W = \begin{bmatrix} 1_m & 0 \\ W' & 1_n \end{bmatrix}, \quad Z = \begin{bmatrix} Z' & Z'' \\ 0 & Z''' \end{bmatrix}.$$

ここで $1_m, 1_n$ はそれぞれ m 次および n 次の単位行列であり, W' は $n \times m$ 行列, Z' は $m \times m$ 行列, Z'' は $m \times n$ 行列, Z''' は $n \times n$ 行列である. このような W, Z を適切に選んで $A = W^{-1}Z$ とできるための必要十分条件を求め, W, Z の A による表示を求めよ. \square

ヒント. Gauss 分解の場合とまったく同様である. この場合の方が Gauss 分解の場合よりずっと簡単になる. \square

略解. A も W, Z と同様に $A = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$ と分割しておく. $W^{-1} = \begin{bmatrix} 1_m & 0 \\ -W' & 1_n \end{bmatrix}$ であるから, $W^{-1}Z = \begin{bmatrix} Z' & Z'' \\ -W'Z' & -W'Z'' + Z''' \end{bmatrix}$. Z が可逆であるための必要十分条件は Z' と Z''' が可逆になることである. そのとき $A = W^{-1}Z$ と次は同値である:

$$Z' = P, \quad Z'' = Q, \quad W' = -RZ'^{-1} = -RP^{-1}, \quad Z''' = S + W'Z'' = S - RP^{-1}Q.$$

したがって $A = W^{-1}Z$ と分解できるための必要十分条件は P と $S - RP^{-1}Q$ が可逆になることである. \square

2.5 行列式と面積や体積の関係

行列式の幾何的な意味を理解するためには行列式と面積や体積の関係について知っておくことが必要である. 私の経験では講義ではこの点をほとんど説明しない場合が多いように思われるので簡単に解説しておこう.

\mathbb{R}^n の二つのベクトル $u = {}^t[u_1, \dots, u_n], v = {}^t[v_1, \dots, v_n]$ の標準的な内積 (inner product) $u \cdot v$ を次のように定義する³⁹:

$$u \cdot v = \sum_{i=1}^n u_i v_i = u_1 v_1 + \dots + u_n v_n.$$

さらに u の長さ (もしくはノルム) $\|u\|$ を次のように定義する:

$$\|u\| = \sqrt{u \cdot u} = \sqrt{u_1^2 + \dots + u_n^2}.$$

以上の記号のもとで **Cauchy-Schwarz** の不等式

$$|u \cdot v| \leq \|u\| \|v\|$$

が成立していることを示せる. この不等式で等号が成立するための必要十分条件は u, v がもう片方の実数倍になっていることである. Cauchy-Schwarz の不等式より $u \neq 0$ かつ

³⁹内積は (u, v) や $\langle u, v \rangle$ のように書かれることが多い.

$v \neq 0$ ならば $|u \cdot v / (||u|| ||v||)| \leq 1$ となるので $u \cdot v / (||u|| ||v||) = \cos \theta$ ($\theta \in \mathbb{R}$) と表わせる. θ をベクトル u とベクトル v のあいだの角度と呼ぶ⁴⁰. このとき

$$u \cdot v = ||u|| ||v|| \cos \theta.$$

必要があれば θ の範囲を $0 \leq \theta \leq \pi$ に制限することができる.

[57] 上の場合における Cauchy-Schwarz の不等式を証明せよ. \square

ヒント. たとえば $n = 3$ の場合には次のようにして証明できる:

$$\begin{aligned} (u \cdot u)(v \cdot v) - (u \cdot v)^2 &= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2 \\ &= (u_1^2v_1^2 + u_1^2v_2^2 + u_1^2v_3^2 \\ &\quad + u_2^2v_1^2 + u_2^2v_2^2 + u_2^2v_3^2 \\ &\quad + u_3^2v_1^2 + u_3^2v_2^2 + u_3^2v_3^2) \\ &\quad - (u_1v_1u_1v_1 + u_1v_1u_2v_2 + u_1v_1u_3v_3 \\ &\quad + u_2v_2u_1v_1 + u_2v_2u_2v_2 + u_2v_2u_3v_3 \\ &\quad + u_3v_3u_1v_1 + u_3v_3u_2v_2 + u_3v_3u_3v_3) \\ &= (u_1^2v_2^2 - 2u_1v_1u_2v_2 + u_2^2v_1^2) \\ &\quad + (u_1^2v_3^2 - 2u_1v_1u_3v_3 + u_3^2v_1^2) \\ &\quad + (u_2^2v_3^2 - 2u_2v_2u_3v_3 + u_3^2v_2^2) \\ &= (u_1v_2 - u_2v_1)^2 + (u_1v_3 - u_3v_1)^2 + (u_2v_3 - u_3v_2)^2 \geq 0. \end{aligned}$$

計算のポイントは3つ目の等号である. 2つ目の等号の後の式の前者の括弧の中と後者の括弧の中の「対角成分」は互いにキャンセルし, 前者と後者の括弧の中の $i < j$ に対する「 (i, j) 成分」と「 (j, i) 成分」をまとめて並べ直せば3つ目の等号が成立することがわかる.

ここまでたどり着けば一般の n の場合も同様の計算が可能であることが容易に想像できるはずである. なぜならば上に説明した3つ目の等号の導き方は n によらない方法だからである⁴¹. \square

[58] (平行四辺形の面積) \mathbb{R}^2 内の2つのベクトル $u = \begin{bmatrix} a \\ c \end{bmatrix}$, $v = \begin{bmatrix} b \\ d \end{bmatrix}$ を任意に取り, 原点 0 と u を結ぶ線分, u と $u+v$ を結ぶ線分, $u+v$ と v を結ぶ線分 v と 0 を結ぶ線分で囲まれた平行四辺形⁴²を考える⁴³. その平行四辺形の面積が次の行列式の絶対値に等しいことを示せ:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \quad \square$$

⁴⁰ この定義による角度の概念と弧度法すなわち円弧の長さで定義した角度の概念が一致していることは証明が必要である. しかしこの問題をこの演習では扱わない. (この点を厳密に遂行するためには三角関数の定義まで戻らなければいけなくなるので結構説明が面倒である. 曲がった曲線の長さの正確な定義を学んだ後に各自十分に考察して欲しい.) この演習では二つの角度の概念が一致していることを自由に用いて良いことにする.

⁴¹ n が一般の場合の結果を得るために n が小さな場合の議論をよく観察して「それがうまく行く仕組み」を見抜くという考え方は極めて重要である.

⁴² parallelogram

⁴³ u, v が特殊な場合には平行四辺形が潰れて線分になったり, 場合によっては一点になってしまうこともあるが, その場合も除外せずに考えることにする.

ヒント. 公式「平行四辺形の面積 = 底辺の長さ × 高さ」を用いて計算してよい. この公式より, u, v のあいだの角度が θ ($0 \leq \theta \leq \pi$) であるとき, 平行四辺形の面積は $\|u\| \|v\| \sin \theta$ であることがわかる. \square

[59] 上の問題の続き. u と v の向きの位置関係の言葉で行列式 $ad - bc$ が正になるための必要十分条件を述べよ. \square

ヒント. $u = \begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} r \cos \theta \\ r \sin \theta \end{bmatrix}$, $v = \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} s \cos \phi \\ s \sin \phi \end{bmatrix}$ ($r, s \geq 0$) と置いて行列式を計算してみよ. \square

略解. まず行列式が 0 にならないための必要十分条件は u と v が一次独立であることである. 行列式が正になるための必要十分条件は u と v が一次独立でかつ u から見て v が反時計回りに 180 度未満の方向を向いていることである. \square

[60] (ベクトル積の定義) \mathbb{R}^3 の 2 つのベクトル $u = {}^t[u_1, u_2, u_3]$, $v = {}^t[v_1, v_2, v_3]$ のベクトル積 (vector product) $u \times v$ を次のように定義する:

$$u \times v := {}^t[u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1].$$

このとき以下が成立する:

1. 第 i 成分だけが 1 で他の成分が 0 であるような 3 次元縦ベクトルを e_i と書くと,

$$\begin{aligned} e_i \times e_j &= e_k, & e_j \times e_i &= -e_k & ((i, j, k) &= (1, 2, 3), (2, 3, 1), (3, 1, 2)), \\ e_i \times e_i &= 0 & (i &= 1, 2, 3). \end{aligned}$$

2. ベクトル $u = {}^t[u_1, u_2, u_3]$ に対して行列 $X(u)$ を次のように定める:

$$X(u) = \begin{bmatrix} 0 & u_1 & u_3 \\ -u_1 & 0 & u_2 \\ -u_3 & -u_2 & 0 \end{bmatrix}.$$

さらに行列 A, B の交換子 (commutator) $[A, B]$ を次のように定義する:

$$[A, B] = AB - BA.$$

このとき $u, v \in \mathbb{R}^3$ に対して

$$[X(u), X(v)] = X(u \times v).$$

3. ベクトル $u = {}^t[u_1, u_2, u_3]$ に対して行列 $Y(u)$ を次のように定める:

$$Y(u) = -\frac{i}{2}(u_1\sigma_1 + u_2\sigma_2 + u_3\sigma_3).$$

ここで $\sigma_1, \sigma_2, \sigma_3$ は次のように定義される **Pauli** 行列と呼ばれる行列である:

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

このとき $u, v \in \mathbb{R}^3$ に対して

$$[Y(u), Y(v)] = Y(u \times v).$$

4. $u, v, w \in \mathbb{R}^3$ に対して以下が成立している⁴⁴:

$$v \times u = -u \times v, \quad (u \times v) \times w = u \times (v \times w) - v \times (u \times w).$$

5. ベクトル積は行列式を用いて形式的に次のように表わされる:

$$u \times v = \begin{vmatrix} u_1 & v_1 & e_1 \\ u_2 & v_2 & e_2 \\ u_3 & v_3 & e_3 \end{vmatrix}.$$

この等式は「右辺の形式的な行列式の第 3 列に関する形式的な余因子展開が左辺に等しい」と読む。□

参考 2.19 上の問題の 2 と 3 はもちろん偶然ではない。実はベクトル積は 3 次元 Euclid 空間の (無限小) 回転を表現しているのである。実は上の問題は 3 次元 Euclid 空間の回転の表現の仕方には様々な方法があることを示していることになっている。

力学の教科書で回転運動の章を見るとベクトル積が登場する。それは回転運動を数学的に表現するためである。また量子物理の教科書を読むと Pauli 行列がよく登場する。それは我々が住んでいる物理的な 3 次元空間の回転対称性を表現するためである。

実は上の問題の 3 は **Hamilton** の四元数体 (quaternion) と関係している。四元数体とは複素数をさらに拡張した非可換体であり、実数体に i, j, k で

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j$$

を満たすものを付け加えることによって構成される。 $I = -i\sigma_1, J = -i\sigma_2, K = -i\sigma_3$ は i, j, k の満たすべき公式と同じ公式を満たしている。したがって、複素数が実 2 次正方行列で表現できたように (問題 [29] を見よ), 四元数は複素 2 次正方行列で表現できる。□

[61] (ベクトル積と平行四辺形の面積) 上の問題の続き。 \mathbb{R}^3 内で原点 0 と u を結ぶ線分, u と $u+v$ を結ぶ線分, $u+v$ と v を結ぶ線分 v と 0 を結ぶ線分で囲まれた平行四辺形を考える。このとき $u \times v$ はその平行四辺形に垂直になり, $u \times v$ の長さはその平行四辺形の面積に等しくなる。□

ヒント. $u \times v$ と u, v の内積が 0 になることが問題 [60] におけるベクトル積の定義もしくは 5 の表示から導かれる。平行四辺形の面積との関係については平行四辺形の面積が $\|u\| \|v\| \sin \theta$ であることを使え。ここで θ は u と v のあいだの角度である。□

[62] (平行六面体の体積) \mathbb{R}^3 の 3 つのベクトル $u = {}^t[u_1, u_2, u_3], v = {}^t[v_1, v_2, v_3], w = {}^t[w_1, w_2, w_3]$ を任意に取り, 0, $u, v, w, u+v, u+w, v+w, u+v+w$ を八つの頂点に持つ平行六面体⁴⁵を考える。次の等式を証明せよ:

$$\det[u, v, w] = (u \times v) \cdot w.$$

この等式を用いてこの等式の両辺の絶対値が平行六面体の体積に等しいことを示せ。□

⁴⁴ ヒント: n 次正方行列 A, B, C に対して $[[A, B], C] = [A, [B, C]] - [B, [A, C]]$ が成立している。これを交換子の **Jacobi 律** と呼ぶ。

⁴⁵ parallelepiped

ヒント. 公式「平行六面体の体積 = 底面の面積 × 高さ」を用いて良い. 底面を $0, u, v, u+v$ を頂点とする平行四辺形に取り, w とその平行四辺形のなす角度を θ ($0 \leq \theta \leq \pi$) とするとき, 平行六面体の体積は $\|u \times v\| \|w\| \sin \theta$ と表示される. \square

[63] $u = {}^t[-1, 2, 2], v = {}^t[2, -1, 2], w = {}^t[2, 2, -1]$ のとき問題 [62] の平行六面体の図を描き, その体積を計算せよ. \square

略解. 体積は 27 になる. \square

参考 2.20 (n 次元空間中の平行 $2n$ 面体の体積) \mathbb{R}^n 内の n 本のベクトル

$$v_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix}, v_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, v_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix}$$

に対して 2^n 個の点

$$\sum_{i \in I} v_i, \quad I \subset \{1, 2, \dots, n\}$$

を頂点とする n 次元平行 $2n$ 面体⁴⁶の体積が行列式

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

の絶対値に等しいことも示せる.

上の行列式が 0 にならないための必要十分条件は v_1, \dots, v_n が \mathbb{R}^n の基底 (basis) になることである. そのとき行列式の正負は順序付けられた基底 (v_1, \dots, v_n) が向き (orientation) が正であるか負であるかを表現している. 「向き」の概念についてここでは解説しないが, 2 次元の場合に関する問題 [59] の結果を見て, 3 次元以上の場合にどうなるかを想像して欲しい. (3 次元の場合には「向き」を「右手系」「左手系」という言葉で区別することが多い.)

行列式と体積 (もしくは面積) のあいだの以上の関係は多重積分の変数変換の公式を求めるときに必要な. \square

[64] (n 次元空間中の平行 $2n$ 面体の体積) 参考 2.20 に書いてあるように \mathbb{R}^n 内の平行 $2n$ 面体の体積が行列式の絶対値で表示できることを示せ. \square

ヒント. まず, 行列式の多重線形性に対応する性質が平行 $2n$ 面体の体積にもあることを示せ. 行列式の絶対値と平行 $2n$ 面体の体積の (v_1, \dots, v_n) の関数としての共通の性質を十分たくさん見付ければ, そのことから二つの関数が等しいことを示せる. \square

[65] (n 次元空間中の平行 $2(n-1)$ 面体の面積) \mathbb{R}^n 内の $n-1$ 本のベクトル

$$v_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix}, v_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, v_{n-1} = \begin{bmatrix} a_{1,n-1} \\ \vdots \\ a_{n,n-1} \end{bmatrix}$$

⁴⁶ 平行体, parallelotope

に対して 2^{n-1} 個の点

$$\sum_{i \in I} v_i, \quad I \subset \{1, 2, \dots, n-1\}$$

を頂点とする $n-1$ 次元平行 $2(n-1)$ 面体を考える. さらに次のような形式的な行列式で定義される \mathbb{R}^n のベクトル u を考える:

$$u = \begin{vmatrix} a_{11} & \cdots & a_{1,n-1} & e_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & e_n \end{vmatrix}.$$

このとき u は平行 $2(n-1)$ 面体に直交し, u の長さは平行 $2(n-1)$ 面体の面積に等しい. \square

ヒント. 直交性の証明は $n=3$ の場合と同様である. 問題は面積との関係である. 問題 [64] の結果を使えば面積との関係も示せる. 問題 [64] の平行 $2n$ 面体の体積を V と書き, 問題 [65] の平行 $2(n-1)$ 面体の面積を B と書き, 平行 $2(n-1)$ 面体を含む超平面と v_n のなす角度を θ ($0 \leq \theta \leq \pi/2$) と書くことにする. 平行 $2n$ 面体の体積が底面の面積と高さの積に等しいことを使えば $V = B \|v_n\| \sin \theta$ が成立していることがわかる. u と v_n を含む直線の角度は $\pi/2 - \theta$ であるから $V = |\det[v_1, \dots, v_n]| = |u \cdot v_n| = \|u\| \|v_n\| \cos(\pi/2 - \theta) = \|u\| \|v_n\| \sin \theta$ である. これより $B = \|u\|$ が出る. \square

参考 2.21 (n 次元空間内の平行 $2r$ 面体の面積) 上の問題のヒントの方法はあまり格好良くない. なぜならば $1 \leq r \leq n$ のとき, 2^r 個の点

$$\sum_{i \in I} v_i, \quad I \subset \{1, 2, \dots, r\}$$

を頂点とする r 次元平行 $2r$ 面体の面積 (もしくは長さや体積) に関する結果を上ヒントの方法を拡張して得るのはかなり面倒だからである. この最も一般の場合に関する結果を述べるためには小行列式と外積代数の言葉を用意するのが便利である. (参考 2.26 を見よ.) 直交変換で面積が保たれることを使えば証明が著しく簡単になる. \square

2.6 小行列式と外積代数

定義 2.22 (小行列式) $A = [a_{ij}]$ は $m \times n$ 行列であるとする:

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

$1 \leq r \leq \min(m, n)$ とし,

$$\begin{aligned} I &= (i_1, \dots, i_r), & 1 \leq i_1, \dots, i_r \leq m, \\ J &= (j_1, \dots, j_r), & 1 \leq j_1, \dots, j_r \leq n \end{aligned}$$

とする. このとき, (I, J) の定める A の小行列式 (**minor**) $A_J^I = A_{j_1 \dots j_r}^{i_1 \dots i_r}$ を次のように定める:

$$A_J^I = A_{j_1 \dots j_r}^{i_1 \dots i_r} = \begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_r} \\ \vdots & & \vdots \\ a_{i_r j_1} & \cdots & a_{i_r j_r} \end{vmatrix}.$$

すなわち行列 A から第 i_1, \dots, i_r 行と第 j_1, \dots, j_r 列の交わりの部分を取り出して行列式を取ったものを $A_J^I = A_{j_1 \dots j_r}^{i_1 \dots i_r}$ と表す.

I, J が共に r 個の元を持つ $\{1, 2, \dots, n\}$ の部分集合であるとき I, J の元を小さな順に並べたものをそれぞれ $i_1 < \dots < i_r, j_1 < \dots < j_r$ と書き, $A_J^I = A_{j_1 \dots j_r}^{i_1 \dots i_r}$ と定めておく. 主としてこの場合の小行列式を以下では扱う. \square

たとえば $r = 1$ のとき

$$A_j^i = a_{ij}$$

であり, $r = 2$ のとき

$$A_{j_1 j_2}^{i_1 i_2} = \begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} \\ a_{i_2 j_1} & a_{i_2 j_2} \end{vmatrix}$$

である. もしも $m = n = r$ で $I = J = (1, \dots, n)$ ならば

$$A_I^I = A_{1 \dots n}^{1 \dots n} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = |A|$$

となり, 小行列式は行列式に一致する. $m = n, r = n-1$ でかつ $I = \{1, \dots, \widehat{i}, \dots, n\} = \{i\}^c, J = \{1, \dots, \widehat{j}, \dots, n\} = \{j\}^c$ のとき (ここで \widehat{i}, \widehat{j} はそれらを取り除くことを意味する), A の (i, j) 余因子を Δ_{ij} と表わすと

$$\Delta_{ij} = (-1)^{i+j} A_J^I = (-1)^{i+j} A_{1 \dots \widehat{j} \dots n}^{1 \dots \widehat{i} \dots n}.$$

さらに問題 [54] の Gauss 分解の行列式表示の結果を

$$w_{ij} = (-1)^{i+j} \frac{A_{1 \dots \widehat{j} \dots n}^{1 \dots \widehat{i} \dots n}}{A_{1 \dots i-1}^{1 \dots i-1}} \quad (i > j), \quad z_{ij} = \frac{A_{1 \dots i-1, j}^{1 \dots i-1, i}}{A_{1 \dots i-1}^{1 \dots i-1}} \quad (i \leq j)$$

とコンパクトに書くこともできる.

これ以後, I の元の個数を $|I|$ と書くことにする.

[66] A, B は n 次正方行列であり, I, J は共に $\{1, 2, \dots, n\}$ の部分集合であり, $|I| = |J| = r$ を満たしているとする. このとき

$$(AB)_J^I = \sum_{|K|=r} A_K^I B_J^K. \quad \square$$

ヒント. 問題 [38] の $|AB| = |A||B|$ の場合と同様に証明可能である. $I = \{i_1 < \dots < i_r\}, J = \{j_1 < \dots < j_r\}$ と表わしておく. A の第 i_1, \dots, i_r 行部分を A' と書き, B の第

j_1, \dots, j_r 列部分を B' と書き, A' の第 j 列を a'_j と表わし, B の (i, j) 成分を b_{ij} と表わし, $A'B'$ の第 ν 列を c_ν と表わすと,

$$c_\nu = \sum_{i=1}^n a'_i b_{ij_\nu} \quad (\nu = 1, \dots, r).$$

AB の第 i_1, \dots, i_r 行と第 j_1, \dots, j_r 列の交わり部分を C と書くと, $C = A'B' = [c_1, \dots, c_r]$ である. よって行列式の多重線形性 [35] より

$$(AB)_J^I = \det C = \det[c_1, \dots, c_r] = \sum_{k_1, \dots, k_r=1}^n \det[a'_{k_1}, \dots, a'_{k_r}] b_{k_1 j_1} \cdots b_{k_r j_r}.$$

さらに行列式の反対称性 [33], [34] より

$$(AB)_J^I = \sum_{1 \leq k_1 < \dots < k_r \leq n} \sum_{\sigma \in S_r} \det[a'_{k_1}, \dots, a'_{k_r}] \text{sign}(\sigma) b_{k_{\sigma(1)} j_1} \cdots b_{k_{\sigma(r)} j_r} = \sum_{|K|=r} A_K^I B_J^K. \quad \square$$

注意 2.23 (長方形型の行列の積の行列式) まず, A, B が次のような形をしている場合について考える:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} A' \\ 0 \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \cdots & b_{1r} & 0 & \cdots & 0 \\ b_{21} & \cdots & b_{2r} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nr} & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} B' & 0 \end{bmatrix}.$$

A の上側の $r \times n$ の部分を A' と書き, B の左側の $n \times r$ の部分を B' と書いた. このとき,

$$AB = \begin{bmatrix} A' \\ 0 \end{bmatrix} \begin{bmatrix} B' & 0 \end{bmatrix} = \begin{bmatrix} A'B' & 0 \\ 0 & 0 \end{bmatrix}$$

なので, $A'B'$ は AB の左上の $r \times r$ の部分に等しい. よって

$$|A'B'| = (AB)_{1 \dots r}^{1 \dots r}$$

が成立する. したがって問題 [66] の結果より, 次が成立することがわかる:

$$|A'B'| = \sum_{1 \leq k_1 < \dots < k_r \leq n} A_{k_1 \dots k_r}^{1 \dots r} B_{1 \dots r}^{k_1 \dots k_r}.$$

次に A と B の立場を逆転させて

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1r} & 0 & \cdots & 0 \\ a_{21} & \cdots & a_{2r} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nr} & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} A'' & 0 \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ \vdots & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} B'' \\ 0 \end{bmatrix}.$$

の場合について考える. A の左側の $n \times r$ 行列部分を A'' と書き, B の上側の $r \times n$ 行列部分を B'' と書いた. このとき, $A''B'' = AB$ が成立する. よって $r < n$ ならば $|A''B''| = |AB| = |A||B| = 0$ である.

結果をまとめよう. $A = [a_{ij}]$ は $m \times n$ 行列であり, $B = [b_{ij}]$ は $n \times m$ 行列であるとすると, 以下が成立する:

1. $m \leq n$ ならば

$$|AB| = \sum_{1 \leq k_1 < \dots < k_m \leq n} \begin{vmatrix} a_{1k_1} & \dots & a_{1k_m} \\ \vdots & & \vdots \\ a_{mk_1} & \dots & a_{mk_m} \end{vmatrix} \begin{vmatrix} b_{k_1 1} & \dots & b_{k_1 m} \\ \vdots & & \vdots \\ b_{k_m 1} & \dots & b_{k_m m} \end{vmatrix}.$$

特に $m = n$ ならば $|AB| = |A||B|$.

2. $m > n$ ならば $|AB| = 0$. \square

定義 2.24 (外積代数) \mathbb{R}^n から生成される外積代数とは

$$e_{i_1} \wedge \dots \wedge e_{i_r}, \quad r = 0, 1, 2, \dots, n, \quad 1 \leq i_1 < \dots < i_r \leq n \quad (*)$$

を基底に持つ \mathbb{R} 上の 2^n 次元のベクトル空間 $\bigwedge(\mathbb{R}^n)$ に以下で説明するような規則で外積 \wedge を定義したもののことである. (e_i は第 i 成分だけが 1 で他の成分が 0 であるような n 次元縦ベクトルである. $r = 0$ のとき $e_{i_1} \wedge \dots \wedge e_{i_r} = 1$ であると約束しておく.)

まず任意の置換 $\sigma \in S_r$ に対して

$$e_{i_{\sigma(1)}} \wedge \dots \wedge e_{i_{\sigma(r)}} = \text{sign}(\sigma) e_{i_1} \wedge \dots \wedge e_{i_r}$$

と定義し, $i_1, \dots, i_r = 1, \dots, n$ が互いに異なるある p, q について $i_p = i_q$ を満たしているとき $e_{i_1} \wedge \dots \wedge e_{i_r} = 0$ であると約束しておく. これで任意の $r = 0, 1, 2, \dots, i_1, \dots, i_r = 1, \dots, n$ に対して $e_{i_1} \wedge \dots \wedge e_{i_r} = 0$ が定義された. 以上の準備のもとで $e_{i_1} \wedge \dots \wedge e_{i_r}$ と $e_{j_1} \wedge \dots \wedge e_{j_s}$ の外積 (exterior product) を

$$(e_{i_1} \wedge \dots \wedge e_{i_r}) \wedge (e_{j_1} \wedge \dots \wedge e_{j_s}) = e_{i_1} \wedge \dots \wedge e_{i_r} \wedge e_{j_1} \wedge \dots \wedge e_{j_s}$$

と定める. 以上の構成は well-defined であり, $\bigwedge(\mathbb{R}^n)$ に 1 を持つ結合代数の構造を定める⁴⁷. $\bigwedge(\mathbb{R}^n)$ は \mathbb{R}^n から生成される外積代数 (exterior algebra) もしくは Grassmann 代数 (Grassmann algebra) と呼ばれている⁴⁸.

⁴⁷ この抽象代数の言葉をこの段階ではあまり気にする必要はない. この演習では外積代数の計算規則だけを認めて自由に使って構わないことにする. 主な計算規則は外積の結合法則と双線形性

$$\begin{aligned} (a \wedge b) \wedge c &= a \wedge (b \wedge c) & (a, b, c \in \bigwedge(\mathbb{R}^n)), \\ (\alpha a + \alpha' a') \wedge b &= \alpha(a \wedge b) + \alpha'(a' \wedge b) & (a, a' b \in \bigwedge(\mathbb{R}^n), \alpha, \alpha' \in \mathbb{R}), \\ (\alpha a) \wedge (\beta b + \beta' b') &= \beta(a \wedge b) + \beta'(a \wedge b') & (a, b, b' \in \bigwedge(\mathbb{R}^n), \beta, \beta' \in \mathbb{R}) \end{aligned}$$

および次の外積代数の基本関係式である:

$$e_i \wedge e_j = -e_j \wedge e_i, \quad e_i \wedge e_i = 0.$$

⁴⁸ 外積代数は幾何学の多様体 (manifold, variety) の理論において微分形式 (differential form) を定義するために使われる. 参考 2.26 で証明抜きで説明するように面積や体積の概念と外積代数は非常に相性が良い. 幾何学において面積や体積の概念は基本的なので必然的に外積代数が幾何学の世界に登場することになる.

さらに $\bigwedge(\mathbb{R}^n)$ に内積 $\langle \cdot, \cdot \rangle$ を $(*)$ が正規直交基底になるように入れておく:

$$\langle e_{i_1} \wedge \cdots \wedge e_{i_r}, e_{j_1} \wedge \cdots \wedge e_{j_s} \rangle = \delta_{rs} \delta_{i_1, j_1} \cdots \delta_{i_r, j_r}.$$

ここで $1 \leq i_1 < \cdots < i_r \leq n$ かつ $1 \leq j_s < \cdots < j_s \leq n$. \square

注意 2.25 上の定義では \mathbb{R}^n から生成される外積代数 $\bigwedge(\mathbb{R}^n)$ を定義したが, 一般に体 K 上のベクトル空間 V から生成される外積代数 $\bigwedge(V)$ を定義することもできる. \square

$I = \{i_1 < \cdots < i_r\} \subset \{1, 2, \dots, n\}$ のとき e_I を次のように定める:

$$e_I = e_{i_1} \wedge \cdots \wedge e_{i_r}$$

[67] (小行列式と外積代数の関係) \mathbb{R} の元を成分に持つ n 次正方形行列 $A = [a_{ij}]$ を用意し, その第 j 列を $a_j = Ae_j = \sum_{i=1}^n a_{ij}e_i \in \mathbb{R}^n$ と書けば, $J = \{j_1 < \cdots < j_r\} \subset \{1, 2, \dots, n\}$ に対して,

$$a_{j_1} \wedge \cdots \wedge a_{j_r} = \sum_{1 \leq i_1 < \cdots < i_r \leq n} A_{j_1 \dots j_r}^{i_1 \dots i_r} e_{i_1} \wedge \cdots \wedge e_{i_r} = \sum_{|I|=r} A_J^I e_I.$$

特に $r = n$, $j_\nu = \nu$ のとき次が成立する:

$$a_1 \wedge \cdots \wedge a_n = |A| e_1 \wedge \cdots \wedge e_n. \quad \square$$

ヒント. 次のように計算される:

$$\begin{aligned} a_{j_1} \wedge \cdots \wedge a_{j_r} &= \sum_{i_1, \dots, i_r=1}^n a_{i_1, j_1} \cdots a_{i_r, j_r} e_{i_1} \wedge \cdots \wedge e_{i_r} \\ &= \sum_{1 \leq i_1 < \cdots < i_r \leq n} \sum_{\sigma \in S_r} a_{i_{\sigma(1)}, j_1} \cdots a_{i_{\sigma(r)}, j_r} e_{i_{\sigma(1)}} \wedge \cdots \wedge e_{i_{\sigma(r)}} \\ &= \sum_{1 \leq i_1 < \cdots < i_r \leq n} \sum_{\sigma \in S_r} a_{i_{\sigma(1)}, j_1} \cdots a_{i_{\sigma(r)}, j_r} \text{sign}(\sigma) e_{i_1} \wedge \cdots \wedge e_{i_r} \\ &= \sum_{1 \leq i_1 < \cdots < i_r \leq n} A_{j_1 \dots j_r}^{i_1 \dots i_r} e_{i_1} \wedge \cdots \wedge e_{i_r}. \end{aligned}$$

各ステップで何を用いたか? \square

参考 2.26 (n 次元空間内の平行 $2r$ 面体の面積) 問題 [67] の記号のもとで 2^r 個の点

$$\sum_{j \in I} a_j, \quad I \subset \{1, 2, \dots, r\}$$

を頂点とする \mathbb{R}^n 内の r 次元平行 $2r$ 面体の面積 (もしくは長さ, 体積) は

$$a_1 \wedge \cdots \wedge a_r = \sum_{1 \leq i_1 < \cdots < i_r \leq n} A_{1 \dots r}^{i_1 \dots i_r} e_{i_1} \wedge \cdots \wedge e_{i_r}$$

のノルムに等しいことを示すことができる. (外積空間 $\bigwedge(\mathbb{R}^n)$ には内積を定めておいたの
でノルムという言葉が意味を持つことに注意せよ.) すなわち, 平行 $2r$ 面体の面積は

$$\|a_{j_1} \wedge \cdots \wedge a_{j_r}\| = \left(\sum_{|I|=r} (A_{1 \dots r}^{i_1 \dots i_r})^2 \right)^{1/2}$$

に等しい. この結果は問題 [58], [61], [62], [64], [65] の結果を含んでいる. 証明するためには, \mathbb{R}^n の直交変換で面積が保たれることと, \mathbb{R}^n の直交変換が $\bigwedge(\mathbb{R}^n)$ の直交変換に自然に持ち上がることを使えば良い⁴⁹ \square

[68] $J = \{j_1 < \cdots < j_r\} \subset \{1, 2, \dots, n\}$ のとき, n 次正方行列 A の $e_{j_1} \wedge \cdots \wedge e_{j_r}$ への左からの積を

$$A(e_{j_1} \wedge \cdots \wedge e_{j_r}) = (Ae_{j_1}) \wedge \cdots \wedge (Ae_{j_r})$$

と定める. このとき n 次正方行列 A, B に対して,

$$A(B(e_{j_1} \wedge \cdots \wedge e_{j_r})) = (AB)(e_{j_1} \wedge \cdots \wedge e_{j_r}). \quad \square$$

ヒント. 問題 [67] の結果を用いて問題 [66] の結果の言い換えてみよ. \square

参考 2.27 実は適切な定式化を行えば上の問題の結論が自明であるようにできる. 適切な定式化のもとでは GL_n の外積代数上での表現が自然に構成される. その立場では問題 [66] の結果はほとんど自明になってしまう. 一見非自明に見える関係式が表現論の立場から見ると表現の matrix elements が満たす自明な関係式になってしまう場合がある. \square

余因子も一般化しておこう.

定義 2.28 (余因子) r 個の元を持つ $\{1, \dots, n\}$ の部分集合 $I = \{i_1 < \cdots < i_r\}$ に対して, その補集合を $I^c = \{i'_1 < \cdots < i'_{n-r}\}$ と表わし, $\text{sign}(I)$ を次のように定める:

$$\text{sign}(I) = \text{sign} \begin{pmatrix} 1 & \cdots & r & r+1 & \cdots & n \\ i_1 & \cdots & i_r & i'_1 & \cdots & i'_{n-r} \end{pmatrix}.$$

I, J が共に r 個の元を持つ $\{1, \dots, n\}$ の部分集合であるとき n 次正方行列 A の (I, J) 余因子 Δ_J^I を次のように定義する:

$$\Delta_J^I := \text{sign}(I) \text{sign}(J) A_{J^c}^{I^c}. \quad \square$$

たとえば A の (i, j) 余因子を Δ_{ij} と表わすと

$$\Delta_{ij} = (-1)^{i+j} A_{\{j\}^c}^{\{i\}^c} = \Delta_{\{j\}}^{\{i\}}.$$

[69] (行列式の Laplace 展開) A が n 次正方行列で I, J は $\{1, 2, \dots, n\}$ の部分集合で $|I| = |J| = r$ を満たしているとき,

$$\sum_{|K|=r} \Delta_I^K A_J^K = \sum_{|K|=r} A_K^I \Delta_K^J = |A| \delta_{IJ}. \quad \square$$

ヒント 1. 行列式の定義に戻れば証明できる. すでに習ったはずの $k = 1$ の場合の Laplace 展開を繰り返し用いて証明することもできる. (高木 [Tkg1] 第 8 章第 52 節の 250–255 頁に詳しい説明がある.) \square

⁴⁹ 申し訳ないが証明に関する詳しい説明は省略する. 外積代数と面積や体積が密接に関係していることを知っておいて欲しいのでこの解説をここに挿入することにした.

ヒント 2. 外積代数を用いた証明の方針が佐武 [St] 第 V 章第 4 節 225 頁の問 6 (略解が 313 頁にある) にある. (横沼 [Ykn] 92–96 頁に詳しい説明がある.)

A の第 j 列を a_j と書き, I, J を $I = \{i_1 < \cdots < i_r\}$, $J = \{j_1 < \cdots < j_r\}$ と表わし, I の補集合を $I^c = \{i'_1 < \cdots < i'_{n-r}\}$ と表わすと, 問題 [67] の結果より,

$$a_{j_1} \wedge \cdots \wedge a_{j_r} \wedge a_{i'_1} \wedge \cdots \wedge a_{i'_{n-r}} = \text{sign}(I) |A| \delta_{IJ} e_{\{1,2,\dots,n\}}. \quad (1)$$

同様に [67] の結果より,

$$a_{j_1} \wedge \cdots \wedge a_{j_r} = \sum_{|K|=r} A_J^K e_K, \quad a_{i'_1} \wedge \cdots \wedge a_{i'_{n-r}} = \sum_{|L|=n-r} A_{I^c}^L e_L.$$

したがって

$$\begin{aligned} (a_{j_1} \wedge \cdots \wedge a_{j_r}) \wedge (a_{i'_1} \wedge \cdots \wedge a_{i'_{n-r}}) &= \sum_{|K|=r} \sum_{|L|=n-r} A_{I^c}^L A_J^K e_K \wedge e_L \\ &= \sum_{|K|=r} A_{I^c}^{K^c} A_J^K e_K \wedge e_{K^c} = \sum_{|K|=r} A_{I^c}^{K^c} A_J^K \text{sign}(K) e_{\{1,2,\dots,n\}}. \end{aligned} \quad (2)$$

(1), (2) を比較すれば $\sum_{|K|=r} \text{sign}(K) A_{I^c}^{K^c} A_J^K = \text{sign}(I) |A| \delta_{IJ}$ が得られる. この等式の右辺の $\text{sign}(I)$ を左辺に移せば示したい公式 $\sum_{|K|=r} \Delta_I^K A_J^K = |A| \delta_{IJ}$ が得られる. 残りの半分も同様に証明されるので実際にやってみよ. \square

3 色々な行列式

行列式についてちょっと進んだ話を知りたいければ『数理科学』1995 年 4 月号の特集「行列式の進化」[Gy] を見よ. たとえば梅田 [Umd] p.25 にある非可換行列式 (その 3) の式と問題 [95] の式を比較してみよ.

第 3.8 節で Pfaffian に関する問題をほんの少しだけ出し, 佐武 [St] pp.81–82 をヒントで紹介した. しかし, それを見るだけでは Pfaffian は行列式自身に負けず劣らず (場合によってはむしろ行列式より) 普遍的な数学的対象であることを納得できないと思われる. Pfaffian について詳しく解説している文献は少ないが, 邦書では広田 [Hr] があり, 特集 [Gy] に所収の若山 [Wky] も良い.

3.1 Vandermonde の行列式 (ヴァンデルモンドの行列式)

[70] (Vandermonde の行列式) 次の公式が成立している:

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

左辺は Vandermonde の行列式 (ヴァンデルモンドの行列式) と呼ばれる. 右辺は x_1, x_2, \dots, x_n の差積 (difference product) と呼ばれ, よく

$$\Delta(x) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

と表わされる. \square

ヒント. 行列式の多重線形性と反対称性のみを用いて n に関する帰納法で証明することもできるし⁵⁰, 大抵の教科書に載っている「両辺が同じ場所に零点を持つ」という原理に基いた巧妙な (しかし他の様々な場面においても常套手段になっている) 証明法もある. \square

[71] A は複素 n 次正方行列であり, $v_1, \dots, v_r \in \mathbb{C}^n$ は

$$Av_i = \alpha_i v_i, \quad \alpha_i \in \mathbb{C}, \quad v_i \neq 0$$

を満たしているとし, $v = v_1 + \dots + v_r$ と置く. もしも $\alpha_1, \dots, \alpha_r$ が互いに異なるならば $v, Av, A^2v, \dots, A^{r-1}v$ の一次結合で v_1, \dots, v_r を表わせる. \square

ヒント. $v = v_1 + \dots + c_r v_r$ の両辺に A^k ($k = 0, 1, 2, \dots, r-1$) をかけると

$$\begin{aligned} v &= v_1 + \dots + v_r, \\ Av &= \alpha_1 v_1 + \dots + \alpha_r v_r, \\ A^2v &= \alpha_1^2 v_1 + \dots + \alpha_r^2 v_r, \\ &\dots\dots\dots \\ A^{r-1}v &= \alpha_1^{r-1} v_1 + \dots + \alpha_r^{r-1} v_r. \end{aligned}$$

これは行列を用いて次のように表わされる:

$$[v, Av, A^2v, \dots, A^{r-1}v] = [v_1, \dots, v_r] \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{r-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha_r & \alpha_r^2 & \dots & \alpha_r^{r-1} \end{bmatrix}.$$

よってもしも $\alpha_1, \dots, \alpha_r$ が互いに異なるならば $\alpha_1, \dots, \alpha_r$ の Vandermonde 行列式は 0 にならないので, 上の等式の右辺の右側の $r \times r$ 行列は可逆になる. よって $[v_1, \dots, v_r]$ は次のように表わされる:

$$[v_1, \dots, v_r] = [v, Av, A^2v, \dots, A^{r-1}v] \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{r-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha_r & \alpha_r^2 & \dots & \alpha_r^{r-1} \end{bmatrix}^{-1}.$$

この等式を各 v_i ごとに見れば v_i が $v, Av, \dots, A^{r-1}v$ の一次結合で表わされていることになる. \square

3.2 Cauchy の行列式と Hilbert の行列式

[72] (Cauchy の行列式)

$$\begin{vmatrix} \frac{1}{x_1-y_1} & \frac{1}{x_1-y_2} & \dots & \frac{1}{x_1-y_n} \\ \frac{1}{x_2-y_1} & \frac{1}{x_2-y_2} & \dots & \frac{1}{x_2-y_n} \\ \vdots & \vdots & & \vdots \\ \frac{1}{x_n-y_1} & \frac{1}{x_n-y_2} & \dots & \frac{1}{x_n-y_n} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} \frac{\prod_{1 \leq i < j \leq n} (x_j - x_i) \cdot \prod_{1 \leq i < j \leq n} (y_j - y_i)}{\prod_{i,j=1}^n (x_i - y_j)}.$$

左辺の行列式を **Cauchy** の行列式と呼ぶ. Cauchy の行列式を $C(x, y)$ と書くことにする. \square

⁵⁰佐武 [St] 第 II 章 §3 例 2 (59 頁) を参照せよ.

ヒント. $f(x, y) := \prod_{i,j=1}^n (x_i - y_j) \cdot C(x, y)$ と置く. このとき $f(x, y)$ は

$$f(x, y) = \begin{vmatrix} \prod_{i \neq 1} (x_i - y_1) & \prod_{i \neq 1} (x_i - y_2) & \cdots & \prod_{i \neq 1} (x_i - y_n) \\ \prod_{i \neq 2} (x_i - y_1) & \prod_{i \neq 2} (x_i - y_2) & \cdots & \prod_{i \neq 2} (x_i - y_n) \\ \vdots & \vdots & & \vdots \\ \prod_{i \neq n} (x_i - y_1) & \prod_{i \neq n} (x_i - y_2) & \cdots & \prod_{i \neq n} (x_i - y_n) \end{vmatrix} \quad (*)$$

という表示を持つので, 互いに異なる j, k で $y_j = y_k$ となるものが存在すれば $f(x, y) = 0$ となる. よって $f(x, y)$ は y_j たちの差積 $\Delta(y) = \prod_{1 \leq i < j \leq n} (y_j - y_i)$ で割り切れる. $f(x, y)$ の別の表示を用いれば同様に $f(x, y)$ が x_i たちの差積 $\Delta(x)$ で割り切れることがわかる. さらに, $f(x, y)$ の次数は $n(n-1)$ であり, $\Delta(x)$ の次数も $\Delta(y)$ の次数も $n(n-1)/2$ であるから, $f(x, y)$ は

$$f(x, y) = C \Delta(x) \Delta(y), \quad C \text{ は定数}$$

の形でなければいけない. ところが (*) より

$$f(x, x) = \prod_{j=1}^n \prod_{i \neq j} (x_i - x_j) = \prod_{i < j} (x_i - x_j) \cdot \prod_{i > j} (x_i - x_j) = (-1)^{\frac{n(n-1)}{2}} \Delta(x)^2.$$

よって $C = (-1)^{\frac{n(n-1)}{2}}$ である. \square

[73] (一般化された Hilbert の行列式)

$$\begin{vmatrix} \frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} & \cdots & \frac{1}{x_1+y_n} \\ \frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} & \cdots & \frac{1}{x_2+y_n} \\ \vdots & \vdots & & \vdots \\ \frac{1}{x_n+y_1} & \frac{1}{x_n+y_2} & \cdots & \frac{1}{x_n+y_n} \end{vmatrix} = \frac{\prod_{1 \leq i < j \leq n} (x_j - x_i) \cdot \prod_{1 \leq i < j \leq n} (y_j - y_i)}{\prod_{i,j=1}^n (x_i + y_j)}.$$

左辺の行列式を 一般化された Hilbert の行列式と呼ぶ. 一般化された Hilbert の行列式を $H(x, y)$ と書くことにする. \square

ヒント. $H(x, y) = C(x, -y)$ である. ところが $\Delta(-y) = (-1)^{\frac{n(n-1)}{2}} \Delta(y)$ なので目標の公式が得られる. \square

[74] (Hilbert の行列式) 次の行列式を n 次の Hilbert の行列式と呼ぶ:

$$H_n := \begin{vmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots & \frac{1}{n+2} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{vmatrix}$$

この値を求めよ. \square

ヒント. $x_i = i, y_j = j - 1$ のとき $H_n = H(x, y)$ である. \square

略解. $H_n = H(x, y)$ の分母と分子が次のように計算される:

$$\prod_{i,j=1}^n (x_i + y_j) = \prod_{i,j=1}^n (i + j - 1) = n! \frac{(n+1)!}{1!} \frac{(n+2)!}{2!} \cdots \frac{(2n-1)!}{(n-1)!},$$

$$\Delta(x) = \Delta(1, 2, \dots, n) = (n-1)! \cdots 3!2!1!,$$

$$\Delta(y) = \Delta(0, 1, \dots, n-1) = (n-1)! \cdots 3!2!1!.$$

したがって

$$H_n = H(x, y) = \frac{(1!2!3! \cdots (n-1)!)^3}{n!(n+1)!(n+2)! \cdots (2n-1)!}. \quad \square$$

3.3 Gram 行列式

[75] (**Gram 行列式**) $r \leq n$ であるとし, $A = [a_{ij}]$ は実 $n \times r$ 行列であるとする. A の第 j 列を a_j と書くと,

$${}^t A A = \begin{bmatrix} {}^t a_1 \\ \vdots \\ {}^t a_r \end{bmatrix} [a_1, \dots, a_r] = \begin{bmatrix} a_1 \cdot a_1 & a_1 \cdot a_2 & \cdots & a_1 \cdot a_r \\ a_2 \cdot a_1 & a_2 \cdot a_2 & \cdots & a_2 \cdot a_r \\ \vdots & \vdots & & \vdots \\ a_r \cdot a_1 & a_r \cdot a_2 & \cdots & a_r \cdot a_r \end{bmatrix}.$$

ここで $a_i \cdot a_j$ は a_i と a_j の内積である. この行列を **Gram 行列** と呼び, その行列式を **Gram 行列式 (Gramian)** と呼ぶ. Gram 行列式に関して次が成立している:

$$\begin{vmatrix} a_1 \cdot a_1 & a_1 \cdot a_2 & \cdots & a_1 \cdot a_r \\ a_2 \cdot a_1 & a_2 \cdot a_2 & \cdots & a_2 \cdot a_r \\ \vdots & \vdots & & \vdots \\ a_r \cdot a_1 & a_r \cdot a_2 & \cdots & a_r \cdot a_r \end{vmatrix} = \sum_{1 \leq i_1 < \cdots < i_r \leq n} \begin{vmatrix} a_{i_1 1} & a_{i_1 2} & \cdots & a_{i_1 r} \\ a_{i_2 1} & a_{i_2 2} & \cdots & a_{i_2 r} \\ \vdots & \vdots & & \vdots \\ a_{i_r 1} & a_{i_r 2} & \cdots & a_{i_r r} \end{vmatrix}^2 \geq 0.$$

特に $r = 2$ で $a_1 = u = {}^t[u_1, \dots, u_n]$, $a_2 = v = {}^t[v_1, \dots, v_n]$ のとき

$$(u \cdot u)(v \cdot v) - (u \cdot v)^2 = \begin{vmatrix} u \cdot u & u \cdot v \\ v \cdot u & v \cdot v \end{vmatrix} = \sum_{1 \leq i < j \leq n} \begin{vmatrix} u_i & v_i \\ u_j & v_j \end{vmatrix}^2 \geq 0$$

が成立する⁵¹. よって次の Cauchy-Schwarz の不等式が出る:

$$|u \cdot v| \leq \|u\| \|v\|. \quad \square$$

ヒント. 注意 2.23 を ${}^t A A$ に適用する. \square

参考 3.1 参考 2.26 で説明したように Gram 行列式の平方根は 2^r 個の点

$$\sum_{j \in I} a_j, \quad I \subset \{1, 2, \dots, r\}$$

を頂点とする \mathbb{R}^n 内の r 次元平行 2^r 面体の面積 (もしくは長さ, 体積) に等しい. \square

⁵¹問題 [57] のヒントと比較せよ.

3.4 Sylvester の行列式と判別式

[76] (Sylvester の行列式) 一変数 x の多項式

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = a_0 \prod_{i=1}^n (x - \alpha_i) \quad (a_0 \neq 0),$$

$$g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_{m-1}x + b_m = b_0 \prod_{j=1}^m (x - \beta_j) \quad (b_0 \neq 0)$$

に対して

$$\begin{vmatrix} a_0 & a_1 & \cdots & a_n & & & 0 \\ & a_0 & a_1 & \cdots & a_n & & \\ & & \ddots & & & \ddots & \\ 0 & & & a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_m & & & 0 \\ & b_0 & b_1 & \cdots & b_m & & \\ & & \ddots & & & \ddots & \\ 0 & & & b_0 & b_1 & \cdots & b_m \end{vmatrix} = a_0^m b_0^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j).$$

左辺の $m+n$ 次の行列式を **Sylvester の行列式** もしくは f, g の終結式 (**resultant**) と呼び, $R(f, g)$ と書くことにする. Sylvester の行列式の中で a_i たちは上から m 段を占めており, b_j たちは下から n 段を占めている. $R(f, g)$ は次のような表示も持つ:

$$R(f, g) = a_0^m \prod_{i=1}^m g(\alpha_i) = (-1)^{mn} b_0^n \prod_{j=1}^n f(\beta_j). \quad \square$$

ヒント. 佐武 [St] 第 II 章 §6 の 1 (70–74 頁). \square

[77] (判別式) 上の問題の続き. f の判別式 (**discriminant**) $D(f)$ を

$$D(f) := a_0 \prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i)^2$$

と定める. このとき

$$D(f) = (-1)^{\frac{n(n-1)}{2}} a_0^{-1} R(f, f').$$

ここで f' は f の導関数である. \square

[78] $f(x) = ax^2 + bx + c$ の判別式を求めよ⁵². \square

略解. $D(f) = -a^{-1} R(f, f') = b^2 - 4ac$. \square

[79] $f(x) = x^3 + ax^2 + bx + c$ の判別式を求めよ. $a = 0$ のときその判別式はどうなるか? \square

略解. $D(f) = -R(f, f') = -4b^3 + a^2b^2 - 4ca^3 + 18abc - 27c^2$. $a = 0$ のとき $D(f) = -4b^3 - 27c^2$. \square

[80] $f(x) = x^n + px + q$ の判別式を求めよ. \square

略解. $D(f) = (-1)^{\frac{n(n-1)}{2}} R(f, f')$, $R(f, f') = (-1)^{n-1} (n-1)^{n-1} p^n + n^n q^{n-1}$. \square

⁵² なつかしい感じのする式.

3.5 Jacobian

[81] (**Jacobian**) \mathbb{R}^n の開部分集合 U から \mathbb{R}^m の開部分集合 V への C^∞ 級写像 f が与えられたとする. f を次のように表わす:

$$y = f(x) = \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{bmatrix}.$$

このとき U 上の函数からなる次の $m \times n$ 行列 $dy/dx = df/dx$ を次のように定義する:

$$\frac{dy}{dx} = \frac{df}{dx} := \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

$m = n$ ならばこれを f の **Jacobi** 行列と呼び, その行列式を f の **Jacobian** と呼ぶ. 以下を示せ.

1. さらに $z = g(y)$ が V から \mathbb{R}^l の部分集合 W への C^∞ 写像であるとする. $y = f(x)$ と $z = g(y)$ の合成函数 $z = g(f(x))$ の偏微分に関する chain rule は上の記号のもとで次のように書くことができる:

$$\frac{dz}{dx} = \frac{dy}{dx} \frac{dz}{dy}.$$

これは形式上一変数函数の合成函数の微分法則と同じ形をしているがその実体は次のような行列の等式である:

$$\begin{bmatrix} \frac{\partial z_1}{\partial x_1} & \dots & \frac{\partial z_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial z_l}{\partial x_1} & \dots & \frac{\partial z_l}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial z_1}{\partial y_1} & \dots & \frac{\partial z_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial z_l}{\partial y_1} & \dots & \frac{\partial z_l}{\partial y_m} \end{bmatrix} \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \dots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}.$$

2. $m = n$ でかつ $U = V = \mathbb{R}^n$ でかつ

$$f(x) = Ax + b, \quad A \in M_n(\mathbb{R}), \quad b \in \mathbb{R}^n$$

の場合について考える. このとき f の Jacobi 行列は A に等しく, f の Jacobian は $|A|$ に等しい:

$$\frac{df}{dx} = A, \quad \det \frac{df}{dx} = |A|. \quad \square$$

参考 3.2 (逆写像定理) 任意の微分可能な f は各 $x_0 \in U$ の近傍で $g(x) = A(x - x_0) + f(x_0)$ (A は df/dx の $x = x_0$ での値) で近似される. (上の問題の 2 と同様に $dg/dx = A$ が成立している.) このように微分という手続きは曲がっていない函数 $g(x)$ で曲がっている函数 $f(x)$ を局所的に近似するという事なのである.

f は $x = x_0$ の近傍において g でよく近似されているのだから f の性質は $x = x_0$ の近傍で g と似ているはずである. もしも $m = n$ でかつ $|A| \neq 0$ ならば $y = g(x)$ は可逆な函数になり, $x = g^{-1}(y) = A^{-1}(y - f(x_0)) + x_0$ が成立する.

実は f が C^∞ 級ならば $|A| \neq 0$ のとき, f は $x = x_0$ の近傍で C^∞ 級の逆写像を持ち, g^{-1} で近似されることを示せる (逆写像定理). このような結果が成立することは「曲がったものを曲がっていないもので近似する」という発想ができれば直観的にかなり自然であることがわかる. \square

3.6 特性多項式の係数

[82] (特性多項式の係数) n 次正方行列 $A = [a_{ij}]$ の特性多項式を

$$p_A(\lambda) = \det(\lambda E - A) = \lambda^n - s_1 \lambda^{n-1} + s_2 \lambda^{n-2} - \cdots + (-1)^{n-1} s_{n-1} \lambda + (-1)^n s_n$$

と表わすと,

$$s_r = \sum_{1 \leq i_1 < \cdots < i_r \leq n} \begin{vmatrix} a_{i_1 i_1} & \cdots & a_{i_1 i_r} \\ \vdots & & \vdots \\ a_{i_r i_1} & \cdots & a_{i_r i_r} \end{vmatrix}. \quad \square$$

ヒント 1. この結果は佐武 [St] p.134 の問 1 にある. 解答は p.297 にある. \square

ヒント 2. A の第 j 列を a_j と表わすと, 問題 [67] の結果より次の公式が成立する:

$$(\lambda e_1 - a_1) \wedge \cdots \wedge (\lambda e_n - a_n) = \det(\lambda E - A) e_1 \wedge \cdots \wedge e_n.$$

たとえば $n = 3$ のとき

$$\begin{aligned} & (\lambda e_1 - a_1) \wedge (\lambda e_2 - a_2) \wedge (\lambda e_3 - a_3) \\ &= \lambda^3 e_1 \wedge e_2 \wedge e_3 + \lambda^2 (a_1 \wedge a_2 \wedge e_3 + a_1 \wedge e_2 \wedge a_3 + e_1 \wedge a_2 \wedge a_3) \\ &+ \lambda (a_1 \wedge e_2 \wedge e_3 + e_1 \wedge a_2 \wedge a_3 + e_1 \wedge e_2 \wedge a_3) + a_1 \wedge a_2 \wedge a_3. \quad \square \end{aligned}$$

3.7 行列式の微分と行列の指数関数の行列式

[83] (行列式の微分) $A(x) = [a_{ij}(x)]$ は x で微分可能な函数を成分に持つ n 次正方行列であるとし, $A(x)$ の余因子行列を $\Delta(x) = [\Delta_{ij}(x)]$ と表わす. そのとき

$$\frac{dA(x)}{dx} = \text{tr} \left({}^t \Delta(x) \frac{dA(x)}{dx} \right).$$

よって $|A(x)| \neq 0$ ならば

$$|A(x)|^{-1} \frac{dA(x)}{dx} = \text{tr} \left(A(x)^{-1} \frac{dA(x)}{dx} \right). \quad \square$$

ヒント. 佐武 [St] pp.84–85 を見よ. \square

[84] (行列の指数関数の行列式) 複素正方行列 X に対して $\det e^X = e^{\text{tr} X}$. \square

3.8 Pfaffian (パフィアン)

[85] (非自明な最も簡単な Pfaffian)

$$\begin{vmatrix} 0 & x_{12} & x_{13} & x_{14} \\ -x_{12} & 0 & x_{23} & x_{24} \\ -x_{13} & -x_{23} & 0 & x_{34} \\ -x_{14} & -x_{24} & -x_{34} & 0 \end{vmatrix} = (x_{12}x_{34} - x_{13}x_{24} + x_{14}x_{23})^2. \quad \square$$

[86] 条件 $x_{ji} = -x_{ij}$, $x_{ii} = 0$ を満たす文字 x_{ij} を成分に持つ n 次の交代行列を $X = [x_{ij}]$ と書く. n が奇数ならば $|X| = 0$ となり, n が偶数ならば $|X|$ は x_{ij} たちの多項式 $P_n(x)$ の二乗の形になる. しかもその $P_n(x)$ は $n = 2p$ のとき次のように表わされる:

$$\begin{aligned} P_n(x) &= \sum_{i_1 < i_2, \dots, i_{n-1} < i_n, i_1 < i_3 < \dots < i_{n-1}} \text{sign} \begin{pmatrix} 1 & \dots & n \\ i_1 & \dots & i_n \end{pmatrix} x_{i_1 i_2} x_{i_3 i_4} \dots x_{i_{n-1} i_n} \\ &= \frac{1}{2^p p!} \sum_{\sigma \in S_n} \text{sign}(\sigma) x_{\sigma(1)\sigma(2)} x_{\sigma(3)\sigma(4)} \dots x_{\sigma(n-1)\sigma(n)}. \end{aligned}$$

$P_n(x)$ を **Pfaffian** (パフィアン) と呼ぶ. \square

ヒント. 佐武 [St] pp.81–82 を見よ. \square

3.9 雑多な問題

[87] A は $m \times m$ 行列であり, B は $m \times n$ 行列であり, C は $n \times n$ 行列であるとする. 次が成立することを示せ:

$$\begin{vmatrix} A & B \\ 0 & C \end{vmatrix} = |A||C|.$$

さらにこの結果を用いて次を示せ:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{nn} \\ & a_{22} & \ddots & \vdots \\ & & \ddots & a_{n-1,n} \\ 0 & & & a_{nn} \end{vmatrix} = a_{11} a_{22} \dots a_{nn}. \quad \square$$

ヒント. 佐武 [St] pp.54–55 の例 4 を見よ. \square

[88] (複素行列を表現する実行列の行列式) A, B が実正方行列のとき

$$\begin{vmatrix} A & -B \\ B & A \end{vmatrix} = |\det(A + iB)|^2.$$

左辺の $||$ は行列式であり, 右辺の $||$ は複素数の絶対値であり, $i = \sqrt{-1}$ である. \square

参考 3.3 複素正方行列 $C = A + iB$ (A, B は実正方行列) に対して,

$$X(C) := \begin{bmatrix} A & -B \\ B & A \end{bmatrix}$$

と置くと, 複素正方行列 C, C' に対して,

$$X(C + C') = X(C) + X(C'), \quad X(CC') = X(C)X(C')$$

などが成立する. 上の問題の結果は $\det X(C) = |\det C|^2$ となることを意味している. \square

[89] A, B が実正方行列のとき

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix} = |A+B||A-B|. \quad \square$$

[90] (巡回行列式) ζ を 1 の原始 n 乗根とすると⁵³,

$$\begin{vmatrix} x_0 & x_1 & x_2 & \cdots & x_{n-1} \\ x_{n-1} & x_0 & x_1 & \cdots & x_{n-2} \\ x_{n-2} & x_{n-1} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & x_0 & x_1 \\ x_1 & \cdots & x_{n-2} & x_{n-1} & x_0 \end{vmatrix} = \prod_{i=1}^{n-1} (x_0 + \zeta^i x_1 + \zeta^{2i} x_2 + \cdots + \zeta^{(n-1)i} x_{n-1}). \quad \square$$

ヒント 1. 佐武 [St] p.79 を覗いてみよ. \square

ヒント 2. $\Lambda = E_{12} + E_{23} + \cdots + E_{n-1,n} + E_{n1}$ と置くと, 左辺の行列式は $X = x_0\Lambda^0 + x_1\Lambda^1 + \cdots + x_{n-1}\Lambda^{n-1}$ の行列式に等しい. ある可逆な行列 P で $P^{-1}\Lambda P = \text{diag}(1, \zeta, \zeta^2, \dots, \zeta^{n-1})$ をみたすものが存在することを示せ. そのとき $P^{-1}XP$ は対角行列になり, $|X| = |P^{-1}XP|$ を用いて X の行列式を容易に計算できる. \square

[91] ω を 1 の原始 3 乗根⁵⁴とすると,

$$a^3 + b^3 + c^3 - 3abc = \begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix} = (a+b+c)(a+\omega b+\omega^2 c)(a+\omega^2 b+\omega c). \quad \square$$

[92]

$$\begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_1 & x_4 & x_3 \\ x_3 & x_4 & x_1 & x_2 \\ x_4 & x_3 & x_2 & x_1 \end{vmatrix} = (x_1+x_2+x_3+x_4)(x_1-x_2+x_3-x_4)(x_1+x_2-x_3-x_4)(x_1-x_2-x_3-x_4). \quad \square$$

[93] (右下の縁に関する展開) n 次正方行列 $A = [a_{ij}]$ の (i, j) 余因子を Δ_{ij} と書き, $x = {}^t[x_1, \dots, x_n]$, $y = {}^t[y_1, \dots, y_n]$ と置くと,

$$\begin{vmatrix} A & x \\ {}^t y & z \end{vmatrix} = |A|z - \sum_{i,j=1}^n \Delta_{ij} x_i y_j. \quad \square$$

ヒント. 佐武 [St] pp.59–60 の例 3 を見よ. \square

⁵³ ζ が 1 の原始 n 乗根であるとは $k = 1, 2, \dots, n-1$ のとき $\zeta^k \neq 1$ であつ $\zeta^n = 1$ となることである. たとえば $\zeta = e^{2\pi i/n}$ は原始 n 乗根である. それ以外の原始 n 乗根は n と互いに素な整数 k によつて $\zeta^k = e^{2\pi i k/n}$ と表わされることが知られている. たとえば 1 の原始 12 乗根全体の集合は $\{e^{2\pi i k/12} \mid k = 1, 5, 7, 11\}$ になる.

⁵⁴1 の 3 乗根が 1 の原始 3 乗根であるための必要十分条件はそれが 1 でないことである. 一般に ω が 1 の原始 n 乗根であるとは $\omega^n = 1$ であつ $k = 1, 2, \dots, n-1$ に対して $\omega^k \neq 1$ であることである.

[94] すべての成分が 1 であるような n 次正方行列を A と書くと,

$$\det(\lambda E - A) = \lambda^{n-1}(\lambda - n). \quad \square$$

[95] A, B, C, D はそれぞれ $m \times m, m \times n, n \times m, n \times n$ 行列であるとする. もしも A が可逆ならば

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A| |D - CA^{-1}B|$$

が成立する. 同様にもしも D が可逆ならば

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A - BD^{-1}C| |D|$$

が成立し, さらに $m = n$ かつ $CD = DC$ ならば

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - BC|$$

が成立する. \square

[96] $|A + B| = |A| + |B|$ が偶然成立する非自明な例と $|A + B| = |A| + |B|$ が成立しない例を構成せよ. \square

[97]

$$\begin{vmatrix} 0 & a^2 & b^2 & 1 \\ a^2 & 0 & c^2 & 1 \\ b^2 & c^2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = (a + b + c)(a + b - c)(a - b + c)(a - b - c). \quad \square$$

4 一次方程式の理論

記号法: 以下において K は実数体 \mathbb{R} または複素数体 \mathbb{C} であるとし, 体 K 係数の一次方程式の理論について説明する⁵⁵. 今まで通り, K の元を成分に持つ $m \times n$ 行列全体の集合を $M_{m,n}(K)$ と書き, K の元を成分に持つ n 次正方行列全体の集合を $M_n(K) = M_{n,n}(K)$ と書き, K の元を成分に持つ n 次元縦ベクトル全体の集合を $K^n = M_{n,1}(K)$ と書くことにする.

4.1 一次方程式とその解空間

定数 $a_{ij}, b_i \in K$ が与えられたとき, x_i たちに関する次の方程式を一次方程式 (線形方程式, linear equation) と呼ぶ:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1,$$

⁵⁵実際には K は任意の体であるとして構わない. しかし, 「任意の体」という抽象代数の言葉に慣れていない方のために K は \mathbb{R} または \mathbb{C} であることにしている.

$$\begin{aligned}
 a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\
 &\dots\dots\dots \\
 a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
 \end{aligned}$$

すべての b_i が 0 であるときこの一次方程式は斉次 (せいじ, 同次, **homogeneous**) であると言い, ある b_i が 0 でないときこの一次方程式は非斉次 (非同次, **inhomogeneous**) であると言う.

$m \times n$ 行列 A と m 次元縦ベクトル b と n 次元縦ベクトル x を次のように定める:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

このとき上の一次方程式を次のように書くことができる:

$$Ax = b.$$

行列 $A \in M_{m,n}(K)$ とベクトル $b \in K^m$ に対して, 一次方程式 $Ax = b$ を解くとは $Ax = b$ を満たす x 全体の集合

$$\mathcal{S} = \{x \in K^n \mid Ax = b\} \subset K^n$$

を求めることである. \mathcal{S} を一次方程式 $Ax = b$ の解空間と呼ぶことにする. A の核 (**kernel**) $\text{Ker } A$ を次のように定める:

$$\text{Ker } A := \{x \in K^n \mid Ax = 0\} \subset K^n.$$

A の核 $\text{Ker } A$ は斉次の一次方程式 $Ax = 0$ の解空間に一致する. さらに A の像 (**image**) $\text{Im } A$ を次のように定める:

$$\text{Im } A := \{Ax \mid x \in K^n\} \subset K^m.$$

一次方程式の理論は抽象的には $\text{Ker } A$ や $\text{Im } A$ に関する理論であるとみなせる. $\text{Ker } A$ は斉次の一次方程式 $Ax = 0$ の解空間そのものであり, $\text{Im } A$ は非斉次の一次方程式 $Ax = b$ が解を持つ b 全体の集合に一致している (次の問題を見よ).

[98] \mathcal{S} は行列 $A \in M_{m,n}(K)$ とベクトル $b \in K^m$ に対する一次方程式 $Ax = b$ の解空間であるとする. このとき以下が成立する:

1. 一次方程式 $Ax = b$ の解が一つ以上存在するための必要十分条件は b が A の像に含まれること (すなわち $b \in \text{Im } A$) が成立することである. 解空間 \mathcal{S} は空集合になり得る (例を挙げよ).
2. 解空間 \mathcal{S} は空集合でないと仮定し, 任意に $v \in \mathcal{S}$ を取る. このとき \mathcal{S} の任意の元は v と $\text{Ker } A$ の元の和の形で一意に表わされる. すなわち, 方程式 $Ax = b$ の解 v が一つ存在すれば, 方程式 $Ax = b$ の任意の解は v と斉次の一次方程式 $Ax = 0$ の解 w の和の形で一意に表わされる. 特に次が成立する:

$$\mathcal{S} = v + \text{Ker } A = \{v + w \mid w \in \text{Ker } A\}. \quad \square$$

ヒント. 1. $\text{Im } A$ の定義より, $b \in \text{Im } A$ とある $x \in K^n$ で $Ax = b$ を満たすものが存在することは同値である. $m = n = 1$ の一次方程式 $\boxed{?}x_1 = 1$ の解は存在しない.

2. $Av = b$, $Av' = b$ とすると $A(v' - v) = 0$ である. よって $w = v' - v$ と置くと $w \in \text{Ker } A$ であつ $v' = v + w$ である. \square

注意 4.1 問題 [98] の結果より, $Ax = b$ を解く手続きを $\text{Im } A$ と $\text{Ker } A$ を用いて以下のように説明することができる:

1. b が A の像に含まれているかどうか (すなわち $b \in \text{Im } A$ であるかどうか) を確かめる. もしも $b \notin \text{Im } A$ ならば解空間は空集合になる.
2. もしも $b \in \text{Im } A$ ならば $Ax = b$ の解 v を一つ求める.
3. 斉次な一次方程式 $Ax = 0$ の解空間 $\text{Ker } A$ を求める.
4. 一次方程式 $Ax = b$ の解空間は $v + \text{Ker } A$ に等しい. \square

$m \times (n + 1)$ 行列 A' と $n + 1$ 次元縦ベクトル x' を

$$A' = [A, b] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}, \quad x' = \begin{bmatrix} x \\ -1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ -1 \end{bmatrix}$$

と定めると一次方程式 $Ax = b$ は次と同値になる:

$$A'x' = 0.$$

これは見かけ上, 斉次な一次方程式の形をしている. 縦ベクトル x' の最後の成分が -1 に固定されているおかげで, 見かけ上斉次な一次方程式 $A'x' = 0$ と斉次とは限らない一次方程式 $Ax = b$ が同値になるのである. 非斉次な一次方程式を扱う場合には $A' = [A, b]$ という行列が非常に役に立つ⁵⁶.

以下の目標は行列 A もしくは $A' = [A, b]$ を用いて一次方程式の理論について説明することである.

⁵⁶ A' と x' を次のように定める流儀もある:

$$A' = [A, -b] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & -b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & -b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & -b_m \end{bmatrix}, \quad x' = \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ 1 \end{bmatrix}.$$

理論的にはこちらの脚註の流儀を採用した方が自然なのだが, 具体的な計算をする場合には本文の流儀を採用した方が便利だと思われる. 具体的な計算に応用するために本文の流儀を採用することにした. (実際にはどちらの流儀を採用しても大した違いはない.)

この脚註の流儀の方が自然な理由を理解するためには n 次元射影空間 $\mathbb{P}^n(K)$ とその上の非斉次座標について学ぶ必要がある. $x' = {}^t[x_1, \dots, x_n, x_{n+1}]$ の x_{n+1} を -1 に固定するよりも 1 に固定する方が理論的には自然であろう.

4.2 Cramér の公式について

$m = n$ すなわち A が正方行列であり、 A が可逆である場合には一次方程式 $Ax = b$ は $x = A^{-1}b$ と一意に解ける。そして、 A^{-1} は A の余因子と行列式を用いて表示可能である。その表示を $x = A^{-1}b$ に代入すれば Cramér の公式が得られる。Cramér の公式については第 2.2 節で詳しく説明してあるのでそちらを参照せよ。

Cramér の公式は一次方程式の解を行列式を用いて表わす理論的に非常に重要な公式である。しかし、第 2.2 節でも説明したように、大きな行列式の計算には大変な手間がかかるので、具体的な一次方程式を解くために Cramér の公式を利用するのは効率が悪い。だから、一次方程式を解くための別の方法を知っておく必要がある。

さらに、Cramér の公式を直接適用できない場合、すなわち $m = n$ だが A が逆行列を持たない場合や $m \neq n$ の場合には一次方程式をどのようにして解けば良いかについても考えなければいけない。

以下では以上で述べたような問題を扱うことになる。

4.3 行列の基本操作と基本変形の導入

一次方程式に限らず方程式を解くための基本は「方程式をうまく変形してより簡単な形にすること」である。一次方程式に関するそのような手続きは行列の**基本操作 (elementary operation)** と**基本変形 (elementary transformation)** の理論の形で整備されている⁵⁷。より抽象的には一般線形群の作用の言葉で整理されることになる。

K の元を成分に持つ可逆な m 次正方行列全体の集合を次のように表わす：

$$GL_m(K) = \{ P \in M_m(K) \mid P \text{ は逆行列を持つ} \}.$$

$GL_m(K)$ は一般線形群 (**general linear group**) と呼ばれる。実際に群であることを示すのが次の問題である。

[99] (一般線形群) $GL_m(K)$ に関して以下が成立している：

1. 任意の $P, Q \in GL_m(K)$ に対して $PQ \in GL_m(K)$ であり、任意の $P, Q, R \in GL_m(K)$ に対して $(PQ)R = P(QR)$ が成立している。
2. m 次の単位行列を E と書くと $E \in GL_m(K)$ であり、任意の $P \in GL_m(K)$ に対して $PE = EP = P$ が成立している。
3. 任意の $P \in GL_m(K)$ に対して $P^{-1} \in GL_m(K)$ であり、 $PP^{-1} = P^{-1}P = E$ が成立している。□

⁵⁷ 行列の基本変形という考え方は後で Jordan 標準形の理論などの基礎になる単因子論 (elementary divisor theory) でも重要になる。単因子論入門には堀田 [H2] がおすすめである。

体 K の元を成分に持つ行列の基本変形の理論は K 係数の一次方程式の理論である。扱う環を体 K から \mathbb{Z} や一変数多項式環 $K[\lambda]$ のような環を含む Euclid 整域 (もしくはさらに一般的に単項イデアル整域) に一般化すると、一次方程式の理論は単因子論に拡張される。

扱う環をさらに一般にすると一次方程式の理論もどんどん複雑になる。現代の数学において環上の一次方程式の理論は環上の加群 (module) の理論として整備されている。環上の加群は体上のベクトル空間の一般化である。

なお堀田 [H1], [H2] では基本操作を**基本変形 (fundamental transformation)** と呼び、基本変形を**初等変形 (elementary transformation)** と呼んでいる。色々調べてみたが標準的な用語法は決まっていないうようである。他の文献を見るときには注意して欲しい。

$i, j = 1, \dots, m$ かつ $i \neq j$ であり, $\alpha \in K$ のとき, $U_{ij}(\alpha) \in GL_m(K)$ を次のように定める:

$$U_{ij}(\alpha) = E + \alpha E_{ij} = \begin{bmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & \cdots & \alpha \\ & & \vdots & \ddots & \vdots \\ & & 0 & \cdots & 1 \\ & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix}.$$

ここで $i < j$ ならばここに書いたように α は対角線の右上に来るが, $i > j$ ならば α は対角線の左下に来ることに注意せよ. E_{ij} は (i, j) 成分のみが 1 で他の成分がすべて 0 であるような行列 (行列単位) であり, E は m 次の単位行列である.

$i, j = 1, \dots, m$ かつ $i \neq j$ のとき, $P_{ij} \in GL_m(K)$ を次のように定める:

$$P_{ij} = E_{ij} + E_{ji} + \sum_{k \neq i, j} E_{kk} = \begin{bmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 0 & \cdots & 1 \\ & & \vdots & \ddots & \vdots \\ & & 1 & \cdots & 0 \\ & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix}.$$

$i = 1, \dots, m$ であり, $\beta \in K$ かつ $\beta \neq 0$ のとき, $D_i(\beta) \in GL_m(K)$ を次のように定める:

$$D_i(\beta) = \beta E_{ii} + \sum_{k \neq i} E_{kk} = \text{diag}(1, \dots, \beta, \dots, 1) = \begin{bmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & \beta & \\ & & & & 1 \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix}.$$

以上で定義した行列 $U_{ij}(\alpha)$, P_{ij} , $D_i(\beta)$ に関して m を陽に示したい場合にはそれぞれを $U_{m;ij}(\alpha)$, $P_{m;ij}$, $D_{m;i}(\beta)$ と書くことにする⁵⁸.

[100] $U_{ij}(\alpha)^{-1} = U_{ij}(-\alpha)$, $P_{ij}^{-1} = P_{ij}$, $D_i(\beta)^{-1} = D_i(\beta^{-1})$ を示せ. \square

体 K の元を成分に持つ $m \times n$ 行列 $A = [a_{ij}] \in M_{m,n}(K)$ が与えられたとする.

A の行に関する基本操作 (elementary operations of rows) とは次の 3 種類の操作のことである:

⁵⁸ $U_{ij}(\alpha)$, P_{ij} , $D_i(\beta)$ の U , P , D はそれぞれ unipotent (巾単), permutation (置換), diagonal (対角) という意味のつもりである. 単位行列と巾零行列 (nilpotent matrix) の和の形の行列を unipotent matrix (巾単行列) と呼ぶ. 正方向行列が巾零であるとはそのある巾が零になることである.

- (a) A のある行の定数倍を他の行に加える.
- (b) A の 2 つの行を交換する.
- (c) A のある行に 0 でない定数をかける.

ここで定数は K の元を意味するものとする. 有限回の行に関する基本操作によって実現できる A の変形を行に関する基本変形 (elementary transformations of rows) と呼ぶ.

A の列に関する基本操作 (elementary operations of columns) とは次の 3 種類の操作のことである:

- (a') A のある列の定数倍を他の列に加える.
- (b') A の 2 つの列を交換する.
- (c') A のある列に 0 でない定数をかける.

ここで定数は K の元を意味するものとする. 有限回の列に関する基本操作によって実現できる A の変形を列に関する基本変形 (elementary transformations of columns) と呼ぶ.

行に関する基本操作と列に関する基本操作を合わせて行列の基本操作 (elementary operations of matrices) と呼び, 行に関する基本変形と列に関する基本変形の合成を行列の基本変形 (elementary transformations of matrices) と呼ぶことにする.

[101] (行列の基本操作の可逆な行列の積による実現) $A \in M_{m,n}(K)$ であるとする. A の行に関する基本操作は上で定義した行列 $U_{m;ij}(\alpha)$, $P_{m;ij}$, $D_{m;i}(\beta)$ のどれかを A に左からかける操作で実現可能である. より正確に言えば以下が成立している:

1. $i, j = 1, \dots, m$ かつ $i \neq j$ であり, $\alpha \in K$ のとき, A の第 j 行の α 倍を第 i 行に加える基本操作は $U_{m;ij}(\alpha)$ を A に左からかける操作に一致する.
2. $i, j = 1, \dots, m$ かつ $i \neq j$ のとき, A の第 i 行と第 j 行を交換する基本操作は $P_{m;ij}$ を A に左からかける操作に一致する.
3. $i = 1, \dots, m$ であり, $\beta \in K$, $\beta \neq 0$ のとき, A の第 i 行を β 倍する基本操作は $D_{m;i}(\beta)$ を A に左からかける操作に一致する.

同様に A の列に関する基本操作は $U_{n;ij}(\alpha)$, $P_{n;ij}$, $D_{n;i}(\beta)$ を A に右からかける操作で実現可能である. \square

ヒント. 実際に $U_{m;ij}(\alpha)A$, $P_{m;ij}A$, $D_{m;i}(\beta)A$ を計算してみればよい. 列に関する基本操作に関する結果は行列の転置を考えれば得られる. \square

注意 4.2 上の問題 [101] の結果より, 行列 A の行に関する基本変形 (行に関する基本操作の有限個の合成) は有限個の $U_{m;ij}(\alpha)$, $P_{m;ij}$, $D_{m;i}(\beta)$ たちの積 $P \in GL_m(K)$ を A に左からかける変換

$$A \mapsto PA, \quad P \in GL_m(K)$$

で実現可能である. 同様に A の列に関する基本変形 (列に関する基本操作の有限個の合成) は有限個の $U_{n;ij}(\alpha)$, $P_{n;ij}$, $D_{n;i}(\beta)$ たちの積 $Q \in GL_n(K)$ を A に右からかける変換

$$A \mapsto AQ, \quad Q \in GL_n(K)$$

で実現可能である.

一次方程式 $Ax = b$ を解くことは $Ax = b$ を満たす x 全体の集合 (解空間) を求めることであった. よって, 解くために施される方程式の変形は同値変形でなければいけない. なぜならば, 方程式の変形によって解全体の集合が増えたり減ったりするとまずいからである.

可逆な $P \in GL_m(K)$ と $Q \in GL_m(K)$ に対して

$$\tilde{A} = PAQ, \quad \tilde{b} = Pb, \quad \tilde{x} = Q^{-1}x$$

と置けば

$$Ax = b \iff PAQQ^{-1}x = Pb \iff \tilde{A}\tilde{x} = \tilde{b}$$

が成立している. したがって, 適当な行列の基本変形によって A をより簡単な形をしている \tilde{A} に変形できればもとの方程式 $Ax = b$ はより簡単な形の方程式 $\tilde{A}\tilde{x} = \tilde{b}$ に同値変形されることになる.

変数変換された結果の \tilde{x} ではなく, もとの x のレベルで解を表示したいことが多い. その場合には行だけに関する基本変形を適用した場合 (そのとき Q が単位行列になる) を考えれば良い.

ここまで説明すれば行列の基本変形が一次方程式論で基本的な役目を果たすことが納得できるだろう.

基本的な問題は次の 2 つである:

- 行にだけ関する基本変形で行列をどれだけ簡単な形にできるか?
- 行と列に関する基本変形を用いて行列をどれだけ簡単な形にできるか?

前者については問題 [109] で扱い, 後者については問題 [111] で扱うことにする. \square

参考 4.3 上の注意への補足. 任意の $P \in GL_m(K)$ が有限個の $U_{ij}(\alpha)$, P_{ij} , $D_i(\beta)$ たちの積で表示可能なことはこの時点では証明されていないので, 任意の $P \in GL_m(K)$ に対する変換 $A \mapsto PA$ が A の行に関する基本変形になっているかどうかはまだわからないということにしなければいけない.

しかし, 任意の $P \in GL_m(K)$ が有限個の $U_{ij}(\alpha)$, P_{ij} , $D_i(\beta)$ たちの積で表示可能である⁵⁹ことを実際に証明できる⁶⁰. したがって, A の行に関する基本変形は $GL_m(K)$ の元を左からかける変換に一致する.

単に可逆な行列を定義するだけでは, 具体的にどのような正方行列が可逆になるのかよくわからない. しかし, 可逆な行列が常に $U_{ij}(\alpha)$, P_{ij} , $D_i(\beta)$ のような基本的な行列たちの有限個の積で表わされることが証明されたならば (実際に証明される), 可逆な行列を系統的に生成する方法が得られたことになる⁶¹. \square

4.4 行列の rank の定義

[102] (一次独立性) $v_1, \dots, v_n \in K^m$ に対して以下の条件は互いに同値である:

- (a) 任意の $\alpha_1, \dots, \alpha_n \in K$ に対して $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ ならば $\alpha_1 = \dots = \alpha_n = 0$ である.

⁵⁹ $U_{ij}(\alpha)$, P_{ij} , $D_i(\beta)$ たちは一般線形群 $GL_m(K)$ の生成元 (generators) であると言う.

⁶⁰証明の粗筋を問題 [112] のヒントで説明する.

⁶¹数学に限らず, 科学のイロハのイは「複雑に見える問題をより単純な問題に分解すること」である.

- (b) 任意の $\alpha_i, \beta_i \in K$ ($i = 1, \dots, n$) に対して $\alpha_1 v_1 + \dots + \alpha_n v_n = \beta_1 v_1 + \dots + \beta_n v_n$ ならば $\alpha_i = \beta_i$ ($i = 1, \dots, n$) である.
 (c) $v_1 \neq 0$ かつ $v_{i+1} \notin K v_1 + \dots + K v_i$ ($i = 1, \dots, n-1$).

ここで $K v_1 + \dots + K v_i$ は次のように定義された集合である:

$$K v_1 + \dots + K v_i = \{ \alpha_1 v_1 + \dots + \alpha_i v_i \mid \alpha_1, \dots, \alpha_i \in K \}.$$

上の同値な条件のどれかが成立するとき, v_1, \dots, v_n は一次独立 (linearly independent) であるという⁶². 一次独立でないときは一次従属 (linearly dependent) であるという. さらに次の条件は v_1, \dots, v_n が一次独立であるという条件に同値である:

- (d) v_1, \dots, v_{n-1} は一次独立でかつ $v_n \notin K v_1 + \dots + K v_{n-1}$. \square

ヒント. (b) に登場する式の右辺を左辺に移項すれば (a) から (b) が出ることがわかる. (a) は (b) の特殊な場合なので (b) ならば (a) である. (c) の否定から (a) の否定が導かれるので (a) ならば (c) である. (d) は (c) の言い換えとみなせる. あとは (c) ならば (a) を示せばよい. その対偶を示すために (a) が成立していないと仮定する. そのとき, ある $\alpha_1, \dots, \alpha_n \in K$ でそれらのうちどれか一つは 0 でなくてかつ $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ を満たすものが存在する. $\alpha_i \neq 0$ であるような最大の i を r と書き, $\beta_i = -\alpha_i / \alpha_r$ と置く. このとき $v_r = \beta_1 v_1 + \dots + \beta_{r-1} v_{r-1} \in K v_1 + \dots + K v_{r-1}$ である. これで (a) の否定から (c) の否定が導かれることがわかった. \square

[103] 以下のベクトルの組は一次独立であるか?

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}. \quad \square$$

略解. $v_3 = v_2 + (v_2 - v_1) = -v_1 + 2v_2$ なので一次独立ではない. \square

[104] 以下のベクトルの組が一次従属になるような $a \in K$ をすべて求めよ:

$$v_1 = \begin{bmatrix} a \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ a \\ 1 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 1 \\ a \\ a \end{bmatrix}. \quad \square$$

略解. $a = 1$ のとき $v_1 = v_2 = v_3$ なので一次従属になる. $a \neq 1$ のとき v_1, v_2 は一次独立になるので v_1, v_2, v_3 が一次従属になるための必要十分条件はある $\alpha, \beta \in K$ で $v_3 = \alpha v_1 + \beta v_2$ を満たすものが存在することである. この条件は $a \neq 1, \alpha, \beta$ に関する連立方程式になる. その方程式の解は唯一であり, $a = -2, \alpha = \beta = -1$ であることがわかる. 以上によって, v_1, v_2, v_3 が一次従属になるための必要十分条件は $a = 1, -2$ であることがわかった. \square

定義 4.4 (行列の rank) $A = [a_{ij}]$ は体 K の元を成分に持つ $m \times n$ 行列であるとし, その第 j 列を $a_j \in K^m$ と書くことにする. このとき, A は $A = [a_1, \dots, a_n]$ と表わされる. a_1, \dots, a_n の一次独立な部分集合の元の個数の最大値を A の階数 (rank) と呼び, $\text{rank } A$ と表わす. \square

⁶²大抵の教科書では (a) の条件で一次独立性を定義している. 他の文献を読むときには注意して欲しい.

[105] $A \in M_m(K)$, $B \in M_n(K)$ のとき

$$\text{rank} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \text{rank } A + \text{rank } B. \quad \square$$

ヒント. A の第 i 列を a_i と書き, B の第 j 列を b_j と書くことにする. $X = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ と置く.

X の第 k 列を x_k と書くと, $k \leq m$ のとき $x_k = \begin{bmatrix} a_k \\ 0 \end{bmatrix}$ であり, $k > m$ のとき $x_k = \begin{bmatrix} 0 \\ b_{k-m} \end{bmatrix}$ である. $a_{i_1}, \dots, a_{i_r} \in K^m$ は一次独立であり, $b_{j_1}, \dots, b_{j_s} \in K^n$ も一次独立であるとする. このとき, $x_{i_1}, \dots, x_{i_r}, x_{j_1+m}, \dots, x_{j_s+m} \in K^{m+n}$ も一次独立であることが直接に確かめられる. よって $\text{rank } X \geq \text{rank } A + \text{rank } B$ である. 逆に $x_{i_1}, \dots, x_{i_r}, x_{j_1+m}, \dots, x_{j_s+m}$ が一次独立ならば, $a_{i_1}, \dots, a_{i_r} \in K^m$ も $b_{j_1}, \dots, b_{j_s} \in K^n$ も一次独立であることがわかる. よって $\text{rank } A + \text{rank } B \geq \text{rank } X$ である. \square

注意 4.5 $\text{rank} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 1$ であるから, $0 \neq C \in M_{m,n}(K)$ のとき, 等式 $\text{rank} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \text{rank } A + \text{rank } B$ は一般には成立しない. \square

[106] (行列の基本変形による rank の不変性) $A = [a_{ij}]$ は体 K の元を成分に持つ $m \times n$ 行列であるとする. 以下を示せ:

1. 行に関する基本変形で行列 A の rank は不変である.
2. 列に関する基本変形で行列 A の rank は不変である. \square

ヒント. 1. 注意 4.2 より任意の $P \in GL_m(K)$ に対して A と PA の rank は等しいことを示せば十分である. $P \in GL_m(K)$ とする. $PA = [Pa_1, \dots, Pa_n]$ である. P が逆行列を持つことより, a_{i_1}, \dots, a_{i_s} が一次独立であれば $Pa_{i_1}, \dots, Pa_{i_s}$ も一次独立であることがわかる. よって PA の rank は少なくとも A の rank 以上である: $\text{rank } A \leq \text{rank } PA$. 同じ議論を A, P の代わりに PA, P^{-1} に適用すると, $\text{rank } PA \leq \text{rank } P^{-1}(PA) = \text{rank } A$ であることがわかる. したがって $\text{rank } A = \text{rank } PA$.

2. $i \neq j$, $\alpha \in K$ とし, A の第 i 列の α 倍を第 j 列に加えてできる行列を A' とし, その第 k 列を $a'_k \in K^m$ と表わす. a_{i_1}, \dots, a_{i_s} は一次独立であると仮定する. もしも j が i_1, \dots, i_s の中に含まれていなければ $a'_{i_1} = a_{i_1}, \dots, a'_{i_s} = a_{i_s}$ は一次独立である. j は i_1, \dots, i_s の中に含まれていると仮定する. $i_1 = j$ と仮定してよい. もしも i が i_2, \dots, i_s の中に含まれているならば $i = i_2$ としてよく, $a'_{i_1} = a_{i_1} + \alpha a_{i_2}, a'_{i_2} = a_{i_2}, \dots, a'_{i_s} = a_{i_s}$ も一次独立であることがわかる. i は i_2, \dots, i_s に含まれていないと仮定する. もしも $a_{i_1}, \dots, a_{i_s}, a_i$ が一次独立ならば $a'_{i_1} = a_{i_1} + \alpha a_i, a'_{i_2} = a_{i_2}, \dots, a'_{i_s} = a_{i_s}, a'_i = a_i$ も一次独立である. $a_{i_1}, \dots, a_{i_s}, a_i$ は一次従属であると仮定する. $\alpha_{i_1}, \dots, \alpha_{i_s}$ は一次独立と仮定したので a_i は $a_i = \beta_1 a_{i_1} + \dots + \beta_s a_{i_s}$ ($\beta_\nu \in K$) と表わされる. もしも $\beta_1 = 0$ ならば $a'_{i_1} = a_{i_1} + \alpha \beta_2 a_{i_2} + \dots + \alpha \beta_s a_{i_s}, a'_{i_2} = a_{i_2}, \dots, a'_{i_s} = a_{i_s}$ は一次独立である. もしも $\beta_1 \neq 0$ ならば $a'_i = a_i = \beta_1 a_{i_1} + \dots + \beta_s a_{i_s}, a'_{i_2} = a_{i_2}, \dots, a'_{i_s} = a_{i_s}$ は一次独立である. 以上によって $\text{rank } A \leq \text{rank } A'$ であることがわかった. A と A' の立場を逆転させれば $\text{rank } A' \leq \text{rank } A$ であることもわかる. これである列の定数倍を他の列に加える基本操作で行列の rank が不変であることがわかった. 2つの列の交換とある列に 0 でない定数をかける基本操作で rank が不変であることは容易に確かめられる. \square

参考 4.6 定義 4.4 における行列の rank の定義はあまり格好良くない. もしも行列が定める線形写像の像の次元の概念を用いて良いならば行列 A の rank を次で定義することもできる:

$$\text{rank } A = \dim \text{Im } A.$$

問題 [127] を見よ. より一般に有限次元ベクトル空間のあいだの線形写像 $f: V \rightarrow W$ に対してその rank が $\text{rank } f = \dim \text{Im } f = \dim f(V)$ と定義される. \square

[107] $A, B \in M_n(K)$ のとき⁶³

$$\text{rank} \begin{bmatrix} A & B \\ B & A \end{bmatrix} = \text{rank}(A+B) + \text{rank}(A-B). \quad \square$$

ヒント 1. 問題 [105], [106] の結果を使う. より簡単な行列式に関する問題 [89] を解いてみるとヒントが得られるかもしれない. それでも駄目なら次のヒントを見よ. \square

ヒント 2. 行列の基本変形によって次のような変形が可能である:

$$\begin{aligned} \begin{bmatrix} A & B \\ B & A \end{bmatrix} &\rightarrow \begin{bmatrix} A & B \\ B-A & A-B \end{bmatrix} \rightarrow \begin{bmatrix} 2A & 2B \\ B-A & A-B \end{bmatrix} \\ &\rightarrow \begin{bmatrix} 2A+(B-A) & 2B+(A-B) \\ B-A & A-B \end{bmatrix} = \begin{bmatrix} A+B & A+B \\ B-A & A-B \end{bmatrix} \\ &\rightarrow \begin{bmatrix} A+B+(A+B) & A+B \\ B-A+(A-B) & A-B \end{bmatrix} = \begin{bmatrix} 2(A+B) & A+B \\ 0 & A-B \end{bmatrix} \\ &\rightarrow \begin{bmatrix} A+B & A+B \\ 0 & A-B \end{bmatrix} \rightarrow \begin{bmatrix} A+B & 0 \\ 0 & A-B \end{bmatrix}. \end{aligned}$$

これは実質的に次の計算をしたことに相当している:

$$\frac{1}{2} \begin{bmatrix} E & E \\ -E & E \end{bmatrix} \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} E & -E \\ E & E \end{bmatrix} = \begin{bmatrix} A+B & 0 \\ 0 & A-B \end{bmatrix}. \quad \square$$

[108] $A, B \in M_n(\mathbb{C})$ のとき

$$\text{rank} \begin{bmatrix} A & -B \\ B & A \end{bmatrix} = \text{rank}(A+iB) + \text{rank}(A-iB). \quad \square$$

ヒント 1. 問題 [105], [106] の結果を使う. より簡単な行列式に関する問題 [88] を解いてみるとヒントが得られるかもしれない. それでも駄目なら次のヒントを見よ. \square

ヒント 2. 行列の基本変形によって次のような変形が可能である:

$$\begin{aligned} \begin{bmatrix} A & -B \\ B & A \end{bmatrix} &\rightarrow \begin{bmatrix} A & -B \\ B+iA & A-iB \end{bmatrix} \rightarrow \begin{bmatrix} 2A & -2B \\ B+iA & A-iB \end{bmatrix} \\ &\rightarrow \begin{bmatrix} 2A+i(B+iA) & -2B+i(A-iB) \\ B+iA & A-iB \end{bmatrix} = \begin{bmatrix} A+iB & iA-B \\ B+iA & A-iB \end{bmatrix} \\ &\rightarrow \begin{bmatrix} A+iB-i(iA-B) & iA-B \\ B+iA-i(A-iB) & A-iB \end{bmatrix} = \begin{bmatrix} 2(A+iB) & iA-B \\ 0 & A-iB \end{bmatrix} \end{aligned}$$

⁶³一般の体を考える場合には K の標数は 2 でないと仮定する. すなわち 2 で自由に割れると仮定する.

$$\rightarrow \begin{bmatrix} A+iB & i(A+iB) \\ 0 & A-iB \end{bmatrix} \rightarrow \begin{bmatrix} A+iB & 0 \\ 0 & A-iB \end{bmatrix}.$$

これは実質的に次の計算をしたことに相当している:

$$\frac{1}{2} \begin{bmatrix} E & iE \\ iE & E \end{bmatrix} \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} E & -iE \\ -iE & E \end{bmatrix} = \begin{bmatrix} A+iB & 0 \\ 0 & A-iB \end{bmatrix}. \quad \square$$

4.5 行列の基本変形による行列の簡単化

[109] (行に関する基本変形による階段行列への変換) $A = [a_{ij}]$ は体 K の元を成分に持つ $m \times n$ 行列であるとする. このとき A はある行の定数倍を他の行に加える基本操作 (a) と 2 つの行を交換する基本操作 (b) の有限回の繰り返しによって次の形に変形可能である:

$$\tilde{A} = \begin{bmatrix} & \begin{array}{c} \boxed{c_1} \end{array} & & & * \\ & & \begin{array}{c} \boxed{c_2} \end{array} & & \\ & & & \ddots & \\ & & & & \begin{array}{c} \boxed{c_r} \end{array} \\ 0 & & & & \end{bmatrix}, \quad c_1, \dots, c_r \neq 0.$$

ここで \tilde{A} の 0 でない成分は右上の (逆さ) 階段状の部分にのみ存在し得る. すべての成分が 0 になる \tilde{A} の左端の数列とすべての成分が 0 になる \tilde{A} の下端の数行が存在しないこともあり得る⁶⁴. 上の \tilde{A} の形の行列を階段行列と呼ぶことにする.

さらにある行に 0 でない定数をかけるという基本操作 (c) を用いることによって, 行だけに関する基本変形で A を次の形に変形できることがわかる:

$$\tilde{\tilde{A}} = \begin{bmatrix} & \begin{array}{c} \boxed{1} \end{array} & & & * \\ & & \begin{array}{c} \boxed{1} \end{array} & & \\ & & & \ddots & \\ & & & & \begin{array}{c} \boxed{1} \end{array} \\ 0 & & & & \end{bmatrix}.$$

この $\tilde{\tilde{A}}$ の形の行列を正規化された階段行列と呼ぶことにする. \square

ヒント. まず, 例 4.7 を読み, 問題 [110] を解いてみて, 感じをつかんでみよ. そうしておけば以下の手続きを納得し易いだろう.

以下のような手続きで A に対して行に関する基本操作 (a),(b) を適用する:

I. A の第 1 列の成分がすべて 0 であるとき,

1. $n = 1$ ならばこの手続きを終える.

⁶⁴たとえば $(m, n) = (3, 7)$ で \tilde{A} が

$$\tilde{A} = \begin{bmatrix} 0 & 0 & c_1 & * & * & * & * \\ 0 & 0 & 0 & 0 & c_2 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & c_3 \end{bmatrix}, \quad c_1, c_2, c_3 \neq 0$$

のような形であれば $r = 3$ であり, すべての成分が 0 になる左端の数列はちょうど 2 列存在し, すべての成分が 0 になる下端の数行は存在しない.

2. $n > 1$ ならば A は次のような形をしている:

$$A = \left[\begin{array}{c|c} 0 & \\ \vdots & B \\ 0 & \end{array} \right], \quad B \in M_{m,n-1}(K).$$

3. B に対してこの手続きを適用する.

II. A の第 1 列の第 i 成分が 0 でないとき,

1. $m = 1$ ならば (このとき $i = 1$ である) この手続きを終える.

2. $m > 1$ ならば A の第 i 行と第 1 行を交換する.

3. さらに第 1 行の定数倍を第 2, ..., m 行に加えて, 第 1 列目の第 2, ..., m 成分をすべて 0 にする. その結果は次のような形になる:

$$A' = \left[\begin{array}{c|c} c & * \cdots * \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array} \right], \quad c \neq 0, \quad B \in M_{m-1,n-1}(K).$$

4. B に対してこの手続きを適用する.

以上の手続きは有限ステップで終了し, A が行のみに関する基本操作 (a),(b) によって \tilde{A} のような階段行列の形に変形できることがわかる.

上の手続きの I の 3 と II の 4 では, 手続き全体に悪影響を及ぼすことなく, B の行に関する基本変形が A もしくは A' の行に関する基本変形によって実現できることを仮定している. その仮定が正しいことを示せ (容易である).

\tilde{A} の第 1, ..., r 行のそれぞれに $c_1^{-1}, \dots, c_r^{-1}$ をかければ \tilde{A} は $\tilde{\tilde{A}}$ の形に変形される. \square

例 4.7 行列 A を次のように定める:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 7 & 10 \\ 3 & 6 & 10 & 16 \end{bmatrix}.$$

この A は行に関する基本変形によって次のように階段行列に変形される:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 7 & 10 \\ 3 & 6 & 10 & 16 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \\ 3 & 6 & 10 & 16 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

ここで 1 つ目の矢印は第 1 行の 2 倍を第 2 行から引き去る基本操作であり, 2 つ目の矢印は第 1 行の 3 倍を第 3 行から引き去る基本操作であり, 3 つ目の矢印は第 2 行を第 3 行から引き去る基本操作である. \square

[110] 行だけに関する基本変形によって次の行列を階段行列に変形せよ:

$$A = \begin{bmatrix} 0 & -3 & 1 & 2 \\ 1 & 3 & -2 & 1 \\ 2 & 3 & -3 & 4 \end{bmatrix}. \quad \square$$

略解. 問題 [109] のヒントの手続きを適用するとき最初に第 1 行と第 2 行を交換すると次のようになる:

$$A = \begin{bmatrix} 0 & -3 & 1 & 2 \\ 1 & 3 & -2 & 1 \\ 2 & 3 & -3 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & -2 & 1 \\ 0 & -3 & 1 & 2 \\ 2 & 3 & -3 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & -2 & 1 \\ 0 & -3 & 1 & 2 \\ 0 & -3 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & -2 & 1 \\ 0 & -3 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad \square$$

[111] (行列の基本変形による行列の簡単化) $A = [a_{ij}]$ は体 K の元を成分に持つ $m \times n$ 行列であるとする. このとき A は行列の基本変形によって次の形に変形可能である:

$$\tilde{A} = \left[\begin{array}{ccc|ccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ \hline & & & & 1 & \\ & & & & & 0 \end{array} \right].$$

ここで \tilde{A} の中に斜めに並んでいる 1 の個数は問題 [109] の r に等しい. \square

ヒント. 問題 [109] の結果より A は行だけに関する基本変形を用いて次の形に変形可能である:

$$\tilde{\tilde{A}} = \left[\begin{array}{cccc|cccc} & & 1 & * & \cdots & * & & & * \\ & & & & & & 1 & * & \cdots & * \\ & & & & & & & \ddots & & \\ & & & & & & & & 1 & * & \cdots & * \\ 0 & & & & & & & & & & & \end{array} \right]. \quad \square$$

階段のかどの 1 は左上から順に第 $(1, j_1), (2, j_2), \dots, (r, j_r)$ 成分にあるとする. このとき第 j_1 列の定数倍をそれより右側の列に加えることによって第 1 行の 0 でない成分が第 $(1, j_1)$ 成分の 1 だけであるようにできる. さらに第 j_2 列の定数倍をそれより右側の列に加えることによって第 1, 2 行の 0 でない成分が第 $(1, j_1), (2, j_2)$ 成分の 1 だけであるようにできる. 同様の作業を続けることによって, 行列全体の 0 でない成分が第 $(1, j_1), (2, j_2), \dots, (r, j_r)$ 成分の 1 だけであるようにできる. つまり $\tilde{\tilde{A}}$ は列に関する基本変形で次の形に変形できる:

$$A' = \left[\begin{array}{cccc|cccc} & & 1 & 0 & \cdots & 0 & & & 0 \\ & & & & & & 1 & 0 & \cdots & 0 \\ & & & & & & & \ddots & & \\ & & & & & & & & 1 & 0 & \cdots & 0 \\ 0 & & & & & & & & & & & \end{array} \right].$$

列の置換によって 0 でない列を左側に寄せることによって, A' は \tilde{A} の形に変形できる. \square

[112] (一般線形群の生成元) $GL_m(K)$ の任意の元は第 4.3 節で定義した行列 $U_{ij}(\alpha)$, P_{ij} , $D_i(\beta)$ たちの有限個の積で表わされる. \square

ヒント. $A \in GL_m(K)$ であるとする. 少し考えれば A の rank は m であることがわかる⁶⁵. 注意 4.2 と問題 [111] の結果より, 行列 $U_{ij}(\alpha)$, P_{ij} , $D_i(\beta)$ たちの有限個の積で表わ

⁶⁵ ヒント: A の第 j 列を a_j と書き, $\alpha = [\alpha_j] \in K^m$ とすると, $A\alpha = \alpha_1 a_1 + \cdots + \alpha_m a_m$ である. A は逆行列を持つので $A\alpha = 0$ ならば $\alpha = A^{-1}A\alpha = 0$ であるから, a_1, \dots, a_m は一次独立である. よって $\text{rank } A = m$. (この議論では A が可逆であることのみを直接用いており, 行列式による正方行列の可逆性の判定法などの他の道具を何も用いていない.) 同様の議論で $A \in M_{m,n}(K)$, $B \in M_{n,m}(K)$ が $BA = E_n$ (E_n は n 次の単位行列) を満たしていれば $\text{rank } A = n$ であることを示せる.

される行列 P, Q で $PAQ = E$ (E は m 次単位行列) を満たすものが存在する. このとき $A = P^{-1}Q^{-1}$ である. 問題 [100] の結果より, P^{-1}, Q^{-1} も行列 $U_{ij}(\alpha), P_{ij}, D_i(\beta)$ たちの有限個の積で表わされる. \square

[113] $A \in M_{m,n}(K), P \in GL_m(K), Q \in GL_n(K)$ に対して

$$\text{rank}(PAQ) = \text{rank } A. \quad \square$$

ヒント. 問題 [106], [112] を使えばただちに得られる. \square

参考 4.8 この演習問題集の議論の流れに沿って $\text{rank}(PAQ) = \text{rank } A$ を証明すると非常に長くなってしまうが, 抽象線形代数を十分に習得すればほとんど自明になってしまう⁶⁶. 別証の概略: P, Q は可逆なのでベクトル空間としての像のあいだの同型 $\text{Im}(PAQ) \cong \text{Im } A$ が成立している. よって $\text{rank}(PAQ) = \dim \text{Im}(PAQ) = \dim \text{Im } A = \text{rank } A$. \square

4.6 K^n の部分空間とその次元

定義 4.9 (K^n の部分空間) W が K^n の部分空間 (線形部分空間, **linear subspace**, ベクトル部分空間, **vector subspace**) であるとは, W が K^n の空でない部分集合であり, 任意の $\alpha \in K, w \in W$ に対して $\alpha w \in W$ であつ任意の $w_1, w_2 \in W$ に対して $w_1 + w_2 \in W$ を満たしていることである⁶⁷. \square

参考 4.10 より一般に K 上の任意の (抽象) ベクトル空間 V の部分空間も定義 4.9 と同様の条件で定義される. \square

[114] (基底) W が K^n の部分空間であるとき, $w_1, \dots, w_r \in W$ に対して以下の条件は互いに同値である:

- (a) w_1, \dots, w_r は一次独立であつ $W = Kw_1 + \dots + Kw_r$ である⁶⁸.
- (b) 任意の W の元は w_1, \dots, w_r の K 係数一次結合で一意に表わされる.

この互いに同値な条件のどちらかが成立しているとき w_1, \dots, w_r は W の基底 (**basis**) であると言う⁶⁹. \square

ヒント. 問題 [102] の一次独立性に関する説明を見よ. この問題の条件 (b) は問題 [102] の条件 (b) に対応している. \square

[115] 以下を示せ:

1. K^n 自身と $0 = \{0\}$ は K^n の部分空間である⁷⁰.

⁶⁶ 数学的一般論に関する主張は抽象化すればするほど自明さが増すことが多い.

⁶⁷ 任意の $\alpha_1, \alpha_2 \in K, w_1, w_2 \in W$ に対して $\alpha_1 w_1 + \alpha_2 w_2 \in W$ を満たしているという条件と同値である. また W が空集合でないという条件の代わりに $0 \in W$ という条件を採用してもよい (むしろ空集合でないという条件より自然かもしれない).

⁶⁸ 条件 $W = Kw_1 + \dots + Kw_r$ は任意の W の元が w_1, \dots, w_r の K 係数一次結合で表わされることと同値である. もちろん条件 $W = Kw_1 + \dots + Kw_r$ だけではその表示の一意性は保証されない. $W = Kw_1 + \dots + Kw_r$ が成立するとき, W は w_1, \dots, w_r で張られると言う.

⁶⁹ basis (基底) と base (基) の複数形はどちらも bases である.

⁷⁰ ベクトルのみで構成される集合をも同じ記号 0 で表わした. 異なるものを同じ記号で表わしているのに注意して欲しい. 習慣的に 0 ベクトルのみで構成されたベクトル空間は単に 0 と書かれることが多い.

2. K^n の部分空間 W と $v \in K^n$ に対して $v + W = \{v + w \mid w \in W\}$ が K^n の部分空間になるための必要十分条件は $v \in W$ である. ($K^n = \mathbb{R}^3$ のときこの小問の様子を図に描いてみよ.) \square

[116] $W \subset \mathbb{R}^3$ を次のように定める:

$$W = \{ {}^t[x, y, z] \in \mathbb{R}^3 \mid x + y + z = 0 \}.$$

このとき W は \mathbb{R}^3 の部分空間であり, ${}^t[1, -1, 0]$, ${}^t[0, 1, -1]$, ${}^t[-1, 0, 1]$ から2つを任意に選ぶと W の基底になっていることを示せ. さらにその様子を図に描け. \square

[117] $A \in M_{m,n}(K)$ に対して, その核 $\text{Ker } A$ と像 $\text{Im } A$ はそれぞれ K^n と K^m の部分空間である⁷¹. \square

ヒント. $\text{Ker } A$ と $\text{Im } A$ が部分空間の定義を満たしていることを機械的に確かめればよい. \square

参考 4.11 線形部分空間の多くがある線形写像の核もしくは像の形で定義される. 線形写像の核および像の構造を調べるのが線形代数の重要な目的の一つである. \square

[118] $v_1, \dots, v_r \in K^n$ に対して

$$W := Kv_1 + \dots + Kv_r = \{ \alpha_1 v_1 + \dots + \alpha_r v_r \mid \alpha_1, \dots, \alpha_r \in K \}$$

は K^n の部分空間である. さらに適当に v_{i_1}, \dots, v_{i_s} を選んで W の基底になるようにできる. \square

ヒント. $W_0 = 0 (= \{0\})$, $W_i = Kv_1 + \dots + Kv_{i_s}$ ($i = 1, \dots, r$) と置く. このとき $0 = W_0 \subset W_1 \subset \dots \subset W_r = W$ である⁷². $W_{i-1} \subsetneq W_i$ を満たす $i \in \{1, \dots, r\}$ の全体を小さい順に i_1, \dots, i_s と書くことにする. このとき $W = W_{i_s}$ である. さらに ν に関して帰納的に $W_{i_\nu} = Kv_{i_1} + \dots + Kv_{i_\nu}$ であることも示せる. そのことを使うと v_{i_1}, \dots, v_{i_s} が W の基底であることがわかる. \square

参考 4.12 実は「任意のベクトル空間は基底を持つ」という結果を証明できる. 有限個のベクトルで張られるベクトル空間の基底の存在の証明は問題 [118] と同様である. 有限個のベクトルで張られない無限次元のベクトル空間の基底の存在証明は Zorn の補題を用いて証明される.

ベクトル空間の理論は基底が存在するおかげで簡単になる. ベクトル空間特有の結果を証明するためには基底の存在を用いることが多い⁷³ \square

⁷¹ $A \in M_{m,n}(K)$ の核 (kernel) $\text{Ker } A$ と像 (image) $\text{Im } A$ は次のように定義されたのであった:

$$\text{Ker } A = \{ x \in K^n \mid Ax = 0 \}, \quad \text{Im } A = \{ Ax \mid x \in K^n \} \subset K^m.$$

⁷² このように部分ベクトル空間の単調な増大列 (もしくは減少列) を考えることは常套手段の一つである. $0 = W_0 \subset W_1 \subset \dots \subset W_r = W$ を W の **filtration** と呼ぶことがある.

⁷³ 体ではない一般の環上では (自由) 基底が存在しないような加群が存在するので理論が複雑になる.

[119] (部分空間の基底の存在) W は K^n の部分空間であるとし, $w_1, \dots, w_s \in W$ は一次独立であると仮定する. このとき w_1, \dots, w_s を拡張して W の基底 $w_1, \dots, w_s, w_{s+1}, \dots, w_r$ ($r \geq s$) を構成できる. \square

ヒント. もしも $W = Kw_1 + \dots + Kw_s$ ならば w_1, \dots, w_s は W の基底である. そうでないならばある $w_{s+1} \in W$ で $w_{s+1} \notin Kw_1 + \dots + Kw_s$ を満たすものが存在する. そのとき w_1, \dots, w_{s+1} は一次独立である. もしも $W = Kw_1 + \dots + Kw_{s+1}$ ならば w_1, \dots, w_{s+1} は W の基底である. そうでないならばある $w_{s+2} \in W$ で $w_{s+2} \notin Kw_1 + \dots + Kw_{s+1}$ を満たすものが存在する. 同様の手続きを続けてどこかで $W = Kw_1 + \dots + Kw_{s+k}$ となることを示せばよい. もしもそうならなければ一次独立な w_1, \dots, w_{n+1} が取れることになる. そのとき行列 $A = [w_1, \dots, w_{n+1}] \in M_{n,n+1}(K)$ の rank は $n+1$ になってしまう. しかし問題 [132] の結果より, それは不可能である. \square

注意 4.13 ヒントによつて w_1, \dots, w_r が K^n の部分空間 W の基底であれば $r \leq n$ であることもわかる. \square

参考 4.14 上の注意への進んだ補足. 上の注意は n 次元ベクトル空間の部分空間の次元 (もうすぐ定義される) が n 以下になるということを意味している. 直観的にも当然そうなるべきだろう. ベクトル空間の次元に関しては「部分は全体よりも小さくなる」という直観が通用する.

しかし, 体上のベクトル空間を離れて, 一般的可換環上の加群について考えるとそのような直観はもはや通用しなくなる. たとえば 2 変数多項式環 $R = K[x, y]$ 上の加群 $M = R$ について考えると, $N = \{f(x, y) \in M \mid f(0, 0) = 0\}$ は $M = R$ の R 部分加群である. $M = R$ は一つ元 (たとえば 1) から R 上生成されるが, N は一つ元 (たとえば x) だけから R 上生成されないことを示せる (N はたとえば x, y から R 上生成される). このように 2 変数多項式環のような可換環を考えると部分加群の生成元の個数がもとの加群よりも増えることがあり得る.

上の説明ではいきなり 2 変数の多項式環を例に説明したが, それでは 1 変数の多項式環 $K[x]$ ではどうだろうか. 実は幸運なことに 1 変数の多項式環上の加群に関しては体上のベクトル空間の場合と同様に部分加群の生成元の個数はもとの加群以下になることを示せる. 可換環として \mathbb{Z} を採用しても同様の事実が成立している. この事実を用いると, $K[x]$ や \mathbb{Z} 上の加群に関して, ベクトル空間の理論とほぼ同じ議論を適用できることがわかる. それを実際に実行して得られるのが単因子論 (elementary divisor theory) もしくは単項イデアル整域上の加群の理論である. 単因子論からは Jordan 標準形の理論や有限生成 Abel 群の基本定理が導かれる. 単因子論入門には堀田 [H2] がおすすめである. \square

[120] $A \in M_{m,n}(K)$, $B \in M_{n,m}(K)$ のとき $AB = E_m$ (E_m は m 次の単位行列) ならば $m \leq n$ である.

ヒント 1. B の第 i 列を b_i と書き, $\beta = [\beta_i] \in K^m$ とすると $B\beta = \beta_1 b_1 + \dots + \beta_m b_m$ である. $AB = E_m$ より $B\beta = 0$ ならば $0 = AB\beta = \beta$ である. よつて b_1, \dots, b_m は一次独立であり, $\text{rank } B = m$ である. 問題 [132] の結果より, $m \leq n$ でなければいけない. \square

ヒント 2. 注意 2.23 より $m > n$ ならば $|AB| = 0$ である. 一方 $AB = E_m$ のとき $|AB| = |E_m| = 1$ である. よつて $AB = E_m$ と $m > n$ は互いに矛盾する. したがつて $AB = E_m$ ならば $m \leq n$ でなければいけない. \square

参考 4.15 このように線形代数学には行列式を使っても使わなくても証明できる結果がたくさんある. 行列式を使う方法も使わない方法もどちらも重要である. \square

[121] (部分空間の次元) W が K^n の部分空間であり, v_1, \dots, v_s と w_1, \dots, w_r がともに W の基底であるならば $s = r$ である. すなわち W の基底に含まれるベクトルの本数は基底の取り方によらずに定まる. W の基底に含まれるベクトルの本数を W の次元 (dimension) と呼び, $\dim W$ と表わす. (基礎体 K を明示したい場合には $\dim_K W$ と表わす.) \square

ヒント. v_l は $v_l = \sum_{i=1}^r a_{il} w_i$ ($a_{il} \in K$) と一意に表わされ, w_j は $w_j = \sum_{k=1}^s b_{kj} v_k$ ($b_{kj} \in K$) と一意に表わされる. $A = [a_{il}] \in M_{r,s}(K)$, $B = [b_{kj}] \in M_{s,r}(K)$ と置く. このとき, $w_j = \sum_{k=1}^s b_{kj} \sum_{i=1}^r a_{ik} w_i = \sum_{i=1}^r (\sum_{k=1}^s a_{ik} b_{kj}) w_i$ なので $\sum_{k=1}^s a_{ik} b_{kj} = \delta_{ij}$ すなわち $AB = E_r$ である. よって問題 [120] の結果より $r \leq s$ である. 同様にして $s \leq r$ も示せる. \square

参考 4.16 「ベクトル空間の基底に含まれるベクトルの本数が基底の取り方によらない」という結果は, K^n の部分空間に限らず, 常に成立している. (有限次元の場合には証明の仕方も上のヒントと同様である. 無限次元の場合には基底の濃度の一意性を証明できる.) だから, ベクトル空間の次元をその基底に含まれるベクトルの本数によって定義できる.

ベクトル空間の次元の概念は極めて基本的である.

たとえば $A \in M_{m,n}(K)$ に対して斉次な一次方程式 $Ax = 0$ の解空間 $\text{Ker } A = \{x \in K^n \mid Ax = 0\}$ は K^n の部分空間になるのであった. 解がどれだけたくさんあるかは $\text{Ker } A$ の次元で測ることができる. 同様に非斉次な方程式 $Ax = b$ の解がどれだけたくさんの $b \in K^m$ に対して存在するかは $\text{Im } A = \{Ax \mid x \in K^n\} \subset K^m$ の次元で測ることができる. \square

[122] W, W' は K^n の部分空間であり, $W' \subset W$ を満たしているとする. このとき $\dim W' \leq \dim W$ であり, 等号が成立するための必要十分条件は $W' = W$ が成立することである. \square

ヒント. W' の基底 $w_1, \dots, w_{\dim W'}$ を取る. 問題 [119] の結果より, $w_1, \dots, w_{\dim W'}$ を拡張して W の基底を構成することができる. よって $\dim W' \leq \dim W$ である. そのとき $\dim W' = \dim W$ は $w_1, \dots, w_{\dim W'}$ を何も拡張せずに W の基底が構成できることを意味するので, そのとき $W = W'$ となる. \square

[123] (2つの部分空間の和と直和) K^n の部分空間 V, W に対して, それらの和 (sum) を

$$V + W = \{v + w \mid v \in V, w \in W\}$$

と定める. $V + W$ と $V \cap W$ は K^n の部分空間であり, 次の等式が成立している:

$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W).$$

特に $\dim(V + W) \leq \dim V + \dim W$ であり, 等号が成立するための必要十分条件は $V \cap W = 0$ が成立することである. $V \cap W = 0$ が成立するとき $V + W$ を V と W の直和 (direct sum) と呼び, $V \oplus W$ と表わす. \square

ヒント. $V \cap W$ の基底を V と W のそれぞれに拡張し, それらを合わせたものを考えると $V + W$ の基底になる. \square

[124] (線形部分空間のモジュラー法則) 以下を示せ:

1. 集合 X, Y, Z について,

$$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z), \quad X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z).$$

この結果は集合の和 (union) と共通部分 (intersection) に関する分配法則 (**distributive law**) と呼ばれる.

2. K^n の部分空間 U, V, W について,

$$U + (V \cap W) \subset (U + V) \cap (U + W), \quad U \cap (V + W) \supset (U \cap V) + (U \cap W).$$

しかし, 等号は一般に成立しない. すなわち線形部分空間について分配法則は一般に成立しない.

3. K^n の部分空間 U, V, W について,

$$U \subset W \implies U + (V \cap W) = (U + V) \cap W.$$

この結果は線形部分空間の和と共通部分に関するモジュラー法則 (**modular law**) と呼ばれる. \square

ヒント. 1. 集合の包含関係の図を描けば容易に確かめられる. しかし念のために厳密な証明を書き下してみよ.

2. 包含関係の証明は基本に戻れば易しい⁷⁴. $K = \mathbb{R}$ であるとし, 2次元の実ベクトル空間 \mathbb{R}^2 の標準的な基底を e_1, e_2 と書き, $U = \mathbb{R}e_1$, $V = \mathbb{R}e_2$, $W = \mathbb{R}(e_1 + e_2)$ と置くとどうなるか? 図を描いてみよ.

3. $U \subset W$ と仮定する. $u \in U$, $v \in V \cap W$ ならば $u + v \in U + V$ かつ $u + v \in W$ であることが容易にわかる. 逆に $u \in U$, $v \in V$ が $w := u + v \in W$ を満たしているならば $v = w - u \in W + U = W$ なので $v \in V \cap W$ である. \square

[125] (有限集合の和と直和) 有限集合 X の元の個数を $|X|$ と書くことにする⁷⁵. 有限集合 X_1, \dots, X_N に対して,

$$|X_1 \cup \dots \cup X_N| = \sum_{p=1}^N (-1)^{p-1} \sum_{1 \leq i_1 < \dots < i_p \leq N} |X_{i_1} \cap \dots \cap X_{i_p}|.$$

特に $|X_1 \cup \dots \cup X_N| \leq |X_1| + \dots + |X_N|$ であり, 等号が成立するための必要十分条件は $X_i \cap X_j = \emptyset$ ($i \neq j$) が成立することである. $X_i \cap X_j = \emptyset$ ($i \neq j$) が成立するとき, $X_1 \cup \dots \cup X_N$ は X_1, \dots, X_N の直和 (**direct sum**) もしくは交わりのない和集合 (**disjoint union**) と呼ばれ, $X_1 \sqcup \dots \sqcup X_N$ と表わされる. \square

⁷⁴ $A \subset B$ を証明するためには, 任意に $a \in A$ を取り, $a \in B$ であることを示せばよい.

⁷⁵ 有限集合 X の元の個数を $\sharp X$ と書くこともある.

ヒント. N に関する数学的帰納法で証明できる. たとえば

$$|X_1 \cup X_2 \cup X_3| = |X_1| + |X_2| + |X_3| - |X_1 \cap X_2| - |X_1 \cap X_3| - |X_2 \cap X_3| + |X_1 \cap X_2 \cap X_3|$$

はどうして成立するか? \square

[126] (線形部分空間の和と直和) K^n の部分空間 V_1, \dots, V_N に対して, それらの和 (sum) を

$$V_1 + \dots + V_N = \{v_1 + \dots + v_N \mid v_i \in V_i \ (i = 1, \dots, N)\}$$

と定める. このとき, $V_1 + \dots + V_N$ と $V_{i_1} \cap \dots \cap V_{i_p}$ ($1 \leq i_1 < \dots < i_p \leq N$) は K^n の部分空間であり, 次が成立している:

$$\dim(V_1 + \dots + V_N) = \sum_{p=1}^N (-1)^{p-1} \sum_{1 \leq i_1 < \dots < i_p \leq N} \dim(V_{i_1} \cap \dots \cap V_{i_p}).$$

特に $\dim(V_1 + \dots + V_N) \leq \dim V_1 + \dots + \dim V_N$ であり, 等号が成立するための必要十分条件は $V_i \cap V_j = 0$ ($i \neq j$) が成立することである. $V_i \cap V_j = 0$ ($i \neq j$) が成立するとき, $V_1 + \dots + V_N$ は V_1, \dots, V_N の直和 (direct sum) と呼ばれ, $V_1 \oplus \dots \oplus V_N$ と表わされる. \square

ヒント. この問題は問題 [123] の一般化である. 問題 [125] とこの問題は類似している⁷⁶. 実際その証明の仕方も似ている. \square

4.7 rank と image と kernel の関係

[127] (rank と image の関係) $A \in M_{m,n}(K)$ に対して $\text{rank } A = \dim \text{Im } A$. \square

ヒント. A の第 j 列を a_j と表わすと, $\alpha = [\alpha_j] \in K^n$ に対して $A\alpha = \alpha_1 a_1 + \dots + \alpha_n a_n$ であるから,

$$\text{Im } A = Ka_1 + \dots + Ka_n$$

が成立する. $r = \text{rank } A$ (すなわち r は一次独立になる a_i たちの組の最大の本数) であるとし, a_{i_1}, \dots, a_{i_r} は一次独立であるとする. このとき a_i ($i \neq i_1, \dots, i_r$) は a_{i_1}, \dots, a_{i_r} の一次結合で表わされる. これより, a_{i_1}, \dots, a_{i_r} は $\text{Im } A$ の基底であることがわかる. したがって $\dim \text{Im } A = \text{rank } A$ である. \square

[128] (image と kernel の関係) $A \in M_{m,n}(K)$ に対して $\dim \text{Ker } A + \dim \text{Im } A = n$. \square

ヒント. K^n の部分空間 $\text{Ker } A$ の基底 v_1, \dots, v_r に v_{r+1}, \dots, v_n を追加して K^n 全体の基底を構成することができる. Av_{r+1}, \dots, Av_n が $\text{Im } A$ の基底になることを示そう. 任意の $x \in K^n$ は $x = \alpha_1 v_1 + \dots + \alpha_n v_n$ ($\alpha_i \in K$) と一意に表わされる. このとき, $Ax = \alpha_{r+1} Av_{r+1} + \dots + \alpha_n Av_n$ であるから $\text{Im } A = KAv_{r+1} + \dots + KAv_n$ である. もしも $\alpha_{r+1}, \dots, \alpha_n \in K$ でかつ $\alpha_{r+1} Av_{r+1} + \dots + \alpha_n Av_n = 0$ ならば $\alpha_{r+1} v_{r+1} + \dots + \alpha_n v_n \in \text{Ker } A$ である. よって $\alpha_{r+1} v_{r+1} + \dots + \alpha_n v_n$ は v_1, \dots, v_r の一次結合で表わされる. しかし v_1, \dots, v_n の一次独立性より $\alpha_{r+1} = \dots = \alpha_n = 0$ でなければいけない. これで Av_{r+1}, \dots, Av_n が一次独立であることがわかった. 以上によって Av_{r+1}, \dots, Av_n が $\text{Im } A$ の基底になることがわかった. したがって $\dim \text{Ker } A + \dim \text{Im } A = r + (n - r) = n$. \square

⁷⁶問題 [126] は問題 [125] の「量子化」であるとみなせる.

注意 4.17 $\text{Ker } A$ は斉次な一次方程式 $Ax = 0$ の解空間であり, $\text{Im } A$ は一次方程式 $Ax = b$ が解を持つ $b \in K^m$ 全体のなす K^m の部分空間なのであった. よって等式 $\dim \text{Ker } A + \dim \text{Im } A = n$ は斉次な一次方程式 $Ax = 0$ の解空間の次元と非斉次な一次方程式 $Ax = b$ が解を持つような b 全体のなす空間の次元のあいだの関係式であると解釈できる. このような形で斉次な一次方程式の解空間の次元と非斉次な一次方程式がいつ解けるかという問題は密接に関係することになる. $m = n$ の場合については問題 [129] を見よ. \square

参考 4.18 上の問題の結果を図 4.1 のように表わすことがある. 図 4.1 は, A の行き先全体が $\text{Im } A$ であり, A で 0 に潰される部分が $\text{Ker } A$ であること, そして $\text{Ker } A$ と $\text{Im } A$ のサイズを合わせると K^n 全体のサイズに等しくなることなどがうまく表現している. (縦方向の線分の長さがベクトル空間の次元の大きさを表わしている.)

A で定まる線形写像 $A: K^n \rightarrow K^m$ が単射, 全射, 全単射である状況を図示してみよ. \square

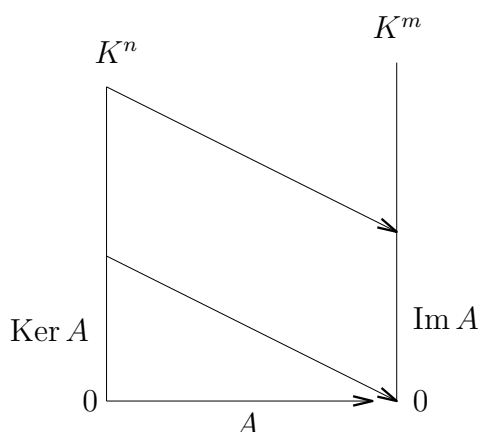


図 4.1: A の kernel と image の図示

[129] 正方行列 $A \in M_n(K)$ に関して以下の条件は互いに同値である:

- (a) $\text{Ker } A = 0$.
- (b) 斉次な一次方程式 $Ax = 0$ の解は $x = 0$ しか存在しない.
- (c) $\text{Im } A = K^n$.
- (d) 一次方程式 $Ax = b$ の解が任意の $b \in K^n$ に対して存在する. \square

ヒント. (a) と (b) の同値性と (c) と (d) の同値性はすぐにわかるので (a) と (c) が同値であることを示せばよい. 問題 [128] の結果より $\dim \text{Ker } A + \dim \text{Im } A = n$ である. よって $\dim \text{Ker } A = 0$ と $\dim \text{Im } A = n$ は同値である. K^n の部分空間 $\text{Im } A$ が K^n に一致するための必要十分条件は問題 [122] の結果より $\dim \text{Im } A = n$ である. \square

[130] $A \in M_{l,m}(K)$, $B \in M_{m,n}(K)$ ならば $\text{rank}(AB) \leq \min\{\text{rank } A, \text{rank } B\}$ である⁷⁷. \square

⁷⁷実数 x_1, \dots, x_k の最小値と最大値をそれぞれ $\min\{x_1, \dots, x_k\}$, $\max\{x_1, \dots, x_k\}$ と表わす. \min , \max はそれぞれ minimum (最小値), maximum (最大値) の略である.

ヒント. 問題 [127] の結果を使う⁷⁸. $\text{Im}(AB) \subset \text{Im } A$ なので $\text{rank}(A) \leq \text{rank } A$ である. 一般に K^m の部分空間 W に対して $\dim(AW) \leq \dim W$ である (それはなぜか?). これを $\text{Im}(AB) = A \text{Im } B$ に適用せよ. \square

[131] $A, B \in M_{m,n}(K)$ のとき

$$\text{rank}(A + B) \leq \text{rank } A + \text{rank } B. \quad \square$$

ヒント. $\text{Im}(A + B) \subset \text{Im } A + \text{Im } B$. \square

4.8 行列の rank の計算の仕方

[132] 問題 [111] の行列 \check{A} の rank は r である (r は問題 [109] の r). よって基本変形前の行列 A の rank も r である. 特に $A \in M_{m,n}(K)$ のとき $\text{rank } A$ は m 以下かつ n 以下である. \square

注意 4.19 (行列の rank の行列の基本変形による計算法) この問題の結果より, 行列 A の rank を求めるためには問題 [109] のヒントの手続きによって A を階段行列に変形し, その階段の段数を数えれば良いことがわかる. \square

ヒント. \check{A} 中の 0 でない列は第 $1, \dots, r$ 列しかなく, それらは一次独立である. 後半は問題 [106] の結果と前半から導かれる. \square

[133] 次の行列の rank を求めよ:

$$A = \begin{bmatrix} 1 & 5 & 3 & -3 \\ 3 & 15 & 10 & -3 \\ -2 & -10 & -6 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 1 \\ 3 & -5 & -5 \\ -2 & 3 & 2 \\ 8 & -11 & -4 \end{bmatrix}.$$

略解. $\text{rank } A = \text{rank } B = 2$. \square

[134] 次の行列の rank が 2 になるような a, b の値を求めよ:

$$A = \begin{bmatrix} 2 & 1 & 0 & 3 \\ -1 & 3 & 5 & 7 \\ 7 & 0 & a & b \end{bmatrix}.$$

解答. A は行に関する基本変形で次のように変形される:

$$A \rightarrow \begin{bmatrix} -1 & 3 & 5 & 7 \\ 2 & 1 & 0 & 3 \\ 7 & 0 & a & b \end{bmatrix} \rightarrow \begin{bmatrix} -1 & 3 & 5 & 7 \\ 0 & 7 & 10 & 17 \\ 0 & 21 & a+35 & b+49 \end{bmatrix} \rightarrow \begin{bmatrix} -1 & 3 & 5 & 7 \\ 0 & 7 & 10 & 17 \\ 0 & 0 & a+5 & b-2 \end{bmatrix}.$$

これより $\text{rank } A = 2$ となるための必要十分条件は $a = -5$ かつ $b = 2$ であることがわかる. \square

⁷⁸ A の rank の定義を $\text{rank } A = \dim \text{Im } A$ だと思って証明を考えよ.

$A = [a_{ij}] \in M_{m,n}(K)$, $1 \leq i_1 < \cdots < i_s \leq m$, $1 \leq j_1 < \cdots < j_s \leq n$ のとき, 行列式

$$A_{j_1 \dots j_s}^{i_1 \dots i_s} = \begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_s} \\ \vdots & & \vdots \\ a_{i_s j_1} & \cdots & a_{i_s j_s} \end{vmatrix}$$

を A の s 次の小行列式と呼ぶ.

[135] (小行列式を用いた rank の特徴付け) $A = [a_{ij}] \in M_{m,n}(K)$ に対して, 非負の整数 $r(A)$ を以下の条件によって定める:

- (a) $s > r(A)$ のとき A の s 次の小行列式はすべて 0 である.
- (b) A の $r(A)$ 次の小行列式で 0 でないものが存在する.

このとき $r(A) = \text{rank } A$ が成立する. \square

ヒント. 問題 [111] の \check{A} に関して $r(\check{A}) = \text{rank } \check{A}$ が成立していることは直接に確かめることができる. 問題 [106] の結果より, $\text{rank } A$ は行列の基本変形で不変なので $\text{rank } \check{A} = \text{rank } A$ である. したがって $r(A)$ が行列の基本変形で不変であることを示すことができれば $r(\check{A}) = r(A)$ であることがわかり, 証明が完結する.

A の s 次の小行列式がすべて 0 であるとき, A に行列の基本操作を施して得られた行列 A' の s 次の小行列式もすべて 0 になることを直接に確かめることができる. これより $r(A) \leq r(A')$ であることがわかる. 逆に A は A' に行列の基本操作を施すことによって得られるので逆向きの不等式 $r(A) \geq r(A')$ も成立している. したがって $r(A') = r(A)$ である. \square

[136] (転置による rank の不変性) $\text{rank } {}^t A = \text{rank } A$. \square

ヒント. 問題 [135] と $\left| ({}^t A)_{i_1 \dots i_s}^{j_1 \dots j_s} \right| = \left| A_{j_1 \dots j_s}^{i_1 \dots i_s} \right|$ を使う. \square

参考 4.20 この演習問題集の議論の流れに沿って $\text{rank } {}^t A = \text{rank } A$ を証明すると非常に長くなってしまうが, 抽象線形代数を十分に習得すれば $\text{rank } {}^t A = \text{rank } A$ はほとんど自明になる⁷⁹. ベクトル空間の双対空間 (dual space) を取る反変関手 (contravariant functor) が完全 (exact) であることの簡単な帰結になってしまう. \square

要約 4.21 (行列の rank に関するまとめ) 問題 [127], [132], [135], [136] を合わせると行列 A の rank には以下のように複数の特徴付け (および計算の仕方) が存在することがわかる:

- $\text{rank } A = \dim \text{Im } A$ である.
- $\text{rank } A$ は A 中の一次独立な列ベクトルの個数の最大値に等しい.
- 行に関する基本変形で A を階段行列に変形できる. $\text{rank } A$ はその階段行列の階段の段数に等しい.

⁷⁹一般に数学的一般論に関する主張は抽象化すればするほど自明さが増すことが多い.

- $\text{rank } A = \text{rank } {}^t A$ である. 特に $\text{rank } A$ は A の中の一次独立な行ベクトルの個数の最大値に等しい.
- $r = \text{rank } A$ であるための必要十分条件は, $s > r$ のとき s 次の小行列式がすべて 0 になり, r 次の小行列式で 0 でないものが存在することである. \square

[137] 行列 A を次のように定める:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 5 & 6 \\ 5 & 6 & 7 & 8 \end{bmatrix}.$$

$\text{rank } A$ を次の 3 通りの方法で計算せよ:

1. A を行に関する基本変形で階段行列に変形する方法,
2. 転置 ${}^t A$ を行に関する基本変形で階段行列に変形する方法,
3. A の小行列式を計算する方法.

行列のサイズを大きくしたとき計算の手間が爆発的に増えそうなのはどの方法であるか? \square

略解. どの方法で計算しても $\text{rank } A = 2$ という結果が得られる. 1 の計算:

$$A \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -2 & -4 & -6 \\ 0 & -4 & -8 & -12 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -2 & -4 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

2 の計算:

$${}^t A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 5 & 7 \\ 4 & 6 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 5 \\ 0 & -2 & -4 \\ 0 & -4 & -8 \\ 0 & -6 & -12 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 5 \\ 0 & -2 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

3 の計算の仕方: まず 4 通りの 3 次の小行列式をすべて計算してどれも 0 になることを確かめる. 2 次の小行列式で 0 にならないものはすぐに見付かる. 小行列式を用いた方法では行列のサイズが大きくなったとき手間が爆発する. \square

[138] 次の行列の rank を求めよ:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ -1 & a-2 & 4 & 6 & 1 & 2 \\ -2 & -4 & a-3 & -8 & -10 & -12 \\ -1 & -2 & -3 & a-1 & -5 & -6 \end{bmatrix}. \quad \square$$

略解. 行に関する基本変形で A は次のように変形される:

$$A \rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & a & 7 & 10 & 6 & 8 \\ 0 & 0 & a+3 & 0 & 0 & 0 \\ 0 & 0 & 0 & a+3 & 0 & 0 \end{bmatrix}.$$

よって $a \neq 0, -3$ のとき $\text{rank } A = 4$ であり, $a = 0$ のとき $\text{rank } A = 3$ であり, $a = -3$ のとき $\text{rank } A = 2$ である. \square

4.9 斉次な一次方程式の解法

$A \in M_{m,n}(K)$ とし, 斉次な一次方程式 $Ax = 0$ の解法について説明しよう. (第 4.1 節で説明したように非斉次な一次方程式を解くときにも斉次な一次方程式を解く必要が生じる.)

問題 [109] の結果より, 行列 A は行だけに関する基本変形によって次の形に変形される⁸⁰:

$$\tilde{A} = \begin{bmatrix} & \boxed{1} & & & * \\ & & \boxed{1} & & \\ & & & \ddots & \\ & & & & \boxed{1} \\ 0 & & & & \end{bmatrix}.$$

問題 [101] の結果より, ある $P \in GL_m(K)$ で $\tilde{A} = PA$ となるものが存在する. このとき,

$$Ax = 0 \iff PAx = 0 \iff \tilde{A}x = 0$$

なので斉次な一次方程式 $Ax = 0$ の解空間と $\tilde{A}x = 0$ の解空間は等しい. したがって, $Ax = 0$ と解く代わりに, より簡単な $\tilde{A}x = 0$ を解けば良い.

正規化された階段行列 \tilde{A} の階段のかどの 1 は左上から順に第 $(1, j_1), \dots, (r, j_r)$ 成分にあるものとし, \tilde{A} の (i, j) 成分を \tilde{a}_{ij} と書くことにすると, 方程式 $\tilde{A}x = 0$ は次と同値である:

$$\begin{aligned} x_{j_1} + \tilde{a}_{1,j_1+1}x_{j_1+1} + \cdots &= 0, \\ x_{j_2} + \tilde{a}_{2,j_2+1}x_{j_2+1} + \cdots &= 0, \\ &\cdots \\ x_{j_r} + \tilde{a}_{r,j_r+1}x_{j_r+1} + \cdots &= 0. \end{aligned}$$

この連立方程式を下から順に $x_{j_r}, \dots, x_{j_2}, x_{j_1}$ に関して解いて次の形に同値変形することは易しい:

$$\begin{aligned} x_{j_r} &= \sum_{j>j_r} c_{rj}x_j, \\ x_{j_{r-1}} &= \sum_{j>j_{r-1}, j \neq j_r} c_{r-1,j}x_j, \\ &\cdots \\ x_{j_2} &= \sum_{j>j_2, j \neq j_3, \dots, j_r} c_{2j}x_j, \\ x_{j_1} &= \sum_{j>j_1, j \neq j_2, \dots, j_r} c_{1j}x_j \quad (c_{\nu j} \in K). \end{aligned}$$

⁸⁰ 以下の一般的な場合に関する説明を読む前に例 4.22 を読んでおいた方が感じがつかみ易いかもしれない. 具体例を試してみて感じをつかむことから始めた方がややこしい計算の一般論の理解が容易になることが多い.

これらの式は x_{j_ν} たちを x_j ($j \neq j_1, \dots, j_r$) の一次結合で表わす式になっている (右辺は x_{j_1}, \dots, x_{j_r} を含まない). よって x_j ($j \neq j_1, \dots, j_r$) は任意の値を取ることができ, その値によって x_{j_1}, \dots, x_{j_r} の値が決定される.

以上の記号のもとで斉次な一次方程式 $Ax = 0$ の解空間 $\text{Ker } A (= \text{Ker } \tilde{A})$ は次のように表わされる:

$$\text{Ker } A = \left\{ x = [x_j]_{j=1}^n \left| \begin{array}{ll} x_j \in K & (j \neq j_1, \dots, j_r), \\ x_{j_\nu} = \sum_{j > j_\nu, j \neq j_{\nu+1}, \dots, j_r} c_{\nu j} x_j & (\nu = 1, \dots, r) \end{array} \right. \right\}.$$

$Ax = 0$ の解空間 $\text{Ker } A$ は K^n の部分空間をなす.

x_j ($j \neq j_1, \dots, j_r$) 中の x_k に 1 を代入して他を 0 に置くと, x_{j_ν} を x_j ($j \neq j_1, \dots, j_r$) の一次結合で表わす式によって, すべての x_j の値が決定される. その値を $w_j^{(k)}$ と書くことにする. すなわち $k \neq j_1, \dots, j_r$ に対して,

$$\begin{aligned} w_j^{(k)} &= \delta_{jk} & (j \neq j_1, \dots, j_r), \\ w_{j_\nu}^{(k)} &= c_{\nu k} & (j_\nu < k), \\ w_{j_\nu}^{(k)} &= 0 & (j_\nu > k) \end{aligned}$$

と置く. このとき $Ax = 0$ の解空間 $\text{Ker } A$ の基底として次が取れる:

$$w^{(k)} = [w_j^{(k)}]_{j=1}^n \in \text{Ker } A, \quad (k \neq j_1, \dots, j_r).$$

例 4.22 A が行に関する基本変形によって次の \tilde{A} に変形されたとする:

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

このとき $\tilde{A}x = 0$ は次と同値である:

$$\begin{aligned} x_1 - x_2 + x_3 - 2x_4 + x_5 + 3x_6 &= 0, \\ x_3 + x_4 - x_5 - 2x_6 &= 0, \\ x_5 - x_6 &= 0. \end{aligned}$$

これは次と同値である:

$$\begin{aligned} x_5 &= x_6, \\ x_3 &= -x_4 + x_5 + 2x_6 = -x_4 + 3x_6, \\ x_1 &= x_2 - x_3 + 2x_4 - x_5 - 3x_6 = x_2 + 3x_4 - 7x_6. \end{aligned}$$

したがって

$$\text{Ker } A = \text{Ker } \tilde{A} = \left\{ \begin{bmatrix} x_2 + 3x_4 - 7x_6 \\ x_2 \\ -x_4 + 3x_6 \\ x_4 \\ x_6 \\ x_6 \end{bmatrix} \left| x_2, x_4, x_6 \in K \right. \right\}.$$

$\text{Ker } A$ の基底を x_2, x_4, x_6 の一つだけに 1 を代入して他を 0 にすることによって得られる次のベクトルの組に取れる:

$$w^{(2)} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad w^{(4)} = \begin{bmatrix} 3 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad w^{(6)} = \begin{bmatrix} -7 \\ 0 \\ 3 \\ 0 \\ 1 \\ 1 \end{bmatrix}. \quad \square$$

[139] 例 4.22 の計算が正しいかどうかを確かめよ. もしも正しいならばそのことを確認し, 誤りが含まれているならばそれを修正せよ. \square

[140] 行列 A が次のように定められているとき, 斉次な一次方程式 $Ax = 0$ を解け:

$$A = \begin{bmatrix} 2 & 1 & 9 & -4 \\ -3 & 2 & -13 & -5 \\ -1 & -3 & -8 & 9 \end{bmatrix}. \quad \square$$

略解. 問題 [144] の略解を見よ. \square

[141] 行列 A が次のように定められているとき, 斉次な一次方程式 $Ax = 0$ を解け:

$$A = \begin{bmatrix} 2 & 2 & 3 & -4 \\ 3 & -1 & 2 & -5 \\ 1 & 5 & 4 & -3 \end{bmatrix}. \quad \square$$

略解. 問題 [145] の略解を見よ. \square

[142] 行列 A が次のように定められているとき, 斉次な一次方程式 $Ax = 0$ を解け:

$$A = \begin{bmatrix} 4 & -8 & 3 & 3 \\ -2 & 4 & 1 & -2 \\ 1 & -2 & 1 & 1 \\ 2 & -4 & 0 & 1 \end{bmatrix}. \quad \square$$

略解. 問題 [146] の略解を見よ. \square

4.10 非斉次な一次方程式の解法

$A \in M_{m,n}(K)$, $b \in K^m$ とし, 非斉次な一次方程式 $Ax = b$ の解法について説明しよう.

問題 [109] の結果より, 行列 A は行だけに関する基本変形によって次の形に変形される⁸¹:

$$\tilde{A} = \begin{bmatrix} & \boxed{1} & & & * \\ & & \boxed{1} & & \\ & & & \ddots & \\ & & & & \boxed{1} \\ 0 & & & & \end{bmatrix}.$$

⁸¹ 以下の一般的な場合に関する説明を読む前に例 4.23 を読んでおいた方が感じがつかみ易いかもしれない. 具体例を試してみて感じをつかむことから始めた方がややこしい計算の一般論の理解が容易になることが多い.

問題 [101] の結果より, ある $P \in GL_m(K)$ で $\tilde{A} = PA$ となるものが存在する. このとき, $b' = Pb$ と置くと,

$$Ax = b \iff PAx = Pb \iff \tilde{A}x = b'$$

なので一次方程式 $Ax = b$ の解空間と $\tilde{A}x = b'$ の解空間は等しい. したがって, $Ax = b$ と解く代わりに, より簡単な $\tilde{A}x = b'$ を解けば良い.

しかし, この方法をそのまま実行するためには, P を求めなければいけない. P を求めるためには A を行に関して基本変形するとき, その途中経過をすべて記録しておかなければいけない. この点を改良するためには A の代わりに $m \times (n+1)$ 行列 $A' = [A, b]$ に対して行に関する基本変形を適用すればよい. A を階段行列に変形するのと同じ行に関する基本変形を A' に適用することは P を A' に左からかけることに一致する. そして $PA' = P[A, b] = [PA, Pb] = [\tilde{A}, b']$ であるから, 我々が必要とする \tilde{A} と b' は A' に対して行に関する基本変形を施せば得られることになる.

上の方針の修正版. 問題 [109] の結果より, 行列 $A' = [A, b]$ は行だけに関する基本変形によって次の形に変形される⁸²:

$$\tilde{A}' = [\tilde{A}, b'] \begin{bmatrix} \begin{array}{ccc|ccc} & 1 & & & & & * \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ 0 & & & & & & \end{array} \end{bmatrix}.$$

問題 [101] の結果より, ある $P \in GL_m(K)$ で $\tilde{A}' = PA'$ となるものが存在する. このとき, $\tilde{A} = PA$, $b' = Pb$ であるから,

$$Ax = b \iff PAx = Pb \iff \tilde{A}x = b'.$$

すなわち一次方程式 $Ax = b$ の解空間と $\tilde{A}x = b'$ の解空間は等しい. したがって, $Ax = b$ と解く代わりに, より簡単な $\tilde{A}x = b'$ を解けば良い.

例 4.23 $A' = [A, b]$ が行に関する基本変形によって次の \tilde{A}' に変形されたとする:

$$\tilde{A}' = [\tilde{A}, b'] \left[\begin{array}{cccccc|c} 1 & -1 & 1 & -2 & 1 & 3 & b'_1 \\ 0 & 0 & 1 & 1 & -1 & -2 & b'_2 \\ 0 & 0 & 0 & 0 & 1 & -1 & b'_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & b'_4 \end{array} \right].$$

このとき $\tilde{A}x = b'$ は次と同値である:

$$\begin{aligned} x_1 - x_2 + x_3 - 2x_4 + x_5 + 3x_6 &= b'_1, \\ x_3 + x_4 - x_5 - 2x_6 &= b'_2, \\ x_5 - x_6 &= b'_3, \end{aligned}$$

⁸² 以下の一般的な場合に関する説明を読む前に例 4.23 を読んでおいた方が感じがつかみ易いかもしれない. 具体例を試してみて感じをつかむことから始めた方がややこしい計算の一般論の理解が容易になることが多い.

$$0 = b'_4.$$

この連立方程式が解を持つための必要十分条件は $b'_4 = 0$ である.

もしも $b'_4 \neq 0$ ならばこの連立方程式は解を持たない.

もしも $b'_4 = 0$ ならば例 4.22 と同様に上の連立方程式を x_1, x_3, x_5 について下から順番に解くことによって, $Ax = b$ の解空間

$$\mathcal{S} = \{x \in K^6 \mid Ax = b\} = \{x \in K^6 \mid \tilde{A}x = b'\}$$

は次のように表わされることがわかる:

$$\mathcal{S} = \left\{ \left[\begin{array}{c} b'_1 - b'_2 - 2b'_3 + x_2 + 3x_4 - 7x_6 \\ x_2 \\ b'_2 + b'_3 - x_4 + 3x_6 \\ x_4 \\ b'_3 + x_6 \\ x_6 \end{array} \right] \mid x_2, x_4, x_6 \in K \right\}. \quad \square$$

[143] 例 4.23 の計算が正しいかどうかを確かめよ. もしも正しいならばそのことを確認し, 誤りが含まれているならばそれを修正せよ. \square

[144] 行列 A とベクトル b が次のように定められているとき, 非斉次な一次方程式 $Ax = b$ を上で説明した方法を用いて解け:

$$A = \begin{bmatrix} 2 & 1 & 9 & -4 \\ -3 & 2 & -13 & -5 \\ -1 & -3 & -8 & 9 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \\ 10 \end{bmatrix}. \quad \square$$

略解. 問題より一般的に $b = {}^t[p, q, r]$ という状況について考える. そのとき $[A, b]$ は行に関する基本変形で次の形に変形できる:

$$\left[\begin{array}{cccc|c} 1 & -3 & 4 & 9 & -p - q \\ 0 & 1 & -3 & -4 & 2p + q + r \\ 0 & 0 & 22 & 6 & -11p - 5q - 7r \end{array} \right].$$

よって $Ax = b$ は次と同値である:

$$\begin{aligned} x_1 - 3x_2 + 4x_3 + 9x_4 &= -p - q, \\ x_2 - 3x_3 - 4x_4 &= 2p + q + r, \\ 22x_3 + 6x_4 &= -11p - 5q - 7r. \end{aligned}$$

これを x_1, x_2, x_3 について解くと,

$$\begin{aligned} x_1 &= \frac{18x_4}{11} + \frac{55p + 19q + 31r}{22}, \\ x_2 &= \frac{35x_4}{11} + \frac{11p + 7q + r}{22}, \\ x_3 &= \frac{-3x_4}{11} + \frac{-11p - 5q - 7r}{22}. \end{aligned}$$

よって $x_4 = 11t$ と置くと, $Ax = b$ の解の全体は $t \in K$ で次のようにパラメーター付けられる:

$$\begin{aligned}x_1 &= 18t + \frac{55p + 19q + 31r}{22}, \\x_2 &= 35t + \frac{11p + 7q + r}{22}, \\x_3 &= -3t + \frac{-11p - 5q - 7r}{22}, \\x_4 &= 11t.\end{aligned}$$

これに $b = {}^t[p, q, r] = {}^t[1, 3, 10]$ を代入すると,

$$x_1 = 18t + \frac{211}{11}, \quad x_2 = 35t + \frac{21}{11}, \quad x_3 = -3t + \frac{-48}{11}, \quad x_4 = 11t. \quad \square$$

[145] 行列 A とベクトル b を次のように定める:

$$A = \begin{bmatrix} 2 & 2 & 3 & -4 \\ 3 & -1 & 2 & -5 \\ 1 & 5 & 4 & -3 \end{bmatrix}, \quad b = \begin{bmatrix} p \\ q \\ r \end{bmatrix}$$

一次方程式 $Ax = b$ の解が存在するための必要十分条件を p, q, r に関する条件の形で述べよ. 解が存在する場合に $Ax = b$ を解け. \square

略解. 行に関する行列の基本で $A' = [A, b]$ を次に変形できる:

$$\left[\begin{array}{cccc|c} 1 & 5 & 4 & -3 & r \\ 0 & 8 & 5 & -2 & -p + 2r \\ 0 & 0 & 0 & 0 & 2p - q - r \end{array} \right].$$

したがって $Ax = b$ の解が存在するための必要十分条件は $2p - q - r = 0$ である. $2p - q - r = 0$ と仮定する. このとき $Ax = b$ は次と同値である:

$$x_1 + 5x_2 + 4x_3 - 3x_4 = r, \quad 8x_2 + 5x_3 - 2x_4 = -p + 2r.$$

これを x_1, x_2 について解くと,

$$x_1 = \frac{5p - 2r}{8} - \frac{7}{8}x_3 + \frac{7}{4}x_4, \quad x_2 = \frac{-p + 2r}{8} - \frac{5}{8}x_3 + \frac{1}{4}x_4.$$

$x_3 = 8\alpha, x_4 = 4\beta$ と置くことによって $Ax = b$ の解空間 \mathcal{S} は次のように表わされる:

$$\mathcal{S} = \left\{ \begin{bmatrix} \frac{5p-2r}{8} - 7\alpha + 7\beta \\ \frac{-p+2r}{8} - 5\alpha + \beta \\ 8\alpha \\ 4\beta \end{bmatrix} \middle| \alpha, \beta \in K \right\}. \quad \square$$

[146] 行列 A とベクトル b を次のように定める:

$$A = \begin{bmatrix} 4 & -8 & 3 & 3 \\ -2 & 4 & 1 & -2 \\ 1 & -2 & 1 & 1 \\ 2 & -4 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix}$$

一次方程式 $Ax = b$ の解が存在するための必要十分条件を p, q, r, s に関する条件の形で述べよ. 解が存在する場合に $Ax = b$ を解け. \square

略解. $A' = [A, b]$ は行に関する基本変形で次の形に変形できる:

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & 1 & r \\ 0 & 0 & 1 & 1 & -p + 4r \\ 0 & 0 & 0 & 1 & -2p + 6r + s \\ 0 & 0 & 0 & 0 & -3p + q + 8r + 3s \end{array} \right].$$

よって $Ax = b$ の解が存在するための必要十分条件は $-3p + q + 8r + 3s = 0$ である. 以下その条件を仮定する. そのとき $Ax = b$ は次と同値である:

$$x_1 - 2x_2 + x_3 + x_4 = r, \quad x_3 + x_4 = -p + 4r, \quad x_4 = -2p + 6r + s.$$

これを x_1, x_3, x_4 について解くと,

$$x_1 = p - 3r + 2x_2, \quad x_3 = p - 2r - s, \quad x_4 = -2p + 6r + s.$$

よって $Ax = b$ の解空間 \mathcal{S} は次のように表わされる:

$$\mathcal{S} = \left\{ \left[\begin{array}{c} p - 3r + 2\alpha \\ \alpha \\ p - 2r - s \\ -2p + 6r + s \end{array} \right] \mid \alpha \in K \right\}. \quad \square$$

5 体上のベクトル空間の理論

5.1 可換環上の加群と体上のベクトル空間の定義

体とそれ上のベクトル空間を一般的に定義する手間と可換環とそれ上の加群を定義する手間にはさほど違いはないので, より一般的な可換環とそれ上の加群をまとめて定義してしまう.

定義 5.1 (可換環と体) R が可換環 (**commutative ring**) であるとは⁸³, R が集合であり, 加法 $+: R \times R \rightarrow R$ と $0 \in R$ と加法の逆元 $-: R \rightarrow R$ と乗法 $\cdot: R \times R \rightarrow R$ と $1 \in R$ が与えられていて, 以下が成立していることである:

1. R は加法に関して可換群をなす. すなわち $a, b, c \in R$ に対して,
 - (a) $(a + b) + c = a + (b + c)$;
 - (b) $0 + a = a + 0 = a$;
 - (c) $(-a) + a = a + (-a) = 0$;
 - (d) $a + b = b + a$.
2. 乗法 $\cdot: R \times R \rightarrow R$ は結合的 (**associative**) かつ双加法的 (**bi-additive**) であり, $1 \in R$ は乗法に関する単位元になる. すなわち $a, b, c \in R$ に対して,
 - (a) $(ab)c = a(bc)$;
 - (b) $a(b + c) = ab + ac$;
 - (c) $(a + b)c = ac + bc$;

⁸³面倒な場合には単に環 (ring) と呼ぶ場合もある.

(d) $1a = a1 = a$.

3. R の乗法は可換 (commutative) である. すなわち $a, b \in R$ に対して

(e) $ab = ba$.

さらに次の条件が成立しているならば R は体 (field) であるという⁸⁴:

4. 任意の $a \in R \setminus \{0\}$ に対してある $b \in R \setminus \{0\}$ が存在して $ba = ab = 1$.

このような b は a に対して一意的に定まる. 実際 $ba = 1, ab' = 1$ ならば $b' = 1b' = (ba)b' = b(ab') = b1 = b$. 要するに 0 でない $a \in R$ に対してその逆元 $a^{-1} = 1/a$ が常に R 自身の中に存在するような可換環を体と呼ぶのである. たとえば $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ は体である. \square

[147] (二元体) 集合 $\mathbb{F}_2 = \{0, 1\}$ に次のように加法と乗法を定めると \mathbb{F}_2 は体をなす:

$$\begin{aligned} 0+0 &= 0, & 0+1 &= 1, & 1+0 &= 1, & 1+1 &= 0; \\ 0 \cdot 0 &= 0, & 0 \cdot 1 &= 0, & 1 \cdot 0 &= 0, & 1 \cdot 1 &= 1. \end{aligned} \quad \square$$

参考 5.2 (有限体) 一般に素数 p が与えられたとき, 集合 $\mathbb{F}_p = \{0, 1, \dots, p-1\}$ に通常の整数の和と積の p で割った余りを考えることによって \mathbb{F}_p の加法と乗法を定めると \mathbb{F}_p は体をなす. 有限個の元しか持たない体を有限体 (finite field) と呼ぶ. 有限体の元の個数 (有限体の位数と呼ばれる) は素数の中 $q = p^e$ (p は素数, e は正の整数) になる. 位数 q の有限体は \mathcal{F}_q と表わされる. \square

参考 5.3 (ガロア) なお, 有限体は Galois 体 (Galois field) と呼ばれ $GF(q)$ と表わされることもある. Évariste Galois (エヴァリスト・ガロア, 1811.10.25–1832.5.31) は 19 世紀の初頭に登場した天才数学者の一人である. 決闘で悲劇的な結末をむかえたことで有名である⁸⁵. 群 (group) を導入して対称性 (symmetry) の概念を数学的に明確にし, 方程式をその対称性を調べることに統制するという考え方を導入したのは Galois である. 19 世紀は天才数学者が次々に登場した世紀であり, その歴史は非常に面白い. \square

参考 5.4 (有限体上の幾何) 慣れ親しんで来た実数体や複素数体の世界と有限体の世界はまるで違っているように見えるかもしれない. しかし, 実際にはそうではないことが知られている. 実数体をもとにして定義された図形には連続性の直観が適用でき, トポロジーの理論が展開される. 有限体上の代数多様体 (これもある種の図形) に対してもトポロジーの理論を展開することができる⁸⁶. このような驚くべき数学の発展が 20 世紀のあいだになされた⁸⁷. \square

[148] (可換環上の多項式環) k が可換環ならば k 上の一変数多項式環 $R = k[x]$ も自然に可換環をなす. \square

ヒント. R が可換環であることを確かめるためには, まず加法 $+: R \times R \rightarrow R$ と $0 \in R$ と加法の逆元 $-: R \rightarrow R$ と乗法 $\cdot: R \times R \rightarrow R$ と $1 \in R$ の定義を明確にし, それらが可換環の公理を満たしていることを証明する. \square

⁸⁴可換体 (commutative field) と呼ぶ場合もある. 非可換な体は斜体 (skew field) と呼ばれる.

⁸⁵おすすめの伝記はインフェルト [I] である.

⁸⁶Grothendieck の étale topology の理論.

⁸⁷有限体上の幾何は純粋数学的に重要なだけでなく, 我々の実生活に関わる応用面でも重要である. 有限体はコンピューターと相性が良い.

定義 5.5 (可換環上の加群と体上のベクトル空間) R は可換環であるとする. 集合 M が R 上の加群 (module over R) もしくは R 加群 (R -module) であるとは加法 $+: M \times M \rightarrow M$, 零元 $0 \in M$ と加法に関する逆元 $-: M \rightarrow M$ と R の元の M の元への作用 $\cdot: R \times M \rightarrow M$ が定義されていて, 以下の R 加群の公理が成立していることである:

1. M は加法に関して可換群をなす. すなわち任意の $u, v, w \in M$ に対して,
 - (a) $(u + v) + w = u + (v + w)$;
 - (b) $0 + u = u + 0 = u$;
 - (c) $(-u) + u = u + (-u) = 0$;
 - (d) $u + v = v + u$.
2. スカラー倍 $\cdot: R \times M \rightarrow M$ は結合的かつ双加法的 (bi-additive) であり, $1 \in R$ の作用は恒等写像になる. すなわち任意の $a, b \in R, u, v \in M$ に対して,
 - (a) $(ab)u = a(bu)$;
 - (b) $a(u + v) = au + av$;
 - (c) $(a + b)u = au + bu$;
 - (d) $1u = u$.

特に R が体 K に等しいならば R 加群を K 上のベクトル空間 (vector space over K) もしくは K 上の線形空間 (linear space over K) もしくは K ベクトル空間 (K -vector space) もしくは K 線形空間 (K -linear space) と呼ぶ. \square

[149] K は体であるとし, 可換環 R は体 K 上の一変数多項式環 $K[\lambda]$ であるとし, $M = K^n$ (縦ベクトルの空間) と置き, 正方行列 $A \in M_n(K)$ を任意に固定する. $f(\lambda) \in K[\lambda]$ の λ に A を代入することによって行列 $f(A) \in M_n(K)$ が自然に定義される. そのことを利用して写像 $\cdot: R \times M \rightarrow M$ を

$$f(\lambda) \cdot v := f(A)v \quad (f(\lambda) \in R = K[\lambda], v \in M = K^n)$$

と定める. このとき M は自然に $R = K[\lambda]$ 上の加群をなす. \square

ヒント. $M = K^n$ は K 上のベクトル空間なので始めから, $+$, 0 , $-$ が定められている. スカラー倍 $R \times M \rightarrow M$ は問題のように定められている. よってそれらが R 加群の公理を満たしているかどうかを確かめればよい. ($M = K^n$ はベクトル空間なので加法に関して可換群をなすことは改めてチェックしなくてよいだろう.) \square

参考 5.6 上の問題で定義した $K[\lambda]$ 上の加群 M は正方行列 $A \in M_n(K)$ の Jordan 標準形の理論を扱うときに重要になる. 正方行列 A の標準形の理論と A に対応する $K[\lambda]$ 加群 M の構造に関する理論は本質的に等しくなる. 代数学の基本は環と加群の理論である. \square

[150] (連続関数全体のなすベクトル空間) 閉区間 $[a, b]$ 上の実数値連続関数全体のなす集合を $C([a, b], \mathbb{R})$ と書くことにする. $C([a, b], \mathbb{R})$ は自然に \mathbb{R} 上のベクトル空間をなす. \square

ヒント. まず, $C([a, b], \mathbb{R})$ に $+$, 0 , $-$ とスカラー倍 $\cdot: \mathbb{R} \times C([a, b], \mathbb{R}) \rightarrow C([a, b], \mathbb{R})$ を定義せよ. それらが well-defined であることを証明し, さらにそれらがベクトル空間の公理を満たしていることを示せ. たとえば $[a, b]$ 上の実数値連続関数 f と g の和もまた連続関数になることなどを証明しなければいけない. \square

参考 5.7 K^n のようなベクトル空間を扱うのではなく, より抽象的に体上のベクトル空間を扱う利点の一つは, 上の問題のようにある種の函数全体の空間 (函数空間) をもベクトル空間として扱えるようになることである. すでに十分習熟しつつあると思われる K^n およびその部分空間でやしなれた直観を函数空間にも拡大するように努力せよ. \square

[151] (連続函数全体のなす可換環) 閉区間 $[a, b]$ 上の実数値連続函数全体のなす集合を $C([a, b], \mathbb{R})$ と書くことにする. $C([a, b], \mathbb{R})$ は自然に可換環をなす. \square

ヒント. $f, g \in C([a, b], \mathbb{R})$ に対して $[a, b]$ 上の函数 fg を $(fg)(x) = f(x)g(x)$ ($x \in [a, b]$) と定めると, fg が $[a, b]$ 上の連続函数になること (すなわち $fg \in C([a, b], \mathbb{R})$ となること) などを証明する必要がある. \square

n 回微分可能でかつ n 階の導函数が連続になる函数を C^n 級函数 (class- C^∞ function) もしくは C^n 函数 (C^∞ -function) と呼ぶ. 任意有限回微分可能な函数を C^∞ 級函数 (class- C^∞ function) もしくは C^∞ 函数 (C^∞ -function) と呼ぶ.

[152] (Leibnitz rule) f, g が开区間 (a, b) 上の実数値 C^∞ 函数であるとき, $h(x) = f(x)g(x)$ ($x \in (a, b)$) と置くと, h も开区間 (a, b) 上の実数値 C^∞ 函数であり, その n 階の導函数 $h^{(n)}$ に関して次の公式が成立している:

$$h^{(n)}(x) = \sum_{k=0}^n \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x) \quad (x \in (a, b)).$$

この公式を積の微分に関する Leibnitz rule (ライプニッツ則) と呼ぶ. \square

ヒント. n に関する帰納法. n から $n+1$ に進むときに $(fg)' = f'g + fg'$ および二項係数 $\binom{n}{k}$ が Pascal の三角形と呼ばれる漸化式を満たしていることを使え. \square

[153] (C^∞ 函数全体のなす可換環) 开区間 (a, b) 上の実数値 C^∞ 函数全体のなす集合を $C^\infty((a, b), \mathbb{R})$ と書くと以下が成立している:

1. $C^\infty((a, b), \mathbb{R})$ は自然に \mathbb{R} 上のベクトル空間をなす.
2. $C^\infty((a, b), \mathbb{R})$ は自然に可換環をなす. \square

ヒント. 記号の簡単のため $A = C^\infty((a, b), \mathbb{R})$ と置く.

1. $+: A \times A \rightarrow A$, $0_A \in A$, $-: A \rightarrow A$, $\cdot: \mathbb{R} \times A \rightarrow A$ を次のように定める: $f, g \in A$, $\alpha \in \mathbb{R}$, $x \in (a, b)$ に対して,

$$(f+g)(x) = f(x) + g(x), \quad 0_A(x) = 0, \quad (-f)(x) = -f(x), \quad (\alpha \cdot f)(x) = \alpha f(x).$$

この定義のもとで A が \mathbb{R} 上のベクトル空間の公理を満たしていることを示せ.

2. さらに, $\cdot: A \times A \rightarrow A$, $1_A \in A$ を次のように定める: $f, g \in A$, $a \in (a, b)$ に対して,

$$(f \cdot g)(x) = f(x)g(x), \quad 1_A(x) = 1.$$

問題 [152] より, $f \cdot g$ もまた C^∞ 函数なので, 写像 $\cdot: A \times A \rightarrow A$ は well-defined である (うまく定義されている). よって以上の定義のもとで A が可換環をなすことを示せばよい. \square

5.2 加群の準同型とベクトル空間の線形写像の定義

定義 5.8 (加群の準同型とベクトル空間の線形写像) R は可換環であり, M, N は R 上の加群であるとし, $f: M \rightarrow N$ は任意の写像であるとする. このとき f が R 加群の準同型もしくは準同型写像 (**homomorphism of R -modules**) であるとは以下の条件が成立していることである:

1. $f(u + v) = f(u) + f(v) \quad (u, v \in M);$
2. $f(\alpha u) = \alpha f(u) \quad (\alpha \in R, u \in M).$

R 加群の準同型写像は R 準同型 (**R -homomorphism**) と呼ばれることも多い.

環 R が体 K に等しいとき, R 加群 ($= K$ 加群) は K 上のベクトル空間と呼ばれるのであった. そのとき R 加群の準同型は K 上の線形写像 (**linear mapping**) と呼ばれる. \square

[154] (準同型の合成) L, M, N は可換環 R 上の加群であり, $f: L \rightarrow M, g: M \rightarrow N$ は R 準同型であるとする. そのとき合成 $g \circ f: L \rightarrow N$ も R 準同型である. \square

[155] (Hom_R) R は可換環であるとし, M, N は R 加群であるとし,

$$\text{Hom}_R(M, N) = \{ f: M \rightarrow N \mid f \text{ は } R \text{ 準同型} \}$$

とおく. $\text{Hom}_R(M, N)$ に加法とスカラー倍の演算を次のように定めることができることを示せ: $f, g \in \text{Hom}_R(M, N), u \in M, \alpha \in R$ に対して,

$$(f + g)(u) := f(u) + g(u), \quad (\alpha \cdot f)(u) := \alpha f(u).$$

これによって $\text{Hom}_R(M, N)$ は自然に R 加群とみなせる. \square

注意 5.9 上の問題で特に R が体 K に等しいとき, M, N は K 上のベクトル空間であり,

$$\text{Hom}_R(M, N) = \text{Hom}_K(M, N) = \{ f: M \rightarrow N \mid f \text{ は } K \text{ 上の線形写像} \}$$

である. $\text{Hom}_K(M, N)$ は自然に K 上のベクトル空間をなす. \square

[156] (行列の定める線形写像) K は任意の体であるとし, K の元を成分に持つ n 次元縦ベクトル全体の空間を K^n と表わし, $m \times n$ 行列全体の空間を $M_{m,n}(K)$ と書くことにする. このとき, 任意の $m \times n$ 行列 $A \in M_{m,n}(K)$ に対して, 写像 $f_A: K^n \rightarrow K^m$ を

$$f_A(u) := Au \in K^m \quad (u \in K^n)$$

と定めると, f_A は K 上の線形写像である. \square

ヒント. この問題の結果はほとんど自明 (trivial) である. 我々は f_A のことを単に A と書いてきたのであった. \square

[157] (積分作用素) 問題 [150] の結果より, 閉区間 $[a, b]$ 上の実数値連続関数全体の集合 $C([a, b], \mathbb{R})$ は自然に \mathbb{R} 上のベクトル空間とみなされる⁸⁸. $K(x, y)$ は $[a, b] \times [a, b]$ 上の任意の実数値連続関数であるとする. このとき, \mathbb{R} 上の線形写像 $T : C([a, b], \mathbb{R}) \rightarrow C([a, b], \mathbb{R})$ を

$$(Tf)(x) := \int_a^b K(x, y)f(y) dy \quad (f \in C([a, b], \mathbb{R}), x \in [a, b])$$

と定めることができることを示せ (Tf もまた $[a, b]$ 上の連続関数になることも示せ). この T は積分作用素 (integral operator) と呼ばれ, $K(x, y)$ はその核関数 (kernel function) と呼ばれる⁸⁹. \square

ヒント. 閉区間 $[a, b]$ 上の積分について以下が成立することを自由に用いてよい⁹⁰:

1. (連続関数の積分可能性) 任意の $f \in C([a, b], \mathbb{R})$ は $[a, b]$ 上で積分可能である.
2. (積分の線形性) 任意の $f, g \in C([a, b], \mathbb{R})$ と $\alpha \in \mathbb{R}$ に対して,

$$\int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x) dx, \quad \int_a^b \alpha f(x) dx = \alpha \int_a^b f(x) dx.$$

3. (積分の単調性) 任意の $f, g \in C([a, b], \mathbb{R})$ に対して,

$$f(x) \leq g(x) \ (x \in [a, b]) \implies \int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

4. (積分の絶対値の評価) $f \in C([a, b], \mathbb{R})$ に対して, その絶対値 $|f|$ の $[a, b]$ での最大値を M と書くと⁹¹,

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx \leq \int_a^b M dx = M(b-a).$$

さらに, $K(x, y)$ の $[a, b] \times [a, b]$ 上での一様連続性を用いてよい⁹².

$f \in C([a, b], \mathbb{R})$ を任意に取り, $|f|$ の $[a, b]$ での最大値を M と書くと, 任意の $x_0, x \in [a, b]$ に対して,

$$\begin{aligned} |(Tf)(x) - (Tf)(x_0)| &= \left| \int_a^b (K(x, y) - K(x_0, y))f(y) dy \right| \\ &\leq \int_a^b |K(x, y) - K(x_0, y)| |f(y)| dy \\ &\leq M \int_a^b |K(x, y) - K(x_0, y)| dy. \end{aligned}$$

⁸⁸「連続な」という形容詞は英語では “continuous” である. 記号 $C([a, b], \mathbb{R})$ の C は「連続」という意味である.

⁸⁹線形写像 $f : U \rightarrow V$ の核 $\text{Ker } f = \{x \in U \mid f(x) = 0\}$ とは無関係であることに注意せよ.

⁹⁰以下の条件は積分の定義の仕方 (Riemann, Lebesgue) によらずに成立している. 以下の条件は積分が x 軸と函数のグラフで囲まれた部分の面積 (x 軸より下の部分の面積は -1 倍する) であるという直観より, 当然成立すべき事柄ばかりである.

⁹¹一般に \mathbb{R}^n の有界閉集合上の実数値連続関数は最大値と最小値を持つ.

⁹²一般に \mathbb{R}^n の有界閉集合上の実数値連続関数は一様連続である.

連続関数 $K(x, y)$ の $[a, b] \times [a, b]$ 上での一様連続性より, 任意の $\varepsilon > 0$ に対して, ある $\delta > 0$ が存在して, $|x - x_0| \leq \delta$ ならば $|K(x, y) - K(x_0, y)| \leq \varepsilon / (M(a - b))$ ($y \in [a, b]$ は任意でよい) となる. よって $|x - x_0| \leq \delta$ ならば

$$|(Tf)(x) - (Tf)(x_0)| \leq M \int_a^b \frac{\varepsilon}{M(a - b)} dy = \varepsilon.$$

これで Tf が $[a, b]$ 上の連続関数であることが示された. \square

注意 5.10 (積分作用素と行列の定める線形写像の類似) 積分作用素 T の定義は行列の定める線形写像の定義と似ている. 実際, $A = [a_{ij}] \in M_{m,n}(\mathbb{R})$, $v = [v_i] \in \mathbb{R}^n$ に対して, $Av \in \mathbb{R}^m$ の第 i 成分を $(Av)_i$ と書くと,

$$(Av)_i = \sum_{j=1}^n a_{ij} v_j.$$

一方, 積分作用素 T は次のように定義されたのであった:

$$(Tf)(x) = \int_a^b K(x, y) f(y) dy.$$

以上の2つの式を比べれば, 以下のような類似関係があることがわかる:

$$\begin{aligned} \text{積分作用素 } T &\leftrightarrow \text{行列 } A, & \text{関数 } f &\leftrightarrow \text{縦ベクトル } v, \\ \text{核関数 } K(x, y) &\leftrightarrow \text{行列の成分 } a_{ij}, & \text{積分 } \int_a^b dy &\leftrightarrow \text{有限和 } \sum_{j=1}^n. \end{aligned} \quad \square$$

[158] (微分作用素) 問題 [153] の結果より, 开区間 (a, b) 上の実数値 C^∞ 関数全体のなす集合 $C^\infty((a, b), \mathbb{R})$ は自然に \mathbb{R} 上のベクトル空間でかつ可換環とみなされる. 記号の簡単のため $A = C^\infty((a, b), \mathbb{R})$ と置く. 以下が成立することを示せ:

1. $f \in A$ に対して, 写像 $\hat{f}: A \rightarrow A$ を

$$\hat{f}(g) := f \cdot g \quad (g \in A)$$

と定めると, \hat{f} は \mathbb{R} 上の線形写像である.

2. 写像 $\partial: A \rightarrow A$ を

$$\partial(f) := f' \quad (f \in A, f' \text{ は } f \text{ の導関数})$$

と定めると, ∂ は \mathbb{R} 上の線形写像である.

3. $a_0, a_1, \dots, a_n \in A$ に対して写像 $P: A \rightarrow A$ を

$$P(f) := a_n f^{(n)} + a_{n-1} f^{(n-1)} + \dots + a_1 f' + a_0 f \quad (f \in A)$$

と定める. ここで $f^{(k)}$ は f の k 階の導関数である. このとき P は \mathbb{R} 上の線形写像である.

4. 任意の $f \in A$ に対して $[\partial, \hat{f}] = \hat{f}'$. ($[A, B] = AB - BA$ である.)

P は線形常微分作用素 (linear ordinary differential operator) と呼ばれ,

$$P = a_n \partial^n + a_{n-1} \partial^{n-1} + \cdots + a_1 \partial + a_0$$

と書かれる. \square

ヒント. 4 は次のようにして証明される. 任意の $g \in A$ を取ると,

$$\partial(\hat{f}(g)) = \partial(fg) = (fg)' = f'g + fg' = \hat{f}'(g) + \hat{f}(\partial(g))$$

なので $\partial(\hat{f}(g)) - \hat{f}(\partial(g)) = \hat{f}'(g)$. これで $[\partial, \hat{f}] = \partial\hat{f} - \hat{f}\partial = \hat{f}'$ が証明された. \square

参考 5.11 積分作用素と微分作用素は線形写像を作るための材料として行列と同じくらい基本的である. 数ベクトルと行列の理論をベクトル空間と線形写像の理論に一般化しておくことのメリットの一つはある種の函数全体のなす空間のあいだの微分作用素や積分作用素も扱えるようになることである. \square

[159] (多項式係数の微分作用素) 複素係数の一変数多項式環 $\mathbb{C}[x]$ は自然に \mathbb{C} 上のベクトル空間をなす. 以下が成立することを示せ:

1. $f \in \mathbb{C}[x]$ に対して, 写像 $\hat{f}: \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ を

$$\hat{f}(g) := f \cdot g \quad (g \in \mathbb{C}[x])$$

と定めると, \hat{f} は \mathbb{C} 上の線形写像である.

2. 写像 $\partial: \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ を

$$\partial(f) := f' \quad (f \in \mathbb{C}[x], f' \text{ は } f \text{ の導関数})$$

と定めると, ∂ は \mathbb{C} 上の線形写像である.

3. $a_0, a_1, \dots, a_n \in \mathbb{C}[x]$ に対して写像 $P: \mathbb{C}[x] \rightarrow \mathbb{C}[x]$ を

$$P(f) := a_n f^{(n)} + a_{n-1} f^{(n-1)} + \cdots + a_1 f' + a_0 f \quad (f \in A)$$

と定める. ここで $f^{(k)}$ は f の k 階の導関数である. このとき P は \mathbb{C} 上の線形写像である.

4. $[\partial, \hat{x}^i] = i\widehat{x^{i-1}}$. 特に $[\partial, \hat{x}] = 1$. ($[A, B] = AB - BA$ である.)

P は多項式係数の線形常微分作用素 (linear ordinary differential operator with polynomial coefficients) と呼ばれ,

$$P = a_n \partial^n + a_{n-1} \partial^{n-1} + \cdots + a_1 \partial + a_0$$

と書かれる. \square

注意 5.12 上の問題 [159] の結果は問題 [175], [176] で応用される. \square

5.3 一次独立性と基底

第4.4節と第4.6節において、体 K に対する K^n およびその部分空間における一次独立性と基底について説明した。それらの結果は体 K 上の一般のベクトル空間に拡張される。詳しい説明については線形代数の教科書を参照せよ。

特にベクトル空間の基底の存在と基底の元の個数の一意性 (ベクトル空間の次元がうまく定義されること) については自前の詳しいノートを作っておくことが望ましい。

K^n およびその部分空間の基底については問題 [119], [120], [121] を見よ。実はそのヒントの議論のかなりの部分は一般のベクトル空間にも適用可能である。手もとにある線形代数の教科書を参考にしながら、別の教科書を書くつもりで理論をまとめると非常に良い勉強になる。

[160] (直和) K 上のベクトル空間 V とその部分空間 V_1, \dots, V_N に関して以下の2条件は互いに同値である:

- (a) 任意の $v \in V$ は $v = v_1 + \dots + v_N$, $v_i \in V_i$ と一意に表わされる。
- (b) 任意の $v \in V$ は $v = v_1 + \dots + v_N$, $v_i \in V_i$ と表わされ、任意の $v_i \in V_i$ ($i = 1, \dots, N$) に対して $v_1 + \dots + v_N = 0$ ならば $v_i = 0$ ($i = 1, \dots, N$) である。

この同値な条件のどちらかが成立するとき、 $V = V_1 \oplus \dots \oplus V_N$ と書き、 V は V_1, \dots, V_N の直和 (direct sum) であると言う。さらに各 V_i が有限次元でかつ $V = V_1 \oplus \dots \oplus V_N$ ならば

$$\dim V = \dim V_1 + \dots + \dim V_N$$

が成立する。□

[161] (補空間の存在) 体 K 上のベクトル空間 U とその部分空間 V に対して、 V の基底を U の基底に拡張できることを用いて、 U の部分空間 W で $U = V \oplus W$ を満たすものが存在することを示せ。そのような W を U における V の (線形) 補空間 (linear complement) と呼ぶ。□

ヒント. V の基底 $\{v_i\}_{i \in I}$ を U の基底 $\{v_i\}_{i \in I} \cup \{w_j\}_{j \in J}$ に拡張して、 W を $\{w_j\}_{j \in J}$ で張られる U の部分空間とすると $U = V \oplus W$ である。□

5.4 同型写像

[162] (同型写像の定義) R は可換環であり、 M, N は R 加群であり、 $f: M \rightarrow N$ は R 準同型であるとする。もしも f が全単射⁹³ならば逆写像 f^{-1} も R 準同型である。そのとき f は R 加群の同型写像 (isomorphism of R -modules) もしくは R 同型 (R -isomorphism) と呼ばれる。□

⁹³一般に集合間の写像 $f: X \rightarrow Y$ の単射性、全射性は以下のように定義される:

- f が単射 (injection) \iff 任意の $x, x' \in X$ に対して $f(x) = f(x')$ ならば $x = x'$.
- f が全射 (surjection) \iff 任意の $y \in Y$ に対してある $x \in X$ で $f(x) = y$ をみたすものが存在する。

全射かつ単射な写像を全単射 (bijection) と呼ぶ。写像が逆写像を持つための必要十分条件はそれが全単射であることである。

定義 5.13 (R 加群の同型性) 可換環 R 上の加群 M, N が R 上同型 (isomorphic) であるとは, M から N への R 加群の同型写像が存在することである. (体 K 上のベクトル空間の同型性も同様に定義する.) \square

[163] (ベクトル空間の同型性と次元の関係) U, V は体 K 上の有限次元ベクトル空間であるとする. そのとき U と V が同型であるための必要十分条件は U と V の次元が等しいことである. \square

ヒント. U の基底 u_1, \dots, u_n を取る. U と V が同型ならばある同型写像 $f: U \xrightarrow{\sim} V$ が存在する. そのとき $f(u_1), \dots, f(u_n)$ が V の基底になることを示せる. U と V の次元が等しいならば, V の基底 v_1, \dots, v_n を任意に取って, 同型写像 $f: U \rightarrow V$ を $f(\alpha_1 u_1 + \dots + \alpha_n u_n) = \alpha_1 v_1 + \dots + \alpha_n v_n$ ($\alpha_i \in K$) と構成できる. \square

[164] (単射性の kernel による判定法) R は可換環であるとし, M, N は R 加群であるとし, $f: M \rightarrow N$ は R 準同型であるとする. f の核 (kernel) $\text{Ker } f$ が $\text{Ker } f = \{u \in M \mid f(u) = 0\}$ と定義される⁹⁴. このとき, 以下の条件は互いに同値である⁹⁵:

- (a) f は単射である.
- (b) 任意の $u \in M$ に対して $f(u) = 0$ ならば $u = 0$.
- (c) $\text{Ker } f = \{0\}$. \square

ヒント. (c) は (b) の言い換えに過ぎないので, (a) と (b) の同値性を示せばよい. (a) が成立するとき, 任意の $u \in M$ に対して, $f(u) = 0 = f(0)$ ならば $u = 0$ なので, (b) が成立する. (b) が成立するとき, 任意の $u, u' \in M$ に対して, $f(u) = f(u')$ ならば $f(u - u') = 0$ なので $u - u' = 0$ すなわち $u = u'$ である. よって (a) が成立する. \square

要約 5.14 M, N が可換環 R 上の加群であり, $f: M \rightarrow N$ が R 準同型であるとき以下が成立する:

- 1. f は単射 $\iff \text{Ker } f = 0$.
- 2. f は全射 $\iff \text{Im } f = N$.
- 3. f は同型 $\iff f$ は全単射 $\iff \text{Ker } f = 0$ かつ $\text{Im } f = N$. \square

5.5 線形写像の行列表示

K は体であるとし, U, V は K 上の有限次元ベクトル空間であるとし, $f: U \rightarrow V$ は K 上の任意の線形写像であるとする. 線形写像 f 自身は極めて抽象的な数学的对象であるが, U と V に基底を定めることによって, f を具体的に行列で表現することができる.

u_1, \dots, u_n は U の基底であり, v_1, \dots, v_m は V の基底であるとする. このとき, 任意の $u \in U, v \in V$ は次のように一意に表わされる:

$$u = \sum_{j=1}^n \alpha_j u_j = \sum_{j=1}^n u_j \alpha_j = [u_1, \dots, u_n] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad (\alpha_j \in K),$$

⁹⁴ f の像 (image) $\text{Im } f$ は $\text{Im } f = \{f(u) \mid u \in M\}$ と定義される.

⁹⁵ この問題の結果は準同型写像もしくは線形写像が単射であるための判定法として今後自由に使われる.

$$v = \sum_{i=1}^m \beta_i v_i = \sum_{i=1}^m v_i \beta_i = [v_1, \dots, v_m] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \quad (\beta_i \in K).$$

これによって $u \in U$ と $\alpha = {}^t[\alpha_1, \dots, \alpha_n] \in K^n$ が一対一に対応し, $v \in V$ と $\beta = {}^t[\beta_1, \dots, \beta_m] \in K^m$ が一対一に対応する. この対応を用いて, 線形写像 $f: U \rightarrow V$ と行列 $A = [a_{ij}] \in M_{m,n}(K)$ の一対一対応を構成可能であることを説明しよう.

まず, 各 $f(u_j) \in V$ は

$$f(u_j) = \sum_{i=1}^m a_{ij} v_i = \sum_{i=1}^m v_i a_{ij} = [v_1, \dots, v_m] \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad (a_{ij} \in K)$$

と一意に表わされるので,

$$[f(u_1), \dots, f(u_n)] = [v_1, \dots, v_m] \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

よって

$$\begin{aligned} f(u) &= \sum_{j=1}^n \alpha_j f(u_j) = \sum_{j=1}^n f(u_j) \alpha_j \\ &= [f(u_1), \dots, f(u_n)] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = [v_1, \dots, v_m] \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}. \end{aligned}$$

以上の記号のもとで線形写像 f は

$$[u_1, \dots, u_n] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in U \text{ を } [v_1, \dots, v_m] \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in V \text{ に}$$

対応させる写像に等しい. 以上のようにして線形写像 f に対応する行列 $A = [a_{ij}]$ が得られる. 逆に行列 $A = [a_{ij}]$ が与えられれば上の対応によって線形写像 $f: U \rightarrow V$ が得られることもわかる. 行列 $A = [a_{ij}]$ を線形写像 f の基底 u_j, v_i に関する行列表示と呼ぶことにする.

要約 5.15 (線形写像の行列表示) U の基底 u_1, \dots, u_n と V の基底 v_1, \dots, v_m に関する線形写像 $f: U \rightarrow V$ の行列表示 $A = [a_{ij}] \in M_{m,n}(K)$ は次の条件によって一意に決定される⁹⁶:

$$[f(u_1), \dots, f(u_n)] = [v_1, \dots, v_m] \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

⁹⁶定義域の基底を横に並べたものに f を左から作用させて, 右側にポコッと出て来る行列 $A = [a_{ij}]$ を計算すれば線形写像 f の行列表示が得られる.

この条件は次と同値である:

$$f(u_j) = \sum_{i=1}^m a_{ij}v_i = \sum_{i=1}^m v_i a_{ij} \quad (j = 1, \dots, n). \quad \square$$

[165] $U = \mathbb{R}^3$, $V = \mathbb{R}^2$ とし, 行列

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$$

の積の定める U から V への線形写像を f と書くことにする (すなわち $f(u) = Au$ ($u \in U = \mathbb{R}^3$)). $u_1, u_2, u_3 \in U$ と $v_1, v_2 \in V$ を次のように定める:

$$u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

このとき, u_1, u_2, u_3 は U の基底であり, v_1, v_2 は V の基底であり, それらに関する f の行列表示を B とすると, B は

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

と簡単な形になることを示せ. \square

ヒント. $[f(u_1), f(u_2), f(u_3)] = [v_1, v_2]B$ を示せ. $[f(u_1), f(u_2), f(u_3)] = [Au_1, Au_2, Au_3] = A[u_1, u_2, u_3]$ なので $A[u_1, u_2, u_3] = [v_1, v_2]B$ が成立することを直接的な計算で示せばよい. というわけでこの問題は非常に簡単な問題である. \square

注意 5.16 (行列の基本変形との関係) 上のような問題の作り方は問題 [111] を理解すればわかる. \square

注意 5.17 (標準的な基底以外のより適切な基底を見付けることの重要性) 上の問題のように行列 A 自身は複雑な形をしていても, 標準的な基底とは別の基底に関して行列表示し直すと簡単な形になることがよくある. 与えられた線形写像の本質を見極めるためには適切な基底を見付けて行列表示してみることが役に立つ.

実は行列の基本変形や (後で習うことになっている) 行列の対角化や Jordan 標準形の理論はどれも「行列もしくは線形写像の本質を見極めるために役に立つ基底の見付け方に関する理論」とみなせる. \square

[166] K は体であるとし, $m \times n$ 行列 $A \in M_{m,n}(K)$ を任意に取る. K^l の標準的基底を $e_1^{(l)}, \dots, e_l^{(l)}$ と書くことにする. すなわち $e_i^{(l)} \in K^l$ は第 i 成分のみが 1 で他の成分は 0 であるとする. 基底 $e_j^{(n)}, e_i^{(m)}$ に関する A の定める線形写像 $A: K^n \rightarrow K^m$ の行列表示は A 自身に等しい. \square

ヒント. $[Ae_1^{(n)}, \dots, Ae_n^{(n)}] = [e_1^{(m)}, \dots, e_m^{(m)}]A$ を示せばよいがほとんど自明である. \square

[167] (基底の変換) K は体であるとし, U, V は K 上の有限次元ベクトル空間であり, u_1, \dots, u_n は U の基底であり, v_1, \dots, v_m は V の基底であるとする. $f: U \rightarrow V$ は線形写像であり, $A \in M_{m,n}(K)$ は基底 u_j, v_i に関する f の行列表示であるとする. u'_1, \dots, u'_n と v'_1, \dots, v'_m はそれぞれ U, V の別の基底であるとする. 以下を示せ.

1. ある可逆な行列 $Q \in GL_n(K), P \in GL_m(K)$ で⁹⁷

$$[u'_1, \dots, u'_n] = [u_1, \dots, u_n]Q, \quad [v'_1, \dots, v'_m] = [v_1, \dots, v_m]P$$

をみたすものが一意に存在する.

2. 基底 u'_j, v'_i に関する f の行列表示は $P^{-1}AQ$ になる. \square

ヒント. 2. $[f(u'_1), \dots, f(u'_n)] = [v'_1, \dots, v'_m]P^{-1}AQ$ を 1 を用いて示せばよい. \square

[168] K は体であるとし, $u_1, \dots, u_n \in K^n$ は K^n の基底であり, $v_1, \dots, v_m \in K^m$ は K^m の基底であるとし, $Q = [u_1, \dots, u_n] \in M_n(K), P = [v_1, \dots, v_m] \in M_m(K)$ とおく. このとき, $m \times n$ 行列 $A \in M_{m,n}(K)$ の定める線形写像 $A: K^n \rightarrow K^m$ の基底 u_j, v_i に関する行列表示は $P^{-1}AQ$ になる. \square

ヒント. 問題 [166], [167] からただちに得られる. もしくは $[Au_1, \dots, Au_n] = AQ = PP^{-1}AQ = [v_1, \dots, v_m]P^{-1}AQ$. \square

[169] $V = \mathbb{R}^2$ とし, 行列

$$A = \frac{1}{5} \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$$

が定める V からそれ自身への線形写像を f と書くことにする. $v_1, v_2 \in V$ を

$$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

と定めると, v_1, v_2 は V の基底である (v_1, v_2 を平面上の図示せよ). 基底 v_i に関する f の行列表示を求めよ. \square

ヒント. $[f(v_1), f(v_2)] = [v_1, v_2]B$ を満たす行列 $B \in M_2(\mathbb{R})$ が答である. \square

略解. $Av_1 = v_1, Av_2 = 2v_2$ なので $B = \text{diag}(1, 2)$. \square

[170] 問題 [169] の結果を用いて, 次の常微分方程式の初期値問題を解け:

$$\frac{d}{dt}u = Au, \quad u(0) = u_0.$$

ここで u は $t \in \mathbb{R}$ の $V = \mathbb{R}^2$ に値を持つ関数であり, $u_0 = e_2 = {}^t[0, 1]$. \square

⁹⁷ $GL_n(K)$ は K の元を成分に持つ可逆な $n \times n$ 行列全体の集合である. $GL_n(K)$ は群をなし, 一般線形群と呼ばれる.

ヒント. まず「行列の指数関数」に関する節 (第 1.9 節⁹⁸) を読め. 特に問題 [21] とそのヒントを見よ. $P = [v_1, v_2]$ と置くと $A = PBP^{-1}$ であるから, 問題 [23] の結果より,

$$e^{tA} = Pe^{tB}P^{-1}.$$

実は B は対角行列になるので e^{tB} は容易に計算される. その結果を用いて $u(t) = e^{tA}u_0$ を整理したものが答になる. \square

略解. $e^{tB} = \text{diag}(e^t, e^{2t})$ であり, $P = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$, $P^{-1} = \frac{1}{5} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} = \frac{1}{5} {}^tP$ なので

$$e^{tA} = Pe^{tA}P^{-1} = \frac{1}{5} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} e^t + 4e^{2t} & 2e^t - 2e^{2t} \\ 2e^t - 2e^{2t} & 4e^t + e^{2t} \end{bmatrix}.$$

よって $u(t) = e^{tA}u_0 = e^{tA}e_2 = \frac{1}{5} \begin{bmatrix} 2e^t - 2e^{2t} \\ 4e^t + e^{2t} \end{bmatrix}$. \square

[171] (一次変換の対角化) V は体 K 上のベクトル空間であり, f は V の一次変換 (すなわち V からそれ自身への線形写像) であるとする. もしも V の基底 v_1, \dots, v_n が $f(v_i) = \alpha_i v_i$ ($\alpha_i \in K$) を満たしているならば, 基底 v_i に関する f の行列表示は対角行列 $D = \text{diag}(\alpha_1, \dots, \alpha_n)$ になる. \square

ヒント. $[f(v_1), \dots, f(v_n)] = [v_1, \dots, v_n]D$ を示せばよいので簡単である. \square

[172] (巡回行列とその行列式) $n \times n$ 行列 Λ を次のように定める:

$$\Lambda = \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ & & 0 & \ddots \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{bmatrix} = E_{12} + E_{23} + \cdots + E_{n-1,n} + E_{n,1} \in M_n(\mathbb{C}).$$

ここで E_{ij} は行列単位 (第 (i, j) 成分だけが 1 で他の成分がすべて 0 であるような行列) である. $\zeta = e^{2\pi i/n}$ (1 の原始 n 乗根) とおき,

$$v_k = \begin{bmatrix} 1 \\ \zeta^k \\ \zeta^{2k} \\ \vdots \\ \zeta^{(n-1)k} \end{bmatrix} \in \mathbb{C}^n \quad (k \in \mathbb{Z})$$

とおく. このとき以下が成立する:

1. $\Lambda^k \neq E$ ($k = 1, \dots, n-1$), $\Lambda^n = E$.
2. $\Lambda v_k = \zeta^k v_k$ ($k \in \mathbb{Z}$).

⁹⁸2003 年 10 月 10 日に渡したプリントでは第 1.3 節になっているが正しくは第 1.9 節である.

3. v_0, v_1, \dots, v_{n-1} は \mathbb{C}^n の基底である.
4. 基底 v_0, v_1, \dots, v_{n-1} に関する Λ の定める \mathbb{C}^n の一次変換の行列表示は対角行列 $D = \text{diag}(1, \zeta, \zeta^2, \dots, \zeta^{n-1})$ になる.
5. $P = [v_0, v_1, \dots, v_{n-1}] \in M_n(\mathbb{C})$ とおくと, P は可逆であり, $\Lambda = PDP^{-1}$.
6. $X = x_0E + x_1\Lambda + x_2\Lambda^2 + \dots + x_{n-1}\Lambda^{n-1}$ とおくと,

$$\det X = \prod_{k=0}^{n-1} (x_0 + x_1\zeta^k + x_2\zeta^{2k} + \dots + x_{n-1}\zeta^{(n-1)k}). \quad \square$$

ヒント. 実は上の問題は問題 [90] のヒント 2 の方針をより詳しくしたものである. 3. $|P| \neq 0$ を Vandermonde の行列式の公式を用いて示せばよい. 4. $[\Lambda v_0, \Lambda v_1, \dots, \Lambda v_{n-1}] = [v_0, v_1, \dots, v_{n-1}]D$ を示せばよい. 6. $X = P(x_0E + x_1D + x_2D^2 + \dots + x_{n-1}D^{n-1})P^{-1} = P \text{diag}(x_0 + x_1\zeta^k + x_2\zeta^{2k} + \dots + x_{n-1}\zeta^{(n-1)k})_{k=0}^{n-1} P^{-1}$. \square

[173] (複素数の実行列表示) 複素数体 \mathbb{C} は自然に実数体 \mathbb{R} 上の 2 次元のベクトル空間とみなせ⁹⁹, $1, i$ は \mathbb{C} の \mathbb{R} 上の基底である. $z = x + iy \in \mathbb{C}$ ($x, y \in \mathbb{R}$) に対して, 写像 $\hat{z}: \mathbb{C} \rightarrow \mathbb{C}$ を

$$\hat{z}(w) := zw \quad (w \in \mathbb{C})$$

と定めると, \hat{z} は \mathbb{R} 上の線形写像である. 基底 $1, i$ に関する \hat{z} の行列表示を $A(z) \in M_2(\mathbb{R})$ と書くと,

$$A(z) = A(x + iy) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}. \quad \square$$

ヒント. $[z1, zi] = [1, i]A(z)$ を示せばよいだけなので非常に簡単である. \square

注意 5.18 上の問題 [173] の $A(z)$ は問題 [29] の $A(z)$ に等しい. さらに $\theta \in \mathbb{R}$ のとき $A(e^{i\theta})$ は第 1.1 節の回転行列 $R(\theta)$ に等しい. このように, 今までに登場した特殊な行列の多くは自然に得られる線形写像の行列表示に等しくなる. \square

[174] (Hamilton の四元数の行列表示) $1, i, j, k$ を基底に持つ \mathbb{R} 上のベクトル空間

$$\mathbb{H} = \{a1 + bi + cj + dj \mid a, b, c, d \in \mathbb{R}\}$$

に積を次の規則で定める:

$$\begin{aligned} 1^2 &= 1, & 1i &= i1 = i, & 1j &= j1 = j, & 1k &= k1 = k, \\ i^2 &= j^2 = k^2 = -1, & ij &= -ji = k, & jk &= -ki = i, & ki &= -ik = j. \end{aligned}$$

このとき \mathbb{H} の元を **Hamilton の四元数 (quaternion)** と呼ぶ. $a, b \in \mathbb{R}$ のとき四元数 $a1 + bi \in \mathbb{H}$ と複素数 $a + bi \in \mathbb{C}$ を同一視することにする. $q = a1 + bi + cj + dj \in \mathbb{H}$ ($a, b, c, d \in \mathbb{R}$) と置く. 写像 $\hat{q}: \mathbb{H} \rightarrow \mathbb{H}$ を

$$\hat{q}(r) = qr \quad (r \in \mathbb{H})$$

と定めると, \hat{q} は \mathbb{R} 上の一次変換である. このとき以下が成立する.

⁹⁹「複素平面」という言葉は複素数全体の集合が実数体上 2 次元のベクトル空間をなすことを含意している.

1. \mathbb{R} 上の基底 $1, i, j, k$ に関する \hat{q} の行列表示を $A(q)$ と書くと,

$$A(q) = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}.$$

2. $z = a + bi$, $w = c + di$, $a, b, c, d \in \mathbb{R}$ とすると, $q = z1 + wj$ であるから, \mathbb{H} は $1, j$ を基底に持つ \mathbb{C} 上の 2 次元のベクトル空間とみなされる. \mathbb{C} 上の基底 $1, j$ に関する \hat{q} の行列表示を $B(q)$ と書くと,

$$B(q) = \begin{bmatrix} z & -w \\ \bar{w} & \bar{z} \end{bmatrix}.$$

ここで \bar{z}, \bar{w} はそれぞれ z, w の複素共役である. \square

ヒント. 1. $[q1, qi, qj, qk] = [1, i, j, k]A(q)$ を示せばよい. 2. $[q1, qj] = [1, j]B(q)$ を示せばよい. $zj = j\bar{z}$ を用いよ. \square

参考 5.19 問題 [60] で定義された Pauli 行列 $\sigma_1, \sigma_2, \sigma_3$ と四元数の複素 2×2 行列表現 $B(q)$ のあいだには $B(i) = i\sigma_3$, $B(j) = -i\sigma_2$, $B(k) = -i\sigma_1$ という関係がある. したがって, $\pm i$ 倍と順序の違いを除けば Pauli 行列と四元数 i, j, k の複素 2×2 行列表示は本質的に一致する. \square

[175] 問題 [159] の記号をそのまま用いる. $v_i = x^i$ と置く. 任意に $\lambda \in \mathbb{C}$ を取り, $\mathbb{C}[x]$ の一次変換 e, f, h を

$$e = \partial, \quad h = -2x\partial + \lambda, \quad f = -x^2\partial + \lambda x$$

と定める. このとき以下が成立している:

1. $hv_i = (\lambda - 2i)v_i, \quad ev_i = iv_{i-1}, \quad fv_i = (\lambda - i)v_{i+1}.$
2. 特に $hv_0 = \lambda v_0, \quad ev_0 = 0.$
3. $[h, e] = 2e, \quad [h, f] = -2f, \quad [e, f] = h.$

ここで $[A, B] = AB - BA$ (交換子) である. \square

ヒント. たとえば

$$fv_4 = (-x^2\partial + \lambda x)(x^4) = -x^2(x^4)' + \lambda x \cdot x^4 = -4x^5 + \lambda x^5 = (\lambda - 4)x^5 = (\lambda - 4)v_5.$$

3 の計算は交換子に関する一般的な公式

$$\begin{aligned} [A, A] &= 0, & [B, A] &= -[A, B], \\ [AB, C] &= [A, C]B + A[B, C], & [A, BC] &= [A, B]C + B[A, C] \end{aligned}$$

と $[\partial, x^i] = ix^{i-1}$ を用いて実行せよ. たとえば

$$[\partial, -x^2\partial] = -[\partial, x^2]\partial - x^2[\partial, \partial] = -2x\partial - x^2 \cdot 0 = -2x\partial. \quad \square$$

参考 5.20 (\mathfrak{sl}_2 -triplet) 2×2 行列 E, F, H を

$$E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

と定めると

$$[H, E] = 2e, \quad [H, F] = -2F, \quad [E, F] = H$$

が成立している. E, F, H を \mathfrak{sl}_2 -triplet (\mathfrak{sl}_2 の三つ組) と呼ぶ. 上の問題 [175] の e, f, h は多項式係数の微分作用素による \mathfrak{sl}_2 -triplet の表現になっている.

Lie 代数 $\mathfrak{sl}_2(\mathbb{C})$ の有限次元表現論は 3 次元空間の回転を司る Lie 群 $SU(2)$ の表現論と同値である. Lie 群および Lie 代数の表現論に関する入門的な解説は山内・杉浦 [YmS] にある. \square

[176] 問題 [175] の続き. $\lambda = \ell \in \mathbb{Z}_{\geq 0}$ と仮定する. 以下を示せ:

1. ℓ 次以下の一変数多項式全体のなす $\mathbb{C}[x]$ の部分集合を V_ℓ と書くことにする:

$$V_\ell = \{ a_0 + a_1x + a_2x^2 + \cdots + a_\ell x^\ell \mid a_0, a_1, \dots, a_\ell \in \mathbb{C} \}.$$

このとき V_ℓ は $\mathbb{C}[x]$ の部分空間であり,

$$v_0 = 1, \quad v_1 = x, \quad v_2 = x^2, \quad \dots, \quad v_\ell = x^\ell$$

は V_ℓ の基底をなす.

2. e, f, h の $\mathbb{C}[x]$ への作用は V_ℓ を保つ. すなわち, 任意の $v \in V_\ell$ に対して $ev, fv, hv \in V_\ell$.
3. e, f, h の定める V_ℓ の一次変換の基底 v_i に関する行列表示をそれぞれ E_ℓ, F_ℓ, H_ℓ と書くこと,

$$E_\ell = \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & 2 & \\ & & 0 & \ddots \\ & & & \ddots & \ell \\ 0 & & & & 0 \end{bmatrix}, \quad F_\ell = \begin{bmatrix} 0 & & & 0 \\ \ell & 0 & & \\ & \ell-1 & 0 & \\ & & \ddots & \ddots \\ 0 & & & 1 & 0 \end{bmatrix},$$

$$H_\ell = \begin{bmatrix} \ell & & & 0 \\ & \ell-2 & & \\ & & \ddots & \\ & & & -\ell+2 \\ 0 & & & & -\ell \end{bmatrix} = \text{diag}(\ell, \ell-2, \ell-4, \dots, -\ell+4, -\ell+2, -\ell).$$

たとえば $\ell = 3$ のとき

$$E_3 = \begin{bmatrix} 0 & 1 & & \\ & 0 & 2 & \\ & & 0 & 3 \\ & & & 0 \end{bmatrix}, \quad F_3 = \begin{bmatrix} 0 & & & \\ 3 & 0 & & \\ & 2 & 0 & \\ & & 1 & 0 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 3 & & & \\ & 1 & & \\ & & -1 & \\ & & & -3 \end{bmatrix}. \quad \square$$

注意 5.21 特に $\ell = 1$ のとき $E_1 = E, F_1 = F, H_1 = H$ である. \square

参考 5.22 実は Lie 代数 $\mathfrak{sl}_2(\mathbb{C})$ の (したがってコンパクト Lie 群 $SU(2)$ の) 有限次元既約表現の同型類の全体は表現 V_ℓ ($\ell = 0, 1, 2, \dots$) で代表される¹⁰⁰. この事実は 3 次元空間の回転を量子論的に実現する方法が非負の整数 ℓ で分類されることを意味している.

H_ℓ の固有値 $\ell, \ell - 2, \dots, -\ell + 2, -\ell$ は表現 V_ℓ のウェイト (weight) と呼ばれており, その最高値の ℓ は表現 V_ℓ の最高ウェイト (highest weight) と呼ばれている.

物理学では \mathfrak{sl}_2 の三つ組 E, F, H の代わりに $\sigma_z = \frac{1}{2}H, \sigma_+ = \frac{1}{\sqrt{2}}E, \sigma_- = \frac{1}{\sqrt{2}}F$ の三つ組を用いることが多い. それらは次の交換関係を満たしている:

$$[\sigma_z, \sigma_\pm] = \pm\sigma_\pm, \quad [\sigma_+, \sigma_-] = \sigma_z.$$

だから, H の作用 H_ℓ の固有値のウェイトではなく, σ_z の作用 $\frac{1}{2}H_\ell$ の固有値を用いることが多い. $j = \ell/2$ の方を用を表現 V_ℓ のスピンと呼ぶ¹⁰¹.

以上のコメントに関する詳しい解説については山内・杉浦 [YmS] を参照せよ. \square

[177] 正の整数 $n \in \mathbb{Z} > 0$ と複素数 $\alpha \in \mathbb{C}$ に対して, $(t - \alpha)^k \neq 0$ ($k = 1, \dots, n - 1$), $(t - \alpha)^n = 0$ を満たす文字 t を用意し¹⁰², $1, t, t^2, \dots, t^{n-1}$ を基底に持つ \mathbb{C} 上のベクトル空間 V を次のように定める:

$$V := \{ \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_{n-1} t^{n-1} \mid \beta_0, \beta_1, \beta_2, \dots, \beta_{n-1} \in \mathbb{C} \}.$$

写像 $f: V \rightarrow V$ を $f(v) = tv$ ($v \in V$) と定めると, f は V の \mathbb{C} 上の一次変換 (V からそれ自身への線形写像) である. 以下が成立することを示せ:

1. 基底 $1, t, t^2, \dots, t^{n-1}$ に関する f の行列表示を A と書くと,

$$A = \begin{bmatrix} 0 & & 0 & -a_{n-1} \\ 1 & 0 & & -a_{n-2} \\ & 1 & \ddots & \vdots \\ & & \ddots & 0 & -a_1 \\ 0 & & & 1 & -a_0 \end{bmatrix}.$$

ここで $a_0, a_1, \dots, a_{n-2}, a_{n-1} \in \mathbb{C}$ は $(\lambda - \alpha)^n$ の展開

$$(\lambda - \alpha)^n = \lambda^n + a_0 \lambda^{n-1} + a_1 \lambda^{n-2} + \dots + a_{n-2} \lambda + a_{n-1}$$

によって定められたものである. 二項定理より,

$$a_{i-1} = \binom{n}{i} (-\alpha)^i \quad (i = 1, \dots, n).$$

よって $a_0 = -n\alpha, a_1 = \frac{n(n-1)}{2}\alpha^2, \dots, a_{n-2} = n(-\alpha)^{n-1}, a_{n-1} = (-\alpha)^n$.

¹⁰⁰ しかも e, f, h が微分作用素で表わされたのも偶然ではない. 半単純 Lie 代数 (もしくは半単純 Lie 群) の表現に関する幾何学的な理論 (Borel-Weil-Bott 理論) が存在し, それを用いれば半単純 Lie 代数の有限次元表現の微分作用素による表示が自然に得られる. この辺の問題は Lie 代数および Lie 群の表現論 (representation theory) という大きな理論の一部分を切り取ることによって作成された.

¹⁰¹ 電子や陽子のスピンは $1/2$ である.

¹⁰² 厳密にはそのような文字 t は多項式環 $\mathbb{C}[\lambda]$ の剰余環 $\mathbb{C}[\lambda]/((\lambda - \alpha)^n)$ の λ で代表される元として構成される ($t = \lambda \bmod (\lambda - \alpha)^n$). 剰余環 $\mathbb{C}[\lambda]/((\lambda - \alpha)^n)$ の構成に関しては問題 [187], [188] を参照せよ.

2. V の \mathbb{C} 上の基底として $1, t - \alpha, (t - \alpha)^2, \dots, (t - \alpha)^{n-1}$ も取れる.
 3. 基底 $1, t - \alpha, (t - \alpha)^2, \dots, (t - \alpha)^{n-1}$ に関する f の行列表示を B と書くと,

$$B = \begin{bmatrix} \alpha & & & 0 \\ 1 & \alpha & & \\ & 1 & \alpha & \\ & & \ddots & \ddots \\ 0 & & & 1 & \alpha \end{bmatrix} \quad (n \times n \text{ 行列}). \quad \square$$

ヒント. 1. $(t - \alpha)^n = 0$ を用いて, $[t1, tt, tt^2, \dots, tt^{n-1}] = [1, t, t^2, \dots, t^{n-1}]A$ を示せばよい.

2. $k = 0, 1, \dots, n-1$ とする. $(t - \alpha)^k$ を展開することによって, $(t - \alpha)^k$ は $1, t, \dots, t^k$ の一次結合で書けることがわかる. 逆に $t^k = ((t - \alpha) + \alpha)^k$ を展開することによって, t^k は $1, t - \alpha, \dots, (t - \alpha)^k$ の一次結合で書けることがわかる. このことより, $1, t - \alpha, \dots, (t - \alpha)^{n-1}$ も V の基底であることがわかる.

3. $[t1, t(t - \alpha), t(t - \alpha)^2, \dots, t(t - \alpha)^{n-1}] = [1, t - \alpha, (t - \alpha)^2, \dots, (t - \alpha)^{n-1}]B$ を示せばよい. そのとき $t(t - \alpha)^k = (\alpha + (t - \alpha))(t - \alpha)^k = \alpha(t - \alpha)^k + (t - \alpha)^{k+1}$ と $(t - \alpha)^n = 0$ を用いよ. \square

注意 5.23 (Jordan 標準形の理論との関係) tA は参考 2.3 のコンパニオン行列の形をしている. tB は問題 [28] の Jordan ブロックの形をしている. 実は上の問題 [177] は単因子論を経由する Jordan 標準形の存在証明の一部になっている.

その方針での Jordan 標準形の理論の解説に関しては堀田 [H2] がおすすめである. \square

5.6 商ベクトル空間

K は体であるとし, V は K 上の任意のベクトル空間であるとし, W は V の部分空間であるとする. 任意の $v \in V$ に対して

$$v + W = \{v + w \mid w \in W\}$$

とおき, 集合の集合 V/W を次のように定める:

$$V/W = \{v + W \mid v \in V\}.$$

[178] $v, v' \in V$ に対して, $v + W = v' + W$ と $v' - v \in W$ は同値である. \square

ヒント. $v + W = v' + W$ ならば $v' \in v' + W$ に対してある $w \in W$ で $v' = v + w$ をみたすものが存在する. そのとき $v' - v = w \in W$ である. 逆に $v' - v \in W$ ならば任意の $w \in W$ に対して $v' + w = v + (v' - v) + w \in v + W$ である. よって $v' + W \subset v + W$ である. 逆向きの包含関係も同様にして示されるので $v + W = v' + W$ である. \square

[179] 写像 $+: (V/W) \times (V/W) \rightarrow (V/W)$ と $\cdot: K \times (V/W) \rightarrow (V/W)$ を

$$(u + W) + (v + W) = (u + v) + W, \quad \alpha(u + W) = (\alpha u) + W \quad (u, v \in V, \alpha \in K)$$

と定義することができることを示せ. \square

ヒント. これは well-definedness (うまく定義されること) を示す問題である. 写像がうまく定義されることを示すためには同じものが同じものに移ることを示さなければいけない. そのためには $u + W = u' + W, v + W = v' + W, u, u', v, v' \in V, \alpha \in K$ のとき,

$$(u + v) + W = (u' + v') + W, \quad (\alpha u) + W = (\alpha u') + W$$

となることを示せばよい. \square

[180] 上の問題で定義された演算 $+, \cdot$ に関して V/W は K 上のベクトル空間をなすことを示せ. \square

ヒント. 写像 $- : V/W \rightarrow V/W$ を $-(u + W) = (-u) + W$ ($u \in V$) と定義することができる. さらに, $0_{V/W} = 0 + W = W$ とおき, ベクトル空間の公理を機械的にチェックすればよい. \square

定義 5.24 (商ベクトル空間) 以上のようにして構成された V/W を V を W で割ってできる V の商ベクトル空間 (quotient vector space) もしくは商空間 (quotient space) と呼ぶ. \square

参考 5.25 (商ベクトル空間の元の記号について) V/W の元 $v + W$ は

$$v + W = v \bmod W = [v] = \bar{v}$$

のように書かれることも多い. $v \bmod W$ は「ベクトル v の W の元による平行移動方向の成分を無視したもの」という意味を持ち, $[v]$ や \bar{v} は v で代表される同値類 (equivalence class) によく使われる記号である. \square

参考 5.26 以上の商ベクトル空間の構成はそのまま一般の環 R 上の加群の商加群の構成に一般化される. \square

参考 5.27 (M/N という記号法について) 代数学において加群 (ベクトル空間も加群の一種であることに注意) M とその部分加群 N に対して, M/N は分子の加群 M の中で分母の部分加群 N をゼロにつぶしてできる商加群を意味している. \square

注意 5.28 商ベクトル空間は集合の集合として定義されたが, V/W が集合の集合であることにこだわりすぎると商ベクトル空間の正しい理解に失敗する. 商ベクトル空間 V/W の元は通常のベクトルだと考えた方がよい.

それでは V/W の元はどのようなベクトルだと考えればよいのだろうか. 問題 [178] によれば, $v, v' \in V$ に対応する商ベクトル空間 V/W の元 $v + W, v' + W$ が互いに等しくなるための必要十分条件は $v' - v \in W$ すなわち $v' \in v + W$ である. よって V の中の v を通り W に平行な部分集合 $v + W$ 上のすべてのベクトルが商ベクトル空間 V/W の一点に対応している. つまり, 直観的に V/W は V を W 方向につぶして¹⁰³できるベクトル空間とみなせる. この点に関しては問題 [181], [182] を参考にせよ. \square

¹⁰³ 「つぶす」という言葉を用いると, 紙屑などを「グシャツ」と潰す様子を想像する人が結構いるようである. しかし, 商ベクトル空間 V/W を作るために V を W 方向につぶす場合には「グシャツ」ではなく「スーツ」と滑らかに潰れる様子を想像しなければならない.

[181] \mathbb{R}^3 の部分空間 Z を $Z = \{(0, 0, z) \mid z \in \mathbb{R}\}$ と定める. このとき, \mathbb{R}^3/Z は \mathbb{R} 上の 2 次元のベクトル空間になる. \square

ヒント. $e_1 + Z, e_2 + Z$ が \mathbb{R}^3/Z の基底をなすことを示せ. \square

注意 5.29 \mathbb{R}^3/Z は直観的に 3 次元空間 \mathbb{R}^3 を z 軸方向に潰してできる 2 次元空間だとみなせる. すなわち \mathbb{R}^3 の中の z 軸 Z に平行な直線を一点に潰してできる 2 次元空間が \mathbb{R}^3/Z である. \square

[182] \mathbb{R}^3 の部分空間 W を $W = \{(x, y, 0) \mid x, y \in \mathbb{R}\}$ と定める. このとき, \mathbb{R}^3/W は \mathbb{R} 上の 1 次元のベクトル空間になる. \square

ヒント. $e_3 + W$ が \mathbb{R}^3/W の基底をなすことを示せ. \square

注意 5.30 \mathbb{R}^3/W は直観的に 3 次元空間 \mathbb{R}^3 を xy 平面方向に潰してできる 1 次元空間だとみなせる. すなわち \mathbb{R}^3 の中の xy 平面 W に平行な平面を一点に潰してできる 1 次元空間が \mathbb{R}^3/W である. \square

[183] (自然な射影) 写像 $p: V \rightarrow V/W$ を

$$p(v) = v + W \quad (v \in V)$$

と定めると, p は K 上の線形写像でかつ全射である. p は V から商空間 V/W への自然な射影 (canonical projection) もしくは自然な写像 (canonical mapping) と呼ばれる. \square

[184] (準同型定理) U, V は体 K 上のベクトル空間であり, $f: U \rightarrow V$ は線形写像であるとする. f の核 (kernel) $\text{Ker } f$ と像 (image) $\text{Im } f$ を

$$\text{Ker } f = \{u \in U \mid f(u) = 0\}, \quad \text{Im } f = \{f(u) \mid u \in U\}$$

と定めると, $\text{Ker } f$ は U の部分空間であり, $\text{Im } f$ は V の部分空間である. 写像 $\phi: U/\text{Ker } f \rightarrow \text{Im } f$ を

$$\phi(u + \text{Ker } f) = f(u) \quad (u \in U)$$

と定義することができ (すなわち $u, u' \in U$ に対して $u + \text{Ker } f = u' + \text{Ker } f$ ならば $f(u) = f(u')$), ϕ は K 上のベクトル空間の同型写像になる. \square

ヒント. 記号の簡単のため $\bar{u} = u + \text{Ker } f$ ($u \in U$) とおく.

ϕ の well-definedness: $u, u' \in U, \bar{u} = \bar{u}'$ と仮定する. そのとき $u - u' \in \text{Ker } f$ である. よって $f(u) - f(u') = f(u - u') = 0$ すなわち $f(u) = f(u')$ である.

ϕ の線形性: $u, u' \in U, \alpha \in K$ に対して, $\phi(\bar{u} + \bar{u}') = \phi(\overline{u + u'}) = f(u + u') = f(u) + f(u') = \phi(\bar{u}) + \phi(\bar{u}')$, $\phi(\alpha \bar{u}) = \phi(\overline{\alpha u}) = f(\alpha u) = \alpha f(u) = \alpha \phi(\bar{u})$.

ϕ の単射性: $u \in U, \phi(\bar{u}) = 0$ と仮定する. $0 = \phi(\bar{u}) = f(u)$ より $u \in \text{Ker } f$ である. よって $\bar{u} = 0$.

ϕ の全射性: $\text{Im } \phi = \{\phi(\bar{u}) \mid u \in U\} = \text{Im } f$. \square

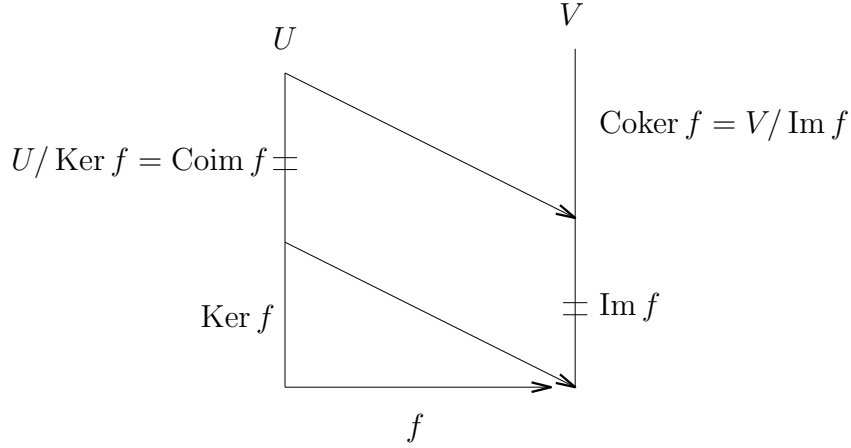


図 5.1: 準同型定理

参考 5.31 $f : U \rightarrow V$ の余核 (cokernel) $\text{Coker } f$ と余像 (coimage) $\text{Coim } f$ が

$$\text{Coker } f = V / \text{Im } f, \quad \text{Coim } f = U / \text{Ker } f$$

と定義される. 準同型定理は余像と像が自然に同型になることを意味している. このことをよく図 5.1 のように描く. \square

参考 5.32 準同型定理は一般の環 R 上の加群にそのまま一般化される. 証明の仕方はベクトル空間の場合とまったく同じである. \square

[185] U は体 K 上のベクトル空間であり, V はその部分空間であるとし, W は U における V の補空間であるとする. このとき自然な射影 $p : U \twoheadrightarrow U/V$ の W への制限 $p|_W : W \rightarrow U/V$ は同型写像になる. よって $(W \oplus V)/V \cong W$ という自然な同型を得る. \square

ヒント. $p|_W$ が単射であることと全射であることを補空間の定義に戻って地道に証明せよ. もしくは写像 $q : U/V \rightarrow W$ を $q((w + v) \bmod V) = w$ ($w \in W, v \in V$) と定めることができ (well-definedness のチェックが必要), q が $p|_W$ の逆写像になることを示せ. \square

[186] (image と kernel の次元の関係) U, V は体 K 上のベクトル空間であり, U は有限次元であると仮定する. このとき任意の線形写像 $f : U \rightarrow V$ に対して $\dim \text{Ker } f + \dim \text{Im } f = \dim_K U$. \square

ヒント. 問題 [128] の一般化. 証明の方針はほとんど同じでよい. もしくは準同型定理と問題 [185] の結果を使えばより簡単に証明できる. \square

[187] 体 K 上の一変数多項式環 $K[\lambda]$ を考え, 任意にゼロでない多項式 $f \in K[\lambda]$ を取る. このとき, $K[\lambda]$ の部分集合 (f) を

$$(f) = K[\lambda]f = \{af \mid a \in K[\lambda]\}$$

と定める¹⁰⁴. 以下を示せ.

¹⁰⁴ (f) は f から生成される $K[\lambda]$ の単項イデアル (principal ideal) と呼ばれる.

1. (f) は $K[\lambda]$ の $K[\lambda]$ 部分加群である. すなわち任意の $g, h \in (f)$ と $a \in K[\lambda]$ に対して $g + h \in (f)$ かつ $af \in (f)$ である. 特に (f) は $K[\lambda]$ の K 上のベクトル部分空間である.
2. $R = K[\lambda]/(f)$ (商ベクトル空間) とおき, $a \in K[\lambda]$ に対する $a + (f) \in R$ を $a \bmod f$ と書くことにする. このとき, 積 $\cdot : R \times R \rightarrow R$ を

$$(a \bmod f) \cdot (b \bmod f) = ab \bmod f \quad (a, b \in K[\lambda])$$

と定めることができる (すなわち $a, b, c, d \in K[\lambda]$ に対して $a \bmod f = c \bmod f$, $b \bmod f = d \bmod f$ ならば $ab \bmod f = cd \bmod f$ が成立する).

3. これによって R は可換環をなす¹⁰⁵. \square

ヒント. 1. $a, b, c \in K[\lambda]$ に対して $af + bf = (a + b)f \in (f)$ であり, $a(bf) = (ab)f \in (f)$.
 2. $a \bmod f = c \bmod f$ と $a - c \in (f)$ は同値であり, $b \bmod f = d \bmod f$ と $b - d \in (f)$ は同値であるから, $ab - cd = ab - ad + ad - cd = a(b - d) + d(a - c) \in (f)$. 3. $1_R = 1 \bmod f$ と置き, 可換環の公理を機械的にチェックすればよい. \square

参考 5.33 $R = K[\lambda]/(f)$ は $K[\lambda]$ の中で f をゼロとみなすことによって得られる可換環である. f がゼロとみなされるならば任意の $a \in K[\lambda]$ に対する af もゼロとみなされなければいけない. (f) はそのような af 全体のなす集合である.

本当は上の問題は可換環とイデアルと剰余環の理論としてより一般的にやるべき事柄である. \square

[188] 上の問題 [187] のつづき. f の次数が n ならば $\dim_K R = \dim_K(K[\lambda]/(f)) = n$ であることを証明せよ. \square

ヒント 1. $t = \lambda \bmod f$ と置くと, $t^i = \lambda^i \bmod f$ である. $1, t, t^2, \dots, t^{n-1}$ が R の K 上の基底になることを示せばよい.

任意の $g \in K[\lambda]$ は g を f で割ることによって $g = qf + r$, $q, r \in K[\lambda]$, $\deg r < n$ と一意に表わされる¹⁰⁶ (商が q で余りが r). そのとき $g \bmod f = r \bmod f$ であり, r は次数が n 未満なので $1, \lambda, \dots, \lambda^{n-1}$ の一次結合で表わされるので, $g \bmod f$ は $1, t, \dots, t^{n-1}$ の一次結合で表わされる.

もしも $g \in K[\lambda]$, $\deg g < n$ かつ $g \bmod f = 0_R = (f)$ ならば $g = af$, $a \in K[\lambda]$ と表わされる. $\deg g < n$ より $a = 0$ でなければいけないので $g = 0$ となる. これより $1, t, \dots, t^{n-1}$ の一次独立性が出る. \square

ヒント 2. 次数が n 未満の λ の多項式全体のなす n 次元のベクトル空間を V と書き, 線形写像 $\phi : V \rightarrow R$ を $\phi(v) = v \bmod f$ ($v \in V$) と定義する. ϕ が同型写像であることを示せば R の次元も n であることがわかる.

任意の $g \in K[\lambda]$ は g を f で割ることによって $g = qf + r$, $q, r \in K[\lambda]$, $\deg r < n$ と一意に表わされるので, $g \bmod f = r \bmod f = \phi(r)$ である. よって ϕ は全射である.

もしも $g \in V$ かつ $\phi(g) = g \bmod f = 0_R = (f)$ ならば $g = af$, $a \in K[\lambda]$ と表わされる. $\deg g < n$ より $a = 0$ でなければいけないので $g = 0$ となる. よって ϕ は単射である. \square

¹⁰⁵ $R = K[\lambda]/(f)$ は $K[\lambda]$ をイデアル (f) で割ってできる剰余環 (residue ring, residue-class ring) と呼ばれる.

¹⁰⁶ $\deg r$ は r の次数である. $r = 0$ のとき $\deg r = -\infty$ と考える.

5.7 双対空間

[189] (双対空間の定義) 体 K 上のベクトル空間 V に対して V から K への線形写像全体のなす集合 V^* は自然に体 K 上のベクトル空間をなすことを示せ. V^* は V の双対ベクトル空間 (dual vector space) もしくは双対空間 (dual space) と呼ばれる. $f \in V^*$ と $v \in V$ に対して $f(v)$ を $\langle f, v \rangle$ と表わすことがある. \square

ヒント. 問題 [155] の特別な場合. $V^* = \text{Hom}_K(V, K)$. \square

[190] (横ベクトルの空間と縦ベクトルの空間の双対性) K は体であるとする. K の元を成分に持つ n 次元縦ベクトル全体のなすベクトル空間を K^n と書き, n 次元横ベクトル全体のなすベクトル空間を仮に ${}^t(K^n)$ と書くことにする. 写像 $\iota: {}^t(K^n) \rightarrow (K^n)^*$ を横ベクトルと縦ベクトルの積によって

$$\left\langle \iota([x_1, \dots, x_n]), \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle := [x_1, \dots, x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n x_i y_i \quad (x_i, y_i \in K)$$

と定義する. このとき ι は同型写像になることを示せ. ι を通して横ベクトルの空間 ${}^t(K^n)$ と縦ベクトルの空間 K^n の双対空間 $(K^n)^*$ は自然に同一視される. \square

参考 5.34 (ブラとケット) 量子力学には, ブラベクトル (bra vector) $\langle v^*|$ やケットベクトル (ket vector) $|v\rangle$ のような記号が登場¹⁰⁷, ブラ $\langle v^*|$ とケット $|v\rangle$ のあいだには $\langle v^*|v\rangle \in \mathbb{C}$ と書かれる内積が定義されている.

実はブラベクトル全体のなすベクトル空間はケットベクトル全体のなすベクトル空間の双対空間と同一視できる. 直観的にブラベクトルは横ベクトルのようなものであり, ケットベクトルは縦ベクトルのようなものだと考えればよい. 横ベクトルと縦ベクトルのあいだには上の問題のように自然に内積が定義される. \square

[191] (基底の定める座標) V は体 K 上の有限次元ベクトル空間であり, v_1, \dots, v_n は V の基底であるとする. 任意の $v \in V$ は $v = \alpha_1 v_1 + \dots + \alpha_n v_n$ ($\alpha_i \in K$) と一意に表わされる. よって v に対して α_i を対応させる写像 x_i が定まる. $x_i \in V^*$ であることを示せ. (x_i を基底 v_i の定める V 上の座標と呼ぶことにする.) \square

[192] (双対基底) V は体 K 上の有限次元ベクトル空間であるとする. V の基底 v_1, \dots, v_n に対して, $v_1^*, \dots, v_n^* \in V^*$ を

$$\langle v_i^*, v_j \rangle = \delta_{ij} \quad (i, j = 1, \dots, n)$$

という条件によって一意に定めることができる. このとき v_1^*, \dots, v_n^* は双対空間 V^* の基底になる. 特に $\dim V^* = \dim V$ である. v_1^*, \dots, v_n^* を v_1, \dots, v_n の双対基底 (dual basis) と呼ぶ. \square

ヒント. 任意の $f \in V^*$ に対して, $g = \sum_{j=1}^n \langle f, v_j \rangle v_j^* \in V^*$ と置くと, $\langle g, v_j \rangle = \langle f, v_j \rangle$ ($j = 1, \dots, n$) であるから, $f = g$ であることがわかる. よって V^* は v_1^*, \dots, v_n^* で張られる. $v^* := \sum_{i=1}^n \alpha_i v_i^* = 0$, $\alpha_i \in K$ と仮定する. このとき $0 = \langle v^*, v_j \rangle = \alpha_j$ ($j = 1, \dots, n$) である. よって v_1^*, \dots, v_n^* は一次独立である. \square

¹⁰⁷Dirac [D] などの量子力学の教科書を参照せよ.

注意 5.35 問題 [191] の x_i と問題 [192] の v_i^* は等しい. \square

[193] (1 の分解) V は体 K 上の有限次元ベクトル空間であるとする. V の基底 v_1, \dots, v_n とその双対基底 $v_1^*, \dots, v_n^* \in V^*$ を任意に取る. V の一次変換 $\sum_{i=1}^n v_i v_i^*$ を次のように定める:

$$\left(\sum_{i=1}^n v_i v_i^* \right) (v) = \sum_{i=1}^n v_i \langle v_i^*, v \rangle \quad (v \in V).$$

この $\sum_{i=1}^n v_i v_i^*$ は V の恒等写像 id_V に等しい. $\text{id}_V = \sum_{i=1}^n v_i v_i^*$ を 1 の分割と呼ぶ. \square

ヒント. $v \in V$ を $v = \sum_{j=1}^n \alpha_j v_j$, $\alpha_j \in K$ と表わし, $\sum_{i=1}^n v_i \langle v_i^*, v \rangle$ を計算してみよ. \square

注意 5.36 $V = K^n$, $v_i = e_i$ ならば $v_i^* = {}^t e_i$ である. $\sum_{i=1}^n e_i {}^t e_i$ が単位行列になることは容易に示される. 上の問題の結果はこれの一般化である. \square

参考 5.37 量子力学では¹⁰⁸, 1 の分割をブラとケットの記号を用いて $1 = \sum_i |i\rangle \langle i|$ のように書くことが多い. $|i\rangle$ はケットベクトル全体の空間の基底であり, $\langle i|$ はその双対基底である. \square

[194] (双対の双対) V は体 K 上の有限次元ベクトル空間であるとする. このとき, 写像 $\iota: V \rightarrow (V^*)^*$ を

$$\langle \iota(v), f \rangle = \iota(v)(f) := \langle f, v \rangle = f(v) \quad (v \in V, f \in V^*)$$

と定めると, ι は同型写像である. $\iota: V \xrightarrow{\sim} (V^*)^*$ を通して, $(V^*)^*$ は V と自然に同一視される. \square

[195] (転置写像) $f: U \rightarrow V$ は体 K 上のベクトル空間のあいだの線形写像であるとする. このとき線形写像 ${}^t f: V^* \rightarrow U^*$ を

$$\langle {}^t f(v^*), u \rangle = \langle v^*, f(u) \rangle \quad (v^* \in V^*, u \in U)$$

と定義できることを示せ. ${}^t f$ を f の転置写像と呼ぶことにする. \square

[196] (行列の転置との関係) K は体であるとし, K の元を成分に持つ n 次元縦ベクトル全体の空間を K^n と表わし, 写像 $\iota: K^n \rightarrow (K^n)^*$ を

$$\langle \iota(x), y \rangle = \iota(x)(y) := {}^t x y = \sum_{i=1}^n x_i y_i \quad (x = [x_i], y = [y_i] \in K^n)$$

と定めると, ι は同型写像である. ι を用いて $(K^n)^*$ と K^n 自身を同一視することにする. そのとき, 任意に $A \in M_{m,n}(K)$ を取ると, A の定める K^n から K^m への線形写像の転置写像が ${}^t A$ の定める K^m から K^n への線形写像になることを示せ. \square

ヒント. $x, y \in K^n$ を任意に取る. $\langle \iota({}^t A x), y \rangle = \langle \iota(x), A y \rangle$ を示せばよい. \square

¹⁰⁸Dirac [D] などを見よ.

[197] (商空間と部分空間の双対) U は体 K 上のベクトル空間であり, V はその部分空間であるとし,

$$V^\perp = \{u^* \in U^* \mid \langle u^*, v \rangle = 0 \ (v \in V)\}$$

とおく¹⁰⁹. V から U への包含写像を i と書き¹¹⁰, U から U/V への自然な射影を p と書くことにする:

$$V \xrightarrow{i} U \xrightarrow{p} U/V.$$

双対空間の移ると次のような転置写像の列ができる:

$$V^* \xleftarrow{^ti} U^* \xleftarrow{^tp} (U/V)^*.$$

以下を示せ:

1. $^tp: (U/V)^* \rightarrow U^*$ は単射である.
2. $\text{Ker } ^ti = V^\perp$.
3. $\text{Ker } ^ti = \text{Im } ^tp$ である.
4. tp は自然な同型 $(U/V)^* \xrightarrow{\sim} V^\perp$, $x^* \mapsto ^tp(x^*)$ を誘導する.
5. $^ti: U^* \rightarrow V^*$ は全射である.
6. ti は自然な同型 $U^*/V^\perp \xrightarrow{\sim} V^*$, $u^* \bmod V^\perp \mapsto ^ti(u^*)$ を誘導する. \square

ヒント. 1. 任意の $u \in U$, $x^* \in (U/V)^*$ に対して, $\langle ^tp(x^*), u \rangle = \langle x^*, u \bmod V \rangle$ であるから, $^tp(x^*) = 0$ ならば $x^* = 0$ である. よって $\text{Ker } ^tp = 0$ である. これで tp は単射であることが示された.

2. 任意の $u^* \in U$, $v \in V$ に対して, $\langle ^ti(u^*), v \rangle = \langle u^*, v \rangle$ であるから, $^ti(u^*) = 0$ と $\langle u^*, v \rangle = 0 \ (v \in V)$ は同値である. これで $\text{Ker } ^ti = V^\perp$ が示された.

3. 任意の $x^* \in (U/V)^*$, $v \in V$ に対して, $\langle ^ti(^tp(x^*)), v \rangle = \langle ^tp(x^*), i(v) \rangle = \langle ^tp(x^*), v \rangle = \langle x^*, p(v) \rangle = \langle x^*, 0 \rangle = 0$ であるから, $\text{Im } ^tp \subset \text{Ker } ^ti$ である. 任意の $u^* \in \text{Ker } ^ti$, $v \in V$ に対して, $0 = \langle ^ti(u^*), v \rangle = \langle u^*, v \rangle$ であるから, $x^* \in (U/V)^*$ を $\langle x^*, u \bmod V \rangle = \langle u^*, u \rangle \ (u \in U)$ と定めることができる. そのとき $^tp(x^*) = u^*$ であるから, $\text{Ker } ^ti \subset \text{Im } ^tp$ である.

4. $^tp: (U/V)^* \rightarrow U^*$ は単射であるから, 同型 $(U/V)^* \xrightarrow{\sim} \text{Im } ^tp = \text{Ker } ^ti = V^\perp$ を誘導する.

5. V の U における補空間 W が存在する (問題 [161] の結果). $v^* \in V^*$ に対して $u^* \in U^*$ を $\langle u^*, v + w \rangle = \langle v^*, v \rangle \ (v \in V, w \in W)$ と定めると, $^ti(u^*) = v^*$ である. よって ti は全射である.

6. 準同型定理を ti に適用すると, 2, 5 より同型 $U^*/V^\perp \xrightarrow{\sim} V^*$, $u^* \bmod V^\perp \mapsto ^ti(u^*)$ が得られる. \square

¹⁰⁹ V^\perp は V の U^* における直交補空間 (orthogonal complement) と呼ばれる. この用語法は計量ベクトル空間における直交補空間の概念を双対空間の場合に一般化したものである.

¹¹⁰ i は $v \in V$ を $v \in U$ に対応させる写像である.

6 計量ベクトル空間

この節では複素ベクトル空間を扱う. 複素行列 A の転置の複素共役を A^* と書くことにする. すなわち

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in M_{m,n}(\mathbb{C})$$

のとき

$$A^* = \overline{A^t} = {}^t \overline{A} = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & \overline{a_{m2}} \\ \vdots & & \vdots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{mn}} \end{bmatrix} \in M_{n,m}(\mathbb{C}).$$

ここで $\overline{a_{ij}}$ は a_{ij} の複素共役である¹¹¹.

6.1 計量ベクトル空間の定義

V は複素ベクトル空間 (\mathbb{C} 上のベクトル空間) であるとする. 写像 $(,) : V \times V \rightarrow \mathbb{C}$ が以下の条件を満たしているとき, $(,)$ を計量 (metric) もしくは内積 (inner product) と呼ぶ: $u, u_1, u_2, v, v_1, v_2 \in V$ と $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ に対して,

$$\text{M1. } (v, u) = \overline{(u, v)}.$$

$$\text{M2. } (u, \beta_1 v_1 + \beta_2 v_2) = \beta_1 (u, v_1) + \beta_2 (u, v_2).$$

$$\text{M3. } (\alpha_1 u_1 + \alpha_2 u_2, v) = \overline{\alpha_1} (u_1, v) + \overline{\alpha_2} (u_2, v).$$

$$\text{M4. } (u, u) \geq 0.$$

$$\text{M5. } (u, u) = 0 \text{ が成立するための必要十分条件は } u = 0 \text{ である}^{112}.$$

V と計量 $(,)$ の組を (複素) 計量ベクトル空間 ((complex) metric vector space) もしくは前ヒルベルト空間¹¹³ (pre-Hilbert space) と呼ぶ.

注意 6.1 条件 M3 は M1 と M2 から導かれる. よって, 与えられた $(,)$ が計量であることを確かめるためには条件 M3 以外の条件が成立していることを示せばよい. \square

V が計量ベクトル空間であるとき, $u, v \in V$ が互いに直交するとは $(u, v) = 0$ が成立することである. そのとき $u \perp v$ と書く.

[198] (部分空間) V は計量ベクトル空間であり, W はその部分空間であるとする. そのとき V の計量の W 上への制限は W における計量をなす. これによって W も計量ベクトル空間とみなせる. \square

¹¹¹物理学では複素数 z の複素共役を z^* と書くことがある. その記号法は z を 1×1 行列だとみなせば上の記号法の特珠な場合だとみなせる.

¹¹² $u = 0$ ならば条件 M2 より $(u, u) = 0$ となるので, この条件を確かめるためには M2 と $(u, u) = 0$ ならば $u = 0$ を示せば十分である.

¹¹³Cauchy 列が常に収束するような空間は完備であると呼ばれる. 完備な前ヒルベクト空間をヒルベルト空間 (Hilbert space) と呼ぶ. 有限次元の前ヒルベルト空間は常に完備なのでヒルベルト空間である. しかし無限次元の前ヒルベルト空間は完備とは限らないので, 完備化などの操作によってヒルベルト空間を構成しなければならない.

[199] $V = \mathbb{C}^n$ (縦ベクトルの空間) のとき, 写像 $(,) : V \times V \rightarrow \mathbb{C}$ を

$$(u, v) = u^* v = \sum_{i=1}^n \overline{u_i} v_i \quad (u = [u_i], v = [v_i] \in \mathbb{C}^n)$$

と定めると, \mathbb{C}^n と $(,)$ の組は計量ベクトル空間である. この $(,)$ を \mathbb{C}^n の標準的な計量もしくは標準的な内積と呼ぶことにする. \square

[200] 問題 [150] と同様にして, 閉区間 $[a, b]$ 上の複素数値連続関数全体の集合 $C([a, b], \mathbb{C})$ は自然に \mathbb{C} 上のベクトル空間とみなされる. $f, g \in C([a, b], \mathbb{C})$ に対して, $(f, g) \in \mathbb{C}$ を

$$(f, g) = \int_a^b \overline{f(x)} g(x) dx$$

と定めると, $C([a, b], \mathbb{C})$ と $(,)$ の組は前 Hilbert 空間 (計量ベクトル空間) である. \square

ヒント. 条件 M5 を示すためには $f \in C([a, b], \mathbb{C})$ の連続性を使わなければいけない. 条件 M5 を示すためには $f \neq 0$ ならば $(f, f) \neq 0$ であることを示せばよい. $f \neq 0$ と仮定する. そのとき, ある $x_0 \in [a, b]$ で $A := |f(x_0)| > 0$ となるものが存在する. f は連続なので, x_0 を含み $[a, b]$ に含まれる長さが正の閉区間 $[a_0, b_0]$ が存在して, $x \in [a_0, b_0]$ ならば $|f(x)| \geq A/2$ となる¹¹⁴. したがって

$$(f, f) = \int_a^b |f(x)|^2 dx \geq \int_{a_0}^{b_0} \left(\frac{A}{2}\right)^2 dx = \frac{A^2(b_0 - a_0)}{4} > 0. \quad \square$$

参考 6.2 前 Hilbert 空間 $C([a, b], \mathbb{C})$ の計量 $(,)$ に関する完備化を $L_2([a, b])$ のように表わし, L_2 空間と呼ぶ. L_2 空間は無限次元の Hilbert 空間の典型的な例である.

積分によって定義された内積も Cauchy-Schwarz の不等式を満たしていることが証明される. 実は問題 [201] を見ればわかるように, 任意の計量ベクトル空間において Cauchy-Schwarz の不等式が成立している.

参考 5.11 で注意しておいたように, 数ベクトルの理論を抽象ベクトル空間に一般化しておく, ある種の関数全体の空間やそのあいだの微分作用素や積分作用素も扱えるようになる, というメリットがあるのであった. 抽象ベクトル空間のレベルで内積 (計量) を定義することによって, ベクトルとベクトルの内積だけではなく, 問題 [200] のように関数と関数の内積を考えることができるようになるのである. これも抽象化のメリットの一つである. \square

6.2 Cauchy-Schwarz の不等式と三角不等式

[201] (Cauchy-Schwarz の不等式) V と $(,)$ の組は複素計量ベクトル空間であるとする. ノルム (norm) $\| \cdot \|$ を

$$\|v\| = \sqrt{(v, v)} \quad (v \in V)$$

¹¹⁴連続関数の ε - δ による定義に戻ってこの主張を確かめてみよ. $\varepsilon = A/2$ としてみよ. さらに, $|f|$ のグラフの図を描いて, この主張が直観的にも正しそうなことを確かめよ. 論理と直観の両方をきたえながら先に進むのがよい.

と定める. このとき, 任意の $u, v \in V$ に対して,

$$|(u, v)| \leq \|u\| \cdot \|v\|$$

が成立し, 等号が成立するための必要十分条件は u, v が一次従属になることである¹¹⁵. この不等式を **Cauchy-Schwarz の不等式 (the Cauchy-Schwarz inequality)** と呼ぶ. \square

ヒント. $u \neq 0$ と仮定してよい. $w = v - \alpha u$ が u と直交するように $\alpha \in \mathbb{C}$ を定め, v のノルムの2乗を u と w で表わしてみよ. \square

[201] の略解. $u \neq 0$ と仮定してよい. そのとき, $\alpha = (u, v)/(u, u)$, $w = v - \alpha u$ とおくと, $(u, w) = (u, v) - \alpha(u, u) = (u, v) - (u, v) = 0$ である. よって

$$\|v\|^2 = (w + \alpha u, w + \alpha u) = (w, w) + |\alpha|^2(u, u) \geq |\alpha|^2(u, u) = \frac{|(u, v)|^2}{\|u\|^2}. \quad (*)$$

ここで, 2つ目の等号で u と w が直交することを用いた. よって $|(u, v)| \leq \|u\| \cdot \|v\|$ が成立する. さらに途中の不等号で等号が成立するための必要十分条件は $w = 0$ すなわち $v = \alpha u$ であり, そのとき u, v は一次従属になる. \square

[202] (ノルムの基本性質) V が計量ベクトル空間であるとき, $u, v \in V$ と $\alpha \in \mathbb{C}$ に対して以下が成立している:

N1. $\|u\| \geq 0$ であつ等号が成立するための必要十分条件は $u = 0$ である.

N2. $\|\alpha u\| = |\alpha| \cdot \|u\|$.

N3. $\|u + v\| \leq \|u\| + \|v\|$ (三角不等式). \square

ヒント. 三角不等式の証明には Cauchy-Schwarz の不等式と $\operatorname{Re}(u, v) \leq |(u, v)|$ を使う. \square

注意 6.3 計量ベクトル空間のノルムに関しては次も成立している. 三角不等式において等号が成立するための必要十分条件はある非負の実数 α で $v = \alpha u$ または $u = \alpha v$ をみたすものが存在すること (直観的には u, v が同じ方向を向いていること) である. 余裕があればこの事実も示してみよ. \square

[202] の略解. 三角不等式のみを証明しよう. $u, v \in V$ に対して,

$$(\|u\| + \|v\|)^2 - \|u + v\|^2 = 2(\|u\| \cdot \|v\| - \operatorname{Re}(u, v)) \leq 2(\|u\| \cdot \|v\| - |(u, v)|) \leq 0.$$

1つ目の不等号は $\operatorname{Re}(u, v) \leq |(u, v)|$ より. 2つ目の不等号は Cauchy-Schwarz の不等式より. $u \neq 0$ のとき, それら2つの不等号が共に等号になるための必要十分条件はある $\alpha \in \mathbb{C}$ が存在して $v = \alpha u$ または $u = \alpha v$ であつ $\operatorname{Re} \alpha = |\alpha|$ (すなわち α は非負の実数) が成立することである. \square

参考 6.4 V が一般の複素ベクトル空間であるとき, 上の問題の条件 N1–3 を満たす V 上の実数値関数 $\|\cdot\|$ を V におけるノルム (norm) と呼び, V とノルムの組をノルム空間 (normed space, normed vector space) と呼ぶ¹¹⁶. \square

¹¹⁵ u, v が一次従属ならば等号が成立することは明らかなので等号が成立するならば u, v が一次従属になることを示せば十分である.

¹¹⁶ 完備なノルム空間はバナッハ空間 (Banach space) と呼ばれる. 実は大学一・二年生で習う \mathbb{R} もしくは \mathbb{R}^n 上の解析学の多くの結果が Banach 空間の場合に一般化される. 実際, 大学一・二年生で習う解析学の結果のほとんどは三角不等式と扱っている空間の完備性を用いて導き出されている.

6.3 正規直交基底

[203] (正規直交基底) V は n 次元計量ベクトル空間であるとする. $u_1, \dots, u_k \in V$ が

$$(u_i, u_j) = \delta_{ij} \quad (i, j = 1, \dots, k)$$

を満たしていれば, u_1, \dots, u_k は一次独立である. (ここで δ_{kl} は Kronecker のデルタである.) このような u_1, \dots, u_k は正規直交系 (orthonormal system) と呼ばれる. 特に $k = n$ のとき u_1, \dots, u_n は V の基底をなす. このような V の基底を V の正規直交基底 (orthonormal basis) と呼ぶ. \square

ヒント. $\alpha_j \in \mathbb{C}$, $v := \sum_{j=1}^k \alpha_j u_j = 0$ とすると $0 = (u_i, v) = \alpha_i$ である. \square

例 6.5 たとえば \mathbb{C}^n の標準的基底 e_1, \dots, e_n は標準的な計量に関する正規直交基底である. もちろん他にも \mathbb{C}^n の正規直交基底はたくさん存在する. \square

[204] (ユニタリ行列) 縦ベクトル $u_1, \dots, u_n \in \mathbb{C}^n$ に対して, $n \times n$ 行列 U を $U = [u_1, \dots, u_n] \in M_n(\mathbb{C})$ と定める. このとき以下の二つの条件は互いに同値である:

- (a) u_1, \dots, u_n は \mathbb{C}^n の正規直交基底である.
- (b) U は可逆であり, $U^* = U^{-1}$ である.

一般に後者の (b) の条件を満たす行列 $U \in M_n(\mathbb{C})$ をユニタリ行列 (unitary matrix) と呼ぶ. \square

[205] (Schmidt の正規直交化法) V は計量ベクトル空間であるとし, 以下を示せ:

1. $u_1, \dots, u_k \in V$ は正規直交系であるとする. そのとき, $v \in V$ に対して \tilde{u} を

$$\tilde{u} = v - \sum_{i=1}^k (u_i, v) u_i$$

と定めると, \tilde{u} は u_1, \dots, u_k と直交する.

2. $v_1, \dots, v_{k+1} \in V$ は一次独立であり, $u_1, \dots, u_k \in V$ は正規直交系であり, u_1, \dots, u_k の張る V の部分空間は v_1, \dots, v_k の張る V の部分空間に等しいと仮定する. $\tilde{u}_{k+1} \in V$ を

$$\tilde{u}_{k+1} = v_{k+1} - \sum_{i=1}^k (u_i, v_{k+1}) u_i \quad (1)$$

と定めると, $\tilde{u}_{k+1} \neq 0$ であるから, $u_{k+1} \in V$ を

$$u_{k+1} = \tilde{u}_{k+1} / \|\tilde{u}_{k+1}\| \quad (2)$$

と定めることができる. このとき, $u_1, \dots, u_{k+1} \in V$ は正規直交系であり, u_1, \dots, u_{k+1} の張る V の部分空間は v_1, \dots, v_{k+1} の張る V の部分空間に等しくなる.

この結果を用いると, $v_1, \dots, v_n \in V$ が一次独立なとき, (1), (2) を帰納的に用いて正規直交系 u_1, \dots, u_n で $k = 1, \dots, n$ に対して u_1, \dots, u_k の張る V の部分空間が v_1, \dots, v_k の張る V の部分空間に等しくなるものが構成できる. この手続きを Schmidt の正規直交化¹¹⁷ (orthonormalization of Schmidt) と呼ぶ. \square

¹¹⁷Gram-Schmidt の正規直交化と呼ばれることも多い.

注意 6.6 参考 5.34 で簡単に説明したブラケットの記号法¹¹⁸を使うと Schmidt の正規直交化法を以下のように書くことができる. $v \in V$ を $|v\rangle$ と書くことにする. $v \in V$ に $(u_i, v) \in \mathbb{C}$ を対応させる関数は V の双対空間の元になるので, $\langle u_i|$ と書くことにする. $\langle u_i|$ と $|v\rangle$ の積 $\langle u_i|v\rangle$ を $\langle u_i|v\rangle = (u_i, v)$ と定める. このとき, 上の問題の小問 1 の式は次のように書き直される:

$$|\tilde{u}\rangle = |v\rangle - \sum_{i=1}^k |u_i\rangle \langle u_i|v\rangle = \left(1 - \sum_{i=1}^k |u_i\rangle \langle u_i|\right) |v\rangle.$$

問題 [193] と同様に $P := \sum_{i=1}^k |u_i\rangle \langle u_i|$ は V の一次変換とみなされる. u_1, \dots, u_k で張られる V の部分空間を W と書くことにすると, この P は W の上への直交射影であり¹¹⁹, $1 - \sum_{i=1}^k |u_i\rangle \langle u_i| = 1 - P$ は W の直交補空間¹²⁰ W^\perp の上への直交射影である. 直交射影の概念を理解すれば Schmidt の正規直交化法が幾何的に何をやっているかがわかる. 一般に任意の数学的構成は, 式のレベルだけではなく, 幾何的直観のレベルでも理解しておくことが望ましい. \square

[206] \mathbb{C}^3 の基底 $v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $v_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $v_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ から, Schmidt の正規直交化によって \mathbb{C}^3 の正規直交基底 u_1, u_2, u_3 を構成せよ. \square

[206] の略解. $u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $u_2 = \sqrt{\frac{2}{3}} \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix}$, $u_3 = \frac{\sqrt{3}}{2} \begin{bmatrix} -2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$. \square

注意 6.7 計量ベクトル空間 V が n 次元であるとき, v_1, \dots, v_n として V の基底を取れば, Schmidt の正規直交化によって V の正規直交基底を構成できる. したがって, 任意の有限次元計量ベクトル空間に正規直交基底が存在する. \square

[207] V は有限次元計量ベクトル空間であるとし, W はその部分空間であるとする. そのとき, W の任意の正規直交基底は V の正規直交基底に拡張可能である. \square

ヒント. W の正規直交基底 u_1, \dots, u_m は V の (単なる) 基底 $u_1, \dots, u_m, v_{m+1}, \dots, v_n$ に拡張可能である. あとはその拡張に Schmidt の正規直交化を適用せよ. \square

[208] (直交補空間) V は有限次元計量ベクトル空間であるとし, W はその部分空間であるとする. そのとき $W^\perp \subset V$ を

$$W^\perp = \{v \in V \mid v \perp W\} = \{v \in V \mid (v, w) = 0 \ (w \in W)\}$$

と定めると W^\perp は V の部分空間をなし, $V = W \oplus W^\perp$ である. W^\perp を W の V における直交補空間 (orthogonal complement) と呼ぶ. \square

ヒント. W の正規直交基底 u_1, \dots, u_m を任意に取り, $v \in V$ に対して, $w, w' \in V$ を

$$w = \sum_{i=1}^m (u_i, v) u_i, \quad w' = v - w$$

と定めると, $w \in W$ かつ $w' \in W^\perp$ である. \square

¹¹⁸数学を広く勉強するためには量子力学を勉強するとよい.

¹¹⁹直交射影の図は線形代数の教科書を見ればどこかに描いてあるはず. P は projection (射影) の頭文字である.

¹²⁰直交補空間については問題 [208] を見よ.

参考文献

- [I] インフェルト, L., ガロアの生涯—神々の愛でし人市井三郎訳, 日本評論社, 新版第3版, 1996
- [Umd] 梅田亨, 非可換行列式との出会い, 『数理科学』1995年4月号 [Gy], 22–25
- [Umm] 梅村浩, 楕円関数論—楕円曲線の解析学, 東京大学出版会, 2000
- [KI] 韓太舜, 伊理正夫, ジョルダン標準形, UP 応用数学選書 8, 東京大学出版会, 1982
- [C] キャッセルズ, J. W., 楕円曲線入門, 徳永浩雄訳, 岩波書店, 1996
- [Gy] 特集「行列式の進化」, 『数理科学』1995年4月号, サイエンス社
- [KO] 小林俊行, 大島利雄, Lie 群と Lie 環 1, 岩波講座現代数学の基礎 12, 岩波書店, 1999
- [St] 佐武一郎, 線型代数学, 数学選書 1, 裳華房, 1974
- [J] 神保道夫, 量子群とヤン・バクスター方程式, シュプリンガー現代数学シリーズ, シュプリンガー・フェアラーク東京, 1990
- [Sh] シャファレヴィッチ, I. R., 代数学とは何か, 蟹江幸博訳, シュプリンガー・フェアラーク東京, 2001
- [ST] シルヴァーマン, J. H., テイト, J., 楕円曲線論入門, 足立恒雄, 木田雅成, 小松啓一, 田谷久雄訳, シュプリンガー・フェアラーク東京, 1995
- [Sg] 杉浦光夫, Jordan 標準形と単因子論 I, II, 岩波講座基礎数学, 線型代数 iii, 1976
- [Tkg1] 高木貞治, 代数学講義, 改定新版, 共立出版, 1965
- [Tkg2] 高木貞治, 初等整数論講義, 第2版, 共立出版, 1971
- [Tkc] 竹内端三, 楕円関数論, 岩波全書, 岩波書店, 1936
- [Ts] 田坂隆士, 2次形式 I, II, 岩波講座基礎数学, 線型代数 iii, 1976
- [Tk] 高崎金久, 可積分系の世界——戸田格子とその仲間——, 共立出版, 2001
- [Tn] 谷崎俊之, リー代数と量子群, 共立叢書 現代数学の潮流, 共立出版, 2002
- [D] デイラック, P. A. M., 量子力学, 朝永振一郎他訳, 岩波書店, 1968 (原書 1958)
- [Tr] 寺沢寛一, 自然科学者のための数学概論, 増訂版, 岩波書店, 1954, 1983, 1986
- [Nk] 中村佳正編, 可積分系の応用数理, 裳華房, 2000
- [Nm] 野海正俊, パンルヴェ方程式——対称性からの入門——, すうがくの風景 4, 朝倉書店, 2000.

- [Hs] 長谷川浩司, 線型代数, 『数学セミナー』における連載, 2001 年 4 月号から 2002 年 10 月号まで, もうすぐ単行本化される予定
- [Ht] 服部昭, 現代代数学, 近代数学講座 1, 朝倉書店, 1968
- [Hr] 広田良吾, 直接法によるソリトンの数理, 岩波書店, 1992
- [Kh] ヒンチン, A. Y., 数論の 3 つの真珠, 蟹江幸博訳, はじめよう数学 4, 日本評論社, 2000
- [H1] 堀田良之, 代数入門——群と加群——, 数学シリーズ, 裳華房, 1987
- [H2] 堀田良之, 加群十話——加群入門——, すうがくぶっくす 3, 朝倉書店, 1988
- [H3] 堀田良之, 環と体 1 — 可換環論, 岩波講座現代数学の基礎 15, 岩波書店, 1997
- [M] 松村英之, 可換環論, 共立出版株式会社, 1980, 2000
- [YmS] 山内恭彦, 杉浦光夫, 連続群論入門, 新数学シリーズ 18, 培風館, 1960
- [Ykt] 横田一郎, 群と位相, 基礎数学選書, 裳華房, 1971
- [Ykn] 横沼健雄, テンソル代数と外積代数, 岩波講座基礎数学, 線型代数 iv, 1976
- [R] リード, M., 可換環論入門, 伊藤由佳理訳, 岩波書店, 2000
- [Wky] 若山正人, 行列式の平方根, 『数理科学』 1995 年 4 月号 [Gy], 32–36
- [Wkm] 脇本実, 無限次元 Lie 環, 岩波講座現代数学の展開 3, 岩波書店, 1999