

線形代数

武井優己

2020 年 5 月 10 日

目次

1	2 次行列の演算	3
1.1	2 次行列の加法	3
1.1.1	零行列	4
1.1.2	2 次行列の減法	4
1.2	2 次行列の乗法	4
1.2.1	2 次行列のスカラー倍	4
1.2.2	2 次行列の積	5
2	一般の行列とベクトル	5
2.1	行列の成分表示	5
2.2	一般の行列の加法、スカラー倍	6
2.3	一般の行列の積	7
2.4	単位行列	7
2.5	逆行列	8
2.6	列ベクトルと行ベクトル	8
2.6.1	$(n, 1)$ 型行列	8
2.6.2	$(1, n)$ 型行列	9
2.6.3	ベクトルの演算	9
2.7	連立一次方程式と行列	9
2.8	転置行列	10
2.8.1	対称と交代	10
2.8.2	内積とテンソル積	10
3	ベクトル空間 (線形空間)	11
3.0.1	単位ベクトル	12
3.1	部分空間	12
3.1.1	部分空間の共通部分	14
3.2	部分空間の和	14

3.3	線形独立と線形従属	15
3.3.1	線形独立と線形従属の定義の成り立ち	16
3.4	部分空間の基底	17
3.5	基底の変換	19
4	線形写像	21
4.1	線形写像の定義	21
4.2	線形写像と行列の対応	22
4.2.1	表現行列	23
4.3	線形写像の合成と行列の積	26
4.4	線形写像の核と像	27
4.5	次元定理	28
4.6	行列の階数	30
4.7	基本変形	31
4.8	基底の延長	33
5	行列式	33
5.1	2次行列における行列式	33
5.1.1	行列式の性質	34
5.2	置換	35
5.2.1	互換の積	37
5.2.2	巡回置換	38
5.3	n 次行列式の定義と基本性質	39
5.3.1	様々な行列式の例	42
5.3.2	余因子	43
5.3.3	余因子展開	45
6	計量ベクトル空間	46
6.1	内積の性質	46
6.1.1	ベクトルのノルム	47
6.1.2	ベクトルのなす角	49
6.1.3	ベクトルの直交条件	51
6.2	正規直交系	52
6.2.1	ノルムが1のベクトルを作る方法	53
6.3	直交行列	55
6.4	直交補空間	58
6.4.1	正射影	59
7	行列の標準化	61
7.1	不変部分空間	61
7.2	固有値、固有空間	62

7.2.1	特性多項式	62
7.2.2	消去法の原理	63
7.3	実対称行列の対角化	67

1 2 次行列の演算

線形代数では、数の拡張としてベクトル、行列という高次元の数を考える。いきなり n 次元に目を向けるのは難しいので、ウォーミングアップとして 2 次元の行列に対して、通常の数と同様の加法と乗法を天下り的に定義する。ここでは、行列は我々の知る実数に似た代数的構造を持つ（実数のような演算ができる）ということが単に分かれればよい。なぜそうなるかはこのレポートを読み進めていくうちに自ずと理解できるようになってくる。

A, A' をそれぞれ以下の 2 次（正方）行列とし、行列の演算を進めていくことにする。2 次正方行列とは行と列がそれぞれ 2×2 の形をした行列のことである。

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, A' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

A であれば、 $(a, b), (c, d)$ の横の並びが行、 $\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}$ の縦の並びが列に該当する。

1.1 2 次行列の加法

A, A' に対し、その和を

$$A + A' = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix}$$

によって定義する。

A, A' の他に、

$$A'' = \begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix}$$

があるとすれば、これらの和は

$$(A + A') + A'' = \begin{pmatrix} (a + a') + a'' & (b + b') + b'' \\ (c + c') + c'' & (d + d') + d'' \end{pmatrix}$$

$$A + (A' + A'') = \begin{pmatrix} a + (a' + a'') & b + (b' + b'') \\ c + (c' + c'') & d + (d' + d'') \end{pmatrix}$$

と書くことができる。

ここで、和で表される各成分は実数における加法の形になっているため、

$$(A + A') + A'' = A + (A' + A'')$$

と結合法則が成立する。同様に、

$$A + A' = A' + A$$

と交換法則も成立する。

1.1.1 零行列

すべての成分が 0 である行列を零行列といい

$$\mathbf{0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

と書く。演算において零行列は単に 0 として表し、 $A + 0 = 0 + A = A$ が成立する。

1.1.2 2 次行列の減法

加法の逆演算として減法も可能である。 $X + A = A'$ となる 2 次行列

$$X = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

が存在する。このような X があったとすれば、 A' のそれぞれの成分は

$$\begin{aligned} x + a &= a', & y + b &= b' \\ z + c &= c', & w + d &= d' \end{aligned}$$

であるから、 X の成分はそれぞれ

$$\begin{aligned} x &= a' - a, & y &= b' - b \\ z &= c' - c, & w &= d' - d \end{aligned}$$

でなければならない。これより、 X の成分は A, A' の差の唯一の解

$$X = \begin{pmatrix} a' - a & b' - b \\ c' - c & d' - d \end{pmatrix}$$

をもつ。この X を $A' - A$ とするのが行列の減法である。

1.2 2 次行列の乗法

1.2.1 2 次行列のスカラー倍

$k \in \mathbb{R}$ に対して

$$kA = \begin{pmatrix} ka & kb \\ kc & kd \end{pmatrix}$$

となるような演算をスカラー倍 (乗法) という。また、

$$A0 = 0A = 0$$

$-A = -1 \times A$ より、

$$A(-A') = (-A)A' = -AA'$$

である。

1.2.2 2 次行列の積

二つの 2 次行列 A, A' の積を次のように定義する。

$$AA' = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$$

A の i 行と A' の j 列の対応する成分の積の和を AA' の (i, j) 成分と定義するのが行列の積である。なぜ加法のような対応ではないのかと疑問に思うが、それは本レポートを読み進めるうちに理解できるようになる。積についても、結合法則が成立する。

$$A'' = \begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix}$$

とすると、

$$(AA')A'' = A(A'A'')$$

である。しかし、2 次行列の積に関しては交換法則

$$AA' = A'A$$

は一般には成立しない。実際、 AA' の $(1, 1)$ 成分は $aa' + bc'$ であるが、 $A'A$ の $(1, 1)$ 成分は、 $a'a + b'c$ となり、 $AA' \neq A'A$ である。

加法と乗法を組み合わせると、

$$\begin{aligned} A(A' + A'') &= AA' + AA'' \\ (A' + A'')A &= A'A + A''A \end{aligned}$$

のような分配法則が成立する。

このように、2 次正方行列は一部の例外を除きほぼ実数と同じような振る舞いをする事が分かる。(これは、群・環・体といった代数的構造の賜である)

2 一般の行列とベクトル

2.1 行列の成分表示

$m \times n$ 個の数を矩形に並べ括弧で囲んだものを m 行 n 列の行列あるいは $m \times n$ 行列という。 $m \times n$ 行列の全体の集合を $M_{m,n}(\mathbb{R})$ と書き、 $m = n$ であるとき、 n 次正方行列といい、 $M_n(\mathbb{R})$ と書く。行列を構成するそれぞれの数を成分と呼ぶ。

定義 2.1. 一般の行列の成分表示

上述の例より、一般の $m \times n$ 行列 A を成分表示すると次のようになる。 A の (i, j) 成分を (a_{ij}) とするとき

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

である。

多くの場合これを省略して、 $A = (a_{ij})$ と表記する。ただし、 $(1 \leq i \leq m), (1 \leq j \leq n)$

例 2.2. 2×2 行列の表示

$m = 2, n = 2$ のとき、 2 次正方行列 A は以下のように表示される

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

例 2.3. 4×5 行列の表示

$m = 4, n = 5$ のとき、 4×5 行列 A は以下のように表示される

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{pmatrix}$$

2.2 一般の行列の加法、スカラー倍

定義 2.4. 一般の行列の加法、スカラー倍

$A, B \in M_{m,n}(\mathbb{R}), A = (a_{ij}), B = (b_{ij})$ とし、加法、スカラー倍を以下のように定義する。

- (1) $A + B = (a_{ij} + b_{ij})$
- (2) $cA = (ca_{ij}) \quad (c \in \mathbb{R})$
- (3) $(-1)A = -A, \quad A + (-B) = A - B$

命題 2.5. 定義 2.4 より、以下の法則を満たす

- (1) $(A + B) + C = A + (B + C)$ (結合法則)
- (2) $A + B = B + A$ (交換法則)
- (3) $c(A + B) = cA + cB$ (分配法則)

である。

証明. (1) a_{ij} と b_{ij} の各成分どうしに結合法則が成り立つことより明らか

(2) a_{ij} と b_{ij} の各成分どうしに交換法則が成り立つことより明らか

(3) $A + B = a_{ij} + b_{ij}$ より、 $c(A + B) = c(a_{ij} + b_{ij})$ である。 $cA = ca_{ij}, cB = cb_{ij}$ であるから、 $cA + cB = ca_{ij} + cb_{ij}$ となる。ここで実数における分配法則により、 $c(a_{ij} + b_{ij}) = ca_{ij} + cb_{ij}$ より、すべての (i, j) について分配法則が適用されるから、 $c(A + B) = cA + cB$ である。

定義 2.6. 零行列との加法、スカラー倍

$A \in M_{m,n}(\mathbb{R})$ とするとき、零行列 0 との加法、スカラー倍で以下が成立する。

- (1) $A + 0 = 0 + A = A$
- (2) $A + (-A) = (-A) + A = 0$
- (3) $1A = A$
- (4) $0A = 0$

2.3 一般の行列の積

定義 2.7. $l, m, n \in \mathbb{Z}$ とし、 $A \in M_{l,m}(R), B \in M_{m,n}(\mathbb{R})$ とするとき、積 AB を以下のように定義する。

$$AB = \sum_{1 \leq k \leq m} a_{ik} b_{kj} \in M_{l,n}(\mathbb{R}) \quad (\text{ただし、} 1 \leq i \leq l, 1 \leq j \leq n)$$

1.2 で紹介した 2×2 行列の積がまさにこのようになっている。このように行列の積を定義する理由は、線形写像の章で説明する。

命題 2.8. 一般の行列の積について以下の法則が成立する

- (1) $A(BC) = (AB)C$
- (2) $A(B + C) = AB + AC$
- (3) $(B + C)A = BA + CA$

ただし、行列 A, B, C はそれぞれの場合にこれらの演算が可能な型である必要があることに注意。

これらも AB の各成分どうしが実数の積と和の形をしていることから、実数の演算における法則が適用できる。これにより、実際に演算することでこれらの証明は容易である。(2) と (3) が同値でない理由として、行列の積は非可換なことが挙げられる。実際、 AB と BA を行列の積の定義に従い演算すると、それぞれ異なる行列が得られ、 $AB \neq BA$ となる。そのため、 $A(B + C) \neq (B + C)A$ となる。

2.4 単位行列

定義 2.9. $A \in M_n(\mathbb{R})$ に対して $AE = EA = A$ となるような行列を n 次単位行列という。単に、単位行列ということもある。

単位行列は、

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

のように対角成分が 1 でそれ以外が 0 から構成される行列である。

2.5 逆行列

n 次行列 A に対し $AX = E, XA = E$ を満たすような n 次行列 X が存在するとき、 A は可逆である、または正則行列という。

定理 2.10. n 次行列 A が可逆であれば、 $AX = E$ および $YA = E$ はそれぞれ唯一の解を持ち、それらの解は一致する。そのような X, Y を A^{-1} と書き、逆行列という。

証明. $AX = XA = E, AY = YA = E$ となる行列 X, Y があるとする。このとき、結合法則により

$$\begin{aligned} X &= XE \\ &= X(AY) \\ &= (XA)Y \\ &= EY \\ &= Y \end{aligned}$$

となり、 $X = Y$ である。仮に、 $AB = BA = E, AC = CA = E$ となる行列 B, C があったとすれば、同様の理由により $X = Y = B = C$ とこれらの方程式の解は一意に定まる。

例 2.11. 単位行列の逆行列

単位行列 E は $EE = E$ であるから $E^{-1} = E$ であり可逆である。 $A, B \in M_n(\mathbb{R})$ が可逆ならば、 AB も可逆で $(AB)^{-1} = B^{-1}A^{-1}$ である。これが言える理由として、 $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = E$ だからである。また、 A が可逆であれば A^{-1} も可逆であり、 $(A^{-1})^{-1} = A$ 。つまり、 $(A^{-1})A = A(A^{-1}) = E$ である。

これは、 $M_n(\mathbb{R})$ の可逆元全体が群の性質を持つことを言っており、この行列全体の集合を一般線形群といい、 $GL_n(\mathbb{R})$ と表される。

2.6 列ベクトルと行ベクトル

2.6.1 $(n, 1)$ 型行列

$$B = \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix}$$

を **n-dim** 列ベクトルという。

本レポートでは、 n -dim 列ベクトルを

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

のようにボードのアルファベットで表記する。

2.6.2 (1, n) 型行列

(b_{11}, \dots, b_{n1}) を **n-dim** 行ベクトルという。

2.6.3 ベクトルの演算

n-dim 列ベクトルは (n,1) 型の行列であるから、やはり加法と乗法が成立する。以降の節の理解を円滑に進めるに当たり、ここではベクトルの演算に関する細かな定義や性質の説明は飛ばす。(3章のベクトル空間にて、この辺りの細かい議論を行う) よって、予めベクトルの演算はできるものとして気楽に考えてほしい。

2次元ベクトルにおいて、

$$\mathbf{a} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{b} = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

を取ってくる。

このとき、 $\mathbf{a} + \mathbf{b}$ は

$$\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x + x' \\ y + y' \end{pmatrix}$$

である。

また、 $c \in \mathbb{R}$ に対して、

$$c\mathbf{a} = c \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} cx \\ cy \end{pmatrix}$$

である。

2.7 連立一次方程式と行列

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}), \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

とすれば、

$$A\mathbf{x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

のように、2次元ベクトル \mathbf{x} に2次行列 A を差乗することができる。この行列算の応用として $A\mathbf{x}$ を連立一次方程式として表示することを考える。 $ax + by$ と $cx + dy$ の和をそれぞれ z, w とすると

$$\begin{cases} ax + by = z \\ cx + dy = w \end{cases} \quad (1)$$

$$\quad (2)$$

のような連立方程式で表すことができる。これより、もし A が逆行列を持つなら、

$$\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1} \begin{pmatrix} z \\ w \end{pmatrix}$$

と計算することができる。

2.8 転置行列

(m, n) 行列

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

に対し対角線 $a_{11}a_{22}\cdots$ に対して、折返してできる (n, m) 行列を tA と書き、 A の転置行列という。

tA を表示すると以下の通りである。

$${}^tA = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$$

命題 2.12. $A, B \in M_{m,n}(R)$ に対して、明らかに ${}^t(A+B) = {}^tA + {}^tB$, ${}^{t({}^tA)} = A$ が成立する。

2.8.1 対称と交代

$A \in M_n(\mathbb{R})$ とする。 $A = {}^tA$ つまり、 $a_{ij} = a_{ji}$ ($1 \leq i, j \leq n$) であるとき、 A を対称または対称行列という。

$A = -{}^tA$ つまり、 $a_{ij} = -a_{ji}$ ($1 \leq i, j \leq n$) であるとき、 A を交代 (的) または交代行列という。

2.8.2 内積とテンソル積

後の計量ベクトル空間の章で内積について詳しく説明するが、ここでは内積の定義とその意味についてだけ軽く触れておく。

二つの n 次元ベクトル

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

に対し、

$${}^t\mathbf{a}\mathbf{b} = \sum_{1 \leq i \leq n} a_i b_i = (a_1, \dots, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

を内積といい $\langle \mathbf{a}, \mathbf{b} \rangle$ と書く。

また、

$$\mathbf{a}{}^t\mathbf{b} = \begin{pmatrix} a_1 b_1 & \cdots & a_1 b_n \\ \vdots & \ddots & \vdots \\ a_n b_1 & \cdots & a_n b_n \end{pmatrix}$$

を \mathbf{a}, \mathbf{b} のテンソル積といい、 $\mathbf{a} \otimes \mathbf{b}$ と書く。

例題 2.13. $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ を n 次元ベクトルとすると、次の等式を証明せよ

- (1) ${}^t(a \otimes b) = b \otimes a$
- (2) $(a \otimes b)c = \langle b, c \rangle a$
- (3) $\langle c, (a \otimes b)d \rangle = \langle a, c \rangle \langle b, d \rangle$

解答

- (1) ${}^t(a \otimes b) = {}^t(a {}^t b) = {}^{tt} b {}^t a = b {}^t a = b \otimes a$
- (2) $(a \otimes b)c = a({}^t b c) = \langle b, c \rangle a$
- (3) $\langle c, (a \otimes b)d \rangle = ({}^t c a)({}^t b d) = \langle c, a \rangle \langle b, d \rangle = \langle a, c \rangle \langle b, d \rangle$

3 ベクトル空間 (線形空間)

n 次元ベクトル全体の集合を \mathbb{R}^n とかく。2 章の時点ですでに 2 次元のベクトルやベクトルと行列間に演算を導入していたが、一般的な \mathbb{R}^n に対して、演算を改めて定義する。

定義 3.1.

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n, a \in \mathbb{R}$$

とするとき、これらの加法 (減法)、スカラー倍はそれぞれ次のように定義される。

$$\boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \boldsymbol{x} - \boldsymbol{y} = \boldsymbol{x} + -(\boldsymbol{y}) = \begin{pmatrix} x_1 + -(y_1) \\ \vdots \\ x_n + -(y_n) \end{pmatrix}, a\boldsymbol{x} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$$

\mathbb{R}^n が次の条件を満たしたとき、単なる集合ではなく、 n 次元ベクトル空間となる。単にベクトル空間ともいう。

定義 3.2. $V = \mathbb{R}^n$ とするとき、 V は以下の条件のもと、ベクトル空間となる

- (1) 任意の $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ に対して、 $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$ (結合法則)
- (2) 任意の $\boldsymbol{x}, \boldsymbol{y} \in V$ に対して、 $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$ (交換法則)
- (3) 任意の $\boldsymbol{x} \in V$ に対して、 $\mathbf{0} + \boldsymbol{x} = \boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$ (零ベクトルの存在)
- (4) 任意の $\boldsymbol{x} \in V$ に対して、逆ベクトル $-\boldsymbol{x}$ が存在して、 $\boldsymbol{x} + (-\boldsymbol{x}) = (-\boldsymbol{x}) + \boldsymbol{x} = \mathbf{0}$ (逆ベクトルの存在)
- (5) 任意の $\boldsymbol{x}, \boldsymbol{y} \in V, a, b \in \mathbb{R}$ に対して
 - (a) $a(\boldsymbol{x} + \boldsymbol{y}) = a\boldsymbol{x} + a\boldsymbol{y}$
 - (b) $(a + b)\boldsymbol{x} = a\boldsymbol{x} + b\boldsymbol{x}$
 - (c) $a(b\boldsymbol{x}) = ab(\boldsymbol{x})$
- (6) 任意の $\boldsymbol{x} \in V$ に対して $1 \cdot \boldsymbol{x} = \boldsymbol{x}$ ($1 \in \mathbb{R}$)

命題 3.3. これらをベクトル空間の公理と認めると、以下が成り立つ

- (1) $\mathbf{0}\boldsymbol{x} = \mathbf{0}$

$$(2) (-1)x = -x$$

証明. (1) (5) より、 $0x = (0 + 0)x = 0x + 0x$ となり、両辺から $0x$ を引くと、 $0 = 0x$

$$(2) 0 = 0x = (1 + (-1))x = 1 \cdot x + (-1) \cdot x = x + (-1)x \text{ となり、両辺から } -x \text{ を引くと、} -x = (-1)x$$

3.0.1 単位ベクトル

特殊なベクトルとして $e_1, \dots, e_n \in V$ があり、それぞれが

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

のように大きさが 1 になるようなベクトルを、単位ベクトルという。この単位ベクトル e_1, \dots, e_n を使えば、任意の V に対して

$$x = \sum_{1 \leq i \leq n} x_i e_i$$

と一意に表現できる。

3.1 部分空間

定義 3.4. V を n 次元ベクトル空間とする。 $V \supset W \neq \phi$ とする。 W が以下の条件を満たすとき、 W を V の部分空間という

$$(1) x, y \in W \Rightarrow x + y \in W$$

$$(2) x \in W, a \in \mathbb{R} \Rightarrow ax \in W$$

これらの条件は、 W の中でベクトルとしての演算が閉じていることを示している。

補題 3.5. W が V の部分空間であれば、 W もベクトル空間である

証明. 仮定により $x \in W$ であるから、定義 3.4(2) を使うと、 $0 \in \mathbb{R}, 0 \cdot x = 0 \in W$ と $-1 \in \mathbb{R}, (-1)x = -x \in W$ である。したがって、任意の x, y に対して、 $x - y = x + (-y) \in W$ である。ベクトルの演算の公理が V で成立しているので、 W においても成立することは明らか。また零ベクトルと任意のベクトル x の逆ベクトルも W に存在することは分かっているので、 W もまたベクトル空間である。

つまり、部分空間はベクトル空間として扱うことができるということである。

命題 3.6. $\{0\}$ 、 $V (= \mathbb{R}^n)$ は部分空間である

証明. $0 + 0 = 0$, $a0 = 0$ ($\forall a \in \mathbb{R}, 0 \in \{0\}$) より、 $\{0\}$ は部分空間である。また、 V はベクトル空間なので、部分空間の条件を満たすことは明らか。これらの部分空間を自明な部分空間と呼ぶ。

例題 3.7. $V = \mathbb{R}^3$ において、以下は部分空間か判定せよ

(1)

$$W_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1 + x_2 + x_3 = 0 \right\}$$

(2)

$$W_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid 2x_1 + x_2 = 5 \right\}$$

解答

(1) 部分空間である。実際、

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in W_1$$

をとってきて、 $\mathbf{x} + \mathbf{y}$ を計算すると、 $(x_1 + x_2 + x_3) = 0, (y_1 + y_2 + y_3) = 0$ より、 $(\mathbf{x}_1 + \mathbf{y}_1) + (\mathbf{x}_2 + \mathbf{y}_2) + (\mathbf{x}_3 + \mathbf{y}_3) = \mathbf{0} \in W_1$ である。同様に、 $\forall a \in \mathbb{R}$ を取ってきて、 $a\mathbf{x} = \mathbf{0} \in W_1$ である。

(2) 部分空間ではない。凡例をあげる。

$$\mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} \frac{1}{2} \\ 4 \\ 0 \end{pmatrix} \in W_2$$

をとってくる。両方とも上の条件を満たすが、

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} \frac{3}{2} \\ 7 \\ 0 \end{pmatrix}$$

$2 \cdot \frac{3}{2} + 7 \neq 5$ であるため、 $\mathbf{x} + \mathbf{y} \notin W_2$ となり、部分空間の必要十分条件に反する。

定義 3.8. $V = \mathbb{R}^n$ とする。任意の $\mathbf{x} \in V$ に対し、そのスカラー倍全体の集合 $\{a\mathbf{x} \mid a \in \mathbb{R}\}$ は明らかに V の部分空間になる。さらに一般的に、 $\mathbf{x}_1, \dots, \mathbf{x}_r \in V$ が与えられたとき、

$$\mathbf{x} = \sum_{1 \leq i \leq r} a_i \mathbf{x}_i \quad (a_1, \dots, a_r \in \mathbb{R})$$

のような形で \mathbf{x} が表されるとき、これを $\mathbf{x}_1, \dots, \mathbf{x}_r$ の線形結合 (linear combination) または 1 次結合という。

$\mathbf{x}_1, \dots, \mathbf{x}_r$ の線形結合全体の集合を $W = \{a_1 \mathbf{x}_1 + \dots + a_r \mathbf{x}_r \mid a_1, \dots, a_r \in \mathbb{R}\}$ とすると、 W は V の一組の部分空間になる。これを $\mathbf{x}_1, \dots, \mathbf{x}_r$ によって張られる部分空間といい、 $\langle \mathbf{x}_1, \dots, \mathbf{x}_r \rangle$ のように書く。(内積と混同しないように注意したい。)

証明.

$$\mathbf{0} = \sum_{1 \leq i \leq r} 0 \mathbf{x}_i \in W$$

より $W \neq \emptyset$ である。 $\mathbf{x}, \mathbf{y} \in W$ とすれば、

$$\mathbf{x} = \sum_{1 \leq i \leq r} a_i \mathbf{x}_i, \quad \mathbf{y} = \sum_{1 \leq i \leq r} b_i \mathbf{x}_i \quad (a_i, b_i \in \mathbb{R})$$

と書くことができる。これより、

$$\mathbf{x} + \mathbf{y} = \sum_{1 \leq i \leq r} (a_i + b_i) \mathbf{x}_i$$

とそれぞれのベクトルを \mathbf{x}_i の線形結合で表すことができる。

また、 $c \in \mathbb{R}$ に対し、

$$c\mathbf{x} = c \left(\sum_{1 \leq i \leq r} a_i \mathbf{x}_i \right) = \sum_{1 \leq i \leq r} (ca_i) \mathbf{x}_i$$

とこれも \mathbf{x}_i の線形結合で表すことができる。

したがって、 $\mathbf{x} + \mathbf{y}, \quad c\mathbf{x} \in W$ となることから、 W は部分空間である。

3.1.1 部分空間の共通部分

命題 3.9. W_1, W_2 を V の二つの部分空間とすると、その共通部分 $W_1 \cap W_2$ は部分空間になる。

証明. $\mathbf{0} \in W_1 \cap W_2$ であるから、 $W_1 \cap W_2 \neq \emptyset$ である。 $\mathbf{x}, \mathbf{y} \in W_1 \cap W_2$ とすれば、 $\mathbf{x} + \mathbf{y} \in W_1$ であり、 $\mathbf{x} + \mathbf{y} \in W_2$ である。 $(\mathbf{x}, \mathbf{y} \in W_1, \mathbf{x}, \mathbf{y} \in W_2$ だから。) したがって、 $\mathbf{x} + \mathbf{y} \in W_1 \cap W_2$ 。同様に、 $a \in \mathbb{R}$ に対して、 $a\mathbf{x} \in W_1 \cap W_2$ である。

3.2 部分空間の和

部分空間の共通部分を考えてときに、今度は和集合（ベクトル空間どうしの和）も部分空間になるかを考えてみる。

$W_1 = \langle \mathbf{e}_1 \rangle, \quad W_2 = \langle \mathbf{e}_2 \rangle$ とおくと、

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in W_1 \cup W_2$$

であるが、すべての $a, b \in \mathbb{R}$ に対して、

$$a\mathbf{e}_1 + b\mathbf{e}_2 = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \notin W_1 \cup W_2$$

となる。（厳密に言うと、和集合の構成次第で含まれているところもあるかもしれないが、含まれていないところもある。）

そこで、 $a\mathbf{e}_1 + b\mathbf{e}_2$ をすべて含むようにすることを考えるために、 $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle$ で張られる空間を考えると良さそうである。

定義 3.10. 一般に二つの部分空間 W_1, W_2 の和を以下のように定義する。

$$W_1 + W_2 = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in W_1, \mathbf{y} \in W_2\}$$

この演算で定義される部分空間は V の最小の部分空間となる。それを次の定理で示す。

定理 3.11. $V = \mathbb{R}^n$ とし、 W_1, W_2 を V の部分空間とする。 $W_1 + W_2$ は W_1 と W_2 を含む最小の部分空間となる。

証明. まず、 $W_1 + W_2$ が部分空間であることを示す。

$\mathbf{0} = \mathbf{0} + \mathbf{0} \in W_1 + W_2$ より、 $W_1 + W_2 \neq \phi$ 。 $\mathbf{x}, \mathbf{y} \in W_1 + W_2$ とすれば、定義 3.10 により、 $\mathbf{x}', \mathbf{y}' \in W_1, \mathbf{x}'', \mathbf{y}'' \in W_2$ と書くことができる。(定義 3.8 で示した形式を思い出すと良い。) よって、 $\mathbf{x} + \mathbf{y} = (\mathbf{x}' + \mathbf{x}'') + (\mathbf{y}' + \mathbf{y}'') = (\mathbf{x}' + \mathbf{y}') + (\mathbf{x}'' + \mathbf{y}'')$ であり、 $a \in \mathbb{R}$ に対して、 $\mathbf{x}' \in W_1$ より $a\mathbf{x}' \in W_1$ また、 $\mathbf{x}'' \in W_2$ より $a\mathbf{x}'' \in W_2$ である。

これらより、 $\mathbf{x}' + \mathbf{y}' \in W_1, \mathbf{x}'' + \mathbf{y}'' \in W_2$ であるから、 $\mathbf{x} + \mathbf{y} \in W_1 \cup W_2$ であり、 $a\mathbf{x} \in W_1 \cup W_2$ である。 $(a\mathbf{x} \in W_1 \text{ であるから、和集合である } W_1 \cup W_2 \text{ に属するのは当然})$ よって、 $W_1 + W_2$ は V の部分空間である。

次に、 $W_1 + W_2$ が最小の部分空間であることを示す。

$\mathbf{x} \in W_1$ とすれば、 $\mathbf{x} = \mathbf{x} + \mathbf{0}$ ($\mathbf{x} \in W_1, \mathbf{0} \in W_2$) より、 $\mathbf{x} \in W_1 + W_2$ である。よって、 $W_1 \subset W_1 + W_2$ 。同様に $W_2 \subset W_1 + W_2$ が言え、 $W_1 + W_2$ は W_1 と W_2 を含むことが分かる。ここで、 W を W_1, W_2 を含む部分空間とすれば、任意の $\mathbf{x} \in W_1, \mathbf{y} \in W_2$ に対し、 $\mathbf{x}, \mathbf{y} \in W$ であるから、 $\mathbf{x} + \mathbf{y} \in W$ 。よって、 $W_1 + W_2 \subset W$ となり、 $W_1 + W_2$ は W_1 と W_2 を含む最小の部分空間である。

例題 3.12. 一般に二つの部分空間 W_1, W_2 に対し、 $W_1 \subset W_2 \iff W_1 + W_2 = W_2$ であることを示せ。

$W_1 \subset W_2$ とすれば、 $W_1 + W_2 \subset W_2 + W_2 = W_2$ 。 $W_1 + W_2 \supset W_2$ は自明であるから、 $W_1 + W_2 = W_2$ 。逆に、 $W_1 + W_2 = W_2$ とすると、 $W_1 \subset W_1 + W_2$ であるから、 $W_1 \subset W_2$ である。

例題 3.13. W_1, W_2, W_3 を $V = \mathbb{R}^n$ の部分空間とすると、 $(W_1 \cap W_3) + (W_2 \cap W_3) \subset (W_1 + W_2) \cap W_3$ を示せ。

$\mathbf{x} \in (W_1 \cap W_3) + (W_2 \cap W_3)$ とする。このとき、 $\mathbf{x} = \mathbf{x}' + \mathbf{x}''$ と分解でき、 $\mathbf{x}' \in W_1 \cap W_3$ & $\mathbf{x}'' \in W_2 \cap W_3$ 。 $\mathbf{x}', \mathbf{x}''$ が両方とも W_3 の元より、 $\mathbf{x} \in W_3$ また、 W_1, W_2 の定義より、 $\mathbf{x} \in W_1 + W_2$ 。したがって、 $\mathbf{x} \in (W_1 + W_2) \cap W_3$ 。

3.3 線形独立と線形従属

定義 3.14. r 個のベクトル $\mathbf{x}_1, \dots, \mathbf{x}_r$ は、次の条件が満たされるとき線形独立 (linearly independent) または一次独立であるという。

$$\sum_{1 \leq i \leq r} a_i \mathbf{x}_i = \mathbf{0} \quad (a_i \in \mathbb{R}) \implies a_1 = \dots = a_r = 0$$

一方、線形独立ではない $\mathbf{x}_1, \dots, \mathbf{x}_r$ を線形従属 (linearly dependent) または 1 次従属であるという。 $\mathbf{x}_1, \dots, \mathbf{x}_r$ が線形従属であるとき以下の 1 次関係式が成立する。

$$\sum_{1 \leq i \leq r} a_i \mathbf{x}_i = \mathbf{0} \quad (a_i \in \mathbb{R}) \text{ \& } \exists a_i \neq 0$$

3.3.1 線形独立と線形従属の定義の成り立ち

線形独立と線形従属の定義は初見では理解が難しいので、この定義の成り立ちを考えてみることにする。

$V = \mathbb{R}^3$ において 3 つの基本ベクトル

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

はそれぞれ独立した方向を向いていると考えられる。つまり、 $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle$ を xy 平面とすると、 \mathbf{e}_3 はそれに属さない、つまり z 軸と考えることができる。これは \mathbf{e}_3 は $\mathbf{e}_1, \mathbf{e}_2$ の線形結合で表されないということである。同様に $\mathbf{e}_1, \mathbf{e}_2$ のいずれかと \mathbf{e}_3 が張る部分空間を考えたときに、残りの単位ベクトルはその部分空間に属さない。つまりこれは

$$a\mathbf{e}_1 + b\mathbf{e}_2 + c\mathbf{e}_3 = \mathbf{0} \quad (a, b, c \in \mathbb{R}) \quad (3)$$

ならば $a = b = c = 0$ と言い換えることができる。これは定義 3.14 そのものである。なぜ $a = b = c = 0$ であれば線形独立なのかを考えたいのであれば、逆に $c \neq 0$ とすれば式 (2) から

$$\mathbf{e}_3 = \left(-\frac{a}{c}\right)\mathbf{e}_1 + \left(-\frac{b}{c}\right)\mathbf{e}_2$$

のように \mathbf{e}_3 を \mathbf{e}_1 と \mathbf{e}_2 の線形結合で表すことが出来てしまう。つまり、 $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle$ に \mathbf{e}_3 が含まれていることを意味する。 $a \neq 0, b \neq 0$ のときも同様である。

命題 3.15. $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r$ が線形独立ならばその 1 部分も線形独立である。線形独立の定義よりこれは明らかなので証明は割愛する。

例題 3.16. $V = \mathbb{R}^3$ において

$$\mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \mathbf{z} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

が線形独立であることを示せ。

$a\mathbf{x} + b\mathbf{y} + c\mathbf{z} = \mathbf{0} \quad (a, b, c \in \mathbb{R})$ であれば、 $a = b = c = 0$ を示す。

$$\begin{cases} a + (-b) = 0 \\ a + (-c) = 0 \\ a + b + c = 0 \end{cases} \quad \begin{matrix} (4) \\ (5) \\ (6) \end{matrix}$$

という連立 1 自方程式を解けば、 $a = b = c = 0$ が唯一の解となるので、 $\mathbf{x}, \mathbf{y}, \mathbf{z}$ は線形独立である。

例題 3.17. $V = \mathbb{R}^3$ において

$$\mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \mathbf{z} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

は線形独立であるか調べよ。

$ax + by + cz = 0$ ($a, b, c \in \mathbb{R}$) とすると、 a, b, c は

$$\begin{cases} a + c = 0 & (7) \\ -a + b = 0 & (8) \\ -b + c = 0 & (9) \end{cases}$$

を満たす。このときこの連立一次方程式は $a = 1, b = 1, c = -1$ の解を持つため、 x, y, z は線形従属である。

補題 3.18. $x_1, x_2, \dots, x_r \in V$ に関して次の 2 つの条件は同値である。

- (1) x_1, \dots, x_r は線形独立である
- (2) x_1, \dots, x_{r-1} は線形独立で、 $x_r \notin \langle x_1, \dots, x_{r-1} \rangle$

証明. (1) \Rightarrow (2) を示す。(1) を仮定すると、 $a_1x_1 + \dots + a_{r-1}x_{r-1} + a_rx_r = 0$ ならば、 $a_i = 0$ ($1 \leq i \leq r$) である。(2) が成り立たないとすれば、 $x_r \in \langle x_1, \dots, x_{r-1} \rangle$ であるから $x_r = a_1x_1 + \dots + a_{r-1}x_{r-1} = 0$ と表せる。よって、 $a_1x_1 + \dots + a_{r-1}x_{r-1} + (-1)x_r = 0$ が成立し $a_r \neq 0$ であるから (1) に反する。よって、 $x_r \notin \langle x_1, \dots, x_{r-1} \rangle$

(2) \Rightarrow (1) を示す。 x_1, \dots, x_r が線形独立でないとする、ある $a_i \neq 0$ ($1 \leq i \leq r$) が存在し、 $a_1x_1 + \dots + a_rx_r = 0$ が成立する。もし $a_r = 0$ とすると a_1, \dots, a_{r-1} のいずれかが 0 でなくなる。これは、(2) の x_1, \dots, x_{r-1} が線形独立であるという仮定に反する。よって、 $a_r \neq 0$ でなければならない。しかしこのとき、

$$\left(-\frac{a_1}{a_r}\right)x_1 + \dots + \left(-\frac{a_{r-1}}{a_r}\right)x_{r-1} = x_r \in \langle x_1 + \dots + x_{r-1} \rangle$$

となり (2) の仮定に反する。 x_1, \dots, x_r は線形独立である。

定理 3.19. x_1, \dots, x_r が線形独立であるための必要十分条件は $x_1 \neq 0$ かつ、任意の i に対して $x_i \notin \langle x_1, \dots, x_{i-1} \rangle$ ($2 \leq i \leq r$) である。

証明. (\Rightarrow): 補題 3.18 により、 x_1, \dots, x_r が線形独立であれば、 x_1, \dots, x_i も線形独立であるから、この主張が成立する。

(\Leftarrow): この主張が成立すると仮定する。 x_1, \dots, x_i が線形独立であることは、補題 3.18 を再帰的に行うことで判断できる。まず $x_1 \neq 0$ から x_1 は線形独立である。次に、 $\langle x_1 \rangle$ に x_2 が含まれていないか、さらにその次に $\langle x_1, x_2 \rangle$ に x_3 が含まれていないかといった判定を x_{i-1} まで再帰的に行っていく。このとき、 x_1, \dots, x_{i-1} が線形独立であるとすれば、補題 3.18 により、 x_1, \dots, x_i も線形独立である。この判定方法は $i = r$ まで成立する。(この操作は高々 r 回で終わる。)

3.4 部分空間の基底

基底の定義をする前に、以下の補題にてベクトル空間を張る線形独立なベクトルは別の線形独立なベクトルに置き換えることができることを示す。

補題 3.20. $W = \langle x_1, \dots, x_r \rangle$ とする。 $y_1, \dots, y_s \in W$ が線形独立であるとすれば、 $s \leq r$ である。

証明. $\{i_1, \dots, i_s\} \subset \{1, \dots, t\}$ をとなる $\{i_1, \dots, i_s\}$ とってくる。 ($\{i_1, \dots, i_s\}$ は、この時点では順不同である。) W を張る $\mathbf{x}_1, \dots, \mathbf{x}_t$ のうち、 $\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_s}$ を $\mathbf{y}_1, \dots, \mathbf{y}_s$ に置き換える、つまり各 \mathbf{y}_i は

$$\mathbf{y}_i = \sum_{1 \leq j \leq s} a_j \mathbf{x}_{i_j}$$

とかけを帰納的に示す。まず、 $s = 0$ のときは自明である。 $s \geq 1$ として、 $\mathbf{y}_1, \dots, \mathbf{y}_{s-1}$ に対してこの主張が言えたとする。つまり $\{i_1, \dots, i_{s-1}\} \subset \{1, \dots, t\}$ をとってきて、 $\mathbf{x}_1, \dots, \mathbf{x}_t$ の中の $\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_{s-1}}$ が $\mathbf{y}_1, \dots, \mathbf{y}_{s-1}$ に替えられたとする。これを、

$$\mathbf{x}'_i = \begin{cases} \mathbf{y}_k & i = i_k (1 \leq k \leq s-1) \\ \mathbf{x}_i & \text{otherwise} \end{cases}$$

とおく。このとき、 $W = \langle \mathbf{x}'_1, \dots, \mathbf{x}'_t \rangle$, $\mathbf{y}_s \in W$ であるから、 $\mathbf{y}_s = a_1 \mathbf{x}'_1 + \dots + a_t \mathbf{x}'_t$ と表される。しかし、補題 3.18 により $\mathbf{y}_s \notin \langle \mathbf{y}_1, \dots, \mathbf{y}_{s-1} \rangle = \langle \mathbf{x}'_{i_1}, \dots, \mathbf{x}'_{i_{s-1}} \rangle$ であるから、ある $i \notin \{i_1, \dots, i_{s-1}\}$ に対して $a_i \neq 0$ である。そのような i の 1 つを i_s とする。そのとき、 $\mathbf{x}'_{i_s} = \mathbf{x}_{i_s}$ は、 $\mathbf{x}'_1, \dots, \mathbf{x}'_{i_{s-1}}, \mathbf{y}_s, \mathbf{x}'_{i_{s+1}}, \dots, \mathbf{x}'_t$ の線形結合で表される。よって、 \mathbf{x}'_{i_s} を \mathbf{y}_s で置き換えることができる。

この補題により 1 から $r = s$ まで順番にこの置き換えを行うことで、次の定理が得られる。

定理 3.21. W を $V = \mathbb{R}^n$ の部分空間とする。 $W = \langle \mathbf{x}_1, \dots, \mathbf{x}_r \rangle = \langle \mathbf{y}_1, \dots, \mathbf{y}_s \rangle$ で、 $\mathbf{x}_1, \dots, \mathbf{x}_r$ と $\mathbf{y}_1, \dots, \mathbf{y}_s$ がともに線形独立であるとすれば、 $r = s$ である。

定義 3.22. W を $V = \mathbb{R}^n$ の部分空間とする。 W に対し、次の条件を満たす順序付けられたベクトルの集合を W の基底 (basis) という。

- (1) $W = \langle \mathbf{x}_1, \dots, \mathbf{x}_r \rangle$
- (2) $\mathbf{x}_1, \dots, \mathbf{x}_r$ は線形独立である。

定理 3.23. $V = \mathbb{R}^n$ の任意の部分空間 W に対して基底 $\mathbf{x}_1, \dots, \mathbf{x}_r$ が存在する。基底をなすベクトルの個数は一定で、 $r \leq n$ となる

証明. $W \subset V = \langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle$ であるから、補題 3.20 により W に含まれる線形独立なベクトルの個数は高々 n である。 $\mathbf{x}_1, \dots, \mathbf{x}_r$ を W に含まれる最大個数の線形独立なベクトルの集合とする。そのとき、 $r \leq n$ で任意の $\mathbf{x} \in W$ に対し、 $\mathbf{x} \in \langle \mathbf{x}_1, \dots, \mathbf{x}_r \rangle$ が成立する。よって、 $W \subset \langle \mathbf{x}_1, \dots, \mathbf{x}_r \rangle$ である。 $\langle \mathbf{x}_1, \dots, \mathbf{x}_r \rangle \subset W$ は明らかであるから、 $W = \langle \mathbf{x}_1, \dots, \mathbf{x}_r \rangle$ 。よって、 $\{\mathbf{x}_1, \dots, \mathbf{x}_r\}$ は W の基底である。補題 3.20 により、 r は一意的に定まる。

この定理により部分空間 W に対して定まる ($r \in \mathbb{N}$) を W の次元 (dimension) といい、 $\dim W$ と書く。

定理 3.24. $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ を線形独立なベクトルの組とする。このとき、 $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ の線形結合としての表し方は一意的である。すなわち、 $\mathbf{x} = a_1 \mathbf{x}_1 + \dots + a_n \mathbf{x}_n$ ($a_i \in \mathbb{R}$) において、 a_i は \mathbf{x} に対して一意的に定まる。

証明. \mathbf{x} の線形結合を a_i と a'_i を使って以下のように表せたとする。

$$\begin{aligned} \mathbf{x} &= a_1 \mathbf{x}_1 + \dots + a_n \mathbf{x}_n \\ \mathbf{x} &= a'_1 \mathbf{x}_1 + \dots + a'_n \mathbf{x}_n \end{aligned}$$

これらについて辺々引くと、

$$0 = (a_1 - a'_1)\mathbf{x}_1 + \cdots + (a_n - a'_n)\mathbf{x}_n$$

である。 $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ は線形独立なベクトルの組であるから、 $a_1 - a'_1 = \cdots = a_n - a'_n = 0$ より、 $a_1 = a'_1, \dots, a_n = a'_n$ である。

例 3.25. 基本ベクトルの集合、 $\langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle$ は明らかに $V = \mathbb{R}^n$ の基底であるから、 $\dim V = n$ である。これを \mathbb{R}^n の標準基底という。

例 3.26. $\{0\}$ は線形独立なベクトルを含まないから、 $\dim\{0\} = 0$ である

例 3.27. $V = \mathbb{R}^n$ とする。 $W, W' \subset V$ のとき、 $W \subset W', \dim W \leq \dim W'$ である。

定理 3.28. $V = \mathbb{R}^n$ とする。 $W, W' \subset V$ とする。 $W \subset W'$ のとき、 $\dim W = \dim W' \implies W = W'$ である。

証明. $r \leq n$ とする。 $W = \langle \mathbf{x}_1, \dots, \mathbf{x}_n \rangle, W' = \langle \mathbf{y}_1, \dots, \mathbf{y}_r \rangle$ とする。 $W \subset W', \dim W = \dim W'$ であれば、 W の基底 $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ と W' の基底 $\{\mathbf{y}_1, \dots, \mathbf{y}_r\}$ の個数は一致するから、補題 3.20 により、

$$\mathbf{y}_i = \sum_{1 \leq j \leq n} a_{ij} \mathbf{x}_j \quad (a_{ij} \in \mathbb{R})$$

と表すことができる。これより、 $W = \langle \mathbf{x}_1, \dots, \mathbf{x}_n \rangle = W' = \langle \mathbf{y}_1, \dots, \mathbf{y}_r \rangle$ とすることができる。

例題 3.29. $V = \mathbb{R}^3$ において、 V の部分空間

$$W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x + 2y + z = 0 \right\}$$

の一组の基底を求めよ。

$x + 2y + z = 0$ を x について解くと、 $x = -2y - z$ であるから、 W の元は

$$\begin{pmatrix} -2y - z \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

と表すことができる。ここで、

$$\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

は線形独立であり、 \mathbb{R}^2 を張ることから、基底である。

3.5 基底の変換

$V = \mathbb{R}^n$ の基底として、 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ と $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ があつたとする。これら 2 組の基底が、

$$(\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n) = (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n) P$$

というような関係式であれば、行列 $P = (p_{ij}) \in M_n(\mathbb{R})$ を $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ から $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ への基底の変換行列という。基底の変換行列は次のように求められる。

まず、それぞれの基底は V に属しているので、 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ の各ベクトルは、 $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ の線形結合で表すことができる。

$$\mathbf{b}_j = \sum_{1 \leq i \leq n} p_{ij} \mathbf{a}_i \quad (1 \leq j \leq n)$$

である。これを書き直すと、

$$\begin{aligned} \mathbf{b}_1 &= P_{11}\mathbf{a}_1 + \cdots + P_{n1}\mathbf{a}_n = (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n) \begin{pmatrix} P_{11} \\ \vdots \\ P_{n1} \end{pmatrix} \\ &\vdots \\ \mathbf{b}_n &= P_{1n}\mathbf{a}_1 + \cdots + P_{nn}\mathbf{a}_n = (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n) \begin{pmatrix} P_{1n} \\ \vdots \\ P_{nn} \end{pmatrix} \end{aligned}$$

となり、これらをまとめると

$$(\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n) = (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n) \begin{pmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{pmatrix}$$

だから、

$$(\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n) = (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n) P$$

が得られた。 P は（次元を変えずに）基底を別の基底に移す線形写像であるから、正則である必要がある。ゆえに $\det(P) \neq 0$ で P^{-1} が存在し、

$$(\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n) P^{-1} = (\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n)$$

のように逆写像としても振る舞う。

例題 3.30. \mathbb{R}^3 の 2 つの底を

$$\left\{ \mathbf{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \quad (10)$$

$$\left\{ \mathbf{b}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \mathbf{b}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \quad (11)$$

とするとき、(10) から (11) への基底の変換行列を求めよ。

基底の変換行列を P とすると、 \mathbf{b}_i は \mathbf{a}_i によって以下のように表される。

$$\mathbf{b}_1 = P_{11}\mathbf{a}_1 + P_{21}\mathbf{a}_2 + P_{31}\mathbf{a}_3$$

$$\mathbf{b}_2 = P_{12}\mathbf{a}_1 + P_{22}\mathbf{a}_2 + P_{32}\mathbf{a}_3$$

$$\mathbf{b}_3 = P_{13}\mathbf{a}_1 + P_{23}\mathbf{a}_2 + P_{33}\mathbf{a}_3$$

これらをそれぞれ連立一次方程式の形式で表し、それらを解くと

$$P = \begin{pmatrix} -\frac{2}{3} & \frac{4}{3} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{4}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

が得られる。

4 線形写像

線形写像とはベクトル空間からベクトル空間への（準同型）写像である。

4.1 線形写像の定義

定義 4.1. $V = \mathbb{R}^n, V' = \mathbb{R}^m$ とする。 V から V' への写像を f とし、以下の 2 つの性質を満たす時 f は線形（一次）であるという。

$$(1) f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \quad (\mathbf{x}, \mathbf{y} \in V)$$

$$(2) f(c\mathbf{x}) = cf(\mathbf{x}) \quad (\mathbf{x} \in V, c \in \mathbb{R})$$

これらの性質は f がベクトル空間上の加法とスカラー倍の 2 つの演算を保存することを意味する。（つまり線形性とは、 f が加法とスカラー倍において準同型であるということである。準同型は群の概念であるため、詳しい話は割愛する。詳しく知りたい方は、群のレポート: <https://github.com/noppoMan/math-report/blob/master/group/group.pdf> を読まれることを推奨する。）

特に、 V から V 自身への線形写像を V の線形変換 (linear transformation) または 一次変換 という

例 4.2. $V = \mathbb{R}^3, V' = \mathbb{R}^2, V'' = \mathbb{R}$ とするとき写像

$$\begin{array}{ccc} V & \xrightarrow{f_1} & V' \\ \cup & & \cup \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} & \mapsto & \begin{pmatrix} x_1 - x_2 \\ x_1 - x_3 \end{pmatrix} \end{array}$$

$$\begin{array}{ccc} V & \xrightarrow{f_2} & V'' \\ \Downarrow & & \Downarrow \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} & \longmapsto & (x_1 + x_2 + x_3) \end{array}$$

などは線形である。 f_2 のように \mathbb{R} への線形写像を線形汎関数ともいう。

例 4.3. \mathbb{R}^2 (平面ベクトル) の原点を中心とした時計、反時計周りに θ だけ回転させるような写像は線形変換である。 $(\mathbb{R}^2$ から \mathbb{R}^2 への線形写像)

$$f_3: \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}$$

この線形変換は三角関数の加法定理から導かれる。

$$\begin{cases} r\cos(\alpha + \beta) = r\cos\alpha\cos\beta - \sin\alpha\sin\beta \\ r\sin(\alpha + \beta) = r\sin\alpha\cos\beta + \cos\alpha\sin\beta \end{cases} \quad \begin{matrix} (12) \\ (13) \end{matrix}$$

変換前の点 $P(x, y)$ 、変換後の点 $P'(r\cos(\alpha + \beta), r\sin(\alpha + \beta))$ とする。

P の座標 (x, y) を (12), (13) にそれぞれ代入して、

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

したがって、これを行列表示すると

$$\begin{pmatrix} x\cos\beta & -y\sin\beta \\ x\sin\beta & y\cos\beta \end{pmatrix}$$

となる。

4.2 線形写像と行列の対応

一般に $A = (a_{ij}) \in M_{m,n}(\mathbb{R})$ が与えられた時、 $V = \mathbb{R}^n, V' = \mathbb{R}^m$ への写像 f_A を

$$f_A = \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto A\mathbf{x} = \begin{pmatrix} \sum_{1 \leq j \leq n} a_{1j}x_j \\ \vdots \\ \sum_{1 \leq j \leq n} a_{mj}x_j \end{pmatrix}$$

によって定義すれば、 f_A は線形である。つまり、

$$\begin{aligned} f_A(\mathbf{x} + \mathbf{y}) &= A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} \quad (\mathbf{x}, \mathbf{y} \in V) \\ f_A(c\mathbf{x}) &= A(c\mathbf{x}) = cA\mathbf{x} \quad (c \in \mathbb{R}, \mathbf{x} \in V) \end{aligned}$$

が成立する。

したがって、上で挙げた例に登場した f_1, f_2, f_3 はそれぞれ行列

$$A_1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \in M_{2,3}(\mathbb{R}), \quad A_2 = (1, 1, 1) \in M_{1,3}(\mathbb{R}), \quad A_3 = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

に対応する線形写像になっている。

4.2.1 表現行列

$V = \mathbb{R}^n, V' = \mathbb{R}^m$ とする。 $f: V \rightarrow V'$ を線形写像とする。 $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}, \{\mathbf{e}'_1, \dots, \mathbf{e}'_m\}$ をそれぞれ V, V' の標準基底とし、

$$f(\mathbf{e}_j) = \sum_{1 \leq i \leq m} a_{ij} \mathbf{e}'_i$$

とする。 $f(\mathbf{e}_j) = \mathbf{a}_j$ とすれば、この関係式は以下の写像 f

$$\begin{array}{ccc} V & \xrightarrow{f} & V' \\ \Psi & & \Psi \\ \mathbf{e}_1 & \mapsto & \mathbf{a}_1 \\ \mathbf{e}_2 & \mapsto & \mathbf{a}_2 \\ \vdots & & \vdots \\ \mathbf{e}_n & \mapsto & \mathbf{a}_n \end{array}$$

を定義する。つまり、

$$f(\mathbf{e}_1) = \mathbf{a}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, f(\mathbf{e}_2) = \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, f(\mathbf{e}_n) = \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

と n 個のベクトルが得られる。これを並べると、

$$A = (\mathbf{a}_1, \dots, \mathbf{a}_n) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

のような行列 A が得られる。つまり、

$$f(\mathbf{e}_1) = A\mathbf{e}_1, f(\mathbf{e}_2) = A\mathbf{e}_2, \dots, f(\mathbf{e}_n) = A\mathbf{e}_n$$

である。

ここで、 V の任意のベクトル \mathbf{x} は

$$\mathbf{x} = \sum_{1 \leq j \leq n} x_j \mathbf{e}_j = x_1 \mathbf{e}_1 + \cdots x_n \mathbf{e}_n$$

で表すことができることを思い出してほしい。これらを利用すると以下の関係式が得られる。

$$\begin{aligned}
 f(\mathbf{x}) &= f\left(\sum_{1 \leq j \leq n} x_j \mathbf{e}_j\right) \\
 &= f(x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n) \\
 &= x_1 f(\mathbf{e}_1) + \cdots + x_n f(\mathbf{e}_n) \\
 &= x_1 A \mathbf{e}_1 + \cdots + x_n A \mathbf{e}_n \\
 &= A x_1 \mathbf{e}_1 + \cdots + A x_n \mathbf{e}_n \\
 &= A(x_1 \mathbf{e}_1 + \cdots + x_n \mathbf{e}_n) \\
 &= A \mathbf{x}
 \end{aligned}$$

つまり、 f は $A = (a_{ij})$ に対応する線形写像 f_A と一致するということである。

定義 4.4. V から V' への線形写像を f とする。 V の基底 $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, V' の基底 $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ に関して、

$$(f(\mathbf{e}_1), f(\mathbf{e}_2), \dots, f(\mathbf{e}_n)) = (\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n) A$$

の形に表される時、行列 A を f の表現行列と呼ぶ。

命題 4.5. 表現行列は一意的である

証明. 定理 3.24 より明らか。表現行列を表すベクトル a_1, \dots, a_n はすべて一意的に表せるので、表現行列も一意的である。

例題 4.6. \mathbb{R}^2 のベクトル \mathbf{x}, \mathbf{y} と線形変換 f の標準基底での表現行列 A を

$$\mathbf{x} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, A = \begin{pmatrix} -3 & 1 \\ 2 & -2 \end{pmatrix}$$

とする。この線形変換 f で $f(a\mathbf{x} + b\mathbf{y}) = a'\mathbf{x} + b'\mathbf{y}$ となるとき、

$$B \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a' \\ b' \end{pmatrix}$$

となる行列を求めよ。

\mathbf{x}, \mathbf{y} は \mathbb{R}^2 を張る標準基底 $\mathbf{e}_1, \mathbf{e}_2$ で

$$(\mathbf{x} \ \mathbf{y}) = (\mathbf{e}_1 \ \mathbf{e}_2) \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$$

と表すことができる。

$$P = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$$

とする。

与式 $f(a\mathbf{x} + b\mathbf{y}) = a'\mathbf{x} + b'\mathbf{y}$ より、

$$f(a\mathbf{x} + b\mathbf{y}) = A(\mathbf{x} \ \mathbf{y}) \begin{pmatrix} a \\ b \end{pmatrix} = AP \begin{pmatrix} a \\ b \end{pmatrix}$$

また、

$$a'x + b'y = (x \ y) \begin{pmatrix} a' \\ b' \end{pmatrix} = P \begin{pmatrix} a' \\ b' \end{pmatrix}$$

の関係式が得られる。これより、

$$P \begin{pmatrix} a' \\ b' \end{pmatrix} = AP \begin{pmatrix} a \\ b \end{pmatrix}$$

である。これに左から P^{-1} を掛けて（ここでは、 P が可逆であることを前提に解答を進める。可逆の判定は行列式の章で解説する。）

$$P^{-1}P \begin{pmatrix} a' \\ b' \end{pmatrix} = P^{-1}AP \begin{pmatrix} a \\ b \end{pmatrix} \iff \begin{pmatrix} a' \\ b' \end{pmatrix} = P^{-1}AP \begin{pmatrix} a \\ b \end{pmatrix}$$

つまり、

$$P^{-1}AP = \frac{1}{2 \cdot 3 - 5 \cdot 1} \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} -3 & -1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ -7 & -8 \end{pmatrix}$$

この例題より以下の定理が得られる。

定理 4.7. 標準基底での \mathbb{R}^n 上の線形変換 f の表現行列を A とする。標準基底 $\{e_1, \dots, e_n\}$ を、基底 $\{x_1, \dots, x_n\}$ に取り替えたとする。基底の変換行列を P とおくと、基底 $\{x_1, \dots, x_n\}$ を座標系とした新座標での線形変換 f の表現行列は $P^{-1}AP$ である。

定義 4.8. 実数係数を持つ x の多項式全体の集合も、通常の加法・スカラー倍に関してベクトル空間である。これを

$$\mathbb{R}[x]_n = \{a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \mid a_i \in \mathbb{R} (1 \leq i \leq n)\}$$

と表記する。

例 4.9. 多項式 $f(x) = a_0 + a_1 x + a_2 x^2$ はベクトルどうしの内積により次のように表せる。

$$f(x) = \begin{pmatrix} 1 & x & x^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = a_0 + a_1 x + a_2 x^2$$

例題 4.10. $\mathbb{R}[x]_4 = \{a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + a_0 \mid a_i \in \mathbb{R}\}$ を実数係数の 4 次以下の多項式全体とする。線形写像 $F: \mathbb{R}[x]_4 \rightarrow \mathbb{R}[x]_4$ を

$$F(f) = \frac{df(x)}{dx}$$

とするとき、基底 $\{1, x, x^2, x^3, x^4\}$ に関する F の表現行列を求めよ。

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix}$$

を \mathbb{R}^5 の標準基底とする。写像 $g: \mathbb{R}^5 \rightarrow \mathbb{R}[x]_4$ を

$$g(\mathbf{a}) = a_1 + a_2 \cdot x + a_3 \cdot x^2 + a_4 \cdot x^3 + a_5 \cdot x^4$$

と定めると、これは \mathbb{R}^5 の標準基底に $\mathbb{R}[x]_4$ の基底 $\{1, x, x^2, x^3, x^4\}$ を対応させる線形写像である。線形写像 g に対し、 F を適用すると

$$\begin{aligned} F(g(\mathbf{a})) &= F(a_1 + a_2 \cdot x + a_3 \cdot x^2 + a_4 \cdot x^3 + a_5 \cdot x^4) \\ &= a_1 F(1) + a_2 F(x) + a_3 F(x^2) + a_4 F(x^3) + a_5 F(x^4) \\ &= a_1 \cdot 0 + a_2 \cdot 1 + a_3 \cdot 2x + a_4 \cdot 3x^2 + a_5 \cdot 4x^3 \end{aligned}$$

となる。この結果を書き直すと

$$(1 \quad x \quad x^2 \quad x^3 \quad x^4) \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{a}$$

となっているので、 F の表現行列

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

が得られる。この表現行列より、

$$(F(1), F(x), F(x^2), F(x^3), F(x^4)) = (1, x, x^2, x^3, x^4)F$$

の関係式が成立する。

4.3 線形写像の合成と行列の積

線形写像 f_A と f_B があるとき、 $f_A \circ f_B = f_{AB}$ が成立する。実際、任意の $\mathbf{x} \in \mathbb{R}^n$ に対して

$$f_A \circ f_B(\mathbf{x}) = f_A(f_B(\mathbf{x})) = A(B\mathbf{x}) = (AB)\mathbf{x} = f_{AB}(\mathbf{x})$$

であるから、 $f_A \circ f_B = f_{AB}$ が得られた。

この関係式を具体的な線形写像 f_A, f_B を使って表すと面白いことが分かる。

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

を表現行列とし、線形写像 f_A と f_B の合成を考える。

$$\begin{aligned} f_A \circ f_B(\mathbf{x}) &= f_A(f_B(\mathbf{x})) = f_A\left(\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\right) \\ &= f_A\left(\begin{pmatrix} xb_{11} + yb_{12} \\ xb_{21} + yb_{22} \end{pmatrix}\right) \\ &= \begin{pmatrix} a_{11} + a_{12} \\ a_{21} + a_{22} \end{pmatrix} \begin{pmatrix} xb_{11} + yb_{12} \\ xb_{21} + yb_{22} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}(xb_{11} + yb_{12}) + a_{12}(xb_{21} + yb_{22}) \\ a_{21}(xb_{11} + yb_{12}) + a_{22}(xb_{21} + yb_{22}) \end{pmatrix} \\ &= \begin{pmatrix} (a_{11}b_{11} + a_{12}b_{21})x + (a_{11}b_{12} + a_{12}b_{22})y \\ (a_{21}b_{11} + a_{22}b_{21})x + (a_{21}b_{12} + a_{22}b_{22})y \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= (AB)\mathbf{x} \end{aligned}$$

が得られる。つまり、行列の積は線形写像の合成に等しいということが分かったのである。

したがって、行列の積は線形写像の合成に対応するように定義されていたのだ。

4.4 線形写像の核と像

$V = \mathbb{R}^n, V' = \mathbb{R}^m$ とし、線形写像 $f: V \rightarrow V'$ が与えられたとする。 f によって、 V' のベクトル $\mathbf{0}_{V'}$ に移される V のベクトル全体の集合

$$\ker(f) = \{\mathbf{x} \in V \mid f(\mathbf{x}) = \mathbf{0}\}$$

を f の核 (Kernel, カーネル) という。また、 V の f による像全体の集合

$$\text{Im}(f) = \{f(\mathbf{x}) \mid \mathbf{x} \in V\}$$

を f の像 (Image, イメージ) という。

この定義から、 f が全射 $\iff \text{Im}(f) = V'$ である。また、 f が単射であれば、 $\ker(f) = \{\mathbf{0}\}$ である。これを証明する。

f が単射であるとする。 $\forall \mathbf{x} \in \ker(f)$ をとると、 $f(\mathbf{x}) = \mathbf{0}$ 。ここで、 f は線形写像より、 $f(\mathbf{0}) = \mathbf{0}$ 。よって、 $f(\mathbf{x}) = f(\mathbf{0})$ 。仮定より、 f は単射なので $\mathbf{x} = \mathbf{0}$ 。よって、 $\ker(f) = \{\mathbf{0}\}$ となる。

逆に、 $f(\mathbf{x}) = f(\mathbf{y})$ となる $\forall \mathbf{x}, \mathbf{y} \in V$ に対し、 $f(\mathbf{x}) = f(\mathbf{y}) = \mathbf{0}$ 。ここで f は線形写像なので、 $f(\mathbf{x} - \mathbf{y}) = \mathbf{0}$ が言え、 $\mathbf{x} - \mathbf{y} \in \ker(f) = \{\mathbf{0}\}$ である。したがって、 $\mathbf{x} - \mathbf{y} = \mathbf{0}$ より、 $\mathbf{x} = \mathbf{y}$ 。

定理 4.11. $\text{Im}(f), \ker(f)$ はそれぞれ V', V の部分空間である。

証明. $\text{Im}(f)$ が V' の部分空間であることを示す。 $\mathbf{x}', \mathbf{y}' \in \text{Im}(f)$ とすれば、ある $\mathbf{x}, \mathbf{y} \in V$ があって、 $\mathbf{x}' = f(\mathbf{x}), \mathbf{y}' = f(\mathbf{y})$ と表せる。よって、

$$\begin{aligned}\mathbf{x}' + \mathbf{y}' &= f(\mathbf{x}) + f(\mathbf{y}) = f(\mathbf{x} + \mathbf{y}) \in \text{Im}(f) \\ c\mathbf{x}' &= cf(\mathbf{x}) = f(c\mathbf{x}) \in \text{Im}(f)\end{aligned}$$

より、 $\text{Im}(f)$ は V' の部分空間である。

$\ker(f)$ が V の部分空間であることを示す。 $\mathbf{x}, \mathbf{y} \in \ker(f)$ とすれば、 $f(\mathbf{x}) = f(\mathbf{y}) = \mathbf{0}$ であるから

$$\begin{aligned}f(\mathbf{x} + \mathbf{y}) &= f(\mathbf{x}) + f(\mathbf{y}) = \mathbf{0} \in \ker(f) \\ cf(\mathbf{x}) &= f(c\mathbf{x}) = \mathbf{0} \in \ker(f)\end{aligned}$$

より、 $\ker(f)$ は V の部分空間である。

4.5 次元定理

定理 4.12. $V = \mathbb{R}^n, \dim V = n$ のとき、次の等式が成立する。

$$\dim \text{Im}(f) = n - \dim \ker(f)$$

証明. $\dim \ker(f) = s$ とし、 $\{\mathbf{x}_1, \dots, \mathbf{x}_s\}$ を $\ker(f)$ の 1 つの基底とする。 $s \leq n$ で V の元 $\mathbf{x}_{s+1}, \dots, \mathbf{x}_n$ を適当に選び、 $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ を V の基底とすることができる。このとき、 $f(\mathbf{x}_i) = \mathbf{0}$ ($1 \leq i \leq s$) であるから、

$$\text{Im}(f) = \langle f(\mathbf{x}_1), \dots, f(\mathbf{x}_n) \rangle = \langle f(\mathbf{x}_{s+1}), \dots, f(\mathbf{x}_n) \rangle$$

である。 $f(\mathbf{x}_{s+1}), \dots, f(\mathbf{x}_n)$ が線形独立であれば、このベクトルの組が $\text{Im}(f)$ の基底となり、 $\dim \text{Im}(f) = n - \dim \ker(f)$ が証明される。 $f(\mathbf{x}_{s+1}), \dots, f(\mathbf{x}_n)$ が線形独立であることを示す。

$$\sum_{s+1 \leq j \leq n} c_j f(\mathbf{x}_j) = \mathbf{0} \quad (c_j \in \mathbb{R})$$

と仮定すると、 f は線形より、

$$f\left(\sum_{s+1 \leq j \leq n} c_j \mathbf{x}_j\right) = \mathbf{0}$$

であるから、

$$\mathbf{y} = \sum_{s+1 \leq j \leq n} c_j \mathbf{x}_j \in \ker(f)$$

したがって、

$$\sum_{s+1 \leq j \leq n} c_j \mathbf{x}_j = \sum_{1 \leq i \leq s} k_i \mathbf{x}_i \quad (k_i \in \mathbb{R})$$

今、 $\mathbf{y} \in \ker(f)$ より $\mathbf{y} = \langle \mathbf{x}_1, \dots, \mathbf{x}_s \rangle$ とかける。上式を移項して

$$\sum_{s+1 \leq j \leq n} c_j \mathbf{x}_j - \sum_{1 \leq i \leq s} k_i \mathbf{x}_i = 0$$

となり、 $\mathbf{x}_1, \dots, \mathbf{x}_n$ が線形独立であることから、 $c_j, k_i = 0$ より $f(\mathbf{x}_{s+1}), \dots, f(\mathbf{x}_n)$ は線形独立となる。
また、この定理により次の等式が成立することが直ちに分かる。

$$\dim V = \dim \ker(f) + \dim \operatorname{Im}(f)$$

例題 4.13.

$$A = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

によって定義される線形写像 $f: \mathbf{x} \mapsto A\mathbf{x}$ ($\mathbf{x} \in \mathbb{R}^3$) に対する $\ker(f)$ と $\operatorname{Im}(f)$ とそれぞれの次元を求めよ。

$\ker(f)$ とその次元を求める。 f のカーネルは f に写すと $\mathbf{0}$ になるベクトルの集合であるから、 $A\mathbf{x} = \mathbf{0}$ 。つまり

$$\begin{cases} -2x + y + z = 0 \\ x - 2y + z = 0 \\ x + y - 2z = 0 \end{cases}$$

を意味する。この連立方程式の解は $x = y = z$ であるから

$$\ker(f) = \left\{ \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x = y = z \right\}$$

である。ここから基底を取り出すことを考える。 \mathbf{x} の成分はすべて同じ値なので、いずれかの成分を固定する（例えば、 x で y, z を置き換える）ことで

$$\mathbf{x} = x \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

と表すことができる。この 1 つのベクトル \mathbf{x} は線形独立であり、カーネルを張る基底であるから、 $\dim \ker(f) = 1$ である。

次に、 $\operatorname{Im}(f)$ とその次元を求める。次元定理を使えば、 $\dim \operatorname{Im}(f) = 2$ となるはずであるから、

$$\operatorname{Im}(f) = \left\{ \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x + y + z = 0 \right\}$$

と仮定する。ここで、 x を $x = -y + (-)z$ とおけば y と z で \mathbf{x} を表現できるから、

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

より、2つの線形独立なベクトル

$$\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

が得られる。これらは \mathbb{R}^2 を張る基底となるから、仮定は正しいことが証明された。

4.6 行列の階数

定義 4.14. 線形写像 $f: V \rightarrow V'$ に対し、 $\dim \operatorname{Im}(f)$ を f の階数 (rank, ランク) といい、 $\operatorname{rank} f$ とかく。 $\operatorname{rank} f = \dim \operatorname{Im}(f)$ である。

また、 $A \in M_{m,n}(\mathbb{R})$ に対してその階数を $\operatorname{rank} A = \operatorname{rank} f_A = \dim A\mathbb{R}^n$ と定義する。したがって、 $\operatorname{rank} A$ は A の列ベクトルの中で線形独立なものの最大個数である。

なお、 $A\mathbb{R}^n$ の定義は次である。 $V = \mathbb{R}^n$ とする。 $f = A$ であることに着目すると

$$\operatorname{Im}(f) = f(V) = \{f(\mathbf{x}) \mid \mathbf{x} \in V\} = AV = \{A\mathbf{x} \mid \mathbf{x} \in V\} \subset V'$$

例 4.15. 例題 4.13 の行列 A のランクは $\operatorname{rank} A = 2$ である。実際、定義 4.14 に従えば、 $\dim \operatorname{Im}(f) = 2$ なので、 $\operatorname{rank} f = \dim \operatorname{Im}(f)$ である。

例 4.16. $A = \mathbf{x}^t \mathbf{y}, \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^n, \mathbf{x}, \mathbf{y} \neq \mathbf{0}$ とすれば、 $\operatorname{rank} A = 1$ である。

$\mathbf{y} = (y_i)$ とすれば、 A の列ベクトルは $y_1 \mathbf{x}, \dots, y_n \mathbf{x}$ であるから、

$$\operatorname{rank} A = \dim \langle y_1 \mathbf{x}, \dots, y_n \mathbf{x} \rangle = \dim \langle \mathbf{x} \rangle = 1$$

である。

定理 4.17. $A \in M_{m,n}(\mathbb{R}), f_A: V \rightarrow V'$ とすれば、

- (1) $\operatorname{rank} A \leq \min(m, n)$
- (2) $\operatorname{rank} A = m \iff f$ は全射 (このとき、 $n \geq m$)
- (3) $\operatorname{rank} A = n \iff f$ は単射 (このとき、 $n \leq m$)

証明. (1) $AV \subset V'$ であるから、 $\dim AV \leq \dim V' = m$. よって、 $\operatorname{rank} A \leq m$ である。しかし、 $\dim AV = n - \dim \ker(f)$ より、 $\dim AV \leq n$. よって、 $\operatorname{rank} A \leq \min(m, n)$

(2) $\operatorname{rank} A = m$ ならば、

$$AV \subset V', \dim AV = \dim V'$$

であるから、定理 3.28 より $AV = V'$ となり、 f_A は全射である。逆に f_A が全射であるとすれば、 $n \geq m$ より、 $\operatorname{rank} A = m$ は明らか。

(3) $\operatorname{rank} A = n$ ならば、定理 4.12 より

$$\dim \ker(f) = 0$$

したがって、 $\ker(f) = \{0\}$. つまり、 f_A は単射である。逆に f_A が単射とする。 $n \leq m$ より、 $\operatorname{rank} A = n$ は明らか。

例 4.18. $\text{rank } A = m$ で f は全射の例

$f: V \rightarrow V'$ を以下の写像とすれば、 f は全射である。

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in M_{2,3}(\mathbb{R})$$

$$\begin{array}{ccc} \mathbb{R}^3 & \xrightarrow{f=A} & \mathbb{R}^2 \\ \Downarrow & & \Downarrow \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} & \longmapsto & \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{pmatrix} \end{array}$$

例 4.19. $\text{rank } A = n$ ならば f は単射の例

$f: V \rightarrow V'$ を以下の写像とすれば、 f は単射である。

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in M_{3,2}(\mathbb{R})$$

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{f=A} & \mathbb{R}^3 \\ \Downarrow & & \Downarrow \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \longmapsto & \begin{pmatrix} x_1 + 4x_2 \\ 2x_1 + 5x_2 \\ 3(x_1 + x_2) + 6(x_1 + x_2) \end{pmatrix} \end{array}$$

4.7 基本変形

具体的な行列の階数は基本変形を用いて求めることができる。ここでは基本変形の定義とそれを使った行列の階数の求め方を示す。

$A = (a_{ij}) \in M_{m,n}(\mathbb{R})$ とする。 A に対する基本変形 (初等変形) とは以下の 3 つの操作とその組み合わせのことである。

- (1) A の i 行 (列) を定数倍する。

$$A \longrightarrow \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ ca_{i1} & ca_{i2} & \cdots & ca_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad (c \in \mathbb{R})$$

の変換を指す。

- (2) A の第 i 行 (列) と第 j 行 (列) を入れ替える。

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

の変換を指す。

(3) A の第 i 行 (列) に第 j 行 (列) の定数倍を加える。

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} + ca_{j1} & a_{i2} + ca_{j2} & \cdots & a_{in} + ca_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad (c \in \mathbb{R})$$

の変換を指す。

行に関する基本変形を行基本変形、列に関する基本変形を列基本変形という。

この基本変形を適当な組み合わせで繰り返し A に適用することで、 A を標準形

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

(ここで、 E_r は A の階数における単位行列である) に変換することができる。 E_r が階数を表すので、標準形に変換する行為は行列の階数を求めることにほかならない。

例 4.20. 行列 $A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}$ の階数を基本変形を用いて求める方法は以下である。

$$\begin{aligned} \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix} &\xrightarrow{\text{第 1 列と第 2 列を入れ替え}} \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 3 \\ -3 & -2 & 0 \end{pmatrix} \\ &\xrightarrow{\text{第 1 行を 3 倍して第 3 行に加える}} \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 3 \\ 0 & -2 & 6 \end{pmatrix} \xrightarrow{\text{第 1 列を -2 倍して第 3 列に加える}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & -2 & 6 \end{pmatrix} \\ &\xrightarrow{\text{第 2 行に -1 倍する}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & -2 & 6 \end{pmatrix} \xrightarrow{\text{第 2 行を -2 倍し、第 3 行に加える}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

この基本変形より、 $\text{rank } A = 2$ である。

この例より、行列の基本変形は連立方程式の変形であることが分かる。よって、連立方程式の変形の際に出来ないような操作 (e.g. 行基本変形と列基本変形が同時に行われる) は行基本変形においても行うことが出来ない。

4.8 基底の延長

$V = \mathbb{R}^n$ とする。 $W \subset V, \text{rank } W = r$ とする。このとき W を張る基底を $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ とすると、 V の基底は $\{\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$ ($\{\mathbf{v}_i\} \in V$) と表すことができる。これを基底の延長という。これは以下により求められる。

$A = (\mathbf{a}_1, \dots, \mathbf{a}_r) \in M_r(\mathbb{R})$ とする。 A の横に n 次単位行列 E を並べて

$$B = (AE) = (\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{e}_1, \dots, \mathbf{e}_n) \in M_{n, (r+n)}(\mathbb{R})$$

を作る。このとき、 B の階数は E の階数と同様であるから $\text{rank } B = n$ である。ここで、 B を基本変形すると先頭から r 個までのベクトル $\{\mathbf{a}_i\}$ は W の基底より、 W の標準基底 $\{\mathbf{e}_1, \dots, \mathbf{e}_r\}$ に変形し

$$C = (\mathbf{e}_1, \dots, \mathbf{e}_r, \mathbf{v}_1, \dots, \mathbf{v}_n)$$

とすることができる。残りの $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形独立ではないが、 B の階数が n であることから、 $\mathbf{e}_1, \dots, \mathbf{e}_r$ とともに V を張る $n - r$ 個の線形独立なベクトルが含まれていなければならない。そのようなベクトルを $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ として選び出すと

$$V = \langle \mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n \rangle$$

として、 W の基底を延長し、 V を作ることができる。

5 行列式

5.1 2 次行列における行列式

定義 5.1. 2 次行列における行列式を次のように定義する。

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

に対して、

$$\det(A) = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

である。また、 A の列ベクトル

$$\mathbf{a}_1 = \begin{pmatrix} a \\ c \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} b \\ d \end{pmatrix}$$

に対して、

$$\det(A) = D(\mathbf{a}_1, \mathbf{a}_2)$$

とする。

5.1.1 行列式の性質

上で定義した行列式 D を $V = \mathbb{R}^2$ 上の 2 変数関数と考えると、次の性質を持つ。

(1) $\mathbf{a}_1, \mathbf{a}_2$ について線形 (多重線形性)

\mathbf{a}_1 について

$$\begin{aligned} D(\mathbf{a}_1 + \mathbf{a}'_1, \mathbf{a}_2) &= D(\mathbf{a}_1, \mathbf{a}_2) + D(\mathbf{a}'_1, \mathbf{a}_2) \\ D(c\mathbf{a}_1, \mathbf{a}_2) &= cD(\mathbf{a}_1, \mathbf{a}_2) \quad (c \in \mathbb{R}) \end{aligned}$$

\mathbf{a}_2 について

$$\begin{aligned} D(\mathbf{a}_1, \mathbf{a}_2 + \mathbf{a}'_2) &= D(\mathbf{a}_1, \mathbf{a}_2) + D(\mathbf{a}_1, \mathbf{a}'_2) \\ D(\mathbf{a}_1, c\mathbf{a}_2) &= cD(\mathbf{a}_1, \mathbf{a}_2) \quad (c \in \mathbb{R}) \end{aligned}$$

(2) $\mathbf{a}_1, \mathbf{a}_2$ について交代的

$$D(\mathbf{a}_2, \mathbf{a}_1) = -D(\mathbf{a}_1, \mathbf{a}_2)$$

これは以下のように展開できる。

$$D(\mathbf{a}_2, \mathbf{a}_1) = \begin{vmatrix} b & a \\ d & c \end{vmatrix} = bc - ad = -D(\mathbf{a}_1, \mathbf{a}_2)$$

(3) \mathbb{R}^2 の単位ベクトルを $\mathbf{e}_1, \mathbf{e}_2$ とすれば、

$$D(\mathbf{e}_1, \mathbf{e}_2) = 1$$

これは以下のように展開できる。

$$D(\mathbf{e}_1, \mathbf{e}_2) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

この性質から 2 次行列式はこれらの性質によって導くことができる。 \mathbb{R}^2 上の 2 変数関数 $D(\mathbf{a}_1, \mathbf{a}_2)$ が (1), (2), (3) の性質を持つとすると、(1) より

$$\begin{aligned} D(\mathbf{a}_1, \mathbf{a}_2) &= D(a\mathbf{e}_1 + c\mathbf{e}_2, b\mathbf{e}_1 + d\mathbf{e}_2) \\ &= aD(\mathbf{e}_1, b\mathbf{e}_1 + d\mathbf{e}_2) + cD(\mathbf{e}_2, b\mathbf{e}_1 + d\mathbf{e}_2) \\ &= abD(\mathbf{e}_1, \mathbf{e}_1) + adD(\mathbf{e}_1, \mathbf{e}_2) + cbD(\mathbf{e}_2, \mathbf{e}_1) + cdD(\mathbf{e}_2, \mathbf{e}_2) \end{aligned}$$

(2), (3) より

$$D(e_1, e_2) = 1, \quad D(e_2, e_1) = -D(e_1, e_2) = -1$$

$D(e_1, e_1) = -D(e_1, e_1)$ であるから、 $D(e_1, e_1) = 0$. 同様に、 $D(e_2, e_2) = 0$ となる。
したがって、(1) で得た関係式に代入すると

$$\begin{aligned} D(a_1, a_2) &= abD(e_1, e_1) + adD(e_1, be_2) + cbD(e_2, e_1) + cdD(e_2, e_2) \\ &= ab \cdot 0 + ad \cdot -1 + cb \cdot +1 + cd \cdot 0 \\ &= ad - bc \end{aligned}$$

である。

5.2 置換

前節で、2 変数の関数 $D(a_1, a_2)$ を紹介したが、一般には 2 次行列式の性質 (1),(2),(3) と類似の性質を持つ n 変数の関数

$$D(a_1, a_2, \dots, a_n) \quad (a_1, a_2, \dots, a_n \in \mathbb{R}^n)$$

が存在する。これを n 次の行列式という。ここでは、この n 次の行列式を導くために必要な置換について概説する。

n 個の文字の集合

$$X = \{1, 2, \dots, n\}$$

を考え、 X のそれ自身への全単射

$$\sigma : X \rightarrow X$$

を X の置換 (permutation) という。(X は有限集合として考える。) この n 文字の置換の集合を S_n とかく。任意の置換 $\sigma \in S_n$ が $1, 2, \dots, n$ から $i_1, i_2, \dots, i_n \in 1$ 対 1 で対応しているとき、つまり $\sigma(k) = i_k \quad (1 \leq k \leq n)$ のとき

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

と書く。特に、 i_1, i_2, \dots, i_n が自分自身であるような置換

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

を恒等置換という。これを単に 1 と表す。

置換に対し、 i, j の 2 つの文字のみを入れ替え、その他の文字は不変にするような置換

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

を互換 (transposition) といい、 (i, j) とかく。具体的には S_4 であれば、

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

のように書き、不変な置換 $\sigma(1) = 1, \sigma(4) = 4$ は省略することができる。この書き方の理由として、 $\sigma(i) = j, \sigma(j) = i$ と $i \rightarrow j, j \rightarrow i$ のように i と j の対応のみが交換されており、その他は不変であることから、 i, j の 2 つの文字だけで σ を表現するに十分だからである。

また、2 つの置換 $\sigma, \gamma \in S_n$ に対し、その合成

$$(\sigma\gamma)(k) = (\sigma(\gamma(k)))$$

によって定義される写像 $\sigma\gamma$ もまた 1 つの置換である。これを σ, γ の積という。また $\sigma \in S_n$ は全単射であるから、その逆写像 σ^{-1} が存在し、それもまた置換となる。これを σ の逆置換という。

例 5.2. 置換の積の例

$$\gamma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

の置換の積を考えると、

$$\gamma(\sigma(1)) = 1, \quad \gamma(\sigma(2)) = 3, \quad \gamma(\sigma(3)) = 2$$

より、

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

となる。

例題 5.3. S_3 のすべての置換を答えよ。

解答

置換は自身から自身への対応の並び替えであるから、その総数は $n!$ 個である。よって、 S_3 では、3! 個の置換がありそれは以下となる。

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

5.2.1 互換の積

$\sigma \in S_3$ を

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

とする。この置換はいくつかの互換の積で

$$\sigma = (1, 3)(1, 2) = (1, 2)(2, 3)$$

のように表すことができる。実際、

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

であるから $\gamma = (1, 2), \rho = (2, 3)$ としてその合成 $\gamma\rho$ を考えると

$$\gamma(\rho(1)) = \gamma(1) = 2$$

$$\gamma(\rho(2)) = \gamma(3) = 3$$

$$\gamma(\rho(3)) = \gamma(3) = 1$$

となることから

$$\gamma\rho = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \sigma$$

である。(互換の合成も写像の合成同様に、右側の置換(互換)から評価する。)

例題 5.4. $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ を互換の積で表わせ。

解答

この置換は

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

および

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

の互換の積で

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

のように表すことができる。

実は、任意の $\sigma \in S_n$ は表示は一定でないにしろ、互換の個数の偶奇は一定であることが知られている。それを示すために、次の補題を証明する。

補題 5.5. S_n 上に定義された関数 $sign(\sigma)$ が次の性質のもと一意に存在する。

- (1) $sign(\sigma) \in \{\pm 1\}$
- (2) σ が互換ならば、 $sign(\sigma) = -1$
- (3) $sign(\sigma\gamma) = sign(\sigma)sign(\gamma)$

証明. このような関数は

$$sign(\sigma) = \prod_{i \neq j} \frac{\sigma(i) - \sigma(j)}{i - j}$$

によって与えられる。

$\sigma \in S_n$ とする。 σ を互換とする。 σ は互換より、 $i \neq j$ のとき、 $\sigma(k) = k, \sigma(j) = i, \sigma(i) = j$. それ以外、つまり $i = j$ ですべての $sign(k) = k$ であるから、右辺の積は

$$\frac{i - j}{i - j} = 1$$

となる。これより、 $i \neq j$ のときに

$$\frac{j - i}{i - j} = -1$$

となることから、(2) を満たす。

したがって、 σ が k 個の互換の積 τ_1, \dots, τ_k の積に等しければ、関数 $sign$ は次のように展開できる。

$$sign(\sigma) = sign(\tau_1) \cdots sign(\tau_k) = (-1)^k \quad (k \in \mathbb{N})$$

これより (1) を満たす。(この関数は well-defined であることが分かる。) また、 $(\tau_1, \dots, \tau_k) = \sigma$ より、 $sign(\tau_1, \dots, \tau_k) = sign(\tau_1) \cdots sign(\tau_k)$ であるから、(3) を満たすことも確かめられる。

$sign(\sigma) = 1$ のとき互換の数は偶数、 $sign(\sigma) = -1$ のとき互換の数は奇数である。 k の偶奇により、 σ を偶置換、奇置換という。

この補題より、次の定理が証明される。

定理 5.6. 任意の $\sigma \in S_n$ は高々 $n - 1$ 個の互換の積に表される。このような表示は一意的ではないが、互換の個数の偶奇は一定である。

第一の主張は n の帰納法で容易に証明され、第二の主張は補題 5.5 により証明される。

例 5.7. $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ は偶置換、 $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ は奇置換である。

このように $\sigma \in S_n$ の偶奇の個数は一致する。(証明は割愛する)

5.2.2 巡回置換

相異なる文字 i_1, i_2, \dots, i_k を i_1 を i_2 に i_2 を i_3 に... と順次次の文字に移し、最後の i_k を i_1 に戻すような(ループを描くような)置換を i_1, i_2, \dots, i_k の巡回置換 (cycle) といい、 (i_1, i_2, \dots, i_k) とかく。互換はこの特別の場合である。よって、巡回置換は

$$(i_1, i_2, \dots, i_k) = (i_1, i_k)(i_1, i_{k-1}) \cdots (i_1, i_2)$$

のように、互換の積に分解できる。したがって、

$$\text{sign}(i_1, i_2, \dots, i_k) = (-1)^{k-1}$$

である。

5.3 n 次行列式の定義と基本性質

定理 5.8. $V = \mathbb{R}^n$ 上に次の性質 (D1), (D2), (D3) をもつような n 変数の関数 $D(\mathbf{a}_1, \dots, \mathbf{a}_n)$ が一意的に存在する。

この関数を n 次の行列式 (**determinant**) といい、 $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ であるとき、

$$D(\mathbf{a}_1, \dots, \mathbf{a}_n) = \det(A) = |A| = |\mathbf{a}_1, \dots, \mathbf{a}_n|$$

などを書く。

この定理の証明の前に、まずは性質 (D1), (D2), (D3) を解説する。

(D1) $D(\mathbf{a}_1, \dots, \mathbf{a}_n)$ は各変数 \mathbf{a}_i に関して線形である。

$$\begin{aligned} D(\mathbf{a}_1 + \mathbf{a}'_1, \dots, \mathbf{a}_n) &= D(\mathbf{a}_1, \dots, \mathbf{a}_n) + D(\mathbf{a}'_1, \dots, \mathbf{a}_n), \\ D(c\mathbf{a}_1, \dots, \mathbf{a}_n) &= cD(\mathbf{a}_1, \dots, \mathbf{a}_n) \end{aligned}$$

(D2) 2 つの変数 $\mathbf{a}_i, \mathbf{a}_j$ ($i \neq j$) を入れ替えると、 D の符号が変わる。つまり、 $\mathbf{a}_1, \dots, \mathbf{a}_n$ に関して交代적이다。

$$D(\dots, \mathbf{a}_j, \dots, \mathbf{a}_i, \dots) = -D(\dots, \mathbf{a}_i, \dots, \mathbf{a}_j, \dots)$$

(D3) \mathbb{R}^n の基本ベクトル $\mathbf{e}_1, \dots, \mathbf{e}_n$ に対して

$$D(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$$

また、(D2) は (D1) の下に、次の (D2') と同値である。

(D2') $\mathbf{a}_i = \mathbf{a}_j$ ($i \neq j$) のとき、

$$D(\dots, \mathbf{a}_i, \dots, \mathbf{a}_i, \dots) = 0$$

実際、(D2) を仮定すると

$$D(\dots, \mathbf{a}_i, \dots, \mathbf{a}_i, \dots) = -D(\dots, \mathbf{a}_i, \dots, \mathbf{a}_i, \dots)$$

であるから、この主張が言える。

逆に (D2') を仮定すると、(D1) により $D(\dots, \mathbf{a}_i, \dots, \mathbf{a}_j, \dots)$ の第 i, j 番目の変数に $\mathbf{a}_i + \mathbf{a}_j$ を代入して

$$\begin{aligned}
D(\cdots, \mathbf{a}_i + \mathbf{a}_j, \cdots, \mathbf{a}_i + \mathbf{a}_j, \cdots) &= D(\cdots, \mathbf{a}_i, \cdots, \mathbf{a}_i, \cdots) + D(\cdots, \mathbf{a}_i, \cdots, \mathbf{a}_j, \cdots) \\
&\quad + D(\cdots, \mathbf{a}_j, \cdots, \mathbf{a}_i, \cdots) + D(\cdots, \mathbf{a}_j, \cdots, \mathbf{a}_j, \cdots) \\
&= 0 + 1 - 1 + 0 \\
&= 0
\end{aligned}$$

(D2') により

$$D(\cdots, \mathbf{a}_i, \cdots, \mathbf{a}_i, \cdots) = D(\cdots, \mathbf{a}_j, \cdots, \mathbf{a}_j, \cdots) = 0$$

であるから、右辺の $\mathbf{a}_j, \mathbf{a}_j$ を入れ替えても

$$D(\cdots, \mathbf{a}_i, \cdots, \mathbf{a}_i, \cdots) = -D(\cdots, \mathbf{a}_j, \cdots, \mathbf{a}_j, \cdots)$$

とでき、(D2) の交代式を得る。よって、(D2) と (D2') はこの条件のもと同値である。

これらの性質より、定理 5.8 は次のように証明される。

証明. (D1), (D2), (D3) の性質をもつ関数 $D(\mathbf{a}_1, \cdots, \mathbf{a}_n)$ ($\mathbf{a}_j \in \mathbb{R}^n$) が存在すると仮定する。また、 $A = (\mathbf{a}_1, \cdots, \mathbf{a}_n) \in M_n(\mathbb{R})$ とする。

A の各ベクトル \mathbf{a}_j は線形結合により

$$\mathbf{a}_j = \sum_{1 \leq i \leq n} a_{ij} \mathbf{e}_i \quad (a_{ij} \in \mathbb{R})$$

と表せる。(D1) より

$$D(\mathbf{a}_1, \cdots, \mathbf{a}_n) = \sum_{i_1, \cdots, i_n=1}^n a_{i_1 1} \cdots a_{i_n n} D(\mathbf{e}_{i_1}, \cdots, \mathbf{e}_{i_n}) \quad (14)$$

ここで、 i_1, \cdots, i_n は独立に 1 から n まで動く。(D2),(D3) より $D(\mathbf{e}_1, \cdots, \mathbf{e}_n)$ は、 (i_1, \cdots, i_n) が $(1, \cdots, n)$ の順列になっているときには、置換 $\sigma = \begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix}$ の符号に等しく、その他に i_1, \cdots, i_n の中に同じ数が含まれているときは 0 となる。したがって、

$$D(\mathbf{e}_1, \cdots, \mathbf{e}_n) = \pm 1$$

であるから、式 (14) は次のように表せる。

$$D(\mathbf{a}_1, \cdots, \mathbf{a}_n) = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \quad (15)$$

このように、 $D(\mathbf{a}_1, \cdots, \mathbf{a}_n)$ は a_{ij} の多項式として明らかに一意的に定まる。 $(A$ に対して $\sigma \in S_n$ は $n!$ 個の組み合わせのみである。)

逆に、式 (14) によって $D(\mathbf{a}_1, \cdots, \mathbf{a}_n)$ を定義すれば、それが (D1),(D2),(D3) の性質を持つことは、補題 5.5 で示したように $\text{sign}(\sigma)$ の性質から確認できる。よって、行列式は式 (14) によって定義でき、それは一意的に存在する。

定理 5.9. $V = \mathbb{R}^n$ 上の関数 F が定理 5.8 における D の性質 (D1),(D2) をもてば、ある定数 c があるとき

$$F(\mathbf{a}_1, \dots, \mathbf{a}_n) = c|A|$$

ただし、 $c = F(\mathbf{e}_1, \dots, \mathbf{e}_n)$

例 5.10.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in M_3(\mathbb{R})$$

の行列式は以下の要領で得られる。

例題 5.3 で得られた置換 $\sigma \in S_n$ を左上から順に、 $\sigma_1, \dots, \sigma_6$ とラベリングすると、

$$\begin{aligned} \det(A) &= \text{sign}(\sigma_1) a_{\sigma_1(1)1} a_{\sigma_1(2)2} a_{\sigma_1(3)3} + \text{sign}(\sigma_2) a_{\sigma_2(1)1} a_{\sigma_2(2)2} a_{\sigma_2(3)3} + \text{sign}(\sigma_3) a_{\sigma_3(1)1} a_{\sigma_3(2)2} a_{\sigma_3(3)3} \\ &\quad + \text{sign}(\sigma_4) a_{\sigma_4(1)1} a_{\sigma_4(2)2} a_{\sigma_4(3)3} + \text{sign}(\sigma_5) a_{\sigma_5(1)1} a_{\sigma_5(2)2} a_{\sigma_5(3)3} + \text{sign}(\sigma_6) a_{\sigma_6(1)1} a_{\sigma_6(2)2} a_{\sigma_6(3)3} \\ &= a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} \\ &\quad - a_{11} a_{32} a_{23} - a_{21} a_{12} a_{33} - a_{31} a_{22} a_{13} \end{aligned}$$

が得られる。

定理 5.11. $A \in M_n(\mathbb{R})$ の転置行列を tA とすれば、

$$|{}^tA| = |A|$$

となる。

証明. tA の列ベクトルを $\mathbf{a}'_1, \dots, \mathbf{a}'_n$ とすれば、 ${}^t\mathbf{a}'_1, \dots, {}^t\mathbf{a}'_n$ は A の行ベクトルである。よって、

$$\mathbf{a}_j = \sum_{i=1}^n a_{ij} \mathbf{e}_i$$

とすれば、

$$\mathbf{a}'_j = \sum_{i=1}^n a_{ij} \mathbf{e}_i$$

である。行列式の定義式を tA に適用して

$$\begin{aligned} |{}^tA| &= D(\mathbf{a}'_1, \dots, \mathbf{a}'_n) \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \end{aligned}$$

σ の逆置換を σ^{-1} とすれば、 $\text{sign}(\sigma) = \text{sign}(\sigma^{-1})$ は明らかであるから、

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

である。これは、 σ が S_n 上を動くとき、 σ^{-1} も S_n の上を動くことから

$$|^tA| = |A|$$

が得られる。

5.3.1 様々な行列式の例

例 5.12. 上半三角、下半三角行列の行列式

n 次行列 $A = (a_{ij})$ が上半三角で、 n 次行列 $B = (b_{ij})$ が下半三角であるとき、つまり

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

の形である時、 A と B の行列式は

$$|A| = a_{11}a_{22} \cdots a_{nn}$$

$$|B| = b_{11}b_{22} \cdots b_{nn}$$

のように対角成分の積に等しい。

証明. A, B の形状より

$$i > j \implies a_{ij} = 0$$

$$i < j \implies b_{ij} = 0$$

したがって、式 (15) において

$$\sigma(i) \leq i \quad (1 \leq i \leq n)$$

$$\sigma(i) \geq i \quad (1 \leq i \leq n)$$

であるような σ に対する項だけが残るが、この条件を満たす σ は恒等置換 $\sigma(i) = i$ のみであるから、ここで主張する $|A|, |B|$ を得る。

例 5.13. n 次行列 $A = (a_{ij})$ が

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},$$

$$A_{11} \in M_{n_1}(\mathbb{R}), \quad A_{12} \in M_{n_1, n_2}(\mathbb{R}), \quad A_{21} \in M_{n_2, n_1}(\mathbb{R}), \quad n_1 + n_2 = n$$

の形ならば、

$$|A| = |A_{11}| |A_{22}| \quad (16)$$

証明. A の形状より、 $i > n_1, j \leq n_1$ のとき $a_{ij} = 0$ となるから、式 (15) において

$$j \leq n_1 \implies \sigma(j) \leq n_1$$

でるような σ に対する項だけが残る。この条件を満たす σ は $\{1, \dots, n_1\}$ の置換 σ_1 と $\{n_1 + 1, \dots, n\}$ の置換 σ_2 の積として表される。 $\{1, \dots, n_1\}$ の置換全体の集合を S_{n_1} , $n_1 + 1, \dots, n$ の置換全体の集合を S_{n-n_1} とおけば、 σ_1, σ_2 は独立にそれぞれ S_{n_1}, S_{n-n_1} の上を動く。また、 $\sigma = \sigma_1 \sigma_2$ であるとき、明らかに

$$\text{sign}(\sigma) = \text{sign}(\sigma_1) \text{sign}(\sigma_2)$$

である。よって、

$$\begin{aligned} |A| &= \sum_{\sigma_1 \in S_{n_1}} \text{sign}(\sigma_1) a_{\sigma_1(1)1} \cdots a_{\sigma_1(n_1)n_1} \\ &\quad \times \sum_{\sigma_2 \in S_{n_2}} \text{sign}(\sigma_2) a_{\sigma_2(n_1+1)n_1+1} \cdots a_{\sigma_2(n)n} \\ &= |A_{11}| |A_{22}| \end{aligned} \quad (17)$$

例題 5.14. $n_1 = 2, n_2 = 3$ として、例 5.13 を確かめよ。

解答

$\{1, 2\}$ の置換全体の集合を S_2 , $\{3, 4, 5\}$ の置換全体の集合を S_3 とする。 $\sigma_1 \in S_2, \sigma_2 \in S_3$ とする。 σ_1 と σ_2 はそれぞれ S_2, S_3 の上を独立に動くから、行列式は以下ようになる。

$$\begin{aligned} |A| &= a_{11}a_{22} - a_{21}a_{12} \\ &\quad + a_{33}a_{44}a_{55} + a_{43}a_{54}a_{35} + a_{53}a_{34}a_{45} \\ &\quad - a_{33}a_{54}a_{45} - a_{43}a_{34}a_{55} - a_{53}a_{44}a_{35} \\ &= |A_{11}| |A_{22}| \end{aligned}$$

したがって、例 5.13 と同様の行列式が得られた。

5.3.2 余因子

n 次行列 $A = (a_{ij}) = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ の第 i 行、第 j 列を取り去ってできる $(n-1)$ 次行列を $A^{(ij)}$ とかくことにする。そのとき、 A の第 j 列を第 i 基本ベクトル \mathbf{e}_i で置きかえて得られる行列の行列式は次の式で与えられる。 $(|\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n|$ の \mathbf{a}_j が \mathbf{e}_i に入れ替えられている。)

$$|\mathbf{a}_1, \dots, \mathbf{e}_i, \dots, \mathbf{a}_n| = (-1)^{i+j} |A^{(ij)}| \quad (18)$$

この式の右边を A の (i, j) 余因子 (cofactor) という。

実際、余因子は次のように求められる。

$\mathbf{a}_1, \dots, \mathbf{a}_n$ の i 行、 j 列を消し去り、第 j 列を基本ベクトル \mathbf{e}_i で置き換えた行列は以下のように表示できる。

$$A' = \begin{pmatrix} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & 0 & a_{i-1j+1} & \cdots & a_{i-1n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+11} & \cdots & a_{i+1j-1} & 0 & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}$$

行列式の性質より、一組の列と列（あるいは行と行）を入れ替えると、符号が反転するところに着目すると、 A' の第 $j-1$ 列と第 j 列（つまり、 \mathbf{e}_i ）を入れ替えた行列 A'' は

$$A' = -A''$$

と表すことができる。この入れ替えを列に関して $j-1$ 回行うと

$$\begin{pmatrix} 0 & a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}$$

の行列が得られるから、この行列の行列式は

$$(-1)^{j-1} |\mathbf{e}_i, \mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \dots|$$

の形になる。さらに、行に関して $i-1$ 回入れ替えを行うと最終的に

$$\begin{pmatrix} 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ 0 & a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}$$

の行列が得られる。この行列の行列式は

$$(-1)^{(i-1)+(j-1)} \begin{vmatrix} a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{vmatrix}$$

に等しい。よって、 $(-1)^{(i-1)+(j-1)} = (-1)^{i+j}$ から

$$|a_1, \dots, e_i, \dots, a_n| = (-1)^{i+j} |A'|$$

が成立する。

5.3.3 余因子展開

この余因子を使った n 次行列式を求めるテクニックを紹介する。

$A = (a_{ij}) \in M_n(\mathbb{R})$ の第 (i, j) 余因子による行列式を $\Delta_{ij} = (-1)^{i+j} |A^{(ij)}|$ とすれば、第 j 列に関する展開を

$$|A| = \sum_{i=1}^n a_{ij} \Delta_{ij}$$

第 i 行に関する展開を

$$|A| = \sum_{j=1}^n a_{ij} \Delta_{ij}$$

とできる。

証明. 列に関する展開式を証明すれば十分である。 A の列ベクトルを a_1, \dots, a_n とすれば、

$$a_j = \sum_{i=1}^n a_{ij} e_i$$

であるから、行列式の線形性 $D(ca_1, a_2) = cD(a_1, a_2)$ により

$$|A| = D(a_1, \dots, a_j, \dots, a_n) = \sum_{i=1}^n a_{ij} D(a_1, \dots, e_i, \dots, a_n)$$

となることより、

$$|A| = \sum_{i=1}^n a_{ij} \Delta_{ij}$$

を得る。

例 5.15. $n = 3$ のとき、第 1 列に関する余因子展開は次となる。

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

定理 5.16. $A, B \in M_n(\mathbb{R})$ に対して

$$|AB| = |A||B|.$$

証明. B の列ベクトルを $\mathbf{b}_1, \dots, \mathbf{b}_n$ とすれば、

$$|AB| = D(A\mathbf{b}_1, \dots, A\mathbf{b}_n)$$

右辺を $F(\mathbf{b}_1, \dots, \mathbf{b}_n)$ とおけば、明らかに行列式の性質 (D1), (D2) をもっている。

$$\begin{aligned} F(\mathbf{b}_1, \dots, \mathbf{b}_n) &= F(\mathbf{e}_1, \dots, \mathbf{e}_n)D(\mathbf{b}_1, \dots, \mathbf{b}_n) \\ &= F(\mathbf{e}_1, \dots, \mathbf{e}_n)|B| \end{aligned}$$

A の列ベクトルを $\mathbf{a}_1, \dots, \mathbf{a}_n$ とすれば、 $A\mathbf{e}_j = \mathbf{a}_j$ であるから

$$F(\mathbf{e}_1, \dots, \mathbf{e}_n) = D(\mathbf{a}_1, \dots, \mathbf{a}_n) = |A|$$

定理 5.17. $A \in M_n(\mathbb{R})$ が可逆であるためには、 $|A| \neq 0$ が必要十分である。このとき、 $|A|^{-1} = |A^{-1}|$ となる。

証明. A が可逆ならば

$$AA^{-1} = E$$

であるから、両辺の行列式をとり

$$|A||A^{-1}| = |E| = 1$$

よって $|A| \neq 0$, $|A|^{-1} = |A^{-1}|$

6 計量ベクトル空間

線形空間では長さ、角、面積、体積といった計量を行うことは出来なかったが、ここに内積という構造を導入することでそのような計量を行える線形空間を計量ベクトル空間（内積空間）と呼ぶ。この章では、計量ベクトル空間における基本的な概念や性質について説明する。

6.1 内積の性質

$V = \mathbb{R}^n$ における内積は 2.8.2 節において定義した。簡単に復習すると $\mathbf{x} = (x_i), \mathbf{y} = (y_i) \in \mathbb{R}^n$ の内積 (inner product) は

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

であった。なお、 \mathbb{R}^n における内積を標準内積という。 $(\mathbb{C}^n$ などの空間にも内積は定義されることから、このように標準と呼称している。)

この定義より、内積は次の性質を持つ。

(1) $\langle \mathbf{x}, \mathbf{y} \rangle$ は \mathbf{x}, \mathbf{y} について線形である。 \mathbf{x} についていえば、任意の $\mathbf{x}, \mathbf{x}', \mathbf{y} \in V, c \in \mathbb{R}$ に対して

$$\begin{aligned}\langle \mathbf{x} + \mathbf{x}', \mathbf{y} \rangle &= \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}', \mathbf{y} \rangle, \\ \langle c\mathbf{x}, \mathbf{y} \rangle &= c\langle \mathbf{x}, \mathbf{y} \rangle\end{aligned}$$

(2) $\langle \mathbf{x}, \mathbf{y} \rangle$ は \mathbf{x}, \mathbf{y} について対称である。

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$$

(3) $\mathbf{x} \neq \mathbf{0}$ ならば

$$\langle \mathbf{x}, \mathbf{x} \rangle > 0$$

実際、 $\mathbf{x} = (x_i) \neq \mathbf{0}$ ならば、

$$\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^n x_i^2 > 0$$

一般に、(1), (2) を満たす V 上の 2 変数関数を $V \times V$ 上の対称双 1 次形式 (symmetric bilinear form) という。特に (3) の性質を持つものを正値定符号であるという。(正定値性を持つという。) したがって、 $\langle \mathbf{x}, \mathbf{y} \rangle$ は正値定符号対称双 1 次形式である。

6.1.1 ベクトルのノルム

内積の正定値性より

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} > 0$$

が定義される。これをベクトル \mathbf{x} の長さあるいはノルムという。ノルムの定義から明らかに

$$\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$$

である。また任意の $\mathbf{x} \in V, c \in \mathbb{R}$ に対して

$$\|c\mathbf{x}\| = |c|\|\mathbf{x}\|$$

が成立する。また同様のベクトルの 2 乗は、内積を利用し

$$\mathbf{x}^2 = \|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$$

と表すことができる。

例題 6.1. 次の等式を証明せよ

$$(1) \langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{2} \{ \|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 \}$$

$$(2) \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

解答

(1)

$$\begin{aligned} t = \|x + y\|^2 - \|x\|^2 + \|y\|^2 &= \langle x + y, x + y \rangle - \langle x, x \rangle + \langle y, y \rangle \\ &= \langle x, x + y \rangle + \langle y, x + y \rangle - \langle x, x \rangle + \langle y, y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle - \langle x, x \rangle + \langle y, y \rangle \\ &= 2\langle x, y \rangle \\ \frac{1}{2} \cdot t &= \frac{1}{2} \cdot 2\langle x, y \rangle = \langle x, y \rangle \end{aligned}$$

(2)

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= \langle x, x + y \rangle + \langle y, x + y \rangle + \langle x, x - y \rangle + \langle -y, x - y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle + \langle x, x \rangle + \langle x, -y \rangle + \langle x, -y \rangle + \langle -y, x \rangle + \langle -y, -y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle + \langle x, x \rangle - \langle x, y \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle \\ &= 2\langle x, x \rangle + 2\langle y, y \rangle \\ &= 2(\|x\|^2 + \|y\|^2) \end{aligned}$$

定理 6.2. 任意の $x, y \in \mathbb{R}^n$ に対して

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

シュワルツの不等式が成立する。

証明. $x = 0$ or $y = 0$ ならば、この不等式は $0 = 0$ として成立する。したがって、 $x, y \neq 0$ とする。

$$f(\lambda) = \|x + \lambda y\|^2 \quad (\lambda \in \mathbb{R})$$

とおくと

$$\begin{aligned} f(\lambda) &= \langle \lambda x + y, \lambda x + y \rangle \\ &= \lambda^2 \langle x, x \rangle + 2\lambda \langle x, y \rangle + \langle y, y \rangle \end{aligned}$$

と展開できる。よって、 f は λ に関する 2 次式で、すべての $\lambda \in \mathbb{R}$ に対して、 $f(\lambda) \geq 0$ であるためには、2 次式の判別式 ($\frac{D}{4} = b^2 - ac$) が負にならないなければならない。したがって、

$$\langle x, y \rangle^2 - \langle x, x \rangle \langle y, y \rangle \leq 0$$

すなわち

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle$$

ここで両辺とも ≥ 0 であるから、平方根をとると

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

が成立する。

系 6.3. $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ に対して三角不等式

$$\|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

が成立する。

証明. 定理 6.2 の証明で $\lambda = 1$ とおけばシュバルツの不等式により

$$\begin{aligned} f(1) &= \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2 \end{aligned}$$

両辺ともに 0 以上であるから、平方根をとって主張の三角不等式のうちの 2 つめの不等式を得る。また、ひとつめの不等式は明らかに成立する。

この三角不等式の成立により計量ベクトル空間においてノルムが距離と呼べるようになる。

6.1.2 ベクトルのなす角

シュバルツの不等式の内積の絶対値を外して式を変形することで、 $\mathbf{x}, \mathbf{y} \neq 0$ のとき、

$$-1 \leq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1 \quad (19)$$

が成立する。もうひとつの内積の定義

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\|\mathbf{y}\| \cos \theta$$

を変形すると

$$\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}$$

を得るから、式 (19) は

$$-1 \leq \cos \theta \leq 1$$

に 1 対 1 で対応する。つまり、 $0 \leq \theta \leq \pi$ が一意的に定まる。この θ を 2 つのベクトル \mathbf{x}, \mathbf{y} のなす角という。

例 6.4. $n = 2$ として、ベクトルのなす角 θ を求める。 $\mathbf{x} = (x_i), \mathbf{y} = (y_i) \in \mathbb{R}^n$,

$$\begin{cases} x_1 = r_1 \cos \theta_1 \\ x_2 = r_1 \sin \theta_1 \end{cases} \quad \begin{cases} y_1 = r_1 \cos \theta_2 \\ y_2 = r_1 \sin \theta_2 \end{cases}$$

とおく。

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= x_1 y_1 + x_2 y_2 \\
&= r_1 r_2 (\cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2) \\
&= r_1 r_2 \cos(\theta_2 - \theta_1)
\end{aligned}$$

したがって

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \cos(\theta_2 - \theta_1)$$

であるから、ある整数 m があって

$$\theta = |\theta_2 - \theta_1 + 2\pi m| \leq \pi$$

となる。よって、この θ は幾何学的な意味での \mathbf{x}, \mathbf{y} のなす角に他ならない。

例 6.5. 2つのベクトルが張る平行四辺形の面積と行列式

$\mathbf{x} = (x_i), \mathbf{y} = (y_i) \in \mathbb{R}^2$ の2つのベクトルが張る平行四辺形の面積と $A = (\mathbf{a}, \mathbf{b})$ の行列式 ($\det(A)$) の絶対値をとった値は一致することを示す。

ベクトル \mathbf{x}, \mathbf{y} が張る平行四辺形に \mathbf{y} の先端から \mathbf{x} の先端に対角線つまり、ベクトル $\mathbf{y} - \mathbf{x}$ を引くと平行四辺形を2分する三角形ができる。

この三角形の面積 S_1 を

$$S_1 = \frac{1}{2} \cdot \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta$$

とすると。高さは $\|\mathbf{y}\| \sin \theta$ であるから、平行四辺形の面積 S_2 は

$$S_2 = \|\mathbf{x}\| \|\mathbf{y}\| \sin \theta$$

この式より

$$\begin{aligned}
\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \sin^2 \theta &= \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle (1 - \cos^2 \theta) \\
&= \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle \cos^2 \theta
\end{aligned}$$

が得られるから、両辺の平方根を取ると

$$\begin{aligned}
\|\mathbf{x}\| \|\mathbf{y}\| \sin \theta &= \|\mathbf{x}\| \|\mathbf{y}\| - \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta \\
&= \|\mathbf{x}\| \|\mathbf{y}\| - \langle \mathbf{x}, \mathbf{y} \rangle
\end{aligned}$$

のような関係式が得られる。

これを展開すると

$$\begin{aligned}
\|\mathbf{x}\|\|\mathbf{y}\| - \langle \mathbf{x}, \mathbf{y} \rangle &= \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle^2} \\
&= \sqrt{(x_1^2 + x_2^2)(y_1^2 + y_2^2) - (x_1y_1 + x_2y_2)^2} \\
&= \sqrt{x_1^2y_1^2 + x_1^2y_2^2 + x_2^2y_1^2 + x_2^2y_2^2 - (x_1^2y_1^2 + x_1y_1x_2y_2 + x_2y_2x_1y_1 + x_2^2y_2^2)} \\
&= \sqrt{x_1^2y_2^2 + x_2^2y_1^2 - x_1y_1x_2y_2 + x_2y_2x_1y_1} \\
&= \sqrt{(x_1y_2 - x_2y_1)^2} \\
&= |x_1y_2 - x_2y_1|
\end{aligned}$$

より $S_2 = x_1y_2 - x_2y_1$ である。

一方

$$\det(A) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1$$

であるから

$$\|\mathbf{x}\|\|\mathbf{y}\| - \langle \mathbf{x}, \mathbf{y} \rangle = \det(A)$$

である。

例題 6.6.

$$\mathbf{x} = \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

に対し、 $\|\mathbf{x}\|, \|\mathbf{y}\|, \theta = \angle(\mathbf{x}, \mathbf{y})$ を求めよ。

ノルムを求める。

$$\|\mathbf{x}\| = 2, \quad \|\mathbf{y}\| = \sqrt{2}$$

角を求める。

$$\begin{aligned}
\angle(\mathbf{x}, \mathbf{y}) &= \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \\
&= \frac{\sqrt{3} + 1}{2\sqrt{2}} \\
&= \frac{1}{2} \cdot \frac{1}{\sqrt{2}} + \frac{\sqrt{3}}{2} \cdot \frac{1}{\sqrt{2}} \\
&= \cos \frac{\pi}{3} \cdot \cos \frac{\pi}{4} + \sin \frac{\pi}{3} \cdot \sin \frac{\pi}{4} \\
&= \cos\left(\frac{\pi}{3} - \frac{\pi}{4}\right)
\end{aligned}$$

6.1.3 ベクトルの直交条件

核の定義から、特別な場合として、ベクトル $\mathbf{x}, \mathbf{y} \neq 0$ に対し

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 \iff \theta = \frac{\pi}{2}$$

となるようなとき、 \mathbf{x}, \mathbf{y} は直交する (orthogonal) という。逆に、シュワルツの不等式で統合が成立する場合、つまり

$$\langle \mathbf{x}, \mathbf{y} \rangle = \pm \|\mathbf{x}\| \|\mathbf{y}\|$$

となるためには、 $\theta = 0$ or π であることが必要十分である。この場合、2つのベクトルは平行となる。

この節で述べた定理やその系により、内積の定義された計量ベクトル空間では、様々な計量を定義することが可能になる。

6.2 正規直交系

ベクトル $\mathbf{e}_1, \dots, \mathbf{e}_r$ が

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} \quad (1 \leq i, j \leq r)$$

を満たすとき、正規直交系 (orthonormal system) であるという。この条件から、各 i に対して $\|\mathbf{e}_i\| = 1$ 。また任意の相異なる i, j に対して、 $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$ 、すなわち $\mathbf{e}_i, \mathbf{e}_j$ は直行する。

例 6.7. $V = \mathbb{R}^n$ の基本ベクトル $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ は標準内積に関して正規直交基底をなす。これまで $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ は基本ベクトルを表してきたが、以降、一般の正規直交基底に関しても $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ で表すことにする。(正規直交基底は基底と名がつく通り、それを構成するベクトルはすべて線形独立である。一方、直交でない通常の基底は線形独立ではあるが、内積は 0 とはならないので、「線形独立である \iff 内積が 0」ではないことに注意したい。)

補題 6.8. $\mathbf{e}_1, \dots, \mathbf{e}_r$ を正規直交系

$$\mathbf{x} = \sum_{i=1}^r x_i \mathbf{e}_i, \quad \mathbf{y} = \sum_{i=1}^r y_i \mathbf{e}_i$$

とすれば

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

である。また

$$\langle \mathbf{x}, \mathbf{e}_i \rangle = x_i$$

証明. これは計算過程により証明できる。

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= \left\langle \sum_{i=1}^r x_i \mathbf{e}_i, \sum_{i=1}^r y_i \mathbf{e}_i \right\rangle \\
&= \sum_{i,j=1}^r x_i y_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle \\
&= \sum_{i,j=1}^r x_i y_j \delta_{ij} = \sum_{i=1}^r x_i y_i
\end{aligned}$$

この計算過程より、 $\mathbf{y} = \mathbf{e}_i$ とすれば

$$\langle \mathbf{x}, \mathbf{e}_i \rangle = x_i$$

となる。(つまり、 \mathbf{x} の i 番目以外はすべて 0 になる)

6.2.1 ノルムが 1 のベクトルを作る方法

$\mathbf{a} \in \mathbb{R}^n$ をノルムが 1 ではないベクトルとする。 \mathbf{a} のノルムを 1 (つまり単位ベクトル) とするには

$$\frac{\mathbf{a}}{\|\mathbf{a}\|}$$

のように、ベクトルをそのベクトルのノルムで割ればよい。

定理 6.9. $V = \mathbb{R}^n$ の任意の部分空間 W は正規直交基底をもつ。

証明

$\dim W = r$ とし、 $\mathbf{a}_1, \dots, \mathbf{a}_r$ を W の任意の底とする。このとき、

$$\langle \mathbf{e}_1, \dots, \mathbf{e}_k \rangle = \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle \quad (1 \leq k \leq r)$$

となるよう正規直交系 $\langle \mathbf{e}_1, \dots, \mathbf{e}_k \rangle$ が取れる (作れる) ことを k に関する帰納法で示す。($k = n$ の場合に定理を与える)

$k = 1$ のとき

$$\mathbf{e}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \tag{20}$$

とおけば、 $\{\mathbf{e}_1\}$ はただ一つのベクトルからなる正規直交系となる。

$k - 1$ のとき式 (20) が成立するような $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}$ が存在したとする。このとき、

$$\mathbf{a}'_k = \sum_{1 \leq i \leq k-1} \langle \mathbf{e}_i, \mathbf{a}_k \rangle \mathbf{e}_i, \quad \mathbf{a}''_k = \mathbf{a}_k - \mathbf{a}'_k$$

とおけば、各 k における式 (20) は

$$\mathbf{e}_i = \frac{\mathbf{a}''_k}{\|\mathbf{a}''_k\|} \quad (2 \leq i \leq k-1)$$

である。 \mathbf{a}'_k は \mathbf{e}_i 方向のベクトルより、(ベクトルどうしの引き算の定義から) \mathbf{a}''_k は \mathbf{e}_i に必ず直行する。つまり、

$$\langle \mathbf{e}_i, \mathbf{a}''_k \rangle = \langle \mathbf{e}_i, \mathbf{a}_k \rangle - \langle \mathbf{e}_i, \mathbf{a}'_k \rangle = 0 \quad (1 \leq i \leq k-1)$$

である。ここで、 $\mathbf{a}_1, \dots, \mathbf{a}_k$ は線形独立であるから

$$\mathbf{a}_k \notin \langle \mathbf{a}_1, \dots, \mathbf{a}_{k-1} \rangle = \langle \mathbf{e}_1, \dots, \mathbf{e}_{k-1} \rangle$$

となるはずであるから、 $\mathbf{a}''_k \neq 0$ 。したがって、

$$\mathbf{e}_k = \frac{\mathbf{a}''_k}{\|\mathbf{a}''_k\|}$$

とおけば、

$$\langle \mathbf{e}_i, \mathbf{e}_k \rangle = \delta_{ik} \quad (1 \leq i \leq k)$$

が成立する。したがって、定理の主張が示せた。

また、このような正規直交基底の構成方法を（グラム）シュミットの正規直交化法 という。

系 6.10. 部分空間 W の任意の正規直交基底は V の正規直交基底に延長される。

これは基底の延長を行い、シュミットの正規直交化法を利用することで証明できる。

例題 6.11.

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

とする。

$$\langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle = \langle \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \rangle \quad (i = 1, 2, 3)$$

となるような正規直交基底 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ を求めよ。

シュミットの正規直交化法より

$$\|\mathbf{a}_1\| = \sqrt{2}, \quad \mathbf{e}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

$$\begin{aligned} \mathbf{a}'_2 &= \langle \mathbf{e}_1, \mathbf{a}_2 \rangle \mathbf{e}_1 \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \end{aligned}$$

$$\mathbf{a}_2'' = \mathbf{a}_2 - \mathbf{a}_2' = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$$

$$\begin{aligned} \mathbf{e}_2 &= \frac{\mathbf{a}_2''}{\|\mathbf{a}_2''\|} \\ &= \left(\frac{\sqrt{6}}{2}\right)^{-1} \cdot \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \mathbf{a}_3' &= \langle \mathbf{e}_1, \mathbf{a}_3 \rangle \mathbf{e}_1 + \langle \mathbf{e}_2, \mathbf{a}_3 \rangle \mathbf{e}_2 \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} + \frac{1}{\sqrt{6}} \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \\ \mathbf{a}_3'' &= \mathbf{a}_3 - \mathbf{a}_3' \\ &= \begin{pmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} \end{aligned}$$

$$\mathbf{e}_3 = \frac{\mathbf{a}_3''}{\|\mathbf{a}_3''\|} = \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

6.3 直交行列

一般に

$$A \in M_{m,n}(\mathbb{R}), \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^n$$

に対し

$$\begin{aligned} \langle \mathbf{x}, A\mathbf{y} \rangle &= {}^t\mathbf{x}A\mathbf{y} = {}^t\mathbf{x}^tA\mathbf{y} \\ &= {}^t({}^tA\mathbf{x})\mathbf{y} = \langle {}^tA\mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

が成立することに注意して、次の定理を見ていくことにする。

定理 6.12. $T \in M_n(\mathbb{R})$ に対して、次の 4 条件は同値である。

(1) T は内積を不変にする。すなわち

$$\langle T\mathbf{x}, T\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \quad (\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n)$$

(2) T はベクトルの長さを不変にする。すなわち

$$\|T\mathbf{x}\| = \|\mathbf{x}\|$$

(3) T は直交行列 (orthogonal matrix) である。すなわち

$${}^tT T = T {}^tT = E$$

これをいいかえると、 T は可逆で $T^{-1} = {}^tT$

(4) $\{\mathbf{e}_i\}$ を V の正規直交基底とすれば、 $\{T\mathbf{e}_i\}$ も正規直交基底である。

証明. (1) \Rightarrow (2), (2) \Rightarrow (3), (3) \Rightarrow (4), (4) \Rightarrow (1) の順で証明する。

(1) \Rightarrow (2)

$$\langle T\mathbf{x}, T\mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle.$$

であるから、

$$\|T\mathbf{x}\| = \sqrt{\langle T\mathbf{x}, T\mathbf{x} \rangle} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \|\mathbf{x}\|$$

(2) \Rightarrow (3)

$T = (t_{ij}), \mathbf{x} = (x_i)$ とおく。

$$\begin{aligned} (\text{左辺}) &= \langle T\mathbf{x}, T\mathbf{x} \rangle = \sum_{i=1}^n \left(\sum_{j=1}^n t_{ij} x_j \right) \left(\sum_{k=1}^n t_{ik} x_k \right) \\ &= \sum_{j,k=1}^n \left(\sum_{i=1}^n t_{ij} t_{ik} \right) x_j x_k \end{aligned}$$

$$(\text{右辺}) = \langle \mathbf{x}, \mathbf{x} \rangle = \sum_{j=1}^n x_j^2$$

両辺の $x_j x_k$ の係数を比べれば

$$\sum_{i=1}^n t_{ij} t_{ik} = \delta_{jk}$$

を得る。よって

$${}^tT T = E$$

がいえ、 T は可逆で $T^{-1} = {}^tT$.

(3) \Rightarrow (4)

(3) を仮定すれば

$$\langle Te_i, Te_j \rangle = \langle {}^tTTe_i, e_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}$$

よって、 $\{Te_i\}$ も正規直交基底である.

(4) \Rightarrow (1)

$\{Te_i\} = \{e_i\}$ を正規直交基底とし $\mathbf{x}, \mathbf{y} \in V$ に対し、

$$\mathbf{x} = \sum_{i=1}^n x_i e_i, \quad \mathbf{y} = \sum_{i=1}^n y_i e_i.$$

とすれば、補題 6.8 を使い

$$\begin{aligned} \langle T\mathbf{x}, T\mathbf{y} \rangle &= \left\langle \sum_{i=1}^n x_i Te_i, \sum_{i=1}^n y_i Te_i \right\rangle \\ &= \sum_{i=1}^n x_i y_i = \langle \mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

□

直交行列 T に対して

$$\det T^2 = \det {}^tTT = \det E = 1$$

であるから、 $\det T = \pm 1$.

なお、 n 次直交行列の全体の集合を $O(n)$ で表す。また、 $T \in O(n)$ で $\det T = 1$ であるようなものの全体の集合を $SO(n)$ で表す。

例題 6.13. $T \in M_n(\mathbb{R})$ の各列ベクトルを $\mathbf{t}_1, \dots, \mathbf{t}_n$ とすれば、

$$T \text{ が直交行列} \iff \mathbf{t}_i \text{ が正規直交基底}$$

であることを示せ.

証明. (方針: T が直交行列であるとは、 ${}^tTT = E$ となることを示し、 $\mathbf{t}_1, \dots, \mathbf{t}_n$ が正規直交基底となるには $\langle \mathbf{t}_i, \mathbf{t}_j \rangle = \delta_{ij}$ を示せば良い.)

$T = (\mathbf{t}_1, \dots, \mathbf{t}_n)$ を直交行列とすると、 $T \in O(n) \iff {}^tTT = E \iff {}^t\mathbf{t}_i \mathbf{t}_j = \langle \mathbf{t}_i, \mathbf{t}_j \rangle = \delta_{ij} \iff \mathbf{t}_1, \dots, \mathbf{t}_n$ は正規直交基底

例 6.14. 次の行列はいずれも直交行列である

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

導出

$$(2) \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} \cos^2 \theta + \sin^2 \theta & \cos \theta \sin \theta - \sin \theta \cos \theta \\ \sin \cos \theta - \cos \theta \sin \theta & \sin^2 \theta + \cos^2 \theta \end{pmatrix} = E$$

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

導出

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

定理 6.15. $\{\mathbf{e}_i\}, \{\mathbf{a}_i\}$ を \mathbb{R}^n の二つの基底、 T を基底の変換 $\{\mathbf{e}_i\} \rightarrow \{\mathbf{a}_i\}$ の表現行列とする。 $\{\mathbf{e}_i\}$ が正規直交基底である時

$$\{\mathbf{a}_i\} \text{ が正規直交基底} \iff T \text{ が直交行列}$$

であることを示せ.

証明. $T = (t_{ij})$ とおけば、 $\mathbf{a}_i = \sum_{k=1}^n t_{ki} \mathbf{e}_k$ であるから、

$$\begin{aligned} \langle \mathbf{a}_i, \mathbf{a}_j \rangle &= \left\langle \sum_{k=1}^n t_{ki} \mathbf{e}_k, \sum_{l=1}^n t_{lj} \mathbf{e}_l \right\rangle \\ &= \sum_{k=1}^n t_{ki} t_{kj} = \delta_{ij} \end{aligned}$$

よって、

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle \iff {}^t T T = E \iff {}^t t_i t_j = \delta_{ij}$$

.

この定理により、 T を直交行列、 $\{\mathbf{e}_i\}$ を任意の正規直交基底とすれば、線形変換 $f_T: \mathbf{x} \rightarrow T\mathbf{x}$ の基底 $\{\mathbf{e}_i\}$ (基底を取り替えた) に関する新座標系での表現行列を T' としてもこれも同様に直交行列である。この仮定のもとに、 $\{\mathbf{e}_i\}, \{T\mathbf{e}_i\}$ はともに正規直交基底であり、 T' は基底の変換 $\{\mathbf{e}_i\} \rightarrow \{T\mathbf{e}_i\}$ の表現行列に他ならない。実際、基底の変換行列を $T_1 = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ とおけば、 $T' = T_1^{-1} T T_1$ である。このように、直交行列 T に対応する線形変換 f_T を直交変換という。 f_T は内積を不変にする線形変換として特徴づけられる。

6.4 直交補空間

$V = \mathbb{R}^n$ の部分空間 W に対し

$$W^\perp = \{\mathbf{x} \in V \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0 \text{ for } \forall \mathbf{y} \in W\}$$

は明らかに V の部分空間になる。これを W の直交補空間 (orthogonal complement) という。

定理 6.16. $V = \mathbb{R}^n$ の部分空間 W の直交補空間を W^\perp とすれば、

$$V = W \oplus W^\perp$$

である。(⊕ は直和を表す論理記号で $W_1, W_2 \subset V$ に対して、 $W_1 \cap W_2 = \{0\}, W_1 \cup W_2 = V$ となるもの) また

$$\dim W^\perp = n - \dim W$$

である。

証明. W の正規直交基底 $\{\mathbf{e}_1, \dots, \mathbf{e}_r\} (r = \dim W)$ をとり、それを V の正規直交基底 $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ に延長する。そのとき、 $W = \langle \mathbf{e}_1, \dots, \mathbf{e}_r \rangle$ であるから、任意の

$$\mathbf{x} \in V, \mathbf{x} = \sum_{1 \leq i \leq n} x_i \mathbf{e}_i$$

に対して、補題 6.8 より

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle = 0 \quad (\forall \mathbf{y} \in W) &\iff \langle \mathbf{x}, \mathbf{e}_i \rangle = 0 \quad (1 \leq i \leq r) \\ &\iff x_i = 0 \quad (1 \leq i \leq r) \\ &\iff \mathbf{x} \in \langle \mathbf{e}_{r+1}, \dots, \mathbf{e}_n \rangle \end{aligned}$$

よって、

$$W^\perp = \langle \mathbf{e}_{r+1}, \dots, \mathbf{e}_n \rangle, \quad \dim W^\perp = n - r$$

である。よって

$$V = \langle \mathbf{e}_1, \dots, \mathbf{e}_r \rangle \oplus \langle \mathbf{e}_{r+1}, \dots, \mathbf{e}_n \rangle$$

は明らか。

6.4.1 正射影

$\{\mathbf{e}_1, \dots, \mathbf{e}_r\}$ を部分空間 W の正規直交基底とし、 $\mathbf{x} \in W$ に対し

$$\mathbf{x}' = \sum_{i=1}^r \langle \mathbf{e}_i, \mathbf{x} \rangle \mathbf{e}_i, \quad \mathbf{x}'' = \mathbf{x} - \mathbf{x}'$$

とおけば、

$$\mathbf{x} = \mathbf{x}' + \mathbf{x}'', \quad \mathbf{x}' \in W, \mathbf{x}'' \in W^\perp$$

である。このとき写像 $p: \mathbf{x} \mapsto \mathbf{x}'$ を W への正射影 (orthogonal projection) という。 $(\mathbf{x}'' \in W^\perp$ となるのは、(シュミットの直交化を思い出すと) ベクトルの引き算 $\mathbf{x} - \mathbf{x}'$ が \mathbf{x}' に直交するベクトルを求める行為に他ならないからである。)

p は V の線形変換で、 W によって一意的に定まる。 $\{\mathbf{e}_1, \dots, \mathbf{e}_r\}$ を V の正規直交基底 $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ に延長すれば、この基底に関する p の行列は

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}, \quad r = \text{rank } p$$

である。

系 6.17. W を $V = \mathbb{R}^n$ の部分空間とすれば

$$W^{\perp\perp} = W$$

証明. $W \subset W^{\perp\perp}$ は定義から明らか。 $\dim W^\perp = n - \dim W$ を使えば

$$\dim W^{\perp\perp} = n - \dim W^\perp = n - (n - \dim W) = \dim W$$

であるから、定理の主張が成立する。

定理 6.18. W_1, W_2 を $V = \mathbb{R}^n$ の部分空間とすれば

$$(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp, \quad (21)$$

$$(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp \quad (22)$$

証明. 式 (21) を示す。 $\mathbf{x} \in (W_1 + W_2)^\perp$ とする。

任意の $\mathbf{y} \in W_1 + W_2$ に対して $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

$$\iff \text{ある } \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2 \text{ に対し、} \mathbf{y} = \mathbf{w}_1 + \mathbf{w}_2 \text{ とおけるから、} \langle \mathbf{x}, \mathbf{w}_1 + \mathbf{w}_2 \rangle = \langle \mathbf{x}, \mathbf{w}_1 \rangle + \langle \mathbf{x}, \mathbf{w}_2 \rangle = 0$$

$$\iff \langle \mathbf{x}, \mathbf{w}_1 \rangle = \langle \mathbf{x}, \mathbf{w}_2 \rangle = 0$$

$$\iff \text{任意の } \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2 \text{ に対し、} \langle \mathbf{x}, \mathbf{w}_1 \rangle = \langle \mathbf{x}, \mathbf{w}_2 \rangle = 0$$

$$\iff \mathbf{x} \in W_1^\perp \cap W_2^\perp$$

$$\text{つまり、} (W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp = \{0\}$$

式 (22) を示す。 $\mathbf{x} \in (W_1 \cap W_2)^\perp$ とする。

任意の $\mathbf{y} \in W_1 \cap W_2$ に対して、 $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

$$\iff \text{ある } \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2 \text{ に対し、} \langle \mathbf{x}, \mathbf{w}_1 \rangle = \langle \mathbf{x}, \mathbf{w}_2 \rangle = 0.$$

$$\iff \langle \mathbf{x}, \mathbf{w}_1 \rangle + \langle \mathbf{x}, \mathbf{w}_2 \rangle = \langle \mathbf{x}, \mathbf{w}_1 + \mathbf{w}_2 \rangle = \langle \mathbf{x}, \mathbf{y} \rangle = 0$$

$$\iff \text{任意の } \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2 \text{ に対し、} \langle \mathbf{x}, \mathbf{w}_1 \rangle + \langle \mathbf{x}, \mathbf{w}_2 \rangle = 0$$

$$\iff \mathbf{x} \in W_1^\perp + W_2^\perp$$

$$\text{つまり、} (W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp = 0$$

7 行列の標準化

7.1 不変部分空間

$V = \mathbb{R}^n, A = (a_{ij}) \in M_n(\mathbb{R})$ とし、 A によって定義される V の線形変換

$$f = f_A : \mathbf{x} \mapsto A\mathbf{x}$$

があるとする。 A は f の標準基底 $\{\mathbf{e}_i\}$ に関する表現行列であるが、 V の他の基底 $\{\mathbf{v}_i\}$ をとり

$$(\mathbf{v}_1, \dots, \mathbf{v}_n) = (\mathbf{e}_1, \dots, \mathbf{e}_n)P, \quad P \in GL_n(\mathbb{R}) \quad (23)$$

とすれば (P は基底の変換行列)、基底 $\{\mathbf{v}_i\}$ に関する f の表現行列は $A' = P^{-1}AP$ である。式 (23) を使えば

$$\begin{aligned} (f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)) &= (f(\mathbf{e}_1), \dots, f(\mathbf{e}_n))P \\ &= (\mathbf{e}_1, \dots, \mathbf{e}_n)AP \\ &= (\mathbf{v}_1, \dots, \mathbf{v}_n)P^{-1}AP \end{aligned}$$

このように、一般にある $P \in GL_n(\mathbb{R})$ があって、 $A' = P^{-1}AP$ となるとき、 A と A' とは相似 (similar) または共益 (conjugate) であるという。相似な行列は同様の線形変換を別の基底によって表す行列と考えられる。線形代数の一つの基本的な利用用途として、与えられた行列 (線形変換) に対し、適当な基底をとり、対応する行列をできるだけ簡単な形にすることである。これを行列の標準化という。この章では行列の標準化に関連した基本的な定義や概念に関して説明する。

まず、部分空間 $W \subset V$ は $AW \subset W$ であるとき、 A - 不変 (A - invariant), または単に不変な部分空間であるという。

定理 7.1. W を A - 不変 な部分空間とし、 V の基底 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ を $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ が W の基底になるようにとる。そのとき基底 $\{\mathbf{v}_i\}$ に関して f_A の表現行列は

$$A' = \begin{pmatrix} A'_1 & A'_{12} \\ 0 & A'_2 \end{pmatrix}$$

の形になる。ここで、 A'_1 は f_A が引き起こす線形変換の基底 $\mathbf{v}_1, \dots, \mathbf{v}_r$ に関する表現行列である。

証明. $A' = (a'_{ij})$ とする。 W が A - 不変 より

$$(f_A(\mathbf{v}_1), \dots, f_A(\mathbf{v}_r)) = (\mathbf{v}_1, \dots, \mathbf{v}_r)A'_1$$

とできるから、基底を延長すれば

$$(f_A(\mathbf{v}_1), \dots, f_A(\mathbf{v}_n)) = (\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \mathbf{v}_n) \begin{pmatrix} A'_1 \\ 0 \end{pmatrix}$$

が得られる。つまり、 $r+1 \leq i \leq n$, $1 \leq j \leq r$ において、 $a'_{ij} = 0$ である。また、 $(f_A(\mathbf{v}_{r+1}), \dots, f_A(\mathbf{v}_n))$ においても

$$(f_A(v_{r+1}), \dots, f_A(v_n)) = (v_1, \dots, v_r, v_{r+1}, \dots, v_n) \begin{pmatrix} A'_{12} \\ A'_2 \end{pmatrix}$$

が成り立つ。よって、 $\begin{pmatrix} A'_1 \\ 0 \end{pmatrix}, \begin{pmatrix} A'_{12} \\ A'_2 \end{pmatrix}$ を並べれば A' の形になる。

この証明より、 A'_1 は f_A が引き起こす線形変換の基底 v_1, \dots, v_r に関する行列であり、 $\text{rank } A'_1 = r$ であることが分かる。

定理 7.2. V が

$$V = W_1 \oplus W_2, \quad W_i : A\text{-invariant}$$

の形に直和分解されるとする。そのとき、 V の基底 $\{v_i\}$ を $\{v_1, \dots, v_r\}, \{v_{r+1}, \dots, v_n\}$ がそれぞれ W_1, W_2 の基底になるようにとれば、対応する f_A の行列は

$$A' = \begin{pmatrix} A'_1 & 0 \\ 0 & A'_2 \end{pmatrix}$$

A'_1, A'_2 は f_A が W_1, W_2 に引き起こす線形変換の行列である。

証明.

$$(f_A(v_1), \dots, f_A(v_r)) = (v_1, \dots, v_r) A'_1 \quad (24)$$

$$(f_A(v_{r+1}), \dots, f_A(v_n)) = (v_{r+1}, \dots, v_n) A'_2 \quad (25)$$

とおけば、 W_1, W_2 は A -不変であるから $(24) \subset W_1, (25) \subset W_2$ でなければならない。よって、

$$(v_1, \dots, v_r, v_{r+1}, \dots, v_n) \begin{pmatrix} A'_1 \\ 0 \end{pmatrix} \subset W_1$$

$$(v_1, \dots, v_r, v_{r+1}, \dots, v_n) \begin{pmatrix} 0 \\ A'_2 \end{pmatrix} \subset W_2$$

より、 f_A の行列は上記の形になる。

これらの定理により、 V を不変部分空間の直和に分解できればそれに対応して行列 A が簡約された形になることがわかる。

7.2 固有値、固有空間

固有値、固有空間に入る前にまずは特性多項式、消去法の原理に関して概説する。

7.2.1 特性多項式

$A \in M_n(\mathbb{R})$ に対し

$$|xE - A| = \begin{vmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x - a_{nn} \end{vmatrix}$$

は x に関する n 次多項式になる。これを A の特性多項式 (characteristic polynominal) あるいは固有方程式といい、 $f(x; A)$ のように書くことにする。これが多項式になることを確認するにはその行列式を展開すれば良い。例えば

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

としたとき、

$$\begin{aligned} f(x; A) &= \begin{vmatrix} x - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & x - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & x - a_{33} \end{vmatrix} \\ &= (x - a_{11})(x - a_{22})(x - a_{33}) + (-a_{12} \cdot -a_{23} \cdot -a_{31}) \\ &\quad + (-a_{13} \cdot -a_{21} \cdot x - a_{32}) - (-a_{12} \cdot -a_{21} \cdot x - a_{33}) \\ &\quad - (x - a_{11} \cdot -a_{23} \cdot -a_{32}) - (-a_{13} \cdot x - a_{22} \cdot -a_{31}) \\ &= x^3 - x^2(a_{33} + a_{22} + a_{11}) + x(a_{22}a_{33} + a_{11}a_{22}) - a_{11}a_{22}a_{33}, \dots \end{aligned} \tag{26}$$

のような 3 次の多項式の形になる。これを一般化すると

$$f(x; A) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$$

のような n 次多項式になる。このとき、この多項式の次数は行列式を求める際に x がある対角成分どうしが掛けられるのがただ一回のみであることから分かるように、行列の rank と一致する。

7.2.2 消去法の原理

$A \in M_n(\mathbb{R})$ に対し、1 次方程式

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

が自明でない解 $\mathbf{x} \neq 0$ をもつためには、 $|A| = 0$ であることが必要十分である。

証明. $|A| \neq 0$ ならば、 A は可逆であるから、 $\mathbf{x} = 0$ が唯一の解である。したがって、自明でない解は存在しない。逆に $|A| = 0$ ならば、次元定理により、解の空間は次元が $n - \text{rank } A > 0$ の部分空間である。よって自明でない解が確かに存在する。

定義 7.3. 固有値、固有空間

まず一番簡単な例として、1 次元部分空間 $W = \langle \mathbf{v}_1 \rangle$ が A - 不変 になっているとする。定理 7.1 によれば、 $\{\mathbf{v}_1\}$ を延長して得られる基底 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ に関する f_A の行列は

$$A' = \begin{pmatrix} \alpha_1 & * \\ 0 & A'_2 \end{pmatrix}$$

の形になる。特に

$$A\mathbf{v}_1 = \alpha_1 \mathbf{v}_1$$

が成立する。一般に $\mathbf{v}_1 \neq 0$ に対して $A\mathbf{v}_1 = \alpha_1\mathbf{v}_1$ が成立する時、 α_1 を A あるいは f_A の固有値 (eigen value), \mathbf{v}_1 をそれに属する固有ベクトル (eigen vector) という。

また

$$V(A; \alpha_1) = \{\mathbf{v} \in V | A\mathbf{v} = \alpha_1\mathbf{v}\}$$

とけば、これは明らかに V の部分空間になる。 $V(A; \alpha_1)$ を α_1 に対応する固有空間 (eigen space) という。

定理 7.4. $\alpha \in \mathbb{R}$ が $A \in M_n(\mathbb{R})$ の固有値になるためには、 α が特性多項式 $f(x; A)$ の根になることが必要十分である。(根は特性方程式における解のこと)

証明. 定義によれば α が A の固有値になるためには、一次方程式

$$(A - \alpha E)\mathbf{x} = 0$$

が自明でない解 ($\mathbf{x} \neq 0$) を持つことが必要十分である。よって、消去法の原理により

$$\alpha : \text{eigen value of } A \iff |A - \alpha E| = 0$$

である。

これを言い換えると次の定理を得る。

定理 7.5. α に対応する固有空間 $V(A; \alpha)$ は一次方程式 $(A - \alpha E)\mathbf{x} = 0$ の解空間となりその次元は

$$\dim V(A; \alpha) = n - \text{rank}(A - \alpha E)$$

である。

複素数の範囲では n 次方程式 $f(x; A) = 0$ は重解まで考慮すると、 n 個の根を持つ。したがって、 A を \mathbb{C}^n の線形変換と考えれば、これも重解を考慮して n 個の固有値を持つ。したがって、固有値について考えるときには \mathbb{R}^n ではなく、 \mathbb{C}^n で考えるほうがより自然である。しかし、もし固有値がすべて実数ならば、 \mathbb{R}^n で考えても十分である。(これは代数学の基本定理により証明されている。)

定理 7.6. $\alpha_1, \dots, \alpha_r$ を A の相異なる固有値 $\mathbf{v}_1, \dots, \mathbf{v}_r$ をそれぞれに属する固有ベクトルとすれば $\mathbf{v}_1, \dots, \mathbf{v}_r$ は線形独立である。

証明. 数学的帰納法で示す。 $r = 1$ の時は明らか。($c_1\mathbf{v}_1 = 0$ となるのは、 $c_1 = 0$ であるから)

$r - 1$ の時

$$c_1\mathbf{v}_1 + \dots + c_{r-1}\mathbf{v}_{r-1} + c_r\mathbf{v}_r = 0, \quad c_i = 0 \quad (27)$$

と仮定する。式 (27) に左から A を掛けて

$$c_1A\mathbf{v}_1 + \dots + c_{r-1}A\mathbf{v}_{r-1} + c_rA\mathbf{v}_r = c_1\alpha_1\mathbf{v}_1 + \dots + c_{r-1}\alpha_{r-1}\mathbf{v}_{r-1} + c_r\alpha_r\mathbf{v}_r = 0 \quad (28)$$

さらに式 (27) に a_r を掛けると

$$c_1\alpha_r\mathbf{v}_1 + \dots + c_{r-1}\alpha_r\mathbf{v}_{r-1} + c_r\alpha_r\mathbf{v}_r = 0 \quad (29)$$

ここで

$$\begin{aligned}
(28) - (29) &= c_1(\alpha_1 - \alpha_r)\mathbf{v}_1 + \cdots + c_{r-1}(\alpha_{r-1} - \alpha_r)\mathbf{v}_{r-1} + c_r(\alpha_r - \alpha_r)\mathbf{v}_r \\
&= c_1(\alpha_1 - \alpha_r)\mathbf{v}_1 + \cdots + c_{r-1}(\alpha_{r-1} - \alpha_r)\mathbf{v}_{r-1} \\
&= 0
\end{aligned}$$

とおくと、仮定より相異なる固有値に属する $r-1$ 個のベクトル $\mathbf{v}_1, \dots, \mathbf{v}_{r-1}$ は線型独立より係数 c_i が 0 になる必要がある。すなわち

$$c_1(\alpha_1 - \alpha_r) + \cdots + c_{r-1}(\alpha_{r-1} - \alpha_r) = 0$$

しかし、固有値はすべて異なるという仮定より

$$c_1 = \cdots = c_{r-1} = 0$$

である。これを式 (27) に代入すれば $c_r = 0$ が得られる。よて、 $r=1, r-1$ で成り立てば、 r でも同様に成り立つ。したがって、 A の相異なる固有値に属する固有ベクトル $\mathbf{v}_1, \dots, \mathbf{v}_r$ は線型独立である。

この結果より（帰納法を $r=1, 2, 3, \dots$ のように進めていけば）

$$V = \sum_{i=1}^r V(A; \alpha_i) = V(A; \alpha_1) \oplus \cdots \oplus V(A; \alpha_r) \quad (\text{直和})$$

であることが分かる。

例 7.7.

$$A = \begin{pmatrix} 6 & -3 & -7 \\ -1 & 2 & 1 \\ 5 & -3 & -6 \end{pmatrix}$$

の固有値、固有ベクトルは以下の手順で求まる。

A の特性多項式

$$|\lambda E - A| = \begin{vmatrix} \lambda - 6 & 3 & 7 \\ 1 & \lambda - 2 & -1 \\ -5 & 3 & \lambda + 6 \end{vmatrix} = 0$$

を解くと

$$\lambda^3 - 2\lambda^2 - \lambda + 2 = (\lambda - 1)(\lambda - 2)(\lambda + 1) = 0$$

より、固有値は $\lambda = 1, 2, -1$ である。以降、各固有ベクトルを求めるには、各固有値に関して

$$(\lambda E - A) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

の連立方程式を解けば良い。（一般的には $(\lambda E - A)$ に基本変形を施して、行列を単純な形にすると連立方程式が簡単に解けるようになる。）

$\lambda = 1$ のとき

$$\begin{pmatrix} -5 & 3 & 7 \\ 1 & -1 & -1 \\ -5 & 3 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

を解くと、 $y = 1, z = 1, x = 2$ より、 $z = \alpha_1$ とすれば

$$\mathbf{x}_1 = \alpha_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

したがって、固有空間は $\left\langle \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right\rangle$ である。これより、固有値 0 に属する固有ベクトルは $\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ である。

$\lambda = 2$ のとき

$$\begin{pmatrix} -4 & 3 & 7 \\ 1 & 0 & -1 \\ -5 & 3 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

を解くと、 $x = z$ これを使って $-4x + 3y + 7x = 3x + 3y = 0$ を解けば、 $y = -x$ である。

$x = \alpha_2$ とすれば

$$\mathbf{x}_2 = \alpha_2 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

であるから、固有値 2 に属する固有ベクトルは

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

$\lambda = -1$ のとき

$$\begin{pmatrix} -7 & 3 & 7 \\ 1 & -3 & -1 \\ -5 & 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

これは、連立方程式が複雑なので簡単のため基本変形を施すと

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

であるから、 $y = 0, x = z$ より $x = \alpha_3$ とおけば

$$\mathbf{x}_3 = \alpha_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

である。よって、固有値 -1 に属する固有ベクトルは

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

なお、 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ は線形独立である。

7.3 実対称行列の対角化

$A \in M_n(\mathbb{R})$ を対称とする。そのとき、 $\langle \mathbf{x}, A\mathbf{y} \rangle = \langle {}^tA\mathbf{x}, \mathbf{y} \rangle$ により任意の $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ に対して

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle = 0. \quad (30)$$

が成立する。

定義 7.8. 実対称行列

$A \in M_n(\mathbb{R})$ が実対称行列であるとは、 A が

- (1) ${}^tA = A$ (対称性)
- (2) $\bar{A} = A$ (複素共役)

を満たすときを言う。特に (1) を満たすものを対称行列、(2) を満たすものを実行列という。

補題 7.9. 実対称行列 A の固有値は実数である。

証明. 複素数の範囲で A の固有値 α をとり、 $\mathbf{v} \neq 0$ をそれに属する一つの固有ベクトルとする。

$A = (a_{ij}), \mathbf{v} = (v_i)$ とし、 $\bar{\mathbf{v}} = (\bar{v}_i)$ とすれば、実対称行列の定義より $A = {}^t\bar{A}$ となるから

$$\begin{aligned} {}^t\bar{\mathbf{v}}A\mathbf{v} &= {}^t\bar{\mathbf{v}}(A\mathbf{v}) = {}^t\bar{\mathbf{v}}(\alpha\mathbf{v}) \\ &= \alpha {}^t\bar{\mathbf{v}}\mathbf{v} = {}^t\bar{\mathbf{v}}{}^tA\mathbf{v} = {}^t(\bar{A}\bar{\mathbf{v}})\mathbf{v} \\ &= {}^t(\alpha\bar{\mathbf{v}})\mathbf{v} = \bar{\alpha} {}^t\bar{\mathbf{v}}\mathbf{v} \end{aligned}$$

よって

$$\alpha - \bar{\alpha} {}^t\bar{\mathbf{v}}\mathbf{v} = 0$$

ここで、

$${}^t\bar{\mathbf{v}}\mathbf{v} = \sum_{i=1}^n \bar{v}_i v_i = \sum_{i=1}^n |v_i|^2 > 0$$

であるから、 $\alpha - \bar{\alpha} = 0$ 。よって、 $\alpha = \bar{\alpha}$ 。つまり α は実数である。

補題 7.10. A を実対称行列とする。 $W \in V$ が A - 不変 ならば、 W^\perp も A - 不変である。

証明. $\mathbf{x} \in W, \mathbf{y} \in W^\perp$ を任意に取ってくる。 $\langle \mathbf{x}, A\mathbf{y} \rangle = 0$ を言えば十分である。 W は A - 不変 であるから、 $A\mathbf{x} \in W$ 。よって、式 (30) により

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A\mathbf{y} \rangle = 0.$$

定理 7.11. 実対称行列 $A \in M_n(\mathbb{R})$ に対して、適当な直交行列 $T \in O(n)$ をとれば、

$$A' = T^{-1}AT = {}^tTAT = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{pmatrix}$$

と対角化される。このことを A の基底の直交変換によって対角化されると言い表す。

証明. A の固有値 $\alpha_i (1 \leq i \leq n)$ に対応する固有ベクトルからなる正規直交基底 $\{t_1, \dots, t_n\}$ が存在することをいえばよい。これを n に関する帰納法で証明する。 $n = 1$ のとき明らかに成立する。よって、 $n \geq 2$ とし、 α_1 を A の固有値とする。実対称行列の固有値は実数であることより、 $\alpha_1 \in \mathbb{R}$ であるから、それに属する固有ベクトル $v_1 \in \mathbb{R}^n$ をとることができる。これを正規化して

$$t_1 = \frac{v_1}{\|v_1\|}$$

とおけば、 t_1 も α_1 に属する固有ベクトルで $\|t_1\| = 1$ である。

$$W_1 = \langle t_1 \rangle, \quad W_2 = W_1^\perp$$

とおく。 W_1 は A - 不変 であるから、 W_2 も A - 不変 で

$$V = W_1 \oplus W_2$$

と直和分解される。(定理 (6.16) より) 次に、 W_2 から適当に基底を取り、それらにシュミットの正規直交化法を適用して $\{t_2, \dots, t_n\}$ とする。

$$T = (t_1, t_2, \dots, t_n)$$

とおけば、 T は直交行列になる。直交行列に関する定理 (6.12) により、 ${}^t T A T = ({}^t t_i A t_j)$ とおけば、その ij 成分は $A \langle t_i, t_j \rangle = \alpha_i \langle t_i, t_j \rangle$ のように固有値と内積の積 (実数) で表される。ここで、 ${}^t T = T^{-1}$, ${}^t T T = E$ より、 ${}^t T A T$ の ij 成分はさらに $\alpha_i \delta_{ij}$ と書ける。特に $i = 1$ のときの固有値を α_1 とすれば、 $j = 1$ のとき $\alpha_1 \cdot 1 = \alpha_1$ それ以外で、 $\alpha_1 \cdot 0 = 0$ になる。よって

$$A' = T^{-1} A T = \begin{pmatrix} \alpha_1 & 0 \\ 0 & A'_2 \end{pmatrix}$$

の形になる。ここで、

$${}^t A' = ({}^t T A T) = {}^t T {}^t A {}^t T = {}^t T A T = A'$$

より、 A'_2 は $(n-1)$ 次対称行列である。ゆえに帰納法の仮定により

$$T_2^{-1} A'_2 T_2 = \begin{pmatrix} \alpha_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_n \end{pmatrix}$$

となるような直交行列 T_2 が存在し

$$A' = \begin{pmatrix} 1 & 0 \\ 0 & T_2 \end{pmatrix}$$

とすると、 A' は直交行列で

$$\begin{aligned}
A' &= T^{-1}AT = \begin{pmatrix} 1 & 0 \\ 0 & T_2^{-1} \end{pmatrix} \begin{pmatrix} \alpha_1 & 0 \\ 0 & A'_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & T_2 \end{pmatrix} \\
&= \begin{pmatrix} \alpha_1 & 0 \\ 0 & T_2^{-1}A'_2T_2 \end{pmatrix} \\
&= \begin{pmatrix} \alpha_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_2 \end{pmatrix}
\end{aligned}$$

よって、 n の場合にも定理は成立する。

例題 7.12. $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ を対角化せよ。

解答. まずは A の固有値、固有ベクトルを求める。 A の特性多項式は

$$\begin{aligned}
\begin{vmatrix} x-2 & -1 & -1 \\ -1 & x-2 & -1 \\ -1 & -1 & x-2 \end{vmatrix} &= (x-2)^3 - 2 - 3(x-2) \\
&= x^3 - 6x^2 + 9x - 4 \\
&= (x-4)(x-1)^2
\end{aligned}$$

より、解は $x = 4, x = 1$ (重解) が得られる。よって、 A の固有値は $4, 1$ である。まず、固有値 4 に対応する固有ベクトルを求める。

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

を解けば、 $x_1 = x_2 = x_3 = 1$ であるから、固有値 4 に対応する固有ベクトルは $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ である。ここで、このベクトルが張る固有空間は

$$W_1 = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

となる。 $\text{rank } A = 3, \text{rank } W_1 = 1$ であるから、 $W_1^\perp = \dim A - \dim W_1 = 2$ にならなければならない。(つまり、このことは n 次方程式が重解を含むときもかならずランクに対応した数の固有ベクトルがあるということ。)

このことを考慮して、 $\lambda = 1$ のときの固有ベクトルを求める。

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff x_1 + x_2 + x_3 = 0$$

これを解くと、 $x_1 = -x_2 - x_3$ より、 $x_2 = -\alpha_1, x_3 = -\alpha_2$ とおけば

$$\begin{pmatrix} \alpha_1 + \alpha_2 \\ -\alpha_1 \\ -\alpha_2 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

である。したがって、固有値 1 に属する固有ベクトルは

$$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

である。(この二つは線形独立であることは容易に確かめられる。)

ここで得られた 3 つの基底 (固有ベクトル)

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

は $\langle \mathbf{x}_2, \mathbf{x}_3 \rangle \neq 0$ であるから、直交していない。これをシュミットの正規直交化法で直交化すると

$$\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix}$$

が得られる。これを並べて

$$T = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}$$

とおけば、 T は直交行列で

$$T^{-1}AT = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

と対角化される。

参考文献

[1] 佐武一郎 [線形代数] 共立出版株式会社