線形代数

1 ベクトルと行列

定義 1.1. 実数全体の集合を $\mathbb{R}$ で表す。

定義 1.2. X を集合とする。このとき、自然数 n に対して、

$$X^n := \{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \dots, x_n \in X \}$$

と定義する。

定義 1.3. n を自然数とする。このとき、$\mathbb{R}^n$ の元を n **次元ベクトル**という。$\mathbb{R}^n$ の元 $(0, \dots, 0)$ を n **次元** 0 **ベクトル**といい、$\mathbf{0}$ で表す。任意の $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n, a \in \mathbb{R}$ に対して、

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad a\mathbf{x} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$$

と定義する。

例 1.1. $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ は 2 次元ベクトルであり、$\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}, 5 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix}$ である。

定義 1.4. m, n を自然数とする。このとき、mn 個の $\mathbb{R}$ の元 $a_{ij} (i = 1, \dots, m, j = 1, \dots, n)$ から成る表

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \tag{1}$$

を $m \times n$ **行列**という。(m, n) を (1) の**型**という。特に、$m = n$ であるとき、(1) を n **次正方行列**という。(1) において、a_{ij} を (1) の (i, j) **成分**という。行列を完全な形で書くには大きなスペースを要するので、(1) を (a_{ij}) で略記することにする。$A = (a_{ij}), B = (b_{ij})$ を $m \times n$ 行列とし、$c \in \mathbb{R}$ とする。このとき、$a_{ij} + b_{ij}$ を (i, j) 成分とする $m \times n$ 行列を $A + B$ で表し、ca_{ij} を (i, j) 成分とする $m \times n$ 行列を cA で表す。$A + B$ を A, B の**和**という。すなわち、

$$A + B = (a_{ij} + b_{ij}), \quad cA = (ca_{ij})$$

である。型が違う行列の間では和は定義されない。

例 1.2. $A = \begin{pmatrix} 2 & -1 & 5 \\ 0 & 3 & -6 \end{pmatrix}, B = \begin{pmatrix} 3 & 0 & -7 \\ 2 & 4 & -2 \end{pmatrix}$ とする。このとき、

$$A + B = \begin{pmatrix} 5 & -1 & -2 \\ 2 & 7 & -8 \end{pmatrix}, \quad 3A = \begin{pmatrix} 6 & -3 & 15 \\ 0 & 9 & -18 \end{pmatrix}$$

である。

定義 1.5. m, n, s, t を自然数とし、$A = (a_{ij})$ を $m \times n$ 行列、$B = (b_{ij})$ を $s \times t$ 行列とする。$n = s$ であるとき、$\sum_{k=1}^{n} a_{ik} b_{kj}$ を (i, j) 成分とする $m \times t$ 行列を AB で表す。AB を A, B **の積**という。すなわち、

$$AB = (\sum_{k=1}^{n} a_{ik} b_{kj})$$

である。$n \neq s$ であるときは、A, B の積は定義されない。

例 1.3. $A = \begin{pmatrix} 2 & -1 \\ 0 & 4 \end{pmatrix}, B = \begin{pmatrix} 1 & -5 & 3 \\ 8 & -2 & 0 \end{pmatrix}$ とする。A は 2×2 行列であり、B は 2×3 行列であるから、AB は定義され、

$$AB = \begin{pmatrix} -6 & -8 & 6 \\ 32 & -8 & 0 \end{pmatrix}$$

である。

定義 1.6. n を自然数とし、$A = (a_{ij})$ を n 次正方行列とする。このとき、$1 \leq i, j \leq n$ に対して、$(a_{i1}, \ldots, a_{in})$ を A **の第** i **行**といい、$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ を A **の第** j **列**という。A の第 i 行と第 j 列を A から取り除くことによって、得られる $(n-1) \times (n-1)$ 行列を A_{ij} で表す。

例 1.4. $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ とする。このとき、

$$A_{11} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}$$

である。

定義 1.7. n を自然数、$A = (a_{ij})$ を n 次正方行列とする。このとき、以下のように実数 $\det(A)$ を定義する。

(1) $n = 1$ であるときは、$\det(A) := a_{11}$ と定義する。
(2) $n > 1$ であるときは、$\det(A) := \sum_{k=1}^{n} (-1)^{k+1} a_{1k} \det(A_{1k})$ と定義する。

例 1.5. n を自然数、$A = (a_{ij})$ を n 次正方行列とする。$n = 2$ であるとき、

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

であり、$n = 3$ であるとき、

$$\det(A) = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}$$

である。

定義 1.8. n を自然数、A を n 次正方行列とし、$\mathbf{x} \in \mathbb{R}^n$ とする。x を $n \times 1$ 行列と見なすことにより、$A, \mathbf{x}$ の積 $A\mathbf{x}$ が定義される。$\lambda \in \mathbb{R}$ とする。$A\mathbf{v} = \lambda\mathbf{v}, \mathbf{v} \neq \mathbf{0}$ となる $\mathbf{v} \in \mathbb{R}^n$ が存在するとき、λ を A の**固有値**という。λ が A の固有値であるとき、$A\mathbf{y} = \lambda\mathbf{y}, \mathbf{y} \neq \mathbf{0}$ となる $\mathbf{y} \in \mathbb{R}^n$ を λ の**固有ベクトル**という。

定理 1.1. n を自然数、A を n 次正方行列とし、$\lambda \in \mathbb{R}$ とする。このとき、以下のことは同等である。

(1) λ は A の固有値である。
(2) $\det(A - \lambda I_n) = 0$

例 1.6. $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ とする。このとき、

$$\det(A - \lambda I_2) = (1-\lambda)(3-\lambda) - 8 = \lambda^2 - 4\lambda - 5 = (\lambda - 5)(\lambda + 1)$$

であるから、定理 1.1 より、A の固有値は $5, -1$ である。

定義 1.9. n を自然数とする。このとき、

$$I_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

と定義する。すなわち、I_n は対角線上の成分が全て 1 であり、他の成分が全て 0 であるような n 次正方行列である。A を n 次正方行列とする。このとき、$AI_n = I_nA = A$ が成り立つ。$AB = I_n, BA = I_n$ となる n 次正方行列 B が存在するとき、A を**正則行列**という。A が正則行列であるとき、$AC = I_n, CA = I_n$ となる n 次正方行列 C がただ一つ存在する。そこで、この C を A^{-1} で表す。

定理 1.2. $A = (a_{ij})$ を n 次正方行列とする。このとき、以下のことは同等である。

(1) A は正則である。
(2) $\det(A) \neq 0$ である。

特に、$n = 2$ かつ、A が正則であるとき、

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

が成り立つ。

定理 1.3. n を自然数、A を n 次正方行列とする。A が互いに相異なる n 個の固有値 $\lambda_1, \dots, \lambda_n$ を持つと仮定し、$\mathbf{v}_1, \dots, \mathbf{v}_n$ をそれぞれ $\lambda_1, \dots, \lambda_n$ の固有ベクトルとする。$\mathbf{v}_1, \dots, \mathbf{v}_n$ を横に並べることによって得られる n 次正方行列を V とおき、

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

とおく。すなわち、Λ は対角線上の成分が A の固有値であり、他の成分が全て 0 であるような n 次正方行列である（このように、対角成分以外の成分が全て 0 である行列を**対角行列**という)。このとき、以下のことが成り立つ。

(1) V は正則である。
(2) $AV = V\Lambda$ が成り立つ。

これらのことから、$A = V\Lambda V^{-1}$ が成り立つ。このように n 次正方行列 A を上述のような 3 つの行列の積に分解することを A を**固有値分解**するという。

例 1.7. $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ とする。A の固有値分解を求める。例 1.6 より、A の固有値は $5, -1$ である。連立方程式

$$A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 5 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

を解くことにより、$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ が 5 の固有ベクトルであり、$\begin{pmatrix} 1 \\ -1/2 \end{pmatrix}$ が -1 の固有ベクトルであることが分かる。

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

とおく。定理 1.3 より、V は正則であり、定理 1.2 より、

$$V^{-1} = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & -2/3 \end{pmatrix}$$

である。よって、A の固有値分解

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1/2 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & -2/3 \end{pmatrix}$$

を得る。

定義 1.10. m, n を自然数とし、$A = (a_{ij})$ を $m \times n$ 行列とする。このとき、a_{ji} を (i, j) 成分とする $n \times m$ 行列を A^t で表す。A^t を A の**転置行列**という。すなわち、

$$A^t = (a_{ji})$$

である。$A = A^t$ であるとき、A を**対称行列**という。

例 1.8. $A = \begin{pmatrix} 1 & -5 & 10 \\ 6 & 3 & -4 \end{pmatrix}$ とする。このとき、

$$A^t = \begin{pmatrix} 1 & 6 \\ -5 & 3 \\ 10 & -4 \end{pmatrix}$$

である。

定義 1.11. n を自然数、A を n 次正方行列とする。A が正則であり、$A^{-1} = A^t$ が成り立つとき、A を**直行行列**という。

定義 1.12. n を自然数とし、$\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ とする。このとき、

$$\|\mathbf{x}\| := \sqrt{x_1^2 + \cdots + x_n^2}$$

と定義する。$\|\mathbf{x}\| = 1$ であるとき、$\mathbf{x}$ を**単位ベクトル**という。任意の $\mathbf{v} \in \mathbb{R}^n$ に対して、$\mathbf{v} \neq 0$ ならば、$\|\mathbf{v}\| \neq 0$ であり、$\|\mathbf{v}\|^{-1}\mathbf{v}$ は単位ベクトルとなる。

定理 1.4. n を自然数、A を n 次正方対称行列とする。A が互いに相異なる n 個の固有値 $\lambda_1, \ldots, \lambda_n$ を持つと仮定し、$\mathbf{v}_1, \ldots, \mathbf{v}_n$ をそれぞれ $\lambda_1, \ldots, \lambda_n$ の固有ベクトルとする。$\mathbf{v}_1/\|\mathbf{v}_1\|, \ldots, \mathbf{v}_n/\|\mathbf{v}_n\|$ を横に並べることによって得られる n 次正方行列を V とおき、

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ldots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

とおく。すなわち、Λ は対角線上の成分が A の固有値であり、他の成分が全て 0 であるような n 次正方行列である。このとき、以下のことが成り立つ。

(1) V は直行行列である。
(2) $AV = V\Lambda$ が成り立つ。

これらのことから、$A = V\Lambda V^t$ が成り立つ。すなわち、A は直行行列を用いて、固有値分解できる。

定理 1.5. m, n を自然数、A を $m \times n$ 行列とする。以下のことが成り立つと仮定する。

(1) AA^t は相異なる m 個の固有値 $\lambda_1, \ldots, \lambda_m$ を持つ。
(2) A^tA は相異なる n 個の固有値 $\mu_1, \ldots, \mu_n$ を持つ。
(3) $m \leq n$

定理 1.4 を用いれば、$AA^t = U\Lambda U^t, AA^t = VMV^t$ となる直行行列 U, V と対角行列 Λ, M が構成できる。ここで、$1 \leq i \leq m$ に対して、$\sigma_i = \sqrt{\lambda_i}$ とおき、

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \ddots & \ldots & \vdots & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 & 0 & \ldots & 0 \\ 0 & \ldots & 0 & \sigma_m & 0 & \ldots & 0 \end{pmatrix}$$

とおく。このとき、$A = U\Sigma V^t$ が成り立つ。このように $m \times n$ 行列 A を上述のような 3 つの行列の積に分解することを A を**特異値分解**するという。$n \leq m$ である場合も同様なことが成り立つ。

例 1.9. $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ とする。A の固有値分解を求める。$AA^t = \begin{pmatrix} 14 & 10 \\ 10 & 14 \end{pmatrix}$ であり、AA^t の固有値は $24, 4$ である。よって、定理 1.4 を用いれば、AA^t の固有値分解

$$AA^t = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 24 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

を得る。一方、$A^tA = \begin{pmatrix} 10 & 8 & 6 \\ 8 & 8 & 8 \\ 6 & 8 & 10 \end{pmatrix}$ であり、A^tA の固有値は $24, 4, 0$ である。よって、定理 1.4 を用いれば、A^tA の固有値分解

$$\begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 24 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{pmatrix}$$

を得る。よって、定理 1.5 より、A の特異値分解

$$A = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2\sqrt{6} & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{pmatrix}$$

を得る。

参考文献

[1] Marc Peter Deisenroth., A. Aldo Faisal., Cheng Soon Ong. (2020). Mathematics for machine learning. Cambridge University Press.

[2] 小縣信也., 斎藤翔汰., 溝口聡., 若杉一幸. (2021). ディープラーニング E 資格エンジニア問題集 インプレス.

[3] 松坂和夫. (1980). 線形代数入門 岩波書店