

ベクトルの内積とその応用

戸瀬 信之

ITOSE PROJECT

MSF2019, Lec 04, 2019年04月30日 (平成最後の講義), V02 2020/03

列ベクトル

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbf{R}^n \quad (1)$$

$\lambda \in \mathbf{R}$ に対して

$$\vec{x} + \vec{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \vec{x} - \vec{y} = \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \\ \vdots \\ x_n - y_n \end{pmatrix}, \quad \lambda \vec{x} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

行ベクトル

n 次元行ベクトル

$$\mathbf{x} = (x_1 \ x_2 \ \cdots \ x_n)$$

(転置)

$${}^t(x_1 \ x_2 \ \cdots \ x_n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad {}^t \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1 \ x_2 \ \cdots \ x_n)$$

列ベクトルと行ベクトルの積

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

$$\vec{a}\vec{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

後に $1 \times n$ 行列と $n \times 1$ 行列の積となる。

ベクトルの内積

$\vec{x}, \vec{y} \in \mathbf{R}^n$ に対して

$$(\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y} = \sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

統計学に関連することを勉強するときは公式

$${}^t \vec{x} \vec{y} = (x_1 \ x_2 \ \cdots \ x_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = (\vec{x}, \vec{y})$$

が有用である。

ベクトルのノルム

$\vec{x} \in \mathbf{R}^n$ に対して

$$\|\vec{x}\|^2 = (\vec{x}, \vec{x}) = \sum_{i=1}^n x_i^2 \quad (2)$$

$$\|\vec{x}\| \geq 0$$

$$\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$$

内積の基本公式

$\vec{x}, \vec{y}, \vec{z} \in \mathbf{R}^n$ と $\lambda \in \mathbf{R}$ に対して

$$(\vec{x} + \vec{y}, \vec{z}) = (\vec{x}, \vec{z}) + (\vec{y}, \vec{z}) \quad (3)$$

$$(\vec{x}, \vec{y} + \vec{z}) = (\vec{x}, \vec{y}) + (\vec{x}, \vec{z}) \quad (4)$$

$$(\lambda \vec{x}, \vec{y}) = (\vec{x}, \lambda \vec{y}) = \lambda (\vec{x}, \vec{y}) \quad (5)$$

$$(\vec{x}, \vec{y}) = (\vec{y}, \vec{x}) \quad (6)$$

内積の有用な公式

$$\|\vec{x} \pm \vec{y}\|^2 = \|\vec{x}\|^2 \pm 2(\vec{x}, \vec{y}) + \|\vec{y}\|^2$$

(証明)

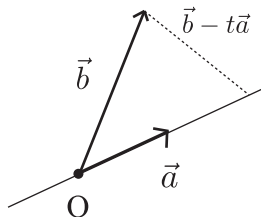
$$\begin{aligned}\|\vec{x} + \vec{y}\|^2 &= (\vec{x} + \vec{y}, \vec{x} + \vec{y}) \quad ((2) \text{ から}) \\ &= (\vec{x}, \vec{x} + \vec{y}) + (\vec{y}, \vec{x} + \vec{y}) \quad ((3) \text{ から}) \\ &= (\vec{x}, \vec{x}) + (\vec{x}, \vec{y}) + (\vec{y}, \vec{x}) + (\vec{y}, \vec{y}) \quad ((4) \text{ から}) \\ &= (\vec{x}, \vec{x}) + (\vec{x}, \vec{y}) + (\vec{x}, \vec{y}) + (\vec{y}, \vec{y}) \quad ((5) \text{ から}) \\ &= \|\vec{x}\|^2 + 2(\vec{x}, \vec{y}) + \|\vec{y}\|^2 \quad ((6) \text{ から})\end{aligned}$$

問題

$\vec{a} \neq \vec{0}$ であるとき

$$\|\vec{b} - t\vec{a}\|^2$$

を最小にする $t \in \mathbf{R}$ を求める。



$$\begin{aligned}\|\vec{b} - t\vec{a}\|^2 &= \|\vec{b}\|^2 - 2t(\vec{a}, \vec{b}) + t^2\|\vec{a}\|^2 \\ &= \|\vec{a}\|^2 \left(t - \frac{(\vec{a}, \vec{b})}{\|\vec{a}\|^2} \right)^2 + \|\vec{b}\|^2 - \frac{(\vec{a}, \vec{b})^2}{\|\vec{a}\|^2}\end{aligned}$$

解答

$$t = t_0 = \frac{(\vec{a}, \vec{b})}{\|\vec{a}\|^2}$$

$\vec{w} = t_0\vec{a} = \frac{(\vec{a}, \vec{b})}{\|\vec{a}\|^2}\vec{a}$ は \vec{b} の \vec{a} 方向の正射影

$$\begin{aligned}(\vec{b} - t_0\vec{a}, \vec{a}) &= (\vec{b}, \vec{a}) - t_0(\vec{a}, \vec{a}) \\ &= (\vec{b}, \vec{a}) - \frac{(\vec{a}, \vec{b})}{\|\vec{a}\|^2} \cdot \|\vec{a}\|^2 = 0\end{aligned}$$

$$(\vec{b} - t_0\vec{a}) \perp \vec{a}$$

データ解析と列ベクトル

2変量のデータを考えます。

$$\vec{x} = \frac{1}{\sqrt{n}} \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}, \vec{y} = \frac{1}{\sqrt{n}} \begin{pmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{pmatrix}$$

x_1	y_1
x_2	y_2
\vdots	\vdots
x_n	y_n

分散・共分散

$$\|\vec{x}\|^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = V(x) \quad (x \text{ の分散})$$

$$(\vec{x}, \vec{y}) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = C_{xy} \quad (x \text{ と } y \text{ の共分散})$$

データ解析と列ベクトル (その2)

新たな変量 $z = ax + by$ (a, b は定数)

$$z_i = ax_i + by_i \quad (i = 1, 2, \dots, n)$$

このとき $\bar{z} = a\bar{x} + b\bar{y}$

$$\begin{aligned}\vec{z} &= \frac{1}{\sqrt{n}} ((ax_i + by_i) - (a\bar{x} + b\bar{y})) \\ &= a \left(\frac{x_i - \bar{x}}{\sqrt{n}} \right) + b \left(\frac{y_i - \bar{y}}{\sqrt{n}} \right) = a\vec{x} + b\vec{y}\end{aligned}$$

z の分散は

$$\begin{aligned}V(z) &= \|\vec{z}\|^2 = \|a\vec{x} + b\vec{y}\|^2 \\ &= a^2\|\vec{x}\|^2 + 2ab(\vec{x}, \vec{y}) + b^2\|\vec{y}\|^2 \\ &= a^2V(x) + 2abC_{xy} + b^2V(y)\end{aligned}$$

Cauchy-Schwartz の不等式

定理 $\vec{x}, \vec{y} \in \mathbf{R}^n$ とすると

$$|(\vec{x}, \vec{y})| \leq \|\vec{x}\| \cdot \|\vec{y}\|$$

証明 すべての $t \in \mathbf{R}$ に対して

$$0 \leq \|t\vec{x} - \vec{y}\|^2 = t^2\|\vec{x}\|^2 - 2t(\vec{x}, \vec{y}) + \|\vec{y}\|^2$$

$\vec{x} = \vec{0}$ のとき「不等式」は明らかに成立。

$\vec{x} \neq \vec{0}$ のとき、従って $\|\vec{x}\| > 0$ のとき

$$(\vec{x}, \vec{y})^2 - \|\vec{x}\|^2 \cdot \|\vec{y}\|^2 \leq 0$$

注意 $A > 0$ のとき、

$$At^2 + 2Bt + C \geq 0 \quad (t \in \mathbf{R}) \Leftrightarrow B^2 - AC \leq 0$$

Cauchy-Schwartz の不等式の応用

相関係数

$$\rho_{xy} = \frac{C_{xy}}{\sqrt{V(x)}\sqrt{V(y)}} = \frac{(\vec{x}, \vec{y})}{\|\vec{x}\| \cdot \|\vec{y}\|}$$

定理

$$-1 \leq \rho_{xy} \leq 1$$

注意 $|\rho_{xy}| \rightarrow 1$ のとき、 $x - y$ 平面上、データはより直線に近く分布する。(これについては次回)

問題 II

$\vec{a}, \vec{b} \in \mathbf{R}^n$ が

$$\vec{a} \nparallel \vec{b}$$

を満たすときに, $\vec{c} \in \mathbf{R}^n$ に対して

$$\|\vec{c} - x\vec{a} - y\vec{b}\|^2$$

を最小化する $x, y \in \mathbf{R}$ を求める.

まず $\vec{a} \perp \vec{b}$ のとき

$\vec{a} \perp \vec{b}$ の場合を考える.

$$\begin{aligned}\|\vec{c} - x\vec{a} - y\vec{b}\|^2 &= \|\vec{c}\|^2 - 2(\vec{c}, x\vec{a} + y\vec{b}) + \|x\vec{a} + y\vec{b}\|^2 \\ &= x^2\|\vec{a}\|^2 + y^2\|\vec{b}\|^2 - 2x(\vec{c}, \vec{a}) - 2y(\vec{c}, \vec{b}) + \|\vec{c}\|^2 \\ &= \|\vec{a}\|^2 \left(x - \frac{(\vec{c}, \vec{a})}{\|\vec{a}\|^2} \right)^2 + \|\vec{b}\|^2 \left(y - \frac{(\vec{c}, \vec{b})}{\|\vec{b}\|^2} \right)^2 \\ &\quad + \|\vec{c}\|^2 - \frac{(\vec{c}, \vec{a})^2}{\|\vec{a}\|^2} - \frac{(\vec{c}, \vec{b})^2}{\|\vec{b}\|^2}\end{aligned}$$

から

$$x = \frac{(\vec{c}, \vec{a})}{\|\vec{a}\|^2}, \quad y = \frac{(\vec{c}, \vec{b})}{\|\vec{b}\|^2}$$

のとき最小となる。この意味は？

グラム・シュミットの直交化一例

$\vec{a} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ に対して \vec{b} の \vec{a} 方向の直交射影は

$$\vec{w} = \frac{(\vec{b}, \vec{a})}{\|\vec{a}\|^2} \vec{a} = \frac{1}{3} \vec{a}$$

$$\vec{b} - \vec{w} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 5 \\ -4 \\ -1 \end{pmatrix}$$

ここで

$$\vec{p} = \frac{1}{\|\vec{a}\|} \vec{a} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{q} = \frac{1}{\|\vec{b} - \vec{w}\|} (\vec{b} - \vec{w}) = \frac{1}{\sqrt{42}} \begin{pmatrix} 5 \\ -4 \\ -1 \end{pmatrix}$$

は $\|\vec{p}\| = \|\vec{q}\| = 1$, $(\vec{p}, \vec{q}) = 0$ を満たし

$$L := \{x\vec{a} + y\vec{b}; x, y \in \mathbf{R}\}$$

の正規直交基底と呼びます。

グラム・シュミットの直交化—例

$$(\vec{p} \ \vec{q}) = (\vec{a} \ \vec{b}) \begin{pmatrix} \frac{1}{\|\vec{a}\|} & -\frac{(\vec{a}, \vec{b})}{\|\vec{b}-\vec{w}\| \cdot \|\vec{a}\|^2} \\ 0 & \frac{1}{\|\vec{b}-\vec{w}\|} \end{pmatrix}$$

から

$$x\vec{a} + y\vec{b} = \xi\vec{p} + \eta\vec{q} \quad \Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\|\vec{a}\|} & -\frac{(\vec{a}, \vec{b})}{\|\vec{b}-\vec{w}\| \cdot \|\vec{a}\|^2} \\ 0 & \frac{1}{\|\vec{b}-\vec{w}\|} \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

となりますから

$$L = \{\xi\vec{p} + \eta\vec{q}; \xi, \eta \in \mathbf{R}\}$$

であることが分かります。

グラム・シュミットの直交化一例

このとき $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ のとき

$$\|\vec{c} - \xi\vec{p} - \eta\vec{q}\|^2 = (\xi - (\vec{c}, \vec{p}))^2 + (\eta - (\vec{c}, \vec{q}))^2 + \|\vec{c}\|^2 - (\vec{c}, \vec{p})^2 - (\vec{c}, \vec{q})^2$$

から最小になるのは

$$\xi = (\vec{c}, \vec{p}) = \frac{1}{\sqrt{3}}, \quad \eta = (\vec{c}, \vec{q}) = \frac{5}{\sqrt{42}}$$

であることが分かります。

問 以上の計算から $\|\vec{c} - x\vec{a} - y\vec{b}\|^2$ を最小化する $x, y \in \mathbf{R}$ を求めましょう。

グラム・シュミットの直交化一例

$\vec{w}_0 = (\vec{c}, \vec{p})\vec{p} + (\vec{c}, \vec{q})\vec{q}$ と定めると

$$\begin{aligned}(\vec{c} - \vec{w}_0, \vec{p}) &= (\vec{c}, \vec{p}) - (\vec{c}, \vec{p})\|\vec{p}\|^2 - (\vec{c}, \vec{q})(\vec{q}, \vec{p}) \\ &= (\vec{c}, \vec{p}) - (\vec{c}, \vec{p}) = 0\end{aligned}$$

$$\begin{aligned}(\vec{c} - \vec{w}_0, \vec{q}) &= (\vec{c}, \vec{q}) - (\vec{c}, \vec{q})(\vec{p}, \vec{q}) - (\vec{c}, \vec{q})\|\vec{q}\|^2 \\ &= (\vec{c}, \vec{q}) - (\vec{c}, \vec{q}) = 0\end{aligned}$$

から

$$\vec{c} - \vec{w}_0 \perp L$$

が分かります。 \vec{w}_0 を \vec{c} の L への直交射影と呼びます。