MA TRẬN ĐẠI SỐ VÀ VECTƠ NGẪU NHIÊN

1 Một số khái niệm cơ bản về ma trận và vectơ đại số

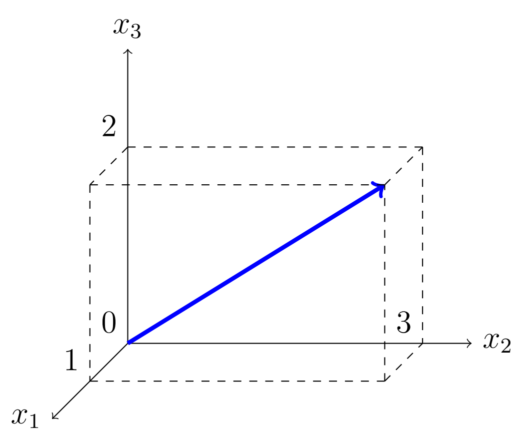
1.1 Vectơ

1.1.1 Định nghĩa

Một điểm $\mathbf{x}$ trong $\mathbb{R}^n$ là một bộ n số thực $x_1, x_2, ..., x_n$ được sắp xếp có thứ tự và được viết dưới dạng cột gọi là một *vectơ*, và được ký hiệu là

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ hoặc } \mathbf{x}^T = [x_1, x_2, ..., x_n]$$

trong đó $\mathbf{x}^T$ được gọi là *chuyển vị* của $\mathbf{x}$ khi xếp các phần tử của $\mathbf{x}$ ở dạng cột thành dòng.



Hình 1: Vectơ $\mathbf{x} = [1, 3, 2]^T$

Một vectơ $\mathbf{x}$ có thể biểu diễn hình học như một đường thẳng có hướng trong không gian n chiều với các phần tử x_1 dọc theo trục đầu tiên, x_2 dọc theo trục thứ 2,... và x_n dọc theo trục thứ n. Hình (1) là một minh hoạ với $n = 3$.

1.1.2 Phép toán trên vectơ

Cho hai vectơ $\mathbf{x}$, $\mathbf{y}$ trong không gian n chiều, ta có các tính chất sau

(1) $c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$ với c là hằng số.

(2) $\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$

1.1.3 Chiều và độ lớn

Vectơ là một đoạn thẳng có hướng biểu thị chiều và độ lớn (chiều dài của vectơ).

Độ lớn của vectơ $\mathbf{x} = [x_1, x_2, ..., x_n]^T$ trong không gian n chiều được định nghĩa bằng

$$L_x = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$$

Phép nhân vectơ $\mathbf{x}$ với hằng số c bất kỳ trong không gian n chiều là một vectơ có gốc và phương trùng với gốc và phương của vectơ $\mathbf{x}$, cùng chiều nếu $c > 0$ và ngược lại nếu $c < 0$, có độ dài bằng

$$\begin{aligned} L_{cx} &= \sqrt{c^2x_1^2 + c^2x_2^2 + ... + c^2x_n^2} \\ &= |c|\sqrt{x_1^2 + x_2^2 + ... + x_n^2} = |c|L_x \end{aligned}$$

1.1.4 Góc

Xét hai vectơ $\mathbf{x}$ và $\mathbf{y}$ trong không gian n chiều. Góc θ giữa hai vectơ được xác định bởi

$$\cos(\theta) = \frac{x_1y_1 + x_2y_2 + ... + x_ny_n}{L_xL_y}$$

1.1.5 Tích vô hướng của hai vectơ

Trong không gian n chiều, tích vô hướng của hai vectơ $\mathbf{x}$ và $\mathbf{y}$ được định nghĩa bởi

$$\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + ... + x_ny_n$$

Tích vô hướng của hai vectơ trên được ký hiệu là $\mathbf{x}^T\mathbf{y}$ hoặc $\mathbf{y}^T\mathbf{x}$.

1.1.6 Không gian vectơ

Không gian của tập hợp **n** số thực với phép nhân vô hướng và phép cộng vectơ được định nghĩa ở mục (1.1.2) được gọi là một *không gian vectơ*.

Tổ hợp tuyến tính. Vectơ $y = a_1x_1 + a_2x_2 + ... + a_kx_k$ là một *tổ hợp tuyến tính* của các vectơ $x_1, x_2, ..., x_k$. Tập hợp các tổ hợp tuyến tính của $x_1, x_2, ..., x_k$ được gọi là *bao tuyến tính* (*linear span*) hay còn gọi là không gian con sinh bởi $S = \{x_1, x_2, ..., x_k\}$, ký hiệu là $span(S)$ hay $\langle S \rangle$.

Độc lập tuyến tính. Tập hợp các vectơ $x_1, x_2, ..., x_k$ được gọi là *độc lập tuyến tính* nếu phương trình sau tồn tại nghiệm duy nhất thoả mãn

$$a_1x_1 + a_2x_2 + ... + a_kx_k = 0 \Leftrightarrow a_1 = a_2 = ... = a_k = 0$$

Ngược lại, tập hợp các vectơ trên gọi là *phụ thuộc tuyến tính*.

Vectơ trực giao. Hai vectơ $\mathbf{x} = [x_1, x_2, ..., x_k]^T$ và $\mathbf{y} = [y_1, y_2, ..., y_k]^T$ được gọi là *trực giao* (*vuông góc*) với nhau nếu

$$\mathbf{x}^T\mathbf{y} = 0$$

Ta ký hiệu là $\mathbf{x} \perp \mathbf{y}$.

Tính chất:

(1) $\mathbf{x}$ trực giao với mọi vectơ khi và chỉ khi $\mathbf{x} = 0$.

(2) Nếu $\mathbf{z}$ vuông góc với mỗi vectơ $x_1, x_2, ..., x_k$, thì $\mathbf{z}$ vuông góc với không gian con sinh với $x_1, x_2, ..., x_k$.

(3) Các vectơ vuông góc với nhau thì độc lập tuyến tính.

Hình chiếu. *Hình chiếu* của vectơ $\mathbf{x}$ lên vectơ $\mathbf{y}$ là

$$proj_{\mathbf{y}}(\mathbf{x}) = \frac{\mathbf{x}^T\mathbf{y}}{L_y^2}\mathbf{y}$$

Nếu $\mathbf{y}$ là vectơ đơn vị, tức là $L_y = 1$ thì

$$proj_{\mathbf{y}}(\mathbf{x}) = (\mathbf{x}^T\mathbf{y})\mathbf{y}$$

Nếu $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_k$ vuông góc với nhau từng đôi một, thì hình chiếu của $\mathbf{x}$ lên không gian con sinh bởi $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_k$ là

$$\frac{(\mathbf{x}^T\mathbf{y}_1)}{L_{y_1}^2}\mathbf{y}_1 + \frac{(\mathbf{x}^T\mathbf{y}_2)}{L_{y_2}^2}\mathbf{y}_2 + ... + \frac{(\mathbf{x}^T\mathbf{y}_k)}{L_{y_k}^2}\mathbf{y}_k$$

Quá trình Gram-Schmidt. Cho các vectơ độc lập tuyến tính $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$, tồn tại các vectơ trực giao $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$ trong cùng không gian sinh. *Quá trình Gram-Schmidt* được tiến hành như sau

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{x}_1 \\
\mathbf{u}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 \\
&\vdots \\
\mathbf{u}_k &= \mathbf{x}_k - \frac{\mathbf{x}_k^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 - \cdots - \frac{\mathbf{x}_k^T\mathbf{u}_{k-1}}{\mathbf{u}_{k-1}^T\mathbf{u}_{k-1}}\mathbf{u}_{k-1}
\end{aligned}$$

Dãy $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$ là dãy vectơ trực giao cần tìm. Ta cũng có thể *trực chuẩn hoá* các vectơ trực giao trên bằng cách $\mathbf{z}_j = \mathbf{u}_j / \sqrt{\mathbf{u}_j^T \mathbf{u}_j}$. Việc tính toán dãy $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k$ gọi là *trực giao hoá Gram-Schmidt* và việc tính toán dãy $\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_k$ gọi là *trực chuẩn hoá Gram-Schmidt*.

1.2 Ma trận

1.2.1 Định nghĩa

Ma trận là một mảng hình chữ nhật các số, ký hiệu, biểu thức sắp xếp theo hàng và cột. Ta biễu diễn một ma trận có n dòng và p cột như sau

$$\underset{(n\times p)}{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{np} \end{bmatrix}$$

Hai ma trận $\underset{(m\times k)}{\mathbf{A}} = \{a_{ij}\}$ và $\underset{(m\times k)}{\mathbf{B}} = \{b_{ij}\}$ được gọi là bằng nhau, ký hiệu $\mathbf{A} = \mathbf{B}$, nếu $a_{ij} = b_{ij}, i = 1, 2, ..., m, j = 1, 2, ..., k$. Tức là

(1) Chiều của chúng phải bằng nhau.

(2) Các phần tử tương ứng bằng nhau.

1.2.2 Các phép toán trên ma trận

- *Ma trận chuyển vị*: Chuyển vị của mà trận $\mathbf{A}_{(n\times p)}$ là ma trận $\mathbf{A}^T_{(p\times n)}$ tạo ra bằng cách chuyển hàng thành cột và cột thành hàng: $(\mathbf{A})^T_{ij} = \mathbf{A}_{ji}$.

- *Phép nhân một số với ma trận*: Tích $c\mathbf{A}$ của hằng số c với ma trận $\mathbf{A}$ được thực hiện bằng cách nhân mỗi phần tử của $\mathbf{A}$ với c.

$$\underset{(n\times p)}{c\mathbf{A}} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1p} \\ ca_{21} & ca_{22} & \cdots & ca_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{np} \end{bmatrix}$$

- *Phép cộng hai ma trận*: Tổng $\mathbf{A} + \mathbf{B}$ của hai ma trận $\mathbf{A}$ và $\mathbf{B}$ có cùng kích thước là một ma trận cùng kích thước, với phần tử trong vị trí tương ứng bằng tổng của hai phần tử tương ứng của mỗi ma trận.

$$\underset{(n\times p)}{\mathbf{A}+\mathbf{B}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{np} \end{bmatrix}$$

trong đó $c_{ij} = a_{ij} + b_{ij}$. *Tính chất:* Cho 3 ma trận $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ có chiều thích hợp và c,d là hai hằng số thực. Khi đó ta có

(1) $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$

(2) $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$

(3) $c(\mathbf{A}+\mathbf{B}) = c\mathbf{A}+c\mathbf{B}$

(4) $(c+d)\mathbf{A} = c\mathbf{A}+d\mathbf{A}$

(5) $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T+\mathbf{B}^T$

(6) $(cd)\mathbf{A} = c(d\mathbf{A})$

(7) $(c\mathbf{A})^T = c\mathbf{A}^T$

- *Phép nhân hai ma trận*: Phép nhân hai ma trận được xác định khi và chỉ khi số cột của ma trận bên trái bằng số hàng của ma trận bên phải. Nếu $\mathbf{A}$ là một ma trận $(m \times n)$ và $\mathbf{B}$ là một ma trận $(n \times p)$, thì ma trận tích $\mathbf{AB}$ là ma trận $(m \times p)$ với các phần tử được xác định theo tích vô hướng của hàng tương ứng trong $\mathbf{A}$ với cột tương ứng trong $\mathbf{B}$:

$$[\mathbf{AB}]_{ij} = \mathbf{A}_{i1}\mathbf{B}_{1j} + \mathbf{A}_{i2}\mathbf{B}_{2j} + ... + \mathbf{A}_{in}\mathbf{B}_{nj} = \sum_{r=1}^{n} \mathbf{A}_{ir}\mathbf{B}_{rj}.$$

Tính chất: Cho 3 ma trận $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ có chiều thích hợp, vectơ $\mathbf{x} = [x_1, x_2, ..., x_n]^T$ và c,d là hai hằng số thực. Khi đó ta có

(1) $c(\mathbf{AB}) =)c\mathbf{A}_{\mathbf{B}}$

(2) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$

(3) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$

(4) $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$

(6) $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

(7) $\displaystyle\sum_{j=1}^{n} \mathbf{A}x_j = \mathbf{A}\sum_{j=1}^{n} x_j$

(8) $\sum_{j=1}^{n}(\mathbf{A}x_j)(\mathbf{A}x_j)^T = \mathbf{A}\left(\sum_{j=1}^{n} x_j x_j^T\right)\mathbf{A}^T$

1.2.3 Định thức

Định thức của ma trận vuông $\underset{(k\times k)}{\mathbf{A}} = \{a_{ij}\}$, ký hiệu là $|\mathbf{A}|$, là một hằng số

$$|\mathbf{A}| = a_{11} \qquad \text{nếu } k = 1$$

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{1j}|\mathbf{A}_{1j}|(-1)^{1+j} \qquad \text{nếu } k > 1$$

trong đó $\mathbf{A}_{1j}$ là ma trận kích thước $(k-1)\times(k-1)$ thu được bằng cách xoá dòng đầu và cột thứ j của $\mathbf{A}$. Tương tự $|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}|\mathbf{A}_{ij}|(-1)^{i+j}$ với hàng thứ i thay cho hàng đầu tiên.

1.2.4 Hạng của ma trận

Hạng hàng (*row rank*) của một ma trận là số hàng độc lập tuyến tính tối đa tương ứng là các vectơ hàng. *Hạng cột* (*column rank*) của một ma trận là hạng của tập hợp các cột của nó ứng với số vectơ cột.

Một kết quả quan trọng trong đại số tuyến tính đó là, hạng cột và hạng hàng luôn luôn bằng nhau. Giá trị các hạng này đơn giản được đồng nhất gọi là *hạng* của ma trận.

Một ma trận vuông $\underset{(k\times k)}{\mathbf{A}}$ là *không suy biến* nếu $\underset{(k\times k)}{\mathbf{A}}\underset{(k\times 1)}{\mathbf{x}} = \underset{(k\times 1)}{0}$ suy ra $\underset{(k\times 1)}{\mathbf{x}} = \underset{(k\times 1)}{0}$. Ngược lại, ma trận trên được gọi là *suy biến*. Tương đương, một ma trận là *không suy biến* nếu hạng của nó bằng số dòng (hoặc cột) mà nó có.

1.2.5 Các loại ma trận đặc biệt

- *Ma trận vuông*: là ma trận có số hàng và số cột bằng nhau. Ma trận $(n\times n)$ còn gọi là ma trận vuông bậc n. Bất kỳ hai ma trận vuông có cùng bậc đều thực hiện được phép cộng và nhân với nhau.

- *Ma trận đối xứng*: một ma trận vuông $\mathbf{A}$ gọi là đối xứng nếu $\mathbf{A} = \mathbf{A}^T$ hay $a_{ij} = a_{ji}$ với mọi i, j.

- *Ma trận đơn vị*: ma trận đơn vị $\mathbf{I}_n$ có số chiều n là một ma trận $(n\times n)$ trong đó mọi phần tử trên đường chéo chính bằng 1 và tất cả những phần tử khác đều bằng 0.

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Nó là ma trận đơn vị bởi vì khi thực hiện nhân một ma trận với nó thì vẫn thu được ma trận đó:

$$\mathbf{A}\mathbf{I}_n = \mathbf{I}_m\mathbf{A} = \mathbf{A} \text{ với ma trận } \mathbf{A}_{m\times n} \text{ bất kỳ.}$$

- *Ma trận khả nghịch*: ma trận vuông $\mathbf{A}$ gọi là *khả nghịch* hay không suy biến nếu tồn tại một ma trận $\mathbf{B}$ sao cho

$$\underset{(n\times n)}{\mathbf{B}}\underset{(n\times n)}{\mathbf{A}} = \underset{(n\times n)}{\mathbf{A}}\underset{(n\times n)}{\mathbf{B}} = \underset{(n\times n)}{\mathbf{I}}$$

Nếu $\mathbf{B}$ tồn tại, thì nó là duy nhất và được gọi là ma trận nghịch đảo của $\mathbf{A}$, ký hiệu bằng $\mathbf{A}^{-1}$. Điều kiện để tồn tại ma trận nghịch đảo là n cột $a_1, a_2, ..., a_n$ của ma trận $\mathbf{A}$ độc lập tuyến tính, tức là

$$c_1a_1 + c_2a_2 + ... + c_na_n = 0 \Leftrightarrow c_1 = c_2 = ... = c_n = 0$$

Kết quả 1. Cho ma trận $\mathbf{A}$ có số chiều $k \times k$, khi đó các điều sau là tương đương:

(1) $\underset{(k\times k)}{\mathbf{A}}\underset{(k\times 1)}{\mathbf{x}} = \underset{(k\times 1)}{0}$ suy ra $\underset{(k\times 1)}{\mathbf{x}} = \underset{(k\times 1)}{0}$ ($\mathbf{A}$ là không suy biến).

(2) $|\mathbf{A}| \neq 0$.

(3) Tồn tại một ma trận $\mathbf{A}^{-1}$ sao cho $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \underset{(k\times k}{I}$.

Kết quả 2. Cho $\mathbf{A}$ và $\mathbf{B}$ là hai ma trận vuông có cùng số chiều và tồn tại các nghịch đảo tương ứng. Ta có các kết quả sau:

(1) $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$.

(2) $(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Kết quả 3. Cho $\mathbf{A}$ và $\mathbf{B}$ là hai ma trận kích thước $k \times k$.

(1) $|\mathbf{A}| = |\mathbf{A}^T|$.

(2) Nếu $\mathbf{A}$ có một hàng hoặc một cột là vectơ $\mathbf{0}$ thì $|\mathbf{A}| = 0$.

(3) Nếu $\mathbf{A}$ có bất kỳ hai hàng (hoặc cột) bằng nhau thì $|\mathbf{A}| = 0$.

(4) Nếu $\mathbf{A}$ không suy biến thì $|\mathbf{A}| = 1/|\mathbf{A}^{-1}|$ hay $|\mathbf{A}||\mathbf{A}^{-1}| = 1$.

(5) $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.

(6) $|c\mathbf{A}| = c^k|\mathbf{A}|$ trong đó c là hằng số.

- *Ma trận trực giao*: là ma trận vuông với các phần tử thực sao cho các cột và hàng là những vectơ đơn vị trực giao (nghĩa là vectơ trực chuẩn). Hay nói tương đương, ma trận $\mathbf{A}$ trực giao nếu và chỉ nếu ma trận chuyển vị của nó bằng ma trận nghịch đảo của nó:

$$\mathbf{A}^T = \mathbf{A}^{-1} \text{ hay } \mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}$$

1.2.6 Vết của ma trận

Cho $\mathbf{A} = \{a_{ij}\}$ là ma trận vuông kích thước $k \times k$. Vết của ma trận $\mathbf{A}$, ký hiệu là $tr(\mathbf{A})$, là tổng các phần tử trên đường chéo chính của $\mathbf{A}$, tức là $tr(\mathbf{A}) = \sum_{i=1}^{k} a_{ii}$.

Kết quả. Cho $\mathbf{A}$ và $\mathbf{B}$ là hai ma trận kích thước $k \times k$ và c là một hằng số.

(1) $tr(c\mathbf{A}) = ctr(\mathbf{A})$.

(2) $tr(\mathbf{A} \pm \mathbf{B}) = tr(\mathbf{A}) \pm tr(\mathbf{B})$.

(3) $tr(\mathbf{AB}) = tr(\mathbf{BA})$.

(4) $tr(\mathbf{B}^{-1}\mathbf{AB}) = tr(\mathbf{A})$.

(5) $tr(\mathbf{A}\mathbf{A}^T) = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}^2$.

1.2.7 Vectơ riêng và trị riêng

Một số λ và một vectơ $x \neq 0$ thoả mãn

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

được gọi lần lượt là giá trị riêng và vectơ riêng của $\mathbf{A}$. Số λ là một trị riêng của một ma trận $\mathbf{A}_{n\times n}$ nếu và chỉ nếu $\mathbf{A} - \lambda\mathbf{I}_n$ là không khả nghịch, mà tương đương với

$$\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0.$$

1.2.8 Dạng toàn phương

Dạng toàn phương $\mathcal{Q}(\mathbf{x})$ trên k biến $x_1, x_2, ..., x_k$ là $\mathcal{Q}(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$, trong đó $\mathbf{x} = [x_1, x_2, ..., x_k]^T$ và $\mathbf{A}$ là ma trận đối xứng kích thước $k \times k$. Lưu ý rằng dạng toàn phương có thể viết dưới dạng $\mathcal{Q}(\mathbf{x}) = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}x_ix_j$.

1.2.9 Phép phân tích suy biến (singular-value decomposition)

Cho $\mathbf{A}$ là ma trận các số thực kích thước $m \times k$. Tồn tại ma trận trực giao $\mathbf{U}$ kích thước $m \times m$ và ma trận trực giao V kích thước $k \times k$ sao cho

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$$

trong đó $\mathbf{\Lambda}$ là ma trận đường chéo kích thước $m \times k$ với các phần tử trên đường chéo chính là $\lambda_i \leq 0, i = 1, 2, ..., \min(m, k)$. Các hằng số dương λ_i được gọi là các giá trị suy biến của $\mathbf{A}$.

2 Ma trận xác định dương

2.1 Định lý phổ và phép phân tích phổ

Định lý phổ: Nếu $\mathbf{A} \in \mathbb{R}^{n\times n}$ là một ma trận đối xứng, thì $\mathbf{A}$ có n trị riêng thoả mãn $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$ và n vectơ riêng đơn vị tương ứng $e_1, e_2, ..., e_n$ đôi một trực giao, tức là

$$e_i^T e_j = 0 \text{ với mọi } i \neq j \text{ và } \|e_i\| = 1 \text{ với mọi } i.$$

Phép phân tích phổ (spectral decomposition) của một ma trận đối xứng $\mathbf{A}$ kích thước $n \times n$ được cho bởi

$$\underset{(n\times n)}{\mathbf{A}} = \lambda_1 \underset{(n\times 1)}{e_1} \underset{(1\times n)}{e_1^T} + \lambda_2 \underset{(n\times 1)}{e_2} \underset{(1\times n)}{e_2^T} + ... + \lambda_n \underset{(n\times 1)}{e_n} \underset{(1\times n)}{e_n^T}$$

trong đó $\lambda_1, \lambda_2, ..., \lambda_n$ là các trị riêng của $\mathbf{A}$ và $e_1, e_2, ..., e_n$ là các vectơ riêng đã được chuẩn hoá tương ứng.

2.2 Ma trận xác định dương

Ma trận đối xứng $\mathbf{A} \in \mathbb{R}^{n\times n}$ được gọi là *xác định dương* nếu

$$x^T\mathbf{A}x > 0, \forall x \in \mathbb{R}^n, x \neq 0.$$

Ma trận đối xứng $\mathbf{A} \in \mathbb{R}^{n\times n}$ được gọi là *bán xác định dương* nếu

$$x^T\mathbf{A}x \geq 0, \forall x \in \mathbb{R}^n, x \neq 0.$$

Ma trận *xác định âm* và *bán xác định âm* cũng được định nghĩa tương tự.

Ký hiệu $\mathbf{A} \succ 0, \succeq 0, \prec 0, \preceq 0$ lần lượt để chỉ một ma trận là xác định dương, bán xác định dương, xác định âm và bán xác định âm. Ký hiệu $\mathbf{A} \succ \mathbf{B}$ cũng được dùng để chỉ ra rằng $\mathbf{A} - \mathbf{B} \succ 0$.

Bằng phép phân tích phổ, ta dễ dàng thấy được ma trận đối xứng $\mathbf{A} \in \mathbb{R}^{n\times n}$ là

- *xác định dương* $\Leftrightarrow \lambda_i > 0, \forall i = 1, ..., n,$
- *bán xác định dương* $\Leftrightarrow \lambda_i \geq 0, \forall i = 1, ..., n.$

3 Ma trận căn bậc hai

3.1 Định nghĩa

Xét ma trận xác định dương $\mathbf{A}_{n\times n}$ với phép phân tích phổ $\mathbf{A} = \sum_{i=1}^{n} \lambda_i e_i e_i^T$ và ma trận $\mathbf{P} = [e_1, e_1, ..., e_n]$. Khi đó

$$\underset{(n\times n)}{\mathbf{A}} = \sum_{i=1}^{n} \lambda_i \underset{(n\times 1)}{e_i} \underset{(1\times n)}{e_i^T} = \underset{(n\times n)}{\mathbf{P}} \underset{(n\times n)}{\Lambda} \underset{(n\times n)}{\mathbf{P}^T}$$

trong đó $\mathbf{P}\mathbf{P}^T = \mathbf{P}^T\mathbf{P} = \mathbf{I}$ và Λ là ma trận đường chéo

$$\underset{(n\times n)}{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \quad \text{với } \lambda_i > 0$$

Do đó

$$\mathbf{A}^{-1} = \mathbf{P}\Lambda^{-1}\mathbf{P}^T = \sum_{i=1}^{n} \frac{1}{\lambda_i} e_i e_i^T$$

vì $(\mathbf{P}\Lambda^{-1}\mathbf{P}^T)\mathbf{P}\Lambda\mathbf{P} = \mathbf{P}\Lambda\mathbf{P}^T(\mathbf{P}\Lambda^{-1}\mathbf{P}^T) = \mathbf{P}\mathbf{P}^T = \mathbf{I}$.

Tiếp theo, ta định nghĩa ma trận căn bậc hai của một ma trận xác định dương $\mathbf{A}$

$$\mathbf{A}^{1/2} = \sum_{i=1}^{n} \sqrt{\lambda_i} e_i e_i^T = \mathbf{P}\Lambda^{1/2}\mathbf{P}^T$$

trong đó $\Lambda^{1/2}$ là ma trận đường chéo với phần tử thứ i trên đường chéo chính là $\sqrt{\lambda_i}$.

3.2 Tính chất

(1) $(\mathbf{A}^{1/2})^T = \mathbf{A}^{1/2}$ ($\mathbf{A}^{1/2}$ là ma trận đối xứng).

(2) $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$.

(3) $(\mathbf{A}^{1/2})^{-1} = \sum_{i=1}^{n} \frac{1}{\sqrt{\lambda_i}} e_i e_i^T = \mathbf{P}\Lambda^{-1/2}\mathbf{P}^T$, trong đó $\Lambda^{-1/2}$ là ma trận đường chéo với $1/\sqrt{\lambda_i}$ là phần tử thứ i trên đường chéo chính.

(4) $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1/2}\mathbf{A}^{1/2} = \mathbf{I}$ và $\mathbf{A}^{-1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1}$, trong đó $\mathbf{A}^{-1/2} = (\mathbf{A}^{1/2})^{-1}$.

4 Ma trận và vectơ ngẫu nhiên

4.1 Khái niệm

Vectơ ngẫu nhiên là một vectơ mà thành phần của nó là các biến ngẫu nhiên. Tương tự, *Ma trận ngẫu nhiên* là ma trận mà các phần tử của nó là các biến ngẫu nhiên.

Kỳ vọng của một ma trận ngẫu nhiên (hay vectơ) là một ma trận (hay vectơ) chứa giá trị kỳ vọng của mỗi phần tử tương ứng của nó. Cụ thể, xét $\mathbf{X} = \{X_{ij}\}$ là một ma trận ngẫu nhiên $n \times p$. Kỳ vọng của $\mathbf{X}$, ký hiệu là $E(\mathbf{X})$, là một ma trận $n \times p$ các số (nếu tồn tại)

$$E(\mathbf{X}) = \begin{bmatrix} E(X_{11}) & E(X_{12}) & \cdots & E(X_{1p}) \\ E(X_{21}) & E(X_{22}) & \cdots & E(X_{2p}) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_{n1}) & E(X_{n2}) & \cdots & E(X_{np}) \end{bmatrix}$$

trong đó mỗi phần tử của ma trận là

$$E(X_{ij}) = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x_{ij} f_{ij}(x_{ij}) dx_{ij} & \text{nếu } X_{ij} \text{ là biến ngẫu nhiên liên tục với} \\ & \text{hàm mật độ xác suất } f_{ij}(x_{ij}). \\ \displaystyle\sum_{\text{all } x_{ij}} x_{ij} p_{ij}(x_{ij}) & \text{nếu } X_{ij} \text{ là một biến ngẫu nhiên rời rạc} \\ & \text{với hàm xác suất } p_{ij}(x_{ij}). \end{cases}$$

4.2 Tính chất

Xét hai ma trận ngẫu nhiên cùng kích thước $\mathbf{X}$, $\mathbf{Y}$ và $\mathbf{A}$, $\mathbf{B}$ là hai ma trận hằng với số chiều thích hợp. Khi đó ta có các tính chất sau:

(1) $E(\mathbf{X} + \mathbf{Y}) = E(\mathbf{X}) + E(\mathbf{Y})$

(2) $E(c\mathbf{X}) = cE(\mathbf{X})$, với c là hằng số

(3) $E(\mathbf{AXB}) = \mathbf{A}E(\mathbf{X})\mathbf{B}$

5 Vectơ trung bình và ma trận hiệp phương sai

5.1 Trung bình và phương sai biên

Giả sử $\mathbf{X} = [X_1, X_2, ..., X_p]^T$ là một vectơ ngẫu nhiên kích thước $p \times 1$ với mỗi phần tử là một biến ngẫu nhiên có phân phối xác suất biên của riêng từng biến. Khi đó, trung bình biên μ_i và phương sai σ_i^2 được định nghĩa như sau

$$\mu_i = E(X_i) \text{ và } \sigma_i^2 = E(X_i - \mu_i)^2 \text{ với } i = 1, 2, ..., p$$

Cụ thể

$$\mu_i = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x_i f_i(x_i) dx_i & \text{nếu } X_i \text{ là biến ngẫu nhiên liên tục với hàm mật độ xác suất } f_i(x_i) \\ \displaystyle\sum_{\text{all } x_i} x_i p_i(x_i) & \text{nếu } X_i \text{ là một biến ngẫu nhiên rời rạc với hàm xác suất } p_i(x_i) \end{cases}$$

$$\sigma_i^2 = \begin{cases} \displaystyle\int_{-\infty}^{\infty} (x_i - \mu_i)^2 f_i(x_i) dx_i & \text{nếu } X_i \text{ là biến ngẫu nhiên liên tục với hàm mật độ xác suất } f_i(x_i) \\ \displaystyle\sum_{\text{all } x_i} (x_i - \mu_i)^2 p_i(x_i) & \text{nếu } X_i \text{ là một biến ngẫu nhiên rời rạc với hàm xác suất } p_i(x_i) \end{cases}$$

5.2 Biến độc lập

Cho p biến ngẫu nhiên $X_1, X_2, ..., X_p$. Ta nói các biến $X_1, X_2, ..., X_p$ độc lập với nhau nếu

$$F_{X_1X_2...X_p}(x_1, x_2, ..., x_p) = F_{X_1}(x_1)F_{X_2}(x_2)...F_{X_p}(x_p), \quad \forall(x_1, x_2, ..., x_p)$$

Với trường hợp các biến ngẫu nhiên rời rạc

$$p_{X_1X_2...X_p}(x_1, x_2, ..., x_p) = p_{X_1}(x_1)p_{X_2}(x_2)...p_{X_p}(x_p), \quad \forall(x_1, x_2, ..., x_p)$$

Với trường hợp các biến ngẫu nhiên liên tục

$$f_{X_1X_2...X_p}(x_1, x_2, ..., x_p) = f_{X_1}(x_1)f_{X_2}(x_2)...f_{X_p}(x_p), \quad \forall(x_1, x_2, ..., x_p)$$

Định lý: Nếu 2 biến ngẫu nhiên $\mathbf{X}$ và $\mathbf{Y}$ độc lập với nhau thì covariance của chúng bằng 0, tức là $Cov(\mathbf{X}, \mathbf{Y}) = 0$.
Chiều ngược lại không đúng.

5.3 Trung bình và hiệp phương sai của vectơ ngẫu nhiên

Cho vectơ ngẫu nhiên $\mathbf{X} = [X_1, X_2, ..., X_n]^T$ có kích thước $n \times 1$. Khi đó ta có vectơ trung bình $\mu = E(\mathbf{X})$ và ma trận hiệp phương sai $\Sigma = E(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T$. Cụ thể

$$E(\mathbf{X}) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} = \mu$$

và

$$\begin{aligned} \mathbf{\Sigma} &= E(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T \\ &= \begin{bmatrix} E(X_1-\mu_1)^2 & E(X_1-\mu_1)(X_2-\mu_2) & \cdots & E(X_1-\mu_1)(X_n-\mu_n) \\ E(X_2-\mu_2)(X_1-\mu_1) & E(X_2-\mu_2)^2 & \cdots & E(X_2-\mu_2)(X_n-\mu_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_n-\mu_n)(X_1-\mu_1)^2 & E(X_n-\mu_n)(X_2-\mu_2) & \cdots & E(X_n-\mu_n)^2 \end{bmatrix} \end{aligned}$$

hay

$$\mathbf{\Sigma} = Cov(\mathbf{X}) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

trong đó $\sigma_{ij} = E(X_i - \mu_i)(X_j - \mu_j)$.

Vì $\sigma_{ik} = E(X_i - \mu_i)(X_k - \mu_k) = \sigma_{ki}$ nên ma trận Σ ở trên có thể viết lại dưới dạng

$$\mathbf{\Sigma} = Cov(\mathbf{X}) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_{nn} \end{bmatrix}$$

Hệ số tương quan ρ_{ik} được xác định thông qua hiệp phương sai σ_{ik} và các phương sai σ_{ii} và σ_{kk} như sau

$$\rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{kk}}}$$

Tổng quát, ta có ma trận tương quan là một ma trận đối xứng kích thước $n \times n$

$$\boldsymbol{\rho} = \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \cdots & \frac{\sigma_{1n}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{nn}}} \\ \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{22}}} & \cdots & \frac{\sigma_{2n}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{nn}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1n}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{nn}}} & \frac{\sigma_{2n}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{nn}}} & \cdots & \frac{\sigma_{nn}}{\sqrt{\sigma_{nn}}\sqrt{\sigma_{nn}}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n} & \rho_{2n} & \cdots & 1 \end{bmatrix}$$

và ma trận độ lệch chuẩn kích thước $n \times n$ là

$$\mathbf{v}^{1/2} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{pp}} \end{bmatrix}$$

trong đó $\sigma_{ii} = E(X_i - \mu_i)^2$.

Từ đây, ta có các kết quả sau

$$\mathbf{v}^{1/2}\boldsymbol{\rho}\mathbf{v}^{1/2} = \boldsymbol{\Sigma}$$

và

$$\boldsymbol{\rho} = (\mathbf{v}^{1/2})^{-1}\boldsymbol{\Sigma}(\mathbf{v}^{1/2})^{-1}$$

5.4 Phân vùng ma trận hiệp phương sai

Cho vectơ ngẫu nhiên $\mathbf{X} = [X_1, X_2, ..., X_n]^T$ kích thước $n \times 1$. Ta có thể phân vùng n đặc trưng trong $\mathbf{X}$ thành nhiều nhóm. Ví dụ, ta chia $\mathbf{X}$ thành hai nhóm có kích thước q và $n - q$ như sau

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_q \\ \hdashline X_{q+1} \\ \vdots \\ X_n \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} X_1 \\ \vdots \\ X_q \end{matrix}}\right\} q \\ \left.\vphantom{\begin{matrix} X_1 \\ \vdots \\ X_q \end{matrix}}\right\} n-q \end{matrix} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \hdashline \mathbf{X}^{(2)} \end{bmatrix} \quad \text{và} \quad \boldsymbol{\mu} = E(\mathbf{X}) = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_q \\ \hdashline \mu_{q+1} \\ \vdots \\ \mu_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}^{(1)} \\ \hdashline \boldsymbol{\mu}^{(2)} \end{bmatrix}$$

Dựa trên kỳ vọng của ma trận $(\mathbf{X}^{(1)} - \boldsymbol{\mu}^{(1)})(\mathbf{X}^{(2)} - \boldsymbol{\mu}^{(2)})^T$, ta có

$$E(\mathbf{X}^{(1)} - \boldsymbol{\mu}^{(1)})(\mathbf{X}^{(2)} - \boldsymbol{\mu}^{(2)})^T = \begin{bmatrix} \sigma_{1,q+1} & \sigma_{1,q+2} & \cdots & \sigma_{1,n} \\ \sigma_{2,q+1} & \sigma_{2,q+2} & \cdots & \sigma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{q,q+1} & \sigma_{q,q+2} & \cdots & \sigma_{q,n} \end{bmatrix} = \boldsymbol{\Sigma}_{12}$$

trong đó $\sigma_{i,j}$ với $i = 1, 2, ..., q; j = q+1, q+2, ..., n$ là hiệp phương sai giữa từng thành phần của $\mathbf{X}^{(1)}$ và $\mathbf{X}^{(2)}$. Lưu ý rằng ma trận $\boldsymbol{\Sigma}_{12}$ có thể không đối xứng hay thậm chí không vuông.

Tổng quát ta có

$$\underset{(n\times n)}{\boldsymbol{\Sigma}} = E(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T = \begin{array}{c} \\ q \\ n-q \end{array} \begin{array}{c} \begin{array}{cc} q & n-q \end{array} \\ \left[\begin{array}{c:c} \boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}} \\ \hdashline \boldsymbol{\Sigma_{21}} & \boldsymbol{\Sigma_{22}} \end{array}\right] \end{array}$$

$$= \left[\begin{array}{ccc:ccc} \sigma_{11} & \cdots & \sigma_{1q} & \sigma_{1,q+1} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} & \sigma_{q,q+1} & \cdots & \sigma_{qn} \\ \hdashline \sigma_{q+1,1} & \cdots & \sigma_{q+1,q} & \sigma_{q+1,q+1} & \cdots & \sigma_{q+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nq} & \sigma_{n,q+1} & \cdots & \sigma_{nn} \end{array}\right]$$

trong đó, ma trận hiệp phương sai của $\mathbf{X}^{(1)}$ là $\boldsymbol{\Sigma}_{11}$, của $\mathbf{X}^{(2)}$ là $\boldsymbol{\Sigma}_{22}$ và của các phần tử từ $\mathbf{X}^{(1)}$ và $\mathbf{X}^{(2)}$ là $\boldsymbol{\Sigma}_{12}$ (hay $\boldsymbol{\Sigma}_{21}$). Lưu ý rằng $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}^T$.

5.5 Vectơ trung bình và ma trận hiệp phương sai của các tổ hợp tuyến tính của các biến ngẫu nhiên

Cho vectơ ngẫu nhiên $\mathbf{X} = [X_1, X_2, ..., X_n]^T$ và vectơ hằng $\mathbf{c} = [c_1, c_2, ..., c_n]^T$. Xét tổ hợp tuyến tính $\mathbf{c}^T\mathbf{X} = c_1X_1 + c_2X_2 + ... + c_nX_n$ ta có

$$\begin{aligned} \text{mean } &= E(\mathbf{c}^T\mathbf{X}) = \mathbf{c}^T\boldsymbol{\mu} \\ \text{variance } &= Var(\mathbf{c}^T\mathbf{X}) = \mathbf{c}^T\boldsymbol{\Sigma}\mathbf{c} \end{aligned}$$

trong đó $\boldsymbol{\mu} = E(\mathbf{X})$ và $\boldsymbol{\Sigma} = Cov(\mathbf{X})$.

Tổng quát, xét q tổ hợp tuyến tính của n biến ngẫu nhiên $X_1, X_2, ..., X_n$:

$$\begin{array}{lcccccccc} Z_1 & = & c_{11}X_1 & + & c_{12}X_2 & + & \cdots & + & c_{1p}X_p \\ Z_2 & = & c_{21}X_1 & + & c_{22}X_2 & + & \cdots & + & c_{2p}X_p \\ \vdots & & & & & & \vdots & & \\ Z_q & = & c_{q1}X_1 & + & c_{q2}X_2 & + & \cdots & + & c_{qp}X_p \end{array}$$

hay

$$\mathbf{Z} = \underset{(q\times 1)}{\begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_q \end{bmatrix}} = \underset{(q\times n)}{\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & \cdots & c_{qn} \end{bmatrix}} \underset{(n\times 1)}{\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}} = \mathbf{CX}$$

Tổ hợp tuyến tính $\mathbf{Z} = \mathbf{CX}$ có

$$\boldsymbol{\mu}_{\mathbf{Z}} = E(\mathbf{Z}) = E(\mathbf{CX}) = \mathbf{C}\boldsymbol{\mu}_{\mathbf{X}}$$
$$\boldsymbol{\Sigma}_{\mathbf{Z}} = Cov(\mathbf{Z}) = Cov(\mathbf{CX}) = \mathbf{C}\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{C}^T$$

trong đó $\boldsymbol{\mu}_{\mathbf{X}}$ và $\boldsymbol{\Sigma}_{\mathbf{X}}$ là vectơ trung bình và ma trận hiệp phương sai của $\mathbf{X}$.

5.6 Phân hoạch vectơ trung bình mẫu và ma trận hiệp phương sai mẫu

Cho $\overline{\mathbf{x}} = [\overline{x}_1, \overline{x}_2, ..., \overline{x}_p]^T$ là vectơ của trung bình mẫu lấy từ n quan sát trên p biến $X_1, X_2, ..., X_p$ và

$$\mathbf{S}_n = \begin{bmatrix} s_{11} & \cdots & s_{1p} \\ \vdots & \ddots & \vdots \\ s_{1p} & \cdots & s_{pp} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{n}\sum_{j=1}^{n}(x_{j1} - \overline{x}_1)^2 & \cdots & \frac{1}{n}\sum_{j=1}^{n}(x_{j1} - \overline{x}_1)(x_{jp} - \overline{x}_p) \\ \vdots & \ddots & \vdots \\ \frac{1}{n}\sum_{j=1}^{n}(x_{j1} - \overline{x}_1)(x_{jp} - \overline{x}_p) & \cdots & \frac{1}{n}\sum_{j=1}^{n}(x_{jp} - \overline{x}_p)^2 \end{bmatrix}$$

là ma trận hiệp phương sai mẫu tương ứng.

Vectơ trung bình mẫu và ma trận hiệp phương sai mẫu được phân hoạch để phân biệt các đại lượng tương ứng với các nhóm biến. Do đó

$$\underset{(p\times 1)}{\overline{x}} = \begin{bmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_q \\ \hdashline \overline{x}_{q+1} \\ \vdots \\ \overline{x}_p \end{bmatrix} = \begin{bmatrix} \overline{x}^{(1)} \\ \hdashline \overline{x}^{(2)} \end{bmatrix}$$

và

$$\underset{(p\times p)}{\mathbf{S}_n} = \left[\begin{array}{ccc|ccc} s_{11} & \cdots & s_{1q} & s_{1,q+1} & \cdots & s_{1p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ s_{q1} & \cdots & s_{qq} & s_{q,q+1} & \cdots & s_{qp} \\ \hline s_{q+1,1} & \cdots & s_{q+1,q} & s_{q+1,q+1} & \cdots & s_{q+1,p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ s_{p1} & \cdots & s_{pq} & s_{p,q+1} & \cdots & s_{pp} \end{array}\right]$$

$$= \begin{array}{c} \\ q \\ p-q \end{array} \begin{array}{c} \begin{array}{cc} q & p-q \end{array} \\ \left[\begin{array}{c|c} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \hline \mathbf{S}_{21} & \mathbf{S}_{22} \end{array}\right] \end{array}$$

trong đó $\overline{\mathbf{x}}^{(1)}$ và $\overline{\mathbf{x}}^{(2)}$ là các vectơ trung bình mẫu được tạo từ các quan sát $\mathbf{x}^{(1)} = [x_1, ..., x_q]^T$ và $\mathbf{x}^{(2)} = [x_{q+1}, ..., x_p]^T$ tương ứng; $\mathbf{S}_{11}$ là ma trận hiệp phương sai mẫu được tạo từ các quan sát $\mathbf{x}^{(1)}$; $\mathbf{S}_{22}$ là ma trận hiệp phương sai mẫu được tạo từ các quan sát $\mathbf{x}^{(2)}$; và $\mathbf{S}_{12} = \mathbf{S}_{21}^T$ là ma trận hiệp phương sai mẫu tương ứng của các phần tử của $\mathbf{x}^{(1)}$ và $\mathbf{x}^{(2)}$.

6 Bất đẳng thức ma trận và cực đại hoá

Bất đẳng thức Cauchy-Schwarz. Cho $\mathbf{b}$ và $\mathbf{d}$ là hai vectơ kích thước $p \times 1$ bất kỳ. Khi đó

$$(\mathbf{b}^T\mathbf{d})^2 \leq (\mathbf{b}^T\mathbf{b})(\mathbf{d}^T\mathbf{d})$$

Dấu "=" xảy ra khi và chỉ khi $\mathbf{b} = c\mathbf{d}$ (hay $\mathbf{d} = c\mathbf{b}$) với c là hằng số.

Bất đẳng thức Cauchy-Schwarz mở rộng. Cho $\underset{(p\times 1)}{\mathbf{b}}$ và $\underset{(p\times 1)}{\mathbf{d}}$ là hai vectơ bất kỳ và $\underset{(p\times p)}{\mathbf{B}}$ là ma trận xác định dương. Khi đó

$$(\mathbf{b}^T\mathbf{d})^2 \leq (\mathbf{b}^T\mathbf{B}\mathbf{b})(\mathbf{d}^T\mathbf{B}^{-1}\mathbf{d})$$

Dấu "=" xảy ra khi và chỉ khi $\mathbf{b} = c\mathbf{B}^{-1}\mathbf{d}$ (hay $\mathbf{d} = c\mathbf{B}\mathbf{b}$) với c là hằng số.

Bổ đề cực đại hoá. Cho $\underset{(p\times p)}{\mathbf{B}}$ là ma trận xác định dương và $\underset{(p\times 1)}{\mathbf{d}}$ là một vectơ bất kỳ. Khi đó, với một vectơ khác không tuỳ ý $\underset{(p\times 1)}{x}$ thì

$$\max_{x\neq 0}\frac{(\mathbf{x}^T\mathbf{d})^2)}{\mathbf{x}^T\mathbf{B}\mathbf{x}} = \mathbf{d}^T\mathbf{B}^{-1}\mathbf{d}$$

với giá trị cực đại đạt được khi $\underset{(p\times 1)}{\mathbf{x}} = c\underset{(p\times p)}{\mathbf{B}^{-1}}\underset{(p\times 1)}{d}$ với mọi hằng số $c \neq 0$.

Cực đại hoá dạng toàn phương cho các điểm trên mặt cầu đơn vị. Cho $\underset{(p\times p)}{\mathbf{B}}$ là một ma trận xác định dương với các trị riêng $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$ và các vectơ riêng chuẩn hoá tương ứng $e_1, e_2, ..., e_p$. Khi đó

$$\max_{x\neq 0} \frac{\mathbf{x}^T\mathbf{B}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_1 \quad (\text{đạt được khi } \mathbf{x} = e_1)$$
$$\min_{(x\neq 0)} \frac{\mathbf{x}^T\mathbf{B}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_p \quad (\text{đạt được khi } \mathbf{x} = e_p)$$

Hơn nữa,

$$\max_{x\perp e_1,...,e_k} \frac{\mathbf{x}^T\mathbf{B}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_{k+1} \quad (\text{đạt được khi } \mathbf{x} = e_{k+1}, k = 1, 2, ..., p-1)$$