

# 1. Lưu ý về kí hiệu

- Số vô hướng - chữ cái in nghiêng:  $x_1, N, y, k$ .
- Vector - chữ thường in đậm:  $\mathbf{y}, \mathbf{x}_1$ .
  - Vector hàng**:  $\mathbf{x} = [x_1, x_2, \dots, x_n]$ .
  - Vector cột**:  $\mathbf{x} = [x_1; x_2; \dots; x_n]$ .
  - LƯU Ý**: Nếu không nói gì thêm, các vector dc mặc định là **vector cột**.
- Ma trận - chữ hoa in đậm:  $\mathbf{X}, \mathbf{Y}, \mathbf{W}$ .
  - Ma trận được tạo thành từ các **vector cột** theo thứ tự **trái qua phải**:  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ .
  - Ma trận được tạo thành từ các **vector hàng** theo thứ tự **trên xuống dưới**:  
 $\mathbf{X} = [\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_n]$ .
  - Phần tử nằm ở dòng  $i$  cột  $j$  của ma trận kí hiệu là  $x_{ij}$ .
  - LƯU Ý**: Cho ma trận  $\mathbf{W}$ , nếu không nói gì thêm thì  $\mathbf{w}_i$  được hiểu là **vector cột** thứ  $i$  của ma trận  $\mathbf{W}$ .

# 2. Chuyển vị và Hermitian

- Tên tiếng anh là **transpose**, kí hiệu là  $T$ .
- Cho  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , ta nói  $\mathbf{B} \in \mathbb{R}^{n \times m}$  là chuyển vị của  $\mathbf{A}$  nếu:

$$b_{ij} = a_{ji}, \forall 1 \leq i \leq n, 1 \leq j \leq m$$

- Ví dụ:
  - Chuyển vị của vector  $\mathbf{x}$  kí hiệu là  $\mathbf{x}^T$ .

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \Rightarrow \mathbf{x}^T = [x_1 \ x_2 \ \dots \ x_m]$$

- Chuyển vị của ma trận  $\mathbf{A}$  kí hiệu là  $\mathbf{A}^T$ .

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \ddots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \Rightarrow \mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \ddots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}$$

- Ta có  $\mathbf{A} \in \mathbb{R}^{m \times n}$  thì  $\mathbf{A}^T \in \mathbb{R}^{n \times m}$ . Nếu  $\mathbf{A} = \mathbf{A}^T$  thì  $\mathbf{A}$  là **ma trận đối xứng** (*symmetric matrix*).
- Trong trường hợp vector hay ma trận có phần tử là **số phức**, thì việc lấy chuyển vị sẽ kèm theo **lấy liên hợp phức** của phần tử đó. Quá trình này được gọi ngắn gọn là **chuyển vị liên hợp** (*conjugate transpose*), kí hiệu là  $H$ .
- Chuyển vị liên hợp của ma trận  $\mathbf{A}$  là  $\mathbf{A}^H$  (còn được gọi là  $\mathbf{A}$  Hermitian).
- Cho  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , ta nói  $\mathbf{B} \in \mathbb{C}^{n \times m}$  là chuyển vị liên hợp của  $\mathbf{A}$  nếu:

$$b_{ij} = \overline{a_{ji}}, \forall 1 \leq i \leq n, 1 \leq j \leq m$$

với  $\overline{a}$  là liên hợp phức của  $a$ .

- Ví dụ:
  - Chuyển vị liên hợp của vector  $\mathbf{x}$  kí hiệu là  $\mathbf{x}^H$ .

$$\mathbf{x} = \begin{bmatrix} 2 + 3i \\ 2i \end{bmatrix} \Rightarrow \mathbf{x}^H = [2 - 3i \quad -2i]$$

- Chuyển vị liên hợp của ma trận  $\mathbf{A}$  kí hiệu là  $\mathbf{A}^H$ .

$$\mathbf{A} = \begin{bmatrix} 1 + 2i & 3 - 4i \\ i & 2 \end{bmatrix} \Rightarrow \mathbf{A}^H = \begin{bmatrix} 1 - 2i & -i \\ 3 + 4i & 2 \end{bmatrix}$$

- Nếu  $\mathbf{A}$ ,  $\mathbf{x}$  lần lượt là các ma trận thực và vector thực thì:

$$\mathbf{A}^H = \mathbf{A}^T, \mathbf{x}^H = \mathbf{x}^T$$

- Nếu chuyển vị liên hợp của một ma trận phức bằng với chính nó,  $\mathbf{A}^H = \mathbf{A}$  thì nói gọi ma trận đó là **Hermittian**.

### 3. Phép nhân hai ma trận

- Cho hai ma trận  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times p}$ , thì tích của hai ma trận này kí hiệu là  $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m \times p}$ , khi đó phần tử tại dòng  $i$  cột  $j$  của ma trận kết quả được tính bởi:

$$c_{ij} = \sum_{k=1}^n a_{ik} \times b_{kj}, \forall 1 \leq i \leq m, 1 \leq j \leq p$$

- Tính chất:
  - Không có tính chất giao hoán.

$$\mathbf{AB} \neq \mathbf{BA}$$

- Có tính chất kết hợp:

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

- Có tính phân phối đối với phép cộng:

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

- Chuyển vị của một tích bằng tích các chuyển vị theo thứ tự **ngược lại**:

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

- Bằng cách coi vector là một trường hợp đặc biệt của ma trận, ta có tích vô hướng của hai vector (*inner product*)  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  được định nghĩa là:

$$\mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x} = \sum_{i=1}^n x_i \times y_i$$

- **LƯU Ý:**

$$\mathbf{x}^H \mathbf{y} = \mathbf{y}^H \mathbf{x} = (\mathbf{y}^H \mathbf{x})^H$$

khi và chỉ khi chúng là các số thực.

$$\mathbf{x}^H \mathbf{x} \geq 0, \forall \mathbf{x} \in \mathbb{C}^n$$

- Nếu tích vô hướng của **hai vector khác không** bằng không, thì hai vector đó **vuông góc với nhau**.

- Phép nhân của một ma trận  $\mathbf{A} \in \mathbb{R}^{m \times n}$  với một vector  $\mathbf{x} \in \mathbb{R}^n$  là một vector  $\mathbf{b} \in \mathbb{R}^m$ :

$$\mathbf{Ax} = \mathbf{b}, \text{ với } b_i = \mathbf{A}_{:,i} \mathbf{x}$$

với  $\mathbf{A}_{:,i}$  là vector hàng thứ  $i$  của ma trận  $\mathbf{A}$ .

- Ngoài ra còn có một phép nhân khác gọi là **Hadamard** (còn được gọi là *Element-wise*) hay được sử dụng trong Machine Learning. Tích Hadamard của hai ma trận có **cùng kích thước**  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$  được định nghĩa là:

$$\mathbf{C} = \mathbf{A} \odot \mathbf{B} \in \mathbb{R}^{m \times n}, \text{ với } c_{ij} = a_{ij} \times b_{ij}$$

## 4. Ma trận đơn vị và ma trận nghịch đảo

### 4.1. Ma trận đơn vị

- **Đường chéo chính** của một ma trận là tập hợp các phần tử có chỉ số hàng và chỉ số cột như nhau. Các định nghĩa này cũng áp dụng trên các ma trận không vuông.
- Nếu ma trận  $\mathbf{A} \in \mathbb{R}^{m \times n}$  thì đường chéo chính của  $\mathbf{A}$  gồm các phần tử:

$$\{a_{11}, a_{22}, \dots, a_{pp}\}, \text{ với } p = \min\{m, n\}$$

- Một ma trận đơn vị bậc  $n$  là một ma trận vuông thuộc  $\mathbb{R}^{n \times n}$  với các phần tử nằm trên đường chéo chính bằng 1, các phần tử còn lại bằng 0 và ma trận này được kí hiệu là  $\mathbf{I}$  (*identity matrix*).
- Ma trận đơn vị bậc  $n$  được kí hiệu là  $\mathbf{I}_n$ . Ví dụ:

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Giả sử có  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times m}$  và  $\mathbf{I}_n$  thì:

$$\mathbf{AI} = \mathbf{A}, \mathbf{IB} = \mathbf{B}$$

- Với mọi vector  $\mathbf{x} \in \mathbb{R}^n$  thì:

$$\mathbf{I}_n \mathbf{x} = \mathbf{x}$$

### 4.2. Ma trận nghịch đảo

- Một ma trận vuông  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , nếu tồn tại một ma trận vuông  $\mathbf{B} \in \mathbb{R}^{n \times n}$  sao cho:

$$\mathbf{AB} = \mathbf{I}$$

- Thì  $\mathbf{B}$  được gọi là **ma trận nghịch đảo** (*inverse matrix*) của  $\mathbf{A}$  và lúc này  $\mathbf{A}$  được gọi là **khả nghịch** (*invertible, nonsingular, nondegenerate*).
- Nếu không tồn tại ma trận  $\mathbf{B}$  thỏa điều kiện trên thì ta nói ma trận  $\mathbf{A}$  **không khả nghịch** (*singular, degenerate*).
- Nếu  $\mathbf{A}$  khả nghịch thì ma trận nghịch đảo của  $\mathbf{A}$  kí hiệu là  $\mathbf{A}^{-1}$ .
- Ma trận nghịch đảo được sử dụng để giải hệ phương trình tuyến tính. Giả sử  $\mathbf{A} \in \mathbb{R}^{n \times n}$  là một ma trận khả nghịch và một vector  $\mathbf{b} \in \mathbb{R}^n$  bất kì. Khi đó, phương trình:

$$\mathbf{Ax} = \mathbf{b}$$

có nghiệm duy nhất  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ .

- Giả sử hai ma trận  $\mathbf{A}$ ,  $\mathbf{B}$  là hai ma trận khả nghịch thì **tích của chúng cũng khả nghịch** và:

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

## 5. Một vài ma trận đặc biệt khác

### 5.1. Ma trận đường chéo

- Ma trận đường chéo** (*diagonal matrix*) là ma trận chỉ có các thành phần trên đường chéo chính **khác không**. Định nghĩa này cũng áp dụng cho các ma trận ko vuông.
- Ma trận không** là ma trận có tất cả các phần tử trong ma trận bằng không.
- Ma trận không và ma trận đơn vị cũng được gọi là các ma trận đường chéo.
- Một vài ví dụ về ma trận đường chéo:

$$\begin{bmatrix} 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$$

- Với các **ma trận đường chéo vuông**, thay vì viết cả ma trận, ta có thể chỉ liệt kê các thành phần nằm trên đường chéo chính. Ví dụ một ma trận đường chéo vuông  $\mathbf{A} \in \mathbb{R}^{m \times m}$  có thể dc kí hiệu là:

$$\text{diag}(a_{11}, a_{22}, \dots, a_{nn})$$

với  $a_{ii}$  là phần tử nằm ở dòng  $i$  cột  $i$  của ma trận  $\mathbf{A}$ .

- Tích, tổng của hai ma trận đường chéo vuông **cùng bậc** cũng là một ma trận đường chéo vuông.
- Một ma trận đường chéo vuông khả nghịch nếu mọi phần tử nằm trên đường chéo chính khác không.
- Nghịch đảo của một ma trận đường chéo vuông khả nghịch cũng là một ma trận đường chéo, cụ thể:

$$(\text{diag}(a_1, a_2, \dots, a_n))^{-1} = \text{diag}(a_1^{-1}, a_2^{-1}, \dots, a_n^{-1})$$

## 5.2. Ma trận tam giác

- Một ma trận vuông dc gọi là **ma trận tam giác trên** (*upper triangular matrix*) nếu tất cả các thành phần nằm dưới đường chéo chính bằng không.
- Một ma trận vuông dc gọi là **ma trận tam giác dưới** (*lower triangular matrix*) nếu tất cả các thành phần nằm trên đường chéo chính bằng không.
- Xét hệ:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1,n-1}x_{n-1} + a_{1n}x_n = b_1 \\ a_{22}x_2 + \dots + a_{2,n-1}x_{n-1} + a_{2n}x_n = b_2 \\ \dots \quad \dots \quad \dots \quad \dots \\ a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1} \\ a_{nn}x_n = b_n \end{cases}$$

- Hệ này có thể dc viết gọn dưới dạng  $\mathbf{Ax} = \mathbf{b}$  với  $\mathbf{A}$  là ma trận tam giác trên.
- Hệ phương trình này có thể được giải mà không cần tìm ma trận nghịch đảo  $\mathbf{A}^{-1}$ .
- Nếu ma trận  $\mathbf{A}$  là ma trận tam giác trên thì có thể áp dụng phương pháp **back substitution**.
- Nếu ma trận  $\mathbf{A}$  là ma trận tam giác dưới thì có thể áp dụng phương pháp **forward substitution**.

## 6. Định thức

### 6.1. Định nghĩa

- Định thức của một ma trận  $\mathbf{A}$  được kí hiệu là  $\det(\mathbf{A})$ .
- Với  $n = 1$ ,  $\det(\mathbf{A})$  chính là phần tử duy nhất của ma trận đó.
- Với ma trận vuông bậc  $n > 1$  thì:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \ddots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \Rightarrow \det(\mathbf{A}) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij})$$

- Trong đó  $i$  là một giá trị bất kì sao cho  $1 \leq i \leq n$  và  $\mathbf{A}_{ij}$  là **phần bù đại số** của  $\mathbf{A}$  bằng cách xóa đi dòng  $i$  và cột  $j$  của ma trận  $\mathbf{A}$  ban đầu.

### 6.2. Tính chất

- $\det(\mathbf{A}) = \det(\mathbf{A}^T)$
- Định thức của một ma trận đường chéo vuông bằng tích các phần tử nằm trên đường chéo chính.
- Định thức của ma trận đơn vị bằng 1.
- Định thức của một tích bằng tích các định thức:

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$$

với  $\mathbf{A}, \mathbf{B}$  là hai ma trận vuông cùng chiều.

- Nếu ma trận có một vector hàng hoặc vector cột là một vector không thì định thức của ma trận này bằng 0.
- Một ma trận khả nghịch khi và chỉ khi định thức của nó khác 0.
- Nếu một ma trận khả nghịch, định thức của ma trận nghịch đảo bằng nghịch đảo định thức của ma trận ban đầu.

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})} \text{ vì } \det(\mathbf{A})\det(\mathbf{A}^{-1}) = \det(\mathbf{AA}^{-1}) = \det(\mathbf{I}) = 1.$$

## 7. Tổ hợp tuyến tính, không gian sinh

### 7.1. Tổ hợp tuyến tính

- Xem thêm tại đây <https://textbooks.math.gatech.edu/ila/linear-independence.html>.
- Cho các **vector khác không**  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$  và **các số thực**  $x_1, \dots, x_n \in \mathbb{R}$ , thì vector:

$$\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n$$

được gọi là một **tổ hợp tuyến tính** (*linear combination*) của  $\mathbf{a}_1, \dots, \mathbf{a}_n$ .

- Xét ma trận  $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n] \in \mathbb{R}^{m \times n}$  và  $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$  thì ta có biểu thức:

$$\mathbf{b} = \mathbf{Ax}$$

lúc này ta nói  $\mathbf{b}$  là một tổ hợp tuyến tính các cột của  $\mathbf{A}$ .

- Tập hợp của **tất cả các vector** có thể dc biểu diễn dưới dạng một **tổ hợp tuyến tính** của các cột của một ma trận dc gọi là **không gian sinh** (*span space, span*) các cột của ma trận đó. Không gian sinh của một hệ các vector dc kí hiệu là  $\text{span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$ .
- Nếu phương trình:

$$0 = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n$$

- Có nghiệm duy nhất  $x_1 = x_2 = \dots = x_n = 0$  thì ta nói rằng hệ  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$  là một hệ **độc lập tuyến tính** (*linear independence*).
- Ngược lại, nếu tồn tại  $x_i \neq 0$  sao cho phương trình trên thỏa mãn, ta nói rằng hệ **phụ thuộc tuyến tính** (*linear dependence*).

## 7.2. Tính chất

- Một hệ là phụ thuộc tuyến tính nếu và chỉ nếu tồn tại một vector trong hệ đó là tổ hợp tuyến tính của các vector còn lại:

$$\mathbf{a}_i = \frac{-x_1}{x_i} \mathbf{a}_1 + \dots + \frac{-x_{i-1}}{x_i} \mathbf{a}_{i-1} + \frac{-x_{i+1}}{x_i} \mathbf{a}_{i+1} + \dots + \frac{-x_n}{x_i} \mathbf{a}_n$$

Lúc này  $\mathbf{a}_i$  là tổ hợp tuyến tính của các vector còn lại.

- Tập con khác rỗng của một hệ độc lập tuyến tính là một hệ độc lập tuyến tính.
- Tập hợp các cột của một ma trận khả nghịch tạo thành một hệ độc lập tuyến tính. Giả sử ma trận  $\mathbf{A}$  khả nghịch thì phương trình  $\mathbf{A}\mathbf{x} = \mathbf{0}$  có nghiệm duy nhất  $\mathbf{x} = \mathbf{A}^{-1} \times \mathbf{0}$ , vì vậy các cột của  $\mathbf{A}$  tạo thành một hệ độc lập tuyến tính.
- Nếu  $\mathbf{A}$  là một **ma trận cao** (*tall matrix*), tức số dòng lớn hơn số cột  $m > n$  thì sẽ tồn tại một vector  $\mathbf{b}$  sao cho  $\mathbf{A}\mathbf{x} = \mathbf{b}$  vô nghiệm.
- Nếu  $n > m$ , thì  $n$  vector bất kì trong không gian  $m$  chiều tạo thành một hệ phụ thuộc tuyến tính.

## 7.3. Cơ sở của một không gian

- Một hệ các vector  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  trong không gian vector  $m$  chiều  $V = \mathbb{R}^m$  dc gọi là một **cơ sở** (*basic*) nếu hai điều kiện sau thỏa mãn:
  - $V \equiv \text{span}(\mathbf{a}_1, \dots, \mathbf{a}_n)$ .
  - $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  là một hệ độc lập tuyến tính.
- Khi đó mọi vector  $\mathbf{b} \in V$  đều có thể biểu diễn **duy nhất** dưới dạng một tổ hợp tuyến tính của các  $\mathbf{a}_i$ .
- Từ đây, kết hợp với hai tính chất cuối của mục 7.2 thì ta có thể suy ra rằng  $m = n$ .

## 7.4. Range và Null space

- Với mỗi ma trận  $\mathbf{A} \in \mathbb{R}^{m \times n}$  thì có hai không gian con quan trọng ứng vs ma trận này:
  - Range của  $\mathbf{A}$ , dc định nghĩa là:

$$\mathcal{R}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^m : \exists \mathbf{x} \in \mathbb{R}^n, \mathbf{A}\mathbf{x} = \mathbf{y}\}$$

- Nói cách khác,  $\mathcal{R}(\mathbf{A})$  là tập hợp các điểm là tổ hợp tuyến tính các cột của  $\mathbf{A}$ .



- $\mathcal{R}(\mathbf{A})$  là một ko gian con của  $\mathbb{R}^m$  vs số chiều là số lượng lớn nhất các cột của  $\mathbf{A}$  độc lập tuyến tính.
- Null của  $\mathbf{A}$ , kí hiệu là  $\mathcal{N}$ , dc định nghĩa là:

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

- Mỗi vector trong  $\mathcal{N}(\mathbf{A})$  chính là một **bộ các hệ số** làm cho tổ hợp tuyến tính của  $\mathbf{A}$  thành một vector  $\mathbf{0}$ .
- $\mathcal{N}(\mathbf{A})$  có thể dc chứng minh là một ko gian con trong  $\mathbb{R}^n$ . Khi các cột của  $\mathbf{A}$  độc lập tuyến tính thì  $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$  (chỉ bao gồm các vector  $\mathbf{0}$ ).
- $\mathcal{R}(\mathbf{A})$  và  $\mathcal{N}(\mathbf{A})$  là các ko gian con vector với số chiều lần lượt là  $\dim(\mathcal{R}(\mathbf{A}))$  và  $\dim(\mathcal{N}(\mathbf{A}))$ , ta có tính chất:

$$\dim(\mathcal{R}(\mathbf{A})) + \dim(\mathcal{N}(\mathbf{A})) = n$$

## 8. Hạng của ma trận

- Xét một ma trận  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Hạng của ma trận  $\mathbf{A}$ , kí hiệu  $\text{rank}(\mathbf{A})$  dc định nghĩa là **số lượng lớn nhất các cột** của nó tạo thành một hệ độc lập tuyến tính.
- Các tính chất quan trọng của hạng:
  - Ma trận có hạng bằng 0 khi nó là ma trận  $\mathbf{0}$ .
  - $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T)$ , tức số lượng lớn nhất các cột độc lập tuyến tính của  $\mathbf{A}$  bằng với số lượng lớn nhất các dòng độc lập tuyến tính của nó.
  - Nếu  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , thì  $\text{rank}(\mathbf{A}) \leq \min(m, n)$ .
  - $\text{rank}(\mathbf{AB}) \leq \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}))$
  - $\text{rank}(\mathbf{A} + \mathbf{B}) \leq \text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B})$
  - Bất đẳng thức Sylvester, nếu  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times k}$  thì:

$$\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{B}) - n \leq \text{rank}(\mathbf{AB})$$

- Xét một ma trận vuông  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , hai điều kiện bất kì dưới đây là tương đương:
  - $\mathbf{A}$  khả nghịch.
  - Các cột của  $\mathbf{A}$  tạo thành một cơ sở trog ko gian  $n$  chiều.
  - $\det(\mathbf{A}) \neq 0$ .
  - $\text{rank}(\mathbf{A}) = n$ .

## 9. Hệ trục chuẩn, ma trận trực giao

### 9.1. Định nghĩa

- Một hệ cơ sở  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m \in \mathbb{R}^m\}$  dc gọi là **trực giao** (*orthogonal*) nếu mỗi vector trong hệ này khác 0 và tích vô hướng của **hai vector khác nhau bất kì bằng 0**.

$$\mathbf{u}_i \neq 0; \mathbf{u}_i^T \mathbf{u}_j = 0, \forall 1 \leq i \neq j \leq m$$

- Một hệ cơ sở  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m \in \mathbb{R}^m\}$  dc gọi là **trực chuẩn** (*orthonormal*) nếu nó **trực giao** và có **độ dài euclidean** ( $l_2$  norm) của mỗi vector bất kì trong hệ bằng 1.

$$\mathbf{u}_i^T \mathbf{u}_j = \begin{cases} 1 & \text{nếu } i = j \\ 0 & \text{otherwise} \end{cases}$$

- Gọi  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m]$  với  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m \in \mathbb{R}^m\}$  là **trực chuẩn**, từ định lí phía trên có thể suy ra:

$$\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$$

trong đó  $\mathbf{I}$  là ma trận đơn vị bậc  $m$ .

- Nếu một ma trận nào đó thỏa mãn điều kiện trên thì ta gọi đó là **ma trận trực giao** (*orthogonal matrix*). Lưu ý ko có định nghĩa cho ma trận trực chuẩn, mà cũng ko tồn tại một khái niệm nào là ma trận trực chuẩn.
- Nếu một **ma trận vuông phức**  $\mathbf{U}$  thỏa mãn  $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$  thì ta nói  $\mathbf{U}$  là ma trận **unitary** (*unitary matrix*).

## 9.2. Tính chất của ma trận trực giao

- $\mathbf{U}^{-1} = \mathbf{U}^T$ : nghịch đảo của một ma trận trực giao chính là chuyển vị của nó.
- Nếu  $\mathbf{U}$  là ma trận trực giao thì  $\mathbf{U}^T$  cũng là ma trận trực giao.
- Nếu  $\mathbf{U}$  là ma trận trực giao thì  $\det(\mathbf{U})$  bằng 1 hoặc  $-1$ . Vì:

$$\det(\mathbf{U}) = \det(\mathbf{U}^T) \text{ và } \det(\mathbf{U})\det(\mathbf{U}^T) = \det(\mathbf{I}) = 1$$

- Nếu một ma trận  $\mathbf{U}$  nào đó trực giao, thì ta có thể thực hiện phép **xoay một vector**. Giả sử có hai vector  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$  và  $\mathbf{U} \in \mathbb{R}^{m \times m}$  là ma trận trực giao. Dùng ma trận này để xoay hai vector  $\mathbf{x}, \mathbf{y}$  ta được hai vector mới là  $\mathbf{U}\mathbf{x}, \mathbf{U}\mathbf{y}$ . Lúc này tích vô hướng của hai vector  $\mathbf{U}\mathbf{x}, \mathbf{U}\mathbf{y}$  là:

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

như vậy **phép xoay ko làm thay đổi tích vô hướng của hai vector**.

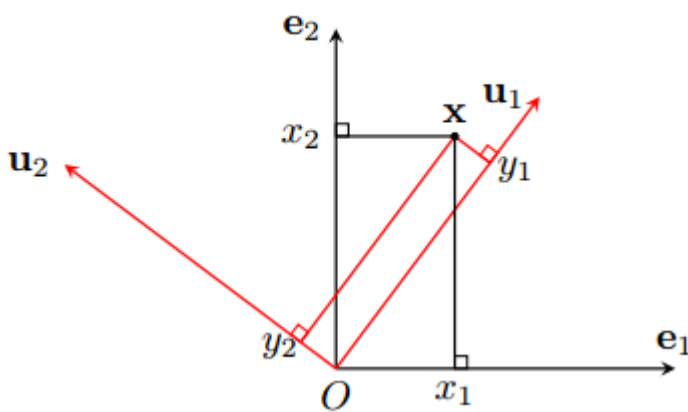
- Giả sử  $\hat{\mathbf{U}} \in \mathbb{R}^{m \times r}$ , với  $r < m$  là một ma trận con dc tạo ra từ  $r$  cột của ma trận trực giao  $\mathbf{U}$ , thì:

$$\hat{\mathbf{U}}^T \hat{\mathbf{U}} = \mathbf{I}_r$$

# 10. Biểu diễn vector trong các hệ cơ sở khác nhau

- Trong không gian  $m$  chiều, tọa độ của mỗi điểm được xác định dựa trên một hệ tọa độ nào đó. Ở các hệ tọa độ khác nhau thì tọa độ của các điểm cũng khác nhau.
- Tập hợp các vector  $\mathbf{e}_1, \dots, \mathbf{e}_m$  mà mỗi vector  $\mathbf{e}_i$  có đúng một phần tử khác 0 ở thành phần thứ  $i$  và phần tử đó bằng 1 được gọi là **hệ cơ sở đơn vị** (hoặc hệ đơn vị, hệ chính tắc) trong không gian  $m$  chiều. Nếu xếp các vector  $\mathbf{e}_i$  với  $i = 1, 2, \dots, m$  theo đúng thứ tự đó thì ta nhận được ma trận đơn vị  $\mathbf{I}$  bao gồm  $m$  chiều.
- Mỗi vector cột  $\mathbf{x} = [x_1, x_2, \dots, x_m] \in \mathbb{R}^m$  có thể được gọi là một tổ hợp tuyến tính của các vector trong hệ cơ sở chính tắc:

$$\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_m \mathbf{e}_m$$



**Hình 1.1:** Chuyển đổi tọa độ trong các hệ cơ sở khác nhau. Trong hệ tọa độ  $O\mathbf{e}_1\mathbf{e}_2$ ,  $\mathbf{x}$  có tọa độ là  $(x_1, x_2)$ . Trong hệ tọa độ  $O\mathbf{u}_1\mathbf{u}_2$ ,  $\mathbf{x}$  có tọa độ là  $(y_1, y_2)$ .

- Giả sử có một hệ cơ sở khác  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$  và các vector này độc lập tuyến tính, thì để biểu diễn một vector  $\mathbf{x}$  nào đó từ hệ cơ sở mới này có dạng:

$$\mathbf{x} = y_1 \mathbf{u}_1 + y_2 \mathbf{u}_2 + \dots + y_m \mathbf{u}_m = \mathbf{U}\mathbf{y}$$

với  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m]$ . Lúc này vector  $\mathbf{y} = [y_1, y_2, \dots, y_m]^T$  chính là biểu diễn của  $\mathbf{x}$  trong hệ cơ sở mới. Biểu diễn này là **duy nhất** vì  $\mathbf{y} = \mathbf{U}^{-1}\mathbf{x}$ .

- Trong **các ma trận đóng vai trò như một hệ cơ sở**, các **ma trận trực giao**, tức  $\mathbf{U}^T\mathbf{U} = \mathbf{I}$  sẽ được quan tâm nhiều hơn vì:
  - Ngược đảo của chúng cũng chính là chuyển vị của chúng**, tức  $\mathbf{U}^{-1} = \mathbf{U}^T$ , từ đây  $\mathbf{y}$  có thể được tính một cách nhanh chóng bằng  $\mathbf{y} = \mathbf{U}^T\mathbf{x}$ . Từ đó suy ra  $y_i = \mathbf{x}^T \mathbf{u}_i = \mathbf{u}_i^T \mathbf{x}$  với  $i = 1, \dots, m$ .
- Việc chuyển đổi hệ cơ sở sử dụng ma trận trực giao có thể được coi như một phép xoay trực tọa độ. Nhìn theo một cách khác thì đây cũng là một phép xoay vector theo chiều ngược lại nếu ta coi các trục tọa độ là cố định.

# 11. Trị riêng và vector riêng

## 11.1. Định nghĩa

- Cho một ma trận vuông  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , một vector  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{x} \neq 0$  và một số vô hướng  $\lambda$  (có thể thực hoặc phức). Nếu  $\mathbf{Ax} = \lambda\mathbf{x}$  thì ta nói  $\lambda$  và  $\mathbf{x}$  là một cặp trị riêng và vector riêng (*eigenvalue* và *eigenvector*) của ma trận  $\mathbf{A}$ .
- Từ định nghĩa trên, ta cũng có  $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$ , tức  $\mathbf{x}$  nằm trong null space của  $(\mathbf{A} - \lambda\mathbf{I})$ , mà vì  $\mathbf{x} \neq 0$  cho nên ma trận  $(\mathbf{A} - \lambda\mathbf{I})$  ko khả nghịch. Nói cách khác  $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ , tức  $\lambda$  là nghiệm của phương trình  $\det(\mathbf{A} - t\mathbf{I}) = 0$ . Định thức này là một đa thức bậc  $n$  của  $t$ , dc gọi là **đa thức đặc trưng** (*characteristic polynomial*) của  $\mathbf{A}$ , kí hiệu là  $p_{\mathbf{A}}(t)$ .
- Tập hợp của tất cả các trị riêng của một ma trận vuông còn được gọi là **phổ** (*spectrum*) của ma trận đó.

## 11.2. Tính chất

- Nếu  $\mathbf{x}$  là vector riêng của  $\mathbf{A}$  ứng với  $\lambda$  thì  $k\mathbf{x}$ ,  $\forall k \neq 0$  cũng là vector riêng ứng với trị riêng  $\lambda$  đó.
  - Nếu  $\mathbf{x}_1, \mathbf{x}_2$  là hai vector riêng ứng với cùng trị riêng  $\lambda$ , thì tổng của chúng cũng là một vector ứng với trị riêng  $\lambda$  này.
  - Từ đây, có thể suy ra tập hợp các vector riêng ứng với một trị riêng của một ma trận vuông tạo thành một không gian vector con, dc gọi là **không gian riêng** (*eigenspace*) ứng với trị riêng đó.
- Mọi ma trận vuông bậc  $n$  đều có  $n$  trị riêng (có khả năng lặp lại) và có thể là các số phức.
- Tích của tất cả các trị riêng** của một ma trận bằng **định thức của ma trận đó**. Tổng tất cả các trị riêng của một ma trận bằng **tổng các phần tử nằm trên đường chéo chính của ma trận đó**.
- Phổ của một ma trận bằng phổ của ma trận chuyển vị của nó.
- Nếu  $\mathbf{A}, \mathbf{B}$  là các ma trận vuông cùng bậc thì  $p_{\mathbf{AB}}(t) = p_{\mathbf{BA}}(t)$ . Điều này có nghĩa là dù tích của hai ma trận **ko có tính chất giao hoán** nhưng đa thức đặc trưng của  $\mathbf{AB}$  và  $\mathbf{BA}$  là như nhau. Tức phổ của hai tích  $\mathbf{AB}$  và  $\mathbf{BA}$  trùng nhau.
- Với ma trận đối xứng (hoặc tổng quát, Hermitian), tất cả các trị riêng của nó đều là các số thực. Giả sử  $\lambda$  là một trị riêng của một ma trận Hermitian  $\mathbf{A}$  và  $\mathbf{x}$  là vector riêng ứng với trị riêng đó. Từ định nghĩa ta suy ra:

$$\mathbf{Ax} = \lambda\mathbf{x} \Rightarrow (\mathbf{Ax})^H = \bar{\lambda}\mathbf{x}^H \Rightarrow \bar{\lambda}\mathbf{x}^H = \mathbf{x}^H \mathbf{A}$$

với  $\bar{\lambda}$  là liên hiệp phức của số vô hướng  $\lambda$ . Nhân cả hai vế vào bên phải với  $\mathbf{x}$  ta có:

$$\bar{\lambda} \mathbf{x}^H \mathbf{x} = \mathbf{x}^H \mathbf{A} \mathbf{x} \Rightarrow (\lambda - \bar{\lambda}) \mathbf{x}^H \mathbf{x} = 0$$

vì  $\mathbf{x} \neq 0$  nên  $\mathbf{x}^H \mathbf{x} \neq 0$ , từ đó suy ra  $\bar{\lambda} = \lambda$ , vậy nên  $\lambda$  phải là một số thực. Từ đó:

$$\bar{\lambda} \mathbf{x}^H \mathbf{x} = \mathbf{x}^H \mathbf{A} \mathbf{x} = \lambda \mathbf{x}^H \mathbf{x}$$

- Nếu  $(\lambda, \mathbf{x})$  là một cặp trị riêng, vector riêng của một ma trận khả nghịch  $\mathbf{A}$  thì  $(\frac{1}{\lambda}, \mathbf{x})$  là một cặp trị riêng, vector riêng của  $\mathbf{A}^{-1}$ , vì  $\mathbf{A} \mathbf{x} = \lambda \mathbf{x} \Rightarrow \frac{1}{\lambda} \mathbf{x} = \mathbf{A}^{-1} \mathbf{x}$

## 12. Chéo hóa ma trận

### 12.1. Định nghĩa

- Giả sử ta có  $\lambda_1, \dots, \lambda_n$  (có thể lặp lại) và  $\mathbf{x}_1, \dots, \mathbf{x}_n \neq 0$  lần lượt là các trị riêng và vector riêng của một ma trận vuông  $\mathbf{A}$ . Tức:

$$\mathbf{A} \mathbf{x}_i = \lambda_i \mathbf{x}_i, \forall i = 1, \dots, n$$

- Đặt  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  (ma trận đường chéo) và  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ , ta sẽ có  $\mathbf{A} \mathbf{X} = \mathbf{X} \mathbf{\Lambda}$
- Hơn nữa, nếu các vector riêng  $\mathbf{x}_1, \dots, \mathbf{x}_n$  độc lập tuyến tính thì  $\mathbf{X}$  khả nghịch. Khi đó ta có thể viết  $\mathbf{A}$  dưới dạng **tích của ba ma trận**:

$$\mathbf{A} = \mathbf{X} \mathbf{\Lambda} \mathbf{X}^{-1}$$

- Các vector riêng  $\mathbf{x}_i$  **thường dc chọn** sao cho  $\mathbf{x}_i^T \mathbf{x}_i = 1$ . Cách biểu diễn ma trận như trên dc gọi là **eigendecomposition** vì nó tách ra thành tích của các ma trận đặc biệt dựa trên vector riêng và trị riêng. Và vì  $\mathbf{\Lambda}$  là ma trận đường chéo cho nên cách triển khai này còn dc gọi là chéo hóa ma trận.

### 12.2. Tính chất

- Khái niệm chéo hóa ma trận chỉ áp dụng với ma trận vuông. Vì ko có định nghĩa vector riêng hay trị riêng cho ma trận ko vuông.
- Ko phải ma trận vuông nào cũng có thể chéo hóa dc (*diagonalizable*). Một ma trận vuông bậc  $n$  chéo hóa dc nếu có đủ  $n$  vector riêng độc lập tuyến tính.
- Nếu một ma trận chéo hóa dc, thì sẽ có nhiều hơn một cách chéo hóa ma trận đó. Chỉ cần đổi vị trí các  $\lambda_i$  và vị trí tương ứng của các cột của  $\mathbf{X}$  ta sẽ có dc một cách chéo hóa mới.

- Nếu  $\mathbf{A}$  có thể biểu diễn dc dưới dạng tích của ba ma trận, thì lúc này lũy thừa của nó cũng chéo hóa dc:

$$\mathbf{A}^2 = (\mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1})(\mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}) = \mathbf{X}\mathbf{\Lambda}^2\mathbf{X}^{-1}$$

$$\mathbf{A}^k = \mathbf{X}\mathbf{\Lambda}^k\mathbf{X}^{-1}, \forall k \in \mathbb{N}$$

- Nếu  $\mathbf{A}$  khả nghịch, thì  $\mathbf{A}^{-1} = (\mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1})^{-1} = \mathbf{X}\mathbf{\Lambda}^{-1}\mathbf{X}^{-1}$ . Vậy có thể áp dụng chéo hóa ma trận để tìm ma trận nghịch đảo.

## 13. Ma trận xác định dương

### 13.1. Định nghĩa

- Nếu một ma trận đối xứng  $\mathbf{A} \in \mathbb{R}^{n \times n}$  dc gọi là **xác định dương** (*positive definite*) nếu:

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

- Một ma trận đối xứng  $\mathbf{A} \in \mathbb{R}^{n \times n}$  dc gọi là **nửa xác định dương** (*positive semidefinite*), thường dc kí hiệu là *PSD* nếu:

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

- Trên thực tế, ma trận nửa xác định dương thường dc sử dụng nhiều hơn.
- Ma trận **xác định âm** (*negative definite*) và **nửa xác định âm** (*negative semidefinite*) cũng dc định nghĩa tương tự.
- Các kí hiệu  $\mathbf{A} \succ 0$ ,  $\mathbf{A} \succeq 0$ ,  $\mathbf{A} \prec 0$ ,  $\mathbf{A} \preceq 0$  lần lượt dc dùng để chỉ một ma trận là xác định dương, nửa xác định dương, xác định âm, nửa xác định âm.
- Kí hiệu  $\mathbf{A} \succ \mathbf{B}$  cũng dc dùng để chỉ ra rằng  $\mathbf{A} - \mathbf{B} \succ 0$ .
- Một ma trận phức, Hermitian  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , dc gọi là xác định dương nếu:

$$\mathbf{x}^H \mathbf{A} \mathbf{x} > 0, \forall \mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq 0$$

- Ví dụ,  $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$  là nửa xác định dương vì với mọi vector  $\mathbf{x} = \begin{bmatrix} u \\ v \end{bmatrix}$ , ta có:

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = u^2 + v^2 - 2vu = (u - v)^2 \geq 0, \forall u, v \in \mathbb{R}$$

### 13.2. Tính chất

- Mọi trị riêng của một ma trận xác định dương đều là một số thực dương
  - Các trị riêng của các ma trận dạng này là số thực vì các ma trận đều là đối xứng (xem lại mục 11.2). Để chứng minh các trị riêng này là các số thực dương, giả sử  $\lambda$  là một trị riêng của một ma trận xác định dương  $\mathbf{A}$  và  $\mathbf{x}$  là vector riêng ứng với trị riêng này. Nhân vào bên trái cả hai vế của  $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$  với  $\mathbf{x}^H$  ta có:

$$\lambda \mathbf{x}^H \mathbf{x} = \mathbf{x}^H \mathbf{A} \mathbf{x} > 0$$

(ở đây Hermitian dc sử dụng để xét tổng quát cho cả trường hợp ma trận phức)

- Vì  $\mathbf{x}^H \mathbf{x}$  luôn dương với mọi  $\mathbf{x}$ . Ban đầu ta giả sử  $\mathbf{A}$  là ma trận xác định dương, nên  $\mathbf{x}^H \mathbf{A} \mathbf{x} > 0$ , từ đây có thể suy ra rằng  $\lambda > 0$ .
- Ta có thể chứng minh dc rằng mọi trị riêng của một **ma trận nửa xác định dương** là **ko âm**.
- Mọi ma trận xác định dương là **khả nghịch**. Ngoài ra, định thức của nó còn là một số dương.
  - Điều này có thể trực tiếp suy ra từ tính chất đầu tiên, vì định thức của một ma trận có thể dc tính bằng tích của tất cả các trị riêng của nó, mà theo tính chất đầu tiên, thì mọi trị riêng đều lớn hơn 0.
- Tiêu chuẩn Sylvester:
  - Một ma trận Hermitian là xác định dương nếu mọi **leading principal minors** của nó là **dương**.
  - Một ma trận Hermitian là nửa xác định dương nếu mọi **principal minors** của nó là **ko âm**.
  - Đây là một tiêu chuẩn để kiểm tra một ma trận Hermitian  $\mathbf{A} \in \mathbb{R}^n$  có phải là xác định dương hoặc nửa xác định dương hay không.
  - Khái niệm về **leading principal minors** và **principal minors** dc định nghĩa như sau:
    - Gọi  $\mathcal{L}$  là một tập con bất kì của  $\{1, 2, 3, \dots, n\}$ ,  $\mathbf{A}_{\mathcal{L}}$  là **ma trận con** của  $\mathbf{A}$  bằng cách trích ra các hàng và cột có chỉ số nằm trong  $\mathcal{L}$  của  $\mathbf{A}$ . Khi đó  $\mathbf{A}_{\mathcal{L}}$  và  $\det(\mathbf{A}_{\mathcal{L}})$  lần lượt dc gọi là **ma trận con chính** (*principal submatrix*) và **principal minor** của  $\mathbf{A}$ .
    - Nếu  $\mathcal{L}$  chỉ bao gồm **các số tự nhiên liên tiếp** từ 1 đến  $k \leq n$ , ta nói  $\mathbf{A}_{\mathcal{L}}$  và  $\det(\mathbf{A}_{\mathcal{L}})$  lần lượt là **leading principal submatrix** và **leading principal minor** bậc  $k$  của  $\mathbf{A}$ .
- Ngoài ra, ta có:

$$\mathbf{A} = \mathbf{B}^H \mathbf{B}$$

tức  $\mathbf{A}$  là ma trận nửa xác định dương với mọi ma trận  $\mathbf{B}$  ( $\mathbf{B}$  ko nhất thiết phải vuông).

- Thật vậy, với mọi vector  $\mathbf{x} \neq 0$  với chiều phù hợp:

$$\mathbf{x}^H \mathbf{A} \mathbf{x} = \mathbf{x}^H \mathbf{B}^H \mathbf{B} \mathbf{x} = (\mathbf{B} \mathbf{x})^H (\mathbf{B} \mathbf{x}) \geq 0$$

- Khai triển Cholesky** (*Cholesky decomposition*).
  - Mọi ma trận Hermitian - nửa xác định dương  $\mathbf{A}$  đều biểu diễn dc dưới dạng:

$$\mathbf{A} = \mathbf{L} \mathbf{L}^H$$

trong đó  $\mathbf{L}$  là một **ma trận tam giác dưới** với các phần tử nằm trên đường chéo là **thực dương**.

- Nếu  $\mathbf{A}$  là một ma trận nửa xác định dương thì:

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = 0 \Leftrightarrow \mathbf{A} \mathbf{x} = \mathbf{0}$$

- Nếu  $\mathbf{A} \mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x}^T \mathbf{A} \mathbf{x} = 0$  một cách hiển nhiên.
- Nếu  $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$ . Với vector  $\mathbf{y} \neq \mathbf{0}$  bất kì **có cùng kích thước** với  $\mathbf{x}$ , xét hàm số sau đây:

$$f(\lambda) = (\mathbf{x} + \lambda \mathbf{y})^T \mathbf{A} (\mathbf{x} + \lambda \mathbf{y})$$

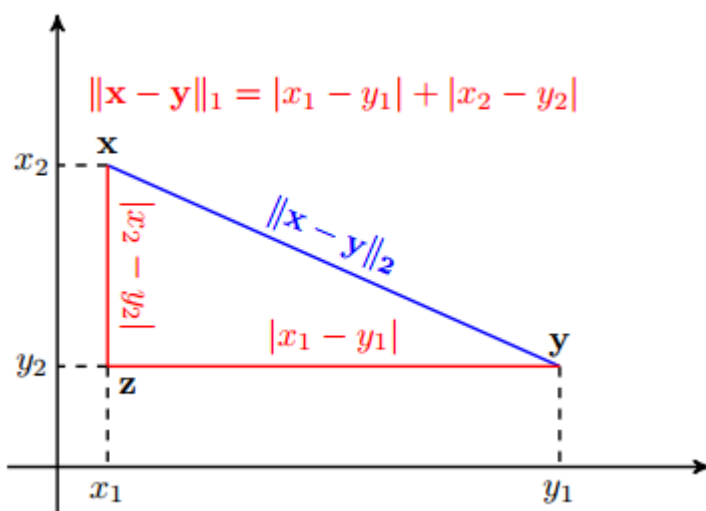
- Hàm số này **ko âm** với mọi  $\lambda$  vì  $\mathbf{A}$  là ma trận nửa xác định dương. Đây là một tam thức bậc hai của  $\lambda$ :

$$f(\lambda) = \mathbf{y}^T \mathbf{A} \mathbf{y} \lambda^2 + 2\mathbf{y}^T \mathbf{A} \mathbf{x} \lambda + \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{A} \mathbf{y} \lambda^2 + 2\mathbf{y}^T \mathbf{A} \mathbf{x} \lambda$$

Xét hai trường hợp:

- $\mathbf{y}^T \mathbf{A} \mathbf{y} = 0$ : Khi đó,  $f(\lambda) = 2\mathbf{y}^T \mathbf{A} \mathbf{x} \lambda \geq 0, \forall \lambda$  nếu và chỉ nếu  $\mathbf{y}^T \mathbf{A} \mathbf{x} = 0$ .
- $\mathbf{y}^T \mathbf{A} \mathbf{y} > 0$ : Khi đó tam thức bậc hai  $f(\lambda) \geq 0, \forall \lambda$  nếu và chỉ nếu  $\Delta' = (\mathbf{y}^T \mathbf{A} \mathbf{x})^2 \leq 0$  vì hệ số ứng với thành phần bậc hai bằng  $\mathbf{y}^T \mathbf{A} \mathbf{y} > 0$ . Điều này cũng đồng nghĩa với việc  $\mathbf{y}^T \mathbf{A} \mathbf{x} = 0$ .
- Tóm lại,  $\mathbf{y}^T \mathbf{A} \mathbf{x} = 0, \forall \mathbf{y} \neq \mathbf{0}$ . Điều này chỉ xảy ra nếu  $\mathbf{A} \mathbf{x} = \mathbf{0}$ .

## 14. Chuẩn của vector và ma trận



**Hình 1.2:** Minh họa  $\ell_1$  norm và  $\ell_2$  norm trong không gian hai chiều.  $\ell_2$  norm chính là khoảng cách giữa hai điểm trong mặt phẳng. Trong khi đó  $\ell_1$  norm là quãng đường ngắn nhất giữa hai điểm nếu chỉ được đi theo các đường song song với các trục tọa độ.

- Norm: Hay còn dc gọi là **chuẩn** dùng để xác định khoảng cách giữa hai vector bất kì.
- Cho một hàm số  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  dc gọi là norm nếu nó thỏa mãn ba điều kiện sau:
  - $f(\mathbf{x}) \geq 0$ , nếu dấu bằng xảy ra  $\Leftrightarrow \mathbf{x} = \mathbf{0}$



- Tính chất này dễ hiểu vì khoảng cách ko thể là số âm. Hơn nữa nếu khoảng cách giữa hai điểm trùng nhau thì điều đó cho thấy rằng hai điểm này trùng nhau.
- $f(\alpha \mathbf{x}) = |\alpha|f(\mathbf{x}), \forall \alpha \in \mathbb{R}$ .
  - Giả sử có ba điểm  $\mathbf{y}, \mathbf{v}, \mathbf{z}$  nằm thẳng hàng, hơn nữa  $\mathbf{v} - \mathbf{y} = \alpha(\mathbf{v} - \mathbf{z})$ , thì điều này có nghĩa rằng khoảng cách giữa  $\mathbf{v}$  và  $\mathbf{y}$  gấp  $|\alpha|$  lần khoảng cách giữa  $\mathbf{v}$  và  $\mathbf{z}$ .
- $f(\mathbf{x}_1) = f(\mathbf{x}_2) \geq f(\mathbf{x}_1 + \mathbf{x}_2), \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ .
  - Đây chính là bất đẳng thức tam giác, nếu ta coi  $\mathbf{x}_1 = \mathbf{y} - \mathbf{w}, \mathbf{x}_2 = \mathbf{w} - \mathbf{z}$  với  $\mathbf{y}, \mathbf{z}, \mathbf{w}$  là ba điểm bất kì nằm trong cùng một không gian.

## 14.1. Một số chuẩn vector thường dùng

- Độ dài Euclid của một vector  $\mathbf{x} \in \mathbb{R}^n$  chính là một norm, norm này dc gọi là  $l_2$  norm hoặc Euclid norm:

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

- Bình phương của  $l_2$  norm chính là **tích vô hướng của một vector với chính nó**,  $\|\mathbf{x}\|_2^2 = \mathbf{x}^T \mathbf{x}$ . Với  $p \geq 1$ , ta có hàm số:

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{\frac{1}{p}}$$

và được gọi là  $l_p$  norm.

- Khi  $p \rightarrow \infty$ , giả sử  $i = \arg \max_{j=1,2,\dots,n} |x_j|$ . Khi đó:

$$\|\mathbf{x}\|_p = |x_i| \left( 1 + \left| \frac{x_1}{x_i} \right|^p + \dots + \left| \frac{x_{i-1}}{x_i} \right|^p + \left| \frac{x_{i+1}}{x_i} \right|^p + \dots + \left| \frac{x_n}{x_i} \right|^p \right)^{\frac{1}{p}}$$

- Ta thấy rằng:

$$\lim_{p \rightarrow \infty} \left( 1 + \left| \frac{x_1}{x_i} \right|^p + \dots + \left| \frac{x_{i-1}}{x_i} \right|^p + \left| \frac{x_{i+1}}{x_i} \right|^p + \dots + \left| \frac{x_n}{x_i} \right|^p \right)^{\frac{1}{p}} = 1$$

và vì đại lượng nằm trong dấu ngoặc đơn ko vượt quá  $n$ , ta sẽ có:

$$\|\mathbf{x}\|_\infty \triangleq \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = |x_i| = \max_{j=1,2,\dots,n} |x_j|$$

## 14.2. Chuẩn Frobenius của ma trận

- Với một ma trận  $\mathbf{A} \in \mathbb{R}^{m \times n}$  chuẩn thường dc dùng nhất là chuẩn Frobenius, kí hiệu là  $\|\mathbf{A}\|_F$  là căn bậc hai của tổng bình phương tất cả các phần tử của ma trận đó:

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$$

- Cần chú ý rằng  $l_2$  norm là một norm khác của ma trận, ko dc phổ biến bằng Forbenius norm.

## 14.3. Vết của ma trận

- **Vết** (*trace*) của một ma trận vuông là tổng của tất cả các phần tử nằm trên đường chéo chính của ma trận đó.
- Vết của một ma trận  $\mathbf{A}$  kí hiệu là  $\text{trace}(\mathbf{A})$ .
- Giả sử rằng các ma trận trong hàm trace là vuông và các phép nhân ma trận thực hiện dc, ta có các tính chất quan trọng sau:
  - Một ma trận vuông bất kì và chuyển vị của nó có trace bằng nhau:

$$\text{trace}(\mathbf{A}) = \text{trace}(\mathbf{A}^T)$$

- Trace của một tổng bằng tổng các trace:

$$\text{trace}\left(\sum_{i=1}^k \mathbf{A}_i\right) = \sum_{i=1}^k \text{trace}(\mathbf{A}_i)$$

- Với  $k$  là một số bất kì, ta có:

$$\text{trace}(k\mathbf{A}) = k \text{trace}(\mathbf{A})$$

- Với ma trận vuông  $\mathbf{A}$  và các trị riêng  $\lambda_i, i = 1, 2, \dots, n$  của nó (có thể phức hoặc lặp), thì:

$$\text{trace}(\mathbf{A}) = \sum_{i=1}^n \lambda_i$$

- $\text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})$
- $\text{trace}(\mathbf{ABC}) = \text{trace}(\mathbf{BCA})$  nhưng  $\text{trace}(\mathbf{ABC})$  ko đồng nhất với  $\text{trace}(\mathbf{ACB})$
- Nếu  $\mathbf{X}$  là một ma trận khả nghịch cùng chiều với  $\mathbf{A}$  thì:

$$\text{trace}(\mathbf{XAX}^{-1}) = \text{trace}(\mathbf{X}^{-1}\mathbf{XA}) = \text{trace}(\mathbf{A})$$

- $\|\mathbf{A}\|_F^2 = \text{trace}(\mathbf{AA}^T) = \text{trace}(\mathbf{A}^T\mathbf{A})$  với  $\mathbf{A}$  là một ma trận bất kì. Từ đây ta có  $\text{trace}(\mathbf{AA}^T) \geq 0$  với mọi ma trận  $\mathbf{A}$ .