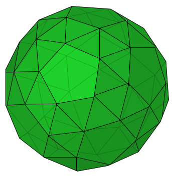
Projects in Mathematics and Applications

ĐẠI SỐ TUYẾN TÍNH

Linh Tran *, Nghia Nguyen †, Trung Can ‡



* National University of Singapore, Singapore
† Ho Chi Minh City University of Science, Vietnam
‡ California Institute of Technology, USA

Mục lục

Mở đầu

Đại số tuyến tính là một ngành của Toán học nghiên cứu các phương trình tuyến tính, các ánh xạ tuyến tính, và biểu diễn của chúng dưới dạng ma trận trong không gian vector.

Ngoài việc đóng vai trò nền tảng trong hầu hết mọi lĩnh vực khác của Toán cao cấp, đại số tuyến tính cung cấp các công cụ cơ bản để nghiên cứu vector, ma trận, hàm số liên tục, giúp ích trong việc mô hình hóa dữ liệu, hình ảnh, tối ưu hóa, và từ đó có nhiều mô hình ứng dụng được trong các bài toán thực tế nói chung và trong Machine Learning nói riêng.

Các tác giả viết tài liệu này với mong muốn vừa giúp bạn đọc hiểu được ý nghĩa toán học của các khái niệm và định lý, qua đó cảm nhận được vẻ đẹp của ĐSTT, vừa thu nhận được những kiến thức cơ bản thường dùng trong Machine Learning (ML). Để dung hòa được cả hai mục tiêu này thực sự khó khăn, các tác giả xin chân thành cảm ơn anh **Lê Việt Hải**[1] đã đóng góp và chỉnh sửa lỗi sai cho tài liệu này.

1 University of Pennsylvania State, U.S.A

Danh pháp tiếng Anh

Danh sách những khái niệm Toán học trong tài liệu và danh pháp tiếng Anh của chúng để tiện cho những bạn tra cứu online:

1. Đại số: Algebra
2. Tuyến tính: Linear
3. Số thực: Real/Real Number
4. Vector: Vector
5. Độ dài: Norm
6. Cộng: Add
7. Tổng: Sum
8. Nhân: Multiply
9. Tích: Product
10. Lũy thừa: Power
11. Tích vô hướng (vector): Dot product
12. Tích của với số thực: Scalar product
13. Giao hoán: Commutative
14. Kết hợp: Associative
15. Phân phối: Distributive
16. Đơn vị: Unit
17. Mặt phẳng: Plane
18. Siêu phẳng: Hyperplane
19. Ma trận: Matrix
20. Dòng: Row
21. Cột: Column
22. Hệ số: Index
23. Đường chéo: Diagonal
24. Ma trận đơn vị: Identity Matrix
25. Ma trận tam giác: Triangle Matrix
26. Chuyển vị: Transpose
27. Nghịch đảo: Inverse
28. Khả nghịch: Invertible
29. Định thức: Determinant
30. Dạng bậc thang: Row-Echelon Form
31. Dạng bậc thang rút gọn: Reduced Row-Echelon Form
32. Phép khử: Elimination
33. Phép biến đổi dòng: Row operation
34. Sơ cấp: Elementary
35. Không gian vector: Vector space
36. Không gian con: Subspace
37. Không gian cột /hàng /hạch: Column /row /nullspace
38. Tổ hợp tuyến tính: Linear combination
39. Phụ thuộc /Độc lập tuyến tính: Linearly dependent /independent
40. Cơ sở: Basis
41. Tập sinh: Span
42. Chiều: Dimension
43. Trực giao: Orthogonal
44. Trực chuẩn: Orthonormal
45. Ánh xạ tuyến tính: Linear transformation
46. Tập ảnh: Image
47. Nhân: Kernel
48. Ma trận biểu diễn: Representation Matrix
49. Đồng dạng: Similar/Conjugate
50. Trị riêng: Eigenvalue
51. Vector riêng: Eigenvector
52. Chéo hóa: Diagonalize

1 Vector

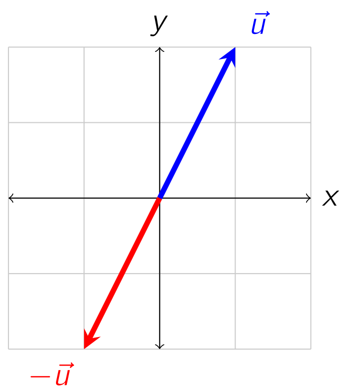
Vector và ma trận là hai đối tượng nghiên cứu chính của đại số tuyến tính. Để thuận tiện cho việc tính toán, người ta đã phát minh ra các phép toán cộng và nhân trên tập hợp các vector, cũng như tập hợp các ma trận, sao cho các tập hợp này đều đóng đối với các phép toán kể trên.

1.1 Vector

Định nghĩa 1.1 (Vector)**.** Trong $\mathbb{R}^n$, một vector là một bộ n số thực có thứ tự:

$$\vec{x} = (x_1, x_2, x_3, \cdots, x_n), \quad x_i \in \mathbb{R}, \quad \forall i \in \{1, 2, \cdots, n\}.$$

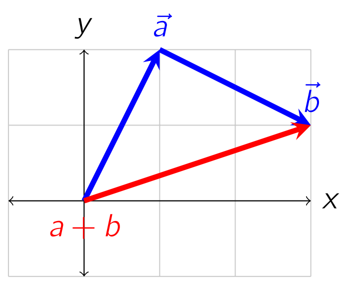
Ví dụ 1.2. $\vec{u} = (1, 2)$ là một vector trong $\mathbb{R}^2$:



Các cách hiểu tương đương nhau:

- ⋆ Mũi tên có hướng và độ dài (Vật Lý).
- ⋆ Một bộ có thứ tự các số (Tin Học và Dữ Liệu), dùng để lưu trữ thông tin của các con người/vật thể, v.v. Ví dụ như nhà ở:

  $(10, 40, 3) =$ Nhà cho thuê 10 triệu, diện tích $40\mathrm{m}^2$ và có 3 lầu.
- ⋆ Toán học: có thể cộng các vector và nhân vector với một hằng số.



Sneak peek:

- ⋆ $\mathbb{R}^n$ là một ví dụ của không gian vector.
- ⋆ Một mặt phẳng trong $\mathbb{R}^3$ đi qua gốc tọa độ O là một không gian con.

1.2 Phép cộng vector, phép nhân hằng số

Sau khi đã giới thiệu khái niệm vector, ta vẫn cần thêm những công cụ để làm việc với chúng một cách hiệu quả. Những công cụ này được gọi là các *phép toán*. Các phép toán cho phép ta tạo ra vector mới từ một số vector cho trước, có ý nghĩa trong việc tính toán và biểu diễn vector.

Ta bắt đầu với 2 phép toán cơ bản: **phép cộng** và **phép nhân hằng số**.

Định nghĩa 1.3 (Phép cộng vector)**.** Cho 2 vector $\vec{x} = (x_1, x_2, \ldots, x_n)$ và $\vec{y} = (y_1, y_2, \ldots, y_n)$ trong $\mathbb{R}^n$. **Tổng** của $\vec{x}$ và $\vec{y}$, ký hiệu là $\vec{x} + \vec{y}$ là một vector trong $\mathbb{R}^n$ có dạng

$$\vec{x} + \vec{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

Ví dụ 1.4. Trong $\mathbb{R}^3$, ta có:

- ⋆ $(2, -1, 3) + (5, 1, -3) = (7, 0, 0)$.
- ⋆ $\forall a, b, c \in \mathbb{R} : (a, 0, 0) + (0, b, 0) + (0, 0, c) = (a, b, c)$.

Định nghĩa 1.5 (Vector không)**.** Trong $\mathbb{R}^n$, **vector không**, ký hiệu là $\vec{0}$ là vector có các thành phần bằng 0: $\vec{0} = (0, 0, \ldots, 0)$.

Định nghĩa 1.6 (Phép nhân hằng số)**.** Cho vector $\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ và số thực $c \in R$. **Tích** của $\vec{x}$ với c, ký hiệu là $c\vec{x}$, là một vector trong $\mathbb{R}^n$ có dạng

$$c\vec{x} = (cx_1, cx_2, \ldots, cx_n).$$

Ví dụ 1.7. Trong $\mathbb{R}^3$, ta có:

- ⋆ $2(2, -1, 3) = (4, -2, 6)$.
- ⋆ $\forall a, b, c \in \mathbb{R} : a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = (a, b, c)$.

Các phép toán vector này có một số tính chất tương đồng với phép cộng và nhân số thực:

Tính chất 1.8 (Tính chất cơ bản của phép toán vector)**.** *Với mọi vector $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ và số thực $a, b \in \mathbb{R}$, ta có:*

- ⋆ $\vec{x} + \vec{y} = \vec{y} + \vec{x}$ *(Giao hoán).*
- ⋆ $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$ *(Kết hợp).*
- ⋆ $\vec{0} + \vec{x} = \vec{x} + \vec{0} = \vec{x}$.
- ⋆ $a(b\vec{x}) = b(a\vec{x}) = (ab)\vec{x}$ *(Kết hợp của phép nhân).*
- ⋆ $(a + b)\vec{x} = a\vec{x} + b\vec{x}$ *(Phân phối trái).*
- ⋆ $a(\vec{x} + \vec{y}) = a\vec{x} + a\vec{y}$ *(Phân phối phải).*
- ⋆ $0\vec{x} = \vec{0}$.

1.3 Phương, hướng và độ dài vector

Trong phần này, ta xem xét vector dưới góc độ hình học: vector $\vec{x}$ là đoạn thẳng có hướng nối gốc tọa độ 0 tới điểm có tọa độ trùng với các thành phần của $\vec{x}$. Dưới dạng hình học này, các vector sẽ có **phương**, **hướng** và **độ dài**.

Định nghĩa 1.9 (Phương và hướng)**.** Cho 2 vector $\vec{x}, \vec{y} \in \mathbb{R}^n$. Khi đó $\vec{x}$ **cùng phương** (**song song**) với $\vec{y}$ nếu tồn tại số thực c sao cho: $\vec{x} = c\vec{y}$.

- ⋆ Nếu $c > 0$, ta nói $\vec{x}$ **cùng hướng** với $\vec{y}$.
- ⋆ Nếu $c < 0$, ta nói $\vec{x}$ **ngược hướng** với $\vec{y}$.
- ⋆ Nếu $c = 0$, $\vec{x} = \vec{0}$, ta nói $\vec{0}$ **không có hướng**.

Định nghĩa 1.10 (Độ dài vector và tính chất cơ bản)**. Độ dài** của vector $\vec{x} = (x_1, x_2, \ldots, x_n)$, kí hiệu $\|x\|$, được tính bởi công thức:

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots x_n^2}.$$

Định nghĩa 1.11 (Vector đơn vị)**.** Khi $\|\vec{x}\| = 1$: $\vec{x}$ được gọi là **vector đơn vị**.

Để liên kết các khái niệm trên với các phép toán trên vector, ta có các tính chất:

Tính chất 1.12. *Với mọi vector $\vec{x}, \vec{y} \in \mathbb{R}^n$, $c \in \mathbb{R}$, ta có:*

- ⋆ *$\vec{0}$ luôn cùng phương với $\vec{x}$.*
- ⋆ *$c\vec{x}$ luôn cùng phương với $\vec{x}$.*
- ⋆ *$\|c\vec{x}\| = |c| \, \|\vec{x}\|$.*
- ⋆ *$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ (**Bất đẳng thức tam giác**)*

  Dấu '=' xảy ra khi và chỉ khi $\vec{y} = k\vec{x}$ với $k \geq 0$ hoặc $\vec{x} = \vec{0}$.

1.4 Tích vô hướng

Bên cạnh phép nhân số thực và phép cộng, **tích vô hướng** cũng là một phép toán nền tảng trên các vector. Sau đây ta sẽ thấy các khái niệm khoảng cách, độ dài và góc giữa các vector đều có thể biểu diễn qua tích vô hướng.

Định nghĩa 1.13 (Tích vô hướng)**.** Cho 2 vector $\vec{x} = (x_1, x_2, \ldots, x_n)$ và $\vec{y} = (y_1, y_2, \ldots, y_n)$ trong $\mathbb{R}^n$. Khi đó **tích vô hướng** của $\vec{x}$ và $\vec{y}$, kí hiệu là $\vec{x} \cdot \vec{y}$ (hoặc $\langle \vec{x}, \vec{y} \rangle$), là một giá trị được cho bởi công thức:

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Từ định nghĩa trên có thể dễ dàng chứng minh các tính chất sau:

Tính chất 1.14 (Tính chất cơ bản của tích vô hướng)**.** *Với mọi vector $\vec{x}, \vec{y} \in \mathbb{R}^n$ và số thực $k \in \mathbb{R}$, ta có:*

- ⋆ *$\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$ (Giao hoán),*

- $(k\vec{x}) \cdot \vec{y} = k(\vec{x} \cdot \vec{y})$ *(Nhân kết hợp một hằng số),*
- $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ *(Phân phối với phép cộng),*
- $\vec{x} \cdot \vec{x} = \|\vec{x}\|^2 \geq 0$.

Định lý 1.15 (Ý nghĩa hình học của tích vô hướng)**.** *Giả sử θ là góc giữa 2 vector $\vec{x}$ và $\vec{y}$ ($0 \leq \theta \leq \pi$), ta có:*

$$\vec{x} \cdot \vec{y} = \|\vec{x}\|\|\vec{y}\| \cos\theta.$$

Từ định nghĩa trên, ta có thêm các tính chất:

Tính chất 1.16. *Với mọi vector $\vec{x}, \vec{y} \in \mathbb{R}^n$, ta có:*

- *Nếu $\vec{x} \cdot \vec{y} = 0$ thì 2 vector này vuông góc với nhau $\left(\theta = \pm\frac{\pi}{2}\right)$, ký hiệu $\vec{x} \perp \vec{y}$,*
- $|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$ *(**Bất đẳng thức Cauchy–Bunyakovsky–Schwarz**).*

  Dấu '$=$' xảy ra khi và chỉ khi $\vec{x}$ và $\vec{y}$ cùng phương.

Định nghĩa 1.17 (Hình chiếu lên vector)**.** Cho $\vec{x}$ và $\vec{y}$ là 2 vector trong $\mathbb{R}^n$ $(\vec{y} \neq \vec{0})$. **Hình chiếu vuông góc** của $\vec{x}$ lên $\vec{y}$, ký hiệu là $\mathrm{proj}_{\vec{y}}(\vec{x})$, là một vector cùng phương với $\vec{y}$ có khoảng cách ngắn nhất đến $\vec{x}$:

$$\mathrm{proj}_{\vec{y}}(\vec{x}) = \vec{z} \in \mathbb{R}^n \iff \exists c \in \mathbb{R} : \vec{z} = c\vec{y}, \quad (\vec{x} - \vec{z}) \cdot \vec{y} = 0.$$

Tính chất 1.18. *Với mọi $\vec{x}, \vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\}$, ta có:*

1. $\vec{x} \cdot \vec{y} = \mathrm{proj}_{\vec{y}}(\vec{x}) \cdot \vec{y} = \vec{x} \cdot \mathrm{proj}_{\vec{x}}(\vec{y})$.
2. $\mathrm{proj}_{\vec{y}}(\vec{x})$ *là duy nhất với mỗi $\vec{x}$ và được tính bằng công thức* $\mathrm{proj}_{\vec{y}}(\vec{x}) = \dfrac{\vec{x} \cdot \vec{y}}{\|y\|^2}\vec{y}$.
3. $\mathrm{proj}_{\vec{x}}(\vec{0}) = \vec{0}$.

Sneak peek. Một cách tổng quát, chúng ta có thể định nghĩa được hình chiếu của một vector lên một *không gian con* (sẽ được học sau).

Bài tập

Bài tập 1.1. Hãy chứng minh toàn bộ định lý/tính chất phía trên.

Bài tập 1.2. Nếu đứng từ gốc tọa độ và bước một đoạn theo vector $\vec{x} = (x_1, x_2, \ldots, x_n,$ ta sẽ đứng ở điểm $(x_1, x_2, \ldots, x_n)$. Tiếp tục bước một đoạn theo vector $\vec{y} = (y_1, y_2, \ldots, y_n)$, ta sẽ đứng ở điểm $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$, cũng tương ứng với $\vec{x} + \vec{y}$. Như vậy việc cộng 2 vector có thể xem như việc "cộng" 2 di chuyển với nhau. Hãy mô tả ý nghĩa của các khái niệm phương, chiều và độ dài của vector dưới góc độ "di chuyển" như trên.

Bài tập 1.3. Tìm hình chiếu của $(1, 2, 3)$ lên $(2, 1, 3)$ trong $\mathbb{R}^3$.

Bài tập 1.4. Chứng minh tính chất 1.18.

2 Ma trận

2.1 Khái niệm

Trong phần này ta sẽ tập trung vào định nghĩa, các tính chất cơ bản của ma trận cũng như các phép toán trên ma trận. Việc ra đời của ma trận đã giúp khám phá các thuật toán giải hệ phương trình tuyến tính một cách tiện lợi (phần 4). Tuy nhiên, ý nghĩa sâu xa của việc phát minh và sử dụng ma trận trong toán học sẽ được làm rõ trong phần 7, Ánh xạ tuyến tính .

Định nghĩa 2.1 (Ma trận)**.** Cho $m, n \in \mathbb{N}$ và trường $\mathbb{F}$. **Ma trận** kích thước $m \times n$ trên $\mathbb{F}$ là một bảng hình chữ nhật gồm m dòng và n cột có dạng

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

trong đó $a_{ij} \in \mathbb{F}$ gọi là phần tử (hệ số) ở dòng i, cột j của ma trận trên.
Dạng rút gọn: $(a_{ij})_{m\times n}$ hoặc (a_{ij}) (nếu đã ngầm hiểu kích thước ma trận)

Nhận xét 2.2.

⋆ Trong bài giảng này, ta chủ yếu làm việc với ma trận trên trường số thực $\mathbb{R}$, và sẽ đề cập đến ma trận trên trường phức $\mathbb{C}$ ở phần Giá trị riêng, Vector riêng và Đa thức đặc trưng.

⋆ Ký hiệu $\mathrm{M}_{m\times n}(\mathbb{R})$ hoặc $\mathbb{R}^{m\times n}$ là tập hợp những ma trận kích thước $m \times n$ trên $\mathbb{R}$.

⋆ Với ma trận A kích thước $m \times n$ cho trước, ta ký hiệu A_{ij} thay cho phần tử dòng i, cột j.

Ví dụ 2.3. Dưới đây là các loại ma trận:

$$\underset{\mathbb{R}^{2\times 2}}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \qquad \underset{\mathbb{R}^{2\times 3}}{\begin{bmatrix} -19 & 4 & 10 \\ 17 & 1 & 2 \end{bmatrix}} \qquad \underset{\mathbb{C}^{3\times 4}}{\begin{bmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5-2i & 0 \end{bmatrix}}$$

Định nghĩa 2.4. Hai ma trận A và B được gọi là **bằng nhau** (ký hiệu $A = B$) nếu và chỉ nếu thỏa cả hai điều kiện dưới đây:

1. A và B cùng kích thước, tức là tồn tại $m, n \in \mathbb{N}$ sao cho $A, B \in \mathbb{R}^{m\times n}$.
2. Mỗi cặp phần tử tương ứng bằng nhau, tức là

$$A_{ij} = B_{ij},$$

với mọi $(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$.

2.2 Vector dòng, vector cột và một số loại ma trận đơn giản

Khi một ma trận có một kích thước $1 \times k$ hay $k \times 1$, nó sẽ có dạng cột hoặc hàng, và thực chất chỉ là một bộ k số. Điều này dẫn đến việc biểu diễn các vector dưới dạng ma trận.

Định nghĩa 2.5 (Vector cột, dòng)**.** Ma trận có kích thước $1 \times n$ được gọi là **vector dòng**. Ma trận có kích thước $m \times 1$ được gọi là **vector cột**.

Một ma trận kích thước $m \times n$ có thể xem như là ghép của m vector dòng (hoặc n vector cột):

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{bmatrix} = \begin{bmatrix} \mathrm{row}_1(A) \\ \mathrm{row}_2(A) \\ \vdots \\ \mathrm{row}_m(A) \end{bmatrix} = \begin{bmatrix} \mathrm{col}_1(A) & \mathrm{col}_2(A) & \dots & \mathrm{col}_n(A) \end{bmatrix},$$

trong đó $\mathrm{row}_i(A) = \begin{bmatrix} A_{i1} & A_{i2} & \dots A_{in} \end{bmatrix}$ và $\mathrm{col}_j(A) = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \dots \\ A_{mj} \end{bmatrix}$.

Lưu ý. Thông thường khi nói về vector trong $\mathbb{R}^n$ mà không nói gì thêm, ta mặc định đây là vector cột.

Sau đây là một số loại ma trận phổ biến và có nhiều ứng dụng.

Định nghĩa 2.6. Ma trận không (ký hiệu $0_{m\times n}$ hoặc 0) kích thước $m \times n$ là ma trận có tất cả hệ số đều bằng 0.

Định nghĩa 2.7. Ma trận vuông là ma trận có số dòng bằng số cột. Ta nói ma trận vuông cấp n ý muốn chỉ đến ma trận kích cỡ $n \times n$.

Trong ma trận vuông A cấp n, các hệ số A_{ii} hợp thành **đường chéo** của ma trận A.

$$\mathrm{diag}(A) = \begin{bmatrix} A_{11} & \dots & A_{ii} & \dots & A_{nn} \end{bmatrix}.$$

Hệ số A_{ij} với $i < j$ gọi là hệ số nằm phía trên đường chéo; hệ số A_{ij} với $j < i$ gọi là hệ số nằm phía dưới đường chéo.

Định nghĩa 2.8. Ma trận đường chéo là ma trận vuông có các hệ số nằm ngoài đường chéo bằng 0.

$$\begin{bmatrix} A_{11} & 0 & \dots & 0 \\ 0 & A_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_{nn} \end{bmatrix}.$$

Khái niệm tiếp theo, **ma trận đơn vị**, đóng vai trò chủ chốt trong rất nhiều phép tính toán trên ma trận. Ý nghĩa của nó sẽ được làm sáng tỏ trong Phần 3.

Định nghĩa 2.9. Ma trận đơn vị (ký hiệu I hoặc I_n) là ma trận đường chéo có các hệ số trên đường chéo chính là 1.

Trong toán học, Kronecker delta được định nghĩa là hàm hai biến (thường xét trên tập hợp các số nguyên không âm $\mathbb{Z}^2_{\geq 0}$):

$$\delta(x, y) = \begin{cases} 1 \text{ nếu } x = y, \\ 0 \text{ nếu } x \neq y. \end{cases}$$

Lúc đó, ta có $(I_n)_{ij} = \delta_{ij}, \quad \forall i, j \in \{1, \dots, n\}$ và

$$I_n = (\delta(i,j))_{n \times n} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Dạng ma trận tiếp theo cũng có vai trò quan trọng trong tính toán ma trận, ứng dụng phổ biến nhất của nó là trong việc giải hệ phương trình tuyến tính thông qua thuật toán **Gauss Jordan**, sẽ được học trong Phần 4.4.

Ma trận tam giác trên (**dưới**) là ma trận vuông mà tất cả hệ số nằm phía **dưới** (**trên**) đường chéo đều bằng 0.

Ma trận tam giác trên: $\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$

Ma trận tam giác dưới: $\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$

Một cách toán học, ta có:

$$A = (A_{ij})_{n \times n} \text{ là ma trận tam giác } \begin{cases} \text{trên khi và khỉ khi } A_{ij} = 0, & \forall j < i, \\ \text{dưới khi và khỉ khi } A_{ij} = 0, & \forall i < j. \end{cases}$$

Bài tập

Bài tập 2.1. Cho $A = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \end{bmatrix}$; $B = \begin{bmatrix} 4 & 5 & 6 \\ 9 & 8 & 7 \end{bmatrix}$; $C = \begin{bmatrix} 10 & 11 \end{bmatrix}$; $D = \begin{bmatrix} 12 \\ 13 \end{bmatrix}$. Hãy xác định các ma trận sau:

$$\left[\begin{array}{c|c} \begin{array}{c|c} \multicolumn{2}{c}{B} \\ \hline D & \begin{array}{c} C \\ \hline C \end{array} \end{array} & A \end{array}\right] \quad \text{và} \quad \left[\begin{array}{c|c|c} A & \begin{array}{c} B \\ \hline B \end{array} & \begin{array}{c|c} \multicolumn{2}{c}{C} \\ \hline D & D \\ \hline \multicolumn{2}{c}{C} \end{array} \end{array}\right].$$

3 Các phép toán trên ma trận

Sau khi đã định nghĩa và xây dựng một số dạng ma trận, chúng ta vẫn chưa thể sử dụng chúng cho đến khi đã tìm được cách tương tác chúng với nhau (tương tự như Hóa học - sau khi tìm ra một chất mới ta phải tìm hiểu cách nó phản ứng với chất khác). Sự tương tác này chính là các **phép toán** trên ma trận. Ngoài các phép cộng và nhân được xây dựng mô phỏng theo số thực, sẽ có một số phép toán đặc thù cho ma trận.

3.1 Phép chuyển vị

Phép chuyển vị là một phép toán quan trọng và đặc thù của ma trận. Chuyển vị cho phép ta đổi thứ tự các trục ngang - dọc của ma trận, hay hoán đổi cột và dòng.

Định nghĩa 3.1 (Phép chuyển vị)**.** Cho $A \in \mathbb{R}^{m\times n}$, **ma trận chuyển vị** của A là một ma trận kích thước $n \times m$ (ký hiệu A^T) thỏa mãn:

$$\left(A^T\right)_{ji} = A_{ij},$$

với mọi $(i, j) \in \{1, \ldots, m\} \times \{1, \ldots, n\}$.

Ví dụ 3.2.

$$A \qquad\qquad\qquad A^T$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad\qquad \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Tính chất 3.3. *Cho $A, B \in \mathbb{R}^{m\times n}$, khi đó:*

- ⋆ $(\mathrm{row}_i(A))^T = \mathrm{col}_i(A^T)$,
  $(\mathrm{col}_j(A))^T = \mathrm{row}_j(A^T)$,
- ⋆ $(A^T)^T = A$,
- ⋆ $(A + B)^T = A^T + B^T$,
- ⋆ *A được gọi là **đối xứng** nếu $A = A^T$.*
  Khi đó điều kiện cần để hai ma trận này bằng nhau là chúng cùng cỡ, tức là $m = n$, nghĩa là điều kiện cần để ma trận đối xứng chính là chúng phải là ma trận vuông.

3.2 Phép nhân vô hướng

Phép nhân cơ bản của ma trận với số thực cũng tương tự như với vector.

Định nghĩa 3.4 (Phép nhân vô hướng)**.** Cho ma trận $A \in \mathbb{R}^{m\times n}$ và $\alpha \in \mathbb{R}$. Ta định nghĩa kết quả của **phép nhân vô hướng** giữa α và A là ma trận kích thước $m \times n$ (ký hiệu αA) có được bằng cách nhân mỗi hệ số của A với α.

$$(\alpha A)_{ij} = \alpha . A_{ij},$$

với mọi $(i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}$.
Ký hiệu: $-A = (-1)A$.

Ví dụ 3.5.

$$3\begin{bmatrix}1-i & 0\\ 0 & 1+i\end{bmatrix} = \begin{bmatrix}3-3i & 0\\ 0 & 3+3i\end{bmatrix},$$
$$2\begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{bmatrix} = \begin{bmatrix}2 & 4\\ 6 & 8\\ 10 & 12\end{bmatrix}.$$

Tính chất 3.6. *Cho $A \in \mathbb{R}^{m\times n}$ và $\alpha,\beta \in \mathbb{R}$, khi đó:*

- ⋆ $\alpha(\beta A) = (\alpha\beta)A = (\beta\alpha)A = \beta(\alpha A)$,
- ⋆ $(\alpha A)^T = \alpha A^T$,
- ⋆ $0A = 0_{m\times n}$,
- ⋆ $1A = A$.

3.3 Phép cộng

Điều kiện cần và đủ để ta có thể cộng hai ma trận là chúng phải cùng kích thước.

Định nghĩa 3.7 (Phép cộng ma trận)**.** Cho $A, B \in \mathbb{R}^{m\times n}$. Ta định nghĩa **tổng** của hai ma trận (ký hiệu $A + B$) là ma trận cùng loại với các hệ số là tổng (tại vị trí tương ứng) của A và B:

$$(A+B)_{ij} = A_{ij} + B_{ij},$$

với mọi $(i,j) \in \{1,\dots,m\} \times \{1,\dots,n\}$.

Ví dụ 3.8.

$$\begin{bmatrix}1 & 5 & 8\\ -2 & 3 & 2\end{bmatrix} + \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \quad : \quad \text{Không xác định},$$
$$\begin{bmatrix}1 & 5 & 8\\ -2 & 3 & 2\end{bmatrix} + \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 2\end{bmatrix} = \begin{bmatrix}2 & 5 & 8\\ -2 & 4 & 4\end{bmatrix}.$$

Trên trường $\mathbb{R}$, ta có những tính chất sau:

Tính chất 3.9. *Cho $A, B, C \in \mathbb{R}^{m\times n}$ và $\alpha,\beta \in \mathbb{R}$. Khi đó:*

- ⋆ $A + B = B + A$,
- ⋆ $(A+B)+C = A+(B+C)$,
- ⋆ $0_{m\times n} + A = A + 0_{m\times n} = A$,
- ⋆ $A + (-A) = 0_{m\times n} = (-A) + A$,
- ⋆ $(A+B)^T = A^T + B^T$,
- ⋆ $\alpha(A+B) = \alpha A + \alpha B$,
- ⋆ $(\alpha+\beta)A = \alpha A + \beta A$,

- $\star$ $(-\alpha)A = \alpha(-A) = -(\alpha A)$.

3.4 Phép nhân ma trận

Cho hai ma trận A và B, điều kiện cần và đủ để ta có thể nhân hai ma trận này (theo đúng thứ tự, ký hiệu là AB) là số cột của A bằng với số dòng của B. Khi đó ta định nghĩa phép nhân này như sau:

Định nghĩa 3.10 (Phép nhân ma trận)**.** Cho $A \in \mathbb{R}^{m\times n}$ và $B \in \mathbb{R}^{n\times p}$. Khi đó **phép nhân ma trận** A và B (ký hiệu AB) cho ra một ma trận loại $m \times p$ (tức $AB \in \mathbb{R}^{m\times p}$) trong đó $(AB)_{ij}$ được xác định bởi tích vô hướng của $(\mathrm{row}_i(A))^T$ và $\mathrm{col}_j(B)$:

$$\begin{aligned}(AB)_{ij} &= \langle (\mathrm{row}_i(A))^T, \mathrm{col}_j(B)\rangle \\ &= \sum_{k=1}^{n} A_{ik}.B_{kj}.\end{aligned}$$

Hình 1 minh họa một cách trực quan phép nhân ma trận. Có thể thấy, việc viết ma trận B lên phía phải trên của ma trận A và tích AB giữa A và B giúp việc thực hiện phép nhân dễ dàng hơn.

Ví dụ 3.11. Cho $A = \begin{bmatrix}1 & 5 & 6\end{bmatrix}$ và $B = \begin{bmatrix}2\\4\\7\end{bmatrix}$. Khi đó:

$$\begin{aligned}BA &= \begin{bmatrix}2 & 10 & 12\\ 4 & 20 & 24\\ 7 & 35 & 42\end{bmatrix},\\ AB &= \begin{bmatrix}1\times 2 + 5\times 4 + 6\times 7\end{bmatrix} = \begin{bmatrix}64\end{bmatrix}.\end{aligned}$$

Nhận xét 3.12. Với hai ma trận A, B bất kỳ, có khả năng ta nhân được A và B nhưng không thể nhân được giữa B và A. Phép nhân cả 2 chiều chỉ có thể thực hiện được khi B có cùng kích thước với A^T.

Điều này có nghĩa phép nhân ma trận nói chung **không** giao hoán. Lưu ý rằng tồn tại vô số cặp ma trận A, B sao cho $AB = BA$. Cặp ma trận nào thỏa mãn điều này được gọi là *giao hoán* với nhau.

Nhận xét 3.13. Với 2 vector cột u, v thuộc $\mathbb{R}^n$, tích $u^T v$ là một ma trận 1×1 và có thể xem như một số thực bằng tích vô hướng $u \cdot v$.

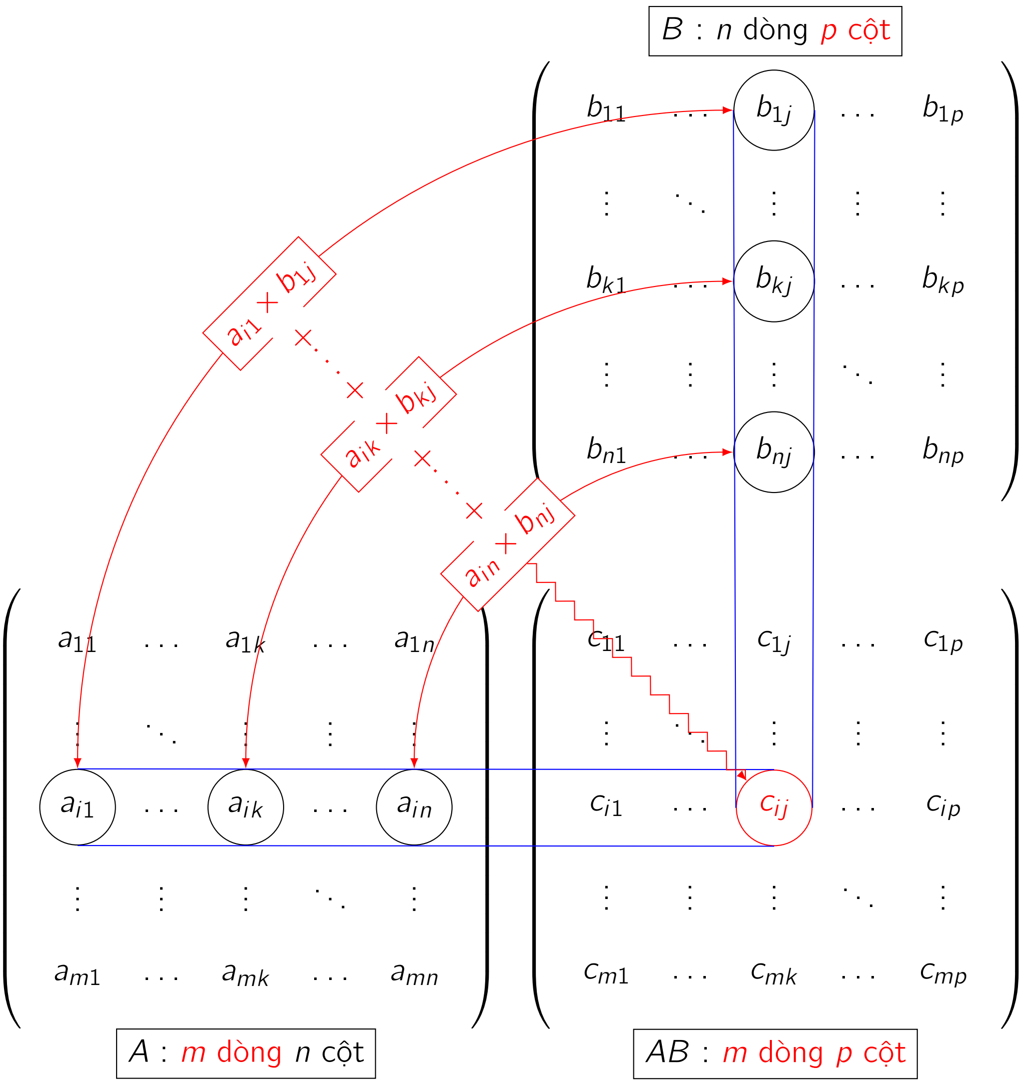
Bổ đề 3.14. *(Quan trọng - Liên hệ giữa tích vô hướng và phép nhân ma trận)*

Cho $A \in \mathbb{R}^{n\times n}$ và $x, y \in \mathbb{R}^n$. Ta có

1. $x \cdot y = x^T y$,
2. $Ax \cdot y = x \cdot A^T y$.

Tính chất 3.15. *Cho $A \in \mathbb{R}^{m\times n}$; $B, B_1, B_2 \in \mathbb{R}^{n\times p}$; $C \in \mathbb{R}^{p\times q}$; $D_1, D_2 \in \mathbb{R}^{p\times m}$:*

- $\star$ $I_m A = A = A I_n$,



Hình 1: Hình minh họa phép nhân ma trận

- ⋆ $0_{p\times m}A = 0$ và $A0_{n\times q} = 0_{m\times q}$,
- ⋆ $(AB)^T = B^T A^T$,
- ⋆ *Phép nhân ma trận có tính kết hợp:* $(AB)C = A(BC)$,
- ⋆ *Phép nhân ma trận có tính phân phối đối với phép cộng:*

$$A(B_1 + B_2) = AB_1 + AB_2,$$
$$(D_1 + D_2)A = D_1A + D_2A.$$

Bài tập

Bài tập 3.1. Hãy chứng minh toàn bộ định lý/tính chất phía trên.

Bài tập 3.2. Tính các tích sau:

a) $\begin{bmatrix} 5 & 1 & 3 & 2 \\ 4 & 5 & 1 & -7 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$

b) $\begin{bmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{bmatrix} \begin{bmatrix} i & i & i \\ 0 & i & i \\ 0 & 0 & i \end{bmatrix}.$

Bài tập 3.3. Hãy tổng quát A^n, trong đó:

a) $A = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}.$

b) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$

c) $A = \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix}.$

d) $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$

Bài tập 3.4. Cho $A \in \mathbb{R}^{n\times n}$. Giả sử $A^{75} = A^{89} = I_n$. Chứng minh rằng $A = I^n$. Từ đó hãy tổng quát hóa bài toán trên.

Bài tập 3.5. Cho $A, B \in \mathbb{R}^{n\times n}$. Liệu $A^2 - B^2 = (A - B)(A + B)$? Nếu câu trả lời là không, hãy tìm điều kiện để đẳng thức trên xảy ra.

Bài tập 3.6. Cho $A, B \in \mathbb{R}^{n\times n}$ là hai ma trận đối xứng. Chứng minh rằng tích AB đối xứng khi và chỉ khi A, B giao hoán.

4 Hệ phương trình tuyến tính và phương pháp giải

Trong phần này, chúng ta phát biểu bài toán hệ phương trình tuyến tính và mô tả thuật toán giải bằng phép khử Gauss-Jordan. Ban đầu bộ môn Đại Số Tuyến Tính ra đời là để phục vụ cho việc giải các hệ phương trình này (trước khi nó được mở rộng thành công cụ trong Hình học Đại số và các ngành khác).

4.1 Hệ phương trình tuyến tính

Một phương trình tuyến tính n biến trên $\mathbb{R}$ có dạng $\sum_{i=1}^{n}\alpha_i x_i = \alpha_1 x_1 + \ldots + \alpha_n x_n = \alpha$ trong đó các $\alpha_i \in \mathbb{R}$ gọi là hệ số ứng với biến x_i. Từ đó ta có định nghĩa sau:

Định nghĩa 4.1. Một hệ m phương trình, n ẩn trên $\mathbb{R}$ có dạng:

$$\begin{cases} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & b_m \end{cases},$$

là một hệ phương trình tuyến tính.
Trong đó các $\alpha_{ij} \in \mathbb{R}$ là các hệ số tương ứng ở hàng i của biến x_j.
Nếu có một bộ $(x_1, \ldots x_n) \in \mathbb{R}^n$ thỏa mãn hệ phương trình trên thì ta gọi chúng là **nghiệm** của hệ phương trình này.

Do cách định nghĩa phép nhân của ma trận, ta có thể viết lại hệ phương trình trên thành:

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Hay đơn giản là $Ax = b$.

Ví dụ 4.2. Giải hệ phương trình:

$$\begin{cases} 2x & + & 3y & = & 19 \\ x & + & y & = & 7 \end{cases}.$$

Để giải một hệ phương trình tuyến tính, ta phải thực hiện các phép biến đổi (không làm ảnh hưởng đến nghiệm của hệ) như đổi hai dòng, nhân một dòng với một vô hướng (khác 0) hoặc là cộng trừ dòng này cho dòng kia.
Đó cũng chính là cơ sở cho việc nghiên cứu cách giải hệ phương trình tuyến tính.

4.2 Phép biến đổi sơ cấp (trên dòng)

Cho ma trận $A \in \mathbb{R}^{m\times n}$:

Định nghĩa 4.3. Ta nói một phép biến đổi được gọi là sơ cấp (trên dòng) trên A khi và chỉ khi nó thuộc một trong ba loại biến đổi dưới đây:

1. Tráo đổi hai dòng i và j:

$$A = \begin{bmatrix} \vdots \\ \mathrm{row}_i(A) \\ \vdots \\ \mathrm{row}_j(A) \\ \vdots \end{bmatrix} \xrightarrow{\mathrm{row}_i(A) \longleftrightarrow \mathrm{row}_j(A)} \begin{bmatrix} \vdots \\ \mathrm{row}_j(A) \\ \vdots \\ \mathrm{row}_i(A) \\ \vdots \end{bmatrix},$$

2. Nhân dòng i với vô hướng $c \in \mathbb{R} \setminus \{0\}$:

$$A = \begin{bmatrix} \vdots \\ \mathrm{row}_i(A) \\ \vdots \end{bmatrix} \xrightarrow{\mathrm{row}_i(A) \longrightarrow c.\mathrm{row}_i(A)} \begin{bmatrix} \vdots \\ c.\mathrm{row}_i(A) \\ \vdots \end{bmatrix},$$

3. Cộng dòng i với $c \in \mathbb{R}$ lần dòng j:

$$A = \begin{bmatrix} \vdots \\ \mathrm{row}_i(A) \\ \vdots \\ \mathrm{row}_j(A) \\ \vdots \end{bmatrix} \xrightarrow{\mathrm{row}_i(A) \longrightarrow \mathrm{row}_i(A) + c.\mathrm{row}_j(A)} \begin{bmatrix} \vdots \\ \mathrm{row}_i(A) + c.\mathrm{row}_j(A) \\ \vdots \\ \mathrm{row}_j(A) \\ \vdots \end{bmatrix}.$$

Định nghĩa 4.4. Cho $A, B \in \mathbb{R}^{m \times n}$. Ta nói A **tương đương dòng** với B (ký hiệu $A \sim B$) nếu qua A ta có thể thực hiện một số phép sơ cấp trên dòng thì được B.

4.3 Ma trận bậc thang và Ma trận bậc thang rút gọn

Định nghĩa 4.5. Một ma trận được gọi là **ma trận dạng bậc thang** nếu và chỉ nếu:

1. Các dòng khác 0 luôn nằm trên các dòng 0,
2. Hệ số cao nhất (hệ số khác 0 đầu tiên tính từ trái qua) của bất kì dòng khác 0 nào luôn nằm bên phải so với hệ số cao nhất của dòng phía trên.

Ví dụ 4.6.

$$\begin{bmatrix} 1 & -1 & 7 & 0 & 2 \\ 0 & 0 & 6 & 8 & 1 \\ 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Các ma trận bậc thang

$$\begin{bmatrix} 1 & -1 & 7 & 0 & 2 \\ 0 & 0 & 6 & 8 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 2 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Các ma trận không có dạng bậc thang

Định nghĩa 4.7. Một ma trận được gọi là **ma trận dạng bậc thang rút gọn** nếu và chỉ nếu:

1. Ma trận này là ma trận bậc thang,
2. Các hệ số cao nhất trên các dòng khác 0 đều là 1,
3. Đối với mỗi cột mà nó chứa hệ số cao nhất (là 1) của dòng khác 0 nào đó, thì số 1 đó là giá trị khác 0 duy nhất trên cột đó.

Các ma trận dạng bậc thang quan trọng vì chúng làm cho việc giải hệ phương trình trở nên đơn giản hơn rất nhiều. Ta xem xét ví dụ sau:

Ví dụ 4.8. Cho hệ phương trình tuyến tính

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

Hệ trên tuy có đến 4 phương trình, 4 ẩn nhưng cách giải thực ra lại cực kỳ đơn giản! Ta có thể bắt đầu từ dưới lên trên như sau:

1. Dòng thứ 4 cho ta: $x_4 = 1$.
2. Dòng thứ 3 cho ta: $x_3 = (3 - 6x_4)/3 = -1$.
3. Dòng thứ 2 cho ta: $x_2 = (2 - 6x_4 - 4x_3)/2 = 0$.
4. Dòng thứ 1 cho ta: $x_1 = 1 - 4x_4 - 3x_3 - 2x_2 = 0$.

Như vậy nếu vế trái là một ma trận dạng bậc thang, việc giải hệ phương trình có thể được thực hiện một cách dễ dàng từ dưới lên trên. Do đó, **hệ phương trình bất kỳ sẽ được giải nếu ta có thể biến đổi nó về dạng bậc thang bằng các biến đổi sơ cấp dòng**. Đó chính là mục tiêu của **thuật toán khử Gauss - Jordan** sau đây.

4.4 Phép khử Gauss - Jordan

Phép khử Gauss - Jordan là tên gọi chung cho các phép biến đổi sơ cấp dòng được thực hiện trong thuật toán Gauss - Jordan dùng để giải hệ phương trình tuyến tính. Các phép biến đổi này được chọn ở từng bước để đảm bảo hệ phương trình được đưa về dạng *bậc thang rút gọn* một cách nhanh nhất.

Chúng ta bắt đầu với định lý sau:

Định lý 4.9. *Mọi ma trận trong $\mathbb{R}^{m\times n}$ đều có thể đưa về ma trận dạng bậc thang rút gọn thông qua một số phép biến đổi sơ cấp. Nói cách khác, với mọi ma trận trong $\mathbb{R}^{m\times n}$ luôn tồn tại một ma trận dạng bậc thang tương đương dòng với ma trận đó.*

Ta có thể dựa vào **thuật toán Gauss - Jordan** để chứng minh định lý trên:
Giả sử $A = (a_{ij})_{m\times n} \in \mathbb{R}^{m\times n}$:

Thuật toán khử Gauss - Jordan

Bắt đầu với $i := 1, j := 1$.

- ⋆ **Bước 0.** Nếu $i > m$ hoặc $j > n$ thì thuật toán kết thúc, không thì chuyển qua bước 1.
- ⋆ **Bước 1.** Nếu $a_{ij} = 0$ thì tráo đổi dòng i với một dòng bên dưới để thu được $a_{ij} \neq 0$, sau đó qua bước 2.
  Trường hợp mà $a_{ik} = 0$ với mọi $j \leq k$, tăng j lên 1 đơn vị, tức là $j := j + 1$ sau đó quay trở lại bước 0.
- ⋆ **Bước 2.** Đưa hệ số cao nhất trên hàng i về 1:
  $$\mathrm{row}_i := \frac{1}{a_{ij}}\mathrm{row}_i.$$
  Sau đó chuyển sang bước 3.
- ⋆ **Bước 3.** Đưa tất cả các giá trị trên cột j (ngoại trừ a_{ij}) về 0:
  $$\mathrm{row}_k := \mathrm{row}_k - a_{kj}\mathrm{row}_i, \qquad \forall k \neq i.$$
  Chuyển sang bước 4.
- ⋆ **Bước 4.** Tăng i và j lên 1 đơn vị:
  $$\begin{aligned} i &:= i + 1, \\ j &:= j + 1. \end{aligned}$$
  Quay về bước 0.

Rõ ràng, ta thấy rằng mỗi ma trận có thể có nhiều ma trận dạng bậc thang tương đương dòng với nó. Tuy nhiên, đối với ma trận dạng bậc thang rút gọn thì đó là duy nhất:

Định lý 4.10. *Cho $A \in \mathbb{R}^{m\times n}$. Khi đó tồn tại duy nhất một ma trận dạng bậc thang rút gọn (ký hiệu là $\mathrm{rref}(A)$) sao cho $A \sim \mathrm{rref}(A)$. Nghĩa là ta có thể xem $\mathrm{rref} : \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{m\times n}$ là một ánh xạ xác định biến mỗi ma trận thành một ma trận dạng bậc thang rút gọn tương đương dòng với nó.*
Số lượng dòng khác 0 của $\mathrm{rref}(A)$ chính là **hạng**[2] *của ma trận A (ký hiệu là $\mathrm{rank}(A)$).*

Nhận xét 4.11. Dạng bậc thang dịch ra tiếng Anh là **Row-echelon Form**, trong khi Rút gọn là **reduced**. Dạng bậc thang rút gọn dịch toàn bộ thành **Reduced Row-echelon Form**, được viết tắt là rref.

2 Sẽ được đề cập rõ hơn ở phần sau

Ví dụ 4.12. Xét ma trận sau:

$$A = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}.$$

Ta có:

$$\begin{aligned}
A = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} & \xrightarrow{r_1 \longleftrightarrow r_3} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \\
& \xrightarrow{r_2 := r_2 - r_1} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \\
& \xrightarrow{r_2 := \frac{1}{2} r_2} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \\
& \xrightarrow{r_3 := r_3 - 3r_2} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \\
& \xrightarrow{r_1 := r_1 + 9r_2} \begin{bmatrix} 3 & 0 & -6 & 9 & 15 & -12 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \\
& \xrightarrow{r_1 := \frac{1}{3} r_1} \begin{bmatrix} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \\
& \xrightarrow{r_1 := r_1 - 5r_3} \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \\
& \xrightarrow{r_2 := r_2 - r_3} \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}.
\end{aligned}$$

Qua đó, $\mathrm{rref}(A) = \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$ và $\mathrm{rank}(A) = 3$.

Thực chất khi ta tác động các phép biến đổi sơ cấp (dòng) lên hệ phương trình cũng chính là tác động tương ứng trên ma trận A và b tương ứng. Vì thế ta có thể ghép nối ma trận b vào ma trận A: gọi là ma trận bổ sung

$$[A|b] = \left[\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right],$$

và thực hiện biến đổi trên ma trận này thay cho hệ phương trình gốc.

Nhận xét 4.13. Để giải được hệ phương trình trên, ta cần đưa chúng về dạng đơn giản nhất, và đó chính là dạng ma trận bậc thang rút gọn.

Ví dụ 4.14. Giải hệ phương trình tuyến tính sau trên $\mathbb{R}$:

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 7 \\ x_2 + 2x_1 + 2x_3 + 3x_4 = 6 \\ 3x_1 + 2x_2 + 2x_4 + x_3 = 7 \\ 4x_1 + 3x_2 + 2x_3 + x_4 = 18 \end{cases}.$$

Viết lại hệ phương trình tuyến tính theo đúng thứ tự:

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 7 \\ 2x_1 + x_2 + 2x_3 + 3x_4 = 6 \\ 3x_1 + 2x_2 + x_3 + 2x_4 = 7 \\ 4x_1 + 3x_2 + 2x_3 + x_4 = 18 \end{cases}.$$

Xét ma trận bổ sung $[A|b]$:

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 7 \\ 2 & 1 & 2 & 3 & 6 \\ 3 & 2 & 1 & 2 & 7 \\ 4 & 3 & 2 & 1 & 18 \end{array}\right] \xrightarrow[r_4-4r_1]{\substack{r_2-2r_1 \\ r_3-3r_1}} \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 7 \\ 0 & -3 & -4 & -5 & -8 \\ 0 & -4 & -8 & -10 & -14 \\ 0 & -5 & -10 & -15 & -10 \end{array}\right],$$

$$\xrightarrow[\frac{-1}{5}r_4]{\substack{r_2-r_3 \\ -\frac{1}{2}r_3}} \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 7 \\ 0 & 1 & 4 & 5 & 6 \\ 0 & 2 & 4 & 5 & 7 \\ 0 & 1 & 2 & 3 & 2 \end{array}\right],$$

$$\xrightarrow[r_4-r2]{\substack{r_1-2r_2 \\ r_3-2r_2}} \left[\begin{array}{cccc|c} 1 & 0 & -5 & -6 & -5 \\ 0 & 1 & 4 & 5 & 6 \\ 0 & 0 & -4 & -5 & -5 \\ 0 & 0 & -2 & -2 & -4 \end{array}\right],$$

$$\xrightarrow[\frac{-1}{2}r_4]{r_3-2r_4} \left[\begin{array}{cccc|c} 1 & 0 & -5 & -6 & -5 \\ 0 & 1 & 4 & 5 & 6 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{array}\right],$$

$$\xrightarrow[r_3\longleftrightarrow r_4]{r_4+r_3} \left[\begin{array}{cccc|c} 1 & 0 & -5 & -6 & -5 \\ 0 & 1 & 4 & 5 & 6 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & -1 & 3 \end{array}\right],$$

$$\xrightarrow[-r_4]{\substack{r_1+5r_3 \\ r_2-4r_3}} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -6 & 20 \\ 0 & 1 & 0 & 5 & -14 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right],$$

$$\xrightarrow[r_2-5r_4]{r_1+6r_4} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right].$$

Phương trình đã cho có duy nhất một nghiệm và hệ phương trình đã cho tương đương với:

$$\begin{cases} x_1 & & & & = & 2 \\ & x_2 & & & = & 1 \\ & & x_3 & & = & 5 \\ & & & x_4 & = & -3 \end{cases}.$$

Bài tập

Bài tập 4.1. Hãy chứng minh toàn bộ định lý/tính chất phía trên.

Bài tập 4.2 (Quan trọng). Cho $A \in \mathbb{R}^{m\times n}$ và φ là một phép biến đổi sơ cấp (dòng). Chứng minh rằng $\varphi(A) = \varphi(I_m)A$.
Ở đây, $\varphi(I_m)$ được gọi là một ma trận sơ cấp cấp n.
Hãy liệt kê các dạng ma trận sơ cấp ứng với mỗi phép biến đổi sơ cấp dòng.

Bài tập 4.3. Chứng minh với mọi ma trận A, tồn tại các ma trận sơ cấp $E_1, E_2, \ldots, E_k$ sao cho:

$$\operatorname{rref}(A) = E_1E_2\cdots E_kA.$$

Bài tập 4.4. Tìm dạng bậc thang rút gọn cũng như hạng của các ma trận sau:

a) $\begin{bmatrix} 3 & 5 & 7 \\ 2 & 4 & 6 \\ 1 & 7 & 11 \end{bmatrix}$.

c) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$.

e) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$.

b) $\begin{bmatrix} 3 & 4 & 6 & -2 \\ 1 & 7 & -1 & 2 \\ 5 & 8 & 9 & 1 \\ 4 & 2 & -3 & 0 \end{bmatrix}$.

d) $\begin{bmatrix} 1 & -1 & 5 & -1 \\ 1 & 1 & -2 & 3 \\ 3 & -1 & 8 & 1 \\ 1 & 3 & -9 & 7 \end{bmatrix}$.

f) $\begin{bmatrix} 1 & 3 & -2 & -1 \\ 2 & 5 & -2 & 1 \\ 1 & 1 & 6 & 13 \\ -2 & -6 & 8 & 10 \end{bmatrix}$.

Từ đó dựa vào bài tập 4.2 hãy tìm các ma trận P sao cho $A = P.\mathrm{rref}(A)$.

Bài tập 4.5. Giải các hệ phương trình sau:

(a) $$\begin{cases} x_1 - 2x_2 + 2x_3 - 3x_4 = 1 \\ x_1 + 4x_2 - x_3 - 2x_4 = -2 \\ x_1 - 4x_2 + 3x_3 - 2x_4 = -2 \\ x_1 - 8x_2 + 5x_3 - 2x_4 = -2 \end{cases},$$

(b) $$\begin{cases} x_1 + 2x_2 + 3x_3 - 2x_4 = 1 \\ 2x_1 - x_2 - 2x_3 - 3x_4 = 2 \\ 3x_1 + 2x_2 - x_3 + 2x_4 = -5 \\ 2x_1 - 3x_2 + 2x_3 + x_4 = 11 \end{cases},$$

(c) $$\begin{cases} 2x_1 - 2x_2 + x_3 - x_4 + x_5 = 1 \\ x_1 + 2x_2 - x_3 + x_4 - 2x_5 = 1 \\ 4x_1 - 10x_2 + 5x_3 - 5x_4 + 7x_5 = 1 \\ 2x_1 - 14x_2 + 7x_3 - 7x_4 + 11x_5 = -1 \end{cases}.$$

Bài tập 4.6. Giả sử B là ma trận có được sau khi biến đổi một số phép sơ cấp (dòng) trên A. Chứng minh rằng ta cũng có thể biến đổi một số phép sơ cấp (dòng) trên B để được A. Từ đó hãy chứng minh $\sim$ là một quan hệ tương đương

Bài tập 4.7. Hãy chứng minh **Thuật toán khử Gauss - Jordan** có thể đưa mọi ma trận về dạng bậc thang rút gọn.

Bài tập 4.8. Chứng minh rằng mỗi ma trận có duy nhất một ma trận dạng bậc thang rút gọn.

Bài tập 4.9. Chứng minh rằng nếu $A \sim B$ thì $\mathrm{rref}(A) = \mathrm{rref}(B)$. Từ đó suy ra rằng nếu $A \sim B$ thì $\mathrm{rank}(A) = \mathrm{rank}(B)$.

Bài tập 4.10. Hãy kiểm tra các tính chất trên khi biến đổi sơ cấp trên cột?

5 Ma trận vuông

Trọng tâm của Đại Số Tuyến Tính sơ khởi chính là các ma trận, trong đó đối tượng được nghiên cứu nhiều nhất là các **ma trận vuông**, tức các ma trận thuộc $\mathbb{R}^{n\times n}$, như đã định nghĩa trong phần 2.2.

Để hiểu lý do ma trận vuông được coi trọng, ta vẫn bắt đầu từ các hệ phương trình tuyến tính. Xét hệ m phương trình gồm n ẩn $Ax = b$, với ma trận bổ sung:

$$[A|b] = \left[\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right].$$

Sử dụng thuật toán Gauss - Jordan đưa $[A|b]$ về dạng bậc thang rút gọn, xảy ra 3 trường hợp:

(1) $m > n$. Như vậy sẽ có ít nhất $m - n$ dòng toàn số 0 trong $\mathrm{rref}(A)$, tương ứng với ít nhất $m - n$ phương trình không có ẩn trong dạng bậc thang rút gọn của hệ. Nếu có 1 số khác 0 trong vế phải của những phương trình này, hệ hiển nhiên vô nghiệm, ngược lại những phương trình này cũng không có ý nghĩa (không chứa đựng thông tin gì về nghiệm). Có thể gọi đây là **Sự thừa thông tin**.

(2) $m < n$. Lần này có ít nhất $n - m$ cột không chứa hệ số đầu khác 0 của dòng nào trong $\mathrm{rref}(A)$. Do vậy nếu hệ có nghiệm thì sẽ có vô số nghiệm, với công thức nghiệm tổng quát sử dụng các ẩn tương ứng với các cột kể trên làm tham số. Có thể gọi đây là **Sự thiếu thông tin** (vì không đủ phương trình để *ràng buộc* tất cả các nghiệm).

(3) $m = n$. A là một ma trận vuông. Đây là trường hợp duy nhất có thể dẫn đến hệ có nghiệm duy nhất mà không bị thừa ra một phương trình "vô dụng" nào.

Như vậy, nếu một hệ phương trình được xem là "đẹp" khi nó không thừa, không thiếu thông tin gì mà vẫn có nghiệm, thì nó phải có n ẩn và n phương trình, nghĩa là được biểu diễn bằng một **ma trận vuông**. Điều này cũng có nghĩa mọi tính chất của các hệ phương trình tuyến tính đều có thể được suy ra từ việc nghiên cứu các hệ "vuông" này.

Một lý do sâu xa cho tầm quan trọng của ma trận vuông chính là sự *đóng* của tập hợp các ma trận vuông đối với phép nhân ma trận:

$$\forall A, B \in \mathbb{R}^{n\times n} : AB \in \mathbb{R}^{n\times n}.$$

Điều này biến tập hợp các ma trận vuông $\mathbb{R}^{n\times n}$ thành một **vành** với nhiều nhóm con có tính chất thú vị[3], có nhiều ứng dụng trong hình học hiện đại.

Tính chất đóng của phép nhân cũng dẫn đến khái niệm **nghịch đảo** và tính đóng của phép nghịch đảo ma trận sẽ được nói đến sau đây.

5.1 Ma trận khả nghịch

Định nghĩa 5.1. Cho $A \in \mathbb{R}^{n\times n}$, A được gọi là **khả nghịch** nếu và chỉ nếu tồn tại $B \in \mathbb{R}^{n\times n}$ sao cho:

$$AB = BA = I_n.$$

3 Khái niệm vành và nhóm không thuộc phạm vi của bài giảng này.

Giả sử B và C là hai ma trận sao cho:

$$\begin{cases} AB = BA = I_n, \\ AC = CA = I_n. \end{cases}$$

Khi đó $B = B.I_n = B.(AC) = (BA).C = I_n.C = C$. Điều này có nghĩa ma trận khả nghịch của A nếu có thì phải là duy nhất.

Ký hiệu: A^{-1}.

Tính chất 5.2. *Giả sử $A, B \in \mathbb{R}^{n\times n}$ khả nghịch và $\alpha \in \mathbb{R}$. Khi đó:*

- ⋆ *I_n khả nghịch và $(I_n)^{-1} = I_n$,*
- ⋆ *(A^{-1}) khả nghịch và $\left(A^{-1}\right)^{-1} = A$,*
- ⋆ *αA khả nghịch và $(\alpha A)^{-1} = \dfrac{1}{\alpha}A^{-1}$,*
- ⋆ *A^T khả nghịch và $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$,*
- ⋆ *AB khả nghịch và $(AB)^{-1} = B^{-1}A^{-1}$.*

Định lý 5.3. *Cho $A \in \mathbb{R}^{n\times n}$. Khi đó các khẳng định sau tương đương:*

- ⋆ *A khả nghịch*
- ⋆ *Tồn tại $B \in \mathbb{R}^{n\times n}$ sao cho $AB = I_n$,*
- ⋆ *Tồn tại $B \in \mathbb{R}^{n\times n}$ sao cho $BA = I_n$,*
- ⋆ *$\mathrm{rank}(A) = n$,*
- ⋆ *$\mathrm{rref}(A) = I_n$.*

Mặt khác, khi đó nếu $\varphi_1, \dots, \varphi_k$ là các phép biến đổi sơ cấp dòng biến A thành I_n thì các phép biến đổi này (theo đúng thứ tự) cũng biến đổi I_n thành A^{-1}.

Từ định lý trên, ta có được phương pháp xác định tính khả nghịch và tìm ma trận khả nghịch (nếu có):

Xác định tính khả nghịch

Cho $A \in \mathbb{R}^{n\times n}$:

- ⋆ **Bước 1.** Thiết lập ma trận bổ sung $[A|I_n]$.
- ⋆ **Bước 2.** Thực hiện các phép khử Gauss - Jordan trên $[A|I_n]$ sao cho A trở về dạng ma trận bậc thang rút gọn:

$$[A|I_n] \xrightarrow{\text{Thuật toán khử Gauss - Jordan}} [\mathrm{rref}(A)|B]\,.$$

- ⋆ **Bước 3.** Nếu $\mathrm{rref}(A) = I_n$ thì ta khẳng định A khả nghịch và $A^{-1} = B$.

  Ngược lại, A không khả nghịch.

Nhận xét 5.4. Nếu $A \in \mathbb{R}^n$ khả nghịch thì hệ phương trình $Ax = b$ có nghiệm duy nhất $x = A^{-1}b$ với mọi $b \in \mathbb{R}^n$ (Chỉ cần nhân cả 2 vế với A^{-1}). Đây là một cách biểu diễn gọn gàng cho nghiệm của hệ phương trình "khả nghịch".

Một loại ma trận khả nghịch thường gặp nhất là ma trận trực giao.

Định nghĩa 5.5. (**Ma trận trực giao**) Ma trận $A \in \mathbb{R}^{n\times n}$ được gọi là **trực giao** nếu

$$A^T A = I_n.$$

Nhận xét 5.6. Các dòng (hoặc cột) của một ma trận trực giao tương ứng với một cơ sở trực chuẩn (mục 5.6) của $\mathbb{R}^n$. Ma trận trực giao tương ứng với một ánh xạ tuyến tính (mục 6) bảo toàn khoảng cách.

Mệnh đề 5.7. *Cho $A \in \mathbb{R}^{n\times n}$ là một ma trận trực giao và $\vec{v} \in \mathbb{R}^n$. Khi đó ta có*

$$\|A\vec{v}\| = \|\vec{v}\|,$$

Khái niệm tiếp theo, **định thức**, từng giữ vai trò quan trọng trong việc xác định số nghiệm của hệ n phương trình n ẩn, cũng như việc tìm các **giá trị riêng** của ma trận - một bài toán thường gặp trong Machine Learning.

5.2 Định thức

5.2.1 Định nghĩa bằng công thức Leibniz

Trước hết, ta bắt đầu với công thức tính định thức sau:

Định nghĩa 5.8 (Công thức Leibniz)**.** Cho ma trận $A \in \mathbb{R}^{n\times n}$, **định thức** của A (ký hiệu là $\det$(A)) được xác định bởi:

$$\det(A) = \sum_{\sigma \in \mathrm{S}_n} \left(\mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \right),$$

trong đó S_n là tập tất cả các hoán vị trên tập $\{1, \ldots, n\}$ và $\mathrm{sgn}(\sigma)$ là dấu của hoán vị σ[4].

Khi ma trận A được viết dưới dạng biểu thức, ta có 1 cách ký hiệu khác cho $\det(A)$:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} := \det \left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \right).$$

Có thể thấy, công thức trên quá phức tạp và tốn nhiều thời gian để tính khi n lớn. Việc tính định thức dựa trên công thức này là quá phức tạp ngay cả với máy tính. Hơn nữa, nó không thể hiện trực quan liên hệ giữa định thức và tính khả nghịch/tính giải được của hệ phương trình.

Sau đây ta giới thiệu thêm hai cách tính khác cho định thức.

4 Định nghĩa hoán vị và dấu của hoán vị bạn đọc có thể tự tìm hiểu.

5.2.2 Khai triển phần phụ

Định nghĩa 5.9 (Ma trận phụ)**.** Cho ma trận $A \in \mathbb{R}^{n\times n}$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Ma trận phụ tại hàng i, cột j của A, ký hiệu $A_{\neq ij}$, là ma trận sau:

$$A_{\neq ij} := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2(j-1)} & a_{2(j+1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{(i-1)1} & a_{(i-1)2} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ a_{(i+1)1} & a_{(i+1)2} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{bmatrix}.$$

Nói cách khác, $A_{\neq ij}$ là định thức của ma trận vuông cấp $n-1$ sau khi xóa dòng i và cột j từ A nhân với $(-1)^{i+j}$.

Định lý 5.10 (Khai triển phần phụ)**.** *Hàm định thức* det *là hàm số duy nhất từ* $\bigcup_{n\in\mathbb{N}} \mathbb{R}^{n\times n}$ *đến* $\mathbb{R}$ *thỏa mãn:*

1. $\det\left(\left[a\right]\right) = a \quad \forall a \in \mathbb{R}$.
2. *Với mọi* $n > 1$, $A \in \mathbb{R}^{n\times n}$, $k \in \{1, 2, \ldots, n\}$

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det\left(A_{\neq ik}\right) = \sum_{i=1}^{n} (-1)^{k+i} a_{ki} \det\left(A_{\neq ki}\right).$$

Biểu thức ở giữa được gọi là **khai triển phần phụ theo cột** k*, biểu thức bên phải được gọi là* **khai triểu phần phụ theo hàng** k*.*

Như vậy, khai triển phần phụ theo một cột hay hàng bất kì trong ma trận đều ra cùng kết quả là định thức ma trận đó.

Chứng minh. Xem bài tập 5.4. □

Từ Định lý 5.10, ta có thể chứng minh những tính chất sau:

Tính chất 5.11. *Cho* $A \in \mathbb{R}^{n\times n}$*. Khi đó:*

- ⋆ *Nếu nhân 1 hàng hoặc cột bất kỳ của A lên k lần thì* $\det(A)$ *cũng tăng lên k lần.*
- ⋆ *Nếu 1 hàng/cột của A gấp k lần 1 hàng/cột khác thì* $\det(A) = 0$.
- ⋆ *Nếu cộng 1 hàng/cột của A với 1 hàng/cột khác,* $row_i(A) \leftarrow row_i(A) + row_j(A)$ *hoặc* $col_i(A) \leftarrow col_i(A) + col_j(A)$*, thì* $\det(A)$ *không đổi.*
- ⋆ *Nếu hoán đổi 2 cột hoặc 2 hàng của A thì* $\det(A)$ *đổi dấu.*

Như vậy, việc thực hiện những phép biến đổi tương đương dòng chỉ gây những thay đổi "trong tầm kiểm soát" đối với giá trị định thức. Từ đó ta có thêm phương pháp tính định thức sau.

5.2.3 Tính định thức bằng phép khử Gauss-Jordan

Thuật toán tính định thức Gauss-Jordan

Bắt đầu với $i := 1, j := 1, D := 1$.

- **Bước 0.** Nếu $i > m$ hoặc $j > n$ thì trả về
$$\det(A) = D,$$
không thì chuyển qua bước 1.

- **Bước 1.** Nếu $a_{ij} = 0$ thì tráo đổi dòng i với một dòng bên dưới để thu được $a_{ij} \neq 0$. Cập nhật
$$D := -D,$$
sau đó qua bước 2.
Trường hợp mà $a_{ik} = 0$ với mọi $j \leq k$, trả về
$$\det(A) = 0.$$

- **Bước 2.** Đưa hệ số cao nhất trên hàng i về 1:
$$\mathrm{row}_i := \frac{1}{a_{ij}}\mathrm{row}_i,$$
đồng thời cập nhật
$$D := a_{ij} \cdot D.$$
Sau đó chuyển sang bước 3.

- **Bước 3.** Đưa tất cả các giá trị trên cột j (ngoại trừ a_{ij}) về 0:
$$\mathrm{row}_k := \mathrm{row}_k - a_{kj}\mathrm{row}_i, \qquad \forall k \neq i.$$
Chuyển sang bước 4.

- **Bước 4.** Tăng i và j lên 1 đơn vị:
$$i := i + 1,$$
$$j := j + 1.$$
Quay về bước 0.

5.2.4 Ví dụ

Ví dụ 5.12. Sử dụng 3 phương pháp tính định thức, chứng minh rằng

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Chứng minh bằng công thức Leibniz. Trên $\{1,2\}$, ký hiệu $\sigma = (a,b)$ là hoán vị mà $\sigma(1) = a, \sigma(2) = b$. Ta có:

$$\operatorname{sgn}((1,2)) = 1, \quad \operatorname{sgn}((2,1)) = -1.$$

Như vậy theo định nghĩa 5.8, ta có

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \operatorname{sgn}((1,2))a_{11}a_{22} + \operatorname{sgn}((2,1))a_{12}a_{21} = a_{11}a_{22} - a_{12}a_{21}.$$

□

Chứng minh bằng khai triển phần phụ. Đặt $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, khi đó

$$A_{\neq 11} = \begin{bmatrix} a_{22} \end{bmatrix}, \quad A_{\neq 21} = \begin{bmatrix} a_{12} \end{bmatrix}.$$

Do đó

$$\det(A) = (-1)^{1+1}a_{11}\det(A_{\neq 11}) + (-1)^{2+1}a_{21}\det(A_{\neq 21}) = a_{11}a_{22} - a_{21}a_{12}.$$

□

Chứng minh bằng phép khử Gauss - Jordan. Ta xét 3 trường hợp:

(1) $a_{11} \neq 0$. Khi đó:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\begin{vmatrix} 1 & \frac{a_{12}}{a_{11}} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\begin{vmatrix} 1 & \frac{a_{12}}{a_{11}} \\ 0 & a_{22} - \frac{a_{12}a_{21}}{a_{11}} \end{vmatrix}.$$

Nếu $a_{22} - \frac{a_{12}a_{21}}{a_{11}} = 0$, ta có

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = 0 = a_{11}a_{22} - a_{12}a_{21}.$$

Ngược lại, ta có:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}a_{22} - a_{12}a_{21})\begin{vmatrix} 1 & \frac{a_{12}}{a_{11}} \\ 0 & 1 \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

(2) $a_{11} = 0 \neq a_{12}$. Sử dụng trường hợp 1, ta có:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = -\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix} = -(a_{21}a_{12} - a_{11}a_{22}) = a_{11}a_{22} - a_{12}a_{21}.$$

(3) $a_{11} = a_{12} = 0$. Ta có:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = 0.$$

Như vậy đẳng thức cần chứng minh đúng trong mọi trường hợp. □

5.3 Quan hệ giữa định thức, các phép toán ma trận và hệ phương trình

Tính chất 5.13. *Cho $A, B \in \mathbb{R}^{n\times n}$ và $\alpha \in \mathbb{R}$, ta có các tính chất sau:*

- ⋆ $\det(I_n) = 1$.
- ⋆ $\det(A) = \det(A^T)$.
- ⋆ *A khả nghịch nếu và chỉ nếu* $\det(A) \neq 0$, *và khi đó* $\det(A^{-1}) = \dfrac{1}{\det(A)}$.
- ⋆ $\det(\alpha A) = \alpha^n \det(A)$.
- ⋆ $\det(AB) = \det(A)\det(B)$.
- ⋆ *AB khả nghịch khi và chỉ khi A và B đều khả nghịch.*

Định lý sau liên kết các khái niệm khả nghịch, định thức và việc khảo sát số nghiệm của hệ phương trình tuyến tính vuông:

Định lý 5.14. *Cho ma trận $A \in \mathbb{R}^{n\times n}$ và $b \in \mathbb{R}^n$. Xét hệ phương trình $Ax = b$, ta có những điều sau:*

- ⋆ *Nếu A khả nghịch* $\iff \det(A) \neq 0$, *hệ có nghiệm duy nhất* $x = A^{-1}b$.
- ⋆ *Nếu A không khả nghịch* $\iff \det(A) = 0$, *hệ vô nghiệm nếu* $\mathrm{rref}([A|b])$ *có một dòng i toàn số 0 ở vế trái trong khi vế phải khác 0. Nếu không tồn tại dòng như vậy thì hệ có vô số nghiệm.*
- ⋆ *Nếu hệ có vô số nghiệm và x_0 là một nghiệm trong đó, mọi nghiệm khác đều có dạng $x = x_0 + z$ với z là một nghiệm của $Az = 0$.*

Bài tập

Bài tập 5.1. Chứng minh tính chất 5.11. Chứng minh thêm rằng định thức không đổi nếu $\text{row}_i \leftarrow \text{row}_i + k\text{row}_j, \quad \forall k \in \mathbb{R}$.

Bài tập 5.2. Tìm công thức tổng quát cho $\det\left((a_{ij})_{3\times 3}\right)$.

Bài tập 5.3 (Khó). Chứng minh $\det(AB) = \det(A)\det(B)$ với mọi $A, B \in \mathbb{R}^{n\times n}$. **Gợi ý:** Sử dụng thuật toán Gauss - Jordan và định thức của ma trận sơ cấp.

Bài tập 5.4 (Khó). Chứng minh công thức Leibniz (Định nghĩa 5.8) và Định lý 5.10 dựa vào định nghĩa truy hồi của định thức trong Định lý 5.10. **Gợi ý:** đầu tiên xây dựng hàm số f sao cho:

1. $f\left([a]\right) = a, \quad \forall a \in \mathbb{R}$.
2. Với mọi $n > 1$, $A \in \mathbb{R}^{n\times n}$,

$$f(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det\left(A_{\neq i1}\right).$$

Rõ ràng $f(A)$ được định nghĩa duy nhất cho mỗi A. Dùng quy nạp để chứng minh công thức Leibniz và Định lý 5.10.

Bài tập 5.5. Cho ma trận $A \in \mathbb{R}^{n\times n}$ khả nghịch. Với mỗi $i, j \le n$, đặt $A_{ij} = (-1)^{i+j} \det(A_{\neq ij})$, gọi là **phần phụ** ở hàng i, cột j của A. Chứng minh rằng

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}.$$

Bài tập 5.6. Chứng minh rằng các ma trận sơ cấp đều khả nghịch.

Bài tập 5.7. Cho $A, B \in \mathbb{R}^{m\times n}$. Chứng minh rằng điều kiện cần và đủ để $A \sim B$ là tồn tại ma trận khả nghịch $P \in \mathbb{R}^{m\times m}$ sao cho $A = PB$. Từ đó suy ra $\operatorname{rank}(PA) = \operatorname{rank}(A)$ nếu P khả nghịch.

Bài tập 5.8. Ma trận nghịch đảo A^{-1} thay đổi thế nào nếu:

1. Hoán vị hai dòng của A?
2. Nhân một dòng với hằng số khác 0?
3. Thêm vào dòng thứ i: α lần dòng j?
4. Tương tự nhưng thực hiện trên cột?

Bài tập 5.9. Cho $A, B \in \mathbb{C}^{n\times n}$ và $AB = xA + yB$ với $x, y \in \mathbb{R}$. Chứng minh rằng $AB = BA$.

Bài tập 5.10. Liệu có tồn tại $A, B \in \mathbb{R}^{n\times n}$ sao cho $AB - BA = I_n$?

Bài tập 5.11. Cho $A, B \in \mathbb{R}^{n\times n}$ thỏa mãn $\operatorname{rank}(AB - BA) = 1$. Chứng minh rằng $(AB - BA)^2 = 0_n$.

Bài tập 5.12. Chứng minh rằng nếu A là ma trận đối xứng và khả nghịch thì A^{-1} cũng đối xứng.

Bài tập 5.13 (Khó). Cho $A \in \mathbb{R}^{m\times n}$ và $B \in \mathbb{R}^{n\times m}$. Chứng minh rằng $I_m + AB$ khả nghịch khi và chỉ khi $I_n + BA$ khả nghịch.

Bài tập 5.14. $A \in \mathbb{R}^{n\times n}$ được gọi là lũy linh nếu tồn tại $k \in \mathbb{N}$ sao cho $A^k = 0$. Chứng minh rằng $I_n - A$ khả nghịch và tìm nghịch đảo của ma trận này.

Bài tập 5.15. Xét tính khả nghịch và tìm nghịch đảo (nếu có) của các ma trận sau:

a) $\begin{bmatrix} 2 & 5 & -1 \\ 6 & -1 & 2 \\ 6 & 4 & 2 \end{bmatrix}$. c) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix}$. d) $\begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 4 \\ 1 & 2 & 5 \end{bmatrix}$.

b) $\begin{bmatrix} 2 & -1 & -1 \\ 3 & 4 & -2 \\ 3 & -2 & 4 \end{bmatrix}$. d) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}$. f) $\begin{bmatrix} 6 & 9 & 6 \\ 9 & 6 & 9 \\ 15 & 15 & 15 \end{bmatrix}$.

Bài tập 5.16. Nếu $A \in \mathbb{R}^{n\times n}$ là một ma trận trực giao thì $\det(A) = \pm 1$.

Bài tập 5.17. Chứng minh Mệnh đề 5.7.

Đọc thêm. Ban đầu trước khi có bộ môn Đại Số Tuyến Tính, Leibniz khi nghiên cứu cách giải cho các hệ phương trình tuyến tính đã phát minh ra phép tính định thức cho các hệ phương trình "vuông", trùng với khái niệm định thức cho các ma trận tương ứng ngày nay. Trên thực tế, có thể viết nghiệm của hệ phương trình $Ax = b$ thành biểu thức chính xác qua các định thức bằng **công thức Cramer**, từ đó suy ra công thức của A^{-1} theo A.

Tuy nhiên, việc tính định thức quá phức tạp và tốn nhiều thời gian ngay cả với máy tính hiện đại. Sự ra đời của Thuật toán Gauss - Jordan chính là bước tiến lớn cho bài toán giải hệ phương trình, đồng thời mang lại phương pháp tính định thức nhanh hơn rất nhiều và phù hợp với máy tính ngày nay.

6 Không gian vector

Không gian vector, cùng các *ánh xạ tuyến tính* (sẽ học ở phần 7), là đối tượng được nghiên cứu chủ yếu trong đại số tuyến tính hiện đại. Không gian vector ban đầu ra đời để phục vụ việc chứng minh một số tính chất cơ bản (nhưng khó) của ma trận, điển hình như:

Ví dụ 6.1. Cho ma trận $A \in \mathbb{R}^{n\times n}$. Nếu tồn tại *nghịch đảo phải* của A, tức B sao cho $AB = I_n$, thì tồn tại nghịch đảo trái C sao cho $CA = I_n$, và $B = C$.

Cùng với sự phát triển ĐSTT theo xu hướng hình học, các không gian vector trở thành phiên bản tổng quát của mặt phẳng $\mathbb{R}^2$ và không gian $\mathbb{R}^3$, giúp mở rộng số chiều của hình học hiện đại lên bất kì hoặc vô hạn.

Hầu hết các đối tượng có nhiều ứng dụng trong thực tế như không gian dữ liệu hay tập hàm liên tục đều là các ví dụ của không gian vector. Không gian vector được xây dựng từ phép lấy tổ hợp tuyến tính và bài toán biểu diễn tuyến tính được là một bài toán trọng yếu của đại số tuyến tính.

Về mặt ứng dụng, các dữ liệu trọng thực tế rất khổng lồ và khó hình dung (số chiều của không gian vector bằng số thông số của dữ liệu). Khái niệm **không gian con** và **chiều** cung cấp cho chúng ta công cụ để giảm kích thước của dữ liệu nhưng vẫn giữ được những thông tin quan trọng.

Ví dụ 6.2. Hải Đông thu thập số liệu về những căn nhà cần được bán với giá x, diện tích y, có z lầu, và đã xây được t năm. Hải Đông phát hiện ra số lầu không ảnh hưởng đến việc nhà có bán được hay không. Như vậy, thực chất dữ liệu chúng ta cần quan tâm ở trong không gian $\mathbb{R}^3$, không phải trong $\mathbb{R}^4$.

6.1 Các quan hệ tuyến tính và tập sinh

Cho tập hợp S gồm các vector trong $\mathbb{R}^n$. Lưu ý là ta có thể xây dựng các định nghĩa tương tự dưới đây cho một không gian vector bất kì, không nhất thiết phải là $\mathbb{R}^n$.

Định nghĩa 6.3 (Tổ hợp tuyến tính)**.** Vector $\vec{v}$ được gọi là một tổ hợp tuyến tính của $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ thuộc S nếu tồn tại các số $c_i \in \mathbb{R}$ sao cho

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k.$$

Lúc đó, ta nói $\vec{v}$ biểu diễn tuyến tính được qua S (trên $\mathbb{R}$).

Ví dụ 6.4.

1. Trong $\mathbb{R}^3$, ta có thể biểu diễn

$$\begin{aligned}(3,5,7) &= 3\times(1,0,0) + 5\times(0,1,0) + 7\times(0,0,1)\\ &= 3\times(1,0,0) + 1\times(0,1,0) + 2\times(0,2,0) + 7\times(0,0,1).\end{aligned}$$

2. Đa thức $x^3 - 2x^2 - 2x + 1 \in \mathbb{R}[x]$ có thể được viết thành:

$$1\times x^3 + (-2)\times x^2 + (-2)\times x + 1\times 1.$$

Nhận xét 6.5. Nếu $A \in \mathbb{R}^{n\times k}$ là ma trận có các vector cột $\vec{v}_i$, và $x = \begin{bmatrix} c_1 & c_2 & \cdots & c_k \end{bmatrix}^T \in \mathbb{R}^k$ thì ta có $A\vec{x} = c_1\vec{v}_1 + \ldots + c_k\vec{v}_k$. Như vậy, kết quả phép nhân ma trận với một vector thực chất là một tổ hợp tuyến tính của các cột của ma trận đó.

Nhận xét 6.6. Cho các vector $v_1, v_2, \cdots, v_k \in \mathbb{R}^n$ và vector $v \in \mathbb{R}^n$ và $x = \begin{bmatrix} x_1 & x_2 & \cdots & x_k \end{bmatrix}^T \in \mathbb{R}^k$. Khi đó 2 điều sau tương đương:

⋆ $v = x_1v_1 + x_2v_2 + \cdots + x_kv_k$.

⋆ x thỏa mãn $Ax = v$ với $A = \begin{bmatrix} v_1 & v_2 & \cdots v_k \end{bmatrix} \in \mathbb{R}^{n\times k}$.

Như vậy việc tìm tổ hợp tuyến tính để biểu diễn một vector thực chất chính là việc giải hệ phương trình tuyến tính.

Định nghĩa 6.7. Tập sinh của S, kí hiệu là $\mathrm{Span}(S)$, bao gồm tất cả những tổ hợp tuyến tính có thể của **hữu hạn** các vectors trong S:

$$\mathrm{Span}(S) = \{\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k \mid \quad \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in S, \quad c_1, c_2, \ldots, c_k \in \mathbb{R}\}.$$

Nhận xét 6.8. Dễ dàng chứng minh các tính chất sau:

⋆ $A\vec{x}$ thuộc tập sinh bởi các cột của A (không gian cột của A).

⋆ Nếu S nằm trong tập nghiệm của $Ax = 0$ thì $\mathrm{Span}(S)$ cũng vậy.

Ví dụ 6.9. Nếu tập $S = \{(1, 0), (0, 1)\}$ thì $\mathrm{Span}(S) = \mathbb{R}^2$.

Định nghĩa 6.10 (Phụ thuộc tuyến tính)**.** Các vector trong S *phụ thuộc tuyến tính* nếu tồn tại các số $c_1, c_2, \ldots, c_k \in \mathbb{R}$, không đồng thời bằng 0, sao cho

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0.$$

Các vector trong S được gọi là *độc lập tuyến tính* nếu chúng **không** phụ thuộc tuyến tính. Như vậy, tập $\{v_1, \ldots, v_k\}$ độc lập tuyến tính nếu

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0 \iff c_i = 0, \qquad \forall i \in \{1, \ldots, k\}.$$

Ví dụ 6.11. Các vector $(1, 2, 1), (6, 5, 3)$ độc lập tuyến tính còn các vector $(1, 2, 0), (-1, 1, 0), (0, 6, 0)$ phụ thuộc tuyến tính trong $\mathbb{R}^3$.

Mệnh đề 6.12. *Cho $S = \{v_1, \ldots, v_k\}$. Khi đó:*

⋆ *S phụ thuộc tuyến tính nếu và chỉ nếu tồn tại vector $v_i \in S$ sao cho v_i là một tổ hợp tuyến tính của những vector còn lại.*

⋆ *S độc lập tuyến tính nếu và chỉ nếu với mọi vector v biểu diễn tuyến tính được qua S thì đó là biểu diễn duy nhất.*

6.2 Không gian con của $\mathbb{R}^n$ và những không gian con cơ bản

Trước khi đưa ra định nghĩa tổng quát của không gian vector, chúng ta làm việc với các không gian vector hay gặp nhất là $\mathbb{R}^n$ và các không gian con của $\mathbb{R}^n$.

Định nghĩa 6.13. Cho một tập con $V \in \mathbb{R}^n$. V được gọi là một không gian con của $\mathbb{R}^n$ nếu V thỏa mãn các điều kiện sau:

1. $\vec{0} \in V$,
2. Nếu $\vec{v} \in V$ và $c \in \mathbb{R}$ thì $c\vec{v} \in V$,
3. Nếu $\vec{v}, \vec{w} \in V$ thì $\vec{v} + \vec{w} \in V$.

Ví dụ 6.14. Mặt phẳng $x - y + z = 0$ là một không gian con của $\mathbb{R}^3$ nhưng $x - y + z = 2$ không phải là một không gian con của $\mathbb{R}^3$ (Tại sao?).

(Nhắc lại) Tích vô hướng của $\vec{x} = (x_1, \ldots, x_n)$ và $\vec{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ (ký hiệu $\cdot$) được xác định như sau:

$$\vec{x} \cdot \vec{y} = x_1 y_1 + \cdots + x_n y_n.$$

Định nghĩa 6.15. Cho V là không gian con của $\mathbb{R}^n$. **Phần bù trực giao** (orthogonal complement) của V trong $\mathbb{R}^n$, ký hiệu là $V^\perp$, được xác định bởi

$$V^\perp = \{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{v} = 0 \quad \forall \vec{v} \in V\}.$$

Lưu ý: hai không gian con V và $W \subset \mathbb{R}^n$ được gọi là trực giao nhau nếu $\vec{v} \cdot \vec{w} = 0$ với mọi $\vec{v} \in V, \vec{w} \in W$.

Bổ đề 6.16.

1. *Nếu V và W là 2 không gian con trực giao của $\mathbb{R}^n$ thì $V \cap W = \{\vec{0}\}$.*
2. *Nếu $V \subset W^\perp$ thì $W \subset V^\perp$.*

Chứng minh. Giả sử $\vec{x} \in V \cap W$, ta có $\vec{x} \cdot \vec{x} = 0$. □

Định lý 6.17. *Nếu V là một không gian con $\mathbb{R}^n$ thì*

$$(V^\perp)^\perp = V.$$

6.3 Bốn không gian con cơ bản sinh bởi một ma trận

Từ Nhận xét 6.8, ta có thể xây dựng 4 tập hợp cơ bản gắn với một ma trận A và chứng minh chúng đều là các không gian vector. Việc phát hiện những không gian này liên kết chúng với hạng của ma trận $\mathrm{rank}(A)$ giúp chứng minh nhiều định lý trong ĐSTT, trong đó có bài toán ở Ví dụ 6.1.

Định nghĩa 6.18. Cho ma trận $A \in \mathbb{R}^{m \times n}$, với $\vec{a_i}$ $(1 \leq i \leq n)$ là các vector cột và $\vec{A_j}$ $(1 \leq j \leq m)$ là các vector dòng. Bốn không gian con cơ bản của $\mathbb{R}^n$ được xác định bởi:

1. Không gian hạch (nullspace):

$$N(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\} \subset \mathbb{R}^n.$$

2. Không gian cột (columnspace):

$$C(A) = \mathrm{Span}\{\vec{a}_1, \ldots, \vec{a}_n\} \subset \mathbb{R}^m.$$

3. Không gian dòng (rowspace):

$$R(A) = \mathrm{Span}\{\vec{A}_1, \ldots, \vec{A}_m\} \subset \mathbb{R}^n.$$

4. Không gian hạch trái (left-nullspace):

$$N(A^T) = \{\vec{x} \in \mathbb{R}^m : \vec{x}^T \cdot A = \vec{0}\} \subset \mathbb{R}^m.$$

Mệnh đề 6.19. *Cho ma trận A kích thước $m \times n$. Ta có các kết quả sau:*

1. $N(A^T) = C(A)^{\perp}$,
2. $N(A) = R(A)^{\perp}$,
3. $N(A^T)^{\perp} = C(A)$,
4. $N(A)^{\perp} = R(A)$.

Bài tập

Bài tập 6.1. Chứng minh rằng tập sinh bởi hữu hạn các vector trong $\mathbb{R}^n$ là một không gian con.

Bài tập 6.2. Chứng minh Mệnh đề 6.12.

Bài tập 6.3. Cho ma trận A vuông cấp n thỏa mãn $A^2 = A$. Chứng minh rằng:
a) $C(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{x}\}$.
b) $N(A) = \{\vec{x} \in \mathbb{R}^n : \vec{x} = \vec{u} - A\vec{u},\ \text{với } \vec{u} \in \mathbb{R}^n\}$.
Nhận xét: Ma trận thỏa mãn $A^2 = A$ tương ứng với ánh xạ tuyến tính là phép chiếu lên một không gian con (không gian cột của A).

6.4 Không gian vector, Cơ sở và Chiều

Trong phần này chúng ta đến với định nghĩa tổng quát của không gian vector và các khái niệm trọng yếu liên quan đến chúng.

Định nghĩa 6.20 (Không gian vector)**.** Một không gian vector V trên trường số thực $\mathbb{R}$ là một tập hợp cùng với hai phép tính: cộng vector $(+)$ và nhân vector với một số thực $(.)$, sao cho với mọi $\vec{u}, \vec{v}, \vec{w} \in V$, với mọi số thực $c, d \in \mathbb{R}$, ta có các tính chất sau:

1. Giao hoán của phép cộng: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$,
2. Kết hợp của phép cộng: $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$,
3. Tồn tại vector 'không', $\vec{0} \in V : \vec{0} + \vec{u} = \vec{u} + \vec{0}$,
4. Tồn tại vector đối, $-\vec{u} \in V : (-\vec{u}) + \vec{u} = 0$,
5. Nhân kết hợp với số thực : $c(d\vec{u}) = (cd)\vec{u}$,
6. Phân phối của phép nhân với phép cộng : $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$,
7. Phân phối của phép cộng với phép nhân : $(c + d)\vec{u} = c\vec{u} + d\vec{u}$,

8. Phép nhân với $1 \in \mathbb{R} : 1\vec{u} = \vec{u}$.

Ví dụ 6.21.

1. Tập $\mathbb{R}^n$ với phép cộng vector thông thường và phép nhân vector với một hằng số thực.
2. Tập hợp các đa thức một biến $f(x)$ có hệ số thực, thường ký hiệu là $\mathbb{R}[x]$, với phép cộng đa thức và nhân đa thức với một hằng số thực.
3. Tập hợp các ma trận $m \times n$ với phép cộng ma trận và nhân ma trận với một hằng số thực.

Định nghĩa 6.22. Cho không gian vector V trên trường số thực $\mathbb{R}$. Một tập con $U \subset V$ được gọi là một **không gian con** của V nếu U thỏa mãn:

1. Vector $\vec{0}$ thuôc U,
2. Nếu $\vec{u}, \vec{v}$ thuôc U thì $\vec{u} + \vec{v}$ thuộc U,
3. Nếu $\vec{v}$ thuộc U thì $c\vec{v}$ thuộc U với mọi $c \in \mathbb{R}$.

Ví dụ 6.23.

1. Các đường thẳng, mặt phẳng, hay siêu phẳng đi qua gốc tọa độ đều là không gian con của $\mathbb{R}^n$. Tập sinh bởi hữu hạn các vector trong $\mathbb{R}^n$ là một không gian con.
2. Tập hợp các đa thức bậc bé hơn hoặc bằng 2 là một không gian con của $\mathbb{R}[x]$.
3. Nếu V là một không gian vector thì V là không gian con của chính nó.

Định nghĩa 6.24 (Cơ sở)**.** Cho không gian vector V và U là một không gian con của V. Một **cơ sở** (basis) của U là một tập hợp S gồm các vectors (có thể vô hạn) thuộc U thỏa mãn:

1. Các vectors trong S độc lập tuyến tính với nhau,
2. Tập sinh bởi các vectors này cũng là U: $\mathrm{Span}(S) = U$.

Ví dụ 6.25.

1. Trong $\mathbb{R}^3$, các vector $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$ tạo thành một cơ sở (chính quy).
2. Trong $\mathbb{R}^2$, các vector $\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}$ tạo thành một cơ sở.

Một không gian vector như $\mathbb{R}^n$ có vô số cơ sở, do đó ta cần tìm một cơ sở dễ dùng nhất làm đại diện để thuận tiện cho việc tính toán.

Định nghĩa 6.26 (Cơ sở chính quy)**.** Trong không gian $\mathbb{R}^n$, hệ **cơ sở chính quy** gồm các vector

$$S_n = \{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}.$$

Trong đó với mỗi $i = 1, 2, \cdots, n$: $e_{ij} = 0$ nếu $j \neq i$ và $e_{ii} = 1$.

Nhận xét 6.27. Ma trận có các cột là các vector trong cơ sở chính quy chính là I_n.

Tuỳ theo từng bài toán mà một cơ sở phù hợp sẽ giúp cho việc biểu diễn các vector đơn giản hơn những cơ sở khác.

Định nghĩa 6.28 (Biểu diễn trong cơ sở)**.** Cho không gian vector U và một cơ sở hữu hạn $S = \{\vec{u}_i : i = 1, 2, \dots, n\}$ của U. Với mỗi vector $\vec{u} \in U$, ký hiệu $[\vec{u}]_S$ là **biểu diễn của $\vec{u}$ qua cơ sở** S. Nói cách khác,

$$[\vec{u}]_S = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T,$$

là một vector trong $\mathbb{R}^n$ thỏa mãn:

$$\vec{u} = x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n.$$

Ví dụ 6.29. Vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ trong $\mathbb{R}^2$ có biểu diễn là $\begin{bmatrix} 2 & 1 \end{bmatrix}$ trong cơ sở chính quy và có biểu diễn là $\begin{bmatrix} 0 & 1 \end{bmatrix}$ trong cơ sở $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Ứng dụng của cơ sở

Nhận xét 6.30. Biểu diễn còn có thể gọi là **tọa độ** trong một cơ sở xác định. Ý tưởng trong ví dụ 6.29 chính là nền tảng cho việc **giảm số chiều dữ liệu** trong Machine Learning hay Data Science.
Thông thường, một cơ sở dữ liệu D với M vector trong $\mathbb{R}^N$ với N rất lớn sẽ cần N bits để lưu trữ mỗi vector. Tuy nhiên nếu ta tìm xấp xỉ được một cơ sở cho $\mathrm{Span}(D)$ chỉ với k vector (không cần phải thuộc M vector trong D), với k nhỏ hơn N rất nhiều, thì có thể biểu diễn mỗi vector dữ liệu chỉ với k bits, giúp giảm kích thước dữ liệu đi rất nhiều.

Tiếp theo, ta sẽ tìm mối tương quan giữa các hệ cơ sở khác nhau và các điểm chung nếu có giữa chúng. Một điểm chung quan trọng là *số phần tử* của chúng, từ đó dẫn đến khái niệm *số chiều* của không gian vector.

Định lý 6.31. *Cho U là một không gian vector.*

1. *Nếu S là một cơ sở của U thì với mọi vector u trong U, tồn tại một cách biểu diễn duy nhất u dưới dạng tổ hợp tuyến tính của hữu hạn vector trong S_1.*
2. *Nếu S_1 và S_2 là 2 cơ sở hữu hạn của U thì chúng có cùng số lượng phần tử.*

Chứng minh.

1. Giả sử điều điều ngược lại: tồn tại một vector u trong U biểu diễn được dưới dạng 2 tổ hợp tuyến tính khác nhau của các vector trong S, ta suy ra các vector dùng để biểu diễn u trên phụ thuộc tuyến tính (mâu thuẫn với định nghĩa của cơ sở).
2. Ta chứng minh bằng phản chứng, không mất tính tổng quát, giả sử $S_1 = \{v_1, \dots, v_k\}$, $S_2 = \{w_1, \dots, w_l\}$, và $k < l$.

   Từ tính chất (1), mỗi vector w_i trong S_2 biểu diễn được duy nhất dưới dạng một tổ hợp tuyến tính của các vector v_j trong S_1, nói cách khác tồn tại các hệ số thực a_{ij}, $\quad 1 \le j \le k$, sao cho $w_i = \sum_j a_{ij} v_j$. Dựng ma trận A gồm l dòng và k thỏa mãn

$$A_{ij} := a_{ij}.$$

Dạng bậc thang rút gọn của A phải có một dòng toàn số 0, hay nói cách khác, các vector dòng của A phụ thuộc tuyến tính. Từ đó, các vector w_i tương ứng với các dòng của A cũng phụ thuộc tuyến tính (mâu thuẫn với định nghĩa của cơ sở).

□

Như vậy, nếu một không gian vector có một cơ sở hữu hạn thì số phần tử trong cơ sở đó xác định một cách duy nhất, và chúng ta xây dựng khái niệm **chiều** của không gian vector như sau.

Định nghĩa 6.32 (Chiều)**.** Cho không gian con U của không gian vector V. Nếu tồn tại một cơ sở hữu hạn S của U, thì số chiều của U, ký hiệu $\dim(U)$, là số phần tử của S:

$$\dim(U) = |S|.$$

Quy ước: $\dim\{\vec{0}\} = 0$.

Mệnh đề 6.33.

1. *Trong một không gian con U với $\dim(U) = k$, một tập hợp gồm $m > k$ vectors bất kỳ trong U phụ thuộc tuyến tính.*
2. $\dim(\mathbb{R}^n) = n$.
3. *Một không gian con U của $\mathbb{R}^n$ có số chiều nhỏ hơn hoặc bằng n.*

Định lý 6.34. *Cho không gian con V với $\dim(V) = n$, và tập $S = \{\vec{v}_1, \dots, \vec{v}_n\}$ gồm n vector trong V. Những khẳng định sau tương đương nhau:*

1. *S là một cơ sở của V,*
2. *S độc lập tuyến tính,*
3. *S sinh ra V: $V \subset \mathrm{Span}(S)$.*

Nhận xét 6.35. Định lý 5.2 là một kết quả rất hữu dụng vì tùy từng bài toán mà việc kiểm tra các một điều kiện cụ thể sẽ dễ hơn các điều kiện còn lại.

Hệ quả 6.36. Nếu S là một không gian con của V và $\dim(V) = n$, thì $\dim(S) \leq n$ và dấu '=' xảy ra khi và chỉ khi $S = V$.

Tiếp theo, chúng ta tính số chiều của các không gian cơ bản sinh bởi một ma trận.

Định lý 6.37. *Nếu $A \in \mathbb{R}^{n\times n}$ là một ma trận khả nghịch $(\mathrm{rank}(A) = n)$ thì:*

1. *Các vector cột (hoặc dòng) của A tạo thành một cơ sở cho $\mathbb{R}^n$.*
2. *Nếu $\{\vec{v}_1, \dots, \vec{v}_n\}$ là một cơ sở của $\mathbb{R}^n$ thì $\{A\vec{v}_1, \dots, A\vec{v}_n\}$ cũng là một cơ sở cho $\mathbb{R}^n$.*

Định lý 6.38. *Cho trận $A \in \mathbb{R}^{m\times n}$ có $\mathrm{rank}\, A = r$. Ta có những kết quả sau:*

1. $\dim R(A) = \dim C(A) = r$,
2. $\dim N(A) = n - r$ *(còn gọi là nullity của A),*
3. $\dim N(A^T) = m - r$.

Định lý 6.39. *(Rank Nullity) Cho ma trận $A \in \mathbb{R}^{m \times n}$. Ta có:*

$$\operatorname{rank} A + \operatorname{null} A = n.$$

Bài tập

Bài tập 6.4. Kiểm tra các tập hợp trong ví dụ 5.5 thỏa mãn các điều kiện của một không gian vector.

Bài tập 6.5. (Cơ sở cho không gian dòng và không gian cột)
Cho ma trận $A \in \mathbb{R}^{m \times n}$ có $R = \operatorname{rref}(A)$ là dạng bậc thang rút gọn.
a) Chứng minh rằng các vector dòng khác 0 của R là một cơ sở cho $R(A)$.
b) Chứng minh rằng các vector cột của R ứng với các hệ số cao nhất khác 0 tạo một cơ sở cho không gian cột của A.

Bài tập 6.6. Cho U và W là 2 không gian con của không gian vector V, giao của U và W được xác định bởi

$$U \cap W = \{\vec{v} \in V \mid \vec{v} \in U \text{ và } \vec{v} \in W\}.$$

Chứng minh rằng $U \cap W$ là không gian con của V.

Bài tập 6.7. Cho U và W là 2 không gian con của $\mathbb{R}^n$, tổng của U và W được xác định bởi

$$U + W = \{\vec{u} + \vec{w} : \forall \vec{u} \in U, \forall \vec{w} \in W\}.$$

a) Chứng minh rằng $U + W$ là không gian con của V.
b) Chưng minh rằng $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$.

6.5 Ma trận chuyển cơ sở

Trong khi làm việc với các không gian vector, thường xuyên nảy sinh tình huống phải chuyển góc nhìn từ cơ sở này sang cơ sở khác. Nhu cầu tìm một công thức biến biểu diễn của một vector bất kì theo cơ sở này thành biểu diễn theo cơ sở khác đến một cách tự nhiên.

Định nghĩa 6.40 (Ma trận chuyển cơ sở)**.** Cho không gian vector n chiều U với các cơ sở $S = \{u_i : i = 1, 2, \cdots, n\}$, $T = \{v_i : i = 1, 2, \cdots, n\}$. **Ma trận chuyển cơ sở** từ S sang T được cho bởi:

$$M_{S \to T} := \begin{bmatrix} [u_1]_T & [u_2]_T & \cdots & [u_n]_T \end{bmatrix}.$$

Tính chất quan trọng sau giúp giải thích cái tên ma trận chuyển cơ sở.

Tính chất 6.41. *Với mỗi vector $u \in U$, ta có: $M_{S \to T}[u]_S = [u]_T$.*

Nhận xét 6.42. Ma trận chuyển cơ sở cung cấp một phương pháp rất nhanh cho máy tính để tìm biểu diễn của một vector qua một cơ sở bất kỳ. Giả sử cần biểu diễn $\mathbb{R}^n$ qua cơ sở $S = \{v_1, v_2, \cdots, v_n\}$, chỉ cần biểu diễn $e_1, e_2, \cdots, e_n$ qua S, sau đó có thể biểu diễn mọi $v \in \mathbb{R}^n$ qua S chỉ với một phép nhân ma trận, thay vì phải giải hệ phương trình tuyến tính.

Tính chất tiếp theo cung cấp một phương pháp tính ngược lại $M_{T \to S}$ sau khi đã có $M_{S \to T}$.

Tính chất 6.43. *Với hai cơ sở S và T của cùng một không gian V, $M_{S\to T}$ luôn khả nghịch và $(M_{S\to T})^{-1} = M_{T\to S}$.*

Ví dụ 6.44. Cho $E = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ là cơ sở chính quy của $\mathbb{R}^3$ và $S = \left\{ \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\}$ là một cơ sở khác của $\mathbb{R}^3$, khi đó:

$$M_{E\to S} = \begin{bmatrix} [e_1]_S & [e_2]_S & [e_3]_S \end{bmatrix} = \frac{1}{9}\begin{bmatrix} -2 & 1 & 4 \\ 4 & -2 & 1 \\ 1 & 4 & -2 \end{bmatrix}.$$

Biểu diễn của vector $v = \begin{bmatrix} 9 & -18 & 27 \end{bmatrix}^T$ qua S sẽ là:

$$\frac{1}{9}\begin{bmatrix} -2 & 1 & 4 \\ 4 & -2 & 1 \\ 1 & 4 & -2 \end{bmatrix} \times \begin{bmatrix} 9 \\ -18 \\ 27 \end{bmatrix} = \begin{bmatrix} 8 \\ 11 \\ -13 \end{bmatrix}.$$

6.6 Siêu phẳng trong $\mathbb{R}^n$

Đại số tuyến tính bắt nguồn từ việc giải các phương trình tuyến tính có dạng

$$a_1x_1 + \cdots + a_nx_n = c. \tag{1}$$

Mỗi phương trình tuyến tính trong $\mathbb{R}^n$ tương ứng với một siêu phẳng. Đường thẳng trong mặt phẳng tọa độ, mặt phẳng trong không gian 3 chiều là các ví dụ về siêu phẳng. Hiểu một cách đại số, siêu phẳng trong không gian $\mathbb{R}^n$ là nghiệm của một phương trình tuyến tính. Chúng ta cũng có thể viết lại (1) dưới dạng $\vec{a}\cdot\vec{x} = c$ để có định nghĩa sau.

Định nghĩa 6.45 (Siêu phẳng)**.** Cho vector $\vec{a} \in \mathbb{R}^n$, $\vec{a} \neq \vec{0}$ và $c \in \mathbb{R}$. Tập nghiệm H của phương trình tuyến tính

$$\vec{a}\cdot\vec{x} = c,$$

được gọi là một **siêu phẳng** trong $\mathbb{R}^n$

$$H = \{\vec{x} \in \mathbb{R}^n \mid \vec{a}\cdot\vec{x} = c\}.$$

Nhận xét 6.46. Cho siêu phẳng $H : \vec{a}\cdot\vec{x} = c$. Khi đó:

1. $\vec{a}$ được gọi là *vector pháp tuyến* của H. Hai siêu phẳng được gọi là *song song* nếu như vector pháp tuyến của chúng song song với nhau.
2. Cố định $x_0 \in H$, phương trình xác định H có thể được viết lại như sau

   $$\vec{a}\cdot(\vec{x} - \vec{x_0}) = 0.$$

   Như vậy mọi siêu phẳng đều có thể được *tịnh tiến*[5] thành một siêu phẳng đi qua gốc tọa độ và song song với nó.
3. Giải hệ phương trình tuyến tính tương đương với tìm miền giao của các siêu phẳng.

5 Phép biến hình trong không gian biến $\vec{x}$ thành $\vec{x} + \vec{c}$.

Ví dụ 6.47. Giao của hai mặt phẳng $H_1 : x + y - z = 0$ và $H_2 : x + y + z = 2$ trong $\mathbb{R}^3$ là đường thẳng $(t, 1 - t, 1)$.

Nhận xét 6.48. Một siêu phẳng chia không gian $\mathbb{R}^n$ ra làm ba miền $\vec{a} \cdot \vec{x} > 0$, $\vec{a} \cdot \vec{x} = 0$, và $\vec{a} \cdot \vec{x} < 0$. Như vậy tích vô hướng còn có thể dùng để xác định vị trí một điểm đối với một siêu phẳng. Tính chất này xuất hiện trong bài toán phân loại tuyến tính, cụ thể là thuật toán **Perceptron** và **Support Vector Machine**.

Nhận xét 6.49. Các siêu phẳng trong không gian $\mathbb{R}^n$ là dạng tổng quát của mặt phẳng trong không gian $\mathbb{R}^3$ và đường thẳng trong $\mathbb{R}^2$. Siêu phẳng đi qua gốc tọa độ là một dạng **không gian con** của **không gian vector** $\mathbb{R}^n$.

6.7 Cơ sở trực chuẩn và hình chiếu lên không gian con

Chúng ta đã bàn về việc một cơ sở thích hợp rất hữu ích cho việc tính toán trên không gian vector. Loại cơ sở thường được dùng nhất là cơ sở trực chuẩn, gồm các vector đơn vị vuông góc đôi một. Cơ sở trực chuẩn giúp việc tính toán biểu diễn của một vector và tích vô hướng giữa các vector dễ dàng hơn.

Cho không gian vector $V \subseteq \mathbb{R}^n$ và một cơ sở $S = \{\vec{v}_1, \dots, \vec{v}_k\}$.

Định nghĩa 6.50. Cơ sở S được gọi là **trực giao** nếu $\vec{v}_i \cdot \vec{v}_j = 0, \quad \forall i \neq j \in \overline{1, k}$.

Định nghĩa 6.51 (Cơ sở trực chuẩn)**.** Một cơ sở của V được gọi là **trực chuẩn** nếu nó trực giao và chỉ gồm các vector đơn vị (tức độ dài bằng 1).

Ví dụ 6.52.

1. Cơ sở chính quy của $\mathbb{R}^n$ là một cơ sở trực chuẩn của $\mathbb{R}^n$.
2. $S = \left\{ \left(\frac{1}{2}, \frac{\sqrt{3}}{2} \right), \left(\frac{-\sqrt{3}}{2}, \frac{1}{2} \right) \right\}$ là một cơ sở trực chuẩn của $\mathbb{R}^2$.

Liệu một không gian con bất kỳ của $\mathbb{R}^n$ có cơ sở trực chuẩn? Làm sao để tìm một cơ sở trực chuẩn nếu có? Thuật toán Gram - Schmidt sau đi tìm cơ sở trực chuẩn $S = \{\vec{e}_1, \dots \vec{e}_k\}$ của V thông qua một cơ sở $T = \{\vec{v}_1, \dots, \vec{v}_k\}$ cho trước.

Phép trực giao Gram - Schmidt

Bước 1: $\vec{u}_1 := \vec{v}_1,$ $\qquad \vec{e}_1 = \frac{u_1}{\|\vec{u}_1\|}$

Bước 2: $\vec{u}_2 := \vec{v}_2 - \mathrm{proj}_{\vec{u}_1}(\vec{v}_2),$ $\qquad \vec{e}_2 := \frac{\vec{u}_2}{\|\vec{u}_2\|}$

Bước 3: $\vec{u}_3 := \vec{v}_3 - \left[\mathrm{proj}_{\vec{u}_1}(\vec{v}_3) + \mathrm{proj}_{\vec{u}_2}(\vec{v}_3)\right],$ $\qquad \vec{e}_3 := \frac{\vec{u}_3}{\|\vec{u}_3\|}$

$\vdots \qquad \vdots$

Bước k: $\vec{u}_k := \vec{v}_k - \left[\mathrm{proj}_{\vec{u}_1}(\vec{v}_k) + \mathrm{proj}_{\vec{u}_2}(\vec{v}_k) + \cdots + \mathrm{proj}_{\vec{u_{k-1}}}(\vec{v}_k)\right],$ $\qquad \vec{e}_k := \frac{\vec{u}_k}{\|\vec{u}_k\|}.$

Khi đó $S = \{\vec{e}_1, \dots \vec{e}_k\}$ là cơ sở trực chuẩn của V.

Ví dụ 6.53. Giả sử $V \subseteq \mathbb{R}^4$ được xác định bởi $V = \mathrm{Span}(\{\vec{v_1}, \vec{v_2}, \vec{v_3}\})$ trong đó:

$$\begin{aligned}\vec{v_1} &= (-7,0,0,0),\\ \vec{v_2} &= (2,-3,0,0),\\ \vec{v_3} &= (1,-2,0,3).\end{aligned}$$

1. Đặt $\vec{u_1} := \vec{v_1} = (-7,0,0,0)$. Chuẩn hóa $\vec{u_1}$, ta có:

$$\vec{e_1} := \frac{\vec{u_1}}{\|\vec{u_1}\|} = (-1,0,0,0).$$

2. Đặt

$$\begin{aligned}\vec{u_2} &:= \vec{v_2} - \mathrm{proj}_{\vec{u_1}}(\vec{v_2})\\ &= \vec{v_2} - \frac{\vec{u_1}\cdot\vec{v_2}}{\vec{u_1}\cdot\vec{u_1}}\vec{u_1}\\ &= (2,-3,0,0) - (2,0,0,0)\\ &= (0,-3,0,0).\end{aligned}$$

Chuẩn hóa u_2, ta được:

$$\vec{e_2} = \frac{\vec{u_2}}{\|\vec{u_2}\|} = (0,-1,0,0).$$

3. Đặt

$$\begin{aligned}\vec{u_3} &:= \vec{v_3} - \mathrm{proj}_{\vec{u_1}}(\vec{v_3}) - \mathrm{proj}_{\vec{u_2}}(\vec{v_3})\\ &= \vec{v_3} - \frac{\vec{u_1}\cdot\vec{v_3}}{\vec{u_1}\cdot\vec{u_1}}\vec{u_1} - \frac{\vec{u_2}\cdot\vec{v_3}}{\vec{u_2}\cdot\vec{u_2}}\vec{u_2}\\ &= (1,-2,0,3) - (1,0,0,0) - (0,-2,0,0)\\ &= (0,0,0,3).\end{aligned}$$

Chuẩn hóa $\vec{u_3}$, ta được:

$$\vec{e_3} := \frac{u_3}{\|u_3\|} = (0,0,0,1).$$

Vậy $S = \{(-1,0,0,0),(0,-1,0,0),(0,0,0,1)\}$ là một cơ sở trực chuẩn của V.

Tính đúng đắn của thuật toán Gram - Schmidt được thể hiện qua định lý sau.

Định lý 6.54. *Cho tập hợp k vector độc lập tuyến tính $T = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}\}$. Gọi $\vec{u_i}, \vec{e_i}$ là kết quả bước thứ i của thuật toán Gram - Schmidt lên T. Khi đó:*

1. *Các tập hợp $\{\vec{u_i}\}_{i=1}^k$ và $\{\vec{e_i}\}_{i=1}^k$ độc lập tuyến tính.*
2. *$\{\vec{u_i}\}_{i=1}^k$ là một cơ sở trực giao và $\{\vec{e_i}\}_{i=1}^k$ là một cơ sở trực chuẩn của span(T).*

Chứng minh. Xem bài tập 6.14. □

Điều này cũng trả lời được câu hỏi về sự tồn tại của cơ sở trực chuẩn.

Hệ quả 6.55. Mọi không gian con $V \subseteq \mathbb{R}^n$ đều có một cơ sở trực chuẩn.

Việc sử dụng các cơ sở trực chuẩn giúp việc tìm biểu diễn của vector thuận tiện hơn rất nhiều. Điều này được thể hiện qua các định lý sau:

Định lý 6.56 (Biểu diễn trong cơ sở trực giao)**.** *Cho vector $\vec{v} \in V$ và $S = \{\vec{v}_1, \ldots, \vec{v}_k\}$ là một cơ sở trực giao của V. Ta có*

$$\vec{v} = \mathrm{proj}_{\vec{v}_1}(\vec{v}) + \mathrm{proj}_{\vec{v}_2}(\vec{v}) + \cdots + \mathrm{proj}_{\vec{v}_k}(\vec{v}).$$

Nhớ lại công thức hình chiếu ở Tính chất 1.18: $\mathrm{proj}_{\vec{y}}(\vec{x}) = \left(\dfrac{\vec{x}\cdot\vec{y}}{\vec{y}\cdot\vec{y}}\right)\vec{y}$, ta có định lý sau cho cơ sở trực chuẩn:

Định lý 6.57 (Biểu diễn trong cơ sở trực chuẩn)**.** *Cho vector $\vec{v} \in V$ và một cơ sở trực giao $S = \{\vec{v}_1, \ldots, \vec{v}_k\}$ của V. Khi đó, ta có*

$$\vec{v} = (\vec{v}\cdot\vec{v}_1)\vec{v}_1 + (\vec{v}\cdot\vec{v}_2)\vec{v}_2 + \cdots + (\vec{v}\cdot\vec{v}_k)\vec{v}_k.$$

Ví dụ 6.58. Với cơ sở trực chuẩn $E = \{e_1, e_2, \cdots, e_n\}$ của $\mathbb{R}^n$, mọi vector $v = \left[x_1, x_2, \cdots, x_n\right]^T \in \mathbb{R}^n$ đều có biểu diễn hiển nhiên:

$$v = x_1 e_1 + x_2 e_2 + \cdots x_n e_n.$$

Đồng thời, với mọi $i = 1, 2, \cdots, n$, $x_i = v \cdot e_i$. Như vậy định lý 6.57 là đúng trong trường hợp này.

Định nghĩa 6.59 (Hình chiếu lên không gian con)**.** Cho không gian con $V \subset \mathbb{R}^n$ và vector $\vec{v} \in \mathbb{R}^n$. **Hình chiếu** của $\vec{v}$ lên V là vector $\vec{u} \in V$ sao cho

$$(\vec{v} - \vec{u}) \in V^{\perp}.$$

Tính chất 6.60.

1. *$\vec{u}$ xác định duy nhất,*
2. $(\vec{v} - \vec{u}) \cdot \vec{u} = 0$,
3. *$\vec{u}$ là nghiệm của bài toán cực trị:* $\min_{\vec{x}\in V} \|\vec{x} - \vec{v}\|$.

Chứng minh. Chúng ta chứng minh sự tồn tại và duy nhất.

V là một không gian con của $\mathbb{R}^n$ nên có số chiều hữu hạn. Giả sự số chiều của V là k và $S = \{\vec{u}_1, \ldots, \vec{u}_k\}$ là một cơ sở trực chuẩn của V. Đặt

$$\vec{u} = \mathrm{proj}_{\vec{u}_1}(\vec{v}) + \cdots + \mathrm{proj}_{\vec{u}_k}(\vec{v}).$$

Ta dễ dàng kiểm tra $\vec{u}$ thỏa mãn định nghĩa hình chiếu. Nếu tồn tại $\vec{w}$ cũng thỏa mãn định nghĩa về hình chiếu thì $\vec{u} - \vec{u}' \in V$ và $\vec{u} - \vec{w} \in V^{\perp}$. Như vậy $\vec{u} - \vec{w} \in V \cap V^{\perp}$ hay $\vec{u} - \vec{w} = \vec{0}$.

□

Bài tập

Bài tập 6.8.

(a) Chứng minh rằng khoảng cách từ gốc tọa độ O đến $H : \vec{a} \cdot \vec{x} = c$ là $\dfrac{|c|}{\|a\|}$.

(b) Tìm công thức tính khoảng cách từ vector $\vec{y} = (y_1, \ldots, y_n)$ đến siêu phẳng $a_1x_1 + \cdots + a_nx_n = c$ trong $\mathbb{R}^n$.

Bài tập 6.9. Tìm hình chiếu của $(1, 2, 3)$ lên siêu phẳng $H : x + 2y + 3z = 0$ trong $\mathbb{R}^3$. Chứng minh rằng mọi vector trong $\mathbb{R}^3$ đều có thể biểu diễn được dưới dạng $c(1, 2, 3) + \vec{v}$ với $c \in \mathbb{R}$, $\vec{v} \in H$.

Bài tập 6.10. (Tổng quát 1.5) Cho siêu phẳng $H : \vec{a} \cdot \vec{x} = 0 \subset \mathbb{R}^n$. Chứng minh rằng mọi vector đều có thể biểu diễn được dưới dạng $c\vec{a} + \vec{b}$ với $\vec{b} \in H$, $c \in \mathbb{R}$.

Bài tập 6.11. Tìm hình chiếu của $(1, 2, 3)$ lên mặt phẳng $x - 2y + z = 0$ và $2x - y + z = 0$ trong $\mathbb{R}^3$.

Bài tập 6.12. Tìm một cơ sở trực chuẩn cho không gian vector

$$\mathrm{Span}\{(1, 2, 4, 0), (0, 1, 1, 0), (0, 3, 1, 4)\}.$$

Bài tập 6.13. Chứng minh định lý 6.56 và định lý 6.57.
Gợi ý: S là một cơ sở của V nên tồn tại $a_1, \ldots, a_k$ sao cho $\vec{v} = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k$, hãy tìm a_i bằng cách xét giá trị của $\vec{v} \cdot \vec{v}_i$.

Bài tập 6.14. Chứng minh Định lý 6.54.

7 Ánh xạ tuyến tính

Như đã đề cập, đại số tuyến tính hiện đại là công cụ để nghiên cứu hình học tổng quát. Trong quá trình tổng quát hóa từ mặt phẳng 2 chiều lên không gian vector, các phép biến hình trở thành các ánh xạ tuyến tính. Việc nghiên cứu các ánh xạ tuyến tính là một nhiệm vụ quan trọng trong hình học cao cấp. Phần này sẽ giới thiệu ánh xạ tuyến tính và các tính chất cơ bản, đồng thời làm rõ mối liên hệ của chúng với các ma trận (như đã đề cập trong phần 2.1).

7.1 Ánh xạ tuyến tính và ma trận biểu diễn

Định nghĩa 7.1. Cho U, V là các không gian vector trên cùng một trường $\mathbb{F}$. Ánh xạ $f : U \to V$ là một **ánh xạ tuyến tính** nếu f thỏa mãn các tính chất:

1. $\forall u_1, u_2 \in U : f(u_1 + u_2) = f(u_1) + f(u_2)$.
2. $\forall u \in U, \alpha \in \mathbb{F} : f(\alpha u) = \alpha f(u)$.

Ánh xạ tuyến tính là công cụ quan trọng để nghiên cứu các mối tương quan giữa những không gian vector. Điều tiện lợi là trong khuôn khổ bài giảng này, chúng ta chỉ quan tâm đến các không gian vector **hữu hạn chiều** như $\mathbb{R}^n$, và khi số chiều là hữu hạn, ta có thể biểu diễn mỗi ánh xạ tuyến tính bằng một ma trận.

Định nghĩa 7.2 (Ma trận biểu diễn)**.** Cho $S = \{v_1, v_2, \ldots, v_m\}$ là một cơ sở của $\mathbb{R}^m$ và $T = \{u_1, u_2, \ldots, u_n\}$ là một cơ sở của $\mathbb{R}^n$. Cho $f : \mathbb{R}^m \to \mathbb{R}^n$ là một ánh xạ tuyến tính. Ma trận biểu diễn của f từ S sang T là ma trận sau:

$$R_{S\to T}(f) = \begin{bmatrix}[f(v_1)]_T & [f(v_2)]_T & \cdots & [f(v_m)]_T\end{bmatrix} \in \mathbb{R}^{n\times m}.$$

Định lý 7.3. *Với mọi cơ sở S, T của $\mathbb{R}^m, \mathbb{R}^n$, mọi ma trận $A \in \mathbb{R}^{n\times m}$, có một ánh xạ tuyến tính $f : \mathbb{R}^m \to \mathbb{R}^n$ sao cho $A = R_{S\to T}(f)$.*

Từ định nghĩa của ánh xạ tuyến tính, ta dễ dàng có các tính chất sau.

Tính chất 7.4. *Với ánh xạ tuyến tính $f : \mathbb{R}^m \to \mathbb{R}^n$ và cơ sở $S = \{v_1, v_2, \cdots, v_m\}$ của $\mathbb{R}^m$, ta có:*

$$\forall v \in \mathbb{R}^m : [f(v)]_T = R_{S\to T}(f)[v]_S.$$

Với $[v]_S$ là biểu diễn của v qua cơ sở S, $[f(v)]_T$ là biểu diễn của $f(v)$ qua cơ sở T.

Tính chất 7.5. *Cho S và T lần lượt là các cơ sở của $\mathbb{R}^m$ và $\mathbb{R}^n$. Với các ánh xạ tuyến tính $f, g : \mathbb{R}^m \to \mathbb{R}^n$, các số thực α, β, ta có:*

1. *$\alpha f + \beta g : \mathbb{R}^m \to \mathbb{R}^n$ là một ánh xạ tuyến tính.*
2. *$R_{S\to T}(\alpha f + \beta g) = \alpha R_{S\to T}(f) + \beta R_{S\to T}(g)$.*

Tính chất 7.6. *Cho M, N, K lần lượt là các cơ sở của $\mathbb{R}^m$, $\mathbb{R}^n$, $\mathbb{R}^k$. Với các ánh xạ tuyến tính $f : \mathbb{R}^m \to \mathbb{R}^n$ và $g : \mathbb{R}^k \to \mathbb{R}^m$, các cơ sở S của $\mathbb{R}^m$ và T của $\mathbb{R}^k$, ta có:*

1. *$f \circ g : \mathbb{R}^k \to \mathbb{R}^n$ là một ánh xạ tuyến tính.*
2. *$R_{K\to N}(f \circ g) = R_{M\to N}(f)R_{K\to M}(g)$.*

Như vậy với một cơ sở cho trước của một không gian vector hữu hạn chiều, mỗi ánh xạ tuyến tính đều có thể biểu diễn bằng một ma trận tương ứng với cơ sở này, và phép hợp hàm số giữa hai ánh xạ tuyến tính thực chất chính là phép nhân ma trận.

7.2 Tập ảnh và nhân của ánh xạ tuyến tính

Hai đối tượng thường được nghiên cứu của các ánh xạ nói chung là tập ảnh và tập hợp "nghiệm", hay còn gọi là "không điểm", tức tập hợp những điểm có ảnh bằng 0. Đối với ánh xạ tuyến tính, tập hợp không điểm thường được gọi là **nhân**.

Định nghĩa 7.7 (Tập ảnh và nhân)**. Tập ảnh** và **nhân** của một ánh xạ tuyến tính $f : U \to V$ lần lượt là các tập hợp sau:

1. $\mathrm{Im}(f) = \{f(u) : u \in U\}$.
2. $\mathrm{Ker}(f) = \{u \in U : f(u) = 0\} = f^{-1}(\{0\})$.

Dễ dàng nhận thấy $\mathrm{Im}(f) \subset V$ và $\mathrm{Ker}(f) \subset U$.

Nhận xét 7.8. Danh pháp tiếng Anh của tập ảnh là **Image** và nhân là **Kernel**, do đó mới có các ký hiệu Im và Ker.

Điều đặc biệt với các ánh xạ tuyến tính là cả tập ảnh và nhân đều là các không gian vector. Điều này được thể hiện qua mệnh đề sau:

Mệnh đề 7.9. *Cho U, V là các không gian vector và $f : U \to V$ là một ánh xạ tuyến tính. Khi đó $\mathrm{Im}(f)$ là một không gian con của V còn $\mathrm{Ker}(f)$ là một không gian con của U.*

Chứng minh. Dựa vào tính chất của không gian con ở định nghĩa 6.22, ta có:

1. Với mọi $v_1, v_2 \in \mathrm{Im}(f)$, $\exists u_1, u_2 \in U : v_1 = f(u_1), v_2 = f(u_2)$. Khi đó $v_1 + v_2 = f(u_1 + u_2)$ theo định nghĩa 7.1. Mà $u_1 + u_2 \in U$, nên $v_1 + v_2 \in \mathrm{Im}(f)$.

   Với mọi $a \in \mathbb{R}$, $v \in \mathrm{Im}(f)$, $\exists u \in U : v = f(u)$. Khi đó $av = af(u) = f(au)$ theo định nghĩa 7.1. Mà $au \in U$, nên $av \in \mathrm{Im}(f)$. Theo định nghĩa 6.22, $\mathrm{Im}(f)$ là không gian con của V.

2. Nếu $u_1, u_2 \in \mathrm{Ker}(f)$, tức là $f(u_1) = f(u_2) = 0$, thì $f(u_1 + u_2) = 0 \implies u_1 + u_2 \in \mathrm{Ker}(f)$. Tương tự $au \in \mathrm{Ker}(f)$ với $a \in \mathbb{R}$ và $u \in U$. Do đó $\mathrm{Ker}(f)$ là không gian con của U theo định nghĩa 6.22.

□

Trong trường hợp f là ánh xạ tuyến tính từ R^m vào R^n, ta có thể đồng nhất f với ma trận biểu diễn của nó (theo một cơ sở cho trước). Một cách thú vị, trong cách chuyển ngữ từ ánh xạ sang ma trận này, tập ảnh và nhân của f cũng được dịch thành các không gian con cơ bản của ma trận biểu diễn, như đã đề cập ở phần 6.3.

Tính chất 7.10. *Cho ánh xạ tuyến tính $f : \mathbb{R}^m \to \mathbb{R}^n$ với ma trận biểu diễn $A := A_{f,S}$ qua cơ sở S. Khi đó $Im(f) = C(A)$, không gian cột của A và $Ker(f) = N(A)$, không gian hạch của A.*

Chứng minh. Suy trực tiếp từ tính chất 7.4 của ma trận biểu diễn A và định nghĩa không gian cột, không gian hạch ở định nghĩa 6.18. ☐

Các tính chất sau cho biết khi nào một ánh xạ tuyến tính trên $\mathbb{R}^n$ là đơn ánh, toàn ánh hoặc song ánh, dựa vào các tính chất của ma trận biểu diễn.

Định lý 7.11. *Cho ánh xạ tuyến tính $f : \mathbb{R}^m \to \mathbb{R}^n$ và 2 cơ sở S và T tương ứng của $\mathbb{R}^m$ và $\mathbb{R}^n$. Khi đó ta có:*

1. $\dim(Im(f)) = \operatorname{rank}(R_{S\to T}(f)), \dim(Ker(f)) = \operatorname{null}(R_{S\to T}(f))$.
2. *f là đơn ánh* $\iff \dim(Ker(f)) = 0 \iff Ker(f) = \{0\}$.
3. *Nếu $m > n$ thì f chắc chắn không phải đơn ánh. f là toàn ánh* $\iff \dim(Im(f)) = n$.
4. *Nếu $m < n$ thì f chắc chắn không phải toàn ánh.*
5. *Nếu $m = n$, f là song ánh $\iff R_{S\to T}(f)$ khả nghịch. Khi đó ánh xạ nghịch đảo $f^{-1} : \mathbb{R}^n \to \mathbb{R}^m$ cũng là song ánh tuyến tính và $R_{T\to S}(f^{-1}) = (R_{S\to T}(f))^{-1}$.*

7.3 Toán tử tuyến tính - Ma trận đồng dạng

Trong phần này, ta quan tâm tới các ánh xạ tuyến tính đi từ $\mathbb{R}^n$ và chính nó. Có thể thấy, các ánh xạ này có ma trận biểu diễn là các *ma trận vuông*, một trong những đối tượng được quan tâm nhiều trong ĐSTT (Phần 5). Việc nghiên cứu các ánh xạ tuyến tính từ $\mathbb{R}^n$ đến $\mathbb{R}^n$ do đó cũng mở ra một hướng tiếp cận mới tới các ma trận vuông.

Định nghĩa 7.12 (Toán tử tuyến tính)**.** Một **toán tử tuyến tính** trong không gian $\mathbb{R}^n$ là một ánh xạ tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$.

Khi làm việc với các toán tử tuyến tính, ta thường sử dụng một cơ sở duy nhất cho $\mathbb{R}^n$, thay vì 2 cơ sở khác nhau để biểu diễn $\vec{v}$ và $f(\vec{v})$. Ta định nghĩa khái niệm **ma trận biểu diễn** cho toán tử tuyến tính như sau:

Định nghĩa 7.13. Cho một cơ sở $S = \{v_1, v_2, \ldots, v_n\}$ của $\mathbb{R}^n$ và toán tử tuyến tính $\mathbb{R}^n \to \mathbb{R}^n$. **Ma trận biểu diễn** của f qua S, ký hiệu $A_S(f)$, là ma trận biểu diễn của f, như một ánh xạ tuyến tính, từ S vào S.

$$A_S(f) = R_{S\to S}(f) = \begin{bmatrix} [f(v_1)]_S & [f(v_2)]_S & \cdots & [f(v_n)]_S \end{bmatrix}.$$

Như một hệ quả của định lý 7.3, ta có:

Định lý 7.14. *Với mọi cơ sở S của $\mathbb{R}^n$ và ma trận $A \in \mathbb{R}^{n\times n}$, tồn tại toán tử tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$ sao cho $A = A_S(f)$.*

Mọi tính chất của ma trận biểu diễn của ánh xạ tuyến tính cũng đúng cho ma trận biểu diễn của toán tử tuyến tính. Thêm vào đó, ta có tính chất đặc trưng sau cho toán tử tuyến tính:

Định lý 7.15. *Cho toán tử tuyến tình f trên $\mathbb{R}^n$, các điều sau là tương đương:*

- ⋆ *$Ker(f) = \{0\}$ (f là đơn ánh).*

- ⋆ *$Im(f) = \mathbb{R}^n$ (f là toàn ánh).*
- ⋆ *f là song ánh.*
- ⋆ *$A_S(f)$ khả nghịch.*

Nếu f là song ánh, f^{-1} cũng là toán tử tuyến tính trên $\mathbb{R}^n$ và có ma trận biểu diễn $A_S(f^{-1}) = (A_S(f))^{-1}$.

Từ định nghĩa trên, có thể thấy ma trận biểu diễn của một toán tử tuyến tính không chỉ phụ thuộc vào toán tử đó, mà còn vào cơ sở đang xét. Một câu hỏi tự nhiên được đặt ra là ma trận biểu diễn thay đổi thế nào nếu thay đổi cơ sở? Khái niệm sau thể hiện mối quan hệ giữa các ma trận biểu diễn của cùng một toán tử tuyến tính dưới các cơ sở khác nhau.

Định nghĩa 7.16 (Ma trận đồng dạng)**.** Hai ma trận vuông thuộc $\mathbb{R}^{n\times n}$, A_1 và A_2 **đồng dạng** với nhau nếu tồn tại một toán tử tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$ và 2 cơ sở S_1, S_2 của $\mathbb{R}^n$ sao cho $A_1 = A_{S_1}(f)$ và $A_2 = A_{S_2}(f)$.

Định lý quan trọng sau làm rõ bản chất của quan hệ đồng dạng:

Định lý 7.17. *Hai ma trận $A, B \in \mathbb{R}^{n\times n}$ đồng dạng khi và chỉ khi tồn tại một ma trận $P \in \mathbb{R}^{n\times n}$ khả nghịch sao cho:*

$$A = PBP^{-1}.$$

Bạn đọc sẽ được chứng minh định lý trên và hiểu mối liên hệ giữa nó với định nghĩa ma trận đồng dạng và việc chuyển cơ sở qua Bài tập 7.5.

Sneak peek. Hai ma trận đồng dạng sẽ có cùng các *trị riêng* và bội của chúng. Phần 8 sẽ đi sâu vào phần này, cụ thể là bài toán tìm ma trận chéo đồng dạng với một ma trận cho trước. Những kiến thức trong phần 8 là nền tảng cho phương pháp xử lý dữ liệu **Principle Component Analysis**.

Đọc thêm. Trong hình học cao cấp, nếu coi các điểm là các vector, thì các ánh xạ chính là các phép biến hình. Có khá nhiều phép biến hình có tính tuyến tính, điển hình như phép chiếu vuông góc lên một không gian con. Các phép biến hình sau còn là các toán tử tuyến tính: phép đối xứng qua một trục đi qua 0, phép đối xứng qua gốc 0, phép vị tự gốc 0, phép quay gốc 0. Một số bài tập trong chương này sẽ dùng các tính chất của ĐSTT để chứng minh các tính chất đơn giản trong hình học thông qua các ma trận biểu diễn, và ngược lại.

Bài tập

Bài tập 7.1. Hãy chứng minh các định lý/tính chất phía trên.

Bài tập 7.2. Cho trước cơ sở $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}\}$ của $\mathbb{R}^n$. Hãy tìm ma trận biểu diễ qua S của các phép biến hình sau:

1. Phép đồng nhất $\mathrm{Id} : \vec{v} \mapsto \vec{v}\ \forall v \in \mathbb{R}^n$.
2. Phép vị tự $\vec{v} \mapsto c\vec{v}\ \forall v \in \mathbb{R}^n$.
3. Phép đối xứng trục qua đường thẳng $\{a\vec{u} : a \in \mathbb{R}\}$.

Bài tập 7.3. Chứng minh ma trận biểu diễn qua một cơ sở bất kỳ của phép quay góc θ trong mặt phẳng $\mathbb{R}^2$ là

$$\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Bài tập 7.4. Chứng minh rằng trong $\mathbb{R}^2$, phép vị tự và phép quay *giao hoán* với nhau (có thể thay đổi thứ tự thực hiện mà không ảnh hưởng đến kết quả).

Bài tập 7.5. Cho cơ sở $S = \{s_1, s_2, \cdots, s_n\}$ và $T = \{t_1, t_2, \cdots, t_n\}$ của $\mathbb{R}^n$. Chứng minh với mọi toán tử tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$, ta có:

$$A_T(f) = M_{S\to T} A_S(f) M_{S\to T}^{-1}.$$

Bài tập 7.6. Chứng minh rằng nếu A và B đồng dạng thì $\mathrm{rank}(A) = \mathrm{B}$ và $\mathrm{null}(A) = \mathrm{null}(B)$.

8 Giá trị riêng, Vector riêng và Đa thức đặc trưng

8.1 Bài toán lũy thừa nhanh ma trận

Như đã giới thiệu ở phần trên, không gian các ma trận cũng có những phép tính cơ bản như không gian của số thực: cộng, trừ, nhân (chỉ riêng phép tính chia chỉ dành cho ma trận khả nghịch). Tuy vậy, vẫn có một điểm khác biệt nho nhỏ, nếu như phép cộng và trừ có thể được thực hiện một cách dễ dàng, phép nhân lại phức tạp hơn một chút. Một trường hợp đặc biệt của phép nhân đó chính là lấy số mũ của một ma trận. Giả sử ta có một ma trận A với kích thước 2×2 như sau

$$A = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}.$$

Một câu hỏi cơ bản là làm sao ta có thể tính nhanh A^2, A^3, hay thậm chí A^{100}?

$$A^2 = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 4 \cdot 1 & 1 \cdot 4 + 4 \cdot 1 \\ 1 \cdot 1 + 1 \cdot 1 & 1 \cdot 4 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ 2 & 5 \end{bmatrix},$$

$$A^3 = A^2 \cdot A = \begin{bmatrix} 5 & 8 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 \cdot 1 + 8 \cdot 1 & 5 \cdot 4 + 8 \cdot 1 \\ 2 \cdot 1 + 5 \cdot 1 & 2 \cdot 4 + 5 \cdot 1 \end{bmatrix} = \begin{bmatrix} 13 & 28 \\ 7 & 13 \end{bmatrix}.$$

Với số mũ nhỏ, việc áp dụng định nghĩa phép nhân để tính ra A^2 và A^3 cho ngay kết quả khá gọn gàng; nhưng với A^{100} ta có thể thử và nhận ra rằng việc nhân 100 lần là không khả thi. Đây mới chỉ là một ví dụ về ma trận 2×2, với ma trận kích thước lớn hơn, việc nhân "từ từ" chắn chắn sẽ không đem lại hiểu quả cao. Từ đây nảy sinh một vấn đề cần được chúng ta giải quyết, với ma trận A có những tính chất nào thì phép tính luỹ thừa có thể được thực hiện dễ dàng?

Một suy luận tự nhiên cho thấy nếu A có nhiều hệ số 0 thì việc nhân với chính A sẽ rút gọn đi rất nhiều. Sau một lúc suy nghĩ, ta xem xét ma trận A với kích thước $n \times n$ chỉ có hệ số trên đường chéo chính như sau

$$A = \begin{bmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & a_n \end{bmatrix}.$$

Và ta có thể thấy ngay rằng, để lấy số mũ của A ta chỉ cần luỹ thừa từng hệ số với số mũ tương ứng, như sau

$$A^k = \begin{bmatrix} a_1^k & 0 & 0 & 0 \\ 0 & a_2^k & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & a_n^k \end{bmatrix}.$$

Vậy là nếu ma trận chỉ có hệ số trên đường chéo, phép luỹ thừa rất là đơn giản! Gọi những ma trận như thế là *ma trận chéo*. Từ đây ta nảy ra một ý tưởng, **liệu có cách nào có thể biến đổi một ma trận A bất kỳ về một ma trận chéo không**? Hay nói cách khác, giả sử ta có ma trận A, liệu có tồn tại hai ma trận B và D trong đó D là ma trận chéo sao cho

$$A = BD?$$

Nhưng ta nhận thấy việc lấy luỹ thừa của BD thực ra không đơn giản hơn việc lấy luỹ thừa của A một chút nào, là vì phép nhân ma trận không có tính giao hoán nên

$$(BD)^k = (BD)(BD)\cdots(BD) \quad (k \text{ lần}),$$

ta không thể gộp D lại để lấy luỹ thừa được. Từ đây, ta nghĩ ra một cách để triệt tiêu B, đó là viết A dưới dạng sau

$$A = BDB^{-1},$$

trong đó B^{-1} là nghịch đảo của B. Ta thấy rằng, mặc dù phép nhân không giao hoán nhưng lại có tính kết hợp,

$$\begin{aligned} A^k &= (BDB^{-1})(BDB^{-1})(BDB^{-1})\cdots(BDB^{-1}) \quad (k \text{ lần}) \\ &= BD(B^{-1}B)D(B^{-1}B)D(B^{-1}B)\cdots(B^{-1}B)DB^{-1} \\ &= BDD\cdots DB^{-1} \\ &= BD^kB^{-1}. \end{aligned}$$

Vì ma trận D là chéo nên việc tính D^k rất đơn giản, và như vậy việc tính A^k trở thành việc nhân ba ma trận B, D^k và B^{-1}. Ta có thể viết $A = BDB^{-1}$ lại thành

$$AB = BD.$$

Giả sử B viết thành các vector cột như sau $B = [v_1\ v_2\ \ldots\ v_n]$ và giả sử các hệ số trên đường chéo của D là $\lambda_1, \lambda_2, \ldots, \lambda_n$, biểu thức trên trở thành

$$A[v_1\ v_2\ \ldots\ v_n] = [v_1\ v_2\ \ldots\ v_n] \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix},$$

hay tương đương

$$[Av_1\ Av_2\ \ldots\ Av_n] = [\lambda_1 v_1\ \lambda_2 v_2\ \ldots\ \lambda_n v_n].$$

Như vậy sự tồn tại của B và D tương đương với sự tồn tại của các vector $v_1, v_2, \ldots, v_n$ và các giá trị $\lambda_1, \lambda_2, \ldots, \lambda_n$ thoả mãn

$$Av_k = \lambda_k v_k, \quad k = 1, 2, \ldots, n.$$

Đây chính là động lực để ta định nghĩa hai khái niệm quan trọng nhất của đại số tuyến tính **giá trị riêng** và **vector riêng**. Trước khi đi vào định nghĩa, quay lại với ví dụ phía trên của ma trận vuông A kích cỡ 2×2. Bạn đọc có thể kiểm chứng đẳng thức sau

$$A = \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -1/4 & 1/2 \\ 1/4 & 1/2 \end{bmatrix},$$

trong đó $B = \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$ và $B^{-1} = \begin{bmatrix} -1/4 & 1/2 \\ 1/4 & 1/2 \end{bmatrix}$. Như vậy, với k bất kỳ

$$A^k = \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (-1)^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} -1/4 & 1/2 \\ 1/4 & 1/2 \end{bmatrix} = \begin{bmatrix} (3^k + (-1)^k)/2 & 3^k - (-1)^k \\ (3^k - (-1)^k)/4 & (3^k + (-1)^k)/2 \end{bmatrix}.$$

Thử với $k = 2$ và $k = 3$ cho ta kết quả giống như ở phía trên.

8.2 Giá trị riêng và vector riêng

Định nghĩa 8.1 (Trị riêng)**.** Cho $f : \mathbb{R}^n \to \mathbb{R}^n$ là một ánh xạ tuyến tính. Ta nói $\lambda \in \mathbb{C}$ là một **trị riêng** (hoặc **giá trị riêng**) của f nếu tồn tại $v \in \mathbb{R}^n \setminus \{0\}$ sao cho:

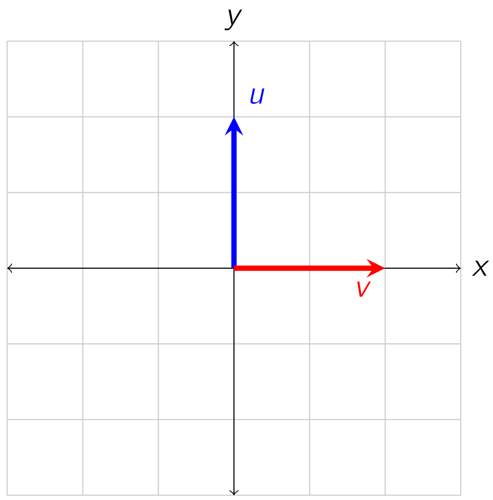
$$f(v) = \lambda v.$$

Lúc đó ta nói v là **vector riêng** ứng với trị riêng λ.

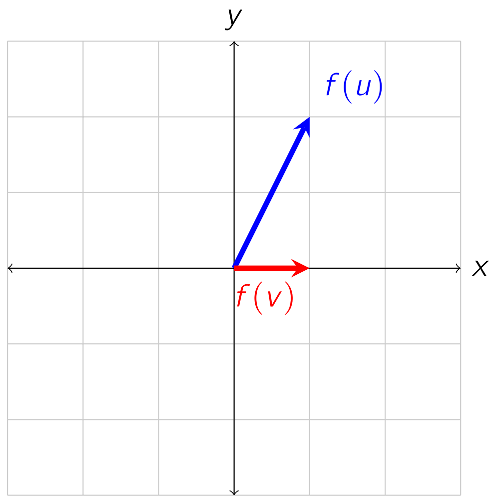
Nhận xét 8.2. Ở định nghĩa trên $v \neq 0$ nhưng λ có thể bằng 0. Khi đó tập tất cả các vector riêng ứng với $\lambda = 0$ chính là $\mathrm{Ker}(f) \setminus \{0\}$.

Thêm vào đó, λ có thể *không phải số thực*. Lý do sẽ được đưa ra trong phần tiếp theo, đa thức đặc trưng.

Một cách tổng quát, tập các vector riêng ứng với trị riêng λ là $\mathrm{Ker}(f - \lambda Id) \setminus \{0\}$.



Trước tác động của f



Sau tác động của f

Ví dụ 8.3.

⋆ Ở hình trên, sau tác động của f,
$$f(v) = \frac{1}{2}v,$$
nên v là một vector riêng ứng với trị riêng $\frac{1}{2}$. Tuy nhiên u không phải vector riêng của f.

⋆ Trong mặt phẳng xét f là một phép quay một góc θ quanh tâm O. Giả sử $\theta \neq k\pi$ với $k \in \mathbb{Z}$ thì mọi vector khác không đều đổi phương, đồng thời cũng không có vector khác 0 nào biến thành 0. Vậy f không có vector riêng.

Ứng với mỗi $\lambda \in \mathbb{R}$, ta đặt $E(\lambda) = \{v \in \mathbb{R}^n \mid f(v) = \lambda v\}$.

Nhận xét 8.4.

⋆ $E(\lambda) = \{v \in \mathbb{R}^n \mid (f - \lambda \mathrm{Id})(v) = 0\} = \mathrm{Ker}(f - \lambda \mathrm{Id})$ là một không gian con của $\mathbb{R}^n$.

⋆ Nếu $E(\lambda) \neq \{0\}$ thì λ chính là một trị riêng của f. Lúc đó ta ký hiệu $E(\lambda) := \mathrm{Eigen}(\lambda)$, gọi là **không gian con riêng** của $\mathbb{R}^n$ ứng với λ.

Ví dụ 8.5.

1. Giả sử $v_\lambda \in \mathbb{R}^n \setminus \{0\}$ là một vector riêng của ánh xạ tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$. Khi đó $\mathrm{Span}(\{v_\lambda\})$ là không gian con một chiều của $\mathbb{R}^n$ (gọi là **đường thẳng**). Lúc đó $\mathrm{Span}(\{v_\lambda\})$ bất biến dưới tác động của f.

2. Xét ánh xạ tuyến tính:

$$\begin{array}{rccc} f & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & \begin{bmatrix}1\\0\end{bmatrix} & \longmapsto & \begin{bmatrix}2\\1\end{bmatrix} \\ & & \begin{bmatrix}0\\1\end{bmatrix} & \longmapsto & \begin{bmatrix}6\\3\end{bmatrix} \end{array}.$$

   Ta sẽ đi tìm các trị riêng cũng như các vector ứng với trị riêng của ánh xạ này:

$$\begin{aligned} f(X) = f\left(\begin{bmatrix}x\\y\end{bmatrix}\right) &= xf\left(\begin{bmatrix}1\\0\end{bmatrix}\right) + yf\left(\begin{bmatrix}0\\1\end{bmatrix}\right) \\ &= \begin{bmatrix}2 & 6\\1 & 3\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} \\ &= A_f X, \end{aligned}$$

   trong đó $A_f = \begin{bmatrix}2 & 6\\1 & 3\end{bmatrix}$ chính là ma trận biểu diễn của f (trên cơ sở chính tắc).

   Ta tìm $X \neq 0$ và λ sao cho $f(X) = \lambda X$, tức là:

$$\begin{aligned} & f(X) &&= \lambda X, \\ \iff & A_f X &&= \lambda I_2 X, \\ \iff & (A_f - \lambda I_2)X &&= 0. \end{aligned}$$

   Do $X \neq 0$ nên $\det(A_f - \lambda I_2) = 0 \iff \lambda(\lambda - 5) = 0$. Khi đó $\lambda = 0$ và $\lambda = 5$ là hai trị riêng của f.

   Ngoài ra, $\mathrm{Eigen}(0) = \mathrm{Ker}(f) = \mathrm{Span}\left(\begin{bmatrix}-3\\1\end{bmatrix}\right)$ và $\mathrm{Eigen}(5) = \mathrm{Ker}(f - 5Id) = \mathrm{Span}\left(\begin{bmatrix}2\\1\end{bmatrix}\right)$.

Từ đó ta có định lý tổng quát sau:

Định lý 8.6. *Cho toán tử tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$ và $A_S(f)$ là ma trận biểu diễn f qua cơ sở S bất kỳ. Khi đó $\lambda \in \mathbb{R}$ là một trị riêng của f khi và chỉ khi $\det(A_S(f) - \lambda I_n) = 0$.*

Nếu A là một ma trận biểu diễn f qua một cơ sở S nào đó, khi đó với mọi $\lambda \in \mathbb{R}$,

$$\begin{aligned} \det(A_S(f) - \lambda I_n) &= \det\left[M_{E\to S}A_S(f)M_{S\to E} - \lambda I_n\right] \\ &= \det\left[M_{E\to S}\cdot(A_E(f) - \lambda I_n)\cdot M_{S\to E}\right] \\ &= \det\left[M_{E\to S}M_{S\to E}\right]\cdot\det(A_E(f) - \lambda I_n) \\ &= \det(A_E(f) - \lambda I_n). \end{aligned}$$

với E là cơ sở chính tắc.

Như vậy $\det(A_S(f) - \lambda I_n)$ không phụ thuộc vào cách chọn cơ sở S mà phụ thuộc vào f.

Định nghĩa 8.7. Ánh xạ $P_f : \mathbb{R} \to \mathbb{R}$ thỏa mãn $P_f(\lambda) = \det(A_E(f) - \lambda I_n), \quad \forall \lambda \in \mathbb{R}$ được gọi là **đa thức đặc trưng** của f.

Do cách định nghĩa định thức, P_f thực chất là một đa thức bậc n.

Nhận xét 8.8. Tập hợp các trị riêng của f thực chất là tập nghiệm của đa thức đặc trưng của nó. Ta biết một đa thức bậc n luôn có n nghiệm thuộc $\mathbb{C}$ (tính cả bội) và không nhất thiết là số thực. Điều này có nghĩa các trị riêng của một ánh xạ tuyến tính có thể không thuộc $\mathbb{R}$.

Ở dạng biến đổi tuyến tính, ánh xạ f còn có thể xem như một ma trận A kích thước $n \times n$ và các vector riêng xem như là một ma trận kích thước $n \times 1$; lúc đó các khái niệm về trị riêng, vector riêng được định nghĩa tương tự, Id đồng nhất với I_n.

Định nghĩa 8.9 (Trị riêng và vector riêng của ma trận)**.** Cho $A \in \mathbb{R}^{n \times n}$. Ta nói $\lambda \in \mathbb{C}$ là một **trị riêng** (hoặc **giá trị riêng**) của A nếu tồn tại $v \in \mathbb{R}^n \setminus \{0\}$ sao cho:

$$Av = \lambda v.$$

Lúc đó ta nói v là **vector riêng** ứng với trị riêng λ.

Đa thức đặc trưng của A được ký hiệu $P_A : \mathbb{R} \to \mathbb{R}$ sao cho $P_A(x) = \det(A - xI_n)$.

Tính chất 8.10. *Với mọi ma trận $A, B \in \mathbb{R}^{n \times n}$, ta có:*

- ⋆ *Nếu A và B đồng dạng (cùng biểu diễn ánh xạ tuyến tính f ở hai cơ sở khác nhau $\Longleftrightarrow$ tồn tại P khả nghịch sao cho $A = PBP^{-1}$) thì $P_A \equiv P_B \equiv P_f$ và chúng có cùng trị riêng (cùng bội số).*
- ⋆ *A khả nghịch khi và chỉ khi $P_A(0) \neq 0$.*
- ⋆ *Ma trận thực đối xứng cấp n có n trị riêng thực.*
- ⋆ *A và A^T có cùng trị riêng (cùng bội).*
- ⋆ *AB và BA có cùng tập các trị riêng (có thể khác bội).*

Xác định trị riêng, vector riêng của một ma trận

- ⋆ **Bước 1.** Xác định đa thức đặc trưng P_A.
- ⋆ **Bước 2.** Giải phương trình $P_A(\lambda) = 0$ để tìm các nghiệm riêng.
- ⋆ **Bước 3.** Ứng với mỗi trị riêng λ, giải phương trình:

$$(A - \lambda I_n)x = 0,$$

ta được $\mathrm{Eigen}(\lambda)$.

8.3 Chéo hóa ma trận

Định nghĩa 8.11. Một ma trận $A \in \mathbb{R}^n$ gọi là **chéo hóa được** trong $\mathbb{R}$ (trong $\mathbb{C}$) nếu tồn tại ma trận đường chéo D_A thuộc $\mathbb{R}^{n\times n}$ (thuộc $\mathbb{C}^{n\times n}$) sao cho A và D_A đồng dạng.
Nghĩa là tồn tại $P \in \mathbb{R}^{n\times n}$ ($P \in \mathbb{C}^{n\times n}$) khả nghịch sao cho $D_A = PAP^{-1}$.

Định lý 8.12. *Cho ánh xạ tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$ và A_f là một ma trận biểu diễn f qua một cơ sở bất kỳ. Khi đó A chéo hóa được trong $\mathbb{R}$ (trong $\mathbb{C}$) khi và chỉ khi tồn tại một cơ sở β của $\mathbb{R}^n$ (của $\mathbb{C}^n$) chứa toàn vector riêng của f.*
Ngoài ra ma trận đường chéo D_A sẽ chứa các trị riêng nằm trên đường chéo.

Qua đó ta thấy không phải ma trận nào cũng có thể chéo hóa được, ngay cả trong $\mathbb{C}$, ta có một định lý mô tả khả năng chéo hóa của một ma trận như sau:

Định lý 8.13. *Cho $A \in \mathbb{R}^{n\times n}$ và $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ là các trị riêng phân biệt của A và giả sử $m_i \in \mathbb{N}$ là số ẩn tự do trong nghiệm của hệ $(A - \lambda_i I_n)x = 0$ ($m_i = \mathrm{Null}(A - \lambda_i I_n)$).*
Khi đó A chéo hóa được khi và chỉ khi $m_1 + \dots + m_k = n$.

Hệ quả 8.14.

⋆ Ma trận vuông cấp n chéo hóa được khi và chỉ khi nó có n vector riêng độc lập tuyến tính.

⋆ Nếu một ma trận vuông cấp n có n trị riêng phân biệt thì nó chéo hóa được.

⋆ Ma trận đối xứng chéo hóa được.

8.4 Ứng dụng

Giả sử ma trận $A \in \mathbb{C}^{n\times n}$ chéo hóa được, nghĩa là tồn tại P khả nghịch sao cho:

$$\begin{aligned} P^{-1}AP &= D_A \\ &= \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}. \end{aligned}$$

Tính lũy thừa của ma trận

Khi đó,

$$\begin{aligned} A^k &= PD_AP^{-1} \\ &= PD_A^kP^{-1} \\ &= P \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix} P^{-1}. \end{aligned}$$

Tìm công thức tổng quát dãy số truy hồi

Giả sử (u_n) và (v_n) là hai dãy số thực thỏa mãn công thức truy hồi:

$$\begin{cases} u_{n+1} = u_n - v_n, \\ v_{n+1} = 2u_n + 4v_n. \end{cases}$$

Trong đó $u_0 = 2$ và $v_0 = 1$.
Nếu ta đặt $X_n = \begin{bmatrix} u_n \\ v_n \end{bmatrix}$ và $A = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$ thì công thức truy hồi có thể viết lại thành:

$$X_{n+1} = AX_n, \quad \forall n \in \mathbb{N} \cup \{0\},$$

trong đó $X_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Từ đó suy ra $X_n = A^n X_0$, từ ý tưởng câu (1), ta tính được A^n tổng quát:

$$\begin{aligned} X_n &= \begin{bmatrix} 2^{n+1} - 3^n & 2^n - 3^n \\ -2^{n+1} + 2.3^n & -2^n + 2.3^n \end{bmatrix} X_0 \\ &= \begin{bmatrix} 5.2^n - 3^{n+1} \\ -5.2^n + 6.3^n \end{bmatrix}. \end{aligned}$$

Tìm cực trị của hàm bậc 2

Định nghĩa 8.15. Hàm bậc 2 n biến có dạng sau, trong đó a_i và b_{ij} là các hằng số.

$$f : \mathbb{R}^n \to \mathbb{R}, \quad (x_1, x_2, \ldots, x_n) \mapsto \sum_{i=1}^{n} a_i x_i^2 + 2\sum_{i<j} b_{ij} x_i x_j.$$

Bài toán tìm cực trị của hàm bậc 2 là bài toán trong đó cho trước hàm 2 $f(\vec{x})$, tìm

$$\min f(\vec{x}) \quad \text{và} \quad \max f(\vec{x}) \quad \text{thỏa điều kiện } \|\vec{x}\| = 1.$$

Để tìm cực trị của hàm bậc 2, ta có định lý tổng quát sau:

Định lý 8.16. *Cho hàm bậc 2 $f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i^2 + 2\sum_{i<j} b_{ij} x_i x_j$. Gọi M_f là ma trận sau:*

$$M_f = \begin{bmatrix} a_1 & b_{12} & b_{13} & \cdots & b_{1n} \\ b_{12} & a_2 & b_{23} & \cdots & b_{2n} \\ b_{13} & b_{23} & a_3 & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{1n} & b_{2n} & b_{3n} & \cdots & a_n \end{bmatrix} = \left(m_{ij}\right)_{i,j=1}^{n}, \quad \text{trong đó } m_{ij} = \begin{cases} a_i & \text{nếu } i = j \\ b_{ij} & \text{nếu } i < j \\ b_{ji} & \text{nếu } i > j \end{cases}.$$

Khi đó M_f là một ma trận đối xứng, tức là tập L các trị riêng của M_f gồm toàn các số thực, và:

$$\min\{f(\vec{x}) \mid \|\vec{x}\| = 1\} = \min L, \quad \max\{f(\vec{x}) \mid \|\vec{x}\| = 1\} = \max L.$$

Ví dụ 8.17. Tìm cực trị của biểu thức $f(x, y) = 4x^2 + 3xy$ với điều kiện $x^2 + y^2 = 1$.

Lời giải. Vector $[x, y]^T$ luôn có norm bằng 1, do đó đây là bài toán tìm cực trị bậc 2. Ta có:

$$M_f = \begin{bmatrix} 4 & 3/2 \\ 3/2 & 0 \end{bmatrix}.$$

M_f có các trị riêng là $\{-1/2, 9/2\}$. Do đó

$$\min\left\{f(x,y) \mid x^2+y^2=1\right\} = -\frac{1}{2}, \quad \max\left\{f(x,y) \mid x^2+y^2=1\right\} = \frac{9}{2}.$$

Kiểm tra lại, ta có:

$$\begin{aligned} 4x^2+3xy &= \frac{9}{2}x^2+3xy+\frac{1}{2}y^2-\frac{1}{2}(x^2+y^2) &&= \frac{1}{2}(3x+y)^2-\frac{1}{2} \geq -\frac{1}{2} \\ &= \frac{9}{2}(x^2+y^2)-\left(\frac{1}{2}x^2-3xy+\frac{9}{2}y^2\right) &&= \frac{9}{2}-\frac{1}{2}(x-3y)^2 \leq \frac{9}{2}. \end{aligned}$$

Vì $f\left(-\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right) = -\frac{1}{2}$, $f\left(\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right) = \frac{9}{2}$, ta kết luận được 2 giá trị cực tiểu và cực đại như trên.

□

Bài tập

Bài tập 8.1. Hãy tìm tất cả trị riêng, cơ sở của không gian con riêng và xác định sự chéo hoá cũng như ma trận chéo của mỗi ma trận vuông dưới đây:

a) $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$.

b) $\begin{bmatrix} -2 & -4 & 2 \\ -2 & 1 & 2 \\ 4 & 2 & 5 \end{bmatrix}$.

c) $\begin{bmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{bmatrix}$.

d) $\begin{bmatrix} 0 & 0 & 2 \\ -3 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}$.

e) $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$.

f) $\begin{bmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{bmatrix}$.

g) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Bài tập 8.2. Chứng minh rằng các toán tử tuyến tính sau chéo hoá được trên $\mathbb{R}$ và tìm cơ sở mà trong đó toán tử có dạng chéo:

a) $f : \mathbb{R}^3 \to \mathbb{R}^3$ thoả mãn:

$$f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} -x-3y-3z \\ 3x+5y+3z \\ -x-y+z \end{bmatrix}.$$

b) $g : \mathbb{R}^4 \to \mathbb{R}^4$ thoả mãn:

$$g\left(\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}\right) = \begin{bmatrix} x+y+z+t \\ x+y-z-t \\ x-y+z-t \\ x-y-z+t \end{bmatrix}.$$

Bài tập 8.3. Cho ma trận

$$\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}.$$

a) A có chéo hoá được trên $\mathbb{R}$ không?

b) A có chéo hoá được trên $\mathbb{C}$ không?

Bài tập 8.4. Tìm công thức tổng quát của dãy số Fibonacci: $u_0 = u_1 = 1$, $u_{n+1} = u_n + u_{n-1}$.

Bài tập 8.5. Cho 2 số thực a, b thỏa $a^2 > 4b$. Tìm công thức tổng quát của dãy số truy hồi $u_{n+1} = au_n + bu_{n-1}$ theo a, b và u_0, u_1.

Bài tập 8.6. Tìm công thức tổng quát của dãy số $u_0 = 0, u_1 = 1, u_2 = 2$, $u_n = 6u_{n-1} - 11u_{n-2} + 6u_{n-3}$.

Bài tập 8.7. Hai công ty P và M là đối thủ cạnh tranh. Mỗi tháng, tỉ lệ khách hàng dùng sản phẩm của công ty P chuyển sang dùng của công ty M là p, với $0 < p < 1$; và tỉ lệ khách hàng của công ty M chuyển sang công ty P là q, với $0 < q < 1$. Tính tỷ lệ giữa lượng khách hàng của 2 công ty về lâu dài.

Bài tập 8.8. Chứng minh Tính chất 8.10.

Bài tập 8.9. Cho $A \in \mathbb{R}^{m\times n}$, chứng minh tập các trị riêng của AA^T và A^TA là như nhau.

Bài tập 8.10. Chứng minh nếu $A \in \mathbb{R}^{n\times n}$ chéo hóa được trong $\mathbb{C}$ và mọi trị riêng của A đều là thuộc $\mathbb{R}$, thì A chéo hóa được trong $\mathbb{R}$.

Bài tập 8.11. Chứng minh một ma trận vuông cấp n có n trị riêng phân biệt thì chéo hóa được.

Bài tập 8.12. Với mỗi ma trận $A = \left(a_{ij}\right) \in \mathbb{C}^{n\times n}$, định nghĩa ma trận *liên hợp* của A, ký hiệu $\overline{A}$ là ma trận gồm các phần tử tương ứng liên hợp với các phần tử của A:

$$\overline{A} = \left(\overline{a_{ij}}\right)_{i,j=1}^{n}.$$

Định nghĩa ma trận *liên hợp chuyển vị* của A, ký hiệu A^*, là

$$A^* = \left(\overline{A}\right)^T.$$

Ma trận A được gọi là *Hermitian* nếu $A = A^*$. Chứng minh rằng nếu $\lambda \in \mathbb{C}$ là một trị riêng của ma trận Hermitian A thì $\lambda \in \mathbb{R}$. Từ đó suy ra mọi trị riêng của ma trận đối xứng trong $\mathbb{R}^{n\times n}$ đều thuộc $\mathbb{R}$.

Bài tập 8.13. Tìm cực đại và cực tiểu của các biểu thức sau:

(a) $x^2 + y^2 + z^3 + xy + yz + zx$, biết $x^2 + 4y^2 + 9z^2 = 1$.

(b) $4x^2 + 2y^2 + z^3 + 2xy + 5yz + 6zx$, biết $2x^2 + 3y^3 + z^2 = 1$.

Bài tập 8.14. Chứng minh định lý 8.16 cho hàm bậc 2 hai biến.

Bài tập 8.15 (Khó). Chứng minh định lý 8.16. **Gợi ý:** Phương pháp nhân tử Lagrange.

Tài liệu

[1] Trinh Thanh Deo, Le Van Luyen, Bui Xuan Hai, and Tran Ngoc Hoi. *Đại số Tuyến tính và Ứng dụng (Tập 1)*. Nhà Xuất bản Đại học Quốc gia Tp.HCM, 2009.

[2] Nguyen Huu Viet Hung. *Đại số tuyến tính*. Nhà Xuất bản Đại học Quốc gia Hà Nội, 2001.