HỒ CÔNG XUÂN VŨ Ý

THIẾT THỰC - HIỆU QUẢ - HÀI HÒA

# Đại Số Tuyến Tính

Tiền Giang - 2013

# Đại Số Tuyến Tính

**Hồ Công Xuân Vũ Ý**

**Trường Đại Học Tiền Giang**

# Mục lục

# Chương I

# Logic, tập hợp, quan hệ, ánh xạ, cấu trúc đại số, phép thế, số phức và đa thức

## § 1 Logic

### 1.1 Mệnh đề và phép toán mệnh đề

**Mệnh đề đơn giản** là một khái niệm nguyên thủy không được định nghĩa, nhưng ta hiểu nó một cách nôm na: mỗi mệnh đề đơn giản là một câu, một khẳng định có duy nhất một trong hai tính chất là *đúng* hoặc *sai*. Đúng hay sai được thấy rõ từ nội dung của mệnh đề. Chẳng hạn, "$2 + 2 = 4$" là mệnh đề đúng và "$\pi$ là số hữu tỉ" là mệnh đề sai. Trong logic mệnh đề người ta không để ý đến ý nghĩa, nội dung và cấu trúc ngữ pháp của mệnh đề mà chỉ quan tâm đến tính chất đúng hoặc sai của nó. Một điều đáng lưu ý ở đây là trong toán học thường chúng ta chỉ quan tâm đến các mệnh đề đúng. Ta ký hiệu mệnh đề bằng các chữ in thường $p$, $q$, ...

Từ hai hay nhiều mệnh đề đơn giản, bằng những phép toán trên mệnh đề (sẽ được trình bày sau) ta xây dựng được những **mệnh đề phức tạp**, mà nó cũng chỉ có duy nhất một trong hai tính chất hoặc *đúng* hoặc *sai*. Hơn nữa tính chất đúng sai của mệnh đề phức tạp chỉ phụ thuộc vào tính đúng sai của các mệnh đề đơn giản tạo nên nó mà không phụ thuộc vào nội dung của chúng.

Cho $p$ là một mệnh đề, ta gọi **phủ định** của $p$ là một mệnh đề, ký hiệu $\neg p$, *đúng* khi $p$ *sai* và *sai* khi $p$ *đúng*. Ví dụ: xét mệnh đề "$2 + 2 = 4$", khi đó $\neg(2 + 2 = 4)$ có nghĩa là "$2 + 2 \neq 4$" là mệnh đề sai; đặt $p$ là mệnh đề "$\pi$ là số hữu tỉ" và ta biết rằng $p$ là mệnh đề sai, khi ấy $\neg p$ được thể hiện "$\pi$ không là số hữu tỉ", đó là mệnh đề đúng.

Cho hai mệnh đề $p$ và $q$. **Hội** của $p$ và $q$, ký hiệu $p \wedge q$, là một mệnh đề đúng khi cả hai $p$ và $q$ đúng, và sai trong các trường hợp khác. Ta có thể thấy $(2 + 2 = 4) \wedge (\pi$ là số hữu tỉ$)$ là mệnh đề sai, và $(2 + 2 = 4) \wedge (\pi$ không là số hữu tỉ$)$ là mệnh đề đúng.

Cho hai mệnh đề $p$ và $q$. **Tuyển** của $p$ và $q$, ký hiệu $p \vee q$ đôi khi được đọc là $p$ "hoặc" $q$, là một mệnh đề sai khi cả hai $p$ và $q$ sai, và đúng trong các trường hợp khác. Chú ý rằng chữ "hoặc" trong "$p$ hoặc $q$" theo nghĩa *không loại trừ* có nghĩa là nếu $p$ hoặc $q$ đúng thì ít nhất một trong hai mệnh đề $p$ và $q$ đúng. Ta có thể thấy $(2 + 2 = 4) \vee (\pi$ là số hữu tỉ$)$ là mệnh đề đúng.

Cho hai mệnh đề $p$ và $q$. Ta gọi $p$ **kéo theo** $q$ là mệnh đề, ký hiệu $p \to q$, mà nó sai khi $p$ đúng và $q$ sai và đúng trong các trường hợp khác. Ta có thể thấy $(2 + 2 = 4) \to (\pi$ là số hữu tỉ$)$ là mệnh đề sai, và $(\pi$ là số hữu tỉ$) \to (2 + 2 = 4)$ là mệnh đề đúng.

Khi mệnh đề $p$ kéo theo mệnh đề $q$ thì ta nói $q$ là một *điều kiện cần* (ắt có) cho $p$. Thay cách nói này ta cũng có thể nói

- ▷ $p$ đúng *chỉ nếu* (hoặc *chỉ khi*) $q$ là đúng
- ▷ Để $p$ là đúng, *cần là* (ắt có) $q$ là đúng
- ▷ *Điều kiện cần* (ắt có) để $p$ là đúng là $q$ là đúng.

Khi mệnh đề $p$ kéo theo mệnh đề $q$, ta cũng nói $p$ là một *điều kiện đủ* cho $q$. Thay vì cách nói này ta còn nói

- ▷ $q$ là đúng *nếu* (hoặc *khi*) $q$ là đúng.
- ▷ Để $q$ là đúng, *đủ* là $p$ là đúng.
- ▷ *Điều kiện đủ* để $q$ đúng là $p$ là đúng.

Trong toán học, các định lý thường có dạng $p \to q$, khi đó $p$ là giả thiết và $q$ là kết luận. Hơn nữa, định lý là mệnh đề đúng, cho nên nếu ta biết được $p$ đúng thì ta thu được $q$ đúng.

Cho hai mệnh đề $p$ và $q$. Ta gọi $p$ **tương đương** $q$ là mệnh đề, ký hiệu $p \leftrightarrow q$, mà nó đúng khi $p$ và $q$ đều đúng hoặc đều sai, và sai trong tất cả các trường hợp còn lại. Ta thấy $(2 + 2 = 4) \leftrightarrow (\pi$ là số hữu tỉ$)$ là mệnh đề sai.

**Ghi chú:** Đôi khi vì thói quen hay có tính chất truyền thống hoặc lịch sử hay tiện trình bày, mà định nghĩa được phát biểu dưới dạng mệnh đề $p \to q$. Nhưng chúng ta phải hiểu là nó ở dạng $p \leftrightarrow q$.

Khi mệnh đề $p$ tương đương với mệnh đề $q$, ta cũng nói

▷ $p$ là đúng *khi và chỉ khi* (hay *nếu và chỉ nếu*) $q$ là đúng.

▷ *Điều kiện cần và đủ* để $p$ là đúng là $q$ là đúng.

Hai mệnh đề phức tạp $P$ và $Q$ được gọi là **tương đương logic** nếu và chỉ nếu chúng có cùng một giá trị chân lý, đối với mọi hệ giá trị của các mệnh đề tạo thành chúng; nói cách khác, mệnh đề tương đương $P \leftrightarrow Q$ luôn đúng với mọi hệ giá trị của các mệnh đề tạo thành chúng. Khi đó chúng ta ký hiệu $P \Leftrightarrow Q$.

Sau đây chúng tôi nêu ra một số luật logic (tương đương logic) thông dụng trong mệnh đề sau

**1.1.1 Mệnh đề.** *Cho $p$, $q$, và $r$ là các mệnh đề. Khi đó chúng ta có*

**1)** $\neg(\neg p) \iff p$

**2)** *Luật giao hoán*

$$p \wedge q \iff q \wedge p, \qquad p \vee q \iff q \vee p.$$

**3)** *Luật kết hợp*

$$(p \wedge q) \wedge r \iff p \wedge (q \wedge r), \quad (p \vee q) \vee r \iff p \vee (q \vee r).$$

**4)** *Luật phân phối*

$$p \wedge (q \vee r) \iff (p \wedge q) \vee (p \wedge r),$$
$$p \vee (q \wedge r) \iff (p \vee q) \wedge (p \vee r).$$

**5)** *Luật De Morgan**

$$\neg(p \wedge q) \iff \neg p \vee \neg q \qquad \neg(p \vee q) \iff \neg p \wedge \neg q.$$

**6)** $p \Rightarrow q \iff \neg p \vee q$

**7)** $p \Rightarrow q \iff \neg q \Rightarrow \neg p.$

## 1.2 Vị từ

Để xây dựng khái niệm *vị từ*, chúng ta cần sử dụng ngôn ngữ tập hợp. Mặc dù khái niệm *tập hợp* và *phần tử* sẽ được trình bày kỷ ở mục sau, nhưng chúng tôi xin nêu một cách tượng trưng các khái niệm ấy. Chúng ta có thể hiểu như sau *tập hợp* qui tụ các đối tượng được phân biệt rõ ràng, và các đối tượng đó gọi là *phần tử*. Để ký hiệu $x$ là phần tử của tập hợp $A$, chúng ta viết $x \in A$.

**Vị từ** $P$ được xác định trên tập hợp $A$ là một khẳng định sao cho với mỗi phần tử $x \in A$ ta có thể xác định được điều khẳng định đó hoặc đúng hoặc sai (nói cách khác $P$ trở thành một mệnh đề với mỗi phần tử của $A$). Nếu ứng phần tử $x$ của $A$ vị từ $P$ là mệnh đề đúng ta viết $P(x)$. Chẳng hạn, "$x \leq 1$" là một vị từ trên tập số thực $\mathbb{R}$. Chúng ta thấy nếu $x$ có giá trị 0, vị từ trở thành mệnh đề đúng "$0 \leq 1$". Nếu $x$ có giá trị 2, vị từ trở thành mệnh đề sai "$2 \leq 1$".

Như đã đề cập ở trên, trong toán học thường chúng ta quan tâm đến mệnh đề đúng, cho nên chúng ta cũng thường quan tâm những phần tử nào làm cho vị từ đúng mà thôi. Chẳng hạn như thí dụ trên khi viết vị từ "$x \leq 1$", chúng ta đặc biệt lưu ý đến những giá trị của số thực nhỏ hơn hoặc bằng 1 bởi vì chúng làm cho vị từ trở nên mệnh đề đúng.

*Augustus De Morgan, (1806-1871) Nhà toán học và logic học Anh. Ông cùng với George Boole đặt nền tảng cho logic ký hiệu hiện đại

Cho vị từ $P$ xác định trên tập $A$. Nếu với mỗi $x \in A$ ta đều có $P(x)$, thì ta có khái niệm **lượng từ toàn thể** hay **lượng từ phổ biến** và được ký hiệu

$$\forall x \in A : P(x) \qquad \text{hay gọn hơn} \qquad \forall x[P(x)]$$

đọc là "với mọi $x \in X$, $P(x)$". Chú ý rằng $\forall x[P(x)]$ là một mệnh đề, nhưng trong định nghĩa này đây là mệnh đề đúng. Ví dụ xét vị từ "$x^2 \geq 0$" trên $\mathbb{R}$. Ta dễ dàng thấy được $\forall x[x^2 \geq 0]$.

Cho vị từ $P$ xác định trên tập $A$. Nếu trong $A$ có phần tử $x$ sao cho $P(x)$, thì ta có khái niệm **lượng từ tồn tại** và được ký hiệu

$$\exists x \in A : P(x) \qquad \text{hay gọn hơn} \qquad \exists x[P(x)]$$

đọc là "tồn tại $x \in A$ để $P(x)$". Chú ý rằng $\exists x[P(x)]$ là một mệnh đề, nhưng trong định nghĩa này đây là mệnh đề đúng. Ví dụ: ta xét phương trình bậc hai $x^2 - 3x + 2 = 0$ trên tập số thực $\mathbb{R}$. Bởi vì $x = 1$ là một nghiệm của phương trình nên ta có thể viết $\exists x[x^2 - 3x + 2 = 0]$.

Chúng ta nêu hai tính chất cơ bản của lượng từ trong mệnh đề sau.

**1.2.1 Mệnh đề.** *Cho $P$ là một vị từ xác định trên tập $A$. Khi đó ta có*

$$\begin{aligned} \neg(\forall x[P(x)]) &\iff \exists x[\neg P(x)] \\ \neg(\exists x[P(x)]) &\iff \forall x[\neg P(x)] \end{aligned}$$

Chẳng hạn, ta xét các vị từ nêu trên, ta viết như sau

$$\begin{aligned} \neg(\forall x[x^2 \geq 0]) &\iff \exists x[x^2 < 0] \\ \neg(\exists x[x^2 - 3x + 2 = 0]) &\iff \forall x[x^2 - 3x + 2 \neq 0] \end{aligned}$$

Vị từ được xét cho đến lúc này là vị từ một biến. Bây giờ ta xét trường hợp vị từ hai biến. **Vị từ** $P$ được xác định trên tập hợp $A$ và $B$ là một khẳng định sao cho với mỗi phần tử $x \in A$ và $y \in B$ ta có $P$ trở thành mệnh đề. Để tránh nhầm lẫn ta thường viết vị từ hai biến $P$ là $P(x, y)$. Cũng như trong trường hợp vị từ một biến với $x_0 \in A$ và $y_0 \in B$ cụ thể và ta viết $P(x_0, y_0)$ mà không nói gì ta quy ước đó là mệnh đề đúng.

Từ vị từ hai biến $P(x, y)$ xác định trên $A$ và $B$ bằng cách lượng từ hóa với các lượng từ với mọi và tồn tại ta xây dựng được các vị từ một biến: $\exists x[P(x, y)]$ và $\forall x[P(x, y)]$ xác định trên $B$; và $\exists y[P(x, y)]$ và $\forall y[P(x, y)]$ xác định trên $A$. Nếu ta lượng từ hóa hai biến cho vị từ $P(x, y)$ sẽ thu được các mệnh đề: $(\forall x)(\forall y)[P(x, y)]$, $(\forall x)(\exists y)[P(x, y)]$, $(\exists x)(\forall y)[P(x, y)]$, $(\exists x)(\exists y)[P(x, y)]$. Chúng ta có tính chất của lượng từ hai ngôi trong mệnh đề sau.

**1.2.2 Mệnh đề.** *Cho $P$ là một vị từ hai ngôi xác định trên $A$ và $B$. Khi đó ta có*

*(i)* $\neg((\exists x)(\forall y)P(x, y)) \iff (\forall x)(\exists y)(\neg P(x, y))$

*(ii)* $\neg((\exists x)(\exists y)P(x, y)) \iff (\forall x)(\forall y)(\neg P(x, y))$

*(iii)* $\neg((\forall x)(\forall y)P(x, y)) \iff (\exists x)(\exists y)(\neg P(x, y))$

*(iv)* $\neg((\forall x)(\exists y)P(x, y)) \iff (\exists x)(\forall y)(\neg P(x, y))$

Tương tự, ta có trường hợp tổng quát cho vị tự $n$ ngôi và tính chất lượng từ cho vị từ ấy. Ta có quy tắc đơn giản và tiện lợi sau đây để phủ định các mệnh đề lượng từ hóa vị từ nhiều ngôi: *Muốn phủ định một mệnh đề lượng từ hóa vị từ nhiều ngôi, ta thay mỗi lượng từ $\exists$ bằng $\forall$ và ngược lại, đồng thời phủ định vị từ bên trong.*

Để kết thúc mục này chúng tôi xin trình bày một số phương pháp chứng minh thường gặp trong toán học.

**1.2.3 (Phép chứng minh bằng cách phân thành các trường hợp nhỏ)** Để chứng minh một mệnh đề (thường là một vị từ đi kèm với lượng từ phổ biến (vị từ này luôn nhận giá trị chân lý *đúng*) ta có thể xét nó trong các trường hợp có thể được. Ví dụ để chứng minh mệnh đề "$\forall n \in \mathbb{N} : P(n)$", ta có thể chứng tỏ rằng $P(n)$ là đúng với tất cả số chẵn, với tất cả số lẻ.

**1.2.4 (Phép chứng minh bằng phản chứng)** Ta cần chứng minh một mệnh đề (hay một kết luận) $P$ là đúng. Tạm thời chúng ta giả thiết ngược lại rằng $P$ là sai, và ta vận dụng các định nghĩa và các kết quả đã biết là đúng và bằng những suy luận logic để suy ra một điều trái với điều mà ta đã biết. Như vậy mâu thuẫn phát sinh là do ta đã giả thiết $P$ là sai. Để tránh điều mâu thuẫn này, ta phải có mệnh đề cần chứng minh $P$ là đúng.

# § 2 Tập hợp

## 2.1 Khái niệm

Trong lý thuyết tập hợp hiện đại, khái niệm **tập hợp**, **sự thuộc vào**, và **sự bằng nhau** là những *khái niệm nguyên thủy* không định nghĩa, tương tự như các khái niệm điểm, đường thẳng trong hình học. Nhưng để làm quen và hình dung khái niệm tập hợp, chúng tôi nêu ra định nghĩa về tập hợp của Georg Cantor*.

**2.1.1 (Định nghĩa tập hợp của Cantor, 1895)** Một **tập hợp** $A$ quy tụ các đối tượng được xác định và phân biệt bởi trực giác hay khả năng lập luận của chúng ta để nhận thức được sự khác biệt giữa chúng. Các đối tượng ấy được gọi là **phần tử** của $A$.

Các danh từ đồng nghĩa với tập hợp: họ, hệ, đám, quần thể, v.v.... Như vậy tập hợp được xác định bởi các phần tử của nó. Nghĩa là khi ta biết được tất cả phần tử của một tập hợp thì tập hợp ấy hoàn toàn được xác định. Mỗi tập hợp đều được cho bởi một quy luật, cho phép xác định một đối tượng nào đó có nằm trong tập hợp đó hay không. Sau đây là một vài ví dụ về tập hợp.

**2.1.2 Thí dụ.** ▷ Tập hợp học sinh trong một trường nào đó.

▷ Tập hợp các nghiệm của phương trình $x^2 - 3x + 2 = 0$.

▷ Tập hợp $\mathbb{N}$ mọi số tự nhiên.

▷ Tập hợp $\mathbb{Z}$ mọi số nguyên. □

*Georg Cantor (1845-1918), nhà toán học Đức sinh ở Nga, ông được coi là người sáng lập môn lý thuyết tập hợp

Các tập hợp thường được ký hiệu bằng các chữ in hoa $A$, $B$, $C$, ..., $X$, $Y$, $Z$, ..., còn các phần tử của một tập thường được ký hiệu bằng các chữ in thường $a$, $b$, $c$, ..., $x$, $y$, $z$, ...

Chúng ta thấy rằng 1 và 2 là nghiệm của phương trình $x^2-3x+2=0$. Vậy ta nói 1 và 2 là phần tử của tập hợp nghiệm của phương trình $x^2-3x+2=0$.

Để biểu thị $x$ là phần tử của tập $A$, ta viết $x \in A$, đọc là $x$ thuộc $A$. Để biểu thị $y$ không là phần tử của $A$, ta viết $y \notin A$, đọc là $y$ không thuộc $A$.

**2.1.3 Thí dụ.** ▷ $5 \in \mathbb{N}$ và $2005 \in \mathbb{N}$ nhưng $-20 \notin \mathbb{N}$.

▷ $-20 \in \mathbb{Z}$ nhưng $\frac{8}{3} \notin \mathbb{Z}$. □

**2.1.4 Định nghĩa.** Cho hai tập hợp $A$ và $B$. Nếu mọi phần tử của $A$ cũng là phần tử của $B$ và ngược lại, thì ta nói $A$ **bằng** $B$ và viết là $A = B$. Như vậy $A = B$ khi và chỉ khi

$$x \in A \iff x \in B.$$

Để kí hiệu hai tập hợp $A$ và $B$ *không* bằng nhau (*khác nhau*) ta viết $A \neq B$. Để chứng minh $A \neq B$, ta tìm một phần tử $a \in A$ sao cho $a \notin B$ hay ta tìm một phần tử $b \in B$ sao cho $b \notin A$.

Để chỉ rõ biểu diễn tất cả các phần tử thuộc vào một tập hợp chúng ta sẽ dùng cặp dấu móc { } (thường được gọi là cái thành lập tập hợp) theo hai cách: liệt kê danh sách tất cả các phần tử của tập hợp và đưa ra quy tắc đặc trưng các phần tử.

**2.1.5** *Liệt kê mọi phần tử* của tập hợp, khi có thể được. Chẳng hạn

$$A = \{a, b, c\}$$

là tập hợp ba chữ đầu của bảng chữ cái tiếng Việt. Khi viết như vậy chúng ta hiểu như sau: $a$, $b$, và $c$ là 3 phần tử của tập hợp $A$. Chúng ta đọc "tập hợp $A$ chứa 3 phần tử $a$, $b$, và $c$".

Những điều cần lưu ý khi liệt kê các phần tử của một tập hợp. Thứ nhất, thứ tự của các phần tử trong khi liệt kê không quan trọng, nghĩa là $\{1, 2, 3\} = \{3, 1, 2\}$. Thứ hai, khi liệt kê các phần tử ta không nên lặp lại phần tử đã được liệt kê, mặc dù $\{1, 2, 3\} = \{1, 2, 3, 2\}$ nhưng ta thấy kí hiệu đầu gọn và hợp lý hơn nên chúng ta luôn viết ở dạng đầu. Do đó ta qui ước khi biểu diễn tập hợp bằng cách liệt kê danh sách các phần tử thì các phần tử trong danh sách phải phân biệt nghia là không có phần tử nào xuất hiện nhiều hơn một lần.

**2.1.6 Thí dụ.** ▷ $\{1, 2\}$ là tập hợp các nghiệm của phương trình $x^2 - 3x + 2 = 0$ bởi vì ngoài 1 và 2 phương trình không nghiệm nào khác.

▷ Chúng ta có thể biểu diễn tập hợp số tự nhiên và số nguyên bằng cách liệt kê các phần tử như sau

$$\mathbb{N} = \{0, 1, 2, 3, 4 \ldots\}$$
$$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \pm 4, \ldots\}. \qquad \square$$

**2.1.7** *Chỉ ra một đặc tính đặc trưng $T$ cho các phần tử* của tập hợp. Nếu $x$ thuộc tập hợp $A$ thì $x$ có thính chất $T$ và $A$ chứa tất cả các phần tử có tính chất $T$ thì ta viết $A = \{x : T(x)\}$. Chẳng hạn

$$A = \{x \in \mathbb{R} : |x| \leq 2\}$$

là tập mọi số thực thỏa mãn tính chất $-2 \leq x \leq 2$, và ta đọc "$A$ là tập tất cả các số thực sao cho $|x| \leq 2$".

**2.1.8 Thí dụ.** ▷ Chúng ta có thể viết lại tập hợp nghiệm của phương trình $x^2 - 3x + 2 = 0$ như sau

$$\left\{x \in \mathbb{R} : x^2 - 3x + 2 = 0\right\}.$$

▷ Chúng ta có thể biểu diễn tập hợp số hữu tỉ bằng cách nêu lên tính chất đặc trưng như sau

$$\mathbb{Q} = \left\{\frac{m}{n} : n \in \mathbb{N}^+, m \in \mathbb{Z}\right\}. \qquad \square$$

**2.1.9 Thí dụ.** ▷ $\{x \in \mathbb{Z} : x^2 - 3x^2 + 2 = 0\} = \{1, 2\}$.

▷ $\{n \in \mathbb{N} : 1 \leq n \leq 6\} = \{1, 2, 3, 4, 5, 6\}$. □

Một tập hợp có thể chỉ gồm một phần tử duy nhất, trong trường hợp này nó được gọi là **đơn tử**. Cũng có khi ta phải xét các tập hợp mà chưa biết chắc nó có phần tử nào không, vì vậy điều hợp lý đưa vào khái niệm tập rỗng.

**2.1.10 Định nghĩa.** Một tập hợp không chứa bất cứ một phần tử nào được gọi là **tập rỗng**, và được kí hiệu là $\varnothing$.

**2.1.11 Thí dụ.** ▷ $\{x : x \neq x\} = \varnothing$.

▷ Chúng ta có $\varnothing \neq \{\varnothing\}$ bởi vì tập $\{\varnothing\}$ chứa một phần tử là tập rỗng trong khi đó tập hợp $\varnothing$ không chứa phần tử nào cả. □

Một trong những khái niệm quan trọng của lý thuyết tập hợp là khái niệm bao hàm. Khái niệm ấy nêu lên mối quan hệ giữa hai tập hợp.

**2.1.12 Định nghĩa.** Cho hai tập hợp $A$ và $B$. Nếu mỗi phần tử của $A$ cũng là phần tử của $B$ thì ta nói $A$ là **tập con** hay *bộ phận* của $B$ và kí hiệu là $A \subseteq B$ hay là $B \supseteq A$ (cũng có thể đọc là $A$ bị chứa trong $B$ hay $B$ chứa $A$).

Nếu tập $A$ không là tập con của tập $B$ thì được kí hiệu như sau $A \nsubseteq B$. Để chứng minh $A \nsubseteq B$ ta tìm một phần tử $a \in A$ sao cho $a \notin B$.

**2.1.13 Thí dụ.** ▷ Chúng ta có thể biểu diễn mối liên hệ giữa các tập hợp $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ như sau $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$.

▷ Xét tập hợp $\{a, \{a\}\}$. Ta thấy rằng $\{a\} \in \{a, \{a\}\}$ và $\{a\} \subseteq \{a, \{a\}\}$. Vậy tồn tại một tập hợp vừa là phần tử vừa là tập con của tập hợp khác.

▷ Tập hợp $A = \{1, 2, 3, 4\}$ không là tập con của tập hợp $B = \{x \in \mathbb{N} : x \text{ là số chẵn}\}$ bởi vì ta biết rằng $1 \in A$ là một số lẻ nên $1 \notin B$. □

Bởi vì tập rỗng không có chứa phần tử nào, nên chúng ta không có phần tử của tập rỗng để kiểm tra có thuộc tập hợp nào đó. Vì vậy tập rỗng là tập con của bất kì tập hợp nào; $\varnothing \subseteq A$ với bất cứ tập $A$ nào. Hơn nữa, cũng từ định nghĩa tập con, chúng ta dễ dàng chứng minh được định lý sau

**2.1.14 Định lý.** *Cho ba tập hợp bất kì $A$, $B$, và $C$. Chúng ta có*

**(1)** $A \subseteq A$.

**(2)** $A = B$ *khi và chỉ khi* $A \subseteq B$ *và* $B \subseteq A$.

**(3)** *Nếu* $A \subseteq B$ *và* $B \subseteq C$ *thì* $A \subseteq C$.

**2.1.15 Thí dụ.** Nếu $C$ là một tập con của tập rỗng $\varnothing$ thì $C = \varnothing$ (bởi vì ta luôn có $\varnothing \subseteq C$). □

**2.1.16 Định nghĩa.** Nếu $A$ là tập con của $B$ và $A \neq B$ thì ta nói $A$ là một **tập con thực sự** của $B$, và được kí hiệu là $A \subset B$. Trong khi đó $B$ và $\varnothing$ được gọi là **tập con tầm thường** của $B$.

Để chứng minh $A \subset B$, ta phải chứng minh $A$ là tập con của $B$ và tìm một phần tử $b \in B$ sao cho $b \notin A$.

**2.1.17 Thí dụ.** ▷ Qua Thí dụ 2.1.3, chúng ta có thể viết được $\mathbb{N} \subset \mathbb{Z}$.

▷ Cho $A_1 = \{x : 0 \leq x \leq 1\}$ và $A_2 = \{x : -1 \leq x \leq 2\}$. Ta thấy rằng $A_1 \subset A_2$.

▷ Xét tập $A = \{1, 2, 3, 4\}$. Ngoài $\varnothing$ và $A$ là các tập con tầm thường của $A$, chúng ta liệt kê các tập con khác của $A$: $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, $\{3, 4\}$, $\{1, 2, 3\}$, $\{1, 2, 4\}$, $\{1, 3, 4\}$, $\{2, 3, 4\}$. □

Thông thường trong phạm vi một môn học cụ thể hay trong vấn đề ta đang đề cập, các tập được nói đến đều là những tập con của một tập $\mathscr{U}$ cố định cho trước. Hơn nữa, việc nêu ra điều kiện này cho phép gạt bỏ được những khó khăn về mặt logic phát sinh trong quá trình xây dựng toán học. Vì vậy chúng tôi đưa ra khái niệm tập phổ dụng sau đây.

**2.1.18 Định nghĩa.** Nếu tất cả những tập hợp mà chúng ta đang đề cập xem xét là tập con của một tập hợp $\mathscr{U}$, khi đó $\mathscr{U}$ được gọi là **tập phổ dụng** hay **tập vạn năng** hay **không gian**.

Chẳng hạn trong giải tích cổ điển, chúng ta có thể nhận ra rằng tập phổ dụng là tập các số thực $\mathbb{R}$, còn trong giải tích hàm biến số phức thì tập phổ dụng là $\mathbb{C}$, và trong môn lý thuyết số thì tập phổ dụng là $\mathbb{Z}$.

**2.1.19 Định nghĩa.** Cho tập hợp $A$. Các bộ phận của $A$ lập thành một tập hợp, ký hiệu $\mathscr{P}(A)$, gọi là **tập hợp các bộ phận** của $A$.

Từ định nghĩa chúng ta thấy rằng các phần tử của $\mathscr{P}(A)$ là tập con của $A$. Tập hợp $\mathscr{P}(A)$ luôn chứa các phần tử $\varnothing$ và $A$, cho nên $\mathscr{P}(A)$ không là tập rỗng với bất kì tập $A$ nào.

**2.1.20 Thí dụ.** Từ Thí dụ 2.1.17, chúng ta thấy rằng tập hợp các bộ phận của $\{1,2,3,4\}$ là

$$\begin{aligned}\mathscr{P}(\{1,2,3,4\}) = \{&\varnothing, \{1\}, \{2\}, \{3\}, \{4\}, \{1,2\}, \{1,3\}, \{1,4\}, \{2,3\},\\ &\{2,4\}, \{3,4\}, \{1,2,3\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}, \{1,2,3,4\}\}.\square\end{aligned}$$

## 2.2 Các phép toán trên các tập hợp

Từ đây trở về sau, chúng tôi xin quy ước các tập hợp mà chúng tôi đang đề cập đều là tập con của một tập phổ dụng $\mathscr{U}$ nào đó, nếu không chúng tôi sẽ nói cụ thể.

**2.2.1 Định nghĩa.** Cho hai tập $A$, $B$. **Hợp** (hay tổng) của $A$ và $B$ là tập gồm và chỉ gồm mọi phần tử hoặc thuộc $A$ hoặc thuộc $B$ và kí hiệu là $A \cup B$ (đọc là $A$ hợp $B$). Ta viết

$$A \cup B = \{x : x \in A \text{ hoặc } x \in B\}.$$

**Giao** của $A$ và $B$ là tập gồm và chỉ gồm mọi phần tử vừa thuộc $A$ vừa thuộc $B$, kí hiệu là $A \cap B$ (đôi khi để cho gọn ta viết $AB$), đọc là $A$ giao $B$.

$$A \cap B = \{x : x \in A \text{ và } x \in B\}.$$

Nếu $A \cap B = \varnothing$ thì ta nói $A$ *không giao* $B$ hay $A$ và $B$ **rời nhau**.

**2.2.2 Thí dụ.** Nếu $A = \{a, b, c\}$ và $B = \{c, d\}$ thì

$$A \cup B = \{a, b, c, d\} \qquad \text{và} \qquad A \cap B = \{c\}\,. \qquad \square$$

**2.2.3 Định nghĩa.** Cho hai tập hợp $A$ và $B$. **Hiệu** của tập $A$ đối với tập $B$ là một tập hợp gồm và chỉ gồm mọi phần tử thuộc $A$ nhưng không thuộc $B$, và kí hiệu là $A \setminus B$.

$$A \setminus B = \{x : x \in A, x \notin B\}\,.$$

Nói riêng, khi $A = \mathscr{U}$, tập phổ dụng, hiệu $\mathscr{U} \setminus B$ được gọi là **phần bù** của tập hợp $B$, và kí hiệu là $\mathscr{U} \setminus B = B^c$.

Từ định nghĩa trên dễ thấy rằng $A \setminus B = A \cap B^c$.

**2.2.4 Thí dụ.** Nếu $A = \{a, b, c, d\}$, $B = \{c, d, e\}$, thì $A \setminus B = \{a, b\}$. $\square$

Từ định nghĩa của các phép hợp, giao, hiệu, suy ra các tính chất được nêu trong định lý sau đây:

**2.2.5 Định lý.** *Cho các tập hợp A, B, và C ứng với tập phổ dụng $\mathscr{U}$. Khi đó, ta có*

**(1)** *Tính giao hoán:*

$$A \cup B = B \cup A; \qquad\qquad A \cap B = B \cap A.$$

**(2)** *Tính kết hợp:*

$$(A \cup B) \cup C = A \cup (B \cup C); \quad (A \cap B) \cap C = A \cap (B \cap C);$$

*do tính chất này mà khi viết hợp hoặc giao của một dãy tập không cần phải để các ngoặc đơn; không có vấn đề gì với việc hợp nào được thực hiện trước.*

**(3)** *Tính phân phối*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C);$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

**(4)** *Công thức De Morgan*

$$(A \cup B)^c = A^c \cap B^c; \qquad (A \cap B)^c = A^c \cup B^c.$$

*Một cách tổng quát ta cũng có*

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$$
$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C).$$

Việc chứng minh các tính chất này dựa vào định nghĩa các phép hợp, giao, phần bù và định nghĩa sự bằng nhau của hai tập hợp. Chẳng hạn ta chứng minh đẳng thức thứ hai của tính chất (3) trong thí dụ sau

**2.2.6 Thí dụ.** Cho ba tập hợp $A$, $B$, và $C$. Chứng minh rằng $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Giả sử $x \in A \cup (B \cap C)$, thế thì $x \in A$ hoặc $x \in B \cap C$. Nếu $x \in A$ thì rõ ràng $x \in A \cup B$ và $x \in A \cup C$ nên $x \in (A \cup B) \cap (A \cup C)$. Còn nếu $x \in B \cap C$ thì $x \in B$ và $x \in C$, do đó $x \in A \cup B$ và $x \in A \cup C$ nên $x \in (A \cup B) \cap (A \cup C)$. Vì vậy ta có nếu $x \in A \cup (B \cap C)$, thì $x \in (A \cup B) \cap (A \cup C)$, hay $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$.

Ngược lại, giả sử $x \in (A \cup B) \cap (A \cup C)$ thì $x \in A \cup B$ và $x \in A \cup C$. Nếu $x \notin A$, ta phải có $x \in B$ và $x \in C$, tức là $x \in B \cap C$. Suy ra $x \in A \cup (B \cap C)$, hay $(A \cup B) \cap (A \cup C) \subseteq A \cup (B \cap C)$. Kết hợp với kết quả ở trên, chúng ta thu được $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. $\square$

## Bài tập

**1)** Liệt kê các phần tử của các tập hợp sau

**(a)** $\{x : x \text{ là số thực sao cho } x^2 = 1\}$

**(b)** $\{x : x \text{ là số nguyên dương nhỏ hơn } 9\}$

**(c)** $\{x : x \text{ là số chính phương và } x < 100\}$

**(d)** $\left\{x : x \text{ là số nguyên sao cho } x^2 = 2\right\}$

**2 )** Xác định giá trị chân lý của các mệnh đề sau:

**(a)** $x \in \{x\}$ **(c)** $\{x\} \in \{x\}$ **(e)** $\{x\} \in \{x, \{x\}\}$

**(b)** $\{x\} \subseteq \{x\}$ **(d)** $\{x\} \in \{\{x\}\}$ **(f)** $\varnothing \in \{x\}$.

**3 )** Nếu $A = \{x : x = 1, 0, 1\}$, $B = \{-1, 0, 1\}$, $C = \{0, 1, -1\}$, $D = \{-1, 1, -1\}$, hãy xác định đúng sai của các khẳng định sau

**(a)** $A = B$ **(d)** $A \subseteq B$ **(g)** $C \subseteq A$ **(j)** $0 \subseteq C$

**(b)** $A = C$ **(e)** $A \subseteq C$ **(h)** $C \subseteq B$ **(k)** $B \subseteq D$

**(c)** $B = C$ **(f)** $B \subseteq C$ **(i)** $0 \in B$ **(l)** $D \in C$.

**4 )** Cho $E = \{x : 0 \leq x \leq 1\}$, $F = \{y : 0 \leq y \leq 1\}$, $G = \{x : 0 < x < 1\}$, $H = \{1\}$, hãy xác định đúng sai của các khẳng định sau.

**(a)** $E = F$ **(c)** $F \subseteq G$ **(e)** $H \subseteq G$ **(g)** $G \in E$

**(b)** $F = G$ **(d)** $G \subseteq F$ **(f)** $H \subseteq E$ **(h)** $H \in F$

**5 )** Xác định tập $\mathscr{P}(X)$ trong các trường hợp sau

**(a)** $X = \varnothing$ **(b)** $X = \{\varnothing, \{\varnothing\}\}$ **(c)** $X = \{a\}$.

**6 )** Cho các tập hợp $A = \{2, 4, 6, 8, 10, 12, 14, 16\}$, $B = \{3, 6, 9, 12, 15\}$, $C = \{1, 2, 4, 7, 11, 16\}$, $\mathscr{U} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16\}$. Hãy xác định các tập

**(a)** $A \cup B$ **(b)** $A \cap C$ **(c)** $C^c$ **(d)** $C \setminus A$.

**7)** Cho một ví dụ chứng tỏ rằng nếu $E \subseteq F$ và $D \subseteq F$ thì chúng ta không nhất thiết phải có $E \subseteq D$ hay $D \subseteq E$.

**8)** Chứng minh Định lý 2.1.14.

**9)** Chứng minh các tính chất trong Định lý 2.2.5.

**10)** Cho $A$, $B$, $C$ là các tập hợp bất kì. Chứng minh rằng

**(a)** $A \cap A = A = A \cup A$

**(b)** $A \cap B \subseteq A \subseteq A \cup B$

**(c)** $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$

**(d)** $A \setminus (A \setminus B) = A \cap B$

**(e)** $A \cup (B \setminus C) = (A \setminus C) \cup (B \setminus C)$

**(f)** $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

**(g)** $(A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$

**(h)** $A \cup (B \setminus A) = A \cup B$

**(i)** $A \cap (B \setminus A) = \varnothing$

**(j)** $A \setminus B = A \setminus (A \cap B) = (A \cup B) \setminus B$

**11)** Chứng minh rằng nếu $B \subseteq A$ thì $A \cap B = B$ và $A \cup B = A$.

**12)** Cho hai tập hợp $A$ và $B$. Chứng minh rằng điều kiện cần và đủ để $A = B$ là $A \cap B = A \cup B$.

**13)** Cho hai tập hợp $A$ và $B$ bất kì. Chứng minh rằng $A \setminus (A \setminus B) = B$ khi và chỉ khi $B \subseteq A$.

**14)** Cho ba tập hợp $A$, $B$, $C$ bất kì. Chứng minh rằng

**(a)** nếu $A \subseteq C$ và $B \subseteq C$ thì $A \cup B \subseteq C$.

**(b)** nếu $A \supseteq C$ và $B \supseteq C$ thì $A \cap B \supseteq C$.

# § 3 Quan hệ và ánh xạ

## 3.1 Quan hệ

**3.1.1 Định nghĩa.** Cho hai tập hợp $A$ và $B$. Tập hợp các cặp sắp thứ tự $(a, b)$ với $a \in A$ và $b \in B$ được gọi là **tích Descartes*** **của** $A$ **và** $B$ và kí hiệu bởi $A \times B$. Ta viết

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

Lưu ý $(a, b) = (a', b')$ khi và chỉ khi $a = a'$ và $b = b'$.

Trường hợp riêng, khi $B = A$ thì tích $A \times B$ được kí hiệu $A^2$. Trong trường hợp $A = \varnothing$ hay $B = \varnothing$, ta thấy hoặc có thể chứng minh được $A \times B = \varnothing$.

**3.1.2 Thí dụ.** Cho hai tập hợp $A = \{a, a'\}$ và $B = \{b, b', b''\}$. Tích Descartes của $A$ và $B$ là

$$A \times B = \big\{(a, b), (a, b'), (a, b''), (a', b), (a', b'), (a', b'')\big\}. \qquad \square$$

**3.1.3** Khái niệm tích Descartes có thể mở rộng cho trường hợp nhiều tập hợp. Cho $A_1$, $A_2$, ..., $A_n$ là $n$ tập hợp. Với mỗi $a_i \in A_i$ ta lập bộ $n$ phần tử có thứ tự $(a_1, a_2, \ldots, a_n)$. Hai bộ $n$ phần tử $(a_1, a_2, \ldots, a_n)$ và $(a'_1, a'_2, \ldots, a'_n)$ được xem là bằng nhau khi và chỉ khi $a_i = a'_i$ với mọi $i = 1, 2, \ldots, n$. Tập tất cả các bộ có thứ tự $n$ phần tử gọi là *tích Descartes* của các tập $A_1$, $A_2$, ..., $A_n$, ký hiệu $A_1 \times A_2 \times \cdots \times A_n$. Như vậy

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) : a_i \in A_i, i = 1, 2, \ldots, n\}$$

*René Descartes (1596-1650), Nhà toán học và triết học Pháp, và ông được xem là cha đẻ của môn học Hình Giải Tích. Một câu nói nổi tiếng của ông là "I think, therefore I am"

Đặc biệt nếu $A_i = A$ với mọi $i = 1, 2, \dots, n$ thì tích Descartes $A \times A \times \dots \times A$ còn được viết $A^n$ gọi là *lũy thừa Descartes* bậc $n$ của $A$. Ta có

$$A^n = \{(a_1, a_2, \dots, a_n) : a_i \in A, i = 1, 2, \dots, n\}.$$

**3.1.4 Định nghĩa.** Giả sử $A$ và $B$ là hai tập hợp. Một **quan hệ hai ngôi từ $A$ đến $B$** là một bộ phận $R$ của tích Descarte $A \times B$. Ta bảo, với hai phần tử $a \in A$ và $b \in B$, rằng *$a$ có quan hệ $R$ với $b$* nếu và chỉ nếu $(a, b) \in R$. Ta viết $aRb$. Tập con của $A$ xác định bởi $\{x \in A :$ tồn tại $y \in B$ sao cho $(x, y) \in R\}$ được gọi là **miền xác định của quan hệ** $R$. Tập con của $B$ xác định bởi $\{y \in B :$ tồn tại $x \in A$ sao cho $(x, y) \in R\}$ được gọi là **miền (giá trị) của quan hệ** $R$.

Trong trường hợp $B = A$, ta nói $R$ là một quan hệ trong $A$.

**3.1.5 Thí dụ.** Cho tập hợp $A$ khác rỗng bất kì. Xét tập $\mathscr{P}(A)$. Khi đó, tập hợp $\big\{(U, V) : U, V \in \mathscr{P}(A), x \in U \Rightarrow x \in V\big\}$ là một quan hệ trong $\mathscr{P}(A)$. Đó chính là quan hệ bao hàm $\subseteq$. $\square$

**3.1.6 Định nghĩa.** Giả sử $R$ là một quan hệ trên một tập hợp $A$. Khi đó $R$ là một **quan hệ tương đương trong** $A$ nếu và chỉ nếu các điều kiện sau đây thỏa mãn

(1) (Phản xạ) Với mọi $x \in A$, $xRx$.

(2) (Đối xứng) Với mọi $x, y \in A$, nếu $xRy$, thì $yRx$.

(3) (Bắc cầu) Với mọi $x, y, z \in A$, nếu $xRy$ và $yRz$, thì $xRz$.

Nếu $R$ là một quan hệ tương đương, thì người ta thường kí hiệu $R$ bằng $\sim$ và thường đọc $aRb$, hay $a \sim b$, là "$a$ tương đương với $b$". Ta cũng có thể suy ra rằng miền xác định và miền giá trị của quan hệ tương đương $R$ chính bằng $A$. Đặt $C_a = \{x \in A : xRa\}$, nó được gọi là *lớp tương đương* của phần tử $a$. Người ta có thể chứng minh được nếu $C_a \cap C_b \neq \varnothing$ thì $C_a = C_b$. Điều đó có nghĩa chúng ta có thể phân hoạch tập $A$ thành các lớp tương đương. Tập các lớp tương đương xác định bởi quan hệ tương đương $R$ được kí hiệu là $A/R$, còn gọi là *tập thương*.

**3.1.7 Định nghĩa.** Giả sử $A$ là một tập hợp khác rỗng, và $R$ là một quan hệ trên $A$. Thế thì $R$ được gọi là một **quan hệ thứ tự trong** $A$ (hay người ta còn gọi $R$ là một quan hệ thứ tự giữa các phần tử của $A$) nếu và chỉ nếu các điều kiện sau đây thỏa mãn:

(1) (Phản xạ) Với mọi $a \in A$, $aRa$.

(2) (Phản đối xứng) Với mọi $a, b \in A$, nếu $aRb$ và $bRa$, thì $a = b$.

(3) (Bắc cầu) Với mọi $a, b, c \in A$, nếu $aRb$ và $bRc$, thì $aRc$.

Hơn nữa, nếu với mọi $a, b \in A$ ta luôn có $aRb$ hoặc $bRa$, thì quan hệ thứ tự $R$ được gọi là **toàn phần**, và $A$ được gọi là **tập được sắp tuyến tính** hay **sắp toàn phần**. Ngược lại, không phải ta luôn có $aRb$ hoặc $bRa$ với mọi $a, b \in A$, thì $R$ được gọi là **quan hệ thứ tự bộ phận**.

Nếu $R$ là một quan hệ thứ tự, thì người ta thường kí hiệu $R$ bằng $\preceq$ và thường đọc $aRb$, hay $a \preceq b$, là "$a$ nhỏ hơn $b$". Và nếu biết thêm rằng $a \neq b$, ta sẽ viết $a \prec b$.

Người ta bảo một tập hợp $A$ là **sắp thứ tự** nếu trong $A$ có một quan hệ thứ tự.

**3.1.8 Thí dụ.** Quan hệ bao hàm, $\subseteq$, trong $\mathscr{P}(A)$ là một quan hệ thứ tự bộ phận. Quan hệ $\leq$ trong tập số thực $\mathbb{R}$ là một quan hệ thứ tự toàn phần. $\square$

**3.1.9 Định nghĩa.** Giả sử $A$ là một tập hợp sắp thứ tự với quan hệ thứ tự $\preceq$ và $E$ là một tập con của $A$. Phần tử $a$ được gọi là **phần tử tối tiểu** của $A$ nếu quan hệ $x \preceq a$ kéo theo $x = a$.

Phần tử $a \in E$ được gọi là **phần tử bé nhất** của $E$ nếu với mọi $x \in E$ ta có $a \preceq x$.

Phần tử $a$ được gọi là **phần tử tối đại** của $A$ nếu quan hệ $a \preceq x$ kéo theo $x = a$.

Phần tử $a \in E$ được gọi là **phần tử lớn nhất** của $E$ nếu với mọi $x \in E$ ta có $x \preceq a$.

Phần từ $b$ được gọi là **cận trên** của tập $E$ nếu $x \leq b$ với mọi $x \in E$. Phần từ $c$ được gọi là **cận dưới** của tập $E$ nếu $c \leq x$ với mọi $x \in E$.

Một tập được sắp thứ tự gọi là **được sắp tốt** nếu mọi tập con khác rỗng của nó đều có phần tử nhỏ nhất.

## 3.2 Ánh xạ

**3.2.1 Định nghĩa.** Giả sử $A$ và $B$ là hai tập hợp đã cho. Ta gọi $f$ là **ánh xạ từ** $A$ **đến** $B$ nếu $f$ là một quan hệ từ $A$ đến $B$ sao cho nếu $(a, b) \in f$ và $(a, b') \in f$ thì $b = b'$ và với mỗi $a \in A$ tồn tại $b \in B$ sao cho $(a, b) \in f$. Ký hiệu $f : A \to B$. Tập $A$ được gọi là **miền xác định** hay *nguồn* của ánh xạ, và tập $B$ gọi là **tập đích** của ánh xạ $f$.

Từ định nghĩa ta nhận thấy rằng với mỗi phần tử $a \in A$ tồn tại *duy nhất* một phần tử $b \in B$ sao cho $(a, b) \in f$, và phần tử duy nhất $b$ đó được gọi là **ảnh** của $a$ qua $f$, kí hiệu $f(a)$ nên ta viết $a \mapsto b = f(a)$. Mặt khác, có thể có hai phần tử $a \neq a'$ của $A$ có cùng một ảnh như nhau, nghĩa là $f(a) = f(a')$. Vì vậy một cách trực giác ta có thể hiểu rằng ánh xạ $f : A \to B$ là một *quy tắc* cho tương ứng mỗi phần tử $a \in A$ với một và chỉ một phần tử $b = f(a)$ thuộc $B$. Ta viết chung

$$\begin{aligned} f : A &\to B \\ a &\mapsto b = f(a) \end{aligned}$$

**3.2.2 Thí dụ.** Cho $A$ là tập con của $B$. Khi đó $j = \{(a, a) : a \in A\}$ là một ánh xạ từ $A$ đến $B$, $j : A \to B$, xác định bởi $a \mapsto j(a) = a$; ánh xạ $j$ được gọi là *ánh xạ chính tắc*. Trường hợp đặc biệt $A = B$, thì ánh xạ ấy gọi là **ánh xạ đồng nhất**, kí hiệu $\mathrm{Id}_A : A \to A$. Vậy với mọi $a \in A$ ta có $\mathrm{Id}_A(a) = a$. □

**3.2.3 Định nghĩa.** Giả sử $f : A \to B$ là một ánh xạ đã cho, $X$ là một bộ phận tùy ý của $A$, và $Y$ là một bộ phận tùy ý của $B$. Khi đó, ta gọi $f(X) = \{b \in B : \text{tồn tại } a \in X \text{ sao cho } f(a) = b\}$ là **ảnh của** $X$ **bởi** $f$.

$f^{-1}(Y) = \{a \in A : f(a) \in Y\}$ là **tạo ảnh của** $Y$ **bởi** $f$. Đặc biệt, với $b \in B$ để đơn giản ta viết $f^{-1}(b)$ thay cho $f^{-1}(\{b\})$ và gọi là tạo ảnh toàn phần của $b$ bởi $f$. Vậy $f^{-1}(b) = \{a \in A : f(a) = b\}$. Tập hợp $f(A) \subseteq B$ được gọi là **tập giá trị** của ánh xạ $f$.

Người ta có thể chứng minh được định lý sau.

**3.2.4 Định lý.** *Cho $f : A \to B$ là một ánh xạ. Giả sử $X, X' \subseteq A$ và $Y, Y' \subseteq B$. Khi đó, ta có*

**(1)** $X \subseteq f^{-1}(f(X)), \qquad Y \supseteq f(f^{-1}(Y))$

**(2)** $f(X \cup X') = f(X) \cup f(X'), \qquad f(X \cap X') \subseteq f(X) \cap f(X')$

**(3)** $f^{-1}(Y \cup Y') = f^{-1}(Y) \cup f^{-1}(Y'),$
$f^{-1}(Y \cap Y') = f^{-1}(Y) \cap f^{-1}(Y').$

**3.2.5 Định nghĩa.** Cho ánh xạ $f : A \to B$.

**(1)** Ta bảo ánh xạ $f$ là một **toàn ánh**. nếu $f(A) = B$. Nói một cách khác, với mọi $b \in B$ có ít nhất một $a \in A$ sao cho $b = f(a)$. Người ta còn gọi một toàn ánh là $f : A \to B$ là một **ánh xạ từ** $A$ **lên** $B$.

**(2)** Ta gọi ánh xạ $f$ là một **đơn ánh** nếu với mọi phần tử $a, a' \in A$, quan hệ $a \neq a'$ kéo theo quan hệ $f(a) \neq f(a')$, hay $f(a) = f(a')$ kéo theo $a = a'$; hay với mọi $b \in B$ có nhiều nhất một $a \in A$ sao cho $b = f(a)$. Người ta còn gọi một đơn ánh $f : A \to B$ là một *ánh xạ một đối một*.

**(3)** Nếu $f$ vừa là toàn ánh vừa là đơn ánh thì nó được gọi là **song ánh** hay một **ánh xạ một đối một từ** $A$ **lên** $B$. Nói một cách khác, $f : A \to B$ là một song ánh nếu và chỉ nếu với mọi $b \in B$ có một và chỉ một $a \in A$ sao cho $b = f(a)$.

**3.2.6 Thí dụ.** Giả sử $A \subseteq B$. Ta dễ dàng thấy được ánh xạ chính tắc $j : A \to B$ (xác định bởi $j(a) = a$) là một đơn ánh. Vì thế nên ta gọi $j$ là **đơn ánh chính tắc**. Ta cũng thấy rằng ánh xạ đồng nhất là một song ánh. $\square$

Giả sử $f : A \to B$ là một song ánh. Như vậy với mỗi phần tử $b \in B$ tồn tại duy nhất một phần tử $a \in A$ sao cho $b = f(a)$. Do đó trong trường hợp này tập hợp $\{(b, a) \in B \times A : b = f(a)\}$ là ánh xạ từ $B$ đến $A$, cụ thể được xác định bởi $b \mapsto a$ nếu $f(a) = b$, kí hiệu $f^{-1} : B \to A$, và được gọi là **ánh xạ ngược của** $f$.

**3.2.7 Định nghĩa.** Cho $f : A \to B$ và $g : B \to C$ là hai ánh xạ. Dễ dàng thấy rằng $\{(x, g(f(x))) \in A \times C : x \in A\}$ là một ánh xạ từ $A$ đến $C$, và nó được gọi là **tích** hay **hợp thành của ánh xạ** $f$ **và** $g$; ký hiệu $g \circ f$. Vậy ta có $g \circ f : A \to C$ và $g \circ f(x) = g(f(x))$ với mọi $x \in A$. Ta có biểu đồ

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & B \arrow[r, "g"] & C \end{tikzcd}$$

**3.2.8 Định lý.** *Giả sử cho ba ánh xạ $f : A \to B$, $g : B \to C$, và $h : C \to D$. Thế thì $h \circ (g \circ f) = (h \circ g) \circ f$.*

*Chứng minh.* Với mọi $x \in A$, ta có

$$h\circ(g\circ f)(x) = h(g\circ f(x)) = h(g(f(x))) = h\circ g(f(x)) = (h\circ g)\circ f(x).$$

Vậy ta có $h \circ (g \circ f) = (h \circ g) \circ f$. □

Do $h \circ (g \circ f) = (h \circ g) \circ f$ nên ta có thể kí hiệu chung bởi $h \circ g \circ f$ và gọi là hợp thành của ba ánh xạ $f$, $g$, và $h$.

Ta biết rằng nếu $f : A \to B$ là một song ánh thì tồn tại một ánh xạ ngược $f^{-1} : B \to A$. Dễ dàng chứng minh được $f^{-1} \circ f = \mathrm{Id}_A$ và $f \circ f^{-1} = \mathrm{Id}_B$. Hơn thế nữa, mệnh đề đảo vẫn đúng, nên ta có định lý sau

**3.2.9 Định lý.** *Cho $f : A \to B$ là một ánh xạ. $f$ là một song ánh khi và chỉ khi tồn tại một ánh xạ $g : B \to A$ sao cho $g \circ f = \mathrm{Id}_A$ và $f \circ g = \mathrm{Id}_B$. Trong trường hợp này chính $g$ là ánh xạ ngược của $f$.*

*Chứng minh.* Ta chỉ cần chứng minh nếu tồn tại ánh xạ $g : B \to A$ sao cho $g \circ f = \mathrm{Id}_A$ và $f \circ g = \mathrm{Id}_B$ thì $f$ là song ánh.

Với mỗi $b \in B$, ta có $\mathrm{Id}_B(b) = b$ hay $f \circ g(b) = b$ bởi vì $f \circ g = \mathrm{Id}_B$. Chỉ việc đặt $a = g(b)$, ta có ngay $f(a) = f(g(b)) = f \circ g(b) = b$. Vậy $f$ là toàn ánh.

Giả sử $a, a' \in A$ sao cho $f(a) = f(a')$. Bởi vì $g$ là một ánh xạ ta có $g(f(a)) = g(f(a'))$ hay $g \circ f(a) = g \circ f(a')$. Bởi vì $g \circ f = \mathrm{Id}_A$, ta suy ra $a = a'$. Do đó $f$ là đơn ánh. Chúng ta kết luận rằng $f$ là song ánh.

Theo trên $a = g(b)$ thì $f(a) = b$, suy ra $f^{-1}(b) = a$. Do đó $g(b) = f^{-1}(b)$ hay $g = f^{-1}$. □

Qua định lý trên ta thấy rằng $f^{-1}$ cũng là một song ánh, và $f$ là ánh xạ ngược của $f^{-1}$. Vậy $(f^{-1})^{-1} = f$.

**3.2.10 Định lý.** *Cho hai song ánh $f : A \to B$ và $g : B \to C$. Khi đó $g \circ f : A \to C$ là một song ánh.*

*Chứng minh.* Theo giả thiết ta suy ra được $g^{-1} : C \to B$ và $f^{-1} : B \to A$ là hai song ánh. Đặc biệt, $f^{-1} \circ g^{-1}$ là một ánh xạ từ $C$ đến $A$. Vận dụng Định lý 3.2.9 ta có

$$\begin{aligned}(g \circ f) \circ (f^{-1} \circ g^{-1}) &= g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \mathrm{Id}_B \circ g^{-1} \\ &= g \circ g^{-1} = \mathrm{Id}_C.\end{aligned}$$

và

$$\begin{aligned}(f^{-1} \circ g^{-1}) \circ (g \circ f) &= f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ \mathrm{Id}_B \circ f \\ &= f^{-1} \circ f = \mathrm{Id}_A.\end{aligned}$$

Vậy $g \circ f$ là một song ánh; hơn nữa, $f^{-1} \circ g^{-1}$ là ánh xạ ngược của nó. Vậy $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. □

**3.2.11 Định nghĩa.** Giả sử $f : A \to B$ là một ánh xạ, và $X$ là một bộ phận của $A$. Khi đó $g = \{(a, b) \in f : a \in X\}$ là một ánh xạ từ $X$ đến $B$. Ánh xạ

$$\begin{aligned} g : X &\to B \\ a &\mapsto g(a) = f(a) \end{aligned}$$

gọi là **cái thu hẹp của ánh xạ $f$ vào bộ phận $X$** và kí hiệu $g = f|_X$. Còn ánh xạ $f$ gọi là **cái mở rộng của $g$ trên tập hợp $A$**.

**3.2.12 Định nghĩa.** Cho $A$ là một tập hợp. Một **phép toán hai ngôi** trên $A$ là một ánh xạ từ $A^2$ vào $A$. Chúng ta thường kí hiệu phép toán hai ngôi bởi $+$ hay $\cdot$. Chẳng hạn, $+ : A^2 \to A$ chúng ta thường viết $a + b$ thay cho $+(a, b)$ với $a, b \in A$.

## Bài tập

**1**) Cho $A, C \subseteq X$ và $B, D \subseteq Y$. Chứng minh rằng

**(a)** $A \times (B \cap D) = (A \times B) \cap (A \times D)$

**(b)** $A \times (B \cup D) = (A \times B) \cup (A \times D)$

**(c)** $A \times (B \setminus D) = (A \times B) \setminus (A \times D)$

**(d)** $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$

**(e)** $(A \times B) \cup (C \times D) \subseteq (A \cup C) \times (B \cup D)$

**(f)** $A \times B = (A \times Y) \cap (X \times B)$

**(g)** $(X \times Y) \setminus (A \times B) = [(X \setminus A) \times Y] \cup [X \times (Y \setminus B)]$

**2**) Trong tập hợp số tự nhiên $\mathbb{N}$, xét các quan hệ sau.

**(a)** $aRb \Leftrightarrow a, b$ không nguyên tố cùng nhau.

**(b)** $aRb \Leftrightarrow a, b$ có ước chung là $p$ (với $p$ là một số nguyên tố cho trước).

**(c)** $aRb \Leftrightarrow a, b$ có các chữ số hàng đơn vị giống nhau.

**(d)** $aRb \Leftrightarrow a + b \,\vdots\, 5$.

Trong các quan hệ trên, quan hệ nào là quan hệ tương đương. Nếu là quan hệ tương đương hãy chỉ ra tập các lớp tương đương.

**3**) Cho $X$ là một tập và $A$ là một tập con của $X$ có nhiều hơn một phần tử. Chứng minh rằng quan hệ $R$ trên $X$ xác định bởi $aRb$ nếu $a = b$ hoặc $a, b \in A$ là một quan hệ tương đương.

**4**) Trong tập $X$ gồm các điểm trong không gian $Oxyz$, xét quan hệ sau: $A\Re B$ khi và chỉ khi $O, A, B$ thẳng hàng.

**(a)** $\Re$ có phải là quan hệ tương đương trong $X$ hay không?

**(b)** $\Re$ có phải là quan hệ tương đương trong $X \setminus \{O\}$ hay không?

**(c)** Xác định tập các lớp tương đương (nếu có) ở (a) và (b).

**5**) Giả sử $R$ là một quan hệ tương đương trên tập $A$ và ký hiệu $C_a$ là lớp tương đương chứa $a$. Chứng minh rằng nếu $C_a \cap C_b \neq \varnothing$ thì $C_a = C_b$.

**6**) Trong tập số thực $\mathbb{R}$, xét quan hệ $S$ và $T$ sau: $xRy$ khi và chỉ khi $x^2 \leq y^2$, và $xTy$ khi và chỉ khi $x^3 \leq y^3$. Trong hai quan hệ $S$ và $T$, quan hệ nào là quan hệ thứ tự.

**7**) Trong tập $\mathbb{R}^2$, xét quan hệ $T$ như sau $(a_1, b_1)T(a_2, b_2)$ khi và chỉ khi $(a_1 < a_2) \vee (a_1 = a_2 \wedge b_1 \leq b_2)$. Chứng minh rằng $T$ là quan hệ thứ tự trong $\mathbb{R}^2$.

**8**) Giả sử tập $E$ được sắp tốt bởi quan hệ $\preceq$ và cho $a \notin E$. Chứng minh rằng quan hệ mở rộng $\preceq$ trên $E \cup \{a\}$ xác định bởi $a \preceq x$ với mọi $x \in E$ xác định $E \cup \{a\}$ được sắp tốt.

**9**) Chứng minh các tính chất trong định lý 3.2.4.

**10**) Cho ánh xạ $f : A \to B$, và $A_1, A_2 \subseteq A$ và $B_1, B_2 \subseteq B$. Chứng minh

**(a)** Nếu $f$ là đơn ánh thì $f(A_1 \cap A_2) = f(A_1) \cap f(A_2)$.

**(b)** $f^{-1}(B_1 \setminus B_2) = f^{-1}(B_1) \setminus f^{-1}(B_2)$.

**11**) Cho $f : X \to Y$ là một ánh xạ. Chứng minh rằng các mệnh đề sau là tương đương.

**(i)** $f$ là một đơn ánh.

**(ii)** $f(A \cap B) = f(A) \cap f(B)$ với mọi $A, B \subseteq X$.

**(iii)** Mọi cặp các tập $A, B \subseteq X$ mà $A \cap B = \varnothing$ thì $f(A) \cap f(B) = \varnothing$.

**12)** $f : X \to Y$ là một toàn ánh khi và chỉ khi $f(f^{-1}(B)) = B$ với mọi $B \subseteq Y$.

**13)** Cho $f : X \to Y$ là một ánh xạ. Chứng tỏ rằng $f$ là một đơn ánh khi và chỉ khi với mọi tập $Z$ và mọi ánh xạ $g_1 : Z \to X$, $g_2 : Z \to X$ sao cho $f \circ g_1 = f \circ g_2$ thì $g_1 = g_2$.

**14)** Xét ánh xạ $f : \mathbb{Z} \to \mathbb{Z}$ được xác định như sau

**(a)** $f(n) = n - 1$ **(b)** $f(n) = n^3$ **(c)** $f(n) = n^2 + 1$

$f$ có là đơn ánh, toàn ánh không?

**15)** Cho ánh xạ $f : \mathbb{N}^2 \to \mathbb{N}$ xác định bởi $f(x, y) = x^2 + y^2$.

**(a)** $f$ có là đơn ánh không?

**(b)** $f$ có là toàn ánh không?

**(c)** Tìm $f^{-1}(0)$ và $f^{-1}(24)$.

**16)** Cho các ánh xạ $f : A \to B$, $g : B \to C$. Chứng minh rằng

**(a)** Nếu $g \circ f$ là đơn ánh thì $f$ là đơn ánh.

**(b)** Nếu $g \circ f$ là toàn ánh thì $g$ là toàn ánh.

**17)** Cho hai ánh xạ $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ xác định bởi $f(x) = \dfrac{1}{x}$ và $g : \mathbb{R} \to \mathbb{R}$ xác định bởi $g(x) = \dfrac{2x}{1 + x^2}$.

**(a)** Ánh xạ nào là đơn ánh, toàn ánh? **(b)** Xác định ánh xạ $g \circ f$.

**18)** Cho các ánh xạ $f : X \to Y$ và $g : Y \to Z$. Chứng minh rằng nếu $f$ là toàn ánh và $g \circ f$ là song ánh thì $f$ và $g$ là song ánh.

# § 4 Tích Descartes và nguyên lý tối đại Hausdorff

## 4.1 Chỉ số hóa và tích Descartes

**4.1.1 Định nghĩa.** Cho $I$ là một tập hợp khác rỗng mà các phần tử của nó ký hiệu là $\alpha$, $\beta$, $\gamma$, …, và $f : I \to X$ là một ánh xạ nào đó từ $I$ đến $X$. Ta ký hiệu $f(\alpha) = x_\alpha$, $f(\beta) = x_\beta$, $f(\gamma) = x_\gamma$, …, và nói các phần tử $x_\alpha$, $x_\beta$, $x_\gamma$, …, lập thành một *họ các phần tử* của $X$ được đánh số (hay chỉ số hóa) bởi tập chỉ số $I$, ký hiệu là $\{x_\alpha\}_{\alpha\in I}$. Nếu các $x_\alpha$, $x_\beta$, $x_\gamma$, …là các tập hợp thì họ $\{x_\alpha\}_{\alpha\in I}$ gọi là họ các tập hợp được đánh số (hay chỉ số hóa) bởi tập chỉ số $I$.

Như vậy, họ $\{x_\alpha\}_{\alpha\in I}$ thực chất là một cách biểu thị khác của ánh xạ $f : I \to X$. Bởi vậy hai họ $\{x_\alpha\}_{\alpha\in I}$ và $\{y_\alpha\}_{\alpha\in I}$ được xem là trùng nhau nếu $x_\alpha = y_\alpha$ với mọi $\alpha \in I$.

**4.1.2 Định nghĩa.** Cho $I$ là tập khác rỗng và $\mathscr{A}$ là tập mà các phần tử của nó là tập hợp. Nếu tồn tại một song ánh $f : I \to \mathscr{A}$, thì ta nói tập $\mathscr{A}$ được **chỉ số hóa** bởi tập $I$ theo ánh xạ $f$. Đặt $f(i) = A_i$ với mọi $i \in I$; khi đó tập $\mathscr{A}$ được viết lại $\{A_i : i \in I\}$ hay $(A_i)_{i\in I}$ và gọi là **họ tập hợp theo chỉ số** $I$.

Chúng ta có thể dễ dàng mở rộng các phép hợp và giao từ hai tập sang phép hợp và giao của một dãy tập hợp nhiều hơn hai. Cụ thể, cho $\{A_j : j \in I\}$ là một *họ các tập hợp* với tập chỉ số $I$; khi đó, hợp và giao của các tập trong họ $\{A_j : j \in I\}$ được kí hiệu $\bigcup_{j\in I} A_j$ và $\bigcap_{j\in I} A_j$ xác định như sau

$$\bigcup_{j\in I} A_j = \{x : \text{tồn tại } k \in I \text{ sao cho } x \in A_k\},$$
$$\bigcap_{j\in I} A_j = \{x : x \in A_k \text{ với mỗi } k \in I\}.$$

Trong trường hợp $I = \{1, 2, \dots, n\}$ ta viết $\bigcup_{j=1}^n A_j$ và $\bigcap_{j=1}^n A_j$ một cách tương ứng là hợp và giao của $n$ tập $A_1$, $A_2$, …, $A_n$ hay $\{A_j\}_{j=1}^n$. Còn nếu $I = \mathbb{N}^+$, ta viết $\bigcup_{j=1}^\infty A_j$, $\bigcap_{j=1}^\infty A_j$ một cách tương ứng là hợp và giao của các tập $A_1$, $A_2$, …, $A_n$, …hay $\{A_j\}_{j=1}^\infty$.

**4.1.3 Thí dụ.** Cho tập hợp $M$, với mỗi phần tử $x$ của $M$, ta đặt $N_x = M \setminus \{x\}$. Từ đó ta thấy rằng $x \notin N_x$ và $N_x \subset M$. Giả sử tồn tại $y \in M$ sao cho $y \in \bigcap_{x\in M} N_x$. Khi ấy chúng ta phải có $y \in N_y$, nhưng điều này mẫu thuẫn với $y \notin N_y$. Vì vậy

$$\bigcap_{x\in M} N_x = \varnothing. \qquad \square$$

Dễ dàng mở rộng hai công thức De Morgan cho hợp, giao của một số tùy ý các tập. Chẳng hạn nếu $A_1, A_2, \ldots, A_n$ là các tập con của $\mathscr{U}$ thì ta có:

$$\Big(\bigcup_{j=1}^{n} A_j\Big)^c = \bigcap_{j=1}^{n} A_j^c, \qquad \Big(\bigcap_{j=1}^{n} A_j\Big)^c = \bigcup_{j=1}^{n} A_j^c.$$

Tương tự như trên chúng ta cũng có công thức De Morgan sau

$$\Big(\bigcup_{j\in I} A_j\Big)^c = \bigcap_{j\in I} A_j^c, \qquad \Big(\bigcap_{j\in I} A_j\Big)^c = \bigcup_{j\in I} A_j^c.$$

Trong thí dụ sau chúng ta chứng minh một trong những công thức trên, phép chứng minh các công thức còn lại xin dành cho bạn đọc.

**4.1.4 Thí dụ.** Cho một họ các tập hợp $\{A_j : j \in I\}$ với chỉ số $I$ trong tập phổ dụng $\mathscr{U}$. Chứng minh rằng

$$\Big(\bigcup_{j\in I} A_j\Big)^c = \bigcap_{j\in I} A_j^c.$$

$x \in (\bigcup_{j\in I} A_j)^c$ tương đương với $x \notin \bigcup_{j\in I} A_j$, điều này tương đương với $x \notin A_j$ với mọi $j \in I$, và nó tương đương với $x \in A_j^c$ với mọi $j \in I$, cuối cùng nó tương đương với $x \in \bigcap_{j\in I} A_j^c$. Do đó,

$$\Big(\bigcup_{j\in I} A_j\Big)^c = \bigcap_{j\in I} A_j^c. \qquad \square$$

Cho $f : X \to Y$ là một ánh xạ và $\{A_j\}_{j\in I}$ là họ tập con của $Y$ theo tập chỉ số $I$. Khi đó ta có

$$f^{-1}\Big(\bigcup_{j\in I} A_j\Big) = \bigcup_{j\in I} f^{-1}(A_j) \quad \text{và} \quad f^{-1}\Big(\bigcap_{j\in I} A_j\Big) = \bigcap_{j\in I} f^{-1}(A_j).$$

**4.1.5 Định nghĩa.** Cho $I$ là tập khác rỗng và $\{A_i : i \in I\}$ là họ tập hợp theo chỉ số $I$. Khi đó tập tất cả các ánh xạ $f : I \to \bigcup_{i\in I} A_i$ thỏa $f(i) \in A_i$ với mọi $i \in I$ được gọi là **tích Descartes** của họ $\{A_i : i \in I\}$, kí hiệu $\prod_{i\in I} A_i$. Đặt $x_i = f(i)$ và kí hiệu $(x_i)_{i\in I}$ thay cho ánh xạ $f$. Khi ấy $(x_i)_{i\in I} \in \prod_{i\in I} A_i$ với $x_i \in A_i$ với mọi $i \in I$.

Các ánh xạ xác định bởi

$$\begin{aligned} p_j : \prod_{i\in I} A_i &\to A_j \\ \{x_i\}_{i\in I} &\mapsto x_j \end{aligned}$$

được gọi là **phép chiếu chính tắc** $\prod_{i\in I} A_i$ lên các thành phần $A_j$ với mỗi $j \in I$, chúng đều là các toàn ánh. Đặc biệt nếu $A_i = A$ với mọi $i \in I$ thì tích Descartes $\prod_{i\in I} A_i$ được gọi là *lũy thừa Descartes* và được viết là $A^I$.

## 4.2 Nguyên lý tối đại Hausdorff

Cho $X$ là một tập hợp được sắp thứ tự. Tập các tập con được sắp thứ tự toàn phần của $X$ được sắp thứ tự bộ phận bởi quan hệ bào hàm $\subseteq$. Tập con $E$ của $X$ gọi là tập **được sắp thứ tự toàn phần tối đại** nếu $E$ được sắp tuyến tính và nếu $D \subseteq X$, $D$ được săp thứ tự toàn phần và $E \subseteq D$ thì $E = D$.

Hiển nhiên trong $X$ có một tập con $E$ được sắp thứ tự toàn phần (chẳng hạn tập chỉ gồm một phần tử). Nếu $E$ không là tập tối đại thì ta có thể thay $E$ bởi một tập con được sắp thứ tự toàn phần lớn hơn. Ta thừa nhận như một tiên đề rằng: trong $X$ có một tập con được sắp thứ tự toàn phần tối đại, đó là nguyên lý tối đại Hausdorff.

**4.2.1 (Nguyên lý tối đại Hausdorff)** Trong mọi tập được sắp thứ tự bộ phận đều tồn tại một tập con được sắp thứ tự toàn phần tối đại.

**4.2.2 Định lý. (Bổ đề Zorn)** *Cho $X$ là một tập hợp sắp thứ tự khác rỗng thỏa tính chất là mỗi tập con sắp thứ tự toàn phần của $A$ đều có cận trên. Khi đó trong $X$ có phần tử tối đại.*

*Chứng minh.* Theo nguyên lý tối đại Hausdorff, tồn tại tập con sắp thứ tự toàn phần tối đại $A$. Theo giả thiết tồn tại cận trên $a$ của $A$. Phần tử $a$ này chính là phần tử tối đại của $X$. Thật vậy, với mỗi $x \in X$ giả sử $a \preceq x$ mà $a \neq x$ thì do $a$ là cận trên của $A$ nên $x \notin A$; hơn nữa ta dễ dàng nhận thấy $A \cup \{a, x\}$ là tập sắp thứ tự toàn phần với $x$ là phần tử lớn nhất. Điều này mâu thuẫn với tính tối đại theo quan hệ bao hàm của tập $A$. □

**4.2.3 Định lý. (Nguyên lý sắp tốt)** *Mọi tập khác rỗng $X$ đều có thể được sắp tốt, tức là tồn tại một quan hệ thứ tự trên $X$ mà $X$ với quan hệ thứ tự này là tập được sắp tốt.*

*Chứng minh.* Ký hiệu $\mathcal{W}$ là họ các quan hệ thứ tự sắp tốt một tập con của $X$. Xét $\mathcal{W}$ như là họ các tập con của $X^2$. Khi đó, $\mathcal{W}$ được sắp thứ tự bởi quan hệ bao hàm $\subseteq$. Từ đó ta nhận thấy mọi tập con sắp thứ tự toàn phần đều có cận trên (đó là hợp của các phần tử trong tập con ấy), nên theo bổ đề Zorn trong $\mathcal{W}$ có phần tử tối đại $E$ với quan hệ sắp tốt $\leq$ trên $E$. Nếu tồn tại $x_0 \in X \setminus E$ thì quan hệ $\leq$ có thể mở rộng thành quan hệ sắp tốt trên $E \cup \{x_0\}$ bằng cách đặt $x_0 \leq x$ với mọi $x \in E$. Ta gặp mâu thuẫn với $E$ tối đại. Vậy $X = E$ là tập được sắp tốt. □

**4.2.4 Định lý. (Tiên đề chọn)** *Nếu $\{X_\alpha\}_{\alpha\in I}$ là một họ khác rỗng các tập khác rỗng thì $\prod\limits_{\alpha\in I} X_\alpha$ khác rỗng. Đặc biệt, nếu $\{X_\alpha\}_{\alpha\in I}$ là họ khác rỗng các tập khác rỗng rời nhau thì tồn tại tập con $Y$ của $\bigcup\limits_{\alpha\in I} X_\alpha$ sao cho với mọi $\alpha \in I$, tập $Y \cap X_\alpha$ có chỉ một phần tử duy nhất.*

*Chứng minh.* Đặt $X = \bigcup\limits_{\alpha\in I} X_\alpha$. Từ giả thiết ta thấy ngay $X$ khác rỗng, cho nên theo nguyên lý sắp tốt $X$ là tập được sắp tốt bởi quan hệ thứ tự nào đó, ta sẽ đề cập đến quan hệ thứ tự sắp tốt này. Với mọi $\alpha \in I$, ta có $X_\alpha$ là tập con khác rỗng của $X$ nên tồn tại phần tử nhỏ nhất đặt nó là $f(\alpha)$. Khi đó $f = (f(\alpha))_{\alpha\in I} \in \prod\limits_{\alpha\in I} X_\alpha$. Bây giờ giả sử thêm rằng các tập $X_\alpha$ rời nhau. Đặt $Y = f(I)$. Ta có $Y \cap X_\alpha = \{f(\alpha)\}$ là tập chỉ chứa một phần tử duy nhất với mỗi $\alpha \in I$. □

## 4.3 Tính chất của tập được sắp tốt

Cho $X$ là một tập được sắp tốt. Khi đó mọi tập con khác rỗng $A$ của $X$ có phần tử nhỏ nhất, ký hiệu là $\min A$; phần tử này cũng là cận dưới lớn nhất của $A$, ký hiệu $\inf A$. Khi đó, ta có

$$\min A = \inf A \in A.$$

Nếu tập $A$ bị chặn trên thì có biên trên nhỏ nhất, ký hiệu là $\sup A$. Chú ý rằng có thể $\sup A \notin A$, nói cách khác có thể có $\sup A$ nhưng không có $\max A$.

**4.3.1 Định nghĩa.** Cho $X$ là một tập được sắp tốt. Với mỗi $x \in X$, ta gọi *khoảng gốc* mút $x$ là tập con $I_x = \{y \in X : y < x\}$. Các phần tử của $I_x$ gọi là các *phần tử đi trước* $x$.

**4.3.2 Định lý. (Nguyên lý quy nạp siêu hạn)** *Cho $X$ là một tập được sắp tốt. Nếu $A$ là một tập con khác rỗng của $X$ thỏa mãn $I_x \subseteq A$ kéo theo $x \in A$ thì $A = X$.*

*Chứng minh.* Nếu trái lại $A \neq X$. Đặt $x = \min(X \setminus A)$. Khi đó, với mọi $y \in I_x$ ta có $y < x$ cho nên $y \notin X \setminus A$ nghĩa là $y \in A$. Vậy $I_x \subseteq A$ nhưng $x \notin A$. Ta gặp mâu thuẫn. □

**4.3.3 Định lý.** *Cho $X$ là một tập được sắp tốt và $A \subseteq X$. Khi đó $J = \bigcup_{x \in A} I_x$ hoặc là một khoảng gốc hoặc bằng $X$.*

*Chứng minh.* Giả sử $J \neq X$, ta sẽ chứng minh $J$ là một khoảng gốc. Vì $X$ được sắp tốt, đặt $b = \min(X \setminus J)$. Nếu $x \in J$ suy ra tồn tại $a \in A$ sao cho $x \in I_a$ hay $x < a$. Giả sử $x \geq b$. Khi đó $b < a$ hay $b \in I_a \subseteq J$ vô lý. Vậy $x < b$, nghĩa là $x \in I_b$. Vậy $J \subseteq I_b$. Ngược lại, với $x \in I_b$ ta có $x < b$, nên $x \notin X \setminus J$ hay $x \in J$. Vậy $I_b \subseteq J$. Do đó, $J = I_b$. □

**4.3.4 Định nghĩa.** Hai tập được sắp thứ tự $X$ và $Y$ gọi là *đẳng cấu thứ tự* với nhau nếu tồn tại một song ánh $f : X \to Y$ thỏa mãn $x_1 \leq x_2$ nếu và chỉ nếu $f(x_1) \leq f(x_2)$ với mọi $x_1, x_2 \in X$.

**4.3.5 Định lý.** *Cho $X$ và $Y$ là hai tập được sắp tốt. Khi đó, hoặc $X$ đẳng cấu thứ tự với $Y$, hoặc $X$ đẳng cấu thứ tự với một khoảng gốc của $Y$, hoặc $Y$ đẳng cấu thứ tự với một khoảng gốc của $X$.*

*Chứng minh.* Xét họ $\mathcal{F}$ các đẳng cấu thứ tự từ một khoảng gốc của $X$ hoặc từ $X$ lên một khoảng gốc của $Y$ hoặc $Y$. Rõ ràng $f : \{\min X\} \to \{\min Y\}$ thuộc $\mathcal{F}$ nên $\mathcal{F} \neq \varnothing$. Sắp thứ tự tập $\mathcal{F}$ bởi quan hệ bao hàm. Khi đó, ta có thể kiểm tra một cách trực tiếp được rằng tập sắp thứ tự $\mathcal{F}$ thỏa điều kiện bổ đề Zorn, cho nên nó có phần tử tối đại là $f : A \to B$. Nếu $A = I_x$ và $B = I_y$. Giả sử $x \notin A$ và $y \notin B$. Khi đó $A \cup \{x\}$ và $B \cup \{y\}$ cũng là khoảng gốc hoặc bằng $X$ hay $Y$ nên $f$ mở rộng được bằng cách đặt thêm $f(x) = y$, ta gặp mâu thuẫn với tính tối đại của $f$. Vậy $x \in A$ hay $y \in B$. Do đó theo quy nạp siêu hạn ta có $A = X$ hoặc $B = Y$. □

### Bài tập

**1**) Cho $A_i = \{1, 2, \dots, i\}$ với $i = 1, 2, 3, \dots$. Tìm $\bigcup\limits_{i=1}^{n} A_i$ và $\bigcap\limits_{i=1}^{n} A_i$.

**2**) Gọi $\{R_i\}_{i\in I}$ với $I \neq \varnothing$ là họ các quan hệ tương đương trên $X$. Chứng minh rằng $\bigcap\limits_{i\in I} R_i$ là một quan hệ tương đương trên $X$.

**3**) Cho $\{R_i\}_{i\in I}$ với $I \neq \varnothing$ là họ các quan hệ thứ tự trong $X$. Chứng minh rằng $R = \bigcap\limits_{i\in I} R_i$ là quan hệ thứ tự trong $X$.

## § 5 Tập số tự nhiên, số nguyên và số hữu tỉ

Trước đây chúng ta đã dùng các tập số tự nhiên, số nguyên, số hữu tỉ và số thực để minh họa một số khái niệm ở các bài trước. Ta sẽ trình bày định nghĩa các tập số tự nhiên, số nguyên và số hữu tỉ ở mục này, tập số thực sẽ trình bày ở mục sau; trong khi đó tập số thực được trình bày một cách đầy đủ trong *Giải tích* nên không được trình bày ở chương này. Những khái niệm và tính chất của các tập hợp số được trình bày ở đây chỉ mang tính sơ lược; trình bày chi tiết và đầy đủ về chúng sẽ được trình bày trong *Đại số đại cương* và *Cơ sở số học*.

### 5.1 Tập số tự nhiên

**5.1.1 Định nghĩa. (Hệ tiên đề Peano)** Tập hợp các số tự nhiên $\mathbb{N}$ là một tập quy nạp. Giả sử $S$ là một bộ phận của $\mathbb{N}$ sao cho hai điều kiện sau thỏa mãn

(1) $\varnothing \in S$

(2) $x \in S$ kéo theo $x^+ \in S$

thế thì $S = \mathbb{N}$.

$$\mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$$

### 5.2 Tập Số Nguyên

$\mathbb{Z} = \{n : n \in \mathbb{N} \text{ hay } -n \in \mathbb{N}\} = \{\dots, -2, -1, 0, 1, 2, \dots\}$

### 5.3 Tập Số Hữu Tỉ

$\mathbb{Q} = \{\frac{m}{n} : m, n \in \mathbb{Z}, n \neq 0\}$. Tập $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$ được gọi là tập số vô tỉ.

### Bài tập

**1)** Dùng quy nạp chứng minh rằng $\displaystyle\sum_{k=4}^{n} \frac{k!}{(k-4)!} = \frac{(n+1)!}{5 \cdot (n-4)!}$.

## § 6 Lực lượng của tập hợp

### 6.1 Lực lượng của tập hợp

Cho các tập $X$ và $Y$. Nếu tồn tại một đơn ánh $f : X \to Y$ thì ta viết $\mathrm{card}(X) \leq \mathrm{card}(Y)$; nếu tồn tại một song ánh $f : X \to Y$ thì ta viết $\mathrm{card}(X) = \mathrm{card}(Y)$; nếu tồn tại một đơn ánh $f : X \to Y$ nhưng không tồn tại một song ánh từ $X$ lên $Y$ thì ta viết $\mathrm{card}(X) < \mathrm{card}(Y)$. Ta gọi $\mathrm{card}(X)$ là *lực lượng* (hay *bản số*) của tập $X$.

Từ định nghĩa, hiển nhiên $X \subset Y$ thì $\mathrm{card}(X) \leq \mathrm{card}(Y)$.

**6.1.1 Định lý.** *Cho $X \neq \varnothing$. Khi đó $\mathrm{card}(X) \leq \mathrm{card}(Y)$ nếu và chỉ nếu tồn tại một toàn ánh $g : Y \to X$.*

*Chứng minh.* Nếu $\mathrm{card}(X) \leq \mathrm{card}(Y)$ thì tồn tại đơn ánh $f : X \to Y$. Lấy cố định $x_0 \in X$. Xác định ánh xạ $g : Y \to X$ với $g(y) = f^{-1}(y)$ nếu $y \in f(X)$ và $g(y) = x_0$ nếu $y \notin f(X)$. Khi đó, $g$ là toàn ánh.

Ngược lại, nếu có $g : Y \to X$ là một toàn ánh thì họ các tập $\{g^{-1}(x)\}_{x \in X}$ là khác rỗng và rời nhau; do đó theo tiên đề chọn với mọi $x \in X$ có phần tử $f(x) \in g^{-1}(x)$. Ta xác định đơn ánh $f : X \to Y$ với $x \mapsto f(x)$. Do đó, $\mathrm{card}(X) \leq \mathrm{card}(Y)$. □

**6.1.2 Định lý.** *Với các tập $X$ và $Y$ thì hoặc $\mathrm{card}(X) \leq \mathrm{card}(Y)$ hoặc $\mathrm{card}(Y) \leq \mathrm{card}(X)$.*

*Chứng minh.* Ký hiệu $\mathcal{F}$ là họ các đơn ánh từ một tập con của $X$ vào $Y$. Ta xem $\mathcal{F}$ là họ tập con của $X \times Y$ và sắp thứ tự $\mathcal{F}$ theo quan hệ bao hàm $\subseteq$. Khi đó $\mathcal{F}$ thỏa mãn giả thiết của bổ đề Zorn, do đó trong $\mathcal{F}$ có phần tử tối đại $f : A \to Y$, $A \subseteq X$. Đặt $B = f(A)$. Nếu đồng thời có $x_0 \in X \setminus A$ và $y_0 \in Y \setminus B$ thì $f$ mở rộng đến một đơn ánh $\overline{f}$ từ $A \cup \{x_0\}$ vào $Y$ có $\overline{f}(A \cup \{x_0\}) = B \cup \{y_0\}$ bằng cách đặt $\overline{f}(x_0) = y_0$. Như vậy ta gặp phải mâu thuẫn với tính tối đại của $f$. Vì vậy hoặc $A = X$, khi đó $f$ là đơn ánh nên $\mathrm{card}(X) \leq \mathrm{card}(Y)$, hoặc $B = Y$ khi đó $f^{-1} : Y \to A \subseteq X$ là đơn ánh nên $\mathrm{card}(Y) \leq \mathrm{card}(X)$. □

**6.1.3 Định lý. (Schöder-Bernstein)** *Cho hai tập hợp $X$ và $Y$. Khi đó, nếu $\mathrm{card}(X) \leq \mathrm{card}(Y)$ và $\mathrm{card}(Y) \leq \mathrm{card}(X)$ thì $\mathrm{card}(X) = \mathrm{card}(Y)$.*

*Chứng minh.* Giả sử $f : X \to Y$ và $g : Y \to X$ là các đơn ánh. Xác định ánh xạ $\varphi : \mathscr{P}(X) \to \mathscr{P}(X)$ bởi $\varphi(E) = X \setminus g(Y \setminus f(E))$ với mọi $E \in \mathscr{P}(X)$. Nếu $D, E \in \mathscr{P}(X)$ thỏa $D \subseteq E$ thì rõ ràng $\varphi(D) \subseteq \varphi(E)$.

Ký hiệu $\mathcal{F} = \{E \in \mathscr{P}(X) : E \subseteq \varphi(E)\}$ và đặt

$$W = \bigcup_{E \in \mathcal{F}} E.$$

Mọi $E \in \mathcal{F}$ ta có $E \subseteq W$ nên $E \subseteq \varphi(E) \subseteq \varphi(W)$. Từ đó suy ra $W \subseteq \varphi(W)$, và suy ra tiếp $\varphi(W) \subseteq \varphi(\varphi(W))$, tức là $\varphi(W) \in \mathcal{F}$. Do đó $\varphi(W) \subseteq W$. Vậy $W = \varphi(W)$. Theo định nghĩa của $\varphi$ ta được

$$W = \varphi(W) = X \setminus g(Y \setminus f(W)) \quad \text{hay} \quad X \setminus W = g(Y \setminus f(W)).$$

Từ kết quả này và do $f$ và $g$ là đơn ánh nên suy ra ánh xạ $h : X \to Y$ xác định bởi

$$h(x) = \begin{cases} f(x) & \text{khi } x \in W \\ g^{-1}(x) & \text{khi } x \in X \setminus W \end{cases}$$

là một song ánh. Vậy $\operatorname{card}(X) = \operatorname{card}(Y)$. □

**6.1.4 Định lý.** *Với mọi tập hợp $X$ đều có* $\operatorname{card}(X) < \operatorname{card}(\mathscr{P}(X))$.

*Chứng minh.* Giả sử ngược lại, nghĩa là tồn tại một tập $X$ sao cho $\operatorname{card}(\mathscr{P}(X)) \leq \operatorname{card}(X)$. Khi đó tồn tại một toàn ánh $f : X \to \mathscr{P}(X)$. Xét tập $E = \{x \in E : x \notin f(x)\}$. Do $f$ là toàn ánh nên tồn tại $x_0 \in X$ sao cho $f(x_0) = E$. Khi đó,

- $x_0 \in E$ thì mâu thuẫn vì $x_0 \notin f(x_0) = E$
- $x_0 \notin E$ thì mâu thuẫn $x_0 \in f(x_0) = E$.

Vậy ta phải có $\operatorname{card}(X) < \operatorname{card}(\mathscr{P}(X))$ với mọi tập $X$. □

## 6.2 Tập đếm được

**6.2.1 Định nghĩa.** Một tập $X$ được gọi là tập **hữu hạn** nếu tồn tại số nguyên dương $n$ và một song ánh $f : X \to \{1, 2, \ldots, n\}$. Số $n$ này là duy nhất và gọi là lực lượng của $X$ hay số phần tử trong $X$. Vậy $\operatorname{card}(X) = n$.

Tập rỗng được quy ước là tập hữu hạn và $\mathrm{card}(\varnothing) = 0$. Một tập không là hữu hạn được gọi là **vô hạn**

**6.2.2 Định nghĩa.** Cho $X$ là một tập hợp vô hạn. Nếu có thể thiết lập một song ánh giữa tập mọi số tự nhiên $\mathbb{N}$ và tập hợp $X$ thì ta nói tập $X$ có lực lượng **vô hạn đếm được** (hay gọi tắt là đếm được).

Một tập hoặc là hữu hạn hoặc là vô hạn đếm được gọi chung là **tập đếm được**. Như vậy nếu tập $X$ là đếm được khi và chỉ khi $\mathrm{card}(X) \leq \mathrm{card}(\mathbb{N})$.

Theo định nghĩa trên, hiển nhiên tập mọi số tự nhiên $\mathbb{N}$ là vô hạn đếm được, ký hiệu $\mathrm{card}(\mathbb{N}) = \mathfrak{w}$. Sau đây là vài ví dụ khác:

Tập hợp mọi số tự nhiên chẵn là vô hạn đếm được. Song ánh $f$ giữa $\mathbb{N}$ và tập hợp này có thể xác định bởi tương ứng

$$n \in \mathbb{N} \qquad \Longleftrightarrow \qquad 2n \in X.$$

Tập hợp mọi số tự nhiên lẻ, tập hợp mọi số nguyên $\mathbb{Z}$, tập hợp mọi số dạng $1/n$, $n \in \mathbb{N}$ là các đếm được. Một song ánh giữa $\mathbb{N}$ và $\mathbb{Z}$ có thể xác lập như sau:

$$n \qquad \Longleftrightarrow \qquad \begin{cases} -k & \text{nếu } n = 2k \quad (k = 1, 2, 3, \ldots) \\ k & \text{nếu } n = 2k + 1 \quad (k = 0, 1, 2, 3, \ldots) \end{cases}$$

**6.2.3 Định lý.** **(a)** *Khi thêm hoặc bớt một tập hữu hạn từ tập vô hạn đếm được thì ta lại được một vô hạn tập đếm được.*

**(b)** *Nếu $X$ và $Y$ đếm được thì $X \times Y$ đếm được.*

**(c)** *Hợp của một số hữu hạn hoặc vô hạn đếm được các tập đếm được là tập đếm được.*

*Chứng minh.*

**(a)** Cho $X$ là một tập vô hạn đếm được. Khi đó, tồn tại một song ánh $f : \mathbb{N} \to X$. Giả sử $A$ là một tập hữu hạn; không mất tính tổng quát ta có thể giả sử $A \cap X = \varnothing$. Khi đó, tồn tại $n$ và một song ánh $g : \{1, 2, \ldots, n\} \to A$. Khi đó, dễ dàng kiểm tra rằng ánh xạ $h : \mathbb{N} \to A \cup X$ xác định bởi

$$h(k) = \begin{cases} f(0) & \text{khi } k = 0 \\ g(k) & \text{khi } k = 1, 2, \ldots, n \\ f(k-n) & \text{khi } k \in \mathbb{N} \setminus \{0, 1, 2, \ldots, n\} \end{cases}$$

là một song ánh, cho nên $\operatorname{card}(X \cup A) = \operatorname{card}(\mathbb{N})$.

Mặt khác, giả sử $B$ là một tập con hữu hạn của $X$. Khi đó, tồn tại một song ánh $\varphi : \{1, 2, \ldots, n\} \to B$. Đặt $\varphi(k) = b_k$ và đặt $i_k = f^{-1}(b_k)$ với $k = 1, 2, \ldots, n$. Khi đó, ánh xạ $\psi : \mathbb{N} \setminus \{i_1, i_2, \ldots, i_n\} \to \mathbb{N}$ xác định bởi

$$\psi(k) = \begin{cases} k & \text{khi } k \leq m = \max\{i_1, i_2, \ldots, i_n\} \\ i_{k-m} & \text{khi } k = m+1, m+2, \ldots, m+n \\ \multicolumn{2}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ k-n & \text{khi } k > m+n \end{cases}$$

là một song ánh. Từ đó ta kiểm tra được ánh xạ $f|_{\mathbb{N}\setminus\{i_1, i_2, \ldots, i_n\}} \circ \psi^{-1} : \mathbb{N} \to X \setminus B$ là một song ánh, cho nên $\operatorname{card} X \setminus B = \operatorname{card}(\mathbb{N}) = \mathfrak{w}$.

**(b)** Ta chỉ cần chứng minh $\mathbb{N}^2$ đếm được. Liệt kê các phần tử $(i, j)$ của $\mathbb{N}^2$ theo thứ tự tăng dần của tổng $i + j$ như sau:

$$(1,1), (1,2), (2,1), (1,3), (2,2), (3,1), (1,4), (2,3), (3,2), (4,1), \ldots$$

Từ đó ta thiết lập được song ánh $f : \mathbb{N} \to \mathbb{N}^2$, trong đó $f(n)$ là phần tử thứ $n$ trong dãy trên. Vậy $\operatorname{card}(\mathbb{N}^2) = \operatorname{card}(\mathbb{N}) = \mathfrak{w}$.

**(c)** Với mọi $\alpha \in I$ tồn tại một toàn ánh $f_\alpha : \mathbb{N} \to X_\alpha$. Từ đó ánh xạ $f : \mathbb{N} \times I \to \bigcup\limits_{\alpha \in I} X_\alpha$ xác định bởi $f(n, \alpha) = f_\alpha(n)$ là toàn ánh. Theo kết quả trên $\mathbb{N} \times I$ là đếm được, suy ra

$$\operatorname{card}\Big(\bigcup_{\alpha \in I} X_\alpha\Big) \leq \operatorname{card}(\mathbb{N} \times I) = \mathfrak{w}. \qquad \square$$

**6.2.4 Hệ quả.** $\mathbb{Q}$ *là các tập đếm được.*

*Chứng minh.* Do $\mathbb{Z}$ đếm được nên $\mathbb{Z}^2$ cũng đếm được. Hơn nữa ánh xạ $f : \mathbb{Z}^2 \to \mathbb{Q}$ xác định bởi

$$f(m,n) = \begin{cases} \dfrac{m}{n} & \text{khi } n \neq 0 \\ 0 & \text{khi } n = 0 \end{cases}$$

là toàn ánh, cho nên $\mathbb{Q}$ là đếm được. Từ đó, dễ thấy rằng $\mathrm{card}(\mathbb{Z}^2) = \mathrm{card}(\mathbb{Q}) = \mathrm{card}(\mathbb{N}) = \mathfrak{w}$. □

## 6.3 Tập continum

Ở trên ta đã gặp các tập vô hạn đếm được. Nếu không tồn tại một song ánh nào giữa tập $\mathbb{N}$ mọi số tự nhiên và tập vô hạn $X$, thì $X$ được gọi là tập **không đếm được**. Người ta chứng minh được tập mọi số thực thuộc đoạn $[0,1]$, hay tổng quát hơn tập hợp mọi số thực thuộc khoảng $[a,b]$ với $a \neq b$, là không đếm được. Người ta gọi chúng là các tập hợp có **lực lượng continuum**. Một cách tổng quát ta có định nghĩa sau.

**6.3.1 Định nghĩa.** Tập $X$ được gọi là có **lực lượng continuun** nếu $\mathrm{card}(X) = \mathrm{card}([0,1])$. Ký hiệu $\mathrm{card}([0,1]) = \mathfrak{c}$.

**6.3.2 Định lý.** $\mathrm{card}(\mathscr{P}(\mathbb{N})) = \mathfrak{c}$.

*Chứng minh.* Xét ánh xạ $f : \mathscr{P}(\mathbb{N}^*) \to [0,2]$ xác định bởi $f(\varnothing) = 0$ và nếu $A \neq \varnothing$ thì

$$f(A) = \begin{cases} \sum\limits_{n\in A} 2^{-n} & \text{khi } \mathbb{N} \setminus A \text{ vô hạn} \\ 1 + \sum\limits_{n\in A} 2^{-n} & \text{khi } \mathbb{N} \setminus A \text{ hữu hạn} \end{cases}$$

Nếu $\mathbb{N}^* \setminus A$ vô hạn thì $f(A)$ có khai triển thập phân cơ số 2 là $0.a_1a_2\ldots$ trong đó $a_i = 1$ nếu $i \in A$ và $a_i = 0$ nếu $i \notin A$ và loại trừ trường hợp từ một lúc nào đó tất cả các $a_i$ đều bằng 1 (do $\mathbb{N}^* \setminus A$ vô hạn). Từ đó các giá trị của $f(A)$ là khác nhau và lấp đầy khoảng $[0,1)$.

Nếu $\mathbb{N}^* \setminus A$ hữu hạn thì $f(A) \in (1, 2]$. Hơn nữa, ta có

$$\sum_{n\in A} 2^{-n} + \sum_{n\in\mathbb{N}^*\setminus A} 2^{-n} = \sum_{n\in\mathbb{N}^*} 2^{-n} = 1,$$

do đó nếu $A_1 \neq A_2$ thì

$$\sum_{n\in\mathbb{N}\setminus A_1} 2^{-n} \neq \sum_{n\in\mathbb{N}^*\setminus A_2} 2^{-n} \qquad \text{nên} \qquad \sum_{n\in A_1} 2^{-n} \neq \sum_{n\in A_2} 2^{-n};$$

vì vậy $f(A_1) \neq f(A_2)$. Do đó $f$ là đơn ánh, và suy ra $\mathrm{card}(\mathscr{P}(\mathbb{N}^*)) \leq \mathrm{card}([0, 2])$ hay $\mathrm{card}(\mathscr{P}(\mathbb{N}^*)) \leq \mathfrak{c}$.

Bây giờ xét ánh xạ $g : \mathscr{P}(\mathbb{N}^*) \to [0, 1]$ xác định bởi $g(\varnothing) = 0$ và nếu $A \neq \varnothing$ thì $g(A) = \sum\limits_{n\in A} 2^{-n}$. Theo phân tích trên, $g$ là toàn ánh do đó, $\mathrm{card}(\mathscr{P}(\mathbb{N}^*)) \geq \mathrm{card}([0, 1]) = \mathfrak{c}$. Do đó theo định lý Schöder-Bernstein ta có $\mathrm{card}(\mathscr{P}(\mathbb{N}^*)) = c$. Từ đó ta suy ra được $\mathrm{card}(\mathscr{P}(\mathbb{N})) = \mathfrak{c}$. □

**6.3.3 Hệ quả.** *Nếu* $\mathrm{card}(X) \geq \mathfrak{c}$ *thì* $X$ *không đếm được.*

*Chứng minh.* Theo các định lý 6.1.4 và 6.3.2 ta có

$$\mathfrak{w} = \mathrm{card}(\mathbb{N}) < \mathrm{card}(\mathscr{P}(\mathbb{N})) = \mathfrak{c}.$$

Do $\mathrm{card}(X) \geq \mathfrak{c}$ thì $\mathrm{card}(X) > \mathfrak{w}$ nên $X$ không đếm được. □

Mệnh đề ngược của hệ quả này nổi tiến với tên gọi giả thiết continum.

**6.3.4 Giả thiết continum.** Nếu $\mathrm{card}(X) < c$ thì $X$ đếm được.

Người ta đã chứng minh giả thiết continum độc lập với các tiên đề khác của lý thuyết tập hợp.

**6.3.5 Định lý.** **(a)** *Nếu* $\mathrm{card}(X) \leq \mathfrak{c}$ *và* $\mathrm{card}(Y) \leq \mathfrak{c}$ *thì* $\mathrm{card}(X \times Y) \leq \mathfrak{c}$.

**(b)** *Nếu* $\mathrm{card}(I) \leq \mathfrak{c}$ *và* $\mathrm{card}(X_\alpha) \leq \mathfrak{c}$ *với mọi* $\alpha \in I$ *thì* $\mathrm{card}\left(\bigcup\limits_{\alpha\in I} X_\alpha\right) \leq \mathfrak{c}$.

*Chứng minh.* **(a)** Ta chỉ cần xét trường hợp $X = Y = \mathscr{P}(\mathbb{N})$. Gọi $\varphi$ và $\psi$ là các ánh xạ từ $\mathbb{N}$ vào $\mathbb{N}$ xác định bởi $\varphi(n) = 2n$ và $\psi(n) = 2n + 1$. Dễ thấy rằng ánh xạ

$$\begin{aligned} f : \mathscr{P}(\mathbb{N})^2 &\to \mathscr{P}(\mathbb{N}) \\ (A, B) &\mapsto \varphi(A) \cup \psi(B) \end{aligned}$$

là đơn ánh do đó $\operatorname{card}(\mathscr{P}(\mathbb{N})^2) \leq \operatorname{card}(\mathbb{N}) = \mathfrak{c}$.

**(b)** Với mọi $\alpha \in I$ tồn tại một toàn ánh $f_\alpha : [0, 1] \to X_\alpha$. Từ đó ta có ánh xạ

$$\begin{aligned} f : [0, 1] \times I &\to \bigcup_{\alpha \in I} X_\alpha \\ (a, \alpha) &\mapsto f_\alpha(a) \end{aligned}$$

là toàn ánh. Vì vậy $\operatorname{card}\left(\bigcup_{\alpha \in I} X_\alpha\right) \leq \operatorname{card}([0, 1] \times I) \leq \mathfrak{c}$ (theo kết quả phần (a)). □

**6.3.6 Hệ quả.** *Tập số thực $\mathbb{R}$, tập số phức $\mathbb{C}$ là tập continum.*

## 6.4 Về lực lượng của tập được sắp tốt

**6.4.1 Định lý.** *Tồn tại một tập sắp tốt không đếm được $\Omega$ sao cho $I_x$ đếm được với mọi $x \in \Omega$. Nếu $\Omega'$ là một tập khác rỗng cũng có tính chất đó thì $\Omega$ và $\Omega'$ đẳng cấu thứ tự.*

*Chứng minh.* Lấy một tập không đếm được $X$ và một quan hệ thứ tự tốt trên nó (tồn tại do nguyên lý sắp tốt). Khi đó hoặc $X$ có tính chất đòi hỏi của định lý hoặc tồn tại phần tử nhỏ nhất $x_0$ sao cho $I_{x_0}$ không đếm được. Chọn $\Omega = I_{x_0}$; khi đó $\Omega$ có tính chất mong muốn (do tính nhỏ nhất của $x_0$). Vì các khoảng gốc của $\Omega$ và $\Omega'$ đều đếm được nên $\Omega$ không đẳng cấu thứ tự với một khoảng gốc nào của $\Omega'$ và $\Omega'$ cũng không đẳng cấu thứ tự với một khoảng gốc của $\Omega$. Theo định lý trên ta phải có $\Omega$ và $\Omega'$ đẳng cấu thứ tự. □

Tập $\Omega$ trong định lý trên gọi là *tập các tự số đếm được*. Ta có tính chất sau đây của $\Omega$.

**6.4.2 Định lý.** *Mọi tập con đếm được của $\Omega$ trong định lý trên đều có cận trên.*

*Chứng minh.* Cho $A \subset \Omega$ là một tập đếm được. khi đó $\bigcup_{a\in A} I_x$ là tập đếm được. Theo Định lý 4.3.3, tồn tại $y \in \Omega$ sao cho $\bigcup_{a\in A} I_x = I_y$. Vậy $A$ có cận trên là $y$, nếu ngược lại tồn tại $x \in A$ sao cho $y < x$ suy ra $y \in I_x$ nên $I_y \subset I_x \subset I_y$ vì thế $y \in I_y$ vậy $y < y$ vô lý. □

Đặt $f(1) = \min \Omega$, $f(n) = \min(\Omega \setminus \{f(1), \ldots, f(n-1)\})$ với mọi số tự nhiên $n > 1$. Ta có một đẳng cấu thứ tự từ tập số tự nhiên khác không $\mathbb{N}^*$ lên khoảng gốc $I_\omega$ của $\Omega$. Ta có $\omega$ là phần tử nhỏ nhất của $\Omega$ sao cho $I_\omega$ vô hạn.

Bổ sung thêm vào $\Omega$ phần tử $\omega_1$. Ký hiệu $\Omega^* = \Omega \cup \{\omega_1\}$, mở rộng thứ tự từ $\Omega$ lên $\Omega^*$ bằng cách đặt $x < \omega_1$ với mọi $x \in \Omega$, $\omega_1$ gọi là *tự số không đếm được đầu tiên*.

## Bài tập

**1 )** Chứng minh rằng các tập sau đây cùng lực lượng bằng cách chỉ ra có một song ánh giữa chúng

**(a)** $[0,1]$ và $(0,1)$ **(b)** $(0,1)$ và $(0,\infty)$

**(c)** $\mathscr{P}(\mathbb{N})$ và tập $2^{\mathbb{N}}$ các ánh xạ từ $\mathbb{N}$ vào tập $\{0,1\}$.

**2 )** Chứng minh rằng $\mathrm{card}(\mathscr{P}(\mathbb{N})) = \mathrm{card}(\mathscr{P}(\mathbb{N}^*))$.

**3 )** Xây dựng họ $\mathcal{U}$ các tập con của $\mathbb{N}$ thỏa mãn các điều kiện.

**(i)** $\mathcal{U}$ có lực lượng continum

**(ii)** $\mathrm{card}(A \cap B) < \mathfrak{w}$ với mọi $A, B \in \mathcal{U}$.

**4 )** Xây dựng họ $\mathcal{U}$ các tập con của $\mathbb{N}$ thỏa mãn

**(i)** $\mathcal{U}$ có lực lượng continum

**(ii)** Mọi $t \in \mathbb{N}$ và $A, B \in \mathcal{U}$ bất đẳng thức $|a-b| < t$ chỉ xảy ra với hữu hạn $a \in A$, $b \in B$.

**5)** Ký hiệu $\mathcal{F}(\mathbb{N})$ là họ các tập con hữu hạn của $\mathbb{N}$. Chứng minh

**(a)** Họ $\mathcal{F}(\mathbb{N})$ đếm được.

**(b)** Tồn tại ánh xạ $\varphi : \mathcal{F}(\mathbb{N}) \to \mathbb{N}$ thỏa mãn $\varphi(A) \leq \varphi(B)$ với mọi $A, B \in \mathcal{F}$, $A \subseteq B$.

**6)** Cho $\{A_n\}$ là dãy các tập con hữu hạn của $\mathbb{N}^*$ thỏa mãn. Mọi tập con vô hạn $B$ của $\mathbb{N}^*$, tồn tại $n \in \mathbb{N}^*$ sao cho $\mathrm{card}(A_n \cap B) \geq 2$.

**(a)** Chứng minh rằng tập $A = \{k \in \mathbb{N}^* : k$ chỉ thuộc hữu hạn tập $A_n\}$ là hữu hạn.

**(b)** Tồn tại tập vô hạn $E \subseteq \mathbb{N}^*$ sao cho $E \setminus \bigcup\limits_{n \in \Delta_k} A_n \subseteq \{1, 2, \ldots, k-1\}$ với mọi $k \in E$, ở đây $\Delta_k = \{n \in \mathbb{N}^* : k \in A_n\}$

**(c)** Giả sử $\{B_n\}$ cũng là dãy có tính chất như dãy $\{A_n\}$. Chứng minh rằng tồn tại $p, q \in \mathbb{N}^*$ sao cho $\mathrm{card}(A_p \cap B_q) \geq 2$.

# § 7 Các cấu trúc đại số cơ bản

## 7.1 Phép toán hai ngôi

**7.1.1 Định nghĩa.** Cho $X$ là một tập hợp không rỗng. Một **phép toán hai ngôi trên** $X$ là một ánh xạ từ tích Descartes $X \times X$ đến $X$. Nói cách khác, một *phép toán hai ngôi trên* $X$ là một qui tắc đặt tương ứng mỗi cặp có thứ tự $(x, y)$ của hai phần tử $x, y$ tùy ý thuộc $X$ với một phần tử hoàn toàn xác định thuộc $X$ mà được gọi là cái *hợp thành* của $x$ và $y$ bởi phép toán đó.

Thông thường, cái hợp thành của $x$ và $y$ được ký hiệu bằng cách viết $x$ và $y$ theo thứ tự đó rồi chen vào giữa chúng một dấu đặc trưng cho phép toán. Chẳng hạn

$$x * y, \qquad x \circ y, \qquad x \cdot y \text{ (hay } xy), \qquad x + y, \ldots$$

Khi dùng ký hiệu $x + y$ (tương ứng $x \cdot y$ hay $xy$), phép toán được gọi là *phép cộng* (tương ứng, *phép nhân*) còn $x + y$ (tương ứng, $xy$) được gọi là *tổng* (tương ứng, *tích*) của $x$ và $y$.

**7.1.2** Giả sử $X$ là một tập hợp không rỗng mà trên nó đã cho một phép toán $*$.

**(i)** Ta bảo $*$ là phép toán có tính chất *kết hợp* nếu với mọi $x, y, z \in X$ ta có

$$(x * y) * z = x * (y * z).$$

**(ii)** Ta bảo $*$ là phép toán có tính chất *giao hoán* nếu với mọi $x, y \in X$ ta có

$$x * y = y * x.$$

**7.1.3** Giả sử $X$ là một tập không rỗng mà trên nó có hai phép toán $*$ và $\circ$. Ta bảo $*$ có tính chất *phân phối bên trái* (tương ứng, *bên phải*) đối với $*$ nếu với mọi $x, y, z \in X$ ta có

$$x * (y \circ z) = (x * y) \circ (x * z) \quad \text{(tương ứng } (x \circ y) * z = (x * z) \circ (y * z)\text{)}.$$

Khi $*$ vừa phân phối bên trái vừa phân phối bên phải đối với $\circ$ thì ta nói $*$ có tính chất *phân phối* đối với $\circ$.

**7.1.4** Cho $*$ là một phép toán trên một tập hợp không rỗng $X$ nào đó.

**(i)** Phần tử $e \in X$ được gọi là *phần tử trung hòa* (đối với $*$) nếu với mọi $x \in X$ ta có

$$x * e = e * x = x.$$

Khi $*$ ký hiệu theo lối cộng (tương ứng, nhân) thì phần tử trung hoàn thường được ký hiệu là 0 (tương ứng, 1) và gọi là *phần tử không* (tương ứng, *phần tử đơn vị*)

**(ii)** Giả sử phép toán $*$ có phần tử trung hòa $e$. Lấy một phần tử tùy ý $x \in X$. Ta nói $x$ là phần tử *khả đối xứng* nếu tồn tại $x' \in X$ sao cho

$$x * x' = x' * x = e.$$

Khi đó, phần tử $x'$ được gọi là *phần tử đối xứng* của $x$ (đối với phép toán $*$). Khi phép toán $*$ ký hiệu theo lối cộng (tương ứng, nhân) thì phần tử khả đối xứng $x$ còn gọi là *khả đối* (tướng ứng, *khả nghịch*), phần tử đối xứng của $x$ được ký hiệu là $-x$ (tương ứng, $x^{-1}$) và được gọi là *phần tử đối* (tương ứng, *phần tử nghịch đảo*) của $x$.

## 7.2 Các cấu trúc đại số cơ bản

**7.2.1 (Nhóm)** Một tập hợp không rỗng cùng với một phép toán đã cho trên nó gọi là một *nhóm* nếu phép toán đã cho kết hợp, có phần tử trung hòa và mọi phần tử để khả đối xứng. Một nhóm được gọi là *nhóm giao hoán* nếu phép toán trên nó có tính chất giao hoán.

**7.2.2 (Vành)** Một tập hợp không rỗng $X$ cùng với hai phép toán ký hiệu theo lối cồng và nhân đã cho trên nó gọi là một *vành* nếu $X$ với phép cộng là một nhóm giao hoán, phép nhân kết hợp và phân phối đối với phép cộng. Vành mà phép nhân có đơn vị gọi là *vành có đơn vị*. Vành mà phép nhân giao hoán gọi là *vành giao hoán*.

**7.2.3 (Trường)** Cho $X$ cùng với hai phép toán cộng và nhân là một vành giao hoán có đơn vị. Ký hiệu 0 và 1 lần lượt là phần tử không (của phép cộng) và phần tử đơn vị (của phép nhân). Giả thiết rằng $0 \neq 1$. Khi đó $X$ gọi là một *trường* nếu mọi phần tử khác 0 trong $X$ đều khả nghịch.

**7.2.4 Thí dụ.** **(i)** Tập hợp các số nguyên $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ là một nhóm cộng nhưng không là nhóm nhân.

Tập hợp các số tự nhiên $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ không là một nhóm cộng cũng không là một nhóm nhân.

**(ii)** Tập hợp các số nguyên $\mathbb{Z}$ với phép cộng và phép nhân là một vành giao hoán có đơn vị nhưng không là một trường. Do đó, $\mathbb{Z}$ được gọi là *vành số nguyên*.

**(iii)** Tập hợp các số hữu tỉ $\mathbb{Q}$ với phép cộng và phép nhân là một trường và được gọi là *trường số hữu tỉ*.

**(iv)** Tập hợp các số thực $\mathbb{R}$ với phép cộng và phép nhân là một trường và được gọi là *trường số thực*.

# § 8 Phép thế

## 8.1 Các khái niệm cơ bản

**8.1.1 Định nghĩa.** Cho $X$ là tập hợp khác rỗng. Một song ánh $\sigma : X \to X$ được gọi là một **phép thế trên** $X$. Tập hợp tất cả các phép thế trên $X$ được kí hiệu là: $S_X$. Vậy

$$S_X = \{\sigma : X \to X \mid \sigma \text{ là song ánh}\}.$$

Đặc biệt, nếu tập $X$ có $n$ phần tử thì phép thế trên $X$ gọi là **phép thế bậc** $n$. Tập tất cả các phép thế bậc $n$ trên $X$ được kí hiệu là $S_n$.

Khi tập $X$ có $n$ phần tử, không mất tính tổng quát ta có thể đánh số các phần tử của $X$, cho nên ta có thể giả sử $X = \{1, 2, 3, \ldots, n\}$. Khi đó, phép thế bậc $n$ là một song ánh $\sigma : \{1, 2, 3, \ldots, n\} \to \{1, 2, 3, \ldots, n\}$. Người ta thường kí hiệu phép thế bậc $n$ dưới dạng:

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

**8.1.2 *Chú ý.*** Từ định nghĩa ta có thể thấy mỗi phép thế bậc $n$ là một hoán vị (permutation) của $n$ phần tử. Do đó $S_n$ có $n!$ phần tử.

Phép thế $\sigma$ bậc $n$ còn có thể được viết dưới nhiều dạng khác nhau bằng cách thay đổi thứ tự các phần tử ở dòng trên và đồng thời thay đổi thứ tự các phần tử tương ứng ở dòng dưới. Tổng quát ta có thể viết phép thế $\sigma$ dưới dạng

$$\sigma = \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ \sigma(i_1) & \sigma(i_2) & \ldots & \sigma(i_n) \end{pmatrix}$$

trong đó $i_1, i_2, \ldots, i_n$ là hoán vị của $n$ phần tử $1, 2, \ldots, n$.

**8.1.3 Thí dụ.** Ta có $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 4 & 3 \\ 4 & 2 & 1 & 3 \end{pmatrix}$ là một phép thế bậc 4 trong đó $1 \mapsto 2$, $2 \mapsto 4$, $3 \mapsto 3$, $4 \mapsto 1$. Ta có $\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}$ cũng là một phép thế bậc 4.

Cho $\mu = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$. Khi đó

$$\sigma \circ \mu = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$$

và

$$\mu \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}.$$

Ta thấy $\sigma \circ \mu \neq \mu \circ \sigma$ là các phép thế bậc 4. □

**8.1.4 Thí dụ.** Trong $S_2$ có hai phần tử là $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ và $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

Trong $S_3$ có sáu phần tử là $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$. □

## 8.2 Vòng xích và chuyển trí

**8.2.1 Định nghĩa.** Một phép thế bậc $n$ được gọi là một **vòng xích độ dài** $k$ (cycle of the length $k$) nếu nó có dạng

$$\sigma = \begin{pmatrix} i_1 & i_2 & \dots & i_k & i_{k+1} & \dots & i_n \\ i_2 & i_3 & \dots & i_1 & i_{k+1} & \dots & i_n \end{pmatrix},$$

kí hiệu $\sigma = (i_1 \ i_2 \ \cdots \ i_k)$. Vòng xích có độ dài 1 chính là ánh xạ đồng nhất và thường kí hiệu là $e$ hay Id.

Hai vòng xích $(i_1 \ i_2 \ \cdots \ i_k)$ và $(j_1 \ j_2 \ \cdots \ j_l)$ trên cùng một tập được gọi là **độc lập** nếu $\{i_1, i_2, \dots, i_k\} \cap \{j_1, j_2, \dots, j_k\} = \varnothing$.

**8.2.2 Nhận xét.** Nếu $\sigma$ là một vòng xích có độ dài là $k$, thì $\sigma^{-1}$ cũng là một vòng xích có độ dài $k$; cụ thể $\sigma = (i_1 \ i_2 \ \cdots \ i_k)$ thì $\sigma^{-1} = (i_k \ \cdots \ i_2 \ i_1)$.

Nếu $\sigma = (i_1\ i_2\ \cdots\ i_k)$ là một vòng xích cấp $k$, thì $\sigma^k = e$.

**8.2.3 Thí dụ.** Cho phép thế bậc 6: $\nu = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 3 & 2 & 5 & 6 \end{pmatrix}$. Khi đó, ta viết lại

$$\nu = \begin{pmatrix} 1 & 4 & 2 & 3 & 5 & 6 \\ 4 & 2 & 1 & 3 & 5 & 6 \end{pmatrix} = (1\ 4\ 2) = (4\ 2\ 1) = (2\ 1\ 4).$$

Vậy $\nu$ là một vòng xích bậc 6 có độ dài 3. □

**8.2.4 Định nghĩa.** Một vòng xích có độ dài 2 thì được gọi là **chuyển trí** (transposition). Nghĩa là phép thế $\sigma$ trên tập $\{1, 2, \ldots, n\}$ là một chuyển trí nếu tồn tại $i \neq j \in \{1, 2, \ldots, n\}$ sao cho $\sigma(i) = j$, $\sigma(j) = i$ và $\sigma(k) = k$ với mọi $k \notin \{i, j\}$.

**8.2.5 Thí dụ.** Ta thấy

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 6 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 6 & 1 & 2 & 3 & 4 \\ 6 & 5 & 1 & 2 & 3 & 4 \end{pmatrix} = (5\ 6) = (6\ 5)$$

là một vòng xích bậc 6 có độ dài 2 tức là một chuyển trí.

Ta cũng thấy vòng xích $\nu$ trong Thí dụ 8.2.3 và vòng xích $\tau$ ở trên là độc lập vì $\{1, 4, 2\} \cap \{5, 6\} = \varnothing$. □

**8.2.6 Mệnh đề.** *Tích của hai vòng xích độc lập có tính chất giao hoán.*

*Chứng minh.* Cho hai vòng xích độc lập $\sigma = (i_1\ \ldots\ i_k)$ và $\mu = (j_1\ \ldots\ j_l)$. Ta có $\{i_1, \ldots, i_k\} \cap \{j_1, \ldots, j_l\} = \varnothing$. Nếu $x \in \{i_1, \ldots, i_k\}$, thì $\sigma(x) \in \{i_1, \ldots, i_k\}$,

$$\sigma \circ \mu(x) = \sigma(\mu(x)) = \sigma(x) \quad \text{và} \quad \mu \circ \sigma(x) = \mu(\sigma(x)) = \sigma(x).$$

Nếu $x \in \{j_1, \ldots, j_l\}$, thì

$$\sigma \circ \mu(x) = \sigma(\mu(x)) = \mu(x) \quad \text{và} \quad \mu \circ \sigma(x) = \mu(\sigma(x)) = \mu(x).$$

Nếu $x \notin \{i_1, \ldots, i_k\} \cup \{j_1, \ldots, j_l\}$, thì

$$\sigma \circ \mu(x) = \sigma(\mu(x)) = \sigma(x) = x \quad \text{và} \quad \mu \circ \sigma(x) = \mu(\sigma(x)) = \mu(x) = x.$$

Vậy $\sigma \circ \mu = \mu \circ \sigma$. □

**8.2.7 Định lý.** *Mỗi phép thế bậc $n$ khác phép thế đồng nhất $e$ đều phân tích được duy nhất thành tích các vòng xích độc lập.*

*Chứng minh.* Cho phép thế bậc $n$ thỏa $\sigma \neq e$. Tồn tại $i_1 \in \{1, 2, \ldots, n\}$ sao cho $\sigma(i_1) = i_2 \neq i_1$. Do $\sigma$ là song ánh nên $i_3 = \sigma(i_2) \neq \sigma(i_1) = i_2$. Nếu $i_3 = i_1$ thì ta dừng lại, ngược lại ta đặt $i_4 = \sigma(i_3)$. Ta cũng có $i_4 \notin \{i_2, i_3\}$. Cứ thế chúng ta tiếp tục và vì ta chỉ có $n$ phần tử nên cuối cùng ta phải có $\sigma(i_{k_1}) = i_1$. Đặt $\sigma_1 = (i_1\ i_2\ \ldots i_{k_1})$, vòng xích này có độ dài lớn hơn hoặc bằng 2.

Nếu trong tập $\{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_{k_1}\}$ có $j_1$ sao cho $\sigma(j_1) = j_2 \neq j_1$ thì tương tự như trên ta thu được vòng xích $\sigma_2 = (j_1\ j_2\ \ldots\ j_{k_2})$ với độ dài lớn hơn hoặc bằng 2, trong đó $\{j_1, j_2, \ldots, j_{k_2}\} \subseteq \{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_{k_1}\}$. Ta thấy hai vòng xích $\sigma_1$ và $\sigma_2$ là độc lập.

Cứ tiếp tục quá trình như vậy cuối cùng ta có vòng xích $\sigma_r$ có độ dài $k_r \geq 2$, với $k_1 + k_2 + \cdots + k_r \leq n$. Khi đó ta có $\sigma = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_r$, trong đó các vòng xích $\sigma_1, \sigma_2, \ldots, \sigma_r$ là độc lập.

Giả sử ta có một phân tích khác $\sigma = \mu_1 \circ \mu_2 \circ \cdots \circ \mu_s$ trong đó các vòng xích $\mu_j$ độc lập. Lấy phần tử $x$ tùy ý. Nếu $x$ không thuộc các vòng xích $\sigma_i$ nào thì nó cũng không thuộc các vòng xích $\mu_j$ nào và ngược lại vì $\sigma(x) = x$. Nếu $x$ thuộc vòng xích $\sigma_i$ thì tồn tại $j$ sao cho $x$ thuộc vòng xích $\mu_j$. Theo mệnh đề trên tích các vòng xích độc lập có tính giao hoán, từ đó ta suy ra được $\sigma(x) = \sigma_i(x) = \mu_j(x)$ và thuộc các vòng xích $\sigma_i$ và $\mu_j$. Từ đó ta suy ra được $\sigma_i^k(x) = \mu_j^k(x)$. Do đó $\sigma_i = \mu_j$. Vì vậy phân tích ở trên của $\sigma$ là duy nhất. $\square$

**8.2.8 Thí dụ.** Cho phép thế bậc 8

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 4 & 5 & 1 & 7 & 6 & 8 \end{pmatrix} = (1\ 2\ 3\ 4\ 5)(6\ 7)$$

Cho phép thế bậc 11

$$\begin{aligned} g &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 3 & 4 & 5 & 7 & 8 & 10 & 2 & 11 & 9 & 6 & 1 \end{pmatrix} \\ &= (1\ 3\ 5\ 8\ 11)(2\ 4\ 7)(6\ 10). \end{aligned}$$

**8.2.9 Hệ quả.** *Mọi phép thế bậc $n$ đều có thể phân tích thành tích của những phép chuyển trí.*

*Chứng minh.* Cho $\sigma$ là một phép thế bậc $n$. Nếu $\sigma = e$ thì ta có thể viết $\sigma = (i\ j)(j\ i)$. Còn nếu $\sigma \neq e$ thì theo Định lý 8.2.7 ta có thể phân tích $\sigma = \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_k$ trong đó các $\sigma_i$ là các vòng xích độc lập với độ dài lớn hơn hoặc bằng 2. Mà các vòng xích thì luôn phân tích được thành tích của các phép chuyển trí, chẳng hạn $(1\ 2\ 3\ \ldots\ k) = (1\ k)(1\ k-1)\cdots(1\ 3)(1\ 2)$. Do đó $\sigma$ cũng phân tích được thành tích của các chuyển trí. □

**8.2.10 *Chú ý.*** Sự phân tích của phép thế thành tích của các chuyển trí không là duy nhất (nhưng số lượng chuyển trí cùng chẵn hoặc cùng lẻ).

**8.2.11 Thí dụ.** Xét vòng xích $(1\,2\,3) \in S_4$. Ta có:

$$(1\,2\,3) = (1\,3)(1\,2) = (2\,3)(1\,3) = (1\,3)(2\,4)(1\,2)(1\,4).$$

## 8.3 Dấu của phép thế

**8.3.1 Định nghĩa.** Cho $\sigma$ là một phép thế trên $\{1, \ldots, n\}$. Nếu $i < j$ nhưng $\sigma(i) > \sigma(j)$, thì ta nói cặp $(\sigma(i), \sigma(j))$ là một **nghịch thế** (inversion) của $\sigma$.

**8.3.2 Thí dụ.** Xét phép thế $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$. Trong phép thế $\sigma$ có 4 nghịch thế, cụ thể là $(3, 2)$, $(3, 1)$, $(2, 1)$, $(4, 1)$. □

**8.3.3 Định nghĩa.** Cho $\sigma$ là một phép thế, và $k$ là số nghịch thế của $\sigma$. Khi đó, đại lượng $(-1)^k$ được gọi là **dấu của phép thế** $\sigma$ (sign of $\sigma$), kí hiệu $\operatorname{sign}\sigma$. Hơn nữa, nếu $\operatorname{sign}\sigma = 1$, ta gọi $\sigma$ là **phép thế chẵn** (even permutation), và nếu $\operatorname{sign}\sigma = -1$, ta gọi $\sigma$ là **phép thế lẻ** (odd permutation).

**8.3.4 Thí dụ.** Phép thế đồng nhất $e$ là một phép thế chẵn bởi vì trong $e$ không có nghịch thế. Nghĩa là $\operatorname{sign} e = 1$. □

**8.3.5 Mệnh đề.** *Nếu hai số hoán vị nhau trong một phép thế, thì dấu của phép thế sẽ thay đổi.*

*Chứng minh.* Giả sử chúng ta có phép thế

$$\sigma = \begin{pmatrix} 1 & \dots & i & \dots & j & \dots & n \\ \sigma(1) & \dots & \sigma(i) & \dots & \sigma(j) & \dots & \sigma(n) \end{pmatrix}.$$

Để cho gọn ta viết $\sigma = (\sigma(1), \dots, \sigma(i), \dots, \sigma(j), \dots, \sigma(n))$. Giả sử chúng ta muốn hoán vị hai số $\sigma(i)$ và $\sigma(j)$ và có phép thế

$$\mu = (\sigma(1), \dots, \sigma(j), \dots, \sigma(i), \dots, \sigma(n)).$$

Ta sẽ chứng minh $\operatorname{sign} \mu = -\operatorname{sign} \sigma$.

Trước hết chúng ta hoán vị $\sigma(i)$ và $\sigma(i+1)$ trong $\sigma$. Ta có phép thế mới

$$(\sigma(1), \dots, \sigma(i-1), \sigma(i+1)\sigma(i), \dots, \sigma(j), \dots, \sigma(n))$$

và khi chúng ta đếm số nghịch thế thì nó chỉ khác với số nghịch thế trong $\sigma$ ở cặp $(\sigma(i+1), \sigma(i))$. Như vậy số nghịch thế của nó và của $\sigma$ sai khác nhau một, cho nên dấu của phép thế này khác với dấu của $\sigma$. Kế tiếp chúng ta hoán vị hai số $\sigma(i)$ và $\sigma(i+2)$ và tiếp tục như thế cho đến khi chúng ta hoán vị $\sigma(i)$ và $\sigma(j)$. Như vậy chúng ta đã thực hiện $j-i$ lần hoán vị từ $\sigma$ để thu được phép thế

$$(\sigma(1), \dots, \sigma(i-1)\sigma(i+1), \dots, \sigma(j-1)\sigma(j)\sigma(i)\sigma(j+1), \dots, \sigma(n)).$$

Bây giờ chúng ta thực hiện một dãy hoán vị theo chiều ngược lại để chuyển $\sigma(j)$ về giữa $\sigma(i-1)$ và $\sigma(i+1)$ (vị trí thứ $i$). Chúng ta cần thực hiện $j-i-1$ lần hoán vị (vì $\sigma(j)$ đang ở vị trí thứ $j-1$). Cuối cùng chúng ta được phép thế mong muốn $\mu$ như sau

$$(\sigma(1), \dots, \sigma(i-1)\sigma(j)\sigma(i+1), \dots, \sigma(j-1)\sigma(i)\sigma(j+1), \dots, \sigma(n)).$$

Như vậy chúng ta đã thực hiện tổng cộng $2(j-i)-1$ lần hoán vị. Như đã biết ở trên mỗi lần hoán vị thì dấu của phép thế sẽ đổi. Ta đã thực hiện *số lẻ lần*, cho nên dấu của phép thế $\mu$ khác với dấu phép thế $\sigma$. □

**8.3.6 Hệ quả.** *Nếu $\tau$ là một chuyển vị, thì* $\operatorname{sign}\tau = -1$.

*Chứng minh.* Giả sử $\tau = (i, j)$. Như vậy phép thế $\tau$ có được từ phép thế đồng nhất $e$ sau khi hoán vị hai số $i$ và $j$. Theo mệnh đề trên ta có $\operatorname{sign}\tau = -\operatorname{sign} e = -1$. □

**8.3.7 Mệnh đề.** *Nếu phép thế $\sigma$ được viết thành tích của các chuyển vị, thì số các chuyển vị là chẵn hay lẻ tùy thuộc và phép thế $\sigma$ chẵn hay lẻ.*

*Chứng minh.* Phép thế đồng nhất là một phép thế chẵn, và mỗi chuyển vị sẽ thay đổi dấu của phép thế. Do đó, số chuyển vị trong tích là chẵn thì $\operatorname{sign}\sigma = 1$ và lẻ nếu $\operatorname{sign}\sigma = -1$. □

**8.3.8 Mệnh đề.** *Cho hai phép thế $\sigma$ và $\mu$ trên cùng một tập hữu hạn. Ta có*

*(1)* $\operatorname{sign}(\sigma \circ \mu) = \operatorname{sign}(\sigma)\operatorname{sign}(\mu)$

*(2)* $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$.

**8.3.9 Mệnh đề.** *Cho $\sigma$ là vòng xích có độ dài $k$. Khi đó* $\operatorname{sign}(\sigma) = (-1)^{k+1}$.

**8.3.10 Thí dụ.** Xét phép thế bậc 11

$$\begin{aligned}\sigma &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 3 & 4 & 5 & 7 & 8 & 10 & 2 & 11 & 9 & 6 & 1 \end{pmatrix}\\ &= (1\ 3\ 5\ 8\ 11)(2\ 4\ 7)(6\ 10)\end{aligned}$$

Vậy $\operatorname{sign}(\sigma) = \operatorname{sign}((1\ 3\ 5\ 8\ 11)\operatorname{sign}(2\ 4\ 7)\operatorname{sign}(6\ 10) = -1$. □

Tập các phép thế chẵn trong $S_n$ kí hiệu là $A_n$.

**8.3.11 Mệnh đề.** *Trong $S_n$ số các phép thế bậc chẵn bằng số các phép thế bậc lẻ và bằng* $\dfrac{n!}{2}$.

*Chứng minh.* Gọi $A_n$ là tập các phép thế bậc chẵn và $\bar{A}_n$ là tập các phép thế bậc lẻ trong $S_n$. Ta xét hai ánh xạ

$$\begin{aligned} \varphi : A_n &\to \bar{A}_n \\ \sigma &\mapsto \sigma \circ (1\ 2) \end{aligned} \qquad\qquad \begin{aligned} \psi : \bar{A}_n &\to A_n \\ \mu &\mapsto \mu \circ (1\ 2) \end{aligned}$$

Khi đó ta có $\varphi \circ \psi = \mathrm{Id}_{\bar{A}_n}$ và $\psi \circ \varphi = \mathrm{Id}_{A_n}$. Suy ra $\varphi$ là song ánh, do đó $|A_n| = |\bar{A}_n|$. □

**Bài tập**

**1**) Cho $\mu$ là một phép thế và $\sigma = (a_1, a_2, \ldots, a_k)$ là một vòng xích có độ dài $k$. Chứng minh rằng $\mu\sigma\mu^{-1}$ là một vòng xích độ dài $k$, cụ thể là $\mu\sigma\mu^{-1} = (\mu(a_1), \mu(a_2), \ldots, \mu(a_k))$.

**2**) Chứng minh Mệnh đề 8.3.8.

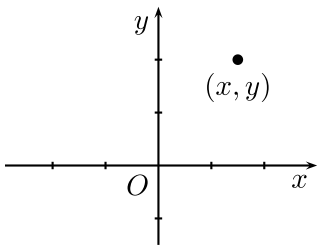
**3**) Chứng minh Mệnh đề 8.3.9.

# § 9 Số phức

## 9.1 Định nghĩa số phức

Ta biết rằng trong số thực không phải mọi số đều có căn bậc chẵn và phương trình bậc lớn hơn một không phải bao giờ cũng có nghiệm. Vì thế xuất hiện sự cần thiết mở rộng khái niệm số thực. Cụ thể là cần đưa đến một loại số mới rộng hơn cùng với các phép toán mà trong trường hợp riêng nó chính là số thực với các phép toán thông thường.

**Số phức** có thể được định nghĩa như là một cặp số thực có thứ tự $(x, y)$. Người ta thường viết số phức bởi các chữ $z$ và $w$. Như vậy với $z = (x, y)$ và $w = (x', y')$ ta có $z = w$ khi và chỉ khi $x = x'$ và $y = y'$. Mặt khác, cặp số $(x, y)$ có thể được hiểu là một điểm trong mặt phẳng tọa độ $Oxy$.



Hình I.1: Mặt phẳng phức $\mathbb{C}$

Khi đó, nếu xem $(x, y)$ là một số phức thì mặt phẳng tọa độ $Oxy$ sẽ được gọi là *mặt phẳng phức $Oxy$* và còn được kí hiệu là $(z)$ hoặc $\mathbb{C}$. Ta có tập hợp số phức

$$\mathbb{C} = \{(x, y) : x, y \in \mathbb{R}\}.$$

Xét số phức $z = (x, y)$. Ta gọi $x$ được gọi là **phần thực** của số phức $z$ và kí hiệu là $\mathrm{Re}z$ còn $y$ được gọi là **phần ảo** của $z$ và ký hiệu là $\mathrm{Im}z$. Trong mặt phẳng phức trục hoành còn được gọi là trục thực và trục tung còn được gọi là trục ảo. Nếu xem trục $Ox$ là một đường thẳng thực, thì mỗi số thực $x$ ứng với điểm $(x, 0)$ trên trục thực $Ox$. Do đó, tập hợp số thực là một tập con của tập số phức, và số phức $z = (x, 0)$ được gọi là số thực và được đồng nhất với $x$, nghĩa là $x \equiv (x, 0)$. Số phức $z = (0, y)$ được gọi là số **thuần ảo**; đặc biệt $(0, 1)$ được gọi là đơn vị ảo và kí hiệu là $i$, nghĩa là $i = (0, 1)$.

Cho số phức $z = (x, y)$ số phức $(x, -y)$ được gọi là **số phức liên hợp** của số phức $z$ và kí hiệu $\bar{z}$. Dễ dàng kiểm tra được

$$\bar{\bar{z}} = z.$$

## 9.2 Các phép toán trên số phức

**Tổng** của hai số phức $z_1 = (x_1, y_1)$ và $z_2 = (x_2, y_2)$ là số phức $z_1 + z_2 = (x_1 + x_2, y_1 + y_2)$ và **tích** của chúng là số phức $z_1 z_2 = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1)$. Người ta chứng minh được phép cộng và phép nhân trên số phức có các tính chất sau.

**9.2.1 Định lý.** *Với mọi $z_1, z_2, z_3 \in \mathbb{C}$, ta có*

*(1)* $z_1 + z_2 = z_2 + z_1$

*(2)* $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$

*(3)* $z_1 + (0, 0) = z_1$

*(4)* $z_1 + z_1' = (0, 0)$*, trong đó* $z_1' = (-x_1, -y_1)$ *nếu* $z_1 = (x_1, y_1)$

*(5)* $z_1 z_2 = z_2 z_1$

*(6)* $z_1(z_2z_3) = (z_1z_2)z_3$

*(7)* $z_1(1,0) = z_1$

*(8) Nếu* $z_1 \neq (0,0)$ *thì* $z_1z_1' = (1,0)$, *trong đó* $z_1' = (\frac{x_1}{x_1^2+y_1^2}, -\frac{y_1}{x_1^2+y_1^2})$ *nếu* $z_1 = (x_1, y_1)$

*(9)* $z_1(z_2+z_3) = z_1z_2 + z_1z_3.$

Với $z = (x,y)$, ta kí hiệu $-z = (-x,-y)$, và nếu $z \neq 0$ kí hiệu $z^{-1} = (\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2})$. Từ đó ta định nghĩa **phép trừ** và **phép chia** như sau.

$$z_1 - z_2 = z_1 + (-z_2) \qquad \frac{z_1}{z_2} = z_1z_2^{-1}, \qquad (z_2 \neq 0).$$

Nhờ có tính chất (6) ta có được định nghĩa sau. **Lũy thừa bậc $n$ của số phức** $z$ là tích $n$ lần của số phức $z$, và kí hiệu $z^n$,

$$z^n = \underbrace{z \cdot z \cdots z}_{n \text{ lần}}.$$

**9.2.2 Thí dụ.** ▷ $i^2 = (0,1)(0,1) = (-1,0) = -1$. Vậy $i$ là nghiệm của $x^2+1=0$ trong $\mathbb{C}$.

▷ Với $z = (x,y)$, ta có $z\bar{z} = (x,y)(x,-y) = (x^2+y^2, 0) = x^2+y^2$.

▷ $(x,y) = (x,0)+(0,y) = x+(0,1)(y,0) = x+iy$. Vậy số phức $(x,y)$ được viết dưới dạng $x+iy$, và gọi là **dạng đại số** của số phức. □

Phép cộng và phép nhân được viết lại ở dạng đại số như sau: với $z_1 = x_1+iy_1$ và $z_2 = x_2+iy_2$ ta có

$$\begin{aligned}(x_1+iy_1)+(x_2+iy_2) &= (x_1+x_2)+i(y_1+y_2)\\ (x_1+iy_1)(x_2+iy_2) &= (x_1x_2-y_1y_2)+i(x_1y_2+x_2y_1).\end{aligned}$$

Như vậy ở dạng đại số phép cộng và nhân được thực hiện như các phép toán đại số trong số thực khi xem $i$ là hằng số và lưu ý đẳng thức $i^2 = -1$.

Ta có thể chứng minh được các phép toán của số phức có các tính chất trong định lý sau.

**9.2.3 Định lý.** *(1)* $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$

*(2)* $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$

*(3)* $z + \bar{z} = 2\,\mathrm{Re}\, z$

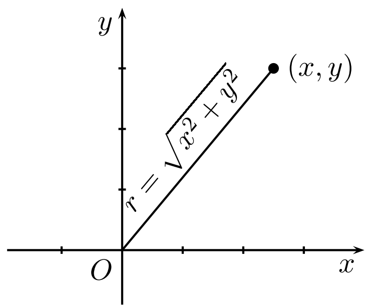
*(4)* $z - \bar{z} = 2i\,\mathrm{Im}\, z$

*(5)* $\overline{\left(\dfrac{z_1}{z_2}\right)} = \dfrac{\bar{z}_1}{\bar{z}_2}$.

## 9.3 Modulus và bất đẳng thức tam giác

Với mỗi $z = (x, y) \in \mathbb{C}$, đặt $|z| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}$, và gọi là **modulus** của $z$. Như vậy $|z|$ chính là khoảng cách của $(x, y)$ đến gốc tọa độ $O(0, 0)$ trong mặt phẳng tọa độ $Oxy$.

Ta biết rằng phương trình của đường tròn có tâm là $(0, 0)$ có bán kính $r$ trong mặt phẳng $Oxy$ là $x^2 + y^2 = r^2$ hay $\sqrt{x^2 + y^2} = r$. Nếu xem $Oxy$ là mặt phẳng phức và $z = (x, y)$ thì phương trình đường tròn tâm $O$ bán kính $r$ được viết lại là $|z| = r$. Lý luận tương tự như trên, ta có *phương trình của đường tròn tâm $z_0 = (x_0, y_0)$ bán kính $r$* là $|z - z_0| = r$.



Hình I.2: Modulus của $z$

Các tính chất về modulus trong định lý sau có thể chứng minh một cách dễ dàng.

**9.3.1 Định lý.** *Với $z, z_1, z_2 \in \mathbb{C}$, ta có*

*(1)* $|z| \geq 0$ *và* $|z| = 0$ *khi và chỉ khi* $z = 0$

*(2)* $|z| \geq |\,\mathrm{Re}\, z|$

*(3)* $|z| \geq |\,\mathrm{Im}\, z|$

*(4)* $|z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|$

*(5)* $|\bar{z}| = |z|$

*(6)* $|z_1 z_2| = |z_1| \cdot |z_2|$

*(7)* $|z_1 + z_2| \leq |z_1| + |z_2|$

*(8)* $|z_1 - z_2| \geq ||z_1| - |z_2||$.

Tính chất (7) ở trên được gọi là **bất đẳng thức tam giác**. Chúng ta sẽ chứng minh bất đẳng thức này. Theo định nghĩa của modulus ta có thể viết

$$\begin{aligned}|z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1 + z_2})\\ &= (z_1 + z_2)(\bar{z}_1 + \bar{z}_2)\\ &= z_1\bar{z}_1 + z_1\bar{z}_2 + z_2\bar{z}_1 + z_2\bar{z}_2.\end{aligned}$$

Mặt khác, ta có

$$z_1\bar{z}_2 + z_2\bar{z}_1 = z_1\bar{z}_2 + \overline{z_1\bar{z}_2} = 2\operatorname{Re}(z_1\bar{z}_2) \leq 2|z_1\bar{z}_2| = 2|z_1||z_2|.$$

Do đó

$$|z_1 + z_2|^2 \leq |z_1|^2 + 2|z_1||z_2| + |z_2|^2 = (|z_1| + |z_2|)^2.$$

**9.3.2 Thí dụ.** Với $z_1$ và $z_2$ là hai số phức tùy ý, chứng minh rằng $|z_1\bar{z}_2 + 1|^2 + |z_1 - z_2|^2 = (1 + |z_1|^2)(1 + |z_2|^2)$.

Thật vậy, ta có

$$\begin{aligned}|z_1\bar{z}_2 + 1|^2 + |z_1 - z_2|^2 &= (z_1\bar{z}_2 + 1)(\overline{z_1\bar{z}_2 + 1}) + (z_1 - z_2)(\overline{z_1 - z_2})\\ &= (z_1\bar{z}_2 + 1)(\bar{z}_1 z_2 + 1) + (z_1 - z_2)(\bar{z}_1 - \bar{z}_2)\\ &= z_1\bar{z}_1\bar{z}_2 z_2 + z_1\bar{z}_2 + \bar{z}_1 z_2 + 1 + z_1\bar{z}_1 - z_1\bar{z}_2 - z_2\bar{z}_1 + z_2\bar{z}_2\\ &= 1 + |z_1|^2 + |z_2|^2 + |z_1|^2|z_2|^2\\ &= (1 + |z_1|^2)(1 + |z_2|^2).\end{aligned}$$

□

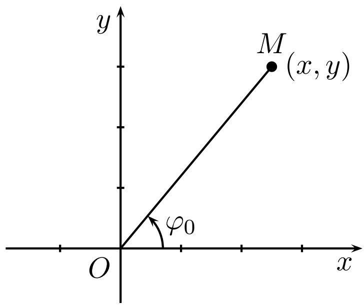
## 9.4 Argument

Cho số phức $z = x + iy \neq 0$. Gọi $\varphi$ là góc có hướng của tia $OM$, với $O$ là gốc tọa độ và $M = (x, y)$, với tia $Ox$ trong mặt phẳng phức $Oxy$. Góc $\varphi$ được gọi là **argument** của số phức $z$ và kí hiệu Arg$z$. Dễ dàng thấy rằng

$$x = |z| \cos \varphi \qquad\qquad y = |z| \sin \varphi.$$

Khi đó, số phức $z$ được viết lại $z = |z| \cos \varphi + i|z| \sin \varphi = |z|(\cos \varphi + i \sin \varphi)$, và ta nói số phức $z$ được viết dưới **dạng lượng giác**.

Trường hợp riêng, với $\varphi_0 \in [0, 2\pi)$ sao cho $x = |z| \cos \varphi_0$ và $y = |z| \sin \varphi_0$, ta nói $\varphi_0$ là **argument chính** của $z$ và ký hiệu $\arg z$.



Hình I.3: Argument của $z$

Argument của số phức được định nghĩa khi nó khác không. Trường hợp số phức bằng 0 thì không xác định argument. Khi số phức khác không, ta có thể xác định argument của một số phức như sau. Xét số phức $z = x + iy$. Nếu $x = 0$ thì

$$\text{Arg}z = \begin{cases} \dfrac{\pi}{2} + 2k\pi & \text{nếu } y > 0, (k \in \mathbb{Z}) \\ -\dfrac{\pi}{2} + 2k\pi & \text{nếu } y < 0, (k \in \mathbb{Z}). \end{cases}$$

Trường hợp $x \neq 0$, ta có

$$\text{Arg}z = \begin{cases} \arctan \dfrac{y}{x} + 2k\pi & \text{nếu } (x, y) \text{ ở góc phần tư thứ I và IV} \\ \arctan \dfrac{y}{x} + (2k+1)\pi & \text{nếu } (x, y) \text{ ở góc phần tư thứ II và III} \end{cases}$$

với $k \in \mathbb{Z}$.

**9.4.1 Thí dụ.** Xét số phức $-4 + 4i$. Ta có $|-4 + 4i| = 4\sqrt{2}$. Vì $-4 + 4i$ ở gốc phần tư thứ hai nên $\text{Arg}z = \arctan(-1) + (2k+1)\pi = \dfrac{3\pi}{4} + 2k\pi$.

Với số phức $3-3\sqrt{3}i$, ta có $|3-3\sqrt{3}i| = 6$ và $\mathrm{Arg}z = \arctan(-\sqrt{3}) + 2k\pi = -\dfrac{\pi}{3} + 2k\pi$. □

Cho hai số phức $z_1$ và $z_2$. Ta có thể chứng minh được khẳng định sau

$$(9.4.2) \qquad z_1 = z_2 \quad \text{khi và chỉ khi} \quad \begin{cases} |z_1| = |z_2| \\ \mathrm{Arg}z_1 = \mathrm{Arg}z_2 + 2k\pi, \quad k \in \mathbb{Z} \end{cases}$$

Dạng lượng giác của số có điểm tiện ích được thể hiện ở định lý sau.

**9.4.3 Định lý.** *Cho hai số phức* $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ *và* $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$. *Khi đó, ta có các hệ thức sau*

$$(9.4.4) \qquad \begin{aligned} z_1 z_2 &= r_1 r_2(\cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2)), \\ \frac{1}{z_1} &= \frac{1}{r_1}(\cos\varphi_1 - i\sin\varphi_1) \end{aligned}$$

Từ định lý trên bằng quy nạp chúng ta có thể chứng minh được công thức sau

$$(9.4.5) \qquad z^n = r^n(\cos n\varphi + i\sin n\varphi).$$

Đặc biệt khi $r = 1$, ta có **công thức Moivre**

$$(9.4.6) \qquad (\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi.$$

Để cho gọn người ta dùng kí hiệu $e^{i\varphi}$ thay cho $\cos\varphi + i\sin\varphi$. Như vậy nếu số phức $z \neq 0$ ta có thể viết lại như sau

$$(9.4.7) \qquad z = |z|e^{i\mathrm{Arg}z}.$$

Số phức $z$ được viết như trên được gọi là **dạng Euler** của $z$. Ta viết lại phép toán nhân và lũy thừa dưới dạng Euler như sau. Với $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ và $z_2 = r_2(\cos\varphi_2 + i\varphi_2)$, ta có

$$(9.4.8) \qquad z_1 z_2 = r_1 r_2 e^{i(\varphi_1+\varphi_2)}, \quad \frac{1}{z_1} = \frac{1}{r_1}e^{i(-\varphi_1)}, \quad (z_1)^n = r_1^n e^{i(n\varphi_1)}.$$

Dạng Euler của số phức rất tiện lợi vì nó phù hợp với các phép toán lũy thừa ở số thực khi xem $i$ như là một hằng số.

## 9.5 Căn bậc $n$ của số phức

**9.5.1 Định nghĩa.** Số phức $w$ được gọi là **căn bậc** $n$ của số phức $z$ nếu $w^n = z$.

Nếu $z = 0$, thì có thể thấy $w = 0$ là căn bậc $n$ duy nhất của $z$. Trường hợp $z \neq 0$, ta cũng thấy $0$ không là căn bậc $n$ của $z$. Ta viết $z$ và $w$ ở Euler như sau

$$z = r_0 e^{i\varphi_0} \qquad\qquad w = re^{i\varphi}.$$

Từ đó suy ra $r^n e^{in\varphi} = r_0 e^{i\varphi_0}$. Vậy ta có

$$\begin{cases} r^n = r_0 \\ n\varphi = \varphi_0 + 2k\pi \end{cases} \qquad \text{suy ra} \qquad \begin{cases} r = \sqrt[n]{r_0} \\ \varphi = \dfrac{\varphi_0 + 2k\pi}{n} \end{cases}$$

hay

$$\text{(9.5.2)} \qquad \begin{cases} |w| = \sqrt[n]{|z|} \\ \mathrm{Arg}w = \dfrac{\mathrm{Arg}z + 2k\pi}{n} \end{cases}$$

trong đó $k \in \mathbb{Z}$. Từ đó có thể thấy được có $n$ căn bậc $n$ của số phức $z$ ứng với các giá trị của $k = 0, 1, \dots, n-1$; cụ thể là $\sqrt[n]{|z|}e^{i\frac{\mathrm{Arg}z+2k\pi}{n}}$ với $k = 0, 1 \dots, n-1$. Người ta dùng kí hiệu $\sqrt[n]{z}$ để chỉ tất cả các căn bậc $n$ của $z$; như vậy cũng có thể hiểu nó là tập hợp tất cả căn bậc $n$ của $z$.

**9.5.3 Thí dụ.** Ta tìm căn bậc 4 của $1 + i\sqrt{3}$. Ta viết lại $1 + i\sqrt{3} = 2e^{i\frac{\pi}{3}}$. Do đó $\sqrt[4]{1 + i\sqrt{3}} = \sqrt[4]{2}e^{i(\frac{\pi}{12} + \frac{k\pi}{2})}$ với $k = 0, 1, 2, 3$. □

# Bài tập

**1)** Tìm các số thực $x, y$ sao cho
**(a)** $(1 + 2i)x + (3 - 5i)y = 1 - 3i$ **(b)** $(2 - 3i)x + (1 + 3i)y = 4 + 5i$

**2)** Thực hiện các phép tính sau đây (viết kết quả ở dạng đại số)

**(a)** $(1+i)(2-i\sqrt{2})$ **(b)** $(3+i\sqrt{3})^2$ **(c)** $\overline{(2+i)^2}$

**3)** Thực hiện các phép tính sau đây (viết kết quả ở dạng đại số)

**(a)** $\dfrac{1-i}{1+2i}$ **(b)** $\dfrac{1}{2-3i}\cdot\dfrac{1}{1+i}$ **(c)** $\dfrac{1+2i}{3-4i}+\dfrac{2-i}{5i}$

**4)** Tìm số thực $a$ và $b$ sao cho $(a+ib)^2=-8+6i$.

**5)** Chứng minh các tính chất trong Định lý 9.2.1.

**6)** Chứng minh các tính chất trong Định lý 9.2.3.

**7)** Chứng minh các tính chất trong Định lý 9.3.1.

**8)** Giải phương trình $z^2+z+1=0$ theo $z=x+iy$ bằng cách viết lại $(x+iy)^2+(x+iy)+1=0$.

**9)** Giải các phương trình sau:
**(a)** $z^2+2iz-5=0$ **(b)** $z^4-3iz^2+4=0$ **(c)** $3z^2+5iz+2=0$

**10)** Với $|z|=2$, chứng minh rằng $\dfrac{2}{15}\le\left|\dfrac{z^2-3z}{4z+7}\right|\le 10$.

**11)** Chứng minh rằng số phức $z$ là số thực hay thuần ảo khi và chỉ khi $(\bar{z})^2=z^2$.

**12)** Hãy tìm tập hợp các điểm trong mặt phẳng phức thỏa (a) $|z+i|\le 3$ và (b) $\mathrm{Re}(\bar{z}-i)=2$.

**13)** Tìm modulus và argument của các số phức sau

**(a)** $1-i$ **(c)** $(-3+i\sqrt{3})^3$ **(e)** $\sqrt{2}-\sqrt{2}i$

**(b)** $\sqrt{3}-i$ **(d)** $(1-i)^{2006}$ **(f)** $\overline{2+2\sqrt{3}i}$

**14)** Giải phương trình $\bar{z}=z^{n-1}$.

**15)** Tìm tất cả các giá trị của các căn sau

**(a)** $\sqrt[4]{-1}$ **(c)** $\sqrt{1+i}$ **(e)** $\sqrt[4]{-8i}$

**(b)** $\sqrt[3]{i}$ **(d)** $\sqrt[4]{-2+i2\sqrt{3}}$ **(f)** $\sqrt[4]{-1-\sqrt{3}i}$

**16)** Dùng công thức Moivre để biểu diễn $\cos nx$ và $\sin nx$ qua các lũy thừa của $\cos x$ và $\sin x$.

**17)** Tính các tổng $1+\cos x+\cdots+\cos nx$ và $\sin x+\cdots+\sin nx$.

# § 10 Đa thức trên trường số

## 10.1 Khái niệm

**10.1.1 Định nghĩa.** Cho $\mathbb{K}$ là một trường số (thường là $\mathbb{Q}$, $\mathbb{R}$ hay $\mathbb{C}$). **Đa thức một biến** $x$ trên $\mathbb{K}$ là một biểu thức dạng

(10.1.2) $$p(x)=a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}+a_nx^n,$$

ở đó $a_0,a_1,\ldots,a_n\in\mathbb{K}$ và được gọi là các **hệ số** của $p(x)$; $x$ gọi là biến, $n$ là một số tự nhiên nào đó. Ký hiệu $\mathbb{K}[x]$ là tập tất cả các đa thức một biến trên $\mathbb{K}$.

Hai đa thức $p(x)$ và $q(x)$ được gọi là bằng nhau, kí hiệu $p(x)=q(x)$, nếu mọi hệ số tương ứng của chúng đều bằng nhau.

**10.1.3 Định nghĩa.** Xét đa thức $p(x)$, (10.1.2), trên $\mathbb{K}$. Nếu $a_0=a_1=\cdots=a_n=0$, thì $p(x)$ gọi là đa thức hằng không, kí hiệu $p(x)=0$.

Nếu $a_0\neq 0$ và $a_1=\cdots=a_n=0$, thì $p(x)$ là đa thức hằng khác không: $p(x)=a_0\neq 0$.

Nếu $a_n\neq 0$, thì $p(x)$ được gọi là **đa thức bậc** $n$, kí hiệu $\deg p(x)=n$. Như vậy, các đa thức hằng khác không $a_0\neq 0$ là các đa thức bậc 0

Quy ước: đa thức không có bậc là $-\infty$.

**10.1.4 Định nghĩa.** Cho hai đa thức $p(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$ và $q(x)=b_0+b_1x+b_2x^2+\cdots+b_mx^m$. **Tổng** của $p(x)$ và $q(x)$, kí hiệu $p(x)+q(x)$, là đa thức $c_0+c_1x+c_2x^2+\cdots+c_lx^l$ ở đó $l=\max\{m,n\}$ và $c_0,c_1,c_2,\ldots,c_l$ được xác định như sau:

- Nếu $m < n = l$, thì $c_k = \begin{cases} a_k + b_k & \text{khi } 0 \leq k \leq m \\ a_k & \text{khi } m < k \leq l. \end{cases}$
- Nếu $m = n = l$, thì $c_k = a_k + b_k$.
- Nếu $n < m = l$, thì $c_k = \begin{cases} a_k + b_k & \text{khi } 0 \leq k \leq n \\ b_k & \text{khi } n < k \leq l. \end{cases}$

**10.1.5 Thí dụ. (OSV07)** Tìm tất cả các đa thức $P(x) \in \mathbb{R}[x]$ thỏa $P(x)+1 = \frac{1}{2}(P(x-1) + P(x+1))$.

*Giải.* Giả sử đa thức $P(x)$ có bậc $n > 2$ thỏa điều kiện đề bài. Khi đó, $P(x)$ có dạng

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \qquad a_n \neq 0.$$

Từ đó hệ số của $x^{n-2}$ phải thỏa $a_{n-2} = \frac{1}{2}(C_n^{n-2} a_n - (n-1)a_{n-1} + a_{n-2} + C_n^{n-2} a_n + (n-1)a_{n-1} + a_{n-2})$, suy ra $a_n = 0$ vô lý. Vậy $P(x)$ chỉ có thể có dạng $P(x) = ax^2 + bx + c$. Thay vào đẳng thức điều kiện ta được

$$\begin{aligned} &ax^2 + bx + c + 1 \\ &= \frac{1}{2}(ax^2 - 2ax + a + bx - b + c + ax^2 + 2ax + a + bx + b + c) \end{aligned}$$

Đồng nhất các hệ số tương ứng của lũy thừa của $x$ ta suy ra được $a = 1$. Vậy $P(x) = x^2 + bx + c$. □

**10.1.6 Định nghĩa.** Cho hai đa thức $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ và $q(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m$. **Tích** của $p(x)$ và $q(x)$, kí hiệu $p(x)q(x)$, là đa thức $c_0 + c_1 x + c_2 x^2 + \cdots + c_{n+m} x^{n+m}$ ở đó

$$c_k = \sum_{i=0}^{k} a_i b_{k-i}, \qquad 0 \leq k \leq n+m.$$

Đặc biệt, với $\lambda \in \mathbb{K}$, ta có $\lambda p(x)$ là đa thức

$$\lambda p(x) = \lambda a_0 + \lambda a_1 x + \lambda a_2 x^2 + \cdots + \lambda a_n x^n.$$

**10.1.7** Ký hiệu $\mathbb{K}[x]$ là tập hợp tất cả các đa thức một biến $x$ trên $\mathbb{K}$. Một điều đáng chú ý là $\mathbb{K}[x]$ cùng với phép cộng và phép nhân định nghĩa như trên lập thành một vành giao hoán có đơn vị. Vành $\mathbb{K}[x]$ được gọi là *vành đa thức trên trường* $\mathbb{K}$.

**10.1.8 Định nghĩa.** Cho $p(x), q(x) \in \mathbb{K}[x]$. Ta nói $q(x)$ **chia hết** $p(x)$ nếu tồn tại $r(x) \in \mathbb{K}[x]$ sao cho $p(x) = q(x)r(x)$. Khi đó, ta nói $q(x)$ là **ước** của $p(x)$ và kí hiệu $q(x) \mid p(x)$.

**10.1.9 Định nghĩa.** Cho $p_1(x), \ldots, p_n(x) \in \mathbb{K}[x]$. Ta nói $q(x)$ là **ước chung lớn nhất** của $p_1(x), \ldots, p_n(x)$ nếu $q(x)$ là ước của $p_i(x)$ với mọi $i = 1, \ldots, n$ và nếu $q_1(x)$ là ước của $p_i(x)$ với mọi $i = 1, \ldots, n$ thì $q_1(x)$ là ước của $q(x)$. Khi đó, ta kí hiệu $q(x) = (p_1(x), \ldots, p_n(x))$.

## 10.2 Định lý cơ bản của đại số học

**10.2.1 Định nghĩa.** Cho $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ là một đa thức trên trường $\mathbb{K}$. Phần tử $\alpha \in \mathbb{K}$ được gọi là một **nghiệm** (hay phần tử không) trên $\mathbb{K}$ của $p(x)$ nếu khi thay $\alpha$ vào chỗ $x$ ta được đẳng thức

$$a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n = 0.$$

**10.2.2 Mệnh đề.** *Nếu $\alpha \in \mathbb{K}$ là một nghiệm của đa thức $p(x) \in \mathbb{K}[x]$ thì luôn tìm được đa thức $q(x)$ sao cho $p(x) = q(x)(x - \alpha)$.*

**10.2.3 Thí dụ. (OVS96)** Cho $P_n(x)$ là đa thức bậc $n$ và cho $m \in \mathbb{N}^*$. Chứng minh rằng nếu $P_n(x^m)$ chia hết cho $x - 1$ thì nó chia hết cho $x^m - 1$.

*Giải.* Do $P_n(x^m)$ chia hết cho $x-1$ nên $1$ là nghiệm của đa thức $P_n(x^m)$, nghĩa là $P_n(1^m) = 0$ hay $P_n(1) = 0$. Điều này cũng có nghĩa $1$ là nghiệm của đa thức $P_n(x)$, cho nên tồn tại một đa thức $q(x)$ sao cho $P_n(x) = q(x)(x - 1)$. Do đó $P_n(x^m) = q(x^m)(x^m - 1)$. Vậy $P_n(x^m)$ chia hết cho $x^m - 1$. □

**10.2.4 Công thức Viete.** Giả sử $x_1, x_2, x_3$ là ba nghiệm của đa thức $x^3 + ax^2 + bx + c = 0$. Bằng cách áp dụng mệnh đề trên nhiều lần ta có thể viết

$$x^3 + ax^2 + bx + c = (x - x_1)(x - x_2)(x - x_3).$$

Bằng cách đồng nhất hệ số của các lũy thừa của $x$ ta được công thức Viete

$$\begin{cases} x_1 + x_2 + x_3 = -a \\ x_1x_2 + x_2x_3 + x_3x_1 = b \\ x_1x_2x_3 = -c \end{cases}$$

Ngược lại người ta cũng chứng minh được nếu $x_1, x_2, x_3$ thỏa hệ thức trên thì chúng là nghiệm của phương trình $x^3 + ax^2 + bx + c = 0$.

**10.2.5 Thí dụ.** Cho $x, y, z$ là các số thực thỏa mãn các đẳng thức sau

$$\begin{cases} x + y + z = 0 \\ x^2 + y^2 + z^2 = 2 \\ x^3 + y^3 + z^3 = 0 \end{cases}$$

Chứng tỏ rằng với mọi số tự nhiên $n$, ta có $x^{2n+1} + y^{2n+1} + z^{2n+1} = 0$.

*Giải.* Từ giả thiết ta có

$$\begin{aligned} 0 = (x + y + z)^2 &= x^2 + y^2 + z^2 + 2(xy + yz + zx) \\ &= 2 + 2(xy + yz + zx) \end{aligned}$$

hay

$$xy + yz + zx = -1.$$

Tương tự ta cũng tính được

$$0 = (x + y + z)^3 = x^3 + y^3 + z^3 + 3(x + y + z)(xy + yz + zx) - 3xyz$$

suy ra

$$xyz = 0.$$

Vậy $x, y, z$ là nghiệm của phương trình $t^3 - t = 0$. Do đó $x, y, z$ là bộ ba nghiệm $1, -1, 0$. Từ đó suy ra $x^{2n+1} + y^{2n+1} + z^{2n+1} = 0$ với mọi $n$ tự nhiên. $\square$

**10.2.6 Mệnh đề.** *Nếu $\alpha \in \mathbb{K}$ là một nghiệm của đa thức $p(x) \in \mathbb{K}[x]$ thì tồn tại một đa thức $r(x) \in K[x]$ và số tự nhiên $k > 0$ sao cho $p(x) = r(x)(x-\alpha)^k$ và $\alpha$ không là nghiệm của $r(x)$. Khi đó, k được gọi là số **bội** của nghiệm $\alpha$ và $\alpha$ được gọi là **nghiệm bội** k của $p(x)$.*

**10.2.7 Mệnh đề.** *Cho $p(x), q(x) \in \mathbb{K}[x]$. Khi đó, $q(x)$ là ước của $p(x)$ khi và chỉ khi nếu $\lambda_0$ là nghiệm bội $k$ của $q(x)$ thì $\lambda_0$ là nghiệm bội $h$ của $p(x)$ với $h \geq k$.*

**10.2.8 Thí dụ. (OSV04)** Xác định đa thức $f(x)$ dạng

$$f(x) = x^5 - 3x^4 + 2x^3 + ax^2 + bx + c$$

biết rằng nó chia hết cho đa thức $(x-1)(x+1)(x-2)$.

*Giải.* Từ giả thiết ta có ngay $1$, $-1$, $2$ là các nghiệm của đa thức $f(x)$. Do đó ta phải có $f(1) = f(-1) = f(2) = 0$, nghĩa là

$$\begin{cases} a+b+c=0 \\ a-b+c=6 \\ 4a+2b+c=0 \end{cases}$$

Dễ dàng suy ra được $a = 1$, $b = -3$, $c = 2$. Vậy ta được $f(x) = x^5 - 3x^4 + 2x^3 + x^2 - 3x + 2$. □

**10.2.9 Định lý. (Định lý cơ bản của đại số học)** *Mọi đa thức có bậc lớn hơn hoặc bằng 1 trên trường số phức, $\mathbb{C}$ đều có ít nhất một nghiệm phức.*

**10.2.10 Hệ quả.** *Mọi đa thức bậc $n > 0$ trên $\mathbb{C}$ đều có đúng $n$ nghiệm phức kể cả bội; mọi đa thức bậc $n > 0$ trên $\mathbb{R}$ có không quá $n$ nghiệm thực kể cả bội.*

**10.2.11 Định lý.** *Giả sử $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n$ là một đa thức bậc $n$ với các hệ số nguyên. Nếu $\alpha = \frac{k}{l}$, với $k, l$ là các số nguyên nguyên tố cùng nhau, là nghiệm hữu tỉ của $p(x)$ thì $k$ chia hết $a_0$ và $l$ chia hết $a_n$.*

*Chứng minh.* Ta có $a_0 + a_1\dfrac{k}{l} + \cdots + a_{n-1}\dfrac{k^{n-1}}{l^{n-1}} + a_n\dfrac{k^n}{l^n} = 0$, suy ra

$$a_0l^n + a_1kl^{n-1} + \cdots + a_{n-1}k^{n-1}l + a_nk^n = 0.$$

Từ đẳng thức trên ta suy ra được $k$ chia hết $a_0l^n$ và $l$ chia hết $a_nk^n$. Do $k$ và $l$ nguyên tố cùng nhau nên suy ra được $k$ chia hết $a_0$ và $l$ chia hết $a_n$. □

## Bài tập

**1)** Chứng minh rằng nếu đa thức thực $p(x)$ có một nghiệm phức $\alpha$ thì nó cũng có một nghiệm phức liên hợp $\bar{\alpha}$. Từ đó hãy suy ra rằng mọi đa thức thực bậc lẻ đều có ít nhất một nghiệm thực.

**2) (OVS96)** Cho $P_n(x)$ là đa thức bậc $n$ và cho $m \in \mathbb{N}^*$. Chứng minh rằng nếu $P_n(x^m)$ chia hết cho $(x-a)^k$ thì nó chia hết cho $(x^m - a^m)^k$ với $a \neq 0$.

# Chương II

# Ma trận và Định thức

*Đại số tuyến tính* là một ngành của Toán học được khởi xướng từ lý thuyết các hệ phương trình tuyến tính và chính lý thuyết các hệ phương trình tuyến tính lại đóng vai trò quan trọng trong đại số tuyến tính. Rất nhiều bài toán của đại số tuyến tính được quy về phép giải các hệ phương trình tuyến tính. Định thức và ma trận, theo thứ tự là hai khái niệm nẩy sinh trong quá trình nghiên cứu lý thuyết các hệ phương trình tuyến tính. Ngày nay, ma trận đã trở thành một công cụ, đồng thời là một ngôn ngữ không thể thiếu của toán học cũng như nhiều ngành khoa học kỹ thuật.

## § 1 Ma trận

Từ nay về sau, $\mathbb{K}$ sẽ luôn là kí hiệu chỉ trường số thực $\mathbb{R}$ hoặc trường số phức $\mathbb{C}$.

### 1.1 Các khái niệm về ma trận

**1.1.1 Định nghĩa.** Cho $m$ và $n$ là hai số nguyên dương. Một **ma trận** (matrix) $A$ cấp $m \times n$ trên $\mathbb{K}$ là một bảng chữ nhật gồm $mn$ số thuộc $\mathbb{K}$ được sắp xếp thành $m$ dòng; $n$ cột dạng:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad \text{với } a_{ij} \in \mathbb{K}.$$

Các số $a_{ij} \in \mathbb{K}$ ở trên được gọi là các **phần tử** (entry) của ma trận $A$, cụ thể hơn $a_{ij}$ gọi là phần tử ở *hàng* hay dòng (row) thứ $i$ và *cột* (column) thứ $j$, hay ở vị trí $(i, j)$ của $A$ với $i = 1, \ldots, m$ và $j = 1, \ldots, n$; nó còn được kí hiệu $(A)_{ij}$. Ma trận $A$ ở trên còn được kí hiệu gọn như sau $(a_{ij})_{m\times n}$ hay $(a_{ij})$ khi cấp của $A$ không bị nhầm lẫn.

Các ma trận thường được kí hiệu bằng các chữ in hoa: $A$, $B$, $C$, .... Tập các ma trận cấp $m \times n$ với các phần tử thuộc $\mathbb{K}$ được kí hiệu là: $\mathcal{M}_{m,n}(\mathbb{K})$.

Xét ma trận $A = (a_{ij})_{m\times n}$. Nếu $m = 1$ thì ta nói $A$ là *ma trận dòng* (row matrix) và ta sẽ xem nó như là vector dòng

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \end{bmatrix}.$$

Khi $n = 1$, $A$ được gọi là *ma trận cột* (column matrix) và ta sẽ xem nó như là vector cột

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}.$$

Khi $m = n = 1$, ta có $A = [a_{11}]$. Khi đó, ta đồng nhất mỗi ma trận cấp $1 \times 1$ với phần tử của nó trong $\mathbb{K}$, nghĩa là $[a_{11}] \equiv a_{11}$.

Trở lại tường hợp tổng quát với ma trận $A = (a_{ij})_{m\times n}$, ma trận dòng $\begin{bmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{bmatrix}$ gọi là ma trận dòng thứ $i$ của ma trận $A$, kí hiệu bởi $A_i$; và ma trận cột $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$ gọi là ma trận cột thứ $j$ của ma trận $A$, kí hiệu bởi $A^j$.

**1.1.2** ***Chú ý.*** Để tiện lợi về sau ta đồng nhất mỗi phần tử $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{K}^n$ với một ma trận cột, tức là $\mathbf{a} \equiv \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$.

**1.1.3 Định nghĩa.** Hai ma trận $A$ và $B$ được gọi là **bằng nhau** nếu chúng có cùng cấp (có cùng số dòng, số cột) $m \times n$ đồng thời các phần tử ở vị trí tương ứng phải bằng nhau ($(A)_{ij} = (B)_{ij}$ với mọi $i = 1, \dots, m$ và $j = 1, \dots, n$).

**1.1.4 Định nghĩa. (Phép chuyển vị)** Cho ma trận $A = (a_{ij})_{m\times n}$ là ma trận cấp $m \times n$. Ta gọi **ma trận chuyển vị** (transpose) của ma trận $A$ là một ma trận cấp $n \times m$, kí hiệu là $A^t$, nhận được từ $A$ bằng cách đổi dòng thành cột,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}; \quad A^t = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix},$$

nghĩa là $(A^t)_{ij} = (A)_{ji}$ với mọi $i, j = 1, \dots, n$.

## 1.2 Một số ma trận đặc biệt

**1.2.1 Định nghĩa. Ma trận không cấp** $m \times n$ (zero matrix), kí hiệu là $O_{m,n}$ hoặc đơn giản là $O$, là một ma trận cấp $m \times n$ mà tất cả các phần tử của nó đều bằng 0.

**1.2.2 Định nghĩa.** Các ma trận cấp $n \times n$ trên $\mathbb{K}$ được gọi là **ma trận vuông** (square matrix) cấp $n$. Các phần tử $a_{11}, a_{22}, a_{33}, \dots, a_{nn}$ gọi là các các phần tử thuộc *đường chéo chính* (main diagonal) của ma trận $A$. Các phần tử $a_{1n}$, $a_{2(n-1)}$, $\dots$, $a_{n1}$ được gọi là các phần tử trên đường chéo phụ của ma trận $A$. Tập hợp các ma trận vuông cấp $n$ trên $\mathbb{K}$ được kí hiệu là $\mathcal{M}_n(\mathbb{K})$.

**1.2.3 Định nghĩa. Ma trận chéo** (diagonal matrix) cấp $n$ là ma trận trận vuông cấp $n$ mà mọi phần tử không nằm trên đường chéo chính đều bằng 0. Như vậy nếu $D$ là một ma trận chéo cấp $n$ thì $D$ có dạng

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{bmatrix}.$$

**1.2.4 Định nghĩa.** **Ma trận đơn vị** (identity matrix) cấp $n$, kí hiệu là $I_n$ hoặc đơn giản là $I$ khi cấp $n$ đã được chỉ rõ, là ma trận chéo cấp $n$ mà các phần tử trên đường chéo chính bằng 1.

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} = (\delta_{ij})_{n\times n}$$

trong đó $\delta_{ij} = \begin{cases} 1 & \text{nếu } i = j \\ 0 & \text{nếu } i \neq j \end{cases}$ là kí hiệu Kronecker.

**1.2.5 Định nghĩa.** Ma trận vuông $A = (a_{ij})_{n\times n}$ được gọi là **ma trận tam giác trên** (upper triangle matrix) cấp $n$ nếu các phần tử nằm phía dưới đường chéo chính đều bằng 0. Nghĩa là $a_{ij} = 0$ với mọi $1 \leq j < i \leq n$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

Ma trận vuông $A = (a_{ij})_{n\times n}$ được gọi là **ma trận tam giác dưới** (lower triangle matrix) cấp $n$ nếu các phần tử nằm phía trên đường chéo chính đều bằng 0. Nghĩa là $a_{ij} = 0$ với mọi $1 \leq i < j \leq n$,

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

**1.2.6 Định nghĩa.** **Ma trận đối xứng** (symmetric matrix) cấp $n$ là ma trận $A = (a_{ij})_{n\times n}$ sao cho các phần tử ở vị trí đối xứng nhau qua đường chéo chính thì bằng nhau nghĩa là $a_{ij} = a_{ji}$ với mọi $i, j = 1, \dots, n$.

**1.2.7 Định nghĩa.** **Ma trận phản đối xứng** (antisymmetric matrix) cấp $n$ là ma trận vuông $A = (a_{ij})$ cấp $n$ sao cho các phần tử ở các vị trí đối xứng nhau qua đường chéo chính thì đối nhau, nghĩa là $a_{ij} = -a_{ji}$ với mọi $i, j = 1, \dots, n$.

**1.2.8 Nhận xét.** Từ định nghĩa ta thấy ma trận $A$ là ma trận phản đối xứng thì phải có các phần tử trên đường chéo chính bằng 0.

**1.2.9 Thí dụ.** Ma trận $\begin{bmatrix} 1 & 2 & 3 & -4 \\ 2 & 5 & 6 & 8 \\ 3 & 6 & 7 & -9 \\ -4 & 8 & -9 & 10 \end{bmatrix}$ là ma trận đối xứng cấp 4.

Ma trận $\begin{bmatrix} 0 & -2 & 3 \\ 2 & 0 & 4 \\ -3 & -4 & 0 \end{bmatrix}$ là ma trận phản đối xứng cấp 3. □

## 1.3 Ma trận sơ cấp

**1.3.1 Ma trận sơ cấp loại I** là ma trận thu được khi ta hoán vị hai dòng $i$ và $j$ của ma trận đơn vị $I$, kí hiệu $E_{I;i,j}$. Vậy ma trận sơ cấp loại I có dạng

$$E_{I;i,j} = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}.$$

**1.3.2 Thí dụ.** Ma trận $E_{I;2,4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ là ma trận sơ cấp loại I cấp 4. □

**1.3.3 Ma trận sơ cấp loại II** là ma trận đường chéo với $a_{kk} = 1$ với mọi $k \neq i$ và $a_{ii} = \lambda$ với $\lambda \neq 0$, kí hiệu $E_{II;i,\lambda}$. Vậy ma trận sơ cấp loại II có dạng

$$E_{II;i,\lambda} = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

**1.3.4 Thí dụ.** Ma trận $E_{II;3,\lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ là ma trận sơ cấp loại II cấp 5. □

**1.3.5 Ma trận sơ cấp loại III** là một ma trận vuông có các phần tử thuộc đường chéo chính bằng một và các phần tử ngoài đường chéo chính đều bằng 0 trừ phần tử $a_{ij} = \lambda$, ký hiệu $E_{III;i,j,\lambda}$. Vậy ma trận sơ cấp loại III có dạng

$$E_{III;i,j,\lambda} = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \dots & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

**1.3.6 Thí dụ.** Ma trận $E_{III;2,3,\lambda} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ là ma trận sơ cấp loại III cấp 4. □

## 1.4 Ma trận chia khối

Cho $A$ là một ma trận cấp $m \times n$ trên $\mathbb{K}$ với $m, n \geq 2$. Dùng một số đường ngang và dọc, ta có thể chia $A$ thành các ma trận cấp nhỏ hơn mà được gọi là các khối của $A$. Khi đó, ma trận $A$ được gọi là một **ma trận chia khối** (hay **ma trận khối**) (partitioned matrix). Một ma trận đã cho có thể được biến thành ma trận chia khối theo nhiều cách khác nhau; chẳng hạn

$$\begin{bmatrix}1&2&0&2&4\\3&1&1&4&3\\2&3&4&6&1\end{bmatrix} = \left[\begin{array}{cc|cc|c}1&2&0&2&4\\3&1&1&4&3\\ \hline 2&3&4&6&1\end{array}\right] = \left[\begin{array}{ccc|cc}1&2&0&2&4\\ \hline 3&1&1&4&3\\ \hline 2&3&4&6&1\end{array}\right]$$

$$A = \left[\begin{array}{c}A_1\\ \hline A_2\\ \hline \vdots\\ \hline A_m\end{array}\right] = \left[A^1 \mid A^2 \mid \cdots \mid A^n\right]$$

Khi $A$ là ma trận chia khối, ta có thể xét $A$ như là một ma trận mới với các "phần tử" là các khối. Trong ví dụ trên ta có

$$B = \begin{bmatrix}1&2&0&2&4\\3&1&1&4&3\\2&3&4&6&1\end{bmatrix} = \left[\begin{array}{cc|cc|c}1&2&0&2&4\\3&1&1&4&3\\ \hline 2&3&4&6&1\end{array}\right] = \begin{bmatrix}B_{11}&B_{12}&B_{13}\\B_{21}&B_{22}&B_{23}\end{bmatrix}$$

với $B_{11} = \begin{bmatrix}1&2\\3&1\end{bmatrix}$, $B_{12} = \begin{bmatrix}0&2\\1&4\end{bmatrix}$, $B_{13} = \begin{bmatrix}4\\3\end{bmatrix}$, $B_{21} = \begin{bmatrix}2&3\end{bmatrix}$, $B_{22} = \begin{bmatrix}4&6\end{bmatrix}$, và $B_{23} = \begin{bmatrix}1\end{bmatrix}$.

## Bài tập

**1**) Tìm chuyển vị của các ma trận sau

**(a)** $\begin{bmatrix}1&0&2&3\\-1&2&0&1\end{bmatrix}$

**(b)** $\begin{bmatrix}1&4&7\\2&5&8\\3&6&9\end{bmatrix}$

**(c)** $\begin{bmatrix}1&4&7&10&13\\2&5&8&11&14\\3&6&9&12&15\end{bmatrix}$

**(d)** $\begin{bmatrix}1&-5&9\\-2&6&-10\\3&-7&11\\-4&8&-12\end{bmatrix}$

**2**) Cho $A$ là một ma trận vuông. Chứng minh rằng $\operatorname{tr}(A) = \operatorname{tr}(A^t)$.

# § 2 Các phép toán trên ma trận

## 2.1 Các phép toán

**2.1.1 Định nghĩa. (Phép cộng ma trận)** (matrix addition) Cho hai ma trận cùng cấp $A = (a_{ij})_{m\times n}$ và $B = (b_{ij})_{m\times n}$. **Tổng** (sum) của $A$ và $B$, kí hiệu $A + B$; là một ma trận cấp $m \times n$ được xác định như sau

$$A + B = (a_{ij} + b_{ij})_{m\times n}$$

tức là $(A + B)_{ij} = a_{ij} + b_{ij}$ với mọi $i = 1, \dots, m$ và $j = 1, \dots, n$.

**2.1.2 Thí dụ.** $\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} + \begin{bmatrix}7 & 8 & 9\\10 & 11 & 12\end{bmatrix} = \begin{bmatrix}8 & 10 & 12\\14 & 16 & 18\end{bmatrix}$. □

**2.1.3 Định nghĩa. (Phép nhân ma trận với một số)** (scalar multiplication) Cho $A = (a_{ij})_{m\times n}$ là một ma trận cấp $m \times n$ và $\lambda \in \mathbb{K}$. Ta gọi **tích của** $\lambda$ **với** $A$, kí hiệu $\lambda A$, là một ma trận cấp $m \times n$ xác định như sau: $\lambda A = (\lambda a_{ij})_{m\times n}$ tức là $(\lambda A)_{ij} = \lambda a_{ij}$.

Đặc biệt, khi $\lambda = -1$, thay cho $(-1)A$ ta sẽ viết $-A$ và gọi nó là **ma trận đối của** $A$ (additive inverse).

**2.1.4 Thí dụ.** $\pi\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} = \begin{bmatrix}\pi & 2\pi & 3\pi\\4\pi & 5\pi & 6\pi\end{bmatrix}$. □

**2.1.5 Định nghĩa. (Phép nhân hai ma trận)** (matrix multiplication) Cho $A$ là một ma trận cấp $m \times n$ trên $\mathbb{K}$ và $B$ là một ma trận cấp $n \times k$ trên $\mathbb{K}$. Ta gọi **tích của** $A$ **với** $B$ (product), kí hiệu $AB$, là một ma trận $C$ cấp $m \times k$ trên $\mathbb{K}$ mà các phần tử của nó được xác định như sau:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{l=1}^{n} a_{il}b_{lj}$$

với $i = 1, \dots, m$ và $j = 1, \dots, k$.

**2.1.6 Thí dụ.** Cho hai ma trận

$$A = \begin{bmatrix}1 & 1 & 2\\1 & 2 & 3\\1 & 4 & 9\end{bmatrix} \qquad B = \begin{bmatrix}0 & 1\\1 & -1\\2 & 0\end{bmatrix}$$

và $C = AB$. Khi đó

$$C = \begin{bmatrix} 1\cdot 0 + 1\cdot 1 + 2\cdot 2 & 1\cdot 1 + 1\cdot -1 + 2\cdot 0 \\ 1\cdot 0 + 2\cdot 1 + 3\cdot 2 & 1\cdot 1 + 2\cdot -1 + 3\cdot 0 \\ 1\cdot 0 + 4\cdot 1 + 9\cdot 2 & 1\cdot 1 + 4\cdot -1 + 9\cdot 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 8 & -1 \\ 22 & -3 \end{bmatrix}. \quad \square$$

**2.1.7 Thí dụ.** Cho các ma trận

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ 3 & -1 \end{bmatrix} \qquad C = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$$

Khi đó ta tính được

$$AB = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 11 & 8 \end{bmatrix}$$

$$BA = \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 10 & 5 & 5 \\ 5 & 5 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

$$BC = \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ 5 & -2 \\ 5 & -3 \end{bmatrix}$$

$$CA = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 3 & -4 \\ 1 & 2 & -1 \end{bmatrix}$$

Ta nhận thấy $AB \neq BA$. Hơn nữa, ta thấy $AC$ và $CB$ không xác định. Vậy phép nhân các ma trận không có tính giao hoán. $\square$

**2.1.8 Thí dụ. (OSV00)** Cho $A$ và $B$ là các ma trận phản đối xứng cấp $n$. Chứng minh rằng $AB$ phản đối xứng khi và chỉ khi $AB = -BA$.

*Giải.* $AB$ phản đối xứng khi và chỉ khi $(AB)_{ij} + (AB)_{ji} = 0$ với mọi $i, j = 1, \ldots, n$. Trong khi đó, ta tính được

$$(AB)_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki} = \sum_{k=1}^{n} (-a_{kj})(-b_{ik}) = \sum_{k=1}^{n} b_{ik} a_{kj} = (BA)_{ij}.$$

Vậy $(AB)_{ij} + (AB)_{ji} = 0$ khi và chỉ khi $(AB)_{ij} = -(BA)_{ij}$. Do đó, $AB$ phản đối xứng khi và chỉ khi $AB = -BA$.

**2.1.9 Định nghĩa.** Với mỗi ma trận vuông $A$ cấp $n$ và mỗi số tự nhiên $k$, ta định nghĩa *lũy thừa bậc k của A* bằng quy nạp như sau

$$A^0 = I_n \qquad\qquad A^k = A^{k-1}A \quad \text{với } k \geq 1$$

**2.1.10 Thí dụ. (OSV93)** Tìm tất cả các ma trận thực $X = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$ sao cho $X^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

*Giải.* Ta tính ma trận

$$X^2 = \begin{bmatrix} x^2 + yz & xy + yt \\ xz + zt & zy + t^2 \end{bmatrix}.$$

Do đó các phần tử của ma trận $X$ là nghiệm của hệ phương trình

$$\begin{cases} x^2 + yz = 1 \\ xy + yt = 0 \\ xz + zt = 0 \\ yz + t^2 = 1 \end{cases} \qquad \text{hay} \qquad \begin{matrix} \begin{cases} x + t = 0 \\ x^2 + zy = 1 \end{cases} \\ \begin{cases} y = z = 0 \\ x = t = 1 \end{cases} \\ \begin{cases} y = z = 0 \\ x = t = -1 \end{cases} \end{matrix}$$

□

**2.1.11 Thí dụ. (OSV11)** Cho ma trận $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Hãy tính $A^{2012}$.

*Giải.* Tính các lũy thừa đầu tiên của $A$ ta được

$$A^2 = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix} \quad A^3 = \begin{bmatrix} -2 & -2 \\ 2 & -2 \end{bmatrix} \quad A^4 = \begin{bmatrix} -4 & 0 \\ 0 & -4 \end{bmatrix} = -4I.$$

Do đó, $A^{2012} = (-4I)^{503} = (-4)^{503}I$. □

**2.1.12 Thí dụ. (OSV09)** Tồn tại hay không một ma trận thực $A$ vuông cấp 2 sao cho $A^{2010} = \begin{bmatrix} -2008 & 2010 \\ 0 & -2009 \end{bmatrix}$.

*Giải.* Giả sử tồn tại ma trận $A$ thỏa mãn yêu cầu. Đặt $A^{1005} = B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Ta có

$$\begin{bmatrix} -2008 & 2010 \\ 0 & -2009 \end{bmatrix} = B^2 = \begin{bmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{bmatrix}$$

Vì vậy ta phải có $ac + cd = 0$, cho nên $c = 0$ hay $a + d = 0$. Nếu $c =$ 0 thì

$$\begin{bmatrix} a^2 & ab+bd \\ 0 & d^2 \end{bmatrix} = \begin{bmatrix} -2008 & 2010 \\ 0 & -2009 \end{bmatrix} \qquad \text{đây là điều vô lý}$$

Nếu $a + d = 0$ thì

$$\begin{bmatrix} a^2+bc & 0 \\ 0 & bc+d^2 \end{bmatrix} = \begin{bmatrix} -2008 & 2010 \\ 0 & -2009 \end{bmatrix} \qquad \text{đây là điều vô lý.}$$

Do đó không tồn tại ma trận $A$ thỏa yêu cầu. □

**2.1.13 Thí dụ. (OSV94)** Tìm $\lim\limits_{x\to 0}\left(\lim\limits_{n\to\infty}\dfrac{1}{x}(A^n - I)\right)$ với $A = \begin{bmatrix} 1 & \frac{x}{n} \\ -\frac{x}{n} & 1 \end{bmatrix}$ với $n$ nguyên dương.

*Giải.* Đặt $\varphi = \arcsin\frac{x}{\sqrt{x^2+n^2}}$. Khi đó $\sin\varphi = \frac{x}{\sqrt{x^2+n^2}}$ và $\cos\varphi = \frac{n}{\sqrt{n^2+x^2}}$. Ta viết lại ma trận $A$ như sau

$$A = \frac{\sqrt{x^2+n^2}}{n}\begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix},$$

dễ dàng tính được

$$A^n = \left(\frac{\sqrt{x^2+n^2}}{n}\right)^n \begin{bmatrix} \cos n\varphi & \sin n\varphi \\ -\sin n\varphi & \cos n\varphi \end{bmatrix}$$

Do đó

$$\begin{aligned}\lim_{n\to\infty} A^n &= \lim_{n\to\infty}\left(1+\frac{x^2}{n^2}\right)^{\frac{n}{2}}\begin{bmatrix} \cos n\arcsin\frac{x}{\sqrt{x^2+n^2}} & \sin n\arcsin\frac{x}{\sqrt{x^2+n^2}} \\ -\sin n\arcsin\frac{x}{\sqrt{x^2+n^2}} & \cos n\arcsin\frac{x}{\sqrt{x^2+n^2}} \end{bmatrix} \\ &= \begin{bmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{bmatrix}\end{aligned}$$

Vậy giới hạn cần tính là

$$\lim_{x\to 0}\left(\lim_{n\to\infty}\frac{1}{x}(A^n - I)\right) = \lim_{x\to 0}\frac{1}{x}\begin{bmatrix}\cos x - 1 & \sin x\\ -\sin x & \cos x - 1\end{bmatrix} = \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}.$$

□

**2.1.14 Thí dụ. (OSV96)** Cho ma trận $A = \begin{bmatrix}a & b\\ 0 & c\end{bmatrix}$ với $a, b, c \in \mathbb{R}$.

**(i)** Chứng minh rằng nếu $A^{1996} = O$ thì $A^2 = O$.

**(ii)** Xác định $a, b, c$ sao cho tồn tại $n \in \mathbb{N}^*$ để $A^n = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$.

*Giải.* **(i)** Dùng quy nạp ta tính được

$$A^2 = \begin{bmatrix}a^2 & ab + bc\\ 0 & c^2\end{bmatrix} \qquad \text{và} \qquad A^{1996} = \begin{bmatrix}a^{1996} & p\\ 0 & c^{1996}\end{bmatrix}$$

trong đó $p$ là biểu thức đa thức theo $a$, $b$, $c$. Do $A^{1996} = O$ nên ta phải có $a^{1996} = c^{1996} = 0$ hay $a = c = 0$. Khi đó $A$ có dạng $A = \begin{bmatrix}0 & b\\ 0 & 0\end{bmatrix}$. Từ đó ta có ngay $A^2 = O$.

**(ii)** Ta có $A^n = \begin{bmatrix}a^n & p\\ 0 & c^n\end{bmatrix}$ trong đó $p$ là biểu thức đa thức theo $a$, $b$, $c$. Do đó ta phải có $a^n = c^n = 1$. Xét trường hợp $a = c = 1$. Khi đó dễ dàng tính được $A^n = \begin{bmatrix}1 & nb\\ 0 & 1\end{bmatrix}$. Ta phải có $nb = 0$ với $n$ nào đó, cho nên $b = 0$. Khi đó $A_1 = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ thỏa điều kiện đề bài với bất kỳ $n$ nguyên dương. Trường hợp $a = c = -1$. Ta cũng tính được $A^n = \begin{bmatrix}(-1)^n & (-1)^{n-1}nb\\ 0 & (-1)^n\end{bmatrix}$. Ta phải có $(-1)^{n-1}nb = 0$ với $n$ chẵn nào đó, cho nên $b = 0$. Khi đó $A_2 = \begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}$ thỏa điều kiện đề bài với bất kỳ $n$ chẵn. Trường hợp $a = 1$, $c = -1$, ta tính được $A^2 = \begin{bmatrix}1 & b\\ 0 & -1\end{bmatrix}^2 = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ Vậy ma trận $A_3 = \begin{bmatrix}1 & b\\ 0 & -1\end{bmatrix}$ thỏa điều kiện đề bài. Tương tự ma trận $A_4 = \begin{bmatrix}-1 & b\\ 0 & 1\end{bmatrix}$ cũng thỏa điều kiện đề bài. □

**2.1.15 Thí dụ. (OSV98)** Đặt $A(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$

**(a)** Chứng minh rằng $A(\alpha)A(\beta) = A(\alpha+\beta)$

**(b)** Tính $A^n(\frac{\pi}{2})$.

*Giải.* **(a)** Ta có

$$\begin{aligned} A(\alpha)A(\beta) &= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \\ &= \begin{bmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\cos\alpha\sin\beta - \sin\alpha\cos\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{bmatrix} \\ &= \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix} = A(\alpha+\beta). \end{aligned}$$

**(b)** Từ kết quả ở (a) và bằng quy nạp ta chứng minh được

$$A^n(\alpha) = \begin{bmatrix} \cos n\alpha & -\sin n\alpha \\ \sin n\alpha & \cos n\alpha \end{bmatrix} \qquad n \in \mathbb{N}^*$$

Do đó ta tính được

$$A^n(\tfrac{\pi}{2}) = \begin{bmatrix} \cos\frac{n\pi}{2} & -\sin\frac{n\pi}{2} \\ \sin\frac{n\pi}{2} & \cos\frac{n\pi}{2} \end{bmatrix} = \begin{cases} \begin{bmatrix} (-1)^k & 0 \\ 0 & (-1)^k \end{bmatrix} & \text{khi } n = 2k \\ \begin{bmatrix} 0 & (-1)^{k+1} \\ (-1)^k & 0 \end{bmatrix} & \text{khi } n = 2k+1 \end{cases}$$

□

**2.1.16 Thí dụ. (OSV99)** Cho ma trận $A = \begin{bmatrix} \frac{x}{1998} & 1999 \\ 0 & \frac{x}{2000} \end{bmatrix}$. Ký hiệu $A^n = \begin{bmatrix} a_{11}(n,x) & a_{12}(n,x) \\ a_{21}(n,x) & a_{22}(n,x) \end{bmatrix}$. Tìm $\lim\limits_{n\to\infty} \lim\limits_{x\to 1} a_{ij}(n,x)$, $i, j = 1, 2$.

*Giải.* Một cách ngắn gọn và tổng quát ta ký hiệu $A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ với $a \neq c$. Khi đó ta có

$$A^2 = \begin{bmatrix} a^2 & b(a+c) \\ 0 & c^2 \end{bmatrix} = \begin{bmatrix} a^2 & \frac{b(a^2-c^2)}{a-c} \\ 0 & c^2 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} a^3 & a^2b + \frac{bc(a^2-c^2)}{a-c} \\ 0 & c^3 \end{bmatrix} = \begin{bmatrix} a^3 & \frac{b(a^3-c^3)}{a-c} \\ 0 & c^3 \end{bmatrix}$$

Bằng quy nạp ta chứng minh được

$$A^n = \begin{bmatrix} a^n & \frac{b(a^n-c^n)}{a-c} \\ 0 & c^n \end{bmatrix} \qquad n \in \mathbb{N}^*$$

Từ đó ta có $a_{11}(n,x) = \frac{x^n}{1998^n}$, $a_{12}(n,x) = \frac{1999(\frac{x}{1998^n} - \frac{x^n}{2000^n})}{\frac{x}{1998} - \frac{x}{2000}}$, $a_{21}(n,x) = 0$, $a_{22}(n,x) = \frac{x^n}{2000^n}$ và ta tính được

$$\lim_{n\to\infty}\lim_{x\to 1}\frac{x^n}{1998^n} = \lim_{n\to\infty}\frac{1}{1998^n} = 0$$

$$\lim_{n\to\infty}\lim_{x\to 1}\frac{x^n}{2000^n} = \lim_{n\to\infty}\frac{1}{2000^n} = 0$$

$$\lim_{n\to\infty}\lim_{x\to 1}\frac{b\big(\frac{x^n}{1998^n} - \frac{x^n}{2000^n}\big)}{\frac{x}{1998} - \frac{x}{2000}} = \lim_{n\to\infty}\frac{b\big(\frac{1}{1998^n} - \frac{1}{2000^n}\big)}{\frac{1}{1998} - \frac{1}{2000}} = 0$$

Vậy $\lim\limits_{n\to\infty}\lim\limits_{x\to 1} a_{ij}(n,x) = 0$ với mọi $i, j = 1, 2$. □

**2.1.17 *Chú ý.*** Nếu các ma trận được cho là ma trận chia khối và ta xem các khối như là phần tử trong ma trận mới. Khi đó, các phép toán trên các ma trận ban đầu vẫn có thể tính khi ta thực hiện các phép toán ấy theo các khối (tất nhiên điều kiện là các phép toán theo khối phải thực hiện được). Cụ thể như sau: Với các ma trận chia khối sau, ta xét các phép toán trên ma trận

$$A = \begin{bmatrix} A_{11} & \dots & A_{1n} \\ \dots & \dots & \dots \\ A_{m1} & \dots & A_{mn} \end{bmatrix} \quad \text{ta có} \quad \lambda A = \begin{bmatrix} \lambda A_{11} & \dots & \lambda A_{1n} \\ \dots & \dots & \dots \\ \lambda A_{m1} & \dots & \lambda A_{mn} \end{bmatrix}$$

cùng với ma trận $B = \begin{bmatrix} B_{11} & \dots & B_{1n} \\ \multicolumn{3}{c}{\dots\dots\dots\dots} \\ B_{m1} & \dots & B_{mn} \end{bmatrix}$ ta có

$$A + B = \begin{bmatrix} A_{11} + B_{11} & \dots & A_{1n} + B_{1n} \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots\dots\dots} \\ A_{m1} + B_{m1} & \dots & A_{mn} + B_{mn} \end{bmatrix}$$

với điều kiện các ma trận $A_{ij}$ và $B_{ij}$ có cùng cấp. Nếu xét thêm ma trận

$$C = \begin{bmatrix} C_{11} & \dots & C_{1p} \\ \multicolumn{3}{c}{\dots\dots\dots\dots} \\ C_{n1} & \dots & C_{np} \end{bmatrix}$$

ta có

$$AC = \begin{bmatrix} A_{11}C_{11} + \cdots + A_{1n}C_{n1} & \dots & A_{11}C_{1p} + \cdots + A_{1n}C_{np} \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ A_{m1}C_{11} + \cdots + A_{mn}C_{n1} & \dots & A_{m1}C_{1p} + \cdots + A_{mn}C_{np} \end{bmatrix}$$

trong đó số cột của ma trận $A_{ij}$ bằng số dòng của ma trận $C_{jl}$.

**2.1.18 Thí dụ.** Cho ma trận $A$ cấp $m \times n$ và ma trận $B$ cấp $n \times p$. Ta có

$$A = \begin{bmatrix} A^1 & \dots & A^n \end{bmatrix} = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \quad \text{và} \quad B = \begin{bmatrix} B^1 & \dots & B^p \end{bmatrix} = \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix}.$$

Khi đó, ta có

$$AB = \begin{bmatrix} A^1 B_1 + \cdots + A^n B_n \end{bmatrix} = \begin{bmatrix} AB^1 & \cdots & AB^p \end{bmatrix} = \begin{bmatrix} A_1 B \\ \vdots \\ A_m B \end{bmatrix}. \quad \square$$

## 2.2 Các tính chất của các phép toán

**2.2.1 Định lý.** *Với mọi $A, B, C \in \mathcal{M}_{m,n}(\mathbb{K})$, $O$ là ma trận không cấp $m \times n$ trên $\mathbb{K}$ và mọi $\lambda, \mu \in \mathbb{K}$, ta có*

*(1)* $A + B = B + A$

*(2)* $(A + B) + C = A + (B + C)$

*(3)* $A + O = A$

*(4)* $A + (-A) = O$

*(5)* $(\lambda + \mu)A = \lambda A + \mu A$

*(6)* $(\lambda\mu)A = \lambda(\mu A)$

*(7)* $\lambda(A + B) = \lambda A + \lambda B$

*(8)* $1A = A$.

*Trong các tính chất sau các ma trận được chỉ rõ cấp của chúng*

*(9)* $(AB)C = A(BC)$, *với* $A \in \mathcal{M}_{m,n}(\mathbb{K})$, $B \in \mathcal{M}_{n,p}(\mathbb{K})$, $C \in \mathcal{M}_{p,q}(\mathbb{K})$

*(10)* $(A + B)C = AC + BC$, *với* $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$, $C \in \mathcal{M}_{n,p}(\mathbb{K})$

$A(B + C) = AB + AC$, *với* $A \in \mathcal{M}_{m,n}(\mathbb{K})$, $B, C \in \mathcal{M}_{n,p}(\mathbb{K})$

*(11)* $\lambda(AB) = (\lambda A)B = A(\lambda B)$, *với* $A \in \mathcal{M}_{m,n}(\mathbb{K})$, $B \in \mathcal{M}_{n,p}(\mathbb{K})$, $\lambda \in \mathbb{K}$

*(12)* $I_m A = AI_n = A$, *với* $A \in \mathcal{M}_{m,n}(\mathbb{K})$, $I_m$ *và* $I_n$ *là các ma trận đơn vị trên* $\mathbb{K}$

*(13)* $(A + B)^t = A^t + B^t$, *với* $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$

$(\lambda A)^t = \lambda A^t$, *với* $A \in \mathcal{M}_{m,n}(\mathbb{K})$ *và* $\lambda \in \mathbb{K}$

*(14)* $(A^t)^t = A$, *với* $A \in \mathcal{M}_{m,n}(\mathbb{K})$

*(15)* $(AB)^t = B^t A^t$, *với* $A \in \mathcal{M}_{m,n}(\mathbb{K})$ *và* $B \in \mathcal{M}_{n,p}(\mathbb{K})$

*(16) Với* $A \in \mathcal{M}_n(\mathbb{K})$, *ta có* $A^t = A$ *khi và chỉ khi* $A$ *đối xứng, và* $A^t = -A$ *khi và chỉ khi* $A$ *phản đối xứng.*

*Chứng minh.* Việc chứng minh các tính chất trên khá đơn giản bằng cách dùng định nghĩa các phép toán. Ta hãy chứng minh tính chất (9). Với $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{n\times p}$, $C = (c_{ij})_{p\times q}$, ta có

$$\begin{aligned}((AB)C)_{ij} &= \sum_{k=1}^{p}(AB)_{ik}c_{kj} = \sum_{k=1}^{p}\Big(\sum_{l=1}^{n} a_{il}b_{lk}\Big)c_{kj} = \sum_{k=1}^{p}\sum_{l=1}^{n} a_{il}b_{lk}c_{kj}\\ &= \sum_{l=1}^{n}\sum_{k=1}^{p} a_{il}b_{lk}c_{kj} = \sum_{l=1}^{n} a_{il}\Big(\sum_{k=1}^{p} b_{lk}c_{kj}\Big) = \sum_{l=1}^{n} a_{il}(BC)_{lj}\\ &= (A(BC))_{ij}.\end{aligned}$$

Từ đó suy ra $(AB)C = A(BC)$. □

**2.2.2 Nhận xét.** **(i)** Nhờ các tính chất kết hợp của phép cộng và phép nhân nên ta có thể xét tổng và tích của một số hữu hạn các ma trận. Riêng tổng $A + (-B)$ được viết lại là $A - B$ và gọi là “$A$ trừ $B$”.

**(ii)** Trên tập hợp $\mathcal{M}_n(\mathbb{K})$, ta có thể thực hiện phép cộng các ma trận và phép nhân các ma trận. Từ các tính chất của phép cộng và phép nhân các ma trận ta nhận thấy tập $\mathcal{M}_n(\mathbb{K})$ cùng với hai phép toán ấy là một vành có đơn vị không giao hoán, và được gọi là *vành các ma trận vuông cấp $n$ trên $\mathbb{K}$.*

**2.2.3 Thí dụ. (OSV04)** Chứng minh rằng với mọi ma trận vuông cấp hai $A, B, C$ ta luôn có $(AB - BA)^{2004}C = C(AB - BA)^{2004}$.

*Giải.* Bằng tính toán trực tiếp ta nhận thấy ma trận $AB - BA$ có dạng $AB - BA = \begin{bmatrix} a & b\\ c & -a\end{bmatrix}$. Do đó, ta có

$$(AB - BA)^2 = \begin{bmatrix} a & b\\ c & -a\end{bmatrix}\begin{bmatrix} a & b\\ c & -a\end{bmatrix} = \begin{bmatrix} a^2+bc & 0\\ 0 & a^2+bc\end{bmatrix} = (a^2+bc)I.$$

Vì vậy ta được $(AB - BA)^{2004} = (a^2 + bc)^{1002}I$. Do đó

$$(AB - BA)^{2004}C = (a^2 + bc)^{1002}IC = C(AB - BA)^{2004}. \qquad \square$$

**2.2.4 Thí dụ. (OSV94)** Cho $A$ là một ma trận vuông cấp hai thỏa mãn điều kiện $A^2 = A$. Chứng minh rằng để $AX - XA = O$ với $X$ là ma trận vuông cấp hai và $O$ là ma trận không, điều kiện cần và đủ là tồn tại ma trận vuông cấp hai $X_0$ sao cho $X = AX_0 + X_0A - X_0$.

Thật vậy, giả sử tồn tại ma trận cấp hai $X_0$. Đặt $X = AX_0 + X_0A - X_0$. Khi đó ta có

$$\begin{aligned} AX - XA &= A(AX_0 + X_0A - X_0) - (AX_0 + X_0A - X_0)A \\ &= AX_0 + AX_0A - AX_0 - (AX_0A + X_0A - X_0A) \\ &= O. \end{aligned}$$

Ngược lại giả sử tồn tại $X$ sao cho $AX - XA = O$. Khi đó đặt $X_0 = 2AX - X$ và

$$\begin{aligned} AX_0 + X_0A - X_0 &= A(2AX - X) + (2AX - X)A - (2AX - X) \\ &= 2AX - AX + 2AX - XA - 2AX - X \\ &= X. \end{aligned}$$

□

**2.2.5 Thí dụ. (OSV01)** Cho ma trận $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{bmatrix}$. Tìm tất cả ma trận vuông $B$ cấp 3 sao cho $AB + BA = O$.

*Giải.* Dễ dàng thấy rằng $A^2 = I$. Giả sử $B$ là ma trận vuông cấp 3 thỏa $AB + BA = O$; khi đó ta suy ra $B = A^2B = -ABA$. Khi đó ta có

$$B = \frac{1}{2}(B + B) = \frac{1}{2}(A^2B - ABA) = \frac{1}{2}A^2B - \frac{1}{2}ABA = AX - XA$$

với $X = \frac{1}{2}AB$. Từ kết quả thu được gợi chúng ta dạng của ma trận $B$, đó là $B = AX - XA$ với $X$ là một ma trận vuông cấp 3 tùy ý. Bây giờ chúng ta kiểm tra lại

$$\begin{aligned} AB + BA &= A(AX - XA) + (AX - XA)A \\ &= A^2X - AXA + AXA - XA^2 \\ &= X - X = O. \end{aligned}$$

Vậy $B = AX - XA$ với $X$ là ma trận vuông cấp 3 tùy ý. □

**2.2.6 Thí dụ. (OSV05)** Cho ma trận $M = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Đặt $M^n = (b_{ij}(n))_{i,j=1,2,3}$, $n \in \mathbb{N}$, $n \geq 2$. Tính $S_n = \sum\limits_{i=1}^{3} \sum\limits_{j=1}^{3} b_{ij}(n)$.

*Giải.* Ta viết lại ma trận $M$ như sau

$$M = I + D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Dễ dàng kiểm tra được rằng $D^n = D$ với $n \geq 1$. Khi đó, ta áp dụng được công thức nhị thức Newton và thu được

$$\begin{aligned} M^n = (I + D)^n &= \sum_{k=0}^{n} C_n^k I^k D^{n-k} = \sum_{k=0}^{n-1} C_n^k D + I \\ &= \begin{bmatrix} 1 + \sum_{k=0}^{n-1} C_n^k & \sum_{k=0}^{n-1} C_n^k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 + \sum_{k=0}^{n-1} C_n^k \end{bmatrix}. \end{aligned}$$

Do $\sum\limits_{k=0}^{n-1} C_n^k = 2^n - 1$, ta có

$$M^n = \begin{bmatrix} 2^n & 2^n - 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^n \end{bmatrix}$$

Từ đó ta tính được $S_n = 3 \cdot 2^n$.

Ta cũng có thể phân tích ma trận $M$ ở dạng khác để tính $S_n$; chẳng hạn ta viết

$$M = A + B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Khi đó ta tính được $AB = 2B$ và $BA = B$. Từ đó bằng quy nạp ta chứng minh được $M^n = (A + B)^n = A^n + (2^n - 1)B$. Do đó $S_n = 2^n + 1 + 2^n + 2^n - 1 = 3 \cdot 2^n$. □

**2.2.7 Thí dụ. (OSV06)** Cho ma trận $A = \begin{bmatrix} 2006 & 1 & -2006 \\ 2005 & 2 & -2006 \\ 2005 & 1 & -2005 \end{bmatrix}$. Xác định các phần tử trên đường chéo chính của ma trận $S = I + A + A^2 + \cdots + A^{2006}$.

*Giải.* Ta viết lại $A = I + B$ với $B = \begin{bmatrix} 2005 & 1 & -2006 \\ 2005 & 1 & -2006 \\ 2005 & 1 & -2006 \end{bmatrix}$. Khi đó, ta có $B^2 = O$. Từ đó ta áp dụng được công thức nhị thức Newton để tính lũy thừa của ma trận $A$ hay có thể chứng minh bằng quy nạp được $A^k = I + kB$ với mọi $k \in \mathbb{N}$. Từ đó suy ra $S = 2007 + 1003 \cdot 2007B$. Do đó các phần tử trên đường chéo chính của $S$ là $s_{11} = 2007 + 1003 \cdot 2007 \cdot 2005$, $s_{22} = 2007 + 1003 \cdot 2007$, $s_{33} = 2007 - 1003 \cdot 2007 \cdot 2006$. □

**2.2.8 Thí dụ. (OSV07)** Cho ma trận $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$. Tìm tất cả các ma trận vuông $X$ cấp 4 sao cho $AX = XA$.

*Giải.* Từ điều kiện $AX = XA$ tương đương với $(A-2I)X = X(A-2I)$ nó được viết rõ ra rằng

$$\begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{bmatrix}$$
$$= \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Đẳng thức cuối tính ra được

$$\begin{bmatrix} -x_{21} & -x_{22} & -x_{23} & -x_{24} \\ -x_{31} & -x_{32} & -x_{33} & -x_{34} \\ -x_{41} & -x_{42} & -x_{43} & -x_{44} \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -x_{11} & -x_{12} & -x_{13} \\ 0 & -x_{21} & -x_{22} & -x_{23} \\ 0 & -x_{31} & -x_{32} & -x_{33} \\ 0 & -x_{41} & -x_{42} & -x_{43} \end{bmatrix}$$

Bằng cách đồng nhất các phần tử tương ứng của hai ma trận ta được

$$X = \begin{bmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{bmatrix} \qquad a, b, c, d \in \mathbb{R}.$$

Ngược lại dễ dàng kiểm tra được mọi ma trận $X$ có dạng như trên đều thỏa mãn điều kiện của bài toán. □

**2.2.9 Định nghĩa.** Cho $A$ là một ma trận vuông trên $\mathbb{K}$ và $p(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{K}[x]$ là một đa thức của biến $x$ với hệ số trên $\mathbb{K}$. Khi đó ma trận

$$a_0I + a_1A + \cdots + a_nA^n,$$

ở đó $I$ là ma trận đơn vị cùng cấp với $A$, được gọi là giá trị của đa thức $p(x)$ tại $x = A$, ký hiệu $p(A)$. Nó cũng gọi là **đa thức ma trận**. $A$ gọi là **nghiệm ma trận** của đa thức $p(x)$ nếu đa thức ma trận $p(A) = O$ (ma trận không cùng cấp với $A$).

**2.2.10 Thí dụ.** Cho $p(x) = x^2 - 2x - 1 \in \mathbb{R}[x]$ và cho $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \in \mathcal{M}_2(\mathbb{R})$. Hãy tính $p(A)$ và $p(B)$.

Ta có

$$\begin{aligned} p(A) &= A^2 - 2A - I_2 \\ &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^2 - 2\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \end{aligned}$$

và

$$\begin{aligned}
p(B) &= B^2 - 2B - I_2 \\
&= \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}^2 - 2\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 4 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} 2 & 4 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\
&= O_2
\end{aligned}$$

Do đó $B$ là một nghiệm ma trận của $p(x)$. □

## 2.3 Vết của ma trận

**2.3.1 Định nghĩa.** Cho ma trận vuông $A = (a_{ij})_n \in \mathcal{M}_n(\mathbb{K})$. Ta gọi tổng các phần tử trên đường chéo chính của $A$ là **vết**, ký hiệu $\operatorname{tr}(A)$, tức là

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

**2.3.2 Định lý.** *Cho* $A \in \mathcal{M}_{n,m}$, $B \in \mathcal{M}_{m,n}$. *Khi đó* $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

*Chứng minh.* Ta có

$$\operatorname{tr}(AB) = \sum_{k=1}^{n}(AB)_{kk} = \sum_{k=1}^{n}\sum_{j=1}^{m} a_{kj}b_{jk} = \sum_{j=1}^{m}\sum_{k=1}^{n} b_{jk}a_{kj} = \sum_{j=1}^{k}(BA)_{jj}$$
$$= \operatorname{tr}(BA).\square$$

## Bài tập

**1**) Cho các ma trận

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 2 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 3 \\ 2 & 1 \\ -3 & -2 \end{bmatrix} \quad C = \begin{bmatrix} 2 & 5 \\ 0 & 3 \\ 4 & 2 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Tính **(a)** $5A - 3B + 2C + 4D$, **(b)** $A + 2B - 3C - 5D$.

**2)** Cho ma trận $A = \begin{bmatrix} 1 & -2 & 6 \\ 4 & 3 & -8 \\ 2 & -2 & 5 \end{bmatrix}$. Tìm ma trận $X$ sao cho

**(a)** $3A + 2X = I$;　　**(b)** $5A - 3X = I$.

**3)** Thực hiện các phép tính sau

**(a)** $\begin{bmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{bmatrix}$　　**(b)** $\begin{bmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{bmatrix}$

**(c)** $\begin{bmatrix} 1 & 2 & -3 \\ 3 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 1 \\ 3 & 2 \\ 4 & 3 \end{bmatrix}$

**4)** Cho các ma trận $A = \begin{bmatrix} 2 & 0 & 1 \\ 3 & 1 & 2 \\ 0 & -1 & 0 \end{bmatrix}$ và $B = \begin{bmatrix} 1 & 0 & 1 & -1 \\ -1 & 1 & 2 & 0 \\ -1 & 1 & 3 & 1 \end{bmatrix}$. Tính các ma trận $AB$, $B^tA$, $B^t(A - A^t)$.

**5)** Cho ma trận $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Tính các ma trận $A^2$, $A^3$, $A^4$, $AA^t$ và $A^tA$.

**6)** Tính các lũy thừa

**(a)** $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}^3$　　**(b)** $\begin{bmatrix} 3 & 2 \\ -4 & -2 \end{bmatrix}^5$　　**(c)** $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^n$.

**(d)** $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n$　　**(e)** $\begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix}^{2008}$　　**(f)** $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^n$

**7)** Cho $p(x) = x^2 - x + 1$ và $q(x) = x^2 - 5x + 3$. Tính các đa thức ma trận $p(A)$ và $q(A)$ biết rằng

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

**(c)** $A = \begin{bmatrix} 2 & 0 & 1 \\ 3 & 1 & 2 \\ 0 & -1 & 0 \end{bmatrix}$ **(d)** $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

**8 )** Chứng minh rằng

**(a)** $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ là một nghiệm của $p(x) = x^3 - 3x^2 + 4$.

**(b)** $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_2(\mathbb{K})$ là nghiệm của $p(x) = x^2 - (a+d)x + (ad - bc) \in \mathbb{K}[x]$.

**9 )** Tìm ma trận $X \in \mathcal{M}_2(\mathbb{R})$ sao cho $AX = XA$ với

**(a)** $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ **(b)** $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ **(c)** $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ **(d)** $A = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$

**10 )** Ma trận $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ có tính chất $A^2 = O$. Có thể có ma trận đối xứng cấp 2 có tính chất này không?

**11 )** Cho $A$ và $B$ là các ma trận đối xứng cấp $n$. Chứng minh rằng $A + B$ và $\alpha A$ cũng là các ma trận đối xứng.

**12 )** Chứng minh rằng nếu $A$ là ma trận đối xứng thì $C^t AC$ cũng là ma trận đối xứng (với điều kiện là phép toán là tương thích).

**13 )** Tích của hai ma trận đối xứng có nhất thiết là ma trận đối xứng không?

**14 )** Cho $A$ là một ma trận. Chứng minh rằng $A^t A$ và $AA^t$ là hai ma trận đối xứng.

**15**) Cho $A$ và $B$ là các ma trận đối xứng cấp $n$. Chứng minh rằng $AB$ đối xứng khi và chỉ khi $AB = BA$.

**16**) Cho $A$ và $B$ là các ma trận phản đối xứng cấp $n$. Chứng minh rằng $A + B$ và $\alpha A$ cũng là các ma trận phản đối xứng.

**17**) Cho $A$ và $B$ là các ma trận tam giác trên cấp $n$. Chứng minh rằng $A + B$ và $\alpha A$ cũng là các ma trận tam giác trên.

**18**) Cho $A$ và $B$ là các ma trận chéo cấp $n$. Chứng minh rằng $A + B$ và $\alpha A$ cũng là các ma trận chéo.

**19**) Với mọi ma trận vuông $A$ cấp $n$ luôn tồn tại các ma trận vuông $B$ và $C$ cùng cấp $n$ sao cho $B$ là ma trận tam giác trên, $C$ là ma trận tam giác dưới thỏa $A = B + C$.

**20**) Chứng minh các tính chất trong Định lý 2.2.1.

**21**) Chứng minh rằng một ma trận vuông luôn là tổng của một ma trận đối xứng và một ma trận phản đối xứng.

**22**) Chứng minh các tính chất sau

**(a)** $\lambda O_{m,n} = O_{m,n} = 0A$ với $A \in \mathcal{M}_{m,n}(\mathbb{K})$.

**(b)** $\lambda A = O_{m,n}$ khi và chỉ khi $\lambda = 0$ hay $A = O_{m,n}$ với mọi $A \in \mathcal{M}_{m,n}(\mathbb{K})$.

**(c)** $(-\lambda)A = -(\lambda A)$ với mọi $A \in \mathcal{M}_{m,n}(\mathbb{K})$, với mọi $\lambda \in \mathbb{K}$.

**(d)** $\lambda(A - B) = \lambda A - \lambda B$ với mọi $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$, với mọi $\lambda, \mu \in \mathbb{K}$.

**(e)** $(\lambda - \mu)A = \lambda A - \mu A$ với mọi $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$, với mọi $\lambda, \mu \in \mathbb{K}$.

**(f)** Với mọi $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$ tồn tại duy nhất $C \in \mathcal{M}_{m,n}(\mathbb{K})$sao cho $A + C = B$.

**(g)** Với mọi $A, B, C \in \mathcal{M}_{m,n}(\mathbb{K})$, ta có $A + B = A + C$ suy ra $B = C$.

**23 )** Chứng minh rằng nếu $AB = BA$ với $A, B \in \mathcal{M}_m(\mathbb{K})$ thì

**(a)** $(A+B)^2 = A^2 + 2AB + B^2$

**(b)** $A^2 - B^2 = (A-B)(A+B)$

**(c)** $(A+B)^n = \sum_{k=0}^{n} C_n^k A^k B^{n-k}$ với $n$ nguyên dương.

**24 )** Chứng minh rằng không tồn tại các ma trận vuông $A, B \in \mathcal{M}_n(\mathbb{K})$ sao cho $AB - BA = I_n$.

**25 ) (OSV95)** Cho ma trận $M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Tính $M^n$ ứng với mọi $n$ nguyên dương cho trước. *Hướng dẫn: Viết lại $M = I + A$.*

**26 ) (OSV96)** Cho $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 2 \end{bmatrix}^n = \begin{bmatrix} a_{11}(n) & a_{12}(n) & a_{13}(n) \\ a_{21}(n) & a_{22}(n) & a_{23}(n) \\ a_{31}(n) & a_{32}(n) & a_{33}(n) \end{bmatrix}$. Tính $\lim\limits_{n\to\infty} \dfrac{a_{22}(n)}{a_{32}(n)}$ theo các bước sau:

**(a)** Đặt $u_n = \dfrac{a_{22}(n)}{a_{32}(n)}$. Chứng minh rằng $u_{n+1} = \dfrac{3u_n}{u_n + 2}$.

**(b)** Chứng minh rằng dãy $\{u_n\}$ giảm nghiêm ngặt và bị chặn dưới bởi 1.

**(c)** Tính giới hạn $\lim\limits_{n\to\infty} u_n$.

# § 3 Phép biến đổi sơ cấp và ma trận bậc thang

## 3.1 Phép biến đổi sơ cấp

**3.1.1 Tích ma trận với ma trận sơ cấp loại I** Cho ma trận $A$ cấp $m \times n$ và $E_{I;i,j}$ là ma trận sơ cấp loại I cấp $m$. Khi đó, tích $E_{I;i,j}A$ là ma trận thu được từ $A$ bằng cách đổi chỗ hai dòng $i$ và $j$, ký hiệu $d_i \leftrightarrow d_j$. Thật vậy, tích $E_{I;i,j}A$ có kết quả như sau

$$\begin{bmatrix}
1 & 0 & \dots & 0 & \dots & 0 & \dots & 0\\
0 & 1 & \dots & 0 & \dots & 0 & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & 0 & \dots & 1 & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & 1 & \dots & 0 & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & 0 & \dots & 0 & \dots & 1
\end{bmatrix}
\begin{bmatrix}
a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n}\\
a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2j} & \dots & a_{2n}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{j1} & a_{j2} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn}
\end{bmatrix}$$

$$= \begin{bmatrix}
a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n}\\
a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2j} & \dots & a_{2n}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{j1} & a_{j2} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn}
\end{bmatrix}$$

Nếu $E_{I;i,j}$ là ma trận sơ cấp loại I cấp $n$, thì tích $AE_{I;i,j}$ là ma trận thu được từ ma trận $A$ bằng cách đổi chỗ hai cột $i$ và $j$, ký hiệu $c_i \leftrightarrow c_j$. Thật vậy, tích $AE_{I;i;j}$ được tính như sau

$$\begin{bmatrix}
a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n}\\
a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2j} & \dots & a_{2n}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{j1} & a_{j2} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn}
\end{bmatrix}
\begin{bmatrix}
1 & 0 & \dots & 0 & \dots & 0 & \dots & 0\\
0 & 1 & \dots & 0 & \dots & 0 & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & 0 & \dots & 1 & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & 1 & \dots & 0 & \dots & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & 0 & \dots & 0 & \dots & 1
\end{bmatrix}$$

$$= \begin{bmatrix}
a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1i} & \dots & a_{1n}\\
a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2i} & \dots & a_{2n}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ii} & \dots & a_{in}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{j1} & a_{j2} & \dots & a_{jj} & \dots & a_{ji} & \dots & a_{jn}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mi} & \dots & a_{mn}
\end{bmatrix}$$

**3.1.2 Tích ma trận với ma trận sơ cấp loại II** Cho ma trận $A$ cấp $m \times n$ và $E_{II;i,\lambda}$ là ma trận sơ cấp loại II cấp $m$. Khi đó, tích $E_{II;i,\lambda}A$ là ma trận thu được từ $A$ bằng cách nhân dòng thứ $i$ cho $\lambda$, ký hiệu $d_i \to \lambda d_i$. Thật vậy,

$$E_{II;i,\lambda}A = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \lambda a_{i1} & \lambda a_{i2} & \dots & \lambda a_{ii} & \dots & \lambda a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mn} \end{bmatrix}.$$

Nếu $E_{II;i,\lambda}$ là ma trận sơ cấp loại II cấp $n$, thì tích $AE_{II;i,\lambda}$ là ma trận thu được từ $A$ bằng cách nhân cột thứ $i$ cho $\lambda$, ký hiệu $c_i \to \lambda c_i$. Thật vậy,

$$AE_{II;i,\lambda} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & \lambda a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \lambda a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & \lambda a_{ii} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & \lambda a_{mi} & \dots & a_{mn} \end{bmatrix}.$$

**3.1.3 Tích ma trận với ma trận sơ cấp loại III** Cho ma trận $A$ cấp $m \times n$ và $E_{III;i,j,\lambda}$ là ma trận sơ cấp loại III cấp $m$. Khi đó, tích $E_{III;i,j,\lambda}A$

là ma trận thu được từ $A$ bằng cách nhân dòng thứ $j$ cho $\lambda$ rồi cộng vào dòng thứ $i$, ký hiệu $d_i \to d_i + \lambda d_j$. Thật vậy, tích $E_{III;i,j,\lambda}A$ được tính như sau

$$\begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \dots & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1}+\lambda a_{j1} & a_{i2}+\lambda a_{j2} & \dots & a_{ii}+\lambda a_{ji} & \dots & a_{ij}+\lambda a_{jj} & \dots & a_{in}+\lambda a_{jn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}$$

Nếu $E_{III;i,j,\lambda}$ là ma trận sơ cấp loại III cấp $n$, thì tích $AE_{III;i,j,\lambda}$ là ma trận thu được từ $A$ bằng cách nhân cột thứ $i$ cho $\lambda$ rồi cộng vào cột thứ $j$, ký hiệu $c_i \to c_i + \lambda c_j$. Thật vậy, tích $AE_{III;i,j,\lambda}$ được tính như sau

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{ji} & \dots & a_{jj} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \dots & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1j}+\lambda a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2j}+\lambda a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{ij}+\lambda a_{ii} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{ji} & \dots & a_{jj}+\lambda a_{ji} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mi} & \dots & a_{mj}+\lambda a_{mi} & \dots & a_{mn} \end{bmatrix}$$

**3.1.4 Các phép biến đổi sơ cấp trên ma trận** Các phép biến đổi trên ma trận sau đây gọi là các phép biến đổi sơ cấp trên dòng của ma trận (elementary row operations on matrix)

**Loại (I):** Đổi chỗ hai dòng cho nhau

**Loại (II):** Nhân (chia) một dòng cho một số khác 0.

**Loại (III):** Nhân một dòng với một số bất kì rồi cộng vào một dòng khác.

Tương tự ta cũng có các phép biến đổi sơ cấp trên cột.

**3.1.5 Nhận xét.** Các phép biến đổi sơ cấp trên dòng của ma trận chính là phép nhân với ma trận sơ cấp với loại tương ứng và cấp thích hợp vào bên trái ma trận ấy.

Các phép biến đổi sơ cấp trên cột của ma trận chính là phép nhân với ma trận sơ cấp với loại tương ứng và cấp thích hợp vào bên phải ma trận ấy.

## 3.2 Ma trận bậc thang

Một dòng của ma trận được gọi là *dòng không* nếu nó chỉ gồm những phần tử 0. Ngược lại, nếu dòng của ma trận mà có ít nhất một phần tử khác 0 thì nó được gọi là *dòng khác không*.

**3.2.1 Định nghĩa.** Ma trận $A$ cấp $m \times n$ (khác ma trận không) được gọi là **ma trận bậc thang theo dòng** (row echelon matrix) nếu thỏa hai tính chất sau

- Các dòng khác không của $A$ luôn ở trên các dòng bằng không (nếu có).
- Trên hai dòng khác không của $A$, phần tử khác 0 đầu tiên của dòng dưới bao giờ cũng ở bên phải cột chứa phần tử khác 0 đầu tiên của dòng trên.

Ta diễn tả bằng kí hiệu như sau: ma trận $A = (a_{ij})_{m\times n}$ được gọi là *ma trận bậc thang theo dòng* nếu có một số nguyên dương $r \leq \min\{m, n\}$ và một dãy các chỉ số cho cột $1 \leq j_1 < j_2 < \cdots < j_r$ sao cho

**(1)** $a_{ij} = 0$ nếu hoặc $r < i \leq m$ hoặc $i \leq r$ và $j \leq j_i$.

**(2)** $a_{1j_1} a_{2j_2} \cdots a_{rj_r} \neq 0$.

Các phần tử $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ gọi là các *phần tử được đánh dấu* của $A$.

Nếu ngoài (1) và (2) còn có

**(3)** $a_{1j_1} = a_{2j_2} = \cdots = a_{rj_r} = 1$

**(4)** $a_{kj_i} = 0$ với $1 \leq k < i \leq r$ (tức là $a_{ij_i} = 1$ là phần tử khác không duy nhất trên cột thứ $j_i$)

thì ma trận $A$ được gọi là **ma trận bậc thang dòng rút gọn** (reduced row echelon matrix).

Tương tự, ta cũng có khái niệm *ma trận bậc thang theo cột* và *ma trận bậc thang cột rút gọn* (khi ma trận chuyển vị của nó là ma trận bậc thang theo dòng và ma trận bậc thang dòng rút gọn, tương ứng). Một ma trận vừa có dạng bậc thang dòng rút gọn vừa có dạng bậc thang cột rút gọn được gọi là *ma trận bậc thang chính tắc*. Như vậy, ma trận khác không $A = (a_{ij})_{m\times n}$ trên $\mathbb{K}$ với $m, n \geq 2$ là ma trận bậc thang chính tắc nếu có một số nguyên dương $r \leq \min\{m, n\}$ sao cho

- $a_{11} = a_{22} = \cdots = a_{rr} = 1$
- $a_{ij} = 0$ với $1 \leq i \leq m$, $1 \leq j \leq n$, $i \neq j$ hoặc $i = j \notin \{1, 2, \ldots, r\}$

Tức là $A$ có dạng

$$A = \left[\begin{array}{c|c} I_r & O_{r,n-r} \\ \hline O_{m-r,r} & O_{m-r,n-r} \end{array}\right]$$

Về sau ta chỉ quan tâm đến các ma trận bậc thang theo dòng, và chỉ nói ngắn gọn là ma trận bậc thang hay ma trận bậc thang thu gọn.

**3.2.2 Thí dụ.** Ma trận $\begin{bmatrix} 3 & 4 & 5 & 6 & 7 & 8 \\ 0 & 0 & 0 & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$ là một ma trận bậc thang theo dòng. Ma trận $\begin{bmatrix} 0 & 1 & 0 & 3 & 0 & 0 & 5 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$ là một ma trận bậc thang dòng rút gọn.

Ma trận $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 4 & 2 & 0 & 0 \\ 5 & 3 & 1 & 0 \\ 6 & 0 & 2 & 0 \end{bmatrix}$ là một ma trận bậc thang theo cột. Ma trận $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ là một ma trận bậc thang cột rút gọn.

Ma trận $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ là ma trận bậc thang chính tắc. □

**3.2.3 Định lý.** *Mọi ma trận khác không cấp $m \times n$ trên $\mathbb{K}$ (với $m, n \geq 2$) đều có thể đưa về ma trận bậc thang theo dòng sau một số hữu hạn các phép biến đổi sơ cấp theo dòng.*

*Chứng minh.* Ta chứng minh bằng phương pháp quy nạp theo số dòng $m$. Cho

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \end{bmatrix}$$

là một ma trận cấp $2 \times n$ khác không tùy ý trên $\mathbb{K}$. Lúc đó, $A$ phải có ít nhất một cột khác không. Gọi cột đầu tiên khác không của $A$ là $j_1$, $1 \leq j_1 \leq n$. Nếu cần có thể đổi chỗ hai dòng của $A$ để luôn có $a_{1j_1} \neq 0$. Ma trận $A$ có dạng

$$A = \begin{bmatrix} 0 & \dots & 0 & a_{1j_1} & a_{1,j_1+1} & \dots & a_{1n} \\ 0 & \dots & 0 & a_{2j_1} & a_{2,j_1+1} & \dots & a_{2n} \end{bmatrix}$$

Bây giờ ta biến đổi sơ cấp theo dòng trên $A$ như sau: nhân dòng thứ nhất cho $-\frac{a_{2j_1}}{a_{1j_1}}$ rồi cộng vào dòng thứ hai

$$A = \begin{bmatrix} 0 & \dots & 0 & a_{1j_1} & a_{1,j_1+1} & \dots & a_{1n} \\ 0 & \dots & 0 & a_{2j_1} & a_{2,j_1+1} & \dots & a_{2n} \end{bmatrix}$$

$$\xrightarrow{d_2 \to d_2 - \frac{a_{2j_1}}{a_{1j_1}} d_1} \begin{bmatrix} 0 & \dots & 0 & a_{1j_1} & a_{1,j_1+1} & \dots & a_{1n} \\ 0 & \dots & 0 & 0 & b_{2,j_1+1} & \dots & b_{2n} \end{bmatrix}$$

trong đó $b_{2k} = a_{2k} - \frac{a_{2j_1}}{a_{1j_1}} a_{1k}$ với $k = j_1 + 1, \dots, n$. Ma trận thu được ở trên chính là ma trận bậc thang. Như vậy sau một số hữu hạn các phép biến đổi sơ cấp ta đã đưa được ma trận $A$ về ma trận bậc thang.

Giả sử định lý đúng với mọi ma trận có cấp $(m-1) \times n'$ trên $\mathbb{K}$ tùy ý; nghĩa là mọi ma trận có số dòng $m-1$ (lớn hơn hoặc bằng 2) sau khi áp dụng một số hữu hạn các phép biến đổi sơ cấp có thể đưa về ma trận bậc thang. Xét một ma trận $A = (a_{ij})_{m \times n}$ khác không trên $\mathbb{K}$. Lúc đó, $A$ phải có ít nhất một cột khác không. Gọi cột đầu tiên khác không của $A$ là $j_1$, $1 \leq j_1 \leq n$. Nếu cần có thể đổi chỗ hai dòng nào đó của $A$ để nhận được $a_{1j_1} \neq 0$. Khi đó, ma trận $A$ có dạng

$$A = \begin{bmatrix} 0 & \dots & 0 & a_{1j_1} & a_{1,j_1+1} & \dots & a_{1n} \\ 0 & \dots & 0 & a_{2j_1} & a_{2,j_1+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{mj_1} & a_{m,j_1+1} & \dots & a_{mn} \end{bmatrix}$$

Bây giờ ta biến đổi sơ cấp theo dòng trên $A$ như sau: nhân dòng thứ nhất cho $-\frac{a_{ij_1}}{a_{1j_1}}$ rồi cộng vào dòng thứ $i$ với $i = 2, \dots, m$

$$A = \begin{bmatrix} 0 & \dots & 0 & a_{1j_1} & a_{1,j_1+1} & \dots & a_{1n} \\ 0 & \dots & 0 & a_{2j_1} & a_{2,j_1+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & a_{mj_1} & a_{m,j_1+1} & \dots & a_{mn} \end{bmatrix}$$

$$\xrightarrow[i=2,\dots,m]{d_i \to d_i - \frac{a_{ij_1}}{a_{1j_1}} d_1} \begin{bmatrix} 0 & \dots & 0 & a_{1j_1} & a_{1,j_1+1} & \dots & a_{1n} \\ 0 & \dots & 0 & 0 & b_{2,j_1+1} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & b_{m,j_1+1} & \dots & b_{mn} \end{bmatrix}$$

trong đó $b_{ik} = a_{ik} - \frac{a_{ij_1}}{a_{1j_1}} a_{1k}$ với $k = j_1 + 1, \ldots, n$ và $i = 2, \ldots, m$. Ma trận thu được nếu bỏ đi dòng thứ nhất thì được ma trận cấp $(m-1) \times n$. Nếu ma trận này là ma trận không thì ma trận thu được ở trên là ma trận bậc thang. Ngược lại ta có ma trận cấp $(m-1) \times n$ ấy có $j_1$ cột đầu tiên bằng không. Theo giả thiết quy nạp sau hữu hạn phép biến đổi sơ cấp ta có thể đưa ma trận này về ma trận bậc thang (có ít nhất $j_1$ cột đầu tiên bằng không). Bây giờ ta ghép lại với dòng đầu bị xóa sẽ được ma trận bậc thang cần tìm từ ma trận $A$. Như vậy chỉ sau hữu hạn phép biến đổi sơ cấp ta có thể đưa ma trận $A$ về ma trận bậc thang theo dòng. $\square$

**3.2.4 Nhận xét.** Từ ma trận bậc thang, sau hữu hạn phép biến đổi sơ cấp ta có thể đưa đến ma trận bậc thang rút gọn. Thật vậy, bắt đầu từ dòng khác không cuối cùng trong ma trận bậc thang. Chia dòng đó cho giá trị phần tử khác không đầu tiên (đây chính là phần tử được đánh dấu cuối cùng trong ma trận bậc thang). Rồi như trong chứng minh định lý trên ta thực hiện phép toán sơ cấp để làm cho các phần tử ở dòng trên của phần tử được đánh dấu cuối cùng đều bằng không. Ta tiếp tục quá trình như vậy cho các dòng trên.

Như ta đã biết mỗi phép biến đổi sơ cấp trên dòng của một ma trận thực chất là phép nhân bên trái của ma trận dòng ấy với ma trận sơ cấp. Cho nên định lý trên có thể phát biểu lại như sau:

> *Mọi ma trận khác không cấp $m \times n$ trên $\mathbb{K}$ (với $m, n \geq 2$) đều có thể đưa về ma trận bậc thang theo dòng sau khi nhân bên trái một số hữu hạn các ma trận sơ cấp.*

Nghĩa là với ma trận $A$ khác không bất kỳ có cấp $m \times n$ với $m, n \geq 2$ tồn tại một dãy các ma trận sơ cấp $E_1, E_2, \ldots E_l$ sao cho $E_l \cdots E_2 E_1 A$ là ma trận bậc thang.

**3.2.5 Thí dụ.** Đưa ma trận $A = \begin{bmatrix} 1 & 2 & -3 & 4 \\ 2 & 4 & 1 & 15 \\ 3 & 6 & -1 & 17 \end{bmatrix}$ về dạng bậc thang và dạng bậc thang rút gọn.

$$A = \begin{bmatrix} 1 & 2 & -3 & 4 \\ 2 & 4 & 1 & 15 \\ 3 & 6 & -1 & 17 \end{bmatrix} \xrightarrow[d_3 \to d_3 - 3d_1]{d_2 \to d_2 - 2d_1} \begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & 0 & 7 & 7 \\ 0 & 0 & 8 & 5 \end{bmatrix}$$

$$\xrightarrow{d_3 \to d_3 - \frac{8}{7}d_2} \begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & 0 & 7 & 7 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

Ma trận sau cùng ở trên là ma trận bậc thang cần tìm. Ta biến đổi tiếp

$$\begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & 0 & 7 & 7 \\ 0 & 0 & 0 & -3 \end{bmatrix} \xrightarrow{d_3 \to -\frac{1}{3}d_3} \begin{bmatrix} 1 & 2 & -3 & 4 \\ 0 & 0 & 7 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[d_2 \to d_2 - 7d_3]{d_1 \to d_1 - 4d_3} \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{d_2 \to \frac{1}{7}d_2} \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{d_1 \to d_1 + 3d_2} \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Ma trận thu được ở trên là ma trận bậc thang rút gọn cần tìm. □

## Bài tập

**1**) Trong mỗi cặp ma trận sau, tìm một ma trận sơ cấp $E$ sao cho $EA = B$.

**(a)** $A = \begin{bmatrix} 2 & -1 \\ 5 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} -4 & 2 \\ 5 & 3 \end{bmatrix}$

**(b)** $A = \begin{bmatrix} 2 & 1 & 3 \\ -2 & 4 & 5 \\ 3 & 1 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 1 & 4 \\ -2 & 4 & 5 \end{bmatrix}$

**(c)** $A = \begin{bmatrix} 4 & -2 & 3 \\ 1 & 0 & 2 \\ -2 & 3 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & -2 & 3 \\ 1 & 0 & 2 \\ 0 & 3 & 5 \end{bmatrix}$

**2**) Đưa các ma trận sau về dạng bậc thang và dạng bậc thang rút gọn:

**(a)** $\begin{bmatrix} 2 & 1 & 11 & 2 \\ 1 & 0 & 4 & -1 \\ 11 & 4 & 56 & 5 \\ 2 & -1 & 5 & -6 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & -3 & 4 & 5 & 2 \\ 0 & 1 & 3 & 4 & 6 \\ -3 & 5 & -2 & -3 & -4 \\ -2 & 3 & 5 & 6 & 4 \end{bmatrix}$

# § 4 Định Thức

## 4.1 Định thức cấp 2 và cấp 3

**4.1.1** Cho $A$ là một ma trận vuông cấp 2,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

**Định thức cấp 2** (determinant) của ma trận vuông $A$ là một số, kí hiệu là $\det A$ hoặc $|A|$, xác định như sau:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = \sum_{\sigma\in S_2} \operatorname{sign}(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}.$$

**4.1.2** Cho $A$ là một ma trận vuông cấp 3,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

**Định thức cấp 3** của ma trận vuông $A$ là một số, kí hiệu là $\det A$ hoặc $|A|$, xác định như sau

$$\begin{aligned} \det A &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\qquad - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \\ &= \sum_{\sigma\in S_3} \operatorname{sign}(\sigma) a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}. \end{aligned}$$

**4.1.3 Thí dụ.** Cho $A$ là ma trận vuông cấp 3 có các phần tử là 1 hoặc $-1$. Chứng minh rằng $|\det A| \leq 4$.

*Giải.* Theo công thức định thức ta có

$$\begin{aligned}\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}\end{aligned}$$

Vì mỗi số hạng chỉ nhận giá trị 1 hoặc $-1$ nên ta thấy ngay $|\det A| \leq 6$. Giả sử tồn tại ma trận $A$ như thế sao cho $|\det A| = 6$. Khi đó mỗi số hạng trong công thức trên hoặc cùng bằng 1 hoặc cùng bằng $-1$. Điều này không thể xảy ra vì tích ba số hạng đầu trái dấu với tích ba số hạng cuối

$$a_{11}a_{22}a_{33}a_{12}a_{23}a_{31}a_{13}a_{21}a_{32} \quad \text{và} \quad -a_{13}a_{22}a_{31}a_{12}a_{21}a_{33}a_{11}a_{23}a_{32}.$$

Vậy 6 số hạng trong công thức định thức ở trên không thể cùng dấu; nghĩa là phải có hai số hạng trái dấu nhau cho nên tổng của hai số hạng ấy triệt tiêu. Do đó $|\det A| \leq 4$. □

## 4.2 Định thức của ma trận

Cho $A = (a_{ij})_{n\times n}$ là ma trận vuông cấp $n$ trên $\mathbb{K}$. **Định thức** (determinant) của ma trận vuông $A$ là một số thuộc $\mathbb{K}$, kí hiệu $\det A$ hoặc $|A|$, xác định như sau

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Các dòng (hay cột) của $A$ cũng được gọi là các dòng (hay cột) của $\det A$. Vì $S_n$ có $n!$ phần tử nên tổng trên có $n!$ số hạng, mỗi số hạng là tích của $n$ phần tử nằm ở các dòng và các cột khác nhau.

**4.2.1 Nhận xét.** Ta nhận thấy trong công thức định thức của ma trận vuông mỗi số hạng đều chứa đúng một phần tử của một dòng hoặc cột tùy ý. Do đó, nếu trong một ma trận vuông có một dòng hay một cột mà các phần tử đều bằng 0, thì định thức của nó bằng 0.

**4.2.2 Thí dụ. (OSV04)** Cho ma trận $A = (a_{ij})_{n\times n}$ thỏa mãn điều kiện

$$a_{ij} = \begin{cases} 0 & \text{khi } i = j \\ \pm 1 & \text{khi } i \neq j \end{cases}$$

Chứng minh rằng

**(a)** Nếu $n = 3$ thì tồn tại ma trận $A$ sao cho $\det A = 0$.

**(b)** Nếu $n = 4$ ta luôn có $\det A \neq 0$.

*Giải.* **(a)** Xét ma trận $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$. Ta có

$$\det A = 0 - 1 + 1 + 0 + 0 + 0 = 0.$$

**(b)** Với $n = 4$ ma trận $A$ có dạng $A = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ a_{21} & 0 & a_{23} & a_{24} \\ a_{31} & a_{32} & 0 & a_{34} \\ a_{41} & a_{42} & a_{43} & 0 \end{bmatrix}$. Ta tính định thức của $A$: theo công thức có đến 24 số hạng ta chỉ ghi lại những số hạng khác không (có 9 số hạng như thế)

$$\begin{aligned} \det A = & a_{12}a_{21}a_{34}a_{43} - a_{12}a_{23}a_{34}a_{41} - a_{12}a_{24}a_{31}a_{43} \\ & - a_{13}a_{21}a_{34}a_{42} + a_{13}a_{24}a_{31}a_{42} - a_{13}a_{24}a_{32}a_{41} \\ & - a_{14}a_{21}a_{32}a_{43} - a_{14}a_{23}a_{31}a_{42} + a_{14}a_{23}a_{32}a_{41} \end{aligned}$$

Theo định nghĩa của ma trận $A$ mỗi số hạng trên bằng 1 hoặc $-1$. Do có 9 số hạng (số lẻ) nên không thể xảy ra trường hợp tổng của biểu thức trên bằng 0. Vậy ta luôn có $\det A \neq 0$. □

**4.2.3 Mệnh đề.** *Định thức của ma trận tam giác trên (dưới) bằng tích các phần tử trên đường chéo chính.*

*Chứng minh.* Giả sử $A = (a_{ij})$ là ma trận tam giác trên cấp $n$. Theo định nghĩa ta có $\det(A) = \sum\limits_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$. Ta tìm điều kiện của

$\sigma$ để có thể có $a_{1\sigma(1)} \cdots a_{n\sigma(n)} \neq 0$. Do $A$ là ma trận tam giác trên nên $a_{nj} = 0$ với mọi $j < n$ nên phải có $\sigma(n) = n$. Ta cũng có $a_{n-1,j} = 0$ với mọi $j < n-1$ và $\sigma(n-1) \in \{1, \ldots, n-1\}$, cho nên $\sigma(n-1) = n-1$. Tiếp tục lập luận như vậy ta sẽ có $\sigma(i) = i$. Như vậy trong công thức định thức của $A$ chỉ có một số hạng có thể khác không là $a_{11} \cdots a_{nn}$ ứng với phép đồng nhất $e$. Do $\operatorname{sign}(e) = 1$, cho nên $\det(A) = a_{11} \cdots a_{nn}$.

Đối với ma trận tam giác dưới ta cũng lập luận tương tự nhưng với chiều ngược lại là chỉ số $i$ đi từ 1 đến $n$. □

**4.2.4 Hệ quả.** *Định thức của ma trận sơ cấp loại II, $E_{II;i,\lambda}$ bằng $\lambda$, và định thức của ma trận sơ cấp loại III, $E_{III;i,j,\lambda}$, bằng* 1.

*Chứng minh.* Theo định nghĩa ta thấy ma trận sơ cấp loại II là những ma trận tam giác trên và ma trận sơ cấp loại III là ma trận tam giác trên hoặc ma trận tam giác dưới, cho nên theo mệnh đề trên ta có được kết quả. □

**4.2.5 Mệnh đề.** *Định thức của ma trận sơ cấp loại I (phải có cấp lớn hơn* 1*) bằng* $-1$.

*Chứng minh.* Trong ma trận sơ cấp loại I cấp $n$, $E_{I;i,j} = (a_{ij})$, chỉ có $n$ phần tử khác không và đều bằng 1, đó là $a_{kk} = 1$ khi $k \neq i, j$ và $a_{ij} = a_{ji} = 1$. Theo định nghĩa của định thức ta thấy chỉ có một số hạng khác không là $a_{11} \cdots a_{ij} \cdots a_{ji} \cdots a_{nn} = 1$ ứng với phép thế $\sigma = (i\ j)$. Do $\operatorname{sign}(i\ j) = -1$ nên $\det(E_{I;i,j}) = -1$. □

## 4.3 Các tính chất cơ bản của định thức

**4.3.1 Định lý.** *Cho $A = (a_{ij})_{n\times n}$ là một ma trận vuông cấp $n$ trên $\mathbb{K}$ và $A^t$ là ma trận chuyển vị của $A$. Khi đó* $\det A^t = \det A$. *Nói cách khác, định thức không thay đổi qua phép chuyển vị.*

*Chứng minh.* Do khi $\sigma$ chạy khắp $S_n$ thì $\sigma^{-1}$ cũng chạy khắp $S_n$ và ngược lại nên từ định nghĩa định thức ta có

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) a^t_{1\sigma^{-1}(1)} a^t_{2\sigma^{-1}(2)} \cdots a^t_{n\sigma^{-1}(n)}$$

$$
\begin{aligned}
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a^t_{\sigma(1)1} a^t_{\sigma(2)2} \cdots a^t_{\sigma(n)n} \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\
&= \det(A)
\end{aligned}
$$

do $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$ và

$$
\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma^{-1}(1) & \sigma^{-1}(2) & \dots & \sigma^{-1}(n) \end{pmatrix} = \begin{pmatrix} \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ 1 & 2 & \dots & n \end{pmatrix}
$$

cho nên ta có được đẳng thức $\operatorname{sign}(\sigma^{-1}) a^t_{1\sigma^{-1}(1)} a^t_{2\sigma^{-1}(2)} \cdots a^t_{n\sigma^{-1}(n)} = \operatorname{sign}(\sigma) a^t_{\sigma(1)1} a^t_{\sigma(2)2} \cdots a^t_{\sigma(n)n}$. □

**4.3.2 Định lý.** *Cho $A$ là một ma trận vuông cấp $n$. Nếu ta đổi chỗ hai dòng bất kì của $A$ được ma trận $B$ thì* $\det B = -\det A$.

*Chứng minh.* Giả sử $A = (a_{ij})$ và $B = (b_{ij})$ trong đó ma trận $B$ có được từ ma trận $A$ khi ta đổi chỗ hai dòng $i$ và $j$. Ta có

$$
\begin{aligned}
\det B &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) b_{1\sigma(1)} \cdots b_{i\sigma(i)} \cdots b_{j\sigma(j)} \cdots b_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{j\sigma(i)} \cdots a_{i\sigma(j)} \cdots a_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} -\operatorname{sign}(\sigma') a_{1\sigma'(1)} \cdots a_{i\sigma'(i)} \cdots a_{j\sigma'(j)} \cdots a_{n\sigma'(n)} \\
&= -\sum_{\sigma' \in S_n} \operatorname{sign}(\sigma') a_{1\sigma'(1)} \cdots a_{i\sigma'(i)} \cdots a_{j\sigma'(j)} \cdots a_{n\sigma'(n)} \\
&= -\det(A)
\end{aligned}
$$

trong đó $\sigma' = \sigma \circ (i\ j)$ và khi $\sigma$ chạy khắp $S_n$ thì $\sigma'$ chạy khắp $S_n$. □

**4.3.3 Hệ quả.** *Nếu ma trận vuông có hai dòng bằng nhau, thì định thức của nó bằng* $0$.

*Chứng minh.* Giả sử ta có ma trận vuông $A$ có hai dòng $i$ và $j$ bằng nhau. Khi đó, ta đổi chỗ hai dòng $i$ và $j$ thì ma trận vẫn không thay đổi. Tuy nhiên theo định lý trên định thức của nó phải đổi dấu. Do đó $\det A = -\det A$ hay $\det A = 0$. $\square$

**4.3.4 Định lý.** *Cho $A = (a_{ij})_{n\times n}$ là một ma trận vuông cấp $n$ trên $\mathbb{K}$ và giả sử dòng thứ $i$ nào đó của $A$ các phần tử có tính chất $a_{ij} = \lambda a'_{ij} + \mu a''_{ij}$. Gọi $B$ là ma trận vuông cấp $n$ thu được bằng cách thay dòng thứ $i$ của ma trận $A$ bởi ma trận dòng $\begin{bmatrix} a'_{i1} & a'_{i2} & \dots & a'_{in}\end{bmatrix}$ và gọi $C$ là ma trận vuông cấp $n$ thu được bằng cách thay dòng thứ $i$ của ma trận $A$ bởi ma trận dòng $\begin{bmatrix} a''_{i1} & a''_{i2} & \dots & a''_{in}\end{bmatrix}$. Khi đó, ta có $\det A = \lambda \det B + \mu \det C$.*

*Chứng minh.* Theo giả thiết định lý ta có

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \qquad B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a'_{i1} & a'_{i2} & \dots & a'_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

$$C = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a''_{i1} & a''_{i2} & \dots & a''_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Do đó

$$\begin{aligned} \det A &= \sum_{\sigma\in S_n} a_{1\sigma(1)}\cdots a_{i\sigma(i)}\cdots a_{n\sigma(n)} \\ &= \sum_{\sigma\in S_n} a_{1\sigma(1)}\cdots(\lambda a'_{i\sigma(i)} + \mu a''_{i\sigma(i)})\cdots a_{n\sigma(n)} \\ &= \lambda \sum_{\sigma\in S_n} a_{1\sigma(1)}\cdots a'_{i\sigma(i)}\cdots a_{n\sigma(n)} \\ &\qquad\qquad + \mu \sum_{\sigma\in S_n} a_{1\sigma(1)}\cdots \mu a''_{i\sigma(i)}\cdots a_{n\sigma(n)} \\ &= \lambda \det B + \mu \det C. \end{aligned}$$ $\square$

**4.3.5 Hệ quả.** *Cho $A$ là một ma trận vuông cấp $n$ trên $\mathbb{K}$.*

*(1) Nếu ta nhân một dòng nào đó của ma trận $A$ với một số $\lambda$ được ma trận $B$ thì* $\det B = \lambda \det A$.

*(2) Với mọi $\lambda \in \mathbb{K}$, ta có* $\det(\lambda A) = \lambda^n \det A$.

*(3) Nếu $A$ có hai dòng tỉ lệ với nhau thì định thức của nó bằng không.*

*(4) Nếu ta nhân một dòng của $A$ với một số bất kì rồi cộng vào dòng khác của $A$ thì định thức của nó không thay đổi.*

*Chứng minh.* Bài tập 4. □

**4.3.6 Thí dụ.** Ta dùng tính chất của định thức trong mục này để biến đổi và tính định thức:

$$\begin{vmatrix} 1+a_1b_1 & a_1b_2 \\ a_2b_1 & 1+a_2b_2 \end{vmatrix} = \begin{vmatrix} 1+a_1b_1 & a_1b_2 \\ 0 & 1 \end{vmatrix} + \begin{vmatrix} 1+a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_2 \end{vmatrix}$$
$$= 1 + a_1b_1 + \begin{vmatrix} 1 & 0 \\ a_2b_1 & a_2b_2 \end{vmatrix} + \begin{vmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_2 \end{vmatrix}$$
$$= 1 + a_1b_1 + a_2b_2 + 0.$$ □

**4.3.7 Thí dụ. (OSV08)** Cho $a_0$, $d$ là các số thực, dãy $a_0, a_1, \dots, a_n$ lập thành một cấp số cộng công sai $d$. Tính định thức của ma trận

$$A = \begin{bmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} & a_n \\ a_1 & a_0 & a_1 & \dots & a_{n-2} & a_{n-1} \\ a_2 & a_1 & a_0 & \dots & a_{n-3} & a_{n-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n-1} & a_{n-2} & a_{n-3} & \dots & a_0 & a_1 \\ a_n & a_{n-1} & a_{n-2} & \dots & a_1 & a_0 \end{bmatrix}$$

*Giải.* Cộng cột thứ nhất vào cột cuối để tính định thức của $A$ ta được

$$\det(A) = \begin{vmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} & 2a_0+nd \\ a_1 & a_0 & a_1 & \dots & a_{n-2} & 2a_0+nd \\ a_2 & a_1 & a_0 & \dots & a_{n-3} & 2a_0+nd \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n-1} & a_{n-2} & a_{n-3} & \dots & a_0 & 2a_0+nd \\ a_n & a_{n-1} & a_{n-2} & \dots & a_1 & 2a_0+nd \end{vmatrix}$$

lưu ý $a_k + a_{n-k} = a_0 + kd + a_0 + (n-k)d = 2a_0 + nd$. Tiếp theo với mỗi $k = 1, 2, \ldots, n-1$ ta lấy dòng $k+1$ trừ dòng $k$ thu được

$$\det(A) = \begin{vmatrix} a_0 & a_1 & a_2 & \ldots & a_{n-1} & 2a_0 + nd \\ d & -d & -d & \ldots & -d & 0 \\ d & d & -d & \ldots & -d & 0 \\ \multicolumn{6}{c}{\dotfill} \\ d & d & d & \ldots & -d & 0 \\ d & d & d & \ldots & d & 0 \end{vmatrix}$$

Cộng dòng thứ $n$ vào dòng thứ $k$ với $k = 2, \ldots, n-1$ ta được

$$\det(A) = \begin{vmatrix} a_0 & a_1 & a_2 & \ldots & a_{n-1} & 2a_0 + nd \\ 2d & 0 & 0 & \ldots & 0 & 0 \\ 2d & 2d & 0 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 2d & 2d & 2d & \ldots & 0 & 0 \\ d & d & d & \ldots & d & 0 \end{vmatrix}$$

Chuyển cột thứ $n$ về cột thứ nhất, cột thứ nhất sang cột thứ hai, ..., cột thứ $n-1$ đến cột cuối, mỗi lần chuyển như thế định thức thu được sẽ đổi dấu, do đó ta được

$$\det(A) = (-1)^n \begin{vmatrix} 2a_0 + nd & a_0 & a_1 & \ldots & a_{n-2} & a_{n-1} \\ 0 & 2d & 0 & \ldots & 0 & 0 \\ 0 & 2d & 2d & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 2d & 2d & \ldots & 2d & 0 \\ 0 & d & d & \ldots & d & d \end{vmatrix}$$

$$= (-1)^n (2n_0 + nd) \begin{vmatrix} 1 & a_0 & a_1 & \ldots & a_{n-2} & a_{n-1} \\ 0 & 2d & 0 & \ldots & 0 & 0 \\ 0 & 2d & 2d & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 2d & 2d & \ldots & 2d & 0 \\ 0 & d & d & \ldots & d & d \end{vmatrix}$$

$$= (-1)^n(2n_0+nd)\begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 2d & 0 & \dots & 0 & 0 \\ 0 & 2d & 2d & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 2d & 2d & \dots & 2d & 0 \\ 0 & d & d & \dots & d & d \end{vmatrix}$$
$$= (-1)^n(2n_0+nd)2^{n-1}d^n.$$

**4.3.8 Thí dụ. (OSV09)** Tìm tất cả các ma trận $A$ vuông cấp $n \geq 2$ sao cho với mọi ma trận vuông $B$ cấp $n$ ta đều có $\det(A+B) = \det A + \det B$.

*Giải.* Chọn $B = A$, ta có $\det(A+B) = \det(2A) = 2^n \det A$ và $\det A + \det B = 2\det A$, suy ra $2^n \det A = 2\det A$. Do $n \geq 2$ nên ta phải có $\det A = 0$. Như vậy ta phải có $\det(A+B) = \det B$ với mọi ma trận $B$. Ta chọn ma trận $B$ theo ma trận $A$ như sau

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -a_{12} & \dots & -a_{1n} \\ 0 & 1-a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1-a_{nn} \end{bmatrix}$$

Khi đó ma trận $A+B$ là ma trận tam giác với các phần tử nằm trên đường chéo lần lượt là $a_{11}, 1, \dots, 1$; và ta tính được $\det(A+B) = a_{11}$ và $\det B = 0$. Do đó, ta có $a_{11} = 0$. Đối với phần tử $a_{ij}$ bất kỳ của ma trận $A$, bằng cách đổi cột 1 với cột $j$ và dòng 1 với dòng $i$ ở các ma trận $A$ và $B$ ta được ma trận $A'$ và $B'$ với $a'_{11} = a_{ij}$. Khi đó ta có $\det(A'+B') = \det(A+B) = \det(B) = \det B'$. Như vậy $\det(A'+B') = \det B'$ với mọi ma trận $B'$. Ta trở lại bài toán ở trên, với cách chọn ma trận $B$ và $B'$ tương ứng thích hợp ta suy ra được $a_{ij} = a'_{11} = 0$. Vậy ma trận $A$ là ma trận không. □

**4.3.9 Thí dụ. (OSV10)** Cho $A, B$ là các ma trận vuông cấp 2010 với hệ số thực sao cho

$$\det A = \det(A+B) = \det(A+2B) = \dots = \det(A+2010B) = 0.$$

**(a)** Chứng minh rằng $\det(xA+yB) = 0$ với mọi $x, y \in \mathbb{R}$.

**(b)** Tìm ví dụ chứng tỏ kết luận trên không còn đúng nếu chỉ có

$$\det A = \det(A+B) = \det(A+2B) = \cdots = \det(A+2009B) = 0.$$

*Giải.* **(a)** Nhận thấy rằng định thức $P(t) = \det(A+tB)$ là một đa thức bậc 2010 theo $t$ có 2011 nghiệm $0, 1, 2, \ldots, 2010$ nên $P(t) \equiv 0$. Định thức $Q(t) = \det(tA+B)$ cũng là một đa thức bậc 2010 của $t$. Với $t \neq 0$ ta có $Q(t) = \det(tA+B) = t^{2010}\det(A+t^{-1}B) = t^{2010}P(t^{-1}) = 0$. Do đó ta cũng có $Q(t) \equiv 0$. Với $x = 0$, ta có $\det(xA+yB) = \det(yB) = y^{2010}\det B = y^{2010}Q(0) = 0$. Với $x \neq 0$, ta có

$$\det(xA+yB) = x^{2010}\det(A+\tfrac{y}{x}B) = x^{2010}P(\tfrac{y}{x}) = 0.$$

Vậy $\det(xA+yB) = 0$ với mọi $x, y \in \mathbb{R}$.

**(b)** Chọn $A = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 0 & \ldots & 0 \\ 0 & 0 & -2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -2009 \end{bmatrix}$ và $B = I$. Khi đó, dễ thấy rằng $\det A = \det(A+B) = \det(A+2B) = \cdots = \det(A+2009B) = 0$ nhưng $\det(A+2010B) = 2010!$. □

**4.3.10 Thí dụ. (OSV12)** Cho $A$ là ma trận vuông cấp $n \geq 2$ có các phần tử là các số chính phương lẻ. Chứng minh rằng $\det(A)$ chia hết cho $8^{n-1}$.

*Giải.* Ta biết rằng hiệu của hai số chính phương lẻ là một số chia hết cho 8; thật vậy $(2k+1)^2 - (2l+1)^2 = (2k+2l+2)(2k-2l) = 4(k-l+1+2l)(k-l)$. Khi tính định thức $\det(A)$ lấy dòng thứ $k$ trừ dòng thứ nhất với $k = 2, \ldots, n$. Lúc đó, từ dòng thứ hai trở đi mỗi phần tử đều chia hết cho 8. Do đó từ dòng thứ hai đến thứ $n$ lấy 8 làm nhân tử chung mỗi dòng. Từ đó suy ra được $\det(A)$ chia hết cho $8^{n-1}$. □

## Bài tập

**1)** Tính các định thức cấp 3 sau đây

**(a)** $\begin{vmatrix} \sin^2\alpha & 1 & \cos^2\alpha \\ \sin^2\beta & 1 & \cos^2\beta \\ \sin^2\gamma & 1 & \cos^2\gamma \end{vmatrix}$ **(b)** $\begin{vmatrix} \sin^2\alpha & \cos 2\alpha & \cos^2\alpha \\ \sin^2\beta & \cos 2\beta & \cos^2\beta \\ \sin^2\gamma & \cos 2\gamma & \cos^2\gamma \end{vmatrix}$

**(c)** $\begin{vmatrix} x & x' & ax+bx' \\ y & y' & ay+by' \\ z & z' & az'+bz' \end{vmatrix}$ **(d)** $\begin{vmatrix} (a_1+b_1)^2 & a_1^2+b_1^2 & a_1b_1 \\ (a_2+b_2)^2 & a_2^2+b_2^2 & a_2b_2 \\ (a_3+b_3)^2 & a_3^2+b_3^2 & a_3b_3 \end{vmatrix}$

**(e)** $\begin{vmatrix} a+b & c & 1 \\ b+c & a & 1 \\ c+a & b & 1 \end{vmatrix}$ **(f)** $\begin{vmatrix} (a^x+a^{-x})^2 & (a^x-a^{-x})^2 & 1 \\ (b^y+b^{-y})^2 & (b^y-b^{-y})^2 & 1 \\ (c^z+c^{-z})^2 & (c^z-c^{-z})^2 & 1 \end{vmatrix}$

**(g)** $\begin{vmatrix} 1 & \varepsilon & \varepsilon^2 \\ \varepsilon^2 & \varepsilon & 1 \\ x & y & z \end{vmatrix}$, ở đây $\varepsilon$ là một căn bậc ba của 1 trong $\mathbb{C}$.

**(h)** $\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \frac{x_1+\lambda x_2}{1+\lambda} & \frac{y_1+\lambda y_2}{1+\lambda} & 1 \end{vmatrix}$ và cho một ý nghĩa hình học của kết quả thu được.

**(i)** $\begin{vmatrix} \alpha & \beta & \gamma \\ \beta & \gamma & \alpha \\ \gamma & \alpha & \beta \end{vmatrix}$ ở đây $\alpha$, $\beta$, $\gamma$ là ba nghiệm của $x^3+px+q=0$.

**2)** Không tính, hãy dùng các tính chất của định thức để chứng minh các đẳng thức sau:

**(a)** $\begin{vmatrix} b+c & c+a & a+b \\ b_1+c_1 & c_1+a_1 & a_1+b_1 \\ b_2+c_2 & c_2+a_2 & a_2+b_2 \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}$

**(b)** $\begin{vmatrix} a_1 & b_1 & a_1x+b_1y+c_1 \\ a_2 & b_2 & a_2x+b_2y+c_2 \\ a_3 & b_3 & a_3x+b_3y+c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

**(c)** $\begin{vmatrix} a_1+b_1x & a_1-b_1x & c_1 \\ a_2+b_2x & a_2-b_2x & c_2 \\ a_3+b_3x & a_3-b_3x & c_3 \end{vmatrix} = -2x\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

**(d)** $$\begin{vmatrix} a_1+b_1i & a_1i+b_1 & c_1 \\ a_2+b_2i & a_2i+b_2 & c_2 \\ a_3+b_3i & a_3i+b_3 & c_3 \end{vmatrix} = 2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

**(e)** $$\begin{vmatrix} a_1+b_1x & a_1x+b_1 & c_1 \\ a_2+b_2x & a_2x+b_2 & c_2 \\ a_3+b_3x & a_3x+b_3 & c_3 \end{vmatrix} = (1-x^2)\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

**(f)** $$\begin{vmatrix} 1 & a & bc \\ 1 & b & ca \\ 1 & c & ab \end{vmatrix} = (b-a)(c-a)(c-b)$$

**3 )** Không tính, dùng tính chất của định thức chứng tỏ rằng các định thức sau bằng 0:

**(a)** $$\begin{vmatrix} 3 & -1 & 6 \\ 7 & 2 & 10 \\ 1 & 4 & -2 \end{vmatrix}$$

**(b)** $$\begin{vmatrix} 4 & 1 & 7 & -3 \\ 2 & 0 & -5 & 4 \\ 5 & 3 & 8 & -7 \\ -1 & 2 & 6 & -8 \end{vmatrix}$$

**4 )** Chứng minh Hệ quả 4.3.5.

# § 5 Định lý Laplace

## 5.1 Định thức con và phần bù

**5.1.1 Định nghĩa.** Cho $A = (a_{ij})_{n\times n}$ là ma trận vuông cấp $n$ trên $\mathbb{K}$ với $n \geq 2$. Nếu ta bỏ đi dòng $i$ cột $j$ của $A$ thì ta được ma trận vuông cấp $n-1$, kí hiệu là $M_{ij}$. Đặt $A_{ij} = (-1)^{i+j}\det M_{ij}$ được gọi là **phần bù đại số của phần tử** $a_{ij}$ (cofactor of $a_{ij}$).

**5.1.2 Mệnh đề.** *Cho $A = (a_{ij})_{n\times n}$. Khi đó,*

$$A_{ij} = \sum_{\substack{\sigma\in S_n \\ \sigma(i)=j}} \operatorname{sign}(\sigma) a_{1\sigma(1)}\cdots a_{i-1\sigma(i-1)}a_{i+1\sigma(i+1)}\cdots a_{n\sigma(n)}.$$

*Chứng minh.* Với kí hiệu như trong định nghĩa trên, ta có

$$M_{ij} = \begin{bmatrix} a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ \multicolumn{6}{c}{\dotfill} \\ a_{i-1,1} & \dots & a_{i-1j-1} & a_{i-1j+1} & \dots & a_{i-1n} \\ a_{i+1,1} & \dots & a_{i+1j-1} & a_{i+1j+1} & \dots & a_{i+1n} \\ \multicolumn{6}{c}{\dotfill} \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{bmatrix},$$

và đặt $M_{ij} = (b_{kl})_{n-1 \times n-1}$. Ta có

$$b_{kl} = \begin{cases} a_{kl} & \text{khi } k < i \text{ và } l < j \\ a_{k,l+1} & \text{khi } k < i \text{ và } l \geq j \\ a_{k+1,l} & \text{khi } k \geq i \text{ và } l < j \\ a_{k+1,l+1} & \text{khi } k \geq i \text{ và } l \geq j \end{cases}$$

Theo định nghĩa định thức ta có

$$\det M_{ij} = \sum_{\mu \in S_{n-1}} \operatorname{sign}(\mu) b_{1\mu(1)} \cdots b_{n-1\mu(n-1)}.$$

Kí hiệu $S(i,j) = \{\sigma \in S_n : \sigma(i) = j\}$. Với mỗi $\sigma \in S(i,j)$, đặt

$$\mu(k) = \begin{cases} \sigma(k) & \text{khi } k < i \text{ và } \sigma(k) < j \\ \sigma(k) - 1 & \text{khi } k < i \text{ và } \sigma(k) > j \\ \sigma(k+1) & \text{khi } k > i \text{ và } \sigma(k+1) < j \\ \sigma(k+1) - 1 & \text{khi } k > i \text{ và } \sigma(k+1) > j \end{cases}$$

Khi đó, ta có thể thấy $\mu \in S_{n-1}$. Xét phép thế $\tau$ cấp $n$ như sau $\tau = (n, n-1, \dots, j)\sigma(i, i+1, \dots, n)$. Ta tính được

$$\begin{aligned} \tau(n) &= (n, n-1, \dots, j)\sigma(i, i+1, \dots, n)(n) \\ &= (n, n-1, \dots, j)\sigma(i) \\ &= (n, n-1, \dots, j)(j) = n. \end{aligned}$$

Vậy $\tau \in S(n,n)$ và khi $\sigma$ chạy khắp $S(i,j)$ thì $\tau$ chạy khắp $S(n,n)$. Hơn nữa, với $1 \le k < n$, ta tính được

$$\tau(k) = \begin{cases} \sigma(k) & \text{khi } k < i \text{ và } \sigma(k) < j \\ \sigma(k) - 1 & \text{khi } k < i \text{ và } \sigma(k) > j \\ \sigma(k+1) & \text{khi } k > i \text{ và } \sigma(k+1) < j \\ \sigma(k+1) - 1 & \text{khi } k > i \text{ và } \sigma(k+1) > j \end{cases}$$

Vậy

$$\tau = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ \mu(1) & \mu(2) & \dots & \mu(n-1) & n \end{pmatrix}$$

hay $\tau|_{\{1,2,\dots,n-1\}} = \mu$. Do đó khi $\tau$ chạy khắp $S(n,n)$ thì $\mu$ chạy khắp $S_{n-1}$, nên khi $\sigma$ chạy khắp $S(i,j)$ thì $\mu$ chạy khắp $S_{n-1}$. Ta có thể thấy $\operatorname{sign}(\tau) = \operatorname{sign}(\mu)$ và

$$\begin{aligned} \operatorname{sign}(\tau) &= \operatorname{sign}(n, n-1, \dots, j) \operatorname{sign}(\sigma) \operatorname{sign}(i, i+1, \dots, n) \\ &= (-1)^{n-j+2+n-i+2} \operatorname{sign}(\sigma) \\ &= (-1)^{i+j} \operatorname{sign}(\sigma), \end{aligned}$$

suy ra $\operatorname{sign}(\mu) = (-1)^{i+j} \operatorname{sign}(\sigma)$. Mặt khác, ta có

$$\begin{aligned} b_{k\mu(k)} &= \begin{cases} a_{k\mu(k)} & \text{khi } k < i,\ \mu(k) < j \\ a_{k,\mu(k)+1} & \text{khi } k < i,\ \mu(k) \ge j \\ a_{k+1,\mu(k)} & \text{khi } k \ge i,\ \mu(k) < j \\ a_{k+1,\mu(k)+1} & \text{khi } k \ge i,\ \mu(k) \ge j \end{cases} \\ &= \begin{cases} a_{k\sigma(k)} & \text{khi } k < i,\ \mu(k) < j \\ a_{k,\sigma(k)+1} & \text{khi } k < i,\ \mu(k) \ge j \\ a_{k+1,\sigma(k+1)} & \text{khi } k \ge i,\ \mu(k) < j \\ a_{k+1,\sigma(k+1)} & \text{khi } k \ge i,\ \mu(k) \ge j \end{cases} \\ &= \begin{cases} a_{k\sigma(k)} & \text{khi } k < i \\ a_{k+1,\sigma(k+1)} & \text{khi } k \ge i \end{cases} \end{aligned}$$

Vậy

$$\det M_{ij} = \sum_{\mu \in S_{n-1}} \operatorname{sign}(\mu) b_{1\mu(1)} \cdots b_{n-1\mu(n-1)}$$
$$= \sum_{\sigma \in S(i,j)} (-1)^{i+j} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{i-1\sigma(i-1)} a_{i+1\sigma(i+1)} \cdots a_{n\sigma(n)}.$$

Do đó, ta có $A_{ij} = \sum\limits_{\substack{\sigma \in S_n \\ \sigma(i)=j}} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{i-1\sigma(i-1)} a_{i+1\sigma(i+1)} \cdots a_{n\sigma(n)}$. □

**5.1.3 Định nghĩa.** Cho ma trận vuông $A = (a_{ij})_{n\times n}$ cấp $n$. Ta lấy ra $k$ dòng bất kì của ma trận $A$: $i_1, i_2, \ldots, i_k$ với $1 \leq i_1 < i_2 < \cdots < i_k \leq n$; và lấy ra $k$ cột bất kì $j_1, j_2, \ldots, j_k$ với $1 \leq j_1 < j_2 \cdots < j_k \leq n$. Các phần tử nằm trên giao của $k$ dòng và $k$ cột đó tạo thành một ma trận vuông cấp $k$ kí hiệu $S_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}$. Ma trận này được gọi là **ma trận con** (submatrix) cấp $k$ của ma trận $A$. Xóa đi $k$ dòng đã chọn $(i_1, i_2, \ldots, i_k)$ và $k$ cột đã chọn $(j_1, j_2, \ldots, j_k)$ thì được một ma trận vuông cấp $n-k$, kí hiệu $M_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}$. Ma trận $M_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}$ là **ma trận con bù** của ma trận $S_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}$. Đặt $D_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k} = \det S_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}$ là **định thức con cấp $k$ của $A$** (minor of order $k$ of $A$) ứng với dòng $i_1, i_2, \ldots, i_k$; cột $j_1, j_2, \ldots j_k$. Đại lượng

$$A_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k} = (-1)^{i_1+i_2+\cdots+i_k+j_1+j_2+\cdots+j_k} \det M_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}$$

gọi là **phần bù đại số** của định thức $D_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}$ (cofactor of the minor $D_{i_1 \ldots i_k, j_1 \ldots j_k}$).

Cho $D(k)$ và $D(h)$ tương ứng là hai định thức con cấp $k$ và $h$ của $D$ với $1 \leq k \leq h \leq n$. Ta nói rằng $D(h)$ *bao* $D(k)$ nếu $D(k)$ là một định thức con của $D(h)$. Như vậy $D$ là định thức duy nhất bao mọi định thức con của nó.

**5.1.4 Nhận xét.** Các khái niệm ma trận con, định thức con có thể mở rộng cho các ma trận trên $\mathbb{K}$ cấp $m \times n$ tùy ý $(m, n \geq 2)$. Giả sử $A$ là một ma trận như thế, còn $k$ và $h$ là hai số nguyên sao cho $1 \leq k, h \leq \min\{m, n\}$. Mỗi ma trận $S$ mà các phần tử nằm trên giao của $k$ dòng và $h$ cột nào đó của $A$ gọi là một **ma trận con** cấp $k \times h$ của $A$; khi $k = h$, định thức $\det S$ cũng gọi là **định thức con** cấp $k$ của $A$. Đặc biệt, mỗi phần tử của $A$ vừa là ma trận con cấp $1 \times 1$, vừa là định thức con cấp 1 của $A$. Định thức con $D(h)$ cấp $h$ gọi là **bao** định thức con $D(k)$ cấp $k$ nếu $D(k)$ là một đnịh thức con của $D(k)$ (với $1 \leq k \leq h \leq \min\{m, n\}$).

**5.1.5 Thí dụ.** Xét ma trận cấp $n = 4$ như sau $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 3 & 0 & 1 & 2 \\ 2 & 3 & 0 & 1 \end{bmatrix}$.

Lúc đó $D_{13,14} = \begin{vmatrix} 1 & 4 \\ 3 & 2 \end{vmatrix}$, $D_{23,23} = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix}$, $D_{14,23} = \begin{vmatrix} 2 & 3 \\ 3 & 0 \end{vmatrix}$ là ba định thức con cấp 2 của $A$ với phần bù đại số lần lượt là $A_{13,14} = (-1)^{1+3+1+4} \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} = 6$, $A_{23,23} = (-1)^{2+3+2+3} \begin{vmatrix} 1 & 4 \\ 2 & 1 \end{vmatrix} = -7$, $A_{14,23} = (-1)^{1+4+2+3} \begin{vmatrix} 0 & 3 \\ 3 & 2 \end{vmatrix} = -9$. □

## 5.2 Khai triển định thức

**5.2.1 Định lý.** *Cho $A = (a_{ij})_{n \times n}$ là ma trận vuông cấp $n$ trên $\mathbb{K}$ với $n \geq 2$, $A_{ij}$ là phần bù đại số của $a_{ij}$ với $i, j = 1, \ldots, n$. Khi đó ta có*

$$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} \tag{5.2.2}$$

$$\det A = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{k=1}^{n} a_{kj}A_{kj}. \tag{5.2.3}$$

*Khi đó, công thức* (5.2.2) *được gọi là khai triển theo dòng thứ $i$ và công thức* (5.2.3) *được gọi là khai triển theo cột thứ $j$ của định thức ma trận $A$.*

*Chứng minh.* Vì định thức của một ma trận và ma trận chuyển vị của nó là như nhau, nên vai trò của dòng và cột là bình đẳng trong phép tính định thức, cho nên ta chỉ chứng minh công thức (5.2.2). Với mỗi $i, j = 1, 2, \ldots, n$, đặt $S(i, j) = \{\sigma \in S_n : \sigma(i) = j\}$. Khi đó

$$S_n = S(i, 1) \cup S(i, 2) \cup \cdots \cup S(i, n)$$

là một phân hoạch của $S_n$ thành $n$ tập con rời nhau $S(i,1), S(i,2), \ldots,$ $S(i,n)$. Do đó, ta có

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{i-1\sigma(i-1)} a_{i\sigma(i)} a_{i+1\sigma(i+1)} \cdots a_{n\sigma(n)} \\
&= \sum_{j=1}^{n} \sum_{\sigma \in S(i,j)} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{i-1\sigma(i-1)} a_{ij} a_{i+1\sigma(i+1)} \cdots a_{n\sigma(n)} \\
&= \sum_{j=1}^{n} a_{ij} \sum_{\sigma \in S(i,j)} \operatorname{sign}(\sigma) a_{1\sigma(1)} \cdots a_{i-1\sigma(i-1)} a_{i+1\sigma(i+1)} \cdots a_{n\sigma(n)} \\
&= \sum_{j=1}^{n} a_{ij} A_{ij}
\end{aligned}$$

do theo Mệnh đề 5.1.2 ta có công thức tính $A_{ij}$. □

**5.2.4 Thí dụ.** Tính định thức $D = \begin{vmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{vmatrix}$.

Khai triển theo dòng thứ nhất ta được

$$\begin{aligned}
D &= 1(-1)^{1+1} \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} + 1(-1)^{1+2} \begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} - 1(-1)^{1+3} \begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} \\
&= 2 + 2 + 0 = 4
\end{aligned}$$

Khai triển theo cột thứ ba ta được

$$\begin{aligned}
D &= -1(-1)^{1+3} \begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} + 1(-1)^{2+3} \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} + 1(-1)^{3+3} \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} \\
&= 0 + 2 + 2 = 4.
\end{aligned}$$

□

**5.2.5 Thí dụ. (OSV08)** Cho các số thực $a_1, a_2, \ldots, a_{2008}$. Chứng minh rằng tồn tại các ma trận thực vuông cấp $n > 1$: $A_1, A_2, \ldots, A_{2008}$ thỏa mãn $\det(A_k) = a_k$ với $k = 1, 2, \ldots, 2008$ và $\det\left(\sum\limits_{k=1}^{2008} A_k\right) = 2009$.

*Giải.* Các ma trận chúng ta cần tìm thỏa mãn điều kiện thường là các ma trận dễ tính được định thức. Chúng ta biết rằng ma trận đơn giản dễ tính định thức chính là ma trận chéo. Tuy nhiên nếu tất cả ma trận cần tìm đều là ma trận chéo thì tổng của chúng là ma trận chéo, và ta có thể thấy rằng định thức ấy không thỏa điều kiện cuối. Do vậy ma trận tổng không là ma trận chéo cũng không là ma trận tam giác, và ma trận đơn giản nhất lúc này là ma trận khi khai triển tính định thức phải có 2 số hạng khác không. Vì những lý do ấy các ma trận đơn giản nhất được đưa ra như sau

$$A_1 = \begin{bmatrix} a_1 & a & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \qquad A_2 = \begin{bmatrix} a_2 & 0 & 0 & \dots & 0 \\ b & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$A_k = \begin{bmatrix} a_k & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \qquad k = 3, 4, \dots, 2008$$

Khi đó $\det(A_k) = a_k$ và

$$\sum_{k=1}^{2008} A_k = \begin{bmatrix} s & a & 0 & \dots & 0 \\ b & 2008 & 0 & \dots & 0 \\ 0 & 0 & 2008 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2008 \end{bmatrix}$$

trong đó $s = \sum\limits_{k=1}^{2008} a_k$. Để tính định thức của tổng này ta khai triển theo dòng thứ nhất được

$$\det\left(\sum_{k=1}^{2008} A_k\right) = s2008^{n-1} - ab2008^{n-2},$$

suy ra ta cần có

$$ab = 2008s - \frac{2009}{2008^{n-2}}.$$

Ta có thể chọn $a = 1$ và $b = 2008s - \frac{2009}{2008^{n-2}}$ cho hai ma trận $A_1$ và $A_2$ ở trên. $\square$

**5.2.6 Thí dụ. (OSV12)** Cho $A$ là ma trận vuông cấp 5 có các phần tử là 1 hoặc $-1$. Chứng minh rằng $|\det A| \leq 64$.

*Giải.* Ta biết rằng định thức của ma trận vuông cấp 3 có các phần tử là 1 hoặc $-1$ có trị tuyệt đối nhỏ hơn hoặc bằng 4. Xét một ma trận $B$ vuông cấp 4 với các phần tử là 1 hoặc $-1$. Khi đó để tính định thức của nó ta khai triển theo một dòng (hay cột) gồm 4 số hạng, trong đó mỗi số hạng gồm $\pm 1$ nhân với một định thức của ma trận vuông cấp 3 có các phần tử là 1 hoặc $-1$. Do đó ta được $|\det B| \leq 4 \cdot 4 = 16$.

Xét một ma trận vuông $A$ cấp 5 với các phần tử là 1 hoặc $-1$. Ta có thể đổi dấu các dòng của $A$ sao cho cột thứ nhất gồm 5 số 1 và được ma trận $C$. Khi đó $|\det C| = |\det A|$.

- Trường hợp cột thứ hai của ma trận $C$ toàn số 1 hoặc toàn số $-1$ thì $\det C = 0$.

- Trường hợp cột thứ hai của ma trận $C$ có bốn số 1 hoặc bốn số $-1$. Khi đó ta khai triển tính định thức ma trận $C$ theo dòng ứng với phần tử trong cột thứ hai khác dấu với các phần tử còn lại trong cột thì chỉ có hai số hạng khác không mà mỗi số hạng là tích của $\pm 1$ với một định thức ma trận cấp 4 có các phần tử là 1 hoặc $-1$, cho nên $|\det A| \leq 2 \cdot 16 = 32$.

- Trường hợp cột thứ hai của ma trận $C$ có ba số 1 hoặc ba số $-1$. Khi đó lấy cột thứ hai nhân với $\pm 1$ rồi cộng vào cột một sao cho kết quả thu được có hai số 2 và các số còn lại bằng 0. Từ ma trận thu được ta khai triển theo cột thứ nhất để tính định thức thì được hai số hạng có thể khác không mà mỗi số hạng là tích của $\pm 2$ với định thức ma trận cấp 4 có các phần tử là 1 hoặc $-1$, cho nên $|\det A| = |\det C| \leq (2 + 2)16 = 64$.

Vậy ta luôn có $|\det A| \leq 64$. $\square$

**5.2.7 Hệ quả.** *Với mỗi ma trận vuông $A = (a_{ij})_n$ trên $\mathbb{K}$ với $n \geq 2$, ta đều có:*

**(1)** $$\sum_{j=1}^{n} a_{ij}A_{kj} = \delta_{ik}\det A = \begin{cases} \det A & \textit{khi } i = k \\ 0 & \textit{khi } i \neq k \end{cases} \textit{ với mọi } i, k = 1, \dots, n.$$

**(2)** $$\sum_{i=1}^{n} a_{ij}A_{ik} = \delta_{jk}\det A = \begin{cases} \det A & \textit{khi } j = k \\ 0 & \textit{khi } j \neq k \end{cases} \textit{ với mọi } i, k = 1, \dots, n.$$

*Chứng minh.* Ta chứng minh đẳng thức (1) ở trên. Khi $i = k$, ta có tổng $\sum\limits_{j=1}^{n} a_{ij}A_{ij}$ chính là khai triển (5.2.2) của định thức $\det A$, cho nên ta có được điều chứng minh. Khi $i \neq k$, xét ma trận $B = (b_{lj})_n$ nhận được từ $A$ bằng cách thay dòng thứ $k$ bởi dòng mới hoàn toàn giống dòng thứ $i$, tức là

$$b_{lj} = \begin{cases} a_{lj} & \text{khi } l \neq k \\ a_{ij} & \text{khi } l = k \end{cases} \qquad l, j = 1, \dots, n$$

Như vậy, $B$ có hai dòng giống nhau là dòng thứ $i$ và thứ $k$ nên $\det B = 0$. Mặt khác, theo định nghĩa ma trận $B$ ta nhận thấy $B_{kj} = A_{kj}$ với $j = 1, 2, \dots, n$. Khai triển $\det B$ theo dòng thứ $k$, ta được

$$0 = \det B = \sum_{j=1}^{n} b_{kj}B_{kj} = \sum_{j=1}^{n} a_{ij}A_{kj}$$

Tóm lại ta đã chứng minh được

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \begin{cases} \det A & \text{khi } i = k \\ 0 & \text{khi } i \neq k \end{cases} \qquad \text{với mọi } i, k = 1, \dots, n.$$

Tương tự ta cũng chứng minh được đẳng thức (2). □

**5.2.8 Hệ quả.** *Định thức của ma trận tam giác bằng tích của các phần tử trên đường chéo chính.*

*Chứng minh.* Ta chỉ việc áp dụng nhiều lần công thức khai triển tính định thức trong định lý khai triển định thức theo dòng hay cột. Ta minh họa cho trường hợp ma trận tam giác trên và khai triển theo những cột đầu tiên

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots \\ 0 & \dots & a_{nn} \end{vmatrix} \\ &= a_{11} \cdots a_{n-2,n-2} \begin{vmatrix} a_{n-1,n-1} & a_{n-1,n} \\ 0 & a_{nn} \end{vmatrix} \\ &= a_{11}a_{22} \cdots a_{nn} \end{aligned}$$

□

**5.2.9 Định lý. (Laplace)** *Cho $A$ là ma trận vuông cấp $n$ trên $\mathbb{K}$. Lấy ra $k$ dòng (tương ứng $k$ cột) bất kì, $1 \leq k \leq n-1$. Khi đó*

$$\det A = \sum D_k A_k$$

*trong đó $D_k$ là định thức con cấp $k$ của $A$ tựa trên $k$ dòng đã lấy và $A_k$ là phần bù đại số của $D_k$ tương ứng, và tổng được lấy theo tất cả các $D_k$ (Chú ý: có $C_n^k = \frac{n!}{k!(n-k)!}$ số hạng trong tổng).*

**5.2.10 Thí dụ.** Tính định thức $D = \begin{vmatrix} 3 & -1 & 5 & 2 \\ 6 & 0 & 3 & 0 \\ 1 & 0 & -2 & 0 \\ 4 & 7 & 0 & -5 \end{vmatrix}$. Chọn cột thứ hai và cột thứ tư (do có chứa nhiều số không). Từ hai cột này thiết lập được sáu định thức con cấp 2; chỉ có một định thức con cấp 2 khác không là $D_{14,24} = \begin{vmatrix} -1 & 2 \\ 7 & -5 \end{vmatrix} = -9$. Phần bù đại số của $D_{14,24}$ là $A_{14,24} = (-1)^{1+4+2+4} \begin{vmatrix} 6 & 3 \\ 1 & -2 \end{vmatrix} = 15$. Vậy $D = -9 \cdot 15 = -135$. □

**5.2.11 Thí dụ. (OSV07)** Giả sử $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ là ma trận vuông cấp 2 khả nghịch. Chứng minh rằng nếu $B$ là ma trận vuông cấp 2 khả nghịch thì ma trận $D$ cấp 4 xác định bởi ma trận khối $D = \begin{bmatrix} aB & bB \\ cA & dA \end{bmatrix}$ cũng khả nghịch.

*Giải.* Chúng ta áp dụng định lý Laplace đối với hai dòng đầu tiên được

$$\begin{aligned}\det(D) &= a^2\det(B)d^2\det(A) + b^2\det(B)c^2\det(A) \\ &\quad + \begin{vmatrix} aa' & bb' \\ ac' & bd' \end{vmatrix} \cdot \begin{vmatrix} cb & da \\ cd & dc \end{vmatrix} + \begin{vmatrix} ab' & ba' \\ ad' & bc' \end{vmatrix} \cdot \begin{vmatrix} ca & db \\ cc & dd \end{vmatrix} \\ &= (a^2d^2 + b^2c^2)\det(A)\det(B) + abcd\det(A)\det(B) \\ &\qquad + abcd\det(A)\det(B) \\ &= (ad - bc)^2\det(A)\det(B) \\ &= (\det(A))^3\det(B) \neq 0\end{aligned}$$

Vậy $D$ khả nghịch. □

**5.2.12 Thí dụ. (OSV93)** Cho hai ma trận thực vuông đồng cấp $A$ và $B$. Giả thiết rằng $\det(A+B) \neq 0$ và $\det(A-B) \neq 0$. Đặt

$$M = \begin{bmatrix} A & B \\ B & A \end{bmatrix}$$

Chứng minh rằng $\det(M) \neq 0$.

*Giải.* Ta thực hiện biến đổi sơ cấp và định lý Laplace để tính định thức ma trận $M$ như sau.

$$\begin{aligned}\det(M) &= \begin{vmatrix} A & B \\ B & A \end{vmatrix} = \begin{vmatrix} A-B & B-A \\ B & A \end{vmatrix} = \begin{vmatrix} A-B & O \\ B & A+B \end{vmatrix} \\ &= \det(A-B)\det(A+B)\end{aligned}$$

trong biến đổi sơ cấp lần thứ nhất ta nhân $-1$ với dòng thứ $n+i$ rồi cộng vào dòng thứ $i$ (với $n$ là cấp của ma trận $A$ và $B$), trong biến đổi sơ cấp lần thứ hai ta lấy cột thứ $j$ cộng vào cột thứ $n+j$, đẳng thức cuối cùng ta áp dụng định lý Laplace cho $n$ dòng đầu tiên. Từ giả thiết ta có ngay $\det(M) \neq 0$. □

**5.2.13 Thí dụ. (OSV95)** Cho hai số thực phân biệt $a$ và $b$ và cho $B = (b_{ij})$ là ma trận vuông cấp 6 được xác định như sau

$$b_{ij} = \begin{cases} x & \text{khi } i = j \\ a & \text{khi } i \neq j, i + j = 2n \\ b & \text{khi } i \neq j, i + j = 2n + 1 \end{cases}$$

Giả sử $\det(B) = \sum\limits_{k=0}^{6} \alpha_k (x - a)^k$. Tính $\alpha_4$.

*Giải.* Ta tính định thức $B$ như sau

$$\begin{aligned}
\det(B) &= \begin{vmatrix} x & b & a & b & a & b \\ b & x & b & a & b & a \\ a & b & x & b & a & b \\ b & a & b & x & b & a \\ a & b & a & b & x & a \\ b & a & b & a & b & x \end{vmatrix} \\
&= \begin{vmatrix} x & b & a & b & a & b \\ b & x & b & a & b & a \\ a - x & 0 & x - a & 0 & 0 & 0 \\ 0 & a - x & 0 & x - a & 0 & 0 \\ a - x & 0 & 0 & 0 & x - a & 0 \\ 0 & a - x & 0 & 0 & 0 & x - a \end{vmatrix} \\
&= (x - a)^4 \begin{vmatrix} x & b & a & b & a & b \\ b & x & b & a & b & a \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{vmatrix} \\
&= (x - a)^4 \begin{vmatrix} x + 2a & 3b & 0 & 0 & 0 & 0 \\ 3b & x + 2a & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{vmatrix} \\
&= ((x + 2a)^2 - 9b^2)(x - a)^4 \\
&= ((x - a)^2 + 6a(x - a) + 9a^2 - 9b^2)(x - a)^4
\end{aligned}$$

Do đó $\alpha_4 = 9(a^2 - b^2)$. $\square$

**Bài tập**

**1)** Tính các định thức sau:

**(a)** $\begin{vmatrix} 5 & -2 & 4 & -1 \\ 3 & 6 & -3 & 1 \\ 8 & 12 & 1 & 0 \\ 11 & -6 & 0 & 0 \end{vmatrix}$

**(b)** $\begin{vmatrix} 2 & 7 & 10 & -4 \\ 1 & 3 & 5 & 3 \\ 0 & 3 & -1 & 6 \\ 1 & 2 & 5 & 10 \end{vmatrix}$

**(c)** $\begin{vmatrix} 1 & 0 & -1 & -1 \\ 0 & -1 & -1 & 1 \\ a & b & c & d \\ -1 & -1 & 1 & 0 \end{vmatrix}$

**(d)** $\begin{vmatrix} 2 & 1 & 1 & x \\ 1 & 2 & 1 & y \\ 1 & 1 & 2 & z \\ 1 & 1 & 1 & t \end{vmatrix}$

**(e)** $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{vmatrix}$

**(f)** $\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix}$

**(g)** $\begin{vmatrix} 0 & 1 & 1 & a \\ 1 & 0 & 1 & b \\ 1 & 1 & 0 & c \\ a & b & c & 0 \end{vmatrix}$

**(h)** $\begin{vmatrix} a & 3 & 0 & 5 \\ 0 & b & 0 & 2 \\ 1 & 2 & c & 3 \\ 0 & 0 & 0 & d \end{vmatrix}$

**(i)** $\begin{vmatrix} 1 & 0 & 2 & a \\ 2 & 0 & b & 0 \\ 3 & c & 4 & 5 \\ d & 0 & 0 & 0 \end{vmatrix}$

**(j)** $\begin{vmatrix} 2 & -3 & 4 & 1 \\ 4 & -2 & 3 & 2 \\ a & b & c & d \\ 3 & -1 & 4 & 3 \end{vmatrix}$

**(k)** $\begin{vmatrix} 5 & a & 2 & -1 \\ 4 & b & 4 & -3 \\ 2 & c & 3 & -2 \\ 4 & d & 5 & -4 \end{vmatrix}$

**(l)** $\begin{vmatrix} 9 & 5 & 0 & 0 & 0 \\ 3 & -4 & 0 & 0 & 0 \\ 7 & 12 & 1 & 5 & 0 \\ 1 & 2 & -1 & -6 & 12 \\ 18 & 15 & 0 & 1 & 5 \end{vmatrix}$

# § 6 Các phương pháp tính định thức

Biểu thức định nghĩa của định thức cấp $n$ hoàn toàn không tiện lợi trong việc tính các định thức cấp $n \geq 4$. Để tính các định thức, nhất là các định thức cấp cao, ta cần sử dụng linh hoạt các tính chất của chúng, kết hợp với việc hạ cấp định thức nhờ vào các công thức khai triển (5.2.2) và (5.2.3) và Định lý Laplace. Các phép biến đổi sơ cấp cũng cho ta một phương pháp tính định thức rất hiệu quả.

## 6.1 Phương pháp khai triển theo các hàng và các cột

Khi thấy một dòng (hay cột) trong định thức có nhiều số 0 thì nên khai triển định thức theo dòng (hay cột) đó.

**6.1.1 Thí dụ.** Tính $D = \begin{vmatrix} 3 & -2 & 5 & 0 \\ 7 & 0 & 6 & 3 \\ 1 & 0 & 0 & 10 \\ -4 & 0 & 2 & 9 \end{vmatrix}$. Ta nhận thấy cột thứ hai trong định thức $D$ có ba số 0, cho nên chúng ta khai triển định thức theo cột thứ hai.

$$
\begin{aligned}
D &= (-1)^{1+2}(-2)\begin{vmatrix} 7 & 6 & 3 \\ 1 & 0 & 10 \\ -4 & 2 & 9 \end{vmatrix} \\
&= 2\left((-1)^{2+1}(1)\begin{vmatrix} 6 & 3 \\ 2 & 9 \end{vmatrix} + (-1)^{2+3}(10)\begin{vmatrix} 7 & 6 \\ -4 & 2 \end{vmatrix}\right) \quad \text{theo hàng thứ hai} \\
&= -2(6\cdot 9 - 2\cdot 3) - 20(7\cdot 2 + 4\cdot 6) \\
&= -856.
\end{aligned}
$$

□

Khi thấy trong định thức có $k$ dòng (hay cột) chứa nhiều số 0 thì nên khai triển định thức theo $k$ dòng (hay cột) đó.

**6.1.2 Thí dụ.** Tính định thức $D = \begin{vmatrix} 0 & 0 & 2 & 1 & 4 \\ 0 & 0 & 3 & -5 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & -4 \\ -2 & 8 & 6 & 3 & -9 \end{vmatrix}$. Ta nhận thấy hai cột đầu tiên có nhiều số 0, cho nên khai triển $D$ theo hai cột đó ta được

$$
D = \begin{vmatrix} 0 & 1 \\ -2 & 8 \end{vmatrix}(-1)^{4+5+1+2}\begin{vmatrix} 2 & 1 & 4 \\ 3 & -5 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 2\begin{vmatrix} 2 & 1 & 4 \\ 3 & -5 & 0 \\ 1 & 0 & 0 \end{vmatrix}
$$

Dòng thứ ba của định thức sau có nhiều số 0, nên khai triển nó theo dòng này ta được

$$D = 2(-1)^{3+1}(1)\begin{vmatrix} 1 & 4 \\ -5 & 0 \end{vmatrix} = 2(20) = 40. \qquad \square$$

## 6.2 Phương pháp biến đổi sơ cấp

Sử dụng các tính chất của định thức đặc biệt là Hệ quả 4.3.5 để biến đổi ma trận trong định thức cần tính về dạng tam giác.

**6.2.1 Thí dụ.** Tính định thức $D = \begin{vmatrix} x & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{vmatrix}$.

Dùng phép biến đổi sơ cấp ta tính được

$$\begin{aligned} D &\overset{d_1 \to d_1 + d_2 + d_3 + d_4}{=} \begin{vmatrix} x+3 & x+3 & x+3 & x+3 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{vmatrix} \\ &= (x+3)\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{vmatrix} \\ &\overset{\substack{d_2 \to d_2 - d_1 \\ d_3 \to d_3 - d_1 \\ d_4 \to d_4 - d_1}}{=} (x+3)\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & x-1 & 0 & 0 \\ 0 & 0 & x-1 & 0 \\ 0 & 0 & 0 & x-1 \end{vmatrix} \\ &= (x+3)(x-1)^3. \end{aligned}$$

Với cách làm như trên ta có được kết quả tổng quát cho định thức cấp $n$

$$\begin{vmatrix} x & 1 & 1 & \dots & 1 \\ 1 & x & 1 & \dots & 1 \\ 1 & 1 & x & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & x \end{vmatrix} = (x+n-1)(x-1)^{n-1}. \qquad \square$$

**6.2.2 Thí dụ. (OSV03)** Tính tổng $S_n = d_2 + d_3 + \cdots + d_n$, trong đó $d_k$ là các định thức cấp $k$, $k = 2, 3, \ldots, n$, dạng

$$d_k = \begin{vmatrix} 0 & 1 & \ldots & 1 & 1 \\ 1 & 0 & \ldots & x & x \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x & \ldots & 0 & x \\ 1 & x & \ldots & x & 0 \end{vmatrix}$$

*Giải.* Với $n = 2$ ta có $S_2 = d_2 = -1$. Với $n > 2$ và $k > 2$, nếu $x = 0$ ta có $d_k = 0$, cho nên $S_n = d_2 + d_3 + \cdots + d_n = d_2 = -1$. Nếu $x \neq 0$ và $k > 2$ thì theo kết quả của ví dụ trên ta có

$$d_k = x^{k-2} \begin{vmatrix} 0 & 1 & \ldots & 1 & 1 \\ 1 & 0 & \ldots & 1 & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & \ldots & 0 & 1 \\ 1 & 1 & \ldots & 1 & 0 \end{vmatrix} = x^{k-2}(k-1)(-1)^{k-1}$$

Vậy $S_n = -1 + 2x - 3x^2 + \cdots + (n-1)(-1)^{n-1}x^{n-2}$. Khi $x = -1$ ta được $S_n = -1 - 2 - \cdots - (n-1) = -\dfrac{n(n-1)}{2}$. Khi $x \neq -1$ ta có

$$\begin{aligned} S_n &= \sum_{k=2}^{n} [(-x)^{k-1}]' \\ &= \Big(\frac{-x - (-x)^n}{1 - (-x)}\Big)' \\ &= \frac{(-1 + n(-x)^{n-1})(1+x) + x + (-x)^n}{(1+x)^2} \\ &= \frac{-1 + n(-x)^{n-1} - (n-1)(-x)^n}{(1+x)^2} \end{aligned}$$

□

**6.2.3 Thí dụ.** Tính $D = \begin{vmatrix} 3 & 1 & 2 & 4 \\ 0 & 0 & -1 & 6 \\ 2 & 1 & 3 & 1 \\ 2 & -2 & 3 & 1 \end{vmatrix}$. Ta tiến hành như sau

$$D \overset{c_1 \leftrightarrow c_2}{=} - \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & 0 & -1 & 6 \\ 1 & 2 & 3 & 1 \\ -2 & 2 & 3 & 1 \end{vmatrix}$$

$$\underset{d_4 \to d_4 + 2d_1}{\overset{d_3 \to d_3 - d_1}{=}} - \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & 0 & -1 & 6 \\ 0 & -1 & 1 & -3 \\ 0 & 8 & 7 & 9 \end{vmatrix}$$

$$\overset{d_2 \leftrightarrow d_3}{=} \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & -1 & 1 & -3 \\ 0 & 0 & -1 & 6 \\ 0 & 8 & 7 & 9 \end{vmatrix}$$

$$\overset{d_4 \to d_4 + 8d_2}{=} \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & -1 & 1 & -3 \\ 0 & 0 & -1 & 6 \\ 0 & 0 & 15 & -15 \end{vmatrix}$$

$$\overset{d_4 \to d_4 + 15d_3}{=} \begin{vmatrix} 1 & 3 & 2 & 4 \\ 0 & -1 & 1 & -3 \\ 0 & 0 & -1 & 6 \\ 0 & 0 & 0 & 75 \end{vmatrix}$$

$$= (1)(-1)(-1)(75) = 75. \qquad \square$$

**6.2.4 Thí dụ.** Tính $D = \begin{vmatrix} 1 & 2 & 2 & \dots & 2 \\ 2 & 2 & 2 & \dots & 2 \\ 2 & 2 & 3 & \dots & 2 \\ \dots & \dots & \dots & \dots & \dots \\ 2 & 2 & 2 & \dots & n \end{vmatrix}$. Với $n = 1$, ta có $D = 1$; với $n = 2$ ta có $D = -2$; và với $n > 2$ ta biến đổi như sau

$$D \underset{d_i \to d_i - d_2, i>2}{\overset{d_2 \to d_2 - 2d_1}{=}} \begin{vmatrix} 1 & 2 & 2 & \dots & 2 \\ 0 & -2 & -2 & \dots & -2 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n-2 \end{vmatrix} = (n-2)!(-2). \qquad \square$$

## 6.3 Phương pháp quy nạp và phương pháp truy hồi

Ta đã biết phương pháp quy nạp, còn nội dung phương pháp truy hồi là biểu diễn định thức cần tính qua những định thức có cấp thấp hơn có dạng xác định và theo một công thức xác định. Tính các định thức cấp thấp ta sẽ lần lượt tính được các định thức cấp cao hơn.

**6.3.1 Thí dụ.** Dùng phương pháp quy nạp, tính định thức

$$D_n = \begin{vmatrix} -a_1 & a_1 & 0 & \dots & 0 & 0 \\ 0 & -a_1 & a_2 & \dots & 0 & 0 \\ 0 & 0 & -a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -a_n & a_n \\ 1 & 1 & 1 & \dots & 1 & 1 \end{vmatrix}.$$

Với $n > 1$, khai triển định thức theo cột cuối ta có:

$$D_n = \begin{vmatrix} -a_1 & a_1 & 0 & \dots & 0 \\ 0 & -a_1 & a_2 & \dots & 0 \\ 0 & 0 & -a_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -a_n \end{vmatrix} - a_n \begin{vmatrix} -a_1 & a_1 & 0 & \dots & 0 \\ 0 & -a_1 & a_2 & \dots & 0 \\ 0 & 0 & -a_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 \end{vmatrix}$$

$$= (-1)^n \prod_{i=1}^{n} a_i - a_n D_{n-1}.$$

Hãy xét vài trường hợp để dự đoán kết quả. Với $n = 1$,

$$D_1 = \begin{vmatrix} -a_1 & a_1 \\ 1 & 1 \end{vmatrix} = -2a_1 = (-1)2a_1.$$

Với $n = 2$,

$$D_2 = (-1)^2 a_1 a_2 - a_2 D_1 = a_1 a_2 + 2a_1 a_2 = 3a_1 a_2 = (-1)^2 3a_1 a_2.$$

Từ đó ta dự đoán $D_n = (-1)^n (n + 1) \prod_{i=1}^{n} a_i$. Ta chứng minh công thức này bằng quy nạp theo $n$. Hiển nhiên công thức đúng với $n = 1$,

$n = 2$. Bây giờ ta giả sử $n > 2$ và công thức đúng với $n-1$; tức là: $D_{n-1} = (-1)^{n-1} n \prod_{i=1}^{n-1} a_i$. Khi đó

$$D_n = (-1)^n \prod_{i=1}^{n} a_i - a_n D_{n-1} = (-1)^n \prod_{i=1}^{n} a_i - a_n (-1)^{n-1} n \prod_{i=1}^{n-1} a_i$$
$$= (-1)^n \Big[ \prod_{i=1}^{n} a_i + n \prod_{i=1}^{n} a_i \Big] = (-1)^n (n+1) \prod_{i=1}^{n} a_i.$$

Vậy $D_n = (-1)^n (n+1) \prod_{i=1}^{n} a_i$. □

**6.3.2 Thí dụ.** Tính định thức cấp $n$

$$D_n = \begin{vmatrix} 5 & 3 & 0 & \dots & 0 & 0 \\ 2 & 5 & 3 & \dots & 0 & 0 \\ 0 & 2 & 5 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 5 & 3 \\ 0 & 0 & 0 & \dots & 2 & 5 \end{vmatrix}$$

Với $n > 2$, khai triển theo dòng thứ nhất ta có

$$D_n = 5 \begin{vmatrix} 5 & 3 & \dots & 0 & 0 \\ 2 & 5 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 5 & 3 \\ 0 & 0 & \dots & 2 & 5 \end{vmatrix} - 3 \begin{vmatrix} 2 & 3 & \dots & 0 & 0 \\ 0 & 5 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 5 & 3 \\ 0 & 0 & \dots & 2 & 5 \end{vmatrix}$$

Trong đẳng thức trên định thức thứ nhất là $D_{n-1}$ ta khai triển định thức thứ hai theo cột thứ nhất và thu được kết quả $D_n = 5D_{n-1} - 6D_{n-2}$. Ta được công thức truy hồi $D_n = 5D_{n-1} - 6D_{n-2}$. Mặt khác, ta có $D_1 = 5$, $D_2 = 19$. Bằng phương pháp truy hồi ta tính được $D_n = 3^{n+1} - 2^{n+1}$. □

**6.3.3 Thí dụ.** Tính định thức Vandermonde (Vandermonde determinant) cấp $n$

$$W_n(a_1, a_2, \dots, a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix}.$$

Với $n > 1$ ta nhân cột thứ $j$ với $-a_1$ rồi cộng vào cột thứ $j+1$ với $j = 1, \ldots, n-1$, ta được

$$W_n(a_1, a_2, \ldots, a_n) = \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & a_2 - a_1 & a_2(a_2 - a_1) & \ldots & a_2^{n-2}(a_2 - a_1) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & a_n - a_1 & a_n(a_n - a_1) & \ldots & a_n^{n-2}(a_n - a_1) \end{vmatrix}.$$

Khai triển ma trận trên theo dòng thứ nhất, sau đó đưa thừa số chung mỗi dòng ra ngoài dấu định thức, ta nhận được

$$W_n(a_1, a_2, \ldots, a_n) = (a_2 - a_1) \ldots (a_n - a_1) \begin{vmatrix} 1 & a_2 & \ldots & a_2^{n-2} \\ \ldots & \ldots & \ldots & \ldots \\ 1 & a_n & \ldots & a_n^{n-2} \end{vmatrix}.$$

Để ý rằng định thức ở vế phải chính là định thức Vandermonde với $n-1$ phần tử $W_{n-1}(a_2, \ldots, a_n)$. Như vậy, ta có

$$W_n(a_1, a_2, \ldots, a_n) = (a_2 - a_1) \ldots (a_n - a_1) W_{n-1}(a_2, \ldots, a_n).$$

Dễ thấy rằng $W_2(a_{n-1}, a_n) = a_n - a_{n-1}$. Từ công thức dạng truy hồi trên ta có $W_3(a_{n-2}, a_{n-1}, a_n) = (a_{n-1} - a_{n-2})(a_n - a_{n-2})(a_n - a_{n-1})$. Ta có thể suy ra được $W_n(a_1, a_2, \ldots, a_n) = \prod\limits_{1 \le j < i \le n} (a_i - a_j)$. $\square$

**6.3.4 Thí dụ. (OSV93)** Cho $2n$ số nguyên $a_1, \ldots, a_n; b_1, \ldots, b_n$, thỏa mãn điều kiện $a_1 b_1 + \cdots + a_n b_n = 0$. Tính định thức

$$\begin{vmatrix} 1 + a_1 b_1 & a_1 b_2 & \ldots & a_1 b_n \\ a_2 b_1 & 1 + a_2 b_2 & \ldots & a_2 b_n \\ \ldots & \ldots & \ldots & \ldots \\ a_n b_1 & a_n b_2 & \ldots & 1 + a_n b_n \end{vmatrix}$$

Đặt định thức cần tính là $\Delta_n$. Khi đó ta biến đổi được

$$\Delta_n = \begin{vmatrix} 1+a_1b_1 & a_1b_2 & \dots & a_1b_{n-1} & a_1b_n \\ a_2b_1 & 1+a_2b_2 & \dots & a_2b_{n-1} & a_2b_n \\ \multicolumn{5}{c}{\dotfill} \\ a_{n-1}b_1 & a_{n-1}b_2 & \dots & 1+a_{n-1}b_{n-1} & a_{n-1}b_n \\ 0 & 0 & \dots & 0 & 1 \end{vmatrix}$$

$$+ \begin{vmatrix} 1+a_1b_1 & a_1b_2 & \dots & a_1b_{n-1} & a_1b_n \\ a_2b_1 & 1+a_2b_2 & \dots & a_2b_{n-1} & a_2b_n \\ \multicolumn{5}{c}{\dotfill} \\ a_{n-1}b_1 & a_{n-1}b_2 & \dots & 1+a_{n-1}b_{n-1} & a_{n-1}b_n \\ a_nb_1 & a_nb_2 & \dots & a_nb_{n-1} & a_nb_n \end{vmatrix}$$

$$= \Delta_{n-1} + a_nb_n \begin{vmatrix} 1+a_1b_1 & a_1b_2 & \dots & a_1b_{n-1} & a_1 \\ a_2b_1 & 1+a_2b_2 & \dots & a_2b_{n-1} & a_2 \\ \multicolumn{5}{c}{\dotfill} \\ a_{n-1}b_1 & a_{n-1}b_2 & \dots & 1+a_{n-1}b_{n-1} & a_{n-1} \\ b_1 & b_2 & \dots & b_{n-1} & 1 \end{vmatrix}$$

$$= \Delta_{n-1} + a_nb_n \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \multicolumn{4}{c}{\dotfill} \\ b_1 & b_2 & \dots & 1 \end{vmatrix}$$

$$= \Delta_{n-1} + a_nb_n$$

Từ công thức truy hồi $\Delta_n = \Delta_{n-1} + a_nb_n$ ta suy ra được $\Delta_n = 1 + a_1b_1 + \cdots + a_nb_n = 1$. □

**6.3.5 Thí dụ. (OSV01)** Cho $a, b \in \mathbb{R}$ với $a \neq b$. Tính định thức của ma trận cấp $n$ sau:

$$\begin{bmatrix} a+b & ab & 0 & \dots & 0 & 0 & 0 \\ 1 & a+b & ab & \dots & 0 & 0 & 0 \\ 0 & 1 & a+b & \dots & 0 & 0 & 0 \\ \multicolumn{7}{c}{\dotfill} \\ 0 & 0 & 0 & \dots & a+b & ab & 0 \\ 0 & 0 & 0 & \dots & 1 & a+b & ab \\ 0 & 0 & 0 & \dots & 0 & 1 & a+b \end{bmatrix}$$

*Giải.* Gọi $\Delta_n$ là đa thức cần tìm. Với $n > 2$ ta có

$$\Delta_n = (a+b)\Delta_{n-1} - ab\begin{vmatrix} 1 & ab & \dots & 0 & 0 & 0 \\ 0 & a+b & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a+b & ab & 0 \\ 0 & 0 & \dots & 1 & a+b & ab \\ 0 & 0 & \dots & 0 & 1 & a+b \end{vmatrix}$$

$$= (a+b)\Delta_{n-1} - ab\Delta_{n-2}.$$

Mặt khác, ta có

$$\Delta_1 = a+b = \frac{a^2-b^2}{a-b} \qquad \Delta_2 = (a+b)^2 - ab = \frac{a^3-b^3}{a-b}.$$

Từ đó ta tính tiếp $\Delta_3 = (a+b)\dfrac{a^3-b^3}{a-b} - ab\dfrac{a^2-b^2}{a-b} = \frac{a^4-b^4}{a-b}$, $\Delta_4 = (a+b)\dfrac{a^4-b^4}{a-b} - ab\dfrac{a^3-b^3}{a-b} = \dfrac{a^5-b^5}{a-b}$. Bằng quy nạp ta chứng minh được

$$\Delta_n = \frac{a^{n+1}-b^{n+1}}{a-b}.$$ □

## 6.4 Phương pháp tách định thức thành nhiều định thức

Từ kết quả trong Định lý 4.3.4 ta có trường hợp cụ thể sau: Giả sử $A = (a_{ij})_{n\times n}$ có cột $j$ thỏa $a_{ij} = a'_{ij} + a''_{ij}$ với $i = 1, \dots, n$, định thức ma trận $A$ có thể tính bởi

$$|A| = \begin{vmatrix} \dots & a'_{1j} + a''_{1j} & \dots \\ \dots & a'_{2j} + a''_{2j} & \dots \\ \dots & \dots & \dots \\ \dots & a'_{nj} + a''_{nj} & \dots \end{vmatrix} = \begin{vmatrix} \dots & a'_{1j} & \dots \\ \dots & a'_{2j} & \dots \\ \dots & \dots & \dots \\ \dots & a'_{nj} & \dots \end{vmatrix} + \begin{vmatrix} \dots & a''_{1j} & \dots \\ \dots & a''_{2j} & \dots \\ \dots & \dots & \dots \\ \dots & a''_{nj} & \dots \end{vmatrix}$$

$$(1) + (2)$$

Khi đó, định thức thứ nhất sau khi phân tích được gọi là định thức có cột thứ $j$ loại (1) và định thức thứ hai là định thức có cột thứ $j$ loại (2).

Phương pháp này rất hữu dụng khi ta tách được nhiều định thức có hai cột tỉ lệ (suy ra có giá trị định thức bằng 0) và các định thức còn lại đơn giản và dễ tính. Sau đây chúng ta xét một số ví dụ minh họa phương pháp này.

**6.4.1 Thí dụ.** Tính định thức sau với $n \geq 1$

$$D = \begin{vmatrix} 2 & 2 & 3 & \dots & n \\ 1 & 3 & 3 & \dots & n \\ 1 & 2 & 4 & \dots & n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n+1 \end{vmatrix}$$

*Giải*. Mỗi cột của $D$ được viết thành tổng của hai cột mà ta kí hiệu cột loại (1) và loại (2) như sau:

$$\begin{array}{c} D = \begin{vmatrix} 1+1 & 0+2 & 0+3 & \dots & 0+n \\ 0+1 & 1+2 & 0+3 & \dots & 0+n \\ 0+1 & 0+2 & 1+3 & \dots & 0+n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0+1 & 0+2 & 0+3 & \dots & 1+n \end{vmatrix} \\ \quad (1)+(2) \quad (1)+(2) \quad (1)+(2) \quad \dots \quad (1)+(2) \end{array}$$

Ta lần lượt tách các cột của định thức, sau khi tách tất cả $n$ cột ta có định thức $D$ bằng tổng của $2^n$ định thức cấp $n$. Cột thứ $i$ của các định thức này chính là cột loại (1) hoặc loại (2) của cột thứ $i$ của định thức ban đầu $D$. Ta chia $2^n$ định thức thành ba dạng:

Dạng 1: Bao gồm các định thức có từ hai cột loại (2) trở lên. Vì các cột loại (2) tỉ lệ với nhau nên tất cả các định thức này có giá trị bằng 0.

Dạng 2: Bao gồm các định thức có đúng một cột loại (2), còn các cột khác là loại (1). Ta có $n$ định thức cấp $n$ dạng 2 như sau:

$$C_1 = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 1 \end{vmatrix} = 1$$

$$C_2 = \begin{vmatrix} 1 & 2 & 0 & \dots & 0 \\ 0 & 2 & 0 & \dots & 0 \\ 0 & 2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 2 & 0 & \dots & 1 \end{vmatrix} = 2$$

Tương tự ta tính được các định thức như vậy cho đến $C_n$

$$C_n = \begin{vmatrix} 1 & 0 & 0 & \dots & n \\ 0 & 1 & 0 & \dots & n \\ 0 & 0 & 1 & \dots & n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & n \end{vmatrix} = n$$

Vậy tổng các định thức dạng 2 là:

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

Dạng 3: Bao gồm các định thức không có cột loại (2) nên tất cả các cột đều là loại (1) và do đó có đúng một định thức dạng 3 là:

$$\begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix} = 1.$$

Vậy $D$ bằng tổng tất cả các định thức ở ba dạng trên hay

$$D = 1 + \frac{n(n+1)}{2} = \frac{n^2+n+2}{2} \qquad (n \geq 1) \qquad \square.$$

**6.4.2 Thí dụ. (OSV03)** Cho ma trận

$$A = \begin{vmatrix} 1+x_1 & 1 & 1 & 1 \\ 1 & 1+x_2 & 1 & 1 \\ 1 & 1 & 1+x_3 & 1 \\ 1 & 1 & 1 & 1+x_4 \end{vmatrix}$$

trong đó $x_1, x_2, x_3, x_4$ là các nghiệm của đa thức $f(x) = x^4 - x + 1$. Tính $\det A$.

*Giải*. Ta có

$$\det A = \begin{vmatrix} 1+x_1 & 1 & 1 & 1 \\ 1 & 1+x_2 & 1 & 1 \\ 1 & 1 & 1+x_3 & 1 \\ 1 & 1 & 1 & 1+x_4 \end{vmatrix}$$

Mỗi cột của $\det A$ được viết thành tổng của hai cột mà ta kí hiệu cột loại (1) và loại (2) như sau:

$$\det A = \begin{array}{c} \begin{vmatrix} x_1+1 & 0+1 & 0+1 & 0+1 \\ 0+1 & x_2+1 & 0+1 & 0+1 \\ 0+1 & 0+1 & x_3+1 & 0+1 \\ 0+1 & 0+1 & 0+1 & x_4+1 \end{vmatrix} \\ \begin{matrix} (1)+(2) & (1)+(2) & (1)+(2) & (1)+(2) \end{matrix} \end{array}$$

Ta lần lượt tách các cột của định thức, sau khi tách tất cả 4 cột ta có định thức $A$ bằng tổng của 16 định thức cấp 4. Cột thứ $i$ của các định thức này chính là cột loại (1) hoặc loại (2) của cột thứ $i$ của định thức ban đầu $\det A$. Ta chia 16 định thức thành ba dạng:

Dạng 1: Bao gồm các định thức có từ hai cột loại (2) trở lên. Vì các cột loại (2) tỉ lệ với nhau nên tất cả các định thức này có giá trị bằng 0.

Dạng 2: Bao gồm các định thức có đúng một cột loại (2), còn các cột khác là loại (1). Ta có 4 định thức cấp 4 dạng 2 như sau:

$$C_1 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & x_2 & 0 & 0 \\ 1 & 0 & x_3 & 0 \\ 1 & 0 & 0 & x_4 \end{vmatrix} = x_2x_3x_4$$

$$C_2 = \begin{vmatrix} x_1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & x_3 & 0 \\ 0 & 1 & 0 & x_4 \end{vmatrix} = x_1x_3x_4$$

$$C_3 = \begin{vmatrix} x_1 & 0 & 1 & 0 \\ 0 & x_2 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & x_4 \end{vmatrix} = x_1x_2x_4$$

$$C_4 = \begin{vmatrix} x_1 & 0 & 0 & 1 \\ 0 & x_2 & 0 & 1 \\ 0 & 0 & x_3 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = x_1x_2x_3$$

Vậy tổng các định thức dạng 2 là:

$$x_2x_3x_4 + x_1x_3x_4 + x_1x_2x_4 + x_1x_2x_3$$

Dạng 3: Bao gồm các định thức không có cột loại (2) nên tất cả các cột đều là loại (1) và do đó có đúng một định thức dạng 3 là:

$$\begin{vmatrix} x_1 & 0 & 0 & 0 \\ 0 & x_2 & 0 & 0 \\ 0 & 0 & x_3 & 0 \\ 0 & 0 & 0 & x_4 \end{vmatrix} = x_1x_2x_3x_4$$

Vậy $D$ bằng tổng tất cả các định thức ở ba dạng trên hay

$$\det A = x_2x_3x_4 + x_1x_3x_4 + x_1x_2x_4 + x_1x_2x_3 + x_1x_2x_3x_4.$$

do $x_1, x_2, x_3, x_4$ là các nghiệm của đa thức $f(x) = x^4 - x + 1$. Áp dụng định lí Vi-et ta có $\det A = 2$. □

## 6.5 Định thức của tích các ma trận

**6.5.1 Định lý.** *Cho $A = (a_{ij})_{n\times n}$ và $B = (b_{ij})_{n\times n}$ là hai ma trận vuông cấp $n$ trên $\mathbb{K}$. Khi đó,* $\det(AB) = \det A \cdot \det B$.

*Chứng minh.* Xét định thức cấp $2n$ sau

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & -1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{vmatrix} = \det \begin{bmatrix} A & O_n \\ -I_n & B \end{bmatrix}.$$

Áp dụng định lý Laplace khai triển định thức $D$ theo $n$ dòng đầu ta có ngay $D = \det A \cdot \det B$. Bây giờ ta tính $D$ theo cách khác: Nhân cột thứ nhất, thứ hai, ..., thứ $n$ lần lượt với $b_{1j}, b_{2j}, \dots, b_{nj}$ rồi cộng vào cột thứ $n+j$ với $j = 1, \dots, n$, ta được

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & \sum_{k=1}^n a_{1k}b_{k1} & \sum_{k=1}^n a_{1k}b_{k2} & \dots & \sum_{k=1}^n a_{1k}b_{kn} \\ a_{21} & a_{22} & \dots & a_{2n} & \sum_{k=1}^n a_{2k}b_{k1} & \sum_{k=1}^n a_{2k}b_{k2} & \dots & \sum_{k=1}^n a_{2k}b_{kn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & \sum_{k=1}^n a_{nk}b_{k1} & \sum_{k=1}^n a_{nk}b_{k2} & \dots & \sum_{k=1}^n a_{nk}b_{kn} \\ -1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & 0 & 0 & \dots & 0 \end{vmatrix}$$

$$= \det \begin{bmatrix} A & AB \\ -I_n & O_n \end{bmatrix}.$$

Đặt $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$, hay $AB = (c_{ij})_{n\times n}$. Áp dụng định lý Laplace khai triển định thức trên theo $n$ dòng cuối ta được

$$D = (-1)^{(n+1)+\dots+2n+1+\dots+n} \begin{vmatrix} -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 \end{vmatrix} \cdot \begin{vmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{vmatrix}$$
$$= (-1)^{n(2n+1)}(-1)^n \det(AB)$$
$$= \det(AB)$$

Vậy ta có $\det(AB) = \det A \cdot \det B$. □

Trong một số trường hợp dùng định lý trên có thể tính được một số định thức (mà nếu biến đổi sơ cấp thì khá phức tạp) bằng cách tách thành tích của hai định thức đơn giản hơn.

**6.5.2 Thí dụ.** Tính định thức của ma trận vuông $A$ cấp $n > 2$

$$A = \begin{bmatrix} 1+x_1y_1 & 1+x_1y_2 & \dots & 1+x_1y_n \\ 1+x_2y_1 & 1+x_2y_2 & \dots & 1+x_2y_n \\ \dots & \dots & \dots & \dots \\ 1+x_ny_1 & 1+x_ny_2 & \dots & 1+x_ny_n \end{bmatrix}$$

Ta nhận thấy

$$A = \begin{bmatrix} 1 & x_1 & 0 & \dots & 0 \\ 1 & x_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ y_1 & y_2 & \dots & y_n \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

Theo định lý định thức của tích các ma trận ta có ngay $\det A = 0$. □

**6.5.3 Thí dụ. (OSV95)** Cho $A$ là ma trận vuông cấp $n$ sao cho $A^{-1} = 3A$. Tính $\det(A^{1995} - A)$.

*Giải.* Từ giả thiết ta có $A^2 = \frac{1}{3}I_n$, cho nên

$$|A|^2 = \frac{1}{3^n} \qquad\qquad A^{1995} = \frac{1}{3^{997}}A.$$

Từ đó ta tính được

$$\det(A^{1995} - A) = \det\left[\left(\frac{1}{3^{997}} - 1\right)A\right] = \left(\frac{1-3^{997}}{3^{997}}\right)^n |A| = \left(\frac{1-3^{997}}{3^{997}\sqrt{3}}\right)^n.$$

□

**6.5.4 Thí dụ. (OSV95)** Cho $B$ là ma trận vuông cấp $n$ và $\alpha$ là một số thực thỏa mãn điều kiện $\det(B - \alpha I_n) = 0$. Chứng minh rằng với mọi $a_0, a_1, \dots, a_n \in \mathbb{R}$, $n \in \mathbb{N}$, ta đều có

$$\det\left(\sum_{k=0}^{n} a_k B^k - \sum_{k=0}^{n} a_k \alpha^k I_n\right) = 0.$$

*Giải.* Với mỗi $k$ ta có

$$B^k - \alpha^k I_n = (B - \alpha I_n)(B^{k-1} + \alpha B^{k-2} + \cdots + \alpha^{k-1} I_n) = (B - \alpha I_n) M_k.$$

Do đó, ta biến đổi được

$$\begin{aligned}
\sum_{k=0}^{n} a_k B^k - \sum_{k=0}^{n} a_k \alpha^k I_n &= \sum_{k=0}^{n} a_k (B^k - \alpha^k I_n) \\
&= \sum_{k=0}^{n} a_k (B - \alpha I_n) M_k \\
&= (B - \alpha I_n) M.
\end{aligned}$$

Vậy $\det \left( \sum_{k=0}^{n} a_k B^k - \sum_{k=0}^{n} a_k \alpha^k I_n \right) = |B - \alpha I_n| \cdot |M| = 0.$ □

**6.5.5 Thí dụ. (OSV01)** Cho các ma trận vuông thực $A$, $B$ thỏa mãn điều kiện sau $A^{2001} = O$, $AB = A + B$. Chứng minh rằng $\det B = 0$.

*Giải.* Từ định thức của tích các ma trận và giả thiết $A^{2001} = O$ ta suy ra $\det A = 0$. Mặt khác, từ giả thiết $AB = A + B$ ta có $B = A(B - I)$. Do đó $\det(B) = \det(A(B - I)) = \det(A)\det(B - I) = 0$. □

**6.5.6 Thí dụ. (OSV04)** Biết rằng các ma trận vuông $A$, $B$ đều là nghiệm của đa thức $f(x) = x^2 - x$ và $AB + BA = O$. Tính $\det(A - B)$.

*Giải.* Từ đề bài ta có $A^2 = A$ và $B^2 = B$. Từ đó ta tính

$$\begin{aligned}
(A + B)^2 &= A^2 + AB + BA + B^2 = A + B \\
(A - B)^2 &= A^2 - AB - BA + B^2 = A + B
\end{aligned}$$

Đặt $\det(A + B) = \alpha$, $\det(A - B) = \beta$. Từ kết quả trên và tính chất định thức của tích hai ma trận ta có

$$\begin{cases} \alpha^2 = \alpha \\ \beta^2 = \alpha \end{cases} \quad \text{suy ra} \quad \begin{cases} \alpha = 0 \\ \beta = 0 \end{cases} \quad \begin{cases} \alpha = 1 \\ \beta = 1 \end{cases} \quad \begin{cases} \alpha = 1 \\ \beta = -1 \end{cases}$$

Ta xét từng trường hợp xem có các ma trận $A$, $B$ nào thỏa hay không

- $\alpha = 0$, $\beta = 0$, ta có thể lấy $A = O$, $B = O$
- $\alpha = 1$, $\beta = 1$, ta có thể lấy $A = I$, $B = 0$.
- $\alpha = 1$, $\beta = -1$, ta có thể lấy $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

Vậy $\det(A - B)$ có thể là 0, 1, $-1$. □

## Bài tập

**1**) Tính các định thức sau:

**(a)** $\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 3 & 4 & 5 \\ 1 & 1 & 1 & 4 & 5 \\ 1 & 1 & 1 & 1 & 5 \end{vmatrix}$ **(b)** $\begin{vmatrix} 1 & 2 & 3 & \dots & n \\ -1 & 0 & 3 & \dots & n \\ -1 & -2 & 0 & \dots & n \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -2 & -3 & \dots & 0 \end{vmatrix}$ **(c)** $\begin{vmatrix} 2 & 3 & 0 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 \\ 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \end{vmatrix}$

**2**) Tính các định thức sau

**(a)** $\begin{vmatrix} a & x & x & \dots & x & x \\ x & a & x & \dots & x & x \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x & x & x & \dots & a & x \\ x & x & x & \dots & x & a \end{vmatrix}$ **(b)** $\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 0 \end{vmatrix}$

**(c)** $\begin{vmatrix} 1 & a_1 & a_2 & \dots & a_n \\ 1 & a_1 + b_1 & a_2 & \dots & a_n \\ 1 & a_1 & a_2 + b_2 & \dots & a_n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_1 & a_2 & \dots & a_n + b_n \end{vmatrix}$

**(d)** $\begin{vmatrix} 1 & a_1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 - a_1 & a_2 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 - a_2 & a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 - a_{n-1} & a_n \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 - a_n \end{vmatrix}$

**(e)** $\begin{vmatrix} 1 & 2 & 3 & 4 & \dots & n-1 & n \\ -1 & x & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & x & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & x & 0 \\ 0 & 0 & 0 & 0 & \dots & -1 & x \end{vmatrix}$

**(f)** $\begin{vmatrix} x & y & 0 & \dots & 0 & 0 \\ 0 & x & y & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x & y \\ 0 & 0 & 0 & \dots & 0 & x \end{vmatrix}$

**(g)** $\begin{vmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ 1 & 3 & 3 & \dots & n-1 & n \\ 1 & 2 & 5 & \dots & n-1 & n \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 2 & 3 & \dots & 2n-3 & n \\ 1 & 2 & 3 & \dots & n-1 & 2n-1 \end{vmatrix}$

**(h)** $\begin{vmatrix} 0 & 1 & 1 & \dots & 1 & 1 \\ 1 & 0 & x & \dots & x & x \\ 1 & x & 0 & \dots & x & x \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x & x & \dots & 0 & x \\ 1 & x & x & \dots & x & 0 \end{vmatrix}$

**(i)** $\begin{vmatrix} 1 & x_1 & x_2 & \dots & x_{n-1} & x_n \\ 1 & x & x_2 & \dots & x_{n-1} & x_n \\ 1 & x_1 & x & \dots & x_{n-1} & x_n \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_1 & x_2 & \dots & x & x_n \\ 1 & x_1 & x_2 & \dots & x_{n-1} & x \end{vmatrix}$

**(j)** $\begin{vmatrix} a_1 & x & x & \dots & x \\ x & a_2 & x & \dots & x \\ x & x & a_3 & \dots & x \\ \dots & \dots & \dots & \dots & \dots \\ x & x & x & \dots & a_n \end{vmatrix}$

**(k)** $\begin{vmatrix} x & a_1 & a_2 & \dots & a_n \\ a_1 & x & a_2 & \dots & a_n \\ a_1 & a_2 & x & \dots & a_n \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_2 & a_3 & \dots & x \end{vmatrix}$

**3)** Tính các định thức sau

**(a)** $\begin{vmatrix} 1-\lambda & 3 & 2 \\ 2 & 1-\lambda & 3 \\ 3 & 2 & 1-\lambda \end{vmatrix}$

**(b)** $\begin{vmatrix} 1-\lambda & 2 & 1 \\ 2 & 3-\lambda & 4 \\ 3 & 4 & 5-\lambda \end{vmatrix}$

**(c)** $\begin{vmatrix} 2-\lambda & 3 & 4 \\ 3 & 4-\lambda & 5 \\ 4 & 5 & 6-\lambda \end{vmatrix}$

**(d)** $\begin{vmatrix} 2-\lambda & 1 & 0 \\ 0 & 3-\lambda & 2 \\ 1 & 4 & -2-\lambda \end{vmatrix}$

**(e)** $\begin{vmatrix} 3-\lambda & 1 & 3 \\ 2 & 2-\lambda & 2 \\ 1 & 3 & 1-\lambda \end{vmatrix}$ **(f)** $\begin{vmatrix} 1-\lambda & 3 & 5 \\ 5 & 3-\lambda & 1 \\ 1 & 5 & 3-\lambda \end{vmatrix}$

**4 )** Tính định thức $\begin{vmatrix} \lambda & 0 & \dots & 0 & a_1 \\ 1 & 0 & \dots & 0 & a_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a_n \end{vmatrix}$.

**5 )** Giải các phương trình sau theo ẩn $x$ trên $\mathbb{R}$

**(a)** $\begin{vmatrix} x^2 & x^3 & x^4 \\ 0 & x^2-1 & 0 \\ 0 & x^3+1 & x^3-1 \end{vmatrix} = 0$ **(b)** $\begin{vmatrix} 1 & x & x^2 & x^3 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{vmatrix} = 0$

**(c)** $\begin{vmatrix} 1 & x & x-1 & x+2 \\ 0 & 0 & x^2-1 & 0 \\ x & 1 & x & x-2 \\ 0 & 0 & x^5+1 & x^{100} \end{vmatrix} = 0$ **(d)** $\begin{vmatrix} x^{10} & x^{20} & x^{30} & x^{40} \\ 0 & 0 & x^{50} & 0 \\ x^{50} & x^{60} & x^{70} & 0 \\ x^2-1 & 0 & x^{90} & 0 \end{vmatrix} = 0$

**(e)** $\begin{vmatrix} x^{100} & 0 & 0 & 0 \\ x^{200} & x^2-1 & x^4 & x^6 \\ x^{300} & 0 & x^3+1 & 0 \\ x^{400} & 0 & x^{500} & x+2 \end{vmatrix} = 0$

**(f)** $\begin{vmatrix} 1 & x & x^2 & \dots & x^{n-1} \\ 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_{n-1} & a_{n-1}^2 & \dots & a_{n-1}^{n-1} \end{vmatrix} = 0$, các hằng số đôi một khác nhau.

**(g)** $\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1-x & 1 & \dots & 1 \\ 1 & 1 & 2-x & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & (n-1)-x \end{vmatrix} = 0$

**(h)** $$\begin{vmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ a_1 & a_1+a_2-x & a_3 & \dots & a_n \\ a_1 & a_2 & a_2+a_3-x & \dots & a_n \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_2 & a_3 & \dots & a_{n-1}+a_n-x \end{vmatrix} = 0.$$

**6 ) (OSV05)** Xét ma trận dạng $A = \begin{bmatrix} x_1^2+1 & x_1x_2 & x_1x_3 & x_1x_4 \\ x_1x_2 & x_2^2+1 & x_2x_3 & x_2x_4 \\ x_1x_3 & x_2x_3 & x_3^2+1 & x_3x_4 \\ x_1x_4 & x_2x_4 & x_3x_4 & x_4^2+1 \end{bmatrix}$

Chứng minh rằng định thức của $A$ là một đa thức đối xứng theo các biến $x_1, x_2, x_3, x_4$. Tính định thức của $A$ khi $x_1, x_2, x_3, x_4$ lần lượt là bộ 4 nghiệm của đa thức $P_4(x) = x^4 - x^3 - 5x^2 + 1$.

**7 )** Cho $A \in \mathcal{M}_n(\mathbb{K})$ với $n \geq 2$. Chứng minh rằng

**(a)** Nếu $A$ là ma trận phản xứng và $n$ lẻ thì $\det A = 0$.

**(b)** Nếu $\mathbb{K} = \mathbb{C}$ và giả sử $A = (a_{ij})_n$ mà $a_{ij} = \overline{a}_{ji}$ với mọi $i, j = 1, \dots, n$ thì $\det A \in \mathbb{R}$.

**(c)** Nếu $B \in \mathcal{M}_m(\mathbb{K})$ và $C = \left[\begin{array}{c|c} A & O \\ \hline O & B \end{array}\right]$ (ma trận chia khối) thì $\det C = \det A \cdot \det B$.

# § 7 Hạng của ma trận

## 7.1 Khái niệm và các tính chất cơ bản

**7.1.1 Định nghĩa.** Cho $A$ là một ma trận cấp $m \times n$ (khác ma trận không), số tự nhiên $r > 0$ gọi **là hạng của ma trận** (rank of matrix) $A$, kí hiệu là $\operatorname{rank} A = r$ hoặc $r(A) = r$, nếu tồn tại một định thức con cấp $r$ khác 0 và mọi định thức con cấp $r + 1$ (nếu có) đều bằng 0. Ta quy ước hạng của ma trận không là 0.

Như vậy hạng của ma trận khác không chính là *cấp cao nhất của các định thức con khác không* của nó.

**7.1.2** Thuật toán tìm hạng của ma trận $A$, $\operatorname{rank} A$, như sau:

- Tìm một định thức con cấp $k$ (kí hiệu $D_k$) khác 0.
- Xét tất cả các định thức con $D_{k+1}$ cấp $k+1$ chứa $D_k$
  - + Nếu tất cả $D_{k+1}$ đều bằng 0 thì $\operatorname{rank} A = k$.
  - + Nếu tồn tại một $D_{k+1}$ khác 0 thì ta lặp lại quá trình trên.

Thuật toán trên cho ta hạng của ma trận là do Định lý 2.1.10 trang 212, Mệnh đề 3.2.3 trang 223, Định lý 3.3.1 trang 225 cùng với phép chứng minh của chúng.

**7.1.3 Thí dụ.** Tìm hạng của ma trận $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & 3 & 0 & 1 \\ 2 & 4 & 1 & 8 \\ 1 & 7 & 6 & 9 \\ 0 & 10 & 1 & 10 \end{bmatrix}$

Ta thấy ma trận $A$ có một định thức con cấp 2 khác không là $D_2 = \begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} \neq 0$. Ta thấy có định thức con cấp 3 chứa $D_2$ khác không là $D_3 = \begin{vmatrix} 1 & 2 & 3 \\ -1 & 3 & 0 \\ 2 & 4 & 1 \end{vmatrix} = -25 \neq 0$. Các định thức con cấp 4 chứa $D_3$ đều bằng 0:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & 3 & 0 & 1 \\ 2 & 4 & 1 & 8 \\ 1 & 7 & 6 & 9 \end{vmatrix} = 0 \qquad \begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & 3 & 0 & 1 \\ 2 & 4 & 1 & 8 \\ 0 & 10 & 1 & 10 \end{vmatrix} = 0.$$

Vậy $\operatorname{rank} A = 3$. □

**7.1.4 Mệnh đề.** *Cho* $A \in \mathcal{M}_{m,n}(\mathbb{K})$. *Khi đó,*

**(i)** $0 \leq \operatorname{rank} A \leq \min\{m, n\}$. *Nếu* $\operatorname{rank} A = \min\{m; n\}$ *thì ta nói* $A$ *có* ***hạng cực đại****.*

**(ii)** *Nếu* $A$ *có một định thức con khác không cấp* $r$*, thì* $\operatorname{rank} A \geq r$.

**(iii)** *Nếu các định thức con cấp k của A đều bằng không, thì* $\operatorname{rank} A < k$.

*Chứng minh.* Các tính chất trên được suy trực tiếp từ định nghĩa và dễ dàng thấy được. □

**7.1.5 Nhận xét.** Nếu $A$ là một ma trận vuông cấp $n$, thì từ hai tính chất trên ta có thể thấy $\operatorname{rank} A = n$ khi và chỉ khi $\det A \neq 0$.

**7.1.6 Định nghĩa.** Ma trận vuông $A$ cấp $n$ được gọi là **ma trận không suy biến** (nonsingular matrix) nếu $\operatorname{rank} A = n$ hay $\det A \neq 0$.

Ma trận vuông $B$ không là ma trận không suy biến được gọi là **ma trận suy biến** (singular matrix). Khi đó $\det B = 0$.

**7.1.7 Thí dụ.** Xét hai ma trận $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 4 \\ 3 & 0 & 5 \end{bmatrix}$ và $B = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 11 \end{bmatrix}$.

Ta có $\det A = (-1)^{1+1}(1) \begin{vmatrix} 3 & 4 \\ 0 & 5 \end{vmatrix} = 15 \neq 0$. Vậy $A$ là ma trận không suy biến. Ta thấy trong ma trận $B$ dòng thứ 3 bằng tổng của hai dòng đầu, cho nên $\det B = 0$. Vậy $B$ là ma trận suy biến. □

**7.1.8 Thí dụ.** Các ma trận sơ cấp là các ma trận không suy biến. □

**7.1.9 Thí dụ. (OSV00)** Cho ma trận $A$ vuông cấp $n$ có các phần tử nằm trên đường chéo chính bằng 0, các phần tử còn lại bằng 1 hoặc bằng 2000. Chứng minh rằng hạng của $A$ hoặc bằng $n$ hoặc bằng $n-1$.

*Giải.* Vì các phần tử ngoài đường chéo chính của ma trận $A$ bằng 1 hoặc 2000 cho nên ta tính định thức của $A$ theo modulo 1999. Ta có

$$\begin{aligned} \det(A) &\equiv \begin{vmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 0 \end{vmatrix} \pmod{1999} \\ &\equiv (n-1)(-1)^{n-1} \pmod{1999} \end{aligned}$$

Khi 1999 không chia hết $n-1$, ta suy ra được $\det(A) \neq 0$ nên $\operatorname{rank}(A) = n$. Khi 1999 chia hết $n-1$ thì 1999 không chia hết $n-2$; do đó theo kết quả trên trong ma trận $A$ có ma trận cấp $n-1$ (bỏ dòng cuối và cột cuối) có định thức khác không, cho nên $\operatorname{rank}(A) \geq n-1$. Vì vậy ta luôn có $\operatorname{rank}(A)$ bằng $n$ hoặc bằng $n-1$. $\square$

**7.1.10 Mệnh đề.** *Hạng của ma trận bậc thang theo dòng chính bằng số dòng khác 0.*

*Chứng minh.* Giả sử ma trận bậc thang $A$ có $r$ dòng khác không. Với bất cứ ma trận con vuông cấp lớn hơn $r$ nào của $A$ đều có chứa ít nhất một dòng bằng không nên định thức của nó sẽ bằng không. Do đó, $\operatorname{rank} A \leq r$. Ta lập một ma trận con vuông cấp $r$ như sau: chọn $r$ dòng khác không ($r$ dòng đầu tiên), ứng với mỗi dòng khác không ta chọn cột ứng với số hạng đầu tiên khác không trong dòng. Như vậy ta sẽ được một ma trận tam giác có các số hạng trên đường chéo khác không nên định thức của nó sẽ khác không. Do đó, $\operatorname{rank} A \geq r$. Vậy $\operatorname{rank} A = r$. $\square$

**7.1.11 Thí dụ.** Ma trận $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 & 10 \end{bmatrix}$ có hạng bằng hai do nó có định thức con cấp 2 $D_{12,12} = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1 \neq 0$ và các định thức con cấp 3 đều bằng không vì luôn có hàng thứ 3 gấp đôi hàng thứ nhất.

Ma trận $B = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 0 & 4 & 5 & 6 \\ 3 & 0 & 0 & 6 & 7 \end{bmatrix}$ có hạng bằng 3 vì có một định thức con cấp 3 bằng không, đó là $D_{123,245} = \begin{vmatrix} 2 & 3 & 4 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} = 2 \cdot 4 \cdot 6 \neq 0$.

Ma trận $C = \begin{bmatrix} 0 & 3 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ có hạng bằng 2 vì $C$ là một ma trận bậc thang với hai dòng khác không. $\square$

## 7.2 Tính chất hạng của ma trận

**7.2.1 Định lý.** *Cho $A$ là một ma trận cấp $m \times n$ trên $\mathbb{K}$. Khi đó* $\operatorname{rank} A^t = \operatorname{rank} A$.

*Chứng minh.* Nếu $A = O_{m,n}$ thì $A^t = O_{n,m}$ và $\operatorname{rank} A^t = 0 = \operatorname{rank} A$. Giả sử $A \neq O$ và $\operatorname{rank} A = \min\{m, n\}$. Lúc đó, trong $A$ có một ma trận con $S(r)$ sao cho $\det(S(r)) \neq 0$. Khi đó $S(r)^t$ là một ma trận con của $A^t$. Theo Định lý 4.3.1 ta có

$$\det(S(r)^t) = \det(S(r)) \neq 0.$$

Như vậy $\operatorname{rank} A^t \geq r = \operatorname{rank} A$. Tương tự ta chứng minh được $\operatorname{rank}(A^t)^t \geq \operatorname{rank} A^t$. Do $(A^t)^t = A$, nên $\operatorname{rank} A^t = \operatorname{rank} A$. □

**7.2.2 Mệnh đề.** *Khi ta nhân bên trái một ma trận với ma trận sơ cấp, thì ma trận thu được có hạng bằng với hạng ma trận ban đầu.*

*Chứng minh.* Giả sử $A$ là ma trận cấp $m \times n$. Gọi $E_{I;i,j}$ là ma trận sơ cấp loại $I$ cấp $m$. Khi đó, ma trận $E_{I;i,j}A$ là ma trận thu được từ ma trận $A$ bằng cách đổi chỗ hai dòng $i$ và $j$. Ta biết rằng trong khi tính định thức của một ma trận nếu ta đổi chỗ hai dòng thì định thức đổi dấu; do đó, tính khác $0$ không thay đổi. Vì vậy nếu trong ma trận $A$ ta có một định thức con khác không thì từ ma trận $E_{I;i,j}A$ ta có thể chọn được một định thức con tương ứng cũng khác không và ngược lại. Do đó, $\operatorname{rank} A = \operatorname{rank}(E_{I;i,j}A)$.

Với $E_{II;i,\lambda}$ là ma trận sơ cấp loại II cấp $m$, ta có $E_{II;i,\lambda}A$ là ma trận thu được từ ma trận $A$ bằng cách nhân $\lambda$ vào dòng thứ $i$. Do $\lambda \neq 0$ và định thức của một ma trận thu được khi ta nhân một hàng nào đó của một ma trận vuông với một số $\lambda$ chính bằng $\lambda$ nhân với định thức của ma trận ban đầu. Nghĩa là tính khác $0$ của định thức không thay đổi khi ta nhân một dòng nào đó với một số khác $0$. Từ đó ta suy ra được hạng của ma trận $E_{II;i,\lambda}A$ lớn hơn hoặc bằng hạng của ma trận $A$, nghĩa là $\operatorname{rank}(E_{II;i,\lambda}A) \geq \operatorname{rank} A$. Tương tự ta có $\operatorname{rank}(E_{II;i,\frac{1}{\lambda}}E_{II;i,\lambda}A) \geq \operatorname{rank}(E_{II;i,\lambda}A)$. Do $E_{II;i,\frac{1}{\lambda}}E_{II;i,\lambda}A = A$, nên $\operatorname{rank}(A) = \operatorname{rank}(E_{II;i,\lambda}A)$.

Với $E_{III;i,j,\lambda}$ là ma trận sơ cấp loại III có cấp $m$, ta có $E_{III;i,j,\lambda}A$ là ma trận thu được từ $A$ bằng cách nhân dòng $j$ cho $\lambda$ rồi cộng vào dòng $i$. Nếu $\operatorname{rank}(A) = r = \min\{m, n\}$, thì trong $A$ tồn tại một ma trận con vuông cấp $r$ có định thức khác không. Với các dòng và các cột tương ứng ta có được một ma trận con vuông của $E_{III;i,j,\lambda}A$ có định thức bằng với định thức của ma trận con của $A$ đang xét (bởi vì giá trị của định thức không thay đổi khi ta nhân một hằng số cho một dòng nào đó rồi cột vào một dòng khác). Do đó $\operatorname{rank}(E_{III;i,j,\lambda}A) \geq r = \min\{m,n\}$. Vậy $\operatorname{rank}(E_{III;i,j,\lambda}A) = \operatorname{rank}(A)$. Xét trường hợp $\operatorname{rank}(A) = r < \min\{m,n\}$. Ta có các định thức con cấp $r+1$ của ma trận $E_{III;i,j,\lambda}A$ có thể có ba trường hợp sau:

- Định thức con ấy không chứa phần tử dòng $i$. Khi đó, định thức này cũng là định thức con cấp $r+1$ của $A$ cho nên nó bằng $0$.

- Định thức con ấy có chứa phần tử dòng $i$ và phần tử dòng $j$. Khi đó, định thức này đúng bằng một định thức con cấp $r+1$ của $A$ (đó là dòng trong định thức tương ứng với dòng $i$ trừ đi $\lambda$ nhân với phần tử tương ứng ở dòng $j$ trong ma trận ban đầu), cho nên định thức ấy bằng $0$.

- Định thức con ấy có chứa phần tử dòng $i$ nhưng không chứa phần tử dòng $j$. Nó sẽ bằng một định thức con cấp $r+1$ của $A$ cộng với $\lambda$ hoặc $-\lambda$ nhân với một định thức con cấp $r+1$ khác của $A$. Cho nên định thức đang xét phải bằng không.

Vậy mọi định thức con cấp $r+1$ của $E_{III;i,j,\lambda}A$ đều bằng không. Do đó, $\operatorname{rank}(E_{III;i,j,\lambda}A) \leq r = \operatorname{rank}(A)$. Với lập luận trên ta suy ra $\operatorname{rank}(A) = \operatorname{rank}(E_{III;i,j,-\lambda}E_{III;i,j,\lambda}A) \leq \operatorname{rank}(E_{III;i,j,\lambda}A)$. Do đó, ta cũng được $\operatorname{rank}(A) = \operatorname{rank}(E_{III;i,j,\lambda}A)$. $\square$

**7.2.3 Định lý.** *Cho $A$ là một ma trận cấp $m \times n$ trên $\mathbb{K}$ và $B$ là ma trận nhận được từ $A$ bằng một số hữu hạn các phép biến đổi sơ cấp. Khi đó,* $\operatorname{rank} B = \operatorname{rank} A$. *Nói cách khác, hạng của ma trận không đổi qua các phép biến đổi sơ cấp.*

*Chứng minh.* Theo mệnh đề trên, hạng của một ma trận không thay đổi khi ta nhân vào bên trái của ma trận với một ma trận sơ cấp. Nói cách khác hạng của ma trận không thay đổi khi ta thực hiện một phép biến đổi sơ cấp. Do $B$ là ma trận nhận được từ $A$ bằng một số hữu hạn các phép biến đổi sơ cấp, nên hạng của chúng phải bằng nhau; nghĩa là $\operatorname{rank} B = \operatorname{rank} A$. □

**7.2.4 Thí dụ. (OSV02)** Cho $B$ là ma trận thực, vuông cấp $n$ có hạng bằng 1. Chứng minh rằng tồn tại duy nhất số thực $\lambda$ sao cho $B^2 = \lambda B$.

*Giải.* Do $\operatorname{rank} B = 1$ nên các dòng của $B$ tỷ lệ nhau và có ít nhất một dòng khác không. Vì vậy tồn tại $a_1, a_2, \ldots, a_n$ không đồng thời bằng không $\lambda_1, \lambda_2, \ldots, \lambda_n$ không đồng thời bằng không sao cho ma trận $B$ có dạng

$$B = \begin{bmatrix} \lambda_1 a_1 & \lambda_1 a_2 & \ldots & \lambda_1 a_n \\ \lambda_2 a_1 & \lambda_2 a_2 & \ldots & \lambda_2 a_n \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_n a_1 & \lambda_n a_2 & \ldots & \lambda_n a_n \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \ldots \\ \lambda_n \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} = \mathbf{u}\mathbf{v}^t$$

Khi đó, ta có $B^2 = \mathbf{u}\mathbf{v}^t\mathbf{u}\mathbf{v}^t = \mathbf{u}(\mathbf{v}^t\mathbf{u})\mathbf{v}^t = \lambda\mathbf{u}\mathbf{v}^t = \lambda B$ với $\lambda = \mathbf{v}^t\mathbf{u} = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n$.

Nếu tồn tại $\lambda'$ thỏa mãn điều kiện $B^2 = \lambda' B$, thì từ

$$O = B^2 - B^2 = (\lambda - \lambda')B \qquad \text{và} \qquad B \neq O$$

suy ra $\lambda' = \lambda$. □

**7.2.5 Nhận xét.** Từ các kết quả của Định lý 7.2.1 và Định lý 7.2.3 ta thấy rằng hạng của ma trận không đổi qua các phép biến đổi sơ cấp theo cột.

**7.2.6 Định lý.** *Cho $A$ là một ma trận khác không cấp $m \times n$ trên $\mathbb{K}$ với $m, n \geq 2$. Khi đó,* $\operatorname{rank}(A) = r$ *khi và chỉ khi dạng bậc thang chính tắc $B$ của $A$ là*

$$B = \left[\begin{array}{c|c} I_r & O_{r,n-r} \\ \hline O_{m-r,r} & O_{m-r,n-r} \end{array}\right]$$

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

Về mặt thực hành để tìm hạng của một ma trận $A$ tùy ý khác không cấp $m \times n$ trên $\mathbb{K}$ với $m, n \geq 2$, trước hết ta dùng các phép biến đổi sơ cấp để đưa $A$ về ma trận bậc thang $B$. Lúc đó hạng của $A$ bằng số dòng khác không của $B$.

**7.2.7 Thí dụ.** Tìm hạng của ma trận $A = \begin{bmatrix} 4 & 3 & -5 & 2 & 3 \\ 8 & 6 & -7 & 4 & 2 \\ 4 & 3 & -8 & 2 & 7 \\ 8 & 6 & -1 & 4 & 6 \end{bmatrix}$.

Ta thực hiện các phép biến đổi sơ cấp trên ma trận $A$

$$A \xrightarrow[\substack{d_3 \to d_3 - d_1 \\ d_4 \to d_4 - 2d_1}]{d_2 \to d_2 - 2d_1} \begin{bmatrix} 4 & 3 & -5 & 2 & 3 \\ 0 & 0 & 3 & 0 & -4 \\ 0 & 0 & -3 & 0 & 4 \\ 0 & 0 & 9 & 0 & 0 \end{bmatrix} \xrightarrow[d_4 \to d_4 - 3d_2]{d_3 \to d_3 + d_2} \begin{bmatrix} 4 & 3 & -5 & 2 & 3 \\ 0 & 0 & 3 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12 \end{bmatrix}$$

$$\xrightarrow{d_3 \leftrightarrow d_4} \begin{bmatrix} 4 & 3 & -5 & 2 & 3 \\ 0 & 0 & 3 & 0 & -4 \\ 0 & 0 & 0 & 0 & 12 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Vậy $\operatorname{rank} A = 3$. □

**7.2.8 Thí dụ.** Tìm hạng của ma trận $A = \begin{bmatrix} 2 & 5 & 7 & 3 & 5 & 8 \\ 1 & 2 & 3 & 1 & 2 & 3 \\ 3 & 7 & 11 & 5 & 8 & 12 \\ 4 & 9 & 14 & 6 & 10 & 15 \\ 5 & 11 & 17 & 7 & 12 & 18 \end{bmatrix}$

Ta thực hiện các phép biến đổi sơ cấp trên các dòng

$$A \xrightarrow{d_1 \leftrightarrow d_2} \begin{bmatrix} 1 & 2 & 3 & 1 & 2 & 3 \\ 2 & 5 & 7 & 3 & 5 & 8 \\ 3 & 7 & 11 & 5 & 8 & 12 \\ 4 & 9 & 14 & 6 & 10 & 15 \\ 5 & 11 & 17 & 7 & 12 & 18 \end{bmatrix}$$

$$\xrightarrow[\substack{d_4 \to d_4 - 4d_1 \\ d_5 \to d_5 - 5d_1}]{\substack{d_2 \to d_2 - 2d_1 \\ d_3 \to d_3 - 3d_1}} \begin{bmatrix} 1 & 2 & 3 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 2 & 2 & 2 & 3 \\ 0 & 1 & 2 & 2 & 2 & 3 \\ 0 & 1 & 2 & 2 & 2 & 3 \end{bmatrix}$$

$$\xrightarrow[d_5 \to d_5 - d_3]{d_4 \to d_4 - d_3} \begin{bmatrix} 1 & 2 & 3 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 2 & 2 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{d_3 \to d_3 - d_2} \begin{bmatrix} 1 & 2 & 3 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Ma trận cuối cùng là ma trận bậc thang theo dòng với ba dòng khác không, cho nên $\operatorname{rank}(A) = 3$. □

## Bài tập

**1)** Tìm hạng của ma trận

**(a)** $\begin{bmatrix} 4 & 3 & -5 & 2 & 3 \\ 8 & 6 & -7 & 4 & 2 \\ 4 & 3 & -8 & 2 & 7 \\ 8 & 6 & -1 & 4 & -6 \end{bmatrix}$

**(b)** $\begin{bmatrix} 2 & 0 & 3 & -1 \\ 1 & -2 & 2 & -3 \\ 3 & -2 & 5 & -4 \\ 5 & -2 & 8 & -5 \end{bmatrix}$

**(c)** $\begin{bmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{bmatrix}$

**(d)** $\begin{bmatrix} 0 & 4 & 10 & 1 \\ 4 & 8 & 18 & 7 \\ 10 & 18 & 40 & 17 \\ 1 & 7 & 17 & 3 \end{bmatrix}$

**(e)** $\begin{bmatrix} 2 & 1 & 11 & 2 \\ 1 & 0 & 4 & -1 \\ 11 & 4 & 56 & 5 \\ 2 & -1 & 5 & -6 \end{bmatrix}$

**(f)** $\begin{bmatrix} 2 & 0 & 3 & -1 \\ 1 & -2 & 2 & -3 \\ 3 & -2 & 5 & -4 \\ 5 & -2 & 8 & -5 \end{bmatrix}$

**(g)** $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$

**(h)** $\begin{bmatrix} 2 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 & 3 & 4 \\ 5 & 5 & 6 & 7 & 5 & 5 \end{bmatrix}$

**2**) Biện luận theo tham số thực $\lambda$ hạng của các ma trận sau

**(a)** $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 6 & 8 & 9 & 10 \\ 5 & 8 & 11 & 13 & 16 \\ 10 & 16 & 22 & 26 & \lambda \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & 1 & 3 & 4 & 2 & 8 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 3 & 4 & 2 & 4 & 1 & -1 \\ 5 & 5 & 5 & 8 & 3 & \lambda \end{bmatrix}$

**(c)** $\begin{bmatrix} -1 & 2 & 1 & -1 & 1 \\ \lambda & -1 & 1 & -1 & -1 \\ 1 & \lambda & 0 & 1 & 1 \\ 1 & 2 & 2 & -1 & 1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & 2 & 1 & 0 & \lambda & 1 \\ 2 & 5 & 3 & 1 & 2\lambda & 3 \\ 3 & 7 & 4 & 1 & 3\lambda & 4 \\ 5 & 12 & 7 & 2 & 5\lambda & \lambda \end{bmatrix}$

**3**) Với giá trị nào của $\lambda \in \mathbb{R}$ thì hạng của các ma trận sau bằng 1.

**(a)** $\begin{bmatrix} -1 & 2 & 1 \\ 2 & \lambda & -2 \\ 3 & -6 & -3 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 6-\lambda & 3 \\ 4 & 8-\lambda & 12 \end{bmatrix}$ **(c)** $\begin{bmatrix} 1 & 3 & 5 \\ 4 & 12 & \lambda+5 \\ 5 & 15 & \lambda+50 \end{bmatrix}$

**4**) Chứng minh Định lý 7.2.6.

# § 8 Ma trận khả nghịch

## 8.1 Khái niệm và tính chất đơn giản

**8.1.1 Định nghĩa.** Ma trận vuông $A$ cấp $n$ trên $\mathbb{K}$ được gọi là **khả nghịch** nếu tồn tại một ma trận vuông $B$ cấp $n$ sao cho $AB = BA = I_n$. Khi đó ma trận $B$ được gọi là **ma trận nghịch đảo** (inverse) của ma trận $A$.

**8.1.2 Nhận xét.** Nếu $A$ khả nghịch thì ma trận nghịch đảo $B$ trong định nghĩa trên là duy nhất. Thật vậy, giả sử có ma trận nghịch đảo $B'$ của $A$. Ta có $AB' = B'A = I_n$. Khi đó

$$B' = I_nB' = (BA)B' = B(AB') = BI_n = B.$$

Nếu $A$ là ma trận khả nghịch thì ma trận nghịch đảo của $A$ kí hiệu $A^{-1}$.

**8.1.3 Thí dụ.** Các ma trận sơ cấp là các ma trận khả nghịch và

$$(E_{I;i,j})^{-1} = E_{I;i,j} \qquad (E_{II;i,\lambda})^{-1} = E_{II;i,\frac{1}{\lambda}}$$
$$(E_{III;i,j,\lambda})^{-1} = E_{III;i,j,-\lambda}.$$ □

**8.1.4 Thí dụ. (OSV94)** Cho $a_1, a_2, a_3, a_4 \in \mathbb{R}$. Chứng minh rằng ma trận

$$A = \begin{bmatrix} 1+a_1^2 & -a_2 & -a_3 & a_4 \\ a_2 & 1+a_1^2 & a_4 & a_3 \\ a_3 & -a_4 & 1+a_1^2 & -a_2 \\ -a_4 & -a_3 & a_2 & 1+a_1^2 \end{bmatrix}$$

khả nghịch. Tìm $A^{-1}$.

*Giải.* Dễ dàng tính được

$$A^tA = ((1+a_1^2)^2 + a_2^2 + a_3^2 + a_4^2) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$AA^t = ((1+a_1^2)^2 + a_2^2 + a_3^2 + a_4^2) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rõ ràng $(1+a_1^2) + a_2^2 + a_3^2 + a_4^2 \geq 1$, cho nên ta suy ra được $A$ khả nghịch và $A^{-1} = \frac{1}{(1+a_1^2)+a_2^2+a_3^2+a_4^2} A^t$. □

**8.1.5 Mệnh đề.** *Cho $A$ và $B$ là hai ma trận khả nghịch cùng cấp. Ta có*

**(i)** $(A^{-1})^{-1} = A$

**(ii)** $(A^t)^{-1} = (A^{-1})^t$

**(iii)** $(AB)^{-1} = B^{-1}A^{-1}$.

*Chứng minh.* Bài tập 3. □

**8.1.6 Định lý.** *Ma trận vuông $A$ cấp $n$ khả nghịch khi và chỉ khi* $\det(A) \neq 0$ *hay* $\operatorname{rank} A = n$ *(nghĩa là $A$ không suy biến).*

*Chứng minh.* ($\Longrightarrow$) Nếu $A$ khả nghịch thì tồn tại ma trận $B$ sao cho: $AB = I_n$. Suy ra $\det A \det B = \det(AB) = \det I_n = 1$. Vậy $\det A \neq 0$.

($\Longleftarrow$) Nếu $\det A \neq 0$ và $A = (a_{ij})_{n\times n}$. Đặt $b_{ij} = \dfrac{A_{ji}}{\det A}$, với $A_{ji}$ là phần bù đại số của phần tử $a_{ji}$. Xét ma trận $B = (b_{ij})_{n\times n}$. Khi đó tích $AB$ có thành phần thứ $(i,j)$ là

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} a_{ik}\frac{A_{jk}}{\det A} = \frac{1}{\det A}\sum_{k=1}^{n} a_{ik}A_{jk}.$$

Theo Hệ quả 5.2.7 trang 131 ta có $\sum\limits_{k=1}^{n} a_{ik}A_{jk} = \begin{cases} \det A & \text{khi } i = j \\ 0 & \text{khi } i \neq j. \end{cases}$

Vậy $c_{ij} = \begin{cases} 1 & \text{khi } i = j \\ 0 & \text{khi } i \neq j. \end{cases}$ Do đó $AB = I_n$.

Tương tự, ta cũng chứng minh được $BA = I_n$. Vậy $A$ khả nghịch. □

**8.1.7 Thí dụ. (OSV96)** Cho $A$ là một ma trận vuông cấp $n$. Chứng minh rằng nếu $A$ không khả nghịch thì có thể thay thế một số phần tử $a_{ij}$ của $A$ bởi số 0 hoặc 1, còn các phần tử khác vẫn giữ nguyên, để nhận được ma trận mới $S$ khả nghịch.

*Giải.* Ta chứng minh bằng qui nạp. Rõ ràng kết quả đúng với $n = 1$. Giả sử ta có kết quả với mọi ma trận vuông có cấp $n - 1$. Với $A$ là ma trận cấp $n$ tùy ý. Nếu $a_{11} = 0$ thì ta thay thế nó bởi 1. Các phần tử $a_{1k}$ và $a_{k1}$ với $k = 2, \dots, n$ được thay thế bởi 0. Ma trận $M_{11}$ thu được từ $A$ bởi bỏ đi cột thứ nhất và dòng thứ nhất. Theo giả thiết qui nạp nếu $M_{11}$ không khả nghịch thì ta có thể thay thế một số phần tử của $M_{11}$ bởi 0 hoặc 1 và giữ nguyên các phần tử khác để được ma trận khả nghịch, cho nên ta có thể giả sử ma trận $M_{11}$ khả nghịch, suy ra $|M_{11}| \neq 0$. Ta khai triển theo dòng để tính định thức $A$ sau khi đã thay thế một số phần tử bởi 0 hoặc 1 và giữ nguyên các phần tử còn lại ta được $a_{11}|M_{11}| \neq 0$. Vậy ma trận thu được khả nghịch. □

**8.1.8 Nhận xét.** Từ Định lý 8.1.6 ta nhận thấy nếu $A, B$ là hai ma trận vuông cấp $n$ thỏa $AB = I$ thì $A$ khả nghịch và $A^{-1} = B$; từ đó ta viết $AB = BA = I$.

**8.1.9 Thí dụ. (OSV07)** Giả sử $A$, $B$ là các ma trận vuông cấp $n \geq 2$ thỏa mãn điều kiện $AB + aA + bB = O$ trong đó $a$, $b$ là hai số thực khác 0. Chứng minh rằng $AB = BA$.

*Giải.* Theo giả thiết ta có $O = AB + aA + bB = A(B + aI) + b(B + aI) - abI = (A + bI)(B + aI) - abI$. Suy ra

$$\frac{1}{ab}(A + bI)(B + aI) = I$$

theo nhận xét trên ta suy ra

$$\frac{1}{ab}(B + aI)(A + bI) = I.$$

Do đó $(A+bI)(B+aI) = (B+aI)(A+bI)$ hay $AB+aA+bB+abI = BA + bB + aA + abI$, cho nên $AB = BA$. □

**8.1.10 Thí dụ. (OSV10)** Cho $A, B, C$ là các ma trận thực, vuông cấp $n$, trong đó $A$ khả nghịch và đồng thời giao hoán với $B, C$. Giả sử $C(A + B) = B$. Chứng minh rằng $B, C$ giao hoán với nhau.

*Giải.* Từ giả thiết ta viết lại $O = C(A + B) - B = C(A + B) - (A + B) + A = (C - I)(A + B) + A$. Do đó, ta được $A = (I - C)(A + B)$. Từ đó và do nhận xét trên suy ra

$$I = (I - C)(A + B)A^{-1} = (A + B)A^{-1}(I - C) = (A + B)(I - C)A^{-1}$$

ở đẳng thức cuối ta có $A^{-1}$ giao hoán với $C$ bởi vì $AC = CA$ và nhân bên trái và bên phải cho $A^{-1}$. Từ đó suy ra $A = (A + B)(I - C)$. Do đó, ta có $(I - C)(A + B) = (A + B)(I - C)$, cho nên

$$A + B - CA - CB = A - AC + B - BC.$$

Từ đó ta thu được $BC = CB$. □

**8.1.11 Nhận xét.** Trong chứng minh Định lý 8.1.6 trên ta nhận thấy nếu $A = (a_{ij})_{n\times n}$ khả nghịch thì ta có công thức ma trận nghịch đảo của nó:

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix} \tag{8.1.12}$$

trong đó $A_{ij}$ là phần bù đại số của phần tử $a_{ij}$. Ma trận

$$\begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}$$

còn được gọi là **ma trận phụ hợp** của $A$, ký hiệu $P_A$.

**8.1.13 Thí dụ.** Tìm ma trận nghịch đảo của ma trận $A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 2 & -1 \\ 3 & 1 & 5 \end{bmatrix}$.

Ta tính $\det A = \begin{vmatrix} 1 & 3 & 0 \\ 0 & 2 & -1 \\ 3 & 1 & 5 \end{vmatrix} = 1\begin{vmatrix} 2 & -1 \\ 1 & 5 \end{vmatrix} - 3\begin{vmatrix} 0 & -1 \\ 3 & 5 \end{vmatrix} 11 - 9 = 2.$

Tìm các phần bù đại số:

$$A_{11} = \begin{vmatrix} 2 & -1 \\ 1 & 5 \end{vmatrix} = 11, \quad A_{12} = -\begin{vmatrix} 0 & -1 \\ 3 & 5 \end{vmatrix} = -3, \quad A_{13} = \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} = -6,$$

$$A_{21} = -\begin{vmatrix} 3 & 0 \\ 1 & 5 \end{vmatrix} = -15, \quad A_{22} = \begin{vmatrix} 1 & 0 \\ 3 & 5 \end{vmatrix} = 5, \quad A_{23} = -\begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} = 8,$$

$$A_{31} = \begin{vmatrix} 3 & 0 \\ 2 & -1 \end{vmatrix} = -3, \quad A_{32} = -\begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} = 1, \quad A_{33} = \begin{vmatrix} 1 & 3 \\ 0 & 2 \end{vmatrix} = 2.$$

Thiết lập ma trận nghịch đảo $A^{-1} = \dfrac{1}{2}\begin{bmatrix} 11 & -15 & -3 \\ -3 & 5 & 1 \\ -6 & 8 & 2 \end{bmatrix}$. □

**8.1.14 Thí dụ.** Với điều kiện nào ma trận $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_2(\mathbb{R})$ khả nghịch? Trong điều kiện ấy, hãy tìm $A^{-1}$.

$A$ khả nghịch khi và chỉ khi $\det A \neq 0$ nghĩa là $ad - bc \neq 0$. Ma trận phụ hợp của $A$ là

$$P_A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Do đó với điều kiện $ad - bc \neq 0$ ma trận nghịch đảo của $A$ là

$$A^{-1} = \begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}.$$ □

**8.1.15 Thí dụ.** Tìm giá trị của tham số thực $a$ để ma trận sau khả nghịch

$$A = \begin{bmatrix} a & 0 & 0 \\ a^{10} & a-1 & a^{100} \\ a^{20} & 0 & a+1 \end{bmatrix}.$$

Ta có $A$ khả nghịch khi và chỉ khi $\det A \neq 0$. Do

$$\det A = \begin{vmatrix} a & 0 & 0 \\ a^{10} & a-1 & a^{100} \\ a^{20} & 0 & a+1 \end{vmatrix} = a \begin{vmatrix} a-1 & a^{100} \\ 0 & a+1 \end{vmatrix} = a(a-1)(a+1)$$

nên $A$ khả nghịch khi và chỉ khi $a \neq 0$ và $a \neq \pm 1$. □

**8.1.16 Mệnh đề.** *Nếu $A$ là một ma trận vuông thỏa $A^k = O$ với $k$ nguyên dương nào đó thì $I - A$ và $I + A$ khả nghịch.*

*Chứng minh.* Ta có đẳng thức sau

$$I = I - A^k = (I - A)(I + A + \cdots + A^{k-1}).$$

Từ đó ta kết luận được $I - A$ khả nghịch; hơn nữa ma trận nghịch đảo của nó là $I + A + \cdots + A^{k-1}$.

Vì $A^k = O$ nên $(-A)^k = O$ nên theo kết quả trên $I + A = I - (-A)$ khả nghịch. □

**8.1.17 Thí dụ. (OSV03)** Cho $P$ và $Q$ là hai ma trận vuông cấp $n$ thỏa mãn các điều kiện sau: $PQ = QP$ và tồn tại các số nguyên dương $s, r$ sao cho $P^s = Q^r = O$. Chứng minh rằng các ma trận $I + (P + Q)$ và $I - (P + Q)$ là các ma trận khả nghịch.

*Giải.* Với $m \geq r + s$ và do $P$, $Q$ giao hoán với nhau ta có

$$(P+Q)^m = \sum_{k=0}^{m} C_m^k P^k Q^{m-k} = O,$$

(vì $m - k \geq r$ hoặc $k \geq s$ với $k = 0, 1, \ldots, m$, nên $P^k = O$ hay $Q^{m-k} = O$ với $k = 0, 1, \ldots, m$). Do đó, theo mệnh đề trên ta được $I - (P + Q)$ và $I + P + Q$ khả nghịch. □

**8.1.18 Định lý.** *Cho $A$ là ma trận vuông cấp $n$ trên $\mathbb{K}$ (với $n \geq 2$). Khi đó các khẳng định sau đây là tương đương:*

**(i)** *$A$ khả nghịch.*

**(ii)** *$I_n$ nhận được từ $A$ bởi một số hữu hạn các phép biến đổi sơ cấp dòng (hay cột).*

**(iii)** *$A$ là tích của một số hữu hạn các ma trận sơ cấp dòng (hay cột).*

*Chứng minh.*

(i) $\Rightarrow$ (ii) Do $A$ khả nghịch nên $\operatorname{rank} A = n$. Từ ma trận $A$ ta có thể áp dụng một số hữu hạn các phép biến đổi sơ cấp theo dòng để được ma trận bậc thang rút gọn. Do ma trận bậc thang vuông cấp $n$ có hạng bằng $n$, nên nó phải là $I_n$.

(ii) $\Rightarrow$ (iii) Vì mỗi phép biến đổi sơ cấp dòng trên $A$ thực chất là phép nhân trái của $A$ với ma trận sơ cấp tương ứng. Vậy theo giả thiết tồn tại một dãy các ma trận sơ cấp $E_1, \ldots, E_m$ sao cho $E_m \cdots E_1 A = I_n$. Từ đó suy ra $A = E_1^{-1} \cdots E_m^{-1}$. Do nghịch đảo của ma trận sơ cấp là ma trận sơ cấp nên $A$ là tích các ma trận sơ cấp.

(iii) $\Rightarrow$ (i) Giả sử có một dãy ma trận sơ cấp $E_1, \ldots, E_k$ sao cho $A = E_1 \cdots E_k$. Vì các ma trận sơ cấp khả nghịch (các ma trận $E_i$ khả nghịch) nên $A$ khả nghịch. □

## 8.2 Tìm ma trận nghịch đảo nhờ các phép biến đổi sơ cấp (phương pháp Gauss)

**8.2.1 (Thuật toán tìm ma trận nghịch đảo)** Muốn tìm ma trận nghịch đảo (nếu có) của ma trận $A$ cấp $n \geq 2$ trên $\mathbb{K}$ ta làm như sau:

**1.** Lập ma trận $B = [A|I_n]$ cấp $n \times 2n$ bằng cách ghép thêm vào bên phải của $A$ ma trận đơn vị $I_n$.

**2.** Chỉ dùng các phép biến đổi sơ cấp trên dòng đưa ma trận $B$ về dạng $[I_n|C]$ (nếu $A$ khả nghịch). Khi đó $C = A^{-1}$. (Nếu không làm được như vậy có nghĩa $A$ là ma trận không khả nghịch.)

**8.2.2 Thí dụ.** Tìm ma trận nghịch đảo (nếu có) của $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \\ 3 & 2 & 5 \end{bmatrix}$.

Ta thực hiện các phép biến đổi sơ cấp trên dòng đối với ma trận sau

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 & 1 & 0 \\ 3 & 2 & 5 & 0 & 0 & 1 \end{array}\right] \xrightarrow{d_3 \to d_3 - 3d_1} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 & 1 & 0 \\ 0 & -4 & -4 & -3 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[\substack{d_2 \to -d_2 \\ d_3 \to d_3 - 4d_1}]{d_1 \to d_1 + 2d_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 7 & 1 & 2 & 0 \\ 0 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & -12 & -3 & -4 & 1 \end{array}\right]$$

$$\xrightarrow[\substack{d_2 \to d_2 - \frac{1}{6}d_3 \\ d_3 \to -\frac{1}{12}d_3}]{d_1 \to d_1 + \frac{7}{12}d_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{3}{4} & -\frac{1}{3} & \frac{7}{12} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{3} & -\frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{3} & -\frac{1}{12} \end{array}\right]$$

Vậy ma trận $A$ khả nghịch và ma trận nghịch đảo của nó là

$$A^{-1} = \begin{bmatrix} -\frac{3}{4} & -\frac{1}{3} & \frac{7}{12} \\ \frac{1}{2} & -\frac{1}{3} & -\frac{1}{6} \\ \frac{1}{4} & \frac{1}{3} & -\frac{1}{12} \end{bmatrix}.$$ □

**8.2.3 Nhận xét.** Thuật toán tìm ma trận nghịch đảo gắn liền với các phép biến đổi sơ cấp dòng nên thường gọi là *phương pháp biến đổi sơ cấp dòng tìm ma trận nghịch đảo*. Ta nhấn mạnh rằng, trong quá trình biến đổi, *chỉ được dùng các phép biến đổi sơ cấp dòng.*

Một thuật toán tương tự mà *chỉ dùng phép biến đổi sơ cấp cột*, nếu xuất phát từ ma trận chia khối $\left[\frac{A}{I_n}\right]$ (cấp $2n \times n$) thay cho ma trận $\left[A \mid I_n\right]$.

## 8.3 Quan hệ tương đương và quan hệ đồng dạng trên các ma trận

**8.3.1 Định nghĩa.** Cho $A$ và $B$ là hai ma trận cùng cấp $m \times n$ trên $\mathbb{K}$. Ta nói rằng $A$ **tương đương** với $B$, ký hiệu $A \sim B$, nếu tồn tại hai ma trận trên $\mathbb{K}$ khả nghịch $C$ và $D$ cấp lần lượt là $m$ và $n$ sao cho $B = CAD$.

**8.3.2 Mệnh đề.** *Quan hệ $\sim$ trong định nghĩa trên là một quan hệ tương đương trên $\mathcal{M}_{m,n}$.*

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

Nhờ kết quả của mệnh đề trên khi ta có $A \sim B$ ta nói rằng $A$ và $B$ *tương đương với nhau.*

**8.3.3 Định lý.** *Cho $A, B \in \mathcal{M}_{m,n}(\mathbb{K})$, $m, n \geq 2$. Khi đó, $A$ và $B$ tương đương với nhau khi và chỉ khi ma trận này nhận được từ ma trận kia bởi một số hữu hạn các phép biến đổi sơ cấp, điều này lại tương đương với $A$ và $B$ có cùng dạng bậc thang chính tắc.*

*Chứng minh.* Suy ra trực tiếp từ Định lý 8.1.18. □

**8.3.4 Hệ quả.** *Hai ma trận $A$ và $B$ cùng cấp tương đương nhau khi và chỉ khi* $\operatorname{rank}(A) = \operatorname{rank}(B)$.

*Chứng minh.* Đây là hệ quả trực tiếp từ định lý trên và các định lý 7.2.1 và 7.2.3. □

**8.3.5 Thí dụ.** Xét hai ma trận thực

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \qquad B = \begin{bmatrix} 11 & 10 & 9 & 12 \\ 7 & 6 & 5 & 8 \\ 3 & 2 & 1 & 4 \end{bmatrix}$$

Khi đó $A \sim B$, vì có hai ma trận sơ cấp (khả nghịch)

$$C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

sao cho $B = CAD$. Thực chất $B$ có được từ $A$ bởi các phép biến đổi sơ cấp sau.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \xrightarrow{d_1 \leftrightarrow d_3} \begin{bmatrix} 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

$$\xrightarrow{c_1 \leftrightarrow c_3} \begin{bmatrix} 11 & 10 & 9 & 12 \\ 7 & 6 & 5 & 8 \\ 3 & 2 & 1 & 4 \end{bmatrix} = B$$

□

**8.3.6 Định nghĩa.** Cho hai ma trận vuông $A$ và $B$ cùng cấp $n$ trên $\mathbb{K}$. Ta bảo $A$ **đồng dạng** với $B$, ký hiệu $A \backsim B$, nếu có một ma trận vuông trên $\mathbb{K}$ khả nghịch $C$ cùng cấp $n$ sao cho $B = C^{-1}AC$.

**8.3.7 Mệnh đề.** *Quan hệ $\backsim$ là một quan hệ tương đương trên $\mathcal{M}_n(\mathbb{K})$. Hơn nữa, $\backsim$ mạnh hơn $\sim$, tức là với mọi $A, B \in \mathcal{M}_n(\mathbb{K})$ nếu $A \backsim B$ thì $A \sim B$.*

**8.3.8 Hệ quả.** *Nếu $A$ và $B$ là hai ma trận vuông đồng dạng thì* $\operatorname{rank}(A) = \operatorname{rank}(B)$.

*Chứng minh.* Do $A \backsim B$ nên $A \sim B$, suy ra $\operatorname{rank}(A) = \operatorname{rank}(B)$. □

**8.3.9 *Chú ý.*** Nếu $A \backsim B$ thì $B$ cũng nhận được từ $A$ bởi một số hữu hạn các phép biến đổi sơ cấp và ngược lại. Hơn nữa, để thu được $B$ từ $A$, ta có thể thực hiện liên tiếp một dãy hữu hạn nào đó các cặp phép biến đổi sơ cấp theo dòng và cột cùng kiểu.

**8.3.10 Thí dụ.** Xét hai ma trận vuông thực

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \qquad B = \begin{bmatrix} 16 & 15 & 14 & 13 \\ 12 & 11 & 10 & 9 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{bmatrix}$$

và ma trận khả nghịch

$$C = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = C^{-1}$$

Tính toán trực tiếp cho ta $C^{-1}AC = B$ tức là $A \backsim B$. Thực sự $B$ nhận được từ $A$ bởi các cặp phép biến đổi sơ cấp dòng cột sau đây.

$$A \xrightarrow{d_2 \leftrightarrow d_3} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 \\ 13 & 14 & 15 & 16 \end{bmatrix} \xrightarrow{c_2 \leftrightarrow c_3} \begin{bmatrix} 1 & 3 & 2 & 4 \\ 9 & 11 & 10 & 12 \\ 5 & 7 & 6 & 8 \\ 13 & 15 & 14 & 16 \end{bmatrix}$$

$$\xrightarrow{d_1 \leftrightarrow d_4} \begin{bmatrix} 13 & 15 & 14 & 16 \\ 9 & 11 & 10 & 12 \\ 5 & 7 & 6 & 8 \\ 1 & 3 & 2 & 4 \end{bmatrix} \xrightarrow{c_1 \leftrightarrow c_4} \begin{bmatrix} 16 & 15 & 14 & 13 \\ 12 & 11 & 10 & 9 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{bmatrix} = B. \quad \square$$

## Bài tập

**1**) Tìm các ma trận nghịch đảo (nếu có) của các ma trận sau

**(a)** $\begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 1 \\ 3 & 2 & 2 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ **(c)** $\begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & 3 & 5 \\ 5 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix}$

**(e)** $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$ **(f)** $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 2 \\ 2 & 0 & 1 & 1 \end{bmatrix}$ **(g)** $\begin{bmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$

**2) (OSV00)** Cho $A$ và $B$ là các ma trận vuông cấp $n$ và thỏa mãn các điều kiện

$$AB = BA, \qquad A^{1999} = O, \qquad B^{2000} = O.$$

Chứng minh rằng ma trận $I + A + B$ khả nghịch.

**3)** Chứng minh Mệnh đề 8.1.5.

**4)** Chứng minh Mệnh đề 8.3.2.

**5)** Chứng minh rằng với mọi $A, B \in \mathcal{M}_n(\mathbb{K})$, nếu $A \backsim B$ thì $\mathrm{tr}(A) = \mathrm{tr}(B)$.

**6)** **(a)** Chứng minh rằng nếu $A$ khả nghịch thì với mọi $m \in \mathbb{N}$, $A^m$ cũng khả nghịch và $\big(A^m\big)^{-1} = \big(A^{-1}\big)^m$.

**(b)** Với mỗi ma trận khả nghịch tùy ý $X$, ta định nghĩa

$$X^{-m} = \big(X^{-1}\big)^m, \qquad \text{với mọi } m \in \mathbb{N} \setminus \{0\}.$$

Chứng minh rằng với mọi $A, B \in \mathcal{M}_n(\mathbb{K})$ ta đều có

**(i)** $\big(A^p\big)^q = A^{pq}$ và $A^p A^q = A^{p+q}$ với mọi $p, q \in \mathbb{Z}$.

**(ii)** Nếu $AB = BA$ thì $(AB)^p = A^p B^p$ với mọi $p \in \mathbb{Z}$.

**7)** Cho $A \in \mathcal{M}_n(\mathbb{K})$ thỏa mãn hệ thức $A^k = O_n$ với $k$ nguyên dương nào đó. Chứng minh rằng $I_n - A$ khả nghịch. Tìm $(I_n - A)^{-1}$.

**8)** Cho $A, B \in \mathcal{M}_n(\mathbb{K})$. Chứng minh rằng nếu $AB = I_n$ thì $A$ khả nghịch và $A^{-1} = B$. Nói riêng $AB = I_n$ khi và chỉ khi $BA = I_n$.

**9)** Chứng minh rằng nếu một trong hai ma trận vuông cùng cấp $A$ và $B$ không suy biến thì $AB$ và $BA$ là đồng dạng.

**10)** Hãy tìm tất cả các ma trận vuông cấp $n$ trên trường số thực mà chỉ đồng dạng với chính nó.

**11 ) (a)** Chứng minh tập hợp $GL_n(\mathbb{K}) = \{A \in \mathcal{M}_n(\mathbb{K}) : A$ khả nghịch$\}$ là một nhóm đối với phép nhân ma trận (gọi là **nhóm các ma trận khả nghịch** cấp $n$ trên $\mathbb{K}$).

**(b)** Ma trận vuông thực $A \in GL_n(\mathbb{R})$ gọi là **ma trận trực giao** nếu $A^t = A^{-1}$. Chứng minh rằng tập hợp $\mathcal{O}(n) = \{A \in GL_n(\mathbb{R}) :$ $A$ là ma trận trực giao$\}$ là một nhóm đối với phép nhân ma trận (gọi là **nhóm trực giao** cấp $n$).

**(c)** Cho $A = (a_{ij})_n \in \mathcal{M}_n(\mathbb{C})$, ma trận liên hợp của $A$, ký hiệu bởi $\overline{A}$ là ma trận vuông phức cùng cấp $n$ sao cho $(\overline{A})_{ij} = \overline{a}_{ij}$ với $i, j = 1, \ldots, n$. Ma trận $A \in GL_n(\mathbb{C})$ gọi là **ma trận unita** nếu $A^{-1} = \overline{A}^t$. Chứng minh rằng $\mathcal{U}(n) = \{A \in GL_n(\mathbb{C}) : A$ là ma trận unita$\}$ là một nhóm đối với phép nhân ma trận (gọi là **nhóm unita** cấp $n$). Hơn nữa, $\mathcal{O}(n) \subset \mathcal{U}(n)$.

**12 )** Cho $A \in \mathcal{M}_n(\mathbb{K})$ với $n \geq 2$ và $P_A$ là ma trận phụ hợp của $A$. Chứng minh rằng

**(a)** $\det(P_A) = (\det A)^{n-1}$

**(b)** Nếu $A$ không suy biến thì $P_A$ cũng vậy

**(c)** Nếu $\text{rank}(A) = n - 1$ thì $\text{rank}(P_A) = 1$

**(d)** Nếu $\text{rank}(A) \leq n - 2$ thì $P_A = O$.

# Chương III

# Hệ phương trình tuyến tính

## § 1 Khái niệm chung về hệ phương trình tuyến tính - Hệ Cramer

### 1.1 Các khái niệm

**1.1.1 Định nghĩa.** Hệ $m$ phương trình tuyến tính $n$ ẩn số $x_1, x_2, \dots, x_n$ là hệ có dạng

$$(1.1.2) \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

trong đó $a_{ij}$, $b_i$ với $i = 1, \dots m$ và $j = 1, \dots n$ là các số cho trước thuộc $\mathbb{K}$. Hệ (1.1.2) được gọi là **hệ phương trình tuyến tính** (system of linear equations) trên $\mathbb{K}$. Số $a_{ij}$ gọi là các **hệ số** (coefficient) của $x_j$ trong phương trình thứ $i$ còn $b_i$ gọi là các **hệ số tự do** của phương trình thứ $i$.

**1.1.3 Định nghĩa.** Xét hệ phương trình (1.1.2). Các hệ số $a_{ij}$ lập thành một ma trận cấp $m \times n$ trên $\mathbb{K}$

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

được gọi là **ma trận hệ số** (coefficient matrix) của hệ (1.1.2). Hạng của ma trận $A$ được gọi là **hạng của hệ phương trình**. Ma trận cột gồm các hệ số tự do

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

gọi là **ma trận hệ số tự do** hay vắn tắt là **cột tự do** của hệ (1.1.2). Ma trận chia khối

$$\bar{A} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right]$$

gọi là **ma trận mở rộng** hay **ma trận bổ sung** (augmented matrix) của hệ (1.1.2). Ma trận cột gồm các ẩn số

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

gọi là **ma trận ẩn số** hay **cột ẩn số** của hệ (1.1.2). Với các kí hiệu ma trận vừa nêu, hệ phương trình tuyến tính (1.1.2) có thể viết dưới dạng ma trận như sau

$$A\mathbf{x} = \mathbf{b} \tag{1.1.4}$$

và được gọi là **dạng ma trận** (matrix representation) của hệ (1.1.2).

**1.1.5 Định nghĩa.** Khi mọi hệ số tự do trong hệ phương trình (1.1.2) triệt tiêu: $b_1 = b_2 = \cdots = b_m = 0$, hệ (1.1.2) trở thành hệ

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0. \end{cases} \tag{1.1.6}$$

và gọi là **hệ phương trình tuyến tính thuần nhất** (homogeneous system of linear equations). Ta cũng gọi (1.1.6) là hệ phương trình tuyến tính thuần nhất *tương ứng* với hệ phương trình tuyến tính (1.1.2). Để phân biệt ta gọi hệ (1.1.2) là **hệ phương trình tuyến tính tổng quát**.

**1.1.7 Định nghĩa.** **Nghiệm** (solution) của hệ (1.1.2) là một bộ $n$ số sắp thứ tự $(c_1, c_2, \dots c_n) \in \mathbb{K}^n$ sao khi thay $x_j = c_j$, $j = 1, \dots, n$ vào các phương trình của hệ (1.1.2) thì ta được các đồng nhất thức trên $\mathbb{K}$. Khi đó, với

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

ta có $A\mathbf{c} = \mathbf{b}$, nghĩa là $\mathbf{c}$ là nghiệm của (1.1.4). Ta cũng bảo $\mathbf{c}$ là nghiệm dạng ma trận của hệ (1.1.2).

Một hệ phương trình tuyến tính có thể có nghiệm (hay tương thích) hoặc vô nghiệm. Hệ có nghiệm lại có thể có nghiệm duy nhất hay nhiều nghiệm. Quá trình tìm tập hợp nghiệm của một hệ phương trình tuyến tính gọi là giải hệ phương trình tuyến tính đó.

**1.1.8 Định nghĩa.** Hai hệ phương trình có cùng số ẩn (số phương trình có thể khác nhau) được gọi là **tương đương** (equivalent) nếu tập nghiệm của chúng là giống nhau.

**1.1.9 Mệnh đề.** *Cho $A\mathbf{x} = \mathbf{b}$ là hệ phương trình tuyến tính $m$ phương trình $n$ ẩn số và $M$ là ma trận vuông cấp $m$ không suy biến. Khi đó hệ $(MA)\mathbf{x} = M\mathbf{b}$ tương đương với hệ $A\mathbf{x} = \mathbf{b}$.*

*Chứng minh.* Giả sử $\mathbf{x}_0$ là nghiệm của hệ $A\mathbf{x} = \mathbf{b}$. Khi đó, $A\mathbf{x}_0 = b$ suy ra $MA\mathbf{x}_0 = M(A\mathbf{x}_0) = M\mathbf{b}$; nghĩa là $\mathbf{x}_0$ là nghiệm của hệ $MA\mathbf{x} = M\mathbf{b}$. Ngược lại, giả sử $\mathbf{x}^*$ là nghiệm của $MA\mathbf{x} = M\mathbf{b}$. Khi đó $MA\mathbf{x}^* = M\mathbf{b}$. Do $M$ không suy biến nên nó có ma trận nghịch đảo $M^{-1}$. Suy ra $M^{-1}MA\mathbf{x}^* = M^{-1}M\mathbf{b}$, cho nên $A\mathbf{x}^* = \mathbf{b}$; nghĩa là $\mathbf{x}^*$ là nghiệm của hệ $A\mathbf{x} = \mathbf{b}$. Vậy hai hệ trên là tương đương. □

## 1.2 Hệ Cramer

**1.2.1 Định nghĩa.** Hệ phương trình tuyến tính gồm $n$ phương trình của $n$ ẩn số được gọi là **hệ Cramer** nếu ma trận các hệ số của nó là ma trận không suy biến.

**1.2.2 Định lý.** *Mọi hệ Cramer $n$ phương trình, $n$ ẩn số đều có duy nhất một nghiệm cho bởi công thức*

(1.2.3)
$$x_j = \frac{D_j}{D}, \qquad j = 1, \ldots, n,$$

*trong đó $D$ là định thức của ma trận hệ số của hệ, $D_j$ là định thức của ma trận nhận được từ ma trận hệ số bằng cách thay cột thứ $j$ bởi cột tự do, $j = 1, \ldots, n$.*

*Chứng minh.* Cho một hệ Cramer $n$ phương trình, $n$ ẩn số với ma trận hệ số là $A$ và cột tự do $\mathbf{b}$. Khi đó hệ có dạng ma trận là

(1.2.4)
$$A\mathbf{x} = \mathbf{b}.$$

Vì $A$ không suy biến (nonsingular matrix) nên $\det A \neq 0$ và $A$ khả nghịch với nghịch đảo là $A^{-1}$. Hệ (1.2.4) tương đương

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \qquad \text{tương đương} \qquad \mathbf{x} = A^{-1}\mathbf{b}.$$

Mặt khác, $A^{-1}$ được xác định bởi

$$A^{-1} = \frac{1}{D}P_A = \frac{1}{D}\begin{bmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{bmatrix}$$

trong đó $A_{ij}$ là phần bù đại số của $a_{ij}$ trong $A$. Vậy nghiệm của hệ Cramer ở dạng ma trận là

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{D}\begin{bmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Do đó $x_j = \dfrac{1}{D}(A_{1j}b_1 + A_{2j}b_2 + \cdots + A_{nj}b_n)$. Ta nhận thấy $A_{1j}b_1 + A_{2j}b_2 + \cdots + A_{nj}b_n$ chính là khai triển định thức $D_j$ (định thức của ma trận nhận được từ $A$ bằng cách thay cột thứ $j$ bởi cột tự do **b**) theo cột thứ $j$. Tức là $x_j = \dfrac{D_j}{D}$ với $j = 1, \ldots, n$. □

**1.2.5 Thí dụ.** Giải hệ phương trình $\begin{cases} 2x_1 - x_2 - 2x_3 = 5 \\ 4x_1 + x_2 + 2x_3 = 1 \\ 8x_1 - x_2 + x_3 = 5 \end{cases}$

Hệ phương trình đã cho là hệ Cramer bởi vì định thức của ma trận hệ số khác không:

$$\begin{vmatrix} 2 & -1 & -2 \\ 4 & 1 & 2 \\ 8 & -1 & 1 \end{vmatrix} = 18.$$

Theo công thức Cramer ta có nghiệm của hệ

$$x_1 = \frac{\begin{vmatrix} 5 & -1 & -2 \\ 1 & 1 & 2 \\ 5 & -1 & 1 \end{vmatrix}}{18} = 1, \qquad x_2 = \frac{\begin{vmatrix} 2 & 5 & -2 \\ 4 & 1 & 2 \\ 8 & 5 & 1 \end{vmatrix}}{18} = 1,$$

$$x_3 = \frac{\begin{vmatrix} 2 & -1 & 5 \\ 4 & 1 & 1 \\ 8 & -1 & 5 \end{vmatrix}}{18} = -2.$$

□

**1.2.6 Thí dụ.** Cho hệ phương trình tuyến tính $\begin{cases} ax + by = c \\ cy + az = b \\ cx + bz = a \end{cases}$ với $a$, $b$, $c$ là ba số thực khác không cho trước. Chứng minh rằng hệ có nghiệm duy nhất. Tìm nghiệm đó.

Ma trận hệ số của hệ phương trình là $A = \begin{bmatrix} a & b & 0 \\ 0 & c & a \\ c & 0 & b \end{bmatrix}$ và định thức của nó là $\det(A) = 2abc \neq 0$. Do đó hệ phương trình đã cho là một hệ Cramer, cho nên nó có nghiệm duy nhất.

$$x = \frac{\begin{bmatrix} c & b & 0 \\ b & c & a \\ a & 0 & b \end{bmatrix}}{2abc} = \frac{a^2 - b^2 + c^2}{2ac}, \quad y = \frac{\begin{bmatrix} a & c & 0 \\ 0 & b & a \\ c & a & b \end{bmatrix}}{2abc} = \frac{-a^2 + b^2 + c^2}{2bc}$$

$$z = \frac{\begin{bmatrix} a & b & c \\ 0 & c & b \\ c & 0 & a \end{bmatrix}}{2abc} = \frac{a^2 + b^2 - c^2}{2ab}. \qquad \square$$

## Bài tập

**1**) Chứng minh rằng các hệ phương trình sau là hệ Cramer và giải chúng

**(a)** $\begin{cases} x_1 + x_2 + 2x_3 = -1 \\ 2x_1 - x_2 + 2x_3 = -4 \\ 4x_1 + x_2 + 4x_3 = -2 \end{cases}$ **(b)** $\begin{cases} 2x_1 - x_2 - x_3 = 4 \\ 3x_1 + 4x_2 - 2x_3 = 11 \\ 3x_1 - 2x_2 + 4x_3 = 11 \end{cases}$

**(c)** $\begin{cases} 3x_1 + 4x_2 - x_3 = 7 \\ x_1 + 2x_2 - 3x_3 = 0 \\ 7x_1 + 10x_2 - 5x_3 = 2 \end{cases}$ **(d)** $\begin{cases} x_1 + 2x_2 + 4x_3 = 31 \\ 5x_1 + x_2 + 2x_3 = 29 \\ 3x_1 - x_2 + x_3 = 10 \end{cases}$

**(e)** $\begin{cases} 2x_1 - x_2 + 3x_3 + 2x_4 = 4 \\ 3x_1 + 3x_2 + 3x_3 + 2x_4 = 6 \\ 3x_1 - x_2 - x_3 + 2x_4 = 6 \\ 3x_1 - x_2 + 3x_3 - x_4 = 6 \end{cases}$ **(f)** $\begin{cases} x_1 + 2x_2 + 3x_3 - 2x_4 = 6 \\ 2x_1 - x_2 - 2x_3 - 3x_4 = 4 \\ 3x_1 + 2x_2 - x_3 + 2x_4 = 4 \\ 2x_1 - 3x_2 + 2x_3 + x_4 = -8 \end{cases}$

**2**) Cho ma trận $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 4 & 3 \\ 1 & 2 & 2 \end{bmatrix}$. Tính cột thứ ba của ma trận $A^{-1}$ bằng cách sử dụng công thức Cramer để giải $A\mathbf{x} = \mathbf{e}_3$.

**3**) Đặt $B_j$ ký hiệu ma trận thu được bằng việc thay cột thứ $j$ của ma trận đơn vị bởi vector $\mathbf{b} = (b_1, \ldots, b_n)$. Dùng công thức Cramer chứng tỏ rằng $b_j = \det(B_j)$ với $j = 1, 2, \ldots, n$.

# § 2 Hệ phương trình tuyến tính tổng quát

Trong mục này, ta sẽ nghiên cứu điều kiện có nghiệm và phương pháp giải một hệ phương trình tuyến tính tổng quát

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

## 2.1 Định lý Kronecker-Capelli

**2.1.1 Định lý.** *Cho hai hệ phương trình tuyến tính trên $\mathbb{K}$ cùng có $m$ phương trình của $n$ ẩn số với ma trận mở rộng lần lượt là $\bar{A} = [A|\mathbf{b}]$ và $\bar{A}' = [A'|\mathbf{b}']$ (với $m \geq 2$). Khi đó, nếu $\bar{A}'$ nhận được từ $\bar{A}$ bởi một số hữu hạn các phép biến đổi sơ cấp dòng thì hai hệ phương trình đã cho tương đương với nhau.*

*Chứng minh.* Vì $\bar{A}'$ nhận được từ $\bar{A}$ bởi một số hữu hạn các phép biến đổi sơ cấp dòng nên tồn tại ma trận không suy biến $M$ cấp $m$ sao cho $\bar{A}' = M\bar{A}$. Ta có $M\bar{A} = M[A|\mathbf{b}] = [MA|M\mathbf{b}]$. Vậy $A' = MA$ và $\mathbf{b}' = M\mathbf{b}$. Theo Mệnh đề 1.1.9 ta có hệ $A\mathbf{x} = \mathbf{b}$ và hệ $MA\mathbf{x} = M\mathbf{b}$ tương đương hay $A\mathbf{x} = \mathbf{b}$ và $A'\mathbf{x} = \mathbf{b}'$ tương đương. □

**2.1.2 Nhận xét.** Thực chất việc thực hiện một phép biến đổi sơ cấp dòng trên ma trận mở rộng của một hệ phương trình tuyến tính chính là việc thực hiện một trong các biến đổi sau đây trên các phương trình của hệ đang xét.

- Đổi chỗ hai phương trình nào đó.
- Nhân một phương trình nào đó với một thừa số khác không.
- Cộng vào một phương trình nào đó một bội của một phương trình khác.

**2.1.3 Định lý. (Kronecker-Capelli)** *Một hệ phương trình tuyến tính có nghiệm khi và chỉ khi ma trận hệ số và ma trận mở rộng của hệ có hạng bằng nhau.*

*Chứng minh.* Xét một hệ phương trình tuyến tính $m$ phương trình và $n$ ẩn với $A$ là ma trận hệ số của hệ phương trình và $\bar{A} = \begin{bmatrix} A \mid \mathbf{b} \end{bmatrix}$ là ma trận mở rộng của hệ. Dùng phép biến đổi sơ cấp dòng đưa ma trận mở rộng $\bar{A}$ về dạng bậc thang thu gọn ta được $\bar{A}' = \begin{bmatrix} A' \mid \mathbf{b}' \end{bmatrix}$. Khi đó, ta có $\operatorname{rank} A = \operatorname{rank} A'$ và $\operatorname{rank} \bar{A} = \operatorname{rank} \bar{A}'$. Giả sử $\operatorname{rank} A = \operatorname{rank} \bar{A} = r$. Khi đó, $\mathbf{b}' = (b'_1, \ldots, b'_r, 0, \ldots, 0)$. Gọi các cột ứng với các phần tử đánh dấu trong ma trận $\bar{A}'$ là $j_1, \ldots, j_r$. Từ Định lý 2.1.1 ta nhận thấy $x_{j_k} = b'_k$ với mọi $k = 1, \ldots, r$ và $x_i = 0$ với mọi $i \in \{1, 2, \ldots, n\} \setminus \{j_1, \ldots, j_r\}$ là một nghiệm của hệ phương trình đã cho.

Giả sử $\operatorname{rank} A \neq \operatorname{rank} \bar{A}$. Khi đó ta phải có $\operatorname{rank} A = r$ và $\operatorname{rank} \bar{A} = r+1$. Dòng thứ $r+1$ trong ma trận $\bar{A}'$ phải là $\begin{bmatrix} 0 & \ldots & 0 \mid 1 \end{bmatrix}$, nó ứng với phương trình $0x_1 + \cdots + 0x_n = 1$; đây là phương trình vô nghiệm. Do đó, theo Định lý 2.1.1 ta suy ra được hệ phương trình đã cho vô nghiệm. □

**2.1.4 Hệ quả.** *Nếu hạng của hệ phương trình tuyến tính đúng bằng số phương trình của nó thì hệ đó luôn có nghiệm.*

*Chứng minh.* Giả sử hệ $A\mathbf{x} = \mathbf{b}$ gồm $m$ phương trình của $n$ ẩn và hạng của hệ là $\operatorname{rank} A = m$. Khi đó

$$m = \operatorname{rank} A \leq \operatorname{rank} \bar{A} \leq \min\{m, n+1\} \leq m.$$

Vậy $\operatorname{rank} A = \operatorname{rank} \bar{A}$, cho nên hệ đã cho có nghiệm. □

**2.1.5** Xét hệ phương trình tuyến tính tổng quát (1.1.2) với $m$ phương trình $n$ ẩn số với dạng ma trận (1.1.4)

$$A\mathbf{x} = \mathbf{b}.$$

Giả sử rằng $A \neq O$ và $m, n \geq 2$ (tức là hệ có ít nhất hai phương trình hai ẩn số); hơn nữa hệ có hạng là $\operatorname{rank}(A) = r > 0$. Khi đó trong $A$ có ít nhất một định thức con khác không cấp $r$. Mỗi định thức con cấp $r$ khác không như thế gọi là một **định thức cơ sở** của $A$ và của hệ đã cho. Rõ ràng $A$ có thể có nhiều định thức cơ sở. Ta có thể tìm một định thức cơ sở bằng cách biến đổi ma trận $A$ về dạng bậc thang ma trận con của $A$ ứng với các hàng khác không ma trận bậc thang và các cột ứng với các cột của các phần tử được đánh dấu trong ma trận bậc thang; khi đó định thức của ma trận con ấy là một định thức cơ sở.

**2.1.6** Cho $D(r) = D_{i_1 i_2 \dots i_r, j_1 j_2 \dots j_r} \neq 0$ là một định thức cơ sở của hệ (1.1.2). Khi đó nếu (1.1.2) có nghiệm thì nó tương đương với hệ

$$(2.1.7)\qquad \begin{cases} a_{i_1 1}x_1 + a_{i_1 2}x_2 + \cdots + a_{i_1 n}x_n = b_{i_1} \\ a_{i_2 1}x_1 + a_{i_2 2}x_2 + \cdots + a_{i_2 n}x_n = b_{i_2} \\ \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{i_r 1}x_1 + a_{i_r 2}x_2 + \cdots + a_{i_r n}x_n = b_{i_r} \end{cases}$$

gồm $r$ phương trình nhận được từ (1.1.2) bằng cách xóa đi $n - r$ phương trình mà các hệ số nằm trên các dòng không chứa phần tử nào của $D(r)$. (Ta có điều tương đương này do ma trận bậc thang rút gọn của hệ (1.1.2) giống với ma trận bậc thang thu gọn của hệ trên trừ các dòng không.)

**2.1.8** Các phương trình của hệ (2.1.7) gọi là các **phương trình chính** của hệ (1.1.2). Các ẩn $x_{j_1}, x_{j_2}, \dots, x_{j_r}$ (tức là các ẩn có hệ số nằm trên các cột của $A$ chứa các phần tử của $D(r)$) gọi là các **ẩn chính**, $n - r$ ẩn còn lại $x_j$ với $j \in J = \{1, 2, \dots, n\} \setminus \{j_1, j_2, \dots, j_r\}$ gọi là **ẩn tự do** của hệ (1.1.2) ứng với định thức con cơ sở $D(r)$. Khi thay $D(r)$ bởi một định thức con cơ sở khác, nói chung các phương trình chính, ẩn chính, ẩn tự do cũng thay đổi.

**2.1.9** Khi $r = n$ thì hệ (2.1.7) trở thành hệ Cramer, cho nên nó có nghiệm duy nhất. Do đó hệ (1.1.2) cũng vậy.

Trường hợp $r < n$, ta viết lại hệ (2.1.7) với các ẩn chính bên trái và đưa các ẩn tự do qua vế phải, ta được một hệ phương trình mới tương đương

$$(2.1.10)\qquad \begin{cases} a_{i_1 j_1}x_{j_1} + \cdots + a_{i_1 j_r} = b_{i_1} - \sum\limits_{j \in J} a_{i_1} x_j \\ a_{i_2 j_1}x_{j_1} + \cdots + a_{i_2 j_r} = b_{i_2} - \sum\limits_{j \in J} a_{i_2} x_j \\ \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{i_r j_1}x_{j_1} + \cdots + a_{i_r j_r} = b_{i_r} - \sum\limits_{j \in J} a_{i_r} x_j \end{cases}$$

Xem các ẩn tự do là các *tham số nhận giá trị tùy ý trên* $\mathbb{K}$, hệ 2.1.10 một lần nữa là hệ Cramer và có nghiệm đối với các ẩn chính $x_{j_1}, \dots, x_{j_r}$ phụ thuộc $n-r$ tham số $x_j$ với $j \in J = \{1, 2, \dots, n\} \setminus \{j_1, \dots, j_r\}$. Vì $n-r$ ẩn tự do có thể nhận giá trị tùy ý trên $\mathbb{K}$ nên hệ (2.1.10) và cũng là hệ (1.1.2) có vô số nghiệm.

**2.1.11** Như vậy từ định lý Kronecker-Capelli ta đã chứng tỏ được các kết quả sau: Cho hệ phương trình tuyến tính tổng quát $A\mathbf{x} = \mathbf{b}$ ($m$ phương trình, $n$ ẩn số). Khi đó, với $A$ là ma trận các hệ số và $\bar{A}$ là ma trận bổ sung, ta có

**(i)** Nếu $\operatorname{rank} A < \operatorname{rank} \bar{A}$ thì hệ vô nghiệm.

**(ii)** Nếu $\operatorname{rank} A = \operatorname{rank} \bar{A} = n$ thì hệ có nghiệm duy nhất.

**(iii)** Nếu $\operatorname{rank} A = \operatorname{rank} \bar{A} = r < n$ thì hệ có vô số nghiệm (phụ thuộc $n-r$ tham số).

**2.1.12 Thí dụ. (OSV96)** Chứng minh rằng nếu $a \neq 0$ thì hệ

$$\begin{cases} ax + (1-b)y + cz + (1-d)t = a \\ (b-1)x + ay + (d-1)z + ct = b \\ -cx + (1-d)y + az + (b-1)t = c \\ (d-1)x - cy + (1-b)z + at = d \end{cases}$$

luôn có nghiệm với mọi $b, c, d \in \mathbb{R}$.

*Giải.* Ma trận hệ số của hệ phương trình đã cho là

$$A = \begin{bmatrix} a & 1-b & c & 1-d \\ b-1 & a & d-1 & c \\ -c & 1-d & a & b-1 \\ d-1 & -c & 1-b & a \end{bmatrix}$$

Dễ dàng kiểm tra được $AA^t = A^tA = [a^2 + (b-1)^2 + c^2 + (d-1)^2]I_4$. Vậy $A$ khả nghịch với $\det(A) = [a^2 + (b-1)^2 + c^2 + (d-1)^2]^2 > 0$ với mọi $b, c, d \in \mathbb{R}$ vì $a \neq 0$. Do đó hệ đã cho có nghiệm duy nhất. $\square$

**2.1.13 Thí dụ. (OSV02)** Cho hệ phương trình

$$\begin{cases} ax_1 + bx_2 + bx_3 + \cdots + bx_{2001} + bx_{2002} = 1 \\ bx_1 + ax_2 + bx_3 + \cdots + bx_{2001} + bx_{2002} = 2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ bx_1 + bx_2 + bx_3 + \cdots + ax_{2001} + bx_{2002} = 2001 \\ bx_1 + bx_2 + bx_3 + \cdots + bx_{2001} + ax_{2002} = 2002 \end{cases}$$

Tìm điều kiện đối với $a$ và $b$ để hệ phương trình đã cho có nghiệm duy nhất.

*Giải.* Đây hệ phương trình tuyến tính gồm 2002 phương trình và 2002 ẩn, cho nên nó có nghiệm duy nhất khi hạng của ma trận hệ số $A$ bằng 2002, nghĩa là $\det A \neq 0$. Ta tính định thức của $A$

$$\det A = \begin{vmatrix} a & b & b & \dots & b & b \\ b & a & b & \dots & b & b \\ \dots & \dots & \dots & \dots & \dots & \dots \\ b & b & b & \dots & a & b \\ b & b & b & \dots & b & a \end{vmatrix}$$

$$= (a + 2001b) \begin{vmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ b & a & b & \dots & b & b \\ \dots & \dots & \dots & \dots & \dots & \dots \\ b & b & b & \dots & a & b \\ b & b & b & \dots & b & a \end{vmatrix}$$

$$= (a + 2001b) \begin{vmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & a-b & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a-b & 0 \\ 0 & 0 & 0 & \dots & 0 & a-b \end{vmatrix}$$

$$= (a + 2001b)(a - b)^{2001}.$$

Vậy $a \neq b$ và $a + 2001b \neq 0$. □

Để giải hệ phương trình (1.1.2) ta phải tìm được các phương trình chính các ẩn chính. Một cách đơn giản để hoàn thành công việc trên là dùng phép biến đổi sơ cấp đưa ma trận mở rộng $\bar{A}$ về ma trận bậc thang rút gọn $\bar{A}''$. Từ ma trận bậc thang rút gọn này ta có thể xác định được hệ đã cho có nghiệm hay không. Hơn thế nữa, từ ma trận bậc thang rút gọn này ta nhận được các ẩn chính của hệ là các phần tử 1 được đánh dấu (các phương trình ứng với các dòng khác 0 của $\bar{A}''$ là các phương trình chính). Từ ma trận $\bar{A}''$ ta viết lại hệ phương trình và chuyển các ẩn tự do qua vế phải ta được nghiệm của hệ (1.1.2). Đó chính là nội dung của phương pháp Gauss-Jordan được trình ở phần sau.

## 2.2 Phương pháp Gauss-Jordan

### 2.2.1 Nội dung phương pháp giải hệ $A\mathbf{x} = \mathbf{b}$

- Dùng các phép biến đổi sơ cấp dòng để đưa ma trận mở rộng $\bar{A}$ về dạng bậc thang dòng $\bar{A}'$. Lúc đó dễ dàng thấy được hạng $\operatorname{rank} A$ và $\operatorname{rank} \bar{A}$, so sánh hai hạng này ta biết ngay hệ vô nghiệm hay có nghiệm.

- Giả sử hệ có nghiệm. Tiếp tục dùng các phép biến đổi sơ cấp dòng đưa ma trận bậc thang dòng $\bar{A}'$ về **ma trận bậc thang dòng thu gọn** (reduced row echelon form) $\bar{A}''$. Các ẩn tương ứng với số hạng đầu tiên khác không (là các số 1) ở các hàng khác không được gọi là **ẩn chính**, các ẩn còn lại gọi là **ẩn tự do**.

- Xét hệ mới (tương đương với hệ đã cho) mà ma trận mở rộng nhận được từ ma trận bậc thang dòng $\bar{A}''$ sau khi xóa đi các dòng không. Chuyển các số hạng chứa ẩn tự do sang vế phải, xem chúng là các tham số, ta được nghiệm của hệ phương trình đầu.

**2.2.2 Thí dụ.** Giải hệ phương trình $\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 7 \\ 2x_1 + x_2 + 2x_3 + 3x_4 = 6 \\ 3x_1 + 2x_2 + x_3 + 2x_4 = 7 \\ 4x_1 + 3x_2 + 2x_3 + x_4 = 18 \end{cases}$

*Giải.* Ta biến đổi sơ cấp ma trận hệ số mở rộng của hệ phương trình

$$\bar{A} = \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 7 \\ 2 & 1 & 2 & 3 & 6 \\ 3 & 2 & 1 & 2 & 7 \\ 4 & 3 & 2 & 1 & 18 \end{array}\right]$$

$$\xrightarrow[\substack{d_3 \to d_3 - 3d_1 \\ d_4 \to d_4 - 4d_1}]{d_2 \to d_2 - 2d_1} \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 7 \\ 0 & -3 & -4 & -5 & -8 \\ 0 & -4 & -8 & -10 & -14 \\ 0 & -5 & -10 & -15 & -10 \end{array}\right]$$

$$\xrightarrow[\substack{d_3 \to d_3 - \frac{4}{3}d_2 \\ d_4 \to d_4 - \frac{5}{3}d_2}]{\substack{d_1 \to d_1 + \frac{2}{3}d_2 \\ d_2 \to -\frac{1}{3}d_2}} \left[\begin{array}{cccc|c} 1 & 0 & \frac{1}{3} & \frac{2}{3} & \frac{5}{3} \\ 0 & 1 & \frac{4}{3} & \frac{5}{3} & \frac{8}{3} \\ 0 & 0 & -\frac{8}{3} & -\frac{10}{3} & -\frac{10}{3} \\ 0 & 0 & -\frac{10}{3} & -\frac{20}{3} & \frac{10}{3} \end{array}\right]$$

$$\xrightarrow[\substack{d_3 \to -\frac{3}{8}d_3 \\ d_4 \to d_4 - \frac{5}{4}d_3}]{\substack{d_1 \to d_1 + \frac{1}{8}d_3 \\ d_2 \to d_2 + \frac{1}{2}d_3}} \left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{4} & \frac{5}{4} \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{5}{4} & \frac{5}{4} \\ 0 & 0 & 0 & -\frac{5}{2} & \frac{15}{2} \end{array}\right]$$

$$\xrightarrow[\substack{d_3 \to d_3 + \frac{1}{2}d_4 \\ d_4 \to -\frac{2}{5}d_4}]{d_1 \to d_1 + \frac{1}{10}d_4} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right]$$

Vậy trong hệ này không có ẩn tự do, tất cả đều là ẩn chính, và nghiệm của hệ là $x_1 = 2$, $x_2 = 1$, $x_3 = 5$, $x_4 = -3$. □

**2.2.3 Thí dụ. (OSV09)** Giải hệ phương trình

$$\begin{cases} 2x_1 + x_2 - x_3 + 2x_4 + x_5 - x_6 = 1 \\ -x_1 + 2x_2 + 2x_3 + x_4 + x_5 - x_6 = 1 \\ x_1 - 2x_2 + 2x_3 + x_4 + x_5 - x_6 = 1 \\ -2x_1 - x_2 - x_3 + 2x_4 + x_5 - x_6 = 1 \\ 2x_1 + x_2 + x_3 - x_4 - x_5 + 2x_6 = 1 \\ -x_1 + 2x_2 + x_3 - x_4 + 2x_5 + x_6 = 1 \end{cases}$$

*Giải.* Ta biến đổi sơ cấp ma trận mở rộng $\bar{A}$

$$\bar{A}=\left[\begin{array}{cccccc|c}2&1&-1&2&1&-1&1\\-1&2&2&1&2&-1&1\\1^*&-2&2&1&1&1&1\\-2&-1&-1&2&1&-1&1\\2&1&1&-1&-1&2&1\\-1&2&1&-1&2&1&1\end{array}\right]$$

$$\xrightarrow[\cdots]{d_3\leftrightarrow d_1}\left[\begin{array}{cccccc|c}1&-2&2&1&1&-1&1\\0&0&4&2&2&-2&2\\0&5^*&-5&0&-1&1&-1\\0&-5&3&4&3&-3&3\\0&5&-3&-3&-3&4&-1\\0&0&3&0&3&0&2\end{array}\right]$$

$$\xrightarrow[\cdots]{d_3\leftrightarrow d_2}\left[\begin{array}{cccccc|c}1&-2&2&1&1&-1&1\\0&5&-5&0&-1&1&-1\\0&0&4&2&2&-2&2\\0&0&-2^*&4&2&-2&2\\0&0&2&-3&-2&3&0\\0&0&3&0&3&0&2\end{array}\right]$$

$$\xrightarrow[\cdots]{d_3\leftrightarrow d_4}\left[\begin{array}{cccccc|c}1&-2&0&5&3&-3&3\\0&5&0&-10&-6&6&-6\\0&0&1&-2&-1&1&-1\\0&0&0&5&3&-3&3\\0&0&0&1^*&0&1&2\\0&0&0&6&6&-3&5\end{array}\right]$$

$$\xrightarrow[\cdots]{d_4\leftrightarrow d_5}\left[\begin{array}{cccccc|c}1&-2&0&0&3&-8&-7\\0&5&0&0&-6&16&14\\0&0&1&0&-1&3&3\\0&0&0&1&0&1&2\\0&0&0&0&3^*&-8&-7\\0&0&0&0&6&-9&-7\end{array}\right]$$

$$\xrightarrow[\cdots]{d_4\leftrightarrow d_5}\left[\begin{array}{cccccc|c}1&-2&0&0&0&0&0\\0&5&0&0&0&0&0\\0&0&1&0&-1&3&3\\0&0&0&1&0&1&2\\0&0&0&0&3&-8&-7\\0&0&0&0&0&7^*&7\end{array}\right]$$

$$\xrightarrow[\cdots]{d_6\to\frac{1}{7}d_6}\left[\begin{array}{cccccc|c}1&0&0&0&0&0&0\\0&1&0&0&0&0&0\\0&0&1&0&-1&0&0\\0&0&0&1&0&0&1\\0&0&0&0&3&0&1\\0&0&0&0&0&1&1\end{array}\right]$$

$$\xrightarrow[\cdots]{d_6\to\frac{1}{7}d_6}\left[\begin{array}{cccccc|c}1&1&0&0&0&0&0\\0&1&0&0&0&0&0\\0&0&1&0&0&0&\frac{1}{3}\\0&0&0&1&0&0&1\\0&0&0&0&1&0&\frac{1}{3}\\0&0&0&0&0&1&1\end{array}\right]$$

Vậy nghiệm của hệ đã cho là $x_1 = 0$, $x_2 = 0$, $x_3 = \frac{1}{3}$, $x_4 = 1$, $x_5 = \frac{1}{3}$, $x_6 = 1$. □

**2.2.4 Thí dụ.** Giải hệ phương trình $\begin{cases} x_1 + 2x_2 - 3x_3 + 5x_4 = 1 \\ x_1 + 3x_2 - 13x_3 + 22x_4 = -1 \\ 3x_1 + 5x_2 + x_3 - 2x_4 = 5 \\ 2x_1 + 3x_2 + 4x_3 - 7x_4 = 4. \end{cases}$

Theo phương pháp Gauss-Jordan ta có

$$\bar{A} = \left[\begin{array}{cccc|c}1&2&-3&5&1\\1&3&-13&22&-1\\3&5&1&-2&5\\2&3&4&-7&4\end{array}\right]$$

$$\xrightarrow[\substack{d_3\to d_3-3d_1\\ d_4\to d_4-2d_1}]{d_2\to d_2-d_1}\left[\begin{array}{cccc|c}1&2&-3&5&1\\0&1&-10&17&-2\\0&-1&10&-17&2\\0&-1&10&-17&2\end{array}\right]$$

$$\xrightarrow[d_3\to d_3+d_2]{d_4\to d_4+d_2}\left[\begin{array}{cccc|c}1&2&-3&5&1\\0&1&-10&17&-2\\0&0&0&0&0\\0&0&0&0&0\end{array}\right]$$

$$\xrightarrow{d_1 \to d_1 - 2d_2} \left[\begin{array}{cccc|c} 1 & 0 & 17 & -29 & 5 \\ 0 & 1 & -10 & 17 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Vậy hệ có vô số nghiệm trong đó $x_1$ và $x_2$ là ẩn chính, $x_3$ và $x_4$ là ẩn tự do, và $x_1 = 5 - 17x_3 + 29x_4$, $x_2 = -2 + 10x_3 - 17x_4$. □

**2.2.5 Thí dụ.** Giải hệ phương trình $\begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 = 2 \\ 3x_1 + 3x_2 - 5x_3 + x_4 = -3 \\ -2x_1 + x_2 + 2x_3 - 3x_4 = 5 \\ 3x_1 + 3x_3 - 10x_4 = 8. \end{cases}$

Dùng phương pháp Gauss-Jordan ta có

$$\bar{A} = \left[\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 2 \\ 3 & 3 & -5 & 1 & -3 \\ -2 & 1 & 2 & -3 & 5 \\ 3 & 0 & 3 & -10 & 8 \end{array}\right]$$

$$\xrightarrow[\substack{d_3 \to d_3 + 2d_1 \\ d_4 \to d_4 - 3d_1}]{d_2 \to d_2 - 3d_1} \left[\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 2 \\ 0 & 9 & -14 & 13 & -9 \\ 0 & -3 & 8 & -11 & 9 \\ 0 & 6 & -6 & 2 & 2 \end{array}\right]$$

$$\xrightarrow{d_2 \leftrightarrow \frac{1}{3}d_4} \left[\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 2 \\ 0 & 1 & -1 & \frac{1}{3} & \frac{1}{3} \\ 0 & -3 & 8 & -11 & 9 \\ 0 & 9 & -14 & 13 & -9 \end{array}\right]$$

$$\xrightarrow[d_4 \to d_4 - 9d_2]{d_3 \to d_3 + 3d_2} \left[\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 2 \\ 0 & 1 & -1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 5 & -10 & 10 \\ 0 & 0 & -5 & 10 & -12 \end{array}\right]$$

$$\xrightarrow[d_3 \to \frac{1}{5}d_3]{d_4 \to d_4 + d_3} \left[\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 2 \\ 0 & 1 & -1 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & -2 \end{array}\right]$$

Vậy hệ vô nghiệm ($\operatorname{rank} A < \operatorname{rank} \bar{A}$). □

**2.2.6 Thí dụ.** Giải và biện luận theo $m$ hệ phương trình

$$\begin{cases} x_1 + 2x_2 + 2x_4 + x_5 = 1 \\ 2x_1 + 4x_2 + x_3 + 3x_4 = 3 \\ 3x_1 + 6x_2 + 2x_3 + 3x_4 + x_5 = m \\ x_1 + 2x_2 + x_3 + x_5 = 2m - 8. \end{cases}$$

Ta có

$$\bar{A} = \begin{bmatrix} 1 & 2 & 0 & 2 & 1 & | & 1 \\ 2 & 4 & 1 & 3 & 0 & | & 3 \\ 3 & 6 & 2 & 3 & 1 & | & m \\ 1 & 2 & 1 & 0 & 1 & | & 2m-8 \end{bmatrix}$$

$$\xrightarrow[\substack{d_3 \to d_3 - 3d_1 \\ d_4 \to d_4 - d_1}]{d_2 \to d_2 - 2d_1} \begin{bmatrix} 1 & 2 & 0 & 2 & 1 & | & 1 \\ 0 & 0 & 1 & -1 & -2 & | & 1 \\ 0 & 0 & 2 & -3 & -2 & | & m-3 \\ 0 & 0 & 1 & -2 & 0 & | & 2m-9 \end{bmatrix}$$

$$\xrightarrow[d_4 \to d_4 - d_2]{d_3 \to d_3 - 2d_2} \begin{bmatrix} 1 & 2 & 0 & 2 & 1 & | & 1 \\ 0 & 0 & 1 & -1 & -2 & | & 1 \\ 0 & 0 & 0 & -1 & 2 & | & m-5 \\ 0 & 0 & 0 & -1 & 2 & | & 2m-10 \end{bmatrix}$$

$$\xrightarrow[d_3 \to -d_3]{d_4 \to d_4 - d_1} \begin{bmatrix} 1 & 2 & 0 & 2 & 1 & | & 1 \\ 0 & 0 & 1 & -1 & -2 & | & 1 \\ 0 & 0 & 0 & 1 & -2 & | & -m+5 \\ 0 & 0 & 0 & 0 & 0 & | & m-5 \end{bmatrix}$$

Ta nhận thấy khi $m \neq 5$ thì hệ vô nghiệm. Trường hợp $m = 5$, ta có

$$\bar{A} \to \begin{bmatrix} 1 & 2 & 0 & 2 & 1 & | & 1 \\ 0 & 0 & 1 & -1 & -2 & | & 1 \\ 0 & 0 & 0 & 1 & -2 & | & 0 \end{bmatrix} \xrightarrow[d_2 \to d_2 + d_3]{d_1 \to d_1 - 2d_3} \begin{bmatrix} 1 & 2 & 0 & 0 & 5 & | & 1 \\ 0 & 0 & 1 & 0 & -4 & | & 1 \\ 0 & 0 & 0 & 1 & -2 & | & 0 \end{bmatrix}$$

Khi đó $x_1$, $x_3$, $x_4$ là ẩn chính còn $x_2$, $x_5$ là ẩn tự do. Ta có $x_1 = 1 - 2x_2 - 5x_5$, $x_3 = 1 + 4x_5$, $x_4 = 2x_5$. □

**2.2.7 Thí dụ. (OSV99)** Giải hệ phương trình

$$\begin{cases} x_1 + 2x_2 + 3x_3 + \cdots + nx_n = 1 \\ x_2 + 2x_3 + 3x_4 + \cdots + nx_1 = 2 \\ \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_n + 2x_1 + 3x_2 + \cdots + nx_{n-1} = n \end{cases}$$

*Giải.* Biến đổi sơ cấp ma trận bổ sung của hệ phương trình

$$\bar{A} = \left[\begin{array}{ccccc|c} 1 & 2 & 3 & \dots & n & 1 \\ n & 1 & 2 & \dots & n-1 & 2 \\ n-1 & n & 1 & \dots & n-2 & 3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 2 & 3 & 4 & \dots & 1 & n \end{array}\right]$$

$$\xrightarrow[d_k \to d_k - d_{k-1}; k=2,\dots,n]{d_1 \to \frac{2}{n(n+1)}(d_1 + \cdots + d_n)} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & \dots & 1 & 1 \\ n-1 & -1 & -1 & \dots & -1 & 1 \\ -1 & n-1 & -1 & \dots & -1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -1 & -1 & -1 & \dots & -1 & 1 \end{array}\right]$$

$$\xrightarrow{d_k \to d_k + d_1; k=2,\dots,n} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & \dots & 1 & 1 \\ n & 0 & 0 & \dots & 0 & 2 \\ 0 & n & 0 & \dots & 0 & 2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 2 \end{array}\right]$$

Do đó, $x_k = \frac{2}{n}$ với $k = 1, \dots, n-1$ và $x_n = 1 - \frac{2(n-1)}{n} = \frac{2-n}{n}$. $\square$

**2.2.8 Nhận xét.** Phương pháp Gauss-Jordan cơ bản và hiệu quả để giải hệ phương trình $A\mathbf{x} = \mathbf{b}$; đặc biệt là nó khá đơn giản để lập trình trên các ngôn ngữ tính toán để giải được hệ một cách nhanh chóng và tổng quát. Tuy nhiên nó không phải là phương pháp tốt nhất cho mọi hệ phương trình. Có một số phương trình khi áp dụng phương pháp Gauss-Jordan khá khó khăn và vất vả để giải được hệ, nhưng nếu ta khai thác được những đặc trưng riêng biệt của hệ phương trình đã cho thì giải được nó khá dễ dàng. Sau đây là một ví dụ thể hiện điều đó.

**2.2.9 Thí dụ. (OSV05)** Giải hệ phương trình

$$\begin{cases} x_1 + x_2 + \cdots + x_{n-1} + x_n = \dfrac{a}{2004} \\ \qquad x_2 + \cdots + x_{n-1} + x_n = \dfrac{a + x_1}{2005^2 - 1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\qquad\qquad x_n = \dfrac{a + x_1 + \cdots + x_{n-1}}{2005^n - 1} \end{cases}$$

*Giải.* Với $2 \le i \le n$, cộng thêm biểu thức $x_1 + \cdots + x_{i-1}$ vào hai vế phương trình thứ $i$ ta được

$$\begin{aligned} & x_1 + \cdots + x_{i-1} + x_i + \cdots + x_n \\ & \qquad = \frac{a + x_1 + \cdots + x_{i-1}}{2005^i - 1} + x_1 + \cdots + x_{i-1}. \end{aligned}$$

Từ đó kết hợp với phương trình thứ nhất ta được

$$\frac{a}{2004} = \frac{a}{2005^i - 1} + (x_1 + \cdots + x_{i-1})\Big(1 + \frac{1}{2005^i - 1}\Big)$$

suy ra

$$x_1 + \cdots + x_{i-1} = \frac{a}{2005^i}\Big(\frac{2005^i - 2005}{2004}\Big).$$

Vậy với $i = 2, 3, \ldots, n-1$, ta được

$$\begin{aligned} x_i &= (x_1 + \cdots + x_{i-1} + x_i) - (x_1 + \cdots + x_{i-1}) \\ &= \frac{a}{2005^{i+1}}\Big(\frac{2005^{i+1} - 2005}{2004}\Big) - \frac{a}{2005^i}\Big(\frac{2005^i - 2005}{2004}\Big) \\ &= \frac{a}{2005^i}. \\ x_n &= (x_1 + \cdots + x_{n-1} + x_n) - (x_1 + \cdots + x_{n-1}) \\ &= \frac{a}{2004} - \frac{a}{2005^n}\Big(\frac{2005^n - 2005}{2004}\Big) \\ &= \frac{a}{2004 \cdot 2005^{n-1}}. \end{aligned}$$

Để tìm $x_1$ lấy phương trình thứ nhất trừ phương trình thứ hai

$$x_1 = \frac{a}{2004} - \frac{a + x_1}{2005^2 - 1} = \frac{2005a}{2005^2 - 1} - \frac{x_1}{2005^2 - 1}$$

suy ra

$$x_1 = \frac{a}{2005}.$$

Vậy nghiệm của hệ đã cho là

$$x_i = \frac{a}{2005^i} \ (i = 1, 2, \dots, n-1); \quad x_n = \frac{a}{2004 \cdot 2005^{n-1}}. \qquad \square$$

## 2.3 Hệ phương trình tuyến tính thuần nhất

Xét hệ phương trình tuyến tính thuần nhất

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0. \end{cases}$$

Hệ này luôn có 1 nghiệm là $(0, 0, \dots, 0)$ được gọi là **nghiệm tầm thường** (trivial solution). Nghiệm (nếu có) $(c_1, c_2, \dots, c_n) \neq (0, 0, \dots, 0)$ của hệ được gọi là **nghiệm không tầm thường** (nontrivial solution).

**2.3.1 Định lý.** *Hệ thuần nhất trên có nghiệm không tầm thường khi và chỉ khi hạng của hệ nhỏ hơn số ẩn, tức là* $\operatorname{rank} A < n$.

*Chứng minh.* Định lý được suy trực tiếp từ kết quả sau định lý Kronecker-Capelli ở phần 2.1.11. $\square$

**2.3.2 Hệ quả.** *Cho hệ phương trình tuyến tính thuần nhất có số phương trình bằng số ẩn với ma trận hệ số là $A$. Khi đó,*

**(1)** *Hệ có nghiệm duy nhất tầm thường khi và chỉ khi* $\det A \neq 0$

**(2)** *Hệ có nghiệm không tầm thường khi và chỉ khi* $\det A = 0$.

*Chứng minh.* Dành cho bạn đọc xem như bài tập. $\square$

**2.3.3 Thí dụ. (OSV12)** Giải hệ phương trình tuyến tính

$$\begin{cases} x_1 = 2(4x_1 + 3x_2 + 2x_3 + x_4) \\ x_2 = 3(x_1 + 4x_2 + 3x_3 + 2x_4) \\ x_3 = 4(2x_1 + x_2 + 4x_3 + 3x_4) \\ x_4 = 5(3x_1 + 2x_2 + x_3 + 4x_4) \end{cases}$$

*Giải.* Hệ phương trình tuyến tính đã cho được viết lại

$$\begin{cases} 7x_1 + 6x_2 + 4x_3 + 2x_4 = 0 \\ 3x_1 + 11x_2 + 9x_3 + 6x_4 = 0 \\ 8x_1 + 4x_2 + 15x_3 + 12x_4 = 0 \\ 15x_1 + 10x_2 + 5x_3 + 19x_4 = 0 \end{cases}$$

Ta tính được định thức của ma trận hệ số của hệ phương trình tuyến tính thuần nhất trên

$$\det(A) = \begin{vmatrix} 7 & 6 & 4 & 2 \\ 3 & 11 & 9 & 6 \\ 8 & 4 & 15 & 12 \\ 15 & 10 & 5 & 19 \end{vmatrix} = 13269$$

Vậy hệ đã cho chỉ có nghiệm tầm thường $x_1 = x_2 = x_3 = x_4 = 0$. □

**2.3.4 Thí dụ. (OSV95)** Cho các số $a_{ij} + a_{ji} = 0$ với mọi $i, j = 1, 2, \ldots, n$ trong đó $n$ là số tự nhiên lẻ. Chứng minh rằng hệ phương trình

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases}$$

có nghiệm không tầm thường.

*Giải.* Gọi $A$ là ma trận hệ số của hệ phương trình đã cho. Theo giả thiết ta có $A^t = -A$. Do đó

$$\det(A) = \det(A^t) = \det(-A) = (-1)^n \det(A) = -\det(A),$$

suy ra $\det(A) = 0$. Vậy hệ phương trình đã cho có nghiệm không tầm thường. □

**2.3.5 Thí dụ. (OSV07)** Cho $A = (a_{ij})$ là ma trận vuông cấp $n$ có các tính chất sau $a_{ii} = 2007$, $a_{ij} = 4$ hoặc 20 khi $i \neq j$ với $i, j = 1, 2, \ldots, n$. Giả hệ phương trình đại số tuyến tính $A\mathbf{x} = \mathbf{0}$.

*Giải.* Ta nhận thấy $A \equiv I \pmod 2$. Từ đó suy ra $\det(A) \equiv 1 \pmod 2$, cho nên $\det(A) \neq 0$. Vậy hệ chỉ có nghiệm tầm thường. $\square$

**2.3.6 Thí dụ. (OSV11)** Tìm điều kiện của các tham số $a, b, c, d$ để hệ phương trình sau có nghiệm duy nhất

$$\begin{cases} (1+a)x_1 + (1+a^2)x_2 + (1+a^3)x_3 + (1+a^4)x_4 = 0 \\ (1+b)x_1 + (1+b^2)x_2 + (1+b^3)x_3 + (1+b^4)x_4 = 0 \\ (1+c)x_1 + (1+c^2)x_2 + (1+c^3)x_3 + (1+c^4)x_4 = 0 \\ (1+d)x_1 + (1+d^2)x_2 + (1+d^3)x_3 + (1+d^4)x_4 = 0 \end{cases}$$

*Giải.* Ta tính định thức ma trận hệ số

$$\begin{aligned}
\det(A) &= \begin{vmatrix} 1+a & 1+a^2 & 1+a^3 & 1+a^4 \\ 1+b & 1+b^2 & 1+b^3 & 1+b^4 \\ 1+c & 1+c^2 & 1+c^3 & 1+d^4 \\ 1+d & 1+d^2 & 1+d^3 & 1+d^4 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1+a & 1+a^2 & 1+a^3 & 1+a^4 \\ 0 & 1+b & 1+b^2 & 1+b^3 & 1+b^4 \\ 0 & 1+c & 1+c^2 & 1+c^3 & 1+d^4 \\ 0 & 1+d & 1+d^2 & 1+d^3 & 1+d^4 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & a & a^2 & a^3 & a^4 \\ -1 & b & b^2 & b^3 & b^4 \\ -1 & c & c^2 & c^3 & c^4 \\ -1 & d & d^2 & d^3 & d^4 \end{vmatrix} \\
&= \begin{vmatrix} 2 & 1 & 1 & 1 & 1 \\ 0 & a & a^2 & a^3 & a^4 \\ 0 & b & b^2 & b^3 & b^4 \\ 0 & c & c^2 & c^3 & c^4 \\ 0 & d & d^2 & d^3 & d^4 \end{vmatrix} + (-1)^5 \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & a & a^2 & a^3 & a^4 \\ 1 & b & b^2 & b^3 & b^4 \\ 1 & c & c^2 & c^3 & c^4 \\ 1 & d & d^2 & d^3 & d^4 \end{vmatrix} \\
&= 2abcd(b-a)(c-a)(d-a)(c-b)(d-b)(d-c) \\
&\quad - (a-1)(b-1)(c-1)(d-1)\times \\
&\qquad \times (b-a)(c-a)(d-a)(c-b)(d-b)(d-c).
\end{aligned}$$

Vậy để hệ có nghiệm duy nhất các số $a, b, c, d$ đôi một khác nhau và $2abcd \neq (a-1)(b-1)(c-1)(d-1)$. $\square$

**2.3.7 Định lý.** *Trong hệ phương trình tuyến tính thuần nhất, tổng hai nghiệm của hệ và tích của một số với nghiệm của hệ cũng là các nghiệm của hệ.*

*Chứng minh.* Xét hệ phương trình tuyến tính thuần nhất $A\mathbf{x} = \mathbf{0}$. Giả sử $\mathbf{a}$ và $\mathbf{b}$ là nghiệm của phương trình và $\lambda \in \mathbb{K}$ tùy ý. Ta có $A\mathbf{a} = \mathbf{0}$ và $A\mathbf{b} = \mathbf{0}$. Do đó, $A(\mathbf{a}+\mathbf{b}) = A\mathbf{a} + A\mathbf{b} = \mathbf{0}$ và $A(\lambda\mathbf{a}) = \lambda A\mathbf{a} = \mathbf{0}$. $\square$

**2.3.8 Định lý.** *Cho hệ phương trình $n$ ẩn dạng ma trận $A\mathbf{x} = \mathbf{b}$, và $S_0$ là tập hợp nghiệm của hệ phương trình tuyến tính thuần nhất tương ứng $A\mathbf{x} = \mathbf{0}$. Nếu $\mathbf{c}$ là một nghiệm của phương trình $A\mathbf{x} = \mathbf{b}$ thì $S = \{\mathbf{a}+\mathbf{c} : \mathbf{a} \in S_0\}$ là tập hợp nghiệm của hệ phương trình $A\mathbf{x} = \mathbf{b}$.*

*Chứng minh.* Theo giả thiết ta có $A\mathbf{c} = \mathbf{b}$ và $A\mathbf{a} = \mathbf{0}$ với mọi $\mathbf{a} \in S_0$. Với mọi $\mathbf{a} \in S_0$, ta có $A(\mathbf{a}+\mathbf{c}) = A\mathbf{a} + A\mathbf{c} = \mathbf{b}$, cho nên $\mathbf{a}+\mathbf{c}$ là một nghiệm của phương trình $A\mathbf{x} = \mathbf{b}$. Ngược lại, giả sử $\mathbf{x}_0$ là một nghiệm khác của hệ phương trình $A\mathbf{x} = \mathbf{b}$. Khi đó $A(\mathbf{x}_0 - \mathbf{c}) = A\mathbf{x}_0 - A\mathbf{c} = \mathbf{0}$, nghĩa là $\mathbf{x}_0 - \mathbf{c}$ là một nghiệm của hệ $A\mathbf{x} = \mathbf{0}$. Vậy $S = \{\mathbf{a}+\mathbf{c} : \mathbf{a} \in S_0\}$ là tập hợp nghiệm của hệ phương trình $A\mathbf{x} = \mathbf{b}$. $\square$

**2.3.9 Thí dụ.** Xét hệ phương trình tuyến tính tổng quát trên $\mathbb{R}$

$$\begin{cases} x + 2y + z - t = 3 \\ 2x + 5y + 3z + t = 11 \\ 3x + 7y + 4z = 14. \end{cases}$$

Trước hết ta giải hệ phương trình tuyến tính thuần nhất tương ứng với hệ đã cho

$$\begin{cases} x + 2y + z - t = 0 \\ 2x + 5y + 3z + t = 0 \\ 3x + 7y + 4z = 0. \end{cases}$$

Ta biến đổi sơ cấp trên dòng của ma trận hệ số

$$\begin{bmatrix} 1 & 2 & 1 & -1 \\ 2 & 5 & 3 & 1 \\ 3 & 7 & 4 & 0 \end{bmatrix} \xrightarrow[d_3 \to d_3 - 3d_1]{d_2 \to d_2 - 2d_2} \begin{bmatrix} 1 & 2 & 1 & -1 \\ 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \end{bmatrix}$$

$$\xrightarrow[d_3 \to d_3 - d_2]{d_1 \to d_1 - 2d_2} \begin{bmatrix} 1 & 0 & -1 & -7 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Vậy nghiệm của hệ thuần nhất là $x = \alpha + 7\beta, y = -\alpha - 3\beta, z = \alpha, t = \beta$ với $\alpha, \beta \in \mathbb{R}$.

Mặt khác ta có thể kiểm tra được $\mathbf{c} = (1, 1, 1, 1)$ là một nghiệm của hệ phương trình tuyến tính đã cho. Vậy tập hợp nghiệm của hệ phương trình ban đầu là $S = \{(1 + \alpha + 7\beta, 1 - \alpha - 3\beta, 1 + \alpha, 1 + \beta) : \alpha, \beta \in \mathbb{R}\}$. □

**2.3.10 Thí dụ. (OSV02)** Cho hệ phương trình tuyến tính có 10 phương trình và 11 ẩn số. Biết rằng (1) bộ số $(1992, 1993, \ldots, 2002)$ là một nghiệm của hệ phương trình đã cho; (2) Khi xóa cột thứ $j$ trong ma trận hệ số của hệ đã cho thì ta được một ma trận vuông có định thức đúng bằng $j$ với $j = 1, 2, \ldots, 11$. Hãy tìm tất cả các nghiệm của hệ phương trình đã cho.

*Giải.* Ta nhận thấy hệ phương trình đã cho có hạng của ma trận hệ số bằng 10, cho nên nó có nghiệm phụ thuộc một tham số. Gọi các biến là $\mathbf{x} = (x_1, x_2, \ldots, x_{11})$ và ma trận hệ số là $A$. Khi đó ta có hệ phương trình tuyến tính thuần nhất tương ứng là $A\mathbf{x} = \mathbf{0}$, suy ra $A'\mathbf{x}' = \mathbf{b}$ với $A'$ là ma trận thu được từ $A$ bỏ đi cột cuối, $\mathbf{x}' = (x_1, \ldots, x_{10})$ và $\mathbf{b} = (-a_{1,11}x_{11}, -a_{2,11}x_{11}, \ldots, -a_{10,11}x_{11})$. Từ giả thiết ta có $\det(A') = 11$ và theo công thức Cramer ta được $x_1 = \frac{1}{11}x_{11}$, $x_2 = -\frac{2}{11}x_{11}$, ..., $x_{10} = -\frac{10}{11}x_{11}$. Vậy hệ đã cho có nghiệm là

$$\mathbf{x} = (1992, 1993, \ldots, 2001, 2002) + (\alpha, -2\alpha, \ldots, -10\alpha, 11\alpha), \quad \alpha \in \mathbb{R}.$$

□

## Bài tập

**1**) Giải hệ phương trình tuyến tính

**(a)** $\begin{cases} 6x_1 + 4x_2 + 5x_3 + 2x_4 + 3x_5 = 1 \\ 3x_1 + 2x_2 + 4x_3 + x_4 + 2x_5 = 3 \\ 3x_1 + 2x_2 - 2x_3 + x_4 = -7 \\ 9x_1 + 6x_2 + x_3 + 3x_4 + 2x_5 = 2 \end{cases}$

**(b)** $\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 7 \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 = -2 \\ x_2 + 2x_3 + 2x_4 + 6x_5 = 23 \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 = 12 \end{cases}$

**(c)** $\begin{cases} 2x_1 + x_2 - x_3 - x_4 + x_5 = 1 \\ x_1 - x_2 + x_3 + x_4 - 2x_5 = 0 \\ 3x_1 + 3x_2 - 3x_3 - 3x_4 + 4x_5 = 2 \\ 4x_1 + 5x_2 - 5x_3 - 5x_4 + 7x_5 = 3 \end{cases}$

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

**(e)** $\begin{cases} 3x_1 + x_2 - 2x_3 + x_4 - x_5 = 1 \\ 2x_1 - x_2 + 7x_3 - 3x_4 + 5x_5 = 2 \\ x_1 + 3x_2 - 2x_3 + 5x_4 - 7x_5 = 3 \\ 3x_1 - 2x_2 + 7x_3 - 5x_4 + 8x_5 = 3 \end{cases}$

**2**) Giải và biện luận hệ phương trình theo $m$:

**(a)** $\begin{cases} mx_1 + x_2 + x_3 = 1 \\ x_1 + mx_2 + x_3 = m \\ x_1 + x_2 + mx_3 = m^2 \end{cases}$

**(b)** $\begin{cases} (1+m)x_1 + x_2 + x_3 = 1 \\ x_1 + (1+m)x_2 + x_3 = m \\ x_1 + x_2 + (1+m)x_3 = 2 \end{cases}$

**(c)** $\begin{cases} x + my + m^2 z = 1 \\ x + 2y + 4z = 2 \\ x + 3y + 9z = 3 \end{cases}$

**(d)** $\begin{cases} x + 2y - z + t = m \\ 2x + 5y - 2z + 2t = 2m + 1 \\ 3x + 7y - 3z + 3t = 1 \end{cases}$

**(e)** $\begin{cases} x - 2y + z + 2t = m \\ x + y - z + t = 2m + 1 \\ x + 7y - 5z - 5 = -m \end{cases}$ **(f)** $\begin{cases} (m+3)x + y + 2z = m \\ mx + (m-1)y + z = 2m \\ 3(m+1)x + my + (m+3)z = 3 \end{cases}$

**(g)** $\begin{cases} (3m-1)x + 2my + (3m+1)z = 1 \\ 2mx + 2my + (3m+1)z = m \\ (m+1)x + (m+1)y + 2(m+1)z = m^2 \end{cases}$

**(h)** $\begin{cases} x - y + 2z - 2t = 0 \\ 2x + y - z + t = 3 \\ 3x + z - t = 3 \\ 5x + y = m \end{cases}$ **(i)** $\begin{cases} 2x - y + z - 2t + 3u = 3 \\ x + y - z - t + u = 1 \\ 3x + y + z - 3t + 4u = 6 \\ 5x + 2z - 5t + 7u = 9 - m \end{cases}$

**(j)** $\begin{cases} 2x + y - z + 2t = 4 \\ x - y + z + 2t = 3 \\ 2x + 2y - 2z + t = 3 \\ x + y - 2z + t = m \end{cases}$ **(k)** $\begin{cases} 2x - y + z + t = 1 \\ x + 2y - z + 4t = 2 \\ x + 7y - 4z + 11t = m \\ 4x + 8y - 4z + 16t = m + 1 \end{cases}$

**3)** Chứng minh Hệ quả 2.3.2.

**4)** Giải hệ phương trình tuyến tính thuần nhất

$$\begin{cases} x_1 + 2x_2 + 3x_3 + \cdots + nx_n = 0 \\ x_2 + 2x_3 + 3x_4 + \cdots + nx_1 = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n + 2x_1 + 3x_2 + \cdots + nx_{n-1} = 0 \end{cases}$$

**5)** Cho các hệ phương trình tuyến tính sau

**(a)** $\begin{cases} x + 3y + 5z - 4t = 1 \\ x + 3y + 2z - 2t + u = -1 \\ x - 2y + z - t - u = 3 \\ x - 4y + z + t - u = 3 \\ x + 2y + z - t + u = -1 \end{cases}$ **(b)** $\begin{cases} x + 2y - 3t + 2u = 1 \\ x - y - 3z + t - 3u = 2 \\ 2x - 3y + 4z - 5t + 2u = 7 \\ 9x - 9y + 7z - 16t + 2u = 25 \end{cases}$

Ứng với mỗi hệ, hãy giải hệ thuần nhất tương ứng, sau đó tìm một nghiệm riêng để suy ra nghiệm tổng quát của hệ đã cho.

**6 )** Trong mặt phẳng tọa độ cho ba điểm đôi một phân biệt $A(a_1, a_2)$, $B(b_1, b_2)$, $C(c_1, c_2)$ và ba đường thẳng đôi một khác nhau $(d_1) : \alpha_1 x + \beta_1 y + c_1$, $(d_2) : \alpha_2 x + \beta_2 y + c_2 = 0$, $(d_3) : \alpha_3 x + \beta_3 y + c_3 = 0$.

**(a)** Tìm điều kiện để $A, B, C$ thẳng hàng.

**(b)** Tìm điều kiện để $(d_1), (d_2), (d_3)$ đồng quy hay song song.

**7 )** Trong không gian tọa độ cho bốn điểm đôi một phân biệt $A(a_1, a_2, a_3)$, $B(b_1, b_2, b_3)$, $C(c_1, c_2, c_3)$, $D(d_1, d_2, d_3)$ và ba mặt phẳng đôi một khác nhau $(P_1) : \alpha_1 x + \beta_1 y + \gamma_1 z + \delta_1 = 0$, $(P_2) : \alpha_2 x + \beta_2 y + \gamma_2 z + \delta_2 = 0$, $(P_3) : \alpha_3 x + \beta_3 y + \gamma_3 z + \delta_3 = 0$.

**(a)** Tìm điều kiện để $A, B, C, D$ đồng phẳng

**(b)** Tìm điều kiện để $(P_1), (P_2), (P_3)$ đồng quy hay song song.

# Chương IV

# Không gian vector

## § 1 Khái niệm về không gian vector

### 1.1 Định nghĩa không gian vector

Cho $V$ là một tập hợp tùy ý khác rỗng. $V$ được gọi là một **không gian vector** (vector space) trên $\mathbb{K}$ hay $\mathbb{K}$**-không gian vector** nếu trong $V$ có hai phép toán:

**(i)** Phép cộng: tức là với hai phần tử bất kì $\mathbf{x}, \mathbf{y} \in V$ thì xác định duy nhất một phần tử $\mathbf{x} + \mathbf{y} \in V$

**(ii)** Phép nhân với vô hướng: tức là cho một số $\alpha \in \mathbb{K}$ và một phần tử $\mathbf{x} \in V$ ta xác định duy nhất một phần tử $\alpha\mathbf{x} \in V$

và thỏa tám tiên đề sau:

**V1.** Phép cộng có tính chất kết hợp: $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$ với mọi $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$

**V2.** Phép cộng có tính chất giao hoán: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ với mọi $\mathbf{x}, \mathbf{y} \in V$

**V3.** Phép cộng có phần tử không: tức là tồn tại $\mathbf{0} \in V$ sao cho $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$

**V4.** Với mỗi $\mathbf{x} \in V$ đều có phần tử đối: tức là tồn tại $\mathbf{x}' \in V$ sao cho $\mathbf{x} + \mathbf{x}' = \mathbf{x}' + \mathbf{x} = \mathbf{0}$

**V5.** $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$ với mọi $\alpha \in \mathbb{K}$, $\mathbf{x}, \mathbf{y} \in V$

**V6.** $(\alpha+\beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$ với mọi $\alpha, \beta \in \mathbb{K}$, $\mathbf{x} \in V$

**V7.** $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$ với mọi $\alpha, \beta \in \mathbb{K}$, $\mathbf{x} \in V$

**V8.** $1\mathbf{x} = \mathbf{x}$ với mọi $\mathbf{x} \in V$.

Trong trường hợp $V$ là một không gian vector ta gọi mỗi phần tử của nó là một **vector**.

**1.1.1 Nhận xét.** Nhờ có tiên đề V1. nên ta viết $\mathbf{x}+\mathbf{y}+\mathbf{z}$ thay cho $(\mathbf{x}+\mathbf{y})+\mathbf{z}$ hay $\mathbf{x}+(\mathbf{y}+\mathbf{z})$. Nghĩa là ta không cần dấu ngoặc trong việc tính tổng của 3 vector. Điều này được tổng quát hóa cho tổng của nhiều vector. Chẳng hạn, xét tổng $\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n$ hay $\sum_{i=1}^{n} \mathbf{x}_i$ với $n \geq 2$. Tổng này không phụ thuộc vào thứ tự lấy tổng.

Sau này chúng ta sẽ chứng minh phần tử đối $\mathbf{x}'$ của $\mathbf{x}$ trong tiên đề V4. là duy nhất và ta dùng kí hiệu $-\mathbf{x}$ để chỉ phần tử đối của $\mathbf{x}$.

Chúng ta nhấn mạnh rằng, trong định nghĩa $\mathbb{K}$-không gian vector $V$, các "vector" của $V$ có thể có bản chất tùy ý, "phép cộng các vector" và "phép nhân với vô hướng" có thể được cho bởi những quy tắc tùy ý; điều quan trọng là 8 tiên đề trong định nghĩa trên phải thỏa mãn.

## 1.2 Một số mô hình không gian vector

**1.2.1** Xét tập $\mathbb{K}^n = \{(a_1, a_2, \ldots, a_n) : a_i \in \mathbb{K}, i = 1, \ldots, n\}$. Ta định nghĩa hai phép toán cộng và nhân vô hướng trên $\mathbb{R}^n$ như sau: với $\mathbf{a} = (a_1, a_2, \ldots, a_n)$, $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, với mọi $\alpha \in \mathbb{K}$ ta có $\mathbf{a}+\mathbf{b} = (a_1+b_1, a_2+b_2, \ldots, a_n+b_n)$ và $\alpha\mathbf{a} = (\alpha a_1, \alpha a_2, \ldots, \alpha a_n)$. Khi đó $\mathbb{K}^n$ là một không gian vector trên $\mathbb{K}$ với vector không là $\mathbf{0} = (0, 0, \ldots, 0)$ gồm $n$ số không, vector đối của $(a_1, a_2, \ldots, a_n)$ là $(-a_1, -a_2, \ldots, -a_n)$.

**1.2.2** Tập hợp $\mathcal{M}_{m,n}(\mathbb{K})$ các ma trận cấp $m \times n$ trên $\mathbb{K}$ cùng với phép cộng hai ma trận và phép nhân một số thuộc $\mathbb{K}$ với một ma trận là một không gian vector trên $\mathbb{K}$, gọi là không gian các ma trận (xem Định lý 2.2.1). Mỗi ma trận cấp $m \times n$ trên $\mathbb{K}$ là một "vector"; "vector không" là ma trận không $O_{m\times n}$, "vector đối" của $A \in \mathcal{M}_{m,n}(\mathbb{K})$ là ma trận đối $-A$.

**1.2.3** Tập $\mathbb{K}[x] = \{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n : n \in \mathbb{N}, a_i \in \mathbb{K}\}$ (tập các đa thức một biến $x$ với các hệ số trên $\mathbb{K}$) với phép cộng hai đa thức và phép nhân một đa thức với một số thuộc $\mathbb{K}$ là một không gian vector trên $\mathbb{K}$. Mỗi đa thức thuộc $\mathbb{K}[x]$ là một "vector", "vector không" là đa thức hằng không, "vector đối" của $p(x) \in \mathbb{K}[x]$ là đa thức đối $-p(x)$.

**1.2.4** Xét tập hợp $\mathcal{C}_{[a,b]}$ các hàm liên tục trên đoạn $[a, b]$. Trên $\mathcal{C}_{[a,b]}$ ta định nghĩa phép cộng và phép nhân với số thực xác định như sau $(f + g)(x) = f(x) + g(x)$ và $(\lambda f)(x) = \lambda f(x)$ với mọi $x \in [a, b]$ và mọi $f, g \in \mathcal{C}_{[a,b]}$. Khi đó, ta có thể chứng minh được $\mathcal{C}_{[a,b]}$ là một không gian vector trên $\mathbb{R}$.

Thay đoạn $[a, b]$ bởi $(a, b)$, $[a, b)$ $(a, b]$, $(-\infty, \infty)$, ... ta cũng có khẳng định tương tự.

Thay vì xét các hàm liên tục, ta xét các hàm khả vi thì cũng thu được các không gian vector thực tương ứng.

## 1.3 Một số tính chất đơn giản

**1.3.1 Định lý.** **(i)** *Vector không là duy nhất.*

**(ii)** *Vector đối là duy nhất.*

**(iii)** *Phép cộng vector có tính giản ước: nghĩa là từ đẳng thức* $\mathbf{x}+\mathbf{z} = \mathbf{y}+\mathbf{z}$ *ta suy ra* $\mathbf{x} = \mathbf{y}$.

**(iv)** $\alpha\mathbf{0} = \mathbf{0}$ *và* $0\mathbf{x} = \mathbf{0}$ *với mọi* $\alpha \in \mathbb{K}$, $\mathbf{x} \in V$.

**(v)** $\alpha\mathbf{x} = \mathbf{0}$ *suy ra* $\alpha = 0$ *hoặc* $\mathbf{x} = \mathbf{0}$.

**(vi)** $-(\lambda\mathbf{x}) = (-\lambda)\mathbf{x} = \lambda(-\mathbf{x})$ *với mọi* $\lambda \in \mathbb{K}$ *và mọi* $\mathbf{x} \in V$.

*Chứng minh.*

**(i)** Giả sử $\mathbf{0}'$ phần tử không khác trong tiên đề V3. Khi đó, ta có

$$\begin{aligned} \mathbf{0} &= \mathbf{0} + \mathbf{0}' && \text{do } \mathbf{0}' \text{ là phần tử không trong tiên đề V3} \\ &= \mathbf{0}' && \text{do } \mathbf{0} \text{ là phần tử không trong tiên đề V3} \end{aligned}$$

**(ii)** Giả sử $\mathbf{x}''$ là vector đối khác của $\mathbf{x}$ trong tiên đề V4. Khi đó ta có

$$\mathbf{x}'' \overset{V3}{=} \mathbf{x}'' + \mathbf{0} = \mathbf{x}'' + (\mathbf{x} + \mathbf{x}') \overset{V1}{=} (\mathbf{x}'' + \mathbf{x}) + \mathbf{x}' = \mathbf{0} + \mathbf{x}' \overset{V3}{=} \mathbf{x}.$$

**(iii)** Từ đẳng thức $\mathbf{x}+\mathbf{z} = \mathbf{y}+\mathbf{z}$ suy ra $(\mathbf{x}+\mathbf{z})+(-\mathbf{z}) = (\mathbf{y}+\mathbf{z})+(-\mathbf{z})$. Theo tiên đề V1 ta có $\mathbf{x} + (\mathbf{z} + (-\mathbf{z})) = \mathbf{y} + (\mathbf{z} + (-z))$, và theo tiên đề V4 ta được $\mathbf{x} + \mathbf{0} = \mathbf{y} + \mathbf{0}$. Theo tiên đề V3 ta có $\mathbf{x} = \mathbf{y}$.

**(iv)** Ta có

$$\begin{aligned}
\alpha\mathbf{0} &\overset{V3,V4}{=} \alpha\mathbf{0} + (\alpha\mathbf{x} + (-\alpha\mathbf{x})) \\
&\overset{V1,V5}{=} \alpha(\mathbf{0} + \mathbf{x}) + (-\alpha\mathbf{x}) \\
&\overset{V3}{=} \alpha\mathbf{x} + (-\alpha\mathbf{x}) \\
&\overset{V4}{=} \mathbf{0}.
\end{aligned}$$

Ta cũng có

$$\begin{aligned}
0\mathbf{x} &\overset{V3,V4}{=} 0\mathbf{x} + (\mathbf{x} + (-\mathbf{x})) \\
&\overset{V1}{=} (0\mathbf{x} + 1\mathbf{x}) + (-\mathbf{x}) \\
&\overset{V6}{=} (0 + 1)\mathbf{x} + (-\mathbf{x}) \\
&\overset{V8}{=} \mathbf{x} + (-\mathbf{x}) \\
&\overset{V4}{=} \mathbf{0}.
\end{aligned}$$

**(v)** Giả sử $\alpha\mathbf{x} = \mathbf{0}$ và $\alpha \neq 0$. Ta có $\mathbf{x} \overset{V8}{=} 1\mathbf{x} = (\frac{1}{\alpha}\alpha)\mathbf{x} \overset{V7}{=} \frac{1}{\alpha}(\alpha\mathbf{x}) = \frac{1}{\alpha}\mathbf{0} \overset{\text{(iv)}}{=} \mathbf{0}$.

**(vi)** Ta có $\lambda\mathbf{x} + (-\lambda)\mathbf{x} \overset{V6}{=} (\lambda + (-\lambda))\mathbf{x} = 0\mathbf{x} \overset{\text{(iv)}}{=} \mathbf{0}$ và $\lambda\mathbf{x} + \lambda(-\mathbf{x}) \overset{V5}{=} \lambda(\mathbf{x} + (-\mathbf{x})) \overset{V4}{=} \lambda\mathbf{0} \overset{\text{(iv)}}{=} \mathbf{0}$. Do phần tử đối của một phần tử là duy nhất, ta suy ra được $(-\lambda)\mathbf{x} = \lambda(\mathbf{x}) = -(\lambda\mathbf{x})$. □

**1.3.2 Nhận xét.** Trong 8 tiên đề của định nghĩa không gian vector, tiên đề V2 về phép cộng giao hoán có thể suy ra từ các tiên đề còn lại. Thật vậy, với mọi $\mathbf{x}, \mathbf{y} \in V$, ta có

$$\begin{aligned}(1+1)(\mathbf{x}+\mathbf{y}) &\overset{V5}{=} (1+1)\mathbf{x} + (1+1)\mathbf{y}\\ &\overset{V6,V8}{=} (\mathbf{x}+\mathbf{x}) + (\mathbf{y}+\mathbf{y})\\ &\overset{V1}{=} \mathbf{x} + (\mathbf{x}+\mathbf{y}) + \mathbf{y}.\end{aligned}$$

Mặt khác, ta có

$$(1+1)(\mathbf{x}+\mathbf{y}) \overset{V6,V8}{=} (\mathbf{x}+\mathbf{y}) + (\mathbf{x}+\mathbf{y}) \overset{V1}{=} \mathbf{x} + (\mathbf{y}+\mathbf{x}) + \mathbf{y}.$$

Suy ra $\mathbf{x} + (\mathbf{x}+\mathbf{y}) + \mathbf{y} = \mathbf{x} + (\mathbf{y}+\mathbf{x}) + \mathbf{y}$. Theo tiên đề V4 tồn tại $-\mathbf{x}$ và $-\mathbf{y}$. Suy ra

$$-\mathbf{x} + [\mathbf{x} + (\mathbf{x}+\mathbf{y}) + \mathbf{y}] + (-\mathbf{y}) = -\mathbf{x} + [\mathbf{x} + (\mathbf{y}+\mathbf{x}) + \mathbf{y}] + (-\mathbf{y})$$

và từ tiên đề V1 ta được

$$(-\mathbf{x}+\mathbf{x}) + (\mathbf{x}+\mathbf{y}) + (\mathbf{y}+(-\mathbf{y})) = (-\mathbf{x}+\mathbf{x}) + (\mathbf{y}+\mathbf{x}) + (\mathbf{y}+(-\mathbf{y}))$$

hay

$$\mathbf{0} + (\mathbf{x}+\mathbf{y}) + \mathbf{0} = \mathbf{0} + (\mathbf{y}+\mathbf{x}) + \mathbf{0}.$$

Dùng tiên đề V3 ta có được $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$.

**1.3.3 *Chú ý.*** Để cho gọn, về sau thay vì viết $\mathbf{x} + (-\mathbf{y})$ ta viết $\mathbf{x} - \mathbf{y}$ và gọi là hiệu của $\mathbf{x}$ và $\mathbf{y}$.

## Bài tập

**1**) Xét xem $\mathbb{R}^2$ có phải là không gian vector với phép cộng và phép nhân vô hướng sau với mọi $(a, b), (c, d) \in \mathbb{R}^2$ và mọi $\alpha \in \mathbb{R}$

**(a)** $(a, b) + (c, d) = (a + c, b + d)$ và $\alpha(a, b) = (\alpha a, 0)$.

**(b)** $(a, b) + (c, d) = (a + c, b + d)$ và $\alpha(a, b) = (\alpha^2 a, \alpha^2 b)$

**(c)** $(a, b) + (c, d) = (a, b)$ và $\alpha(a, b) = (\alpha a, \alpha b)$

**2**) Cho $S$ là một tập hợp khác rỗng và $V$ là không gian vector trên $\mathbb{K}$.

**(a)** Xét tập $\mathbb{K}^S = \{f : S \to \mathbb{K}\}$ các ánh xạ từ $S$ đến $\mathbb{K}$. Ta định nghĩa phép cộng và trên $\mathbb{K}^S$ và phép nhân với vô hướng như sau

$$(f+g)(s) = f(s) + g(s) \qquad (\lambda f)(s) = \lambda \cdot f(s)$$

với mọi $f, g \in \mathbb{K}^S$ và mọi $\lambda \in \mathbb{K}$. Chứng minh rằng $\mathbb{K}^S$ là không gian vector trên $\mathbb{K}$.

**(b)** Bằng cách tương tự, hãy làm cho tập hợp $V^S = \{f : S \to V\}$ các ánh xạ từ $S$ đến $V$ trở thành không gian vector trênn $\mathbb{K}$.

**3**) Cho $\mathbb{R}^+$ là tập hợp các số thực dương. Trong $\mathbb{R}^+$ ta định nghĩa phép cộng và phép nhân vô hướng: Phép cộng $x \oplus y = xy$ với mọi $x, y \in \mathbb{R}^+$; Phép nhân vô hướng $\alpha * x = x^\alpha$ với mọi $x \in \mathbb{R}^+$ với mọi $\alpha \in \mathbb{R}$. Chứng minh rằng $\mathbb{R}^+$ cùng với hai phép toán trên là một không gian vector trên $\mathbb{R}$.

**4**) Chứng minh rằng một không gian vector hoặc có một vector hoặc có vô số vector.

**5**) Chứng minh rằng với mọi $\lambda, \mu \in \mathbb{K}$ và mọi $\mathbf{x}, \mathbf{y} \in V$ ta luôn có $(\lambda - \mu)\mathbf{x} = \lambda\mathbf{x} - \mu\mathbf{x}$ và $\lambda(\mathbf{x} - \mathbf{y}) = \lambda\mathbf{x} - \lambda\mathbf{y}$.

# § 2 Độc lập tuyến tính và phụ thuộc tuyến tính

## 2.1 Các khái niệm

**2.1.1 Định nghĩa.** Cho $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ là $n$ vector của không gian vector $V$ và $\lambda_1, \lambda_2, \dots, \lambda_n$ là $n$ vô hướng trong $\mathbb{K}$. Vector

$$\mathbf{x} = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \cdots + \lambda_n\mathbf{x}_n = \sum_{i=1}^{n} \lambda_i\mathbf{x}_i$$

được gọi là một **tổ hợp tuyến tính** (linear combination) của hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ hay $(\mathbf{x}_i)_{i=1,\dots,n}$ với họ hệ số $(\lambda_1, \lambda_2, \dots, \lambda_n)$ hay $(\lambda_i)_{i=1,\dots,n}$.

Khi vector $\mathbf{x}$ là một tổ hợp tuyến tính của hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ thì ta bảo $\mathbf{x}$ **biểu thị tuyến tính** được qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$.

Như vậy $\mathbf{x}$ biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ nếu có một họ vô hướng $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ trong $\mathbb{K}$ sao cho $\mathbf{x} = \sum\limits_{i=1}^{n} \lambda_i \mathbf{x}_i$. Tuy nhiên cách biểu diễn $\mathbf{x} = \sum\limits_{i=1}^{n} \lambda_i \mathbf{x}_i$ nói chung không duy nhất, nghĩa là còn có thể có $\mathbf{x} = \sum\limits_{i=1}^{n} \mu_i \mathbf{x}_i$ mà không phải $\lambda_i = \mu_i$ với mọi $i = 1, 2, \ldots, n$.

**2.1.2 Thí dụ.** Trong không gian vector $\mathbb{R}^2$, cho 3 vector $\mathbf{x}_1 = (-1, 0)$, $\mathbf{x}_2 = (0, -1)$, $\mathbf{x}_3 = (1, 2)$. Khi đó, vector $\mathbf{0} = (0, 0)$ biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ ít nhất hai cách sau: $\mathbf{0} = 0\mathbf{x}_1 + 0\mathbf{x}_2 + 0\mathbf{x}_3$ và $\mathbf{0} = \mathbf{x}_1 + 2\mathbf{x}_2 + \mathbf{x}_3$. $\square$

**2.1.3 Thí dụ.** Trong không gian $\mathbb{R}[x]$ cho các đa thức $\mathbf{u}_1 = 3x^3 - 2x + 1$, $\mathbf{u}_2 = x^2 - x + 2$, $\mathbf{u}_3 = x^3 + x + 1$. Ta có $\mathbf{u} = 2x^3 + 2x^2 - 5x + 4$ biểu thị tuyến tính được qua hệ $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$. Cụ thể là $\mathbf{u} = \mathbf{u}_1 + 2\mathbf{u}_2 - \mathbf{u}_3$. $\square$

**2.1.4 Định nghĩa.** Cho $V$ là không gian vector trên $\mathbb{K}$ và $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ là hệ $n$ vector thuộc $V$. Hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ được gọi là **độc lập tuyến tính** (linear independent) nếu với mọi $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ sao cho $\sum_{i=1}^{n} \lambda_i \mathbf{x}_i = \mathbf{0}$ thì $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Hệ các vector không độc lập tuyến tính được gọi là **phụ thuộc tuyến tính**. Nghĩa là hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ là phụ thuộc tuyến tính nếu tồn tại $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ không đồng thời bằng 0 mà $\sum_{i=1}^{n} \lambda_i \mathbf{x}_i = \mathbf{0}$.

**2.1.5 Thí dụ.** Trong không gian vector thực $\mathbb{R}^2$ cho 3 vector $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$ và $\mathbf{x} = (1, 1)$.

**(a)** Hệ $(\mathbf{e}_1, \mathbf{e}_2)$ độc lập tuyến tính. Thật vậy, với mọi $\lambda, \mu \in \mathbb{R}$ ta có

$$\lambda\mathbf{e}_1 + \mu\mathbf{e}_2 = \lambda(1, 0) + \mu(0, 1) = (\lambda, \mu).$$

Do đó $\lambda\mathbf{e}_1 + \mu\mathbf{e}_2 = \mathbf{0}$ khi và chỉ khi $(\lambda, \mu) = (0, 0)$ hay $\lambda = 0$ và $\mu = 0$.

**(b)** Hệ $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{x})$ phụ thuộc tuyến tính. Thật vậy, ta thấy rằng $\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{x} = \mathbf{0}$ là một tổ hợp tuyến tính bằng $\mathbf{0}$ không tầm thường của hệ $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{x})$. □

**2.1.6 Thí dụ. (OSV11)** Xét không gian trên trường số thực $\mathbb{R}$, chứng minh rằng tập hợp $\{e^x, e^{x^2}, \dots, e^{x^n}\}$ độc lập tuyến tính trong không gian các hàm liên tục $\mathcal{C}_{(0,\infty)}$.

*Giải.* Xét tổ hợp tuyến tính $a_1e^x + a_2e^{x^2} + \cdots + a_ne^{x^n} = 0$. Chia hai vế cho $e^{x^n}$ và lấy giới hạn khi $x \to \infty$

$$\lim_{x\to\infty}(a_1e^{x-x^n} + a_2e^{x^2-x^n} + \cdots + a_n) = a_n = 0.$$

Do đó, ta được $a_1e^x + a_2e^{x^2} + \cdots + a_{n-1}e^{x^{n-1}} = 0$. Bằng cách trên ta suy ra được $a_{n-1} = 0$. Tiếp tục như thế ta suy ra được $a_i = 0$ với mọi $i = 1, 2, \dots, n$. Do đó, hệ đã cho độc lập tuyến tính. □

**2.1.7 Thí dụ. (OSV12)** Chứng minh rằng các hàm số

$$\sin x, \sin 2x, \sin 3x, \sin|x-\pi|, \sin|x-2\pi|, \sin|x-3\pi|$$

độc lập tuyến tính trong không gian các hàm liên tục $\mathcal{C}_{(-\infty,\infty)}$ trên trường số thực.

*Giải.* Xét tổ hợp tuyến tính

$$\begin{aligned} a_1\sin x + a_2\sin 2x + a_3\sin 3x + b_1\sin|x-\pi| \\ + b_2\sin|x-2\pi| + b_3\sin|x-3\pi| = 0. \end{aligned}$$

Dùng định nghĩa đạo hàm ta chứng minh được hàm $\sin|x-k\pi|$ không khả vi tại $k\pi$ với $k = 1, 2, 3$. Ta thấy rằng các số hạng ở vế trái khả vi tại $x = k\pi$ trừ số hạng $b_k\sin|x-k\pi|$, cho nên ta phải có $b_k = 0$ với $k = 1, 2, 3$. Bây giờ chỉ còn $a_1\sin x + a_2\sin 2x + a_3\sin 3x = 0$. Ta có thể lấy ba giá trị cho $x$ là $\frac{\pi}{2}$, $\frac{\pi}{3}$, $\frac{\pi}{4}$, ta được hệ sau

$$\begin{cases} a_1 - a_3 = 0 \\ \frac{\sqrt{3}}{2}a_1 + \frac{\sqrt{3}}{2}a_2 = 0 \\ \frac{\sqrt{2}}{2}a_1 + a_2 + \frac{\sqrt{2}}{2}a_3 = 0 \end{cases}$$

Từ đó suy ra $a_1 = a_2 = a_3 = 0$. Vậy các hàm đã cho độc lập tuyến tính. □

**2.1.8** ***Chú ý.*** Ta quy ước hệ $\varnothing$ (không chứa vector nào) là hệ độc lập tuyến tính. Vector **0** là tổ hợp tuyến tính tầm thường của hệ $\varnothing$ và là vector duy nhất biểu thị tuyến tính qua hệ $\varnothing$.

**2.1.9** Xét hệ phương trình tuyến tính thuần nhất

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{n2}x_2 + \cdots + a_{mn}x_n = 0. \end{cases}$$

Ký hiệu $\mathbf{a}_i = (a_{1i}, a_{2i}, \dots, a_{mi})$ với $i = 1, \dots, n$ và $\mathbf{0} = (0, 0, \dots, 0)$ với $m$ số 0. Khi đó hệ trên có thể viết lại như sau

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

Từ đó ta kết luận được hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ trong $\mathbb{R}^m$ độc lập tuyến tính khi và chỉ khi hệ phương trình tuyến tính thuần nhất đang xét có nghiệm duy nhất là $(0, 0, \dots, 0)$. Nói cách khác, hệ $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ phụ thuộc tuyến tính khi và chỉ khi hệ phương trình tuyến tính thuần nhất trên có ít nhất một nghiệm không tầm thường (nghiệm khác **0**). Với kí hiệu và kết quả trên cùng với Định lý 2.3.1 (trang 196) ta có thể phát biểu lại trong định lý sau.

**2.1.10 Định lý.** *Hệ vector* $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ *trong* $\mathbb{R}^m$ *độc lập tuyến tính khi và chỉ khi* $\operatorname{rank} A = n$.

**2.1.11 Thí dụ. (OSV06)** Cho $A$ là ma trận vuông cấp $n$ sao cho mỗi dòng của nó chứa đúng 2 phần tử khác 0, trong đó phần tử nằm ở đường chéo chính là 2006, phần tử còn lại là 1. Chứng minh rằng ma trận $A$ khả nghịch.

*Giải* Gọi $\mathbf{a}_1, \dots, \mathbf{a}_n$ là các vector cột của $A$. Giả sử $A$ suy biến; khi đó các vector $\mathbf{a}_1, \dots, \mathbf{a}_n$ phụ thuộc tuyến tính. Cho nên tồn tại các hệ số $\lambda_1, \dots, \lambda_n$ không đồng thời bằng 0 sao cho

$$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n = \mathbf{0}.$$

Giả sử $|\lambda_m| = \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$. Rõ ràng $|\lambda_m| > 0$. Giả sử hai phần tử khác không ở dòng thứ $m$ là $a_{mm} = 2006$ và $a_{mp} = 1$ với $p \neq m$. Từ đẳng thức phụ thuộc trên ứng với tọa độ thứ $m$ ta có

$$\lambda_1 a_{m1} + \lambda_2 a_{m2} + \dots + \lambda_n a_{mn} = 2006\lambda_m + \lambda_p = 0.$$

Suy ra $|\lambda_p| = 2006|\lambda_m| > |\lambda_m|$, mâu thuẫn với định nghĩa $\lambda_m$. Vậy $A$ khả nghịch. $\square$

**2.1.12** Tương tự như trên ta xét hệ phương trình tuyến tính

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{n2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}$$

Với kí hiệu đã dùng ở trên và $\mathbf{b} = (b_1, b_2, \dots, b_m)$, hệ phương trình trên được viết gọn lại như sau

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

Như vậy hệ phương trình tuyến tính đang xét có nghiệm khi và chỉ khi vector $\mathbf{b}$ biểu thị tuyến tính được qua hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$. Từ kết quả trên và Định lý 2.1.3 (Kronecker-Capelli) ta có định lý sau.

**2.1.13 Định lý.** *Trong $\mathbb{R}^m$, vector $\mathbf{b}$ biểu thị tuyến tính được qua hệ $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ khi và chỉ khi* $\operatorname{rank} A = \operatorname{rank} \bar{A}$.

**2.1.14 Thí dụ.** Cho $\mathbf{a}_1 = (1, 0, 1, 0)$; $\mathbf{a}_2 = (1, 1, -1, 2)$; $\mathbf{a}_3 = (0, 1, 2, 3)$ là các vector trong $\mathbb{R}^4$. Hệ $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ độc lập tuyến tính hay phụ thuộc tuyến tính? Vector $\mathbf{b} = (1, 0, -1, 2)$ có biểu thị tuyến tính qua hệ trên được hay không?

Xét hệ phương trình

$$\begin{cases} x_1 + x_2 = 0 \\ x_2 + x_3 = 0 \\ x_1 - x_2 + 2x_3 = 0 \\ 2x_2 + 3x_3 = 0 \end{cases}$$

Dùng phương pháp Gauss ta giải hệ phương trình này. Ta có

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 2 \\ 0 & 2 & 3 \end{bmatrix} \xrightarrow{d_3 \to d_3 - d_1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & 2 \\ 0 & 2 & 3 \end{bmatrix} \xrightarrow[d_4 \to d_4 - 2d_1]{d_3 \to d_3 + 2d_2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{d_4 \to d_4 - \frac{1}{4} d_3} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}.$$

Vậy $\operatorname{rank} A = 3$, cho nên hệ phương trình trên chỉ có nghiệm tầm thường. Do đó hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ độc lập tuyến tính.

Xét hệ phương trình tuyến tính

$$\begin{cases} x_1 + x_2 = 1 \\ x_2 + x_3 = 0 \\ x_1 - x_2 + 2x_3 = -1 \\ 2x_2 + 3x_3 = 2 \end{cases}$$

Dùng các phép biến đổi sơ cấp trên ma trận nhưng trên ta được

$$\bar{A} = \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & -1 & 2 & -1 \\ 0 & 2 & 3 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & 0 & \frac{5}{2} \end{array}\right].$$

Vậy hệ phương trình đang xét vô nghiệm, cho nên vector $\mathbf{b}$ không thể biểu thị tuyến tính qua hệ $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$. □

## 2.2 Các tính chất cơ bản

**2.2.1 Mệnh đề.** *Hệ chứa vector không luôn là hệ phụ thuộc tuyến tính.*

*Chứng minh.* Giả sử hệ vector $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ có vector $\mathbf{x}_k = \mathbf{0}$. Khi đó, lấy $\lambda_i = 0$ khi $i \neq k$ và $\lambda_k = 1$; ta có $\sum_{i=1}^{n} \lambda_i \mathbf{x}_i = \mathbf{0}$. Vậy hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ phụ thuộc tuyến tính. □

**2.2.2 Mệnh đề.** *Hệ gồm một vector phụ thuộc tuyến tính khi và chỉ khi vector đó là vector không.*

*Chứng minh.* Rõ ràng theo mệnh đề trên hệ $(\mathbf{0})$ phụ thuộc tuyến tính. Ngược lại, nếu hệ $(\mathbf{x})$ phụ thuộc tuyến tính, thì tồn tại $\lambda \neq 0$ sao cho $\lambda\mathbf{x} = \mathbf{0}$. Theo Tính chất 1.3.1 (v) của không gian vector ta có $\mathbf{x} = \mathbf{0}$. □

**2.2.3 Mệnh đề.** *Nếu một hệ vector độc lập tuyến tính thì mọi hệ con của nó cũng độc lập tuyến tính. Hay ta có phát biểu tương đương: nếu một hệ vector có một hệ con phụ thuộc tuyến tính thì hệ đó phụ thuộc tuyến tính.*

*Chứng minh.* Xét hệ vector $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ độc lập tuyến tính. Giả sử $(\mathbf{x}_i)_{i\in I}$ là hệ con bất kì của hệ trên. Giả sử ta có tổ hợp tuyến tính sau $\sum_{i\in I} \lambda_i\mathbf{x}_i = \mathbf{0}$. Lấy $\lambda_i = 0$ khi $i \in \{1, \dots, n\}\setminus I$. Suy ra $\sum_{i=1}^{n} \lambda_i\mathbf{x}_i = \mathbf{0}$. Vì hệ $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ độc lập tuyến tính, nên $\lambda_i = 0$ với mọi $i = 1, \dots, n$, suy ra $\lambda_i = 0$ với mọi $i \in I$. Vậy hệ $(\mathbf{x}_i)_{i\in I}$ độc lập tuyến tính. □

**2.2.4 Định lý. (Đặc trưng của hệ phụ thuộc tuyến tính)** *Một hệ $n \geq 2$ vector phụ thuộc tuyến tính khi và chỉ khi có ít nhất một vector trong hệ biểu thị tuyến tính được qua các vector còn lại.*

*Chứng minh.* Giả sử hệ vector $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ phụ thuộc tuyến tính. Khi đó, tồn tại các hệ số $\lambda_1, \dots, \lambda_n$ không đồng thời bằng không sao cho $\sum_{i=1}^{n} \lambda_i\mathbf{x}_i = \mathbf{0}$. Giả sử $\lambda_j \neq 0$, ta có $\mathbf{x}_j = \sum_{i\neq j}(-\frac{\lambda_i}{\lambda_j})\mathbf{x}_i$ hay $\mathbf{x}_j$ biểu thị tuyến tính được qua các vector còn lại trong hệ $(\mathbf{x}_1, \dots, \mathbf{x}_n)$.

Ngược lại, giả sử trong hệ $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ có vector $\mathbf{x}_k$ biểu thị tuyến tính được qua các vector còn lại trong hệ. Nghĩa là tồn tại $\lambda_1, \dots, \lambda_{k-1}$, $\lambda_{k+1}, \dots, \lambda_n$ sao cho $\mathbf{x}_k = \sum_{i\neq k} \lambda_i\mathbf{x}_i$ hay $\sum_{i=1}^{n} \lambda_i\mathbf{x}_i = \mathbf{0}$ với $\lambda_k = -1$. Vậy hệ $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ phụ thuộc tuyến tính. □

**2.2.5 Hệ quả.** *Hệ gồm hai vector phụ thuộc tuyến tính khi và chỉ khi hai vector đó tỷ lệ với nhau.*

*Chứng minh.* Xét hệ $(\mathbf{x}_1, \mathbf{x}_2)$ phụ thuộc tuyến tính. Khi đó, theo định lý trên tồn tại một vector giả sử đó là $\mathbf{x}_1$ biểu thị tuyến tính được qua các vector còn lại là $\mathbf{x}_2$; nghĩa là tồn tại $\lambda$ sao cho $\mathbf{x}_1 = \lambda\mathbf{x}_2$. Ngược lại, giả sử trong hệ $(\mathbf{x}_1, \mathbf{x}_2)$ có $\mathbf{x}_2 = \lambda\mathbf{x}_1$ suy ra $\lambda\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$. Vậy $(\mathbf{x}_1, \mathbf{x}_2)$ phụ thuộc tuyến tính. □

**2.2.6 Định lý.** *Cho hệ vector độc lập tuyến tính* $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ *và* $\mathbf{y}$ *là một vector trong cùng một không gian. Khi đó hệ* $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m, \mathbf{y})$ *độc lập tuyến tính khi và chỉ khi* $\mathbf{y}$ *không biểu thị tuyến tính được qua* $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$.

*Chứng minh.* Nếu hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y})$ độc lập tuyến tính thì theo định lý trên ta có $\mathbf{y}$ không biểu thị tuyến tính được qua $(\mathbf{x}_1, \ldots, \mathbf{x}_m)$.

Ngược lại, giả sử $\mathbf{y}$ không biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_m)$. Xét tổ hợp tuyến tính $\sum_{i=1}^m \lambda_i\mathbf{x}_i + \lambda\mathbf{y} = \mathbf{0}$. Nếu $\lambda \neq 0$ ta suy ra $\mathbf{y} = \sum_{i=1}^m (-\frac{\lambda_i}{\lambda})\mathbf{x}_i$ vô lý vì $\mathbf{y}$ không biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_m)$; do đó, $\lambda = 0$, suy ra $\sum_{i=1}^m \lambda_i\mathbf{x}_i = \mathbf{0}$. Vì hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_m)$ độc lập tuyến tính nên $\lambda_i = 0$ với mọi $i = 1, \ldots, m$. Vậy hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y})$ độc lập tuyến tính. □

**2.2.7 Định lý.** *Cho* $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ *là một hệ độc lập tuyến tính trong không gian vector* $V$. *Khi đó với mọi* $\mathbf{v} \in V$ *đều có không quá một cách biểu thị tuyến tính qua hệ* $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$.

*Chứng minh.* Lấy $\mathbf{v} \in V$ tùy ý. Giả sử $\mathbf{v}$ biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ bởi hai tổ hợp tuyến tính sau $\mathbf{v} = \sum\limits_{i=1}^n \lambda_i\mathbf{x}_i$ và $\mathbf{v} = \sum\limits_{i=1}^n \mu_i\mathbf{x}_i$. Do đó

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = \sum_{i=1}^n \lambda_i\mathbf{x}_i - \sum_{i=1}^n \mu_i\mathbf{x}_i = \sum_{i=1}^n (\lambda_i - \mu_i)\mathbf{x}_i.$$

Vì hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ độc lập tuyến tính nên $\lambda_i - \mu_i = 0$ với mọi $i = 1, \ldots, n$ hay $\lambda_i = \mu_i$ với mọi $i = 1, \ldots, n$. Tức là $\mathbf{v}$ biểu thị tuyến tính một cách duy nhất qua hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. □

## 2.3 Sự độc lập tuyến tính đối với các hệ vô hạn

Các khái niệm về biểu thị tuyến tính và độc lập tuyến tính có thể mở rộng cho các hệ bất kỳ vector. Hệ vector trong không gian vector $V$ được chỉ số hóa bởi tập $I$, ký hiệu $(\mathbf{x}_i)_{i\in I}$, là một ánh xạ từ $I$ đến $V$ với $i \mapsto \mathbf{x}_i$. Tập $I$ có thể là tập vô hạn.

**2.3.1 Định nghĩa.** Vector $\mathbf{x}$ được gọi là *biểu thị tuyến tính* được qua hệ $(\mathbf{x}_i)_{i\in I}$ nếu $\mathbf{x}$ biểu thị tuyến tính được qua một hệ con hữu hạn của $(\mathbf{x}_i)_{i\in I}$; nghĩa là tồn tại ít nhất một họ vô hướng $(\lambda_i)_{i\in I}$ sao cho $\lambda_i = 0$ trừ một số hữu hạn các chỉ số $i \in I$ và $\mathbf{x} = \sum\limits_{i\in I} \lambda \mathbf{x}_i$.

**2.3.2 Định nghĩa.** Hệ $(\mathbf{x}_i)_{i\in I}$ được gọi là *độc lập tuyến tính* nếu mọi hệ con hữu hạn của nó đều độc lập tuyến tính. Hệ không độc lập tuyến tính gọi là hệ *phụ thuộc tuyến tính*. Như vậy, hệ $(\mathbf{x}_i)_{i\in I}$ phụ thuộc tuyến tính nếu có ít nhất một hệ con hữu hạn của nó phụ thuộc tuyến tính.

**2.3.3 Nhận xét.** Các tính chất về hệ hữu hạn độc lập tuyến tính hay phụ thuộc tuyến tính đã trình bày ở trên vẫn đúng đối với các hệ vô hạn độc lập tuyến tính hay phụ thuộc tuyến tính.

## Bài tập

**1)** Trong $\mathbb{R}^3$ cho hai vector $\mathbf{u}_1 = (1, -2, 3)$, $\mathbf{u}_2 = (0, 1, -3)$.

**(a)** Vector $\mathbf{u} = (2, -3, 3)$ có biểu thị tuyến tính được qua $(\mathbf{u}_1, \mathbf{u}_2)$ không?

**(b)** Tìm $m$ để $\mathbf{v} = (1, m, -3)$ biểu thị tuyến tính được qua $(\mathbf{u}_1, \mathbf{u}_2)$.

**2)** Trong $\mathbb{R}^4$ cho các vector $\mathbf{u}_1 = (1, 1, 1, 1)$, $\mathbf{u}_2 = (2, 3, -1, 0)$, $\mathbf{u}_3 = (-1, -1, 1, 1)$.

**(a)** Tìm điều kiện để vector $\mathbf{b} = (b_1, b_2, b_3, b_4) \in \mathbb{R}^4$ biểu thị tuyến tính được qua hệ $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$.

**(b)** Xét xem các vector $\mathbf{v}_1 = (3, 2, 3, 1)$, $\mathbf{v}_2 = (3, 4, 2, 3)$ có biểu thị tuyến tính được qua hệ $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ hay không? Nếu được, hãy tìm các cách biểu diễn tuyến tính có thể có.

**3)** Trong $\mathbb{R}^4$ cho các vector $\mathbf{u}_1 = (1, 1, 1, 1)$, $\mathbf{u}_2 = (2, 3, -1, 0)$, $\mathbf{u}_3 = (-1, -1, 1, 1)$, $\mathbf{u}_4 = (1, 2, 1, -1)$. Tìm điều kiện để vector $\mathbf{u} = (x_1, x_2, x_3, x_4)$ là tổ hợp tuyến tính của

**(a)** $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ **(b)** $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4)$.

**4)** Trong $\mathbb{R}_2[x]$ cho các đa thức $\mathbf{u}_1 = x^2 + 3x + 1$, $\mathbf{u}_2 = -2x^2 - x - 1$, $\mathbf{u}_3 = -2x^2 + x + m$. Tìm $m$ để $\mathbf{u}_3$ biểu thị tuyến tính được qua $(\mathbf{u}_1, \mathbf{u}_2)$.

**5)** Trong $\mathbb{R}_3[x]$ cho các đa thức $\mathbf{u}_1 = x^3 + 2x^2 + x + 1$, $\mathbf{u}_2 = 2x^3 + x^2 - x + 1$, $\mathbf{u}_3 = 3x^3 + 3x^2 - x + 2$. Tìm điều kiện để $\mathbf{u} = ax^3 + bx^2 + cx + d$ là tổ hợp tuyến tính của $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$.

**6)** Trong $\mathbb{R}^3$, xét xem các hệ vector sau đây là độc lập tuyến tính hay phụ thuộc tuyến tính?

**(a)** $(\mathbf{u}_1 = (1, 0, 1), \mathbf{u}_2 = (1, 2, 3), \mathbf{u}_3 = (10, 11, 12), \mathbf{u}_4 = (4, 5, 6))$

**(b)** $(\mathbf{u}_1 = (1, 0, 1), \mathbf{u}_2 = (1, 2, 3), \mathbf{u}_3 = (2, 2, 4))$

**(c)** $(\mathbf{u}_1 = (1, 0, 1), \mathbf{u}_2 = (1, 2, 3), \mathbf{u}_3 = (2, 2, 5))$

**(d)** $(\mathbf{u}_1 = (1, 0, 1), \mathbf{u}_2 = (1, 2, 3))$

**7)** Xét sự độc lập tuyến tính và phụ thuộc tuyến tính của các hệ vector trong $\mathbb{R}^4$.

**(a)** $(\mathbf{x}_1 = (1, 2, 3, 4)$, $\mathbf{x}_2 = (2, 3, 4, 5)$, $\mathbf{x}_3 = (3, 4, 5, 6)$, $\mathbf{x}_4 = (4, 5, 6, 7))$

**(b)** $(\mathbf{x}_1 = (1, 1, 1, 1)$, $\mathbf{x}_2 = (2, 3, -1, 0)$, $\mathbf{x}_3 = (-1, -1, 1, 1)$, $\mathbf{x}_4 = (1, 2, 1, -1))$

**8)** Xét tính độc lập tuyến tính hay phụ thuộc tuyến tính của các hệ vector sau đây trong không gian đã chỉ ra.

**(a)** $(\mathbf{u}_1 = (1, 1, 0), \mathbf{u}_2 = (0, 1, 1), \mathbf{u}_3 = (1, 0, 1))$ trong $\mathbb{R}^3$.

**(b)** $(\mathbf{v}_1 = (1, 1, 1), \mathbf{v}_2 = (1, -1, 1), \mathbf{v}_3 = (3, 1, 3))$ trong $\mathbb{R}^3$.

**(c)** $(\mathbf{w}_1 = (1,1,0,0), \mathbf{w}_2 = (0,1,1,0), \mathbf{w}_3 = (2,3,1,0))$ trong $\mathbb{R}^4$.

**9)** Xét sự độc lập tuyến tính và phụ thuộc tuyến tính của các hệ vector trong $\mathcal{M}_2(\mathbb{R})$

**(a)** $\left(A_1 = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 1 \\ -2 & 5 \end{bmatrix}\right)$

**(b)** $\left(A_1 = \begin{bmatrix} 1 & 4 \\ -1 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} 2 & 1 \\ -3 & -1 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 2 \\ 1 & -5 \end{bmatrix}\right)$

**10)** Trong $\mathbb{R}_3[x]$ xét xem các hệ sau đây độc lập tuyến tính hay phụ thuộc tuyến tính?

**(a)** $(\mathbf{u}_1 = x^3 - 2x + 3, \mathbf{u}_2 = x^2 + 1)$

**(b)** $(\mathbf{u}_1 = x^3 - 2x + 3, \mathbf{u}_2 = x^2 + 1, \mathbf{u}_3 = 2x^3 + x^2 - 4x + 10)$

**(c)** $(\mathbf{u}_1 = x^3 - 2x + 3, \mathbf{u}_2 = x^2 + 2x - 1, \mathbf{u}_3 = x^3 + x^2 + 2)$

**(d)** $(\mathbf{u}_1 = x^3, \mathbf{u}_2 = 2x^2, \mathbf{u}_3 = 3x, \mathbf{u}_4 = 2x^2 + 3x, \mathbf{u}_5 = 1)$

**11)** Trong không gian vector $V$ cho ba vector $\mathbf{x}, \mathbf{y}, \mathbf{z}$. Chứng minh rằng $(\mathbf{x} + \mathbf{y}, \mathbf{y} + \mathbf{z}, \mathbf{z} + \mathbf{x})$ độc lập tuyến tính khi và chỉ khi $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ độc lập tuyến tính.

**12)** Chứng minh rằng nếu hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ độc lập tuyến tính, thì hệ $\mathbf{y}_1 = \mathbf{x}_1, \mathbf{y}_2 = \mathbf{x}_1 + \mathbf{x}_2, \ldots, \mathbf{y}_n = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_n$ cũng độc lập tuyến tính.

**13)** Cho $\mathbf{x}_1, \ldots, \mathbf{x}_k$ là các vector độc lập tuyến tính trong $\mathbb{R}^n$ và cho $A$ là một ma trận vuông cấp $n$ không suy biến. Đặt $\mathbf{y}_i = A\mathbf{x}_i$ với $i = 1, \ldots, k$. Chứng minh rằng các vector $\mathbf{y}_1, \ldots, \mathbf{y}_k$ độc lập tuyến tính.

**14)** Cho một hệ vector độc lập tuyến tính $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ trong không gian vector $V$. Chứng minh rằng mọi vector $\mathbf{y} \in V$ đều có không quá một cách biểu thị tuyến tính qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$.

**15)** Chứng minh rằng các hàm $\varphi_0(x) = 1$, $\varphi_1(x) = x$, ..., $\varphi_n(x) = x^n$, ... trên $[0,1]$ là độc lập tuyến tính trong $\mathcal{C}_{[0,1]}$.

**16)** Chứng minh rằng các hàm $\varphi_0(x) = 1$, $\varphi_1(x) = \sin x$, ..., $\varphi_n(x) = \sin nx$, ... trên $[0,\pi]$ là độc lập tuyến tính trong $\mathcal{C}_{[0,\pi]}$.

# § 3 Hạng của một hệ vector

## 3.1 Hai hệ vector tương đương

**3.1.1 Định nghĩa.** Trong không gian vector $V$ cho hai hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ và $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Ta nói hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ **biểu thị tuyến tính** được qua hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ nếu mỗi vector $\mathbf{x}_i$ đều biểu thị tuyến tính được qua hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$.

**3.1.2 Mệnh đề.** *Hệ vector con của một hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ bao giờ cũng biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$.*

*Chứng minh.* Xét hệ con $(\mathbf{x}_i)_{i\in I}$ của hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$. Với mọi $k \in I$, ta có $\mathbf{x}_k = \sum_{i=1}^{m} \lambda_i \mathbf{x}_i$ với $\lambda_k = 1$ và $\lambda_i = 0$ khi $i \neq k$. Vậy hệ $(\mathbf{x}_i)_{i\in I}$ biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$. □

**3.1.3 Mệnh đề.** *Nếu hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ biểu thị tuyến tính được qua hệ vector $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ và hệ vector $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ biểu thị tuyến tính được qua hệ vector $(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_l)$ thì hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ biểu thị tuyến tính được qua hệ $(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_l)$.*

*Chứng minh.* Vì hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ biểu thị tuyến tính được qua hệ vector $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$, nên với mỗi $i = 1, \ldots, m$, ta có thể biểu diễn $\mathbf{x}_i = \sum_{j=1}^{n} \lambda_{ij} \mathbf{y}_j$. Hơn nữa, do hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ biểu thị tuyến tính được qua hệ $(\mathbf{z}_1, \mathbf{x}_2, \ldots, \mathbf{y}_l)$ nên với mỗi $j = 1, \ldots, n$ ta có biễu diễn $\mathbf{y}_j = \sum_{k=1}^{l} \alpha_{jk} \mathbf{z}_k$. Do đó,

$$\mathbf{x}_i = \sum_{j=1}^{n} \lambda_{ij} \mathbf{y}_j \sum_{k=1}^{l} \alpha_{jk} \mathbf{z}_k = \sum_{j=1}^{n} \sum_{k=1}^{l} \lambda_{ij} \alpha_{jk} \mathbf{z}_k = \sum_{k=1}^{l} \Big( \sum_{j=1}^{n} \lambda_{ij} \alpha_{jk} \Big) \mathbf{z}_k.$$

Vậy hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ biểu thị tuyến tính được qua hệ $(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_l)$. □

**3.1.4 Định nghĩa.** Hai hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ và $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ được gọi là **tương đương** nếu hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ biểu thị tuyến tính được qua hệ vector $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ và ngược lại.

**3.1.5 Định lý. (Bổ đề cơ bản về sự phụ thuộc tuyến tính)** *Trong không gian vector $V$ cho 2 hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ và $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Nếu hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ độc lập tuyến tính và biểu thị tuyến tính được qua hệ vector $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ thì $m \leq n$ và ta có thể thay thế $m$ vector nào đó của hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ bởi các vector $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ để được một hệ vector mới tương đương với hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$.*

*Chứng minh.* Do $\mathbf{x}_1$ biểu thị tuyến tính được qua $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ nên $\mathbf{x}_1 = k_1\mathbf{y}_1 + k_2\mathbf{y}_2 + \cdots + k_n\mathbf{y}_n$ với $k_i \in \mathbb{K}$. Do hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ độc lập tuyến tính nên $\mathbf{x}_1 \neq \mathbf{0}$; do đó, tồn tại ít nhất một $k_i \neq 0$. Không mất tính tổng quát ta có thể giả sử $k_1 \neq 0$. Khi đó

$$\mathbf{y}_1 = \frac{1}{k_1}\mathbf{x}_1 - \frac{k_2}{k_1}\mathbf{y}_2 - \cdots - \frac{k_n}{k_1}\mathbf{y}_n.$$

Từ đẳng thức trên suy ra hệ vector $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ tương đương với hệ vector $(\mathbf{x}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$.

Do $\mathbf{x}_2$ biểu thị tuyến tính được qua hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ nên $\mathbf{x}_2$ cũng biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ nên

$$\mathbf{x}_2 = h_1\mathbf{x}_1 + h_2\mathbf{y}_2 + h_3\mathbf{y}_3 + \cdots + h_n\mathbf{y}_n.$$

Khi đó tồn tại ít nhất một $h_i \neq 0$ với $i \geq 2$ vì nếu $h_i = 0$ với mọi $i \geq 2$ thì $\mathbf{x}_2 = h_1\mathbf{x}_1$ mâu thuẫn với hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ độc lập tuyến tính. Không mất tính tổng quát ta giả sử $h_2 \neq 0$. Suy ra

$$\mathbf{y}_2 = -\frac{h_1}{h_2}\mathbf{x}_1 + \frac{1}{h_2}\mathbf{x}_2 - \frac{h_3}{h_2}\mathbf{y}_3 - \cdots - \frac{h_n}{h_2}\mathbf{y}_n.$$

Từ đó ta suy ra được hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_3, \ldots, \mathbf{y}_n)$ tương đương với hệ vector $(\mathbf{x}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ nên cũng tương đương với $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$.

Tiếp tục tiến hành lặp lại quá trình trên. Ta thực hiện tất cả là $\min\{m; n\}$ lần. Giả sử $m > n$. Ta phải thực hiện quá trình trên là $n$

lần và thu được hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ (sau khi đã thay $\mathbf{y}_i$ bởi $\mathbf{x}_i$ trong hệ $(\mathbf{y}_1, \ldots, \mathbf{y}_n)$) tương đương với hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Vì $\mathbf{x}_{n+1}$ biểu thị tuyến tính được qua hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ cho nên cũng biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ mâu thuẫn với giả thiết hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ độc lập tuyến tính. Do đó ta không thể có $m > n$ cho nên $m \leq n$. Vậy ta phải thực hiện quá trình trên $m$ lần và thu được hệ $(\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y}_{m+1}, \ldots, \mathbf{y}_n)$ tương đương với hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Ta có điều phải chứng minh. □

**3.1.6 Hệ quả.** *Hai hệ vector độc lập tuyến tính và tương đương nhau thì có số vector bằng nhau.*

*Chứng minh.* Cho hai hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ và $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ độc lập tuyến tính và tương đương nhau. Áp dụng bổ đề trên ta có: hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ độc lập tuyến tính và biểu thị tuyến tính được qua hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ suy ra $m \leq n$, và hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$ độc lập tuyến tính và biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ suy ra $n \leq m$. Vậy $m = n$. □

## 3.2 Hệ con độc lập tuyến tính tối đại

**3.2.1 Định nghĩa.** Cho hệ vector hữu hạn $(\mathbf{x}_i)_{i \in I}$ trong không gian vector $V$ và $J \subset I$. Hệ con $(\mathbf{x}_j)_{j \in J}$ được gọi là **hệ con độc lập tuyến tính tối đại** của hệ đã cho nếu nó là một hệ độc lập tuyến tính và khi thêm bất cứ vector $\mathbf{x}_i$ nào với $i \in I \setminus J$ vào hệ con đó ta đều nhận được một hệ phụ thuộc tuyến tính.

**3.2.2 Mệnh đề.** *Một hệ con độc lập tuyến tính tối đại của hệ cho trước thì tương đương với chính hệ ấy.*

*Chứng minh.* Cho hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. Giả sử $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_r})$ là một hệ con độc lập tuyến tính tối đại của hệ trên. Theo Mệnh đề 3.1.2 ta thấy hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_r})$ biểu thị tuyến tính được qua hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$.

Ngược lại, lấy vector $\mathbf{x}_k$ tùy ý thuộc hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. Nếu $k \in \{i_1, \ldots, i_r\}$, thì cũng theo Mệnh đề 3.1.2 ta có vector $\mathbf{x}_k$ biểu thị tuyến tính được qua hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_r})$. Nếu $k \in \{1, 2, \ldots, n\} \setminus \{i_1, \ldots, i_r\}$, thì hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_r}, \mathbf{x}_k)$ phụ thuộc tuyến tính vì hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_r})$ độc lập tuyến tính tối đại. Theo Định lý 2.2.6 ta có $\mathbf{x}_k$ phải biểu thị tuyến tính được qua hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_r})$. Vậy hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ biểu thị tuyến tính được qua hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_r})$. Do đó hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \ldots, \mathbf{x}_{i_r})$ tương đương với hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. $\square$

**3.2.3 Mệnh đề.** *Một hệ vector hữu hạn có chứa vector khác không thì luôn có hệ con độc lập tuyến tính tối đại.*

*Chứng minh.* Giả sử hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ thỏa giả thiết. Lấy $\mathbf{x}_{i_1} \neq \mathbf{0}$. Nếu với mọi $\mathbf{x}_i$ của hệ vector đều biểu thị tuyến tính được qua $\mathbf{x}_{i_1}$ thì hệ $(\mathbf{x}_{i_1})$ là hệ con độc lập tuyến tính tối đại.

Ngược lại, tồn tại $\mathbf{x}_{i_2}$ không biểu thị tuyến tính được qua $\mathbf{x}_{i_1}$ khi đó hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2})$ độc lập tuyến tính. Nếu với mọi $\mathbf{x}_i$ đều biểu thị tuyến tính được qua hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2})$ thì hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2})$ độc lập tuyến tính tối đại.

Ngược lại, tồn tại $\mathbf{x}_{i_3}$, không biểu thị tuyến tính được qua hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2})$ thì hệ $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \mathbf{x}_{i_3})$ độc lập tuyến tính.

Làm tiếp tục như vậy ta sẽ tìm được hệ con độc lập tuyến tính tối đại vì hệ ban đầu chỉ có $n$ vector. $\square$

**3.2.4 Nhận xét.** Một hệ vector cho trước có thể có nhiều hệ con độc lập tuyến tính tối đại khác nhau, nhưng các hệ con độc lập tuyến tính tối đại ấy phải tương đương nhau (vì cùng tương đương với hệ ban đầu). Do đó theo Hệ quả 3.1.6 tất cả các hệ con độc lập tuyến tính tối đại của hệ vector ban đầu đều có số vector bằng nhau. Số đó ta gọi là **hạng** của hệ vector.

**3.2.5 Định nghĩa.** **Hạng** của hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$ là số vector của hệ con độc lập tuyến tính tối đại bất kì của nó, kí hiệu $\operatorname{rank}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$. Qui ước: Hệ gồm toàn vectơ không có hạng bằng $0$.

**3.2.6 Thí dụ.** Trong $\mathbb{R}^4$ xét ba vector $\mathbf{x}_1 = (1, 0, 1, 2)$; $\mathbf{x}_2 = (2, -1, 0, 3)$; $\mathbf{x}_3 = (-1, 2, 3, 0)$ có $(\mathbf{x}_1, \mathbf{x}_2)$ là hệ con độc lập tuyến tính tối đại vì hai vector $\mathbf{x}_1$ và $\mathbf{x}_2$ không tỷ lệ. Tương tự ta cũng thấy các hệ $(\mathbf{x}_1, \mathbf{x}_3)$ và $(\mathbf{x}_2, \mathbf{x}_3)$ là độc lập tuyến tính. Hơn nữa, ta thấy $\mathbf{x}_3 = 3\mathbf{x}_1 - 2\mathbf{x}_2$, cho nên hệ $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ phụ thuộc tuyến tính. Vậy $\operatorname{rank}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = 2$. $\square$

**3.2.7 Mệnh đề.** *Hạng của hệ vector* $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ *bằng* $n$ *khi và chỉ khi hệ* $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ *độc lập tuyến tính. Ta có phát biểu tương đương: hạng của hệ vector* $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ *nhỏ hơn* $n$ *khi và chỉ khi hệ* $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ *phụ thuộc tuyến tính.*

*Chứng minh.* Nếu hạng của hệ $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ bằng $n$, thì nó có hệ con độc lập tuyến tính tối đại có $n$ vector. Do đó hệ con độc lập tuyến tính tối đại chính là hệ $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$. Vậy hệ $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ độc lập tuyến tính.

Ngược lại, nếu hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ độc lập tuyến tính, thì $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ chính là hệ con độc lập tuyến tính tối đại của nó. Do đó, hạng của hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ bằng $n$. □

**3.2.8 Nhận xét.** Đối với hệ vector vô hạn ta cũng có khái niệm hệ con độc lập tuyến tính tối đại, được định nghĩa như hệ vector hữu hạn. Các tính chất về hệ cong độc lập tuyến tính tối đại vẫn đúng cho hệ tùy ý, tuy nhiên phép chứng minh phức tạp hơn.

**3.2.9 Định nghĩa.** Hệ $(\mathbf{x}_i)_{i \in I}$ được gọi là có *hạng hữu hạn*, nếu nó có ít nhất một hệ con hữu hạn độc lập tuyến tính tối đại. Khi đó mọi hệ con độc lập tuyến tính tối đại của hệ đó đều hữu hạn và có cùng số vector. Số đó gọi là *hạng* của hệ đã cho, ký hiệu $\operatorname{rank}(\mathbf{x}_i)_{i \in I}$.

Hệ không có hạng hữu hạn gọi là hệ có *hạng vô hạn*. Hệ loại này luôn chứa những hệ con độc lập tuyến tính có số vector tùy ý.

Các tính chất về hạng của hệ vector hữu hạn cũng đúng đối với hệ vector tùy ý có hạng hữu hạn.

## 3.3 Hệ vector trong $\mathbb{K}^n$

Trong mục nhỏ này ta xét các hệ vector trong không gian $\mathbb{K}^n$. Xét hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ với $\mathbf{a}_i = (a_{1i}, a_{2i}, \dots, a_{ni})$. Ta lập ma trận cấp

$n \times m$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}.$$

Nếu ta xem mỗi vector $\mathbf{a}_i$ là ma trận cột $\begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix}$, thì ma trận $A$ thu được bằng cách ghép các ma trận cột $\mathbf{a}_1, \dots, \mathbf{a}_m$.

**3.3.1 Định lý.** *Với kí hiệu trên, hạng của hệ vector* $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ *bằng hạng của ma trận* $A$, $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m) = \operatorname{rank} A$. *Hơn nữa, một hệ con độc lập tuyến tính tối đại của hệ* $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ *là các vector tương ứng với các dòng khác không trong ma trận bậc thang sau khi ta biến đổi theo dòng ma trận* $A^t$.

*Chứng minh.* Trường hợp $A = O_{n\times m}$ hay $m = 1$ hoặc $n = 1$ thì định lý hiển nhiên đúng.

Giả sử $\operatorname{rank} A = \operatorname{rank} A^t = r$ với $1 \leq r \leq \min\{m, n\}$. Khi đó, $A^t$ có một định thức con $D_r \neq 0$ cấp $r$. Không mất tính tổng quát ta giả sử $D_r$ chính là định thức con nằm ở góc trên bên trái của $A^t$, tức là

$$D_r = \begin{vmatrix} a_{11} & a_{21} & \dots & a_{r1} \\ a_{12} & a_{22} & \dots & a_{r2} \\ \dots & \dots & \dots & \dots \\ a_{1r} & a_{2r} & \dots & a_{rr} \end{vmatrix} \neq 0.$$

Khi đó, ma trận cấp $r \times n$ sau đây

$$A_r^t = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{r1} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{r2} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{1r} & a_{2r} & \dots & a_{rr} & \dots & a_{nr} \end{bmatrix}$$

có hạng $r$, suy ra $\operatorname{rank}(A_r) = r$. Theo Định lý 2.1.10 ta có hệ $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r)$ độc lập tuyến tính.

Lấy vector $\mathbf{a}_j$ tùy ý với $r < j \leq m$. Với mỗi $i = 1, \ldots, n$, ta xét định thức cấp $r+1$ sau

$$D_{r+1}(i) = \begin{vmatrix} a_{11} & a_{21} & \ldots & a_{r1} & a_{i1} \\ a_{12} & a_{22} & \ldots & a_{r2} & a_{i2} \\ \multicolumn{5}{c}{\dotfill} \\ a_{1r} & a_{2r} & \ldots & a_{rr} & a_{ir} \\ a_{1j} & a_{2j} & \ldots & a_{rj} & a_{ij} \end{vmatrix}$$

Khi $1 \leq i \leq r$ thì $D_{r+1}(i) = 0$ vì nó có hai cột như nhau cột thứ $i$ và cột cuối. Khi $r < i \leq n$ thì $D_{r+1}(i) = 0$ vì nó là định thức con của $A^t$ cấp $r+1$. Như vậy $D_{r+1}(i) = 0$ với mọi $i = 1, \ldots, n$.

Mặt khác, khai triển định thức trên theo cột cuối ta có

$$D_{r+1}(i) = a_{ij} D_r + \sum_{l=1}^{r} a_{il} K_l$$

trong đó $K_l$ là $(-1)^{r+1+l}$ nhân với định thức của ma trận thu được từ ma trận sau bỏ đi hàng thứ $l$

$$\begin{bmatrix} a_{11} & a_{21} & \ldots & a_{r1} \\ a_{12} & a_{22} & \ldots & a_{r2} \\ \multicolumn{4}{c}{\dotfill} \\ a_{1r} & a_{2r} & \ldots & a_{rr} \\ a_{1j} & a_{2j} & \ldots & a_{rj} \end{bmatrix}$$

Vậy

$$a_{ij} = \sum_{l=1}^{r} \Big( - \frac{K_l}{D_r} \Big) a_{il} \qquad i = 1, \ldots, n, \qquad r < j \leq m.$$

Do đó $\mathbf{a}_j = \sum_{l=1}^{r} \big( - \frac{K_l}{D_r} \big) \mathbf{a}_l$. Suy ra hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r, \mathbf{a}_j)$ phụ thuộc tuyến tính với mọi $j = r+1, \ldots, m$. Vậy hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r)$ là hệ độc lập tuyến tính tối đại. Từ đó ta có $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m) = \operatorname{rank} A$. $\square$

**3.3.2 Hệ quả.** *Hệ* $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ *độc lập tuyến tính khi và chỉ khi* $\det A \neq 0$.

*Chứng minh.* Ta biết rằng $\det A \neq 0$ khi và chỉ khi $\operatorname{rank} A = n$ hay $\operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_n) = n$. Điều này tương đương với hệ $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ độc lập tuyến tính. □

**3.3.3 Thí dụ.** Trong $\mathbb{R}^4$ xét ba vector $\mathbf{a}_1 = (4, 1, 0, 2)$; $\mathbf{a}_2 = (3, 0, 1, 2)$; $\mathbf{a}_3 = (1, 1, -1, 0)$. Lập ma trận

$$A^t = \begin{bmatrix} 4 & 1 & 0 & 2 \\ 3 & 0 & 1 & 2 \\ 1 & 1 & -1 & 0 \end{bmatrix}.$$

Thực hiện các phép biến đổi sơ cấp trên dòng để đưa ma trận $A^t$ về ma trận bậc thang tương ứng.

$$A^t \xrightarrow{d_1 \leftrightarrow d_3} \begin{bmatrix} 1 & 1 & -1 & 0 \\ 3 & 0 & 1 & 2 \\ 4 & 1 & 0 & 2 \end{bmatrix} \xrightarrow[d_3 \to d_3 - 4d_1]{d_2 \to d_2 - 3d_1} \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & -3 & 4 & 2 \\ 0 & -3 & 4 & 2 \end{bmatrix}$$
$$\xrightarrow{d_4 \to d_4 - d_3} \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & -3 & 4 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Vậy $\operatorname{rank} A = 2$. Do đó $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = 2$ và hệ con độc lập tuyến tính tối đại là $(\mathbf{a}_2, \mathbf{a}_3)$. □

**3.3.4 Thí dụ.** Xét $\mathbb{R}^2$ là mặt phẳng tọa độ $Oxy$ và 3 vector $\overrightarrow{OA} = (a_1, a_2)$, $\overrightarrow{OB} = (b_1, b_2)$, $\overrightarrow{OC} = (c_1, c_2)$. Hệ $((a_1, a_2), (b_1, b_2)) = (\overrightarrow{OA}, \overrightarrow{OB})$ độc lập tuyến tính khi và chỉ khi

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0 \iff \frac{a_1}{b_1} \neq \frac{a_2}{b_2} \quad (\text{nếu } b_1 \neq 0 \text{ và } b_2 \neq 0)$$

Điều đó tương đương với $\overrightarrow{OA}$, $\overrightarrow{OB}$ không cùng phương hay $A, O, B$ không thẳng hàng.

Do đó ta cũng có hệ $((a_1, a_2), (c_1, c_2)) = (\overrightarrow{OA}, \overrightarrow{OC})$ phụ thuộc tuyến tính khi và chỉ khi $\overrightarrow{OA}$, $\overrightarrow{OC}$ cùng phương hay $A, O, C$ thẳng hàng. □

**3.3.5 Thí dụ.** Xét $\mathbb{R}^3$ là không gian tọa độ $Oxyz$ và 4 vector $\overrightarrow{OA} = (a_1, a_2, a_3)$, $\overrightarrow{OB} = (b_1, b_2, b_3)$, $\overrightarrow{OC} = (c_1, c_2, c_3)$, $\overrightarrow{OD} = (d_1, d_2, d_3)$.

Hệ $((a_1, a_2, a_3), (b_1, b_2, b_3)) = (\overrightarrow{OA}, \overrightarrow{OB})$ độc lập tuyến tính khi và chỉ khi

$$\begin{aligned}
&\operatorname{rank}\left(\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\right) = 2 \\
\iff & \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2 + \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2 \neq 0 \\
\iff & (a_1 : a_2 : a_3) \neq (b_1 : b_2 : b_3) \qquad (\text{nếu } a_1, a_2, a_3, b_1, b_2, b_3 \text{ khác } 0)
\end{aligned}$$

Điều đó tương đương với $\overrightarrow{OA}$, $\overrightarrow{OB}$ không cùng phương.

Giả sử $\overrightarrow{OA}$ và $\overrightarrow{OB}$ không cùng phương. Khi đó hệ

$$((a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3))$$

phụ thuộc tuyến tính khi và chỉ khi

$$\begin{aligned}
\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0 &\iff c_i = \lambda a_i + \mu b_i; \quad i = 1, 2, 3 \\
&\iff \overrightarrow{OC} \text{ thuộc mặt phẳng } (OAB)
\end{aligned}$$

Điều đó tương đương $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ đồng phẳng hay $O, A, B, C$ đồng phẳng.

Do đó ta có hệ $((a_1, a_2, a_3), (b_1, b_2, b_3), (d_1, d_2, d_3))$ độc lập tuyến tính khi và chỉ khi $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ không đồng phẳng hay $O, A, B, C$ không đồng phẳng. □

## Bài tập

**1**) Tìm một hệ con độc lập tuyến tính tối đại và hạng của các hệ vector sau trong không gian $\mathbb{R}^3$.

**(a)** $(\mathbf{u}_1 = (1, 1, 1), \mathbf{u}_2 = (1, 2, 1))$

**(b)** $(\mathbf{u}_1 = (1, 0, -1), \mathbf{u}_2 = (0, 1, -1), \mathbf{u}_3 = (1, -1, 0))$

**(c)** $(\mathbf{u}_1 = (2,1,0), \mathbf{u}_2 = (0,-2,1), \mathbf{u}_3 = (2,-1,2))$

**(d)** $(\mathbf{u}_1 = (1,-1,0), \mathbf{u}_2 = (2,-1,-1), \mathbf{u}_3 = (0,1,-1), \mathbf{u}_4 = (2,0,-2))$.

**2 )** Tìm hạng và hệ con độc lập tuyến tính tối đại của các hệ vector sau:

**(a)** $(\mathbf{a}_1 = (1,0,0,-1), \quad \mathbf{a}_2 = (2,1,1,0), \quad \mathbf{a}_3 = (1,1,1,1), \mathbf{a}_4 = (1,2,3,4), \mathbf{a}_5 = (0,1,2,3))$ trong $\mathbb{R}^4$.

**(b)** $(\mathbf{a}_1 = (1,1,1,1,0), \quad \mathbf{a}_2 = (1,1,-1,-1,-1), \quad \mathbf{a}_3 = (2,2,0,0,1), \mathbf{a}_4 = (1,1,5,5,2), \quad \mathbf{a}_5 = (1,-1,-1,0,0))$ trong $\mathbb{R}^5$.

**3 )** Chứng minh rằng nếu hệ $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ biểu thị tuyến tính được qua hệ $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m)$ thì hạng của hệ đầu không vượt quá hạng của hệ sau.

**4 )** Cho hai hệ vector cùng hạng, hệ đầu biểu thị tuyến tính được qua hệ sau. Chứng minh rằng hai hệ vector tương đương.

**5 )** Chứng minh rằng nếu vector $\mathbf{y}$ biểu thị tuyến tính được qua hệ vector $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ thì $\operatorname{rank}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \operatorname{rank}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{y})$.

**6 )** Chứng minh bổ đề cơ bản về sự phụ thuộc tuyến tính (Định lý 3.1.5) bằng quy nạp theo $n$.

# § 4 Cơ sở, số chiều, và tọa độ

## 4.1 Cơ sở và số chiều của không gian vector

**4.1.1 Định nghĩa.** Cho không gian vector $V$ và $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là hệ vector trong $V$. Hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ được gọi là **hệ sinh** của $V$ nếu mọi vector của $V$ đều biểu thị tuyến tính được qua hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. Nghĩa là với mọi $\mathbf{x} \in V$ tồn tại $\lambda_i \in \mathbb{K}$ với $i = 1, \ldots, n$ sao cho $\mathbf{x} = \sum_{i=1}^{n} \lambda_i \mathbf{a}_i$.

Hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ được gọi là **cơ sở** (basis) của không gian vector $V$ nếu nó vừa là hệ sinh vừa độc lập tuyến tính.

**4.1.2 Mệnh đề.** *Các hệ sinh của một không gian vector là tương đương.*

*Chứng minh.* Giả sử $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ và $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ là hai hệ sinh của không gian vector $V$. Theo định nghĩa của hệ sinh ta thấy: hệ $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ biểu thị tuyến tính được qua hệ $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ và hệ $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ biểu thị tuyến tính được qua hệ $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$. Vậy hai hệ $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ và $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ tương đương. □

**4.1.3 Mệnh đề.** *Nếu trong một không gian vector có hai cơ sở hữu hạn, thì chúng có số vector bằng nhau.*

*Chứng minh.* Giả sử $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ và $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ là hai cơ sở của không gian vector $V$. Theo mệnh đề trên ta có hai hệ $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ và $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ tương đương. Hơn nữa, hai hệ trên là độc lập tuyến tính. Do đó theo Hệ quả 3.1.6 ta có $m = n$. □

**4.1.4 Định nghĩa.** Ta bảo một không gian vector $V$ là một **không gian $n$ chiều** nếu trong $V$ có ít nhất một cơ sở có $n$ vector. Ta cũng bảo số chiều của $V$ là $n$ và kí hiệu $\dim V = n$.

Không gian vector chỉ có một vector duy nhất (vector không): $V = \{\mathbf{0}\}$ không gian này không có cơ sở ta quy ước $\dim V = 0$. Các không gian $n$ chiều ($n \in \mathbb{N}$ tùy ý) được gọi chung là các *không gian hữu hạn chiều*.

Không gian vector $V$ được gọi là *không gian vô hạn chiều*, ký hiệu $\dim V = +\infty$, nếu nó không hữu hạn chiều, tức là không tồn tại cơ sở có hữu hạn vector.

**4.1.5 Thí dụ.** Trong không gian $\mathbb{K}^n$, xét hệ vector $\mathbf{e}_1 = (1, 0, \dots, 0)$; $\mathbf{e}_2 = (0, 1, \dots, 0)$; …; $\mathbf{e}_n = (0, 0, \dots, 1)$. Ta có $\lambda_1\mathbf{e}_1 + \lambda_2\mathbf{e}_2 + \cdots + \lambda_n\mathbf{e}_n = \mathbf{0}$ khi và chỉ khi $(\lambda_1, \lambda_2, \dots, \lambda_n) = (0, 0, \dots, 0)$ hay $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. Vậy hệ $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ độc lập tuyến tính.

Mặt khác, ta cũng có $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ là hệ sinh của $\mathbb{K}^n$ vì với $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{K}^n$ tùy ý đều biểu thị tuyến tính được qua hệ vector $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ cụ thể là $\mathbf{x} = (x_1, x_2, \dots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$. Vậy $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ là một cơ sở của $\mathbb{K}^n$ và $\dim \mathbb{K}^n = n$. Cơ sở này gọi là cơ sở chính tắc của $\mathbb{K}^n$. □

**4.1.6 Thí dụ.** Xét không gian vector $V = \mathcal{M}_{m\times n}(\mathbb{K})$, không gian các ma trận cấp $m \times n$ với các hệ số trong $\mathbb{K}$ và với phép toán cộng hai ma trận thông thường và nhân một số với ma trận. Xét các ma trận

$$E_{ij} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

trong đó tất cả các phần tử đều bằng không chỉ có phần tử ở hàng $i$ cột $j$ là bằng 1. Khi đó ta có thể chứng minh được hệ ma trận $\mathscr{E} = \{E_{ij} : i = 1, \dots, m; j = 1, \dots, n\}$ là một cơ sở của không gian $V = \mathcal{M}_{m\times n}(\mathbb{K})$. Vậy $\dim \mathcal{M}_{m\times n}(\mathbb{K}) = mn$. $\square$

**4.1.7 Thí dụ.** Không gian vector $\mathbb{K}[x]$, các đa thức trên $\mathbb{K}$, là không gian vô hạn chiều. Thật vậy, giả sử $(p_1(x), p_2(x), \dots, p_n(x))$ là một cơ sở của $\mathbb{K}[x]$. Đặt $m = \max\{\deg p_1(x), \deg p_2(x), \dots, \deg p_n(x)\}$. Xét đa thức $p(x) = x^{m+1}$. Đa thức này không biểu thị tuyến tính được qua hệ $(p_1(x), p_2(x), \dots, p_n(x))$ bởi vì các tổ hợp tuyến tính của hệ $(p_1(x), p_2(x), \dots, p_n(x))$ đều có bậc nhỏ hơn hoặc bằng $m$ trong khi đó $\deg x^{m+1} = m + 1$. Điều đó mâu thuẫn với giả thiết hệ $(p_1(x), p_2(x), \dots, p_n(x))$ là cơ sở của $\mathbb{K}[x]$. Vậy $\mathbb{K}[x]$ không có cơ sở hữu hạn. Do đó $\dim \mathbb{K}[x] = +\infty$. $\square$

**4.1.8 Định lý.** *Cho $V$ là không gian vector $n$ chiều với $n > 0$. Khi đó,*

**(a)** *Mọi hệ có nhiều hơn $n$ vector đều phụ thuộc tuyến tính.*

**(b)** *Một hệ vector là cơ sở của $V$ khi và chỉ khi hệ vector đó độc lập tuyến tính và có đúng $n$ vector.*

**(c)** *Một hệ vector là cơ sở của $V$ khi và chỉ khi hệ vector đó là hệ sinh và có đúng $n$ vector.*

**(d)** *Một hệ vector độc lập tuyến tính gồm $k$ (với $k < n$) vector đều có thể bổ sung $n - k$ vector nữa để được một cơ sở của $V$.*

*Chứng minh.* Vì $\dim V = n$ nên trong $V$ có cơ sở gồm $n$ vectơ. Giả sử trong $V$ có cơ sở là $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$.

**(a)** Giả sử $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$ với $m > n$ là một hệ vector độc lập tuyến tính của $V$. Do $\mathscr{A}$ là cơ sở của $V$ nên hệ $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$ biểu thị tuyến tính được qua hệ $\mathscr{A}$. Theo Định lý 3.1.5, bổ đề cơ bản về phụ thuộc tuyến tính, ta phải có $m \leq n$. Điều này mâu thuẫn với giả thiết $m > n$, cho nên không có hệ độc lập tuyến tính có số vector nhiều hơn $n$.

**(b)** Giả sử hệ vector $\mathscr{B}$ là một cơ sở của $V$. Vì $\dim V = n$, nên trong $\mathscr{B}$ có $n$ vector và hệ đó độc lập tuyến tính (theo định nghĩa cơ sở).

Ngược lại, giả sử hệ vector $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ là một hệ độc lập tuyến tính trong $V$. Vì $\mathscr{A}$ là một cơ sở của $V$ nên hệ $\mathscr{B}$ vừa độc lập tuyến tính vừa biểu thị tuyến tính được qua hệ $\mathscr{A}$. Theo Định lý 3.1.5, bổ đề cơ bản về sự phụ thuộc tuyến tính, ta suy ra hệ $\mathscr{B}$ tương đương hệ $\mathscr{A}$. Mà hệ $\mathscr{A}$ là hệ sinh của $V$ nên hệ $\mathscr{B}$ là hệ sinh của $V$. Vậy hệ $\mathscr{B}$ là cơ sở của $V$.

**(c)** Giả sử hệ vector $\mathscr{B}$ là một cơ sở của $V$. Vì $\dim V = n$, nên trong $\mathscr{B}$ có $n$ vector và hệ đó là hệ sinh của $V$ (theo định nghĩa cơ sở).

Ngược lại, giả sử $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ là một hệ sinh của $V$. Ta chứng minh nó là cơ sở của $V$. Vì $\mathscr{B}$ là hệ sinh của $V$ nên trong $\mathscr{B}$ phải có vector khác không. Từ đó ta tìm được trong $\mathscr{B}$ một hệ con độc lập tuyến tính tối đại, kí hiệu là $\mathscr{B}'$. Ta phải có $\mathscr{B}'$ tương đương với $\mathscr{B}$. Do $\mathscr{B}$ là hệ sinh của $V$ nên $\mathscr{B}'$ là cũng hệ sinh của $V$. Suy ra $\mathscr{B}'$ là cơ sở của $V$. Vì $\dim V = n$ nên trong $\mathscr{B}'$ có $n$ vector, cho nên $\mathscr{B}' = \mathscr{B}$ hay $\mathscr{B}$ cơ sở của $V$.

**(d)** Giả sử $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k)$ là hệ độc lập tuyến tính của $V$ (với $k < n$). Ta thấy hệ $\mathscr{B}$ độc lập tuyến tính và biểu thị tuyến tính được qua hệ $(\mathscr{A})$. Theo Định lý 3.1.5, bổ đề cơ bản về sự phụ thuộc tuyến tính, ta có thể thay $\mathbf{b}_i$ vào $\mathbf{a}_{i_l}$ nào đó với $l = 1, \dots, k$ để được hệ mới $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k, \mathbf{a}_{i_{k+1}}, \dots, \mathbf{a}_{i_n})$ tương đương với hệ $\mathscr{A}$. Vì hệ $\mathscr{A}$ là hệ sinh của không gian vector $V$ nên hệ $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k, \mathbf{a}_{i_{k+1}}, \dots, \mathbf{a}_{i_n})$

cũng là hệ sinh của $V$. Hơn nữa nó có $n$ vector nên theo tính chất trên hệ $(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k, \mathbf{a}_{i_{k+1}}, \ldots, \mathbf{a}_{i_n})$ là cơ sở của $V$. $\square$

Khái niệm cơ sở có thể được định nghĩa cho không gian vector tùy ý như sau.

**4.1.9 Định nghĩa.** Hệ $\mathscr{B} = (\mathbf{x}_i)_{i \in I}$ được gọi là một *cơ sở* của không gian vector $V$ nếu $\mathscr{B}$ độc lập tuyến tính và mọi vector của $V$ đều biểu thị tuyến tính được qua $\mathscr{B}$. Không gian vector $V$ có số chiều hữu hạn hay vô hạn tùy vào $\mathscr{B}$ hữu hạn hay vô hạn.

**4.1.10 Thí dụ.** Dễ dàng chứng minh được hệ $\mathscr{C} = (1, x, x^2, \ldots, x^n, \ldots)$ là một cơ sở của $\mathbb{K}[x]$. $\square$

## 4.2 Tọa độ

**4.2.1 Mệnh đề.** *Cho $V$ là một không gian vector có cơ sở $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. Với $\mathbf{x} \in V$ tùy ý, $\mathbf{x}$ được biểu diễn duy nhất dưới dạng $\mathbf{x} = k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_n\mathbf{a}_n$, với $k_i \in \mathbb{K}$.*

*Chứng minh.* Vì $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một cơ sở của $V$ nên $\mathbf{x}$ biểu thị tuyến tính được qua hệ đó. Giả sử chúng ta có hai biểu thị tuyến tính

$$\mathbf{x} = \sum_{i=1}^{n} k_i\mathbf{a}_i \qquad \text{và} \qquad \mathbf{x} = \sum_{i=1}^{n} \lambda_i\mathbf{a}_i.$$

Thực hiện phép trừ ta được $\mathbf{0} = \mathbf{x} - \mathbf{x} = \sum_{i=1}^{n}(k_i - \lambda_i)\mathbf{a}_i$. Vì hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là cơ sở nên nó độc lập tuyến tính, suy ra $k_i - \lambda_i = 0$ với mọi $i = 1, \ldots, n$ hay $k_i = \lambda_i$ với mọi $i = 1, \ldots, n$. $\square$

**4.2.2 Định nghĩa.** Bộ số $(k_1, k_2, \ldots, k_n)$ trong mệnh đề trên gọi là **tọa độ của $\mathbf{x}$ đối với cơ sở** $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. Ta kí hiệu tọa độ của $\mathbf{x}$ đối với cơ sở $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là: $\mathbf{x}|_{(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)} = (k_1, k_2, \ldots, k_n)$. Mỗi $k_i$ gọi là *tọa độ thứ $i$* của $\mathbf{x}$ với $i = 1, 2, \ldots, n$

**4.2.3 Thí dụ.** Trong $\mathbb{R}^3$ cho ba vector $\mathbf{a}_1 = (1,1,0)$, $\mathbf{a}_2 = (0,1,1)$, $\mathbf{a}_3 = (1,0,1)$. Hệ $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ là một cơ sở của $\mathbb{R}^3$ vì nó gồm có ba vector độc lập tuyến tính do

$$\begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{vmatrix} = 2.$$

Ta hãy tìm tọa độ của vector $\mathbf{x} = (4,3,5)$ đối với cơ sở $\mathscr{A}$. Nghĩa là ta tìm các hệ số $k_1, k_2, k_3$ thỏa $\mathbf{x} = k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + k_3\mathbf{a}_3$. Ta phải giải hệ phương trình tuyến tính

$$\begin{cases} k_1 + k_3 = 4 \\ k_1 + k_2 = 3 \\ k_2 + k_3 = 5 \end{cases}$$

Ta biến đổi ma trận bổ sung

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 5 \end{array}\right] \xrightarrow{d_2 \to d_2 - d_1} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & 1 & 5 \end{array}\right]$$

$$\xrightarrow{d_3 \to d_3 - d_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 2 & 6 \end{array}\right]$$

Từ đó ta được $k_3 = 3$, $k_2 = 2$, $k_1 = 1$. Vậy $\mathbf{x}|_{\mathscr{A}} = (1,2,3)$. □

**4.2.4** Giả sử $\mathscr{B} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một cơ sở trong không gian $\mathbb{K}^n$ với $\mathbf{a}_i = (a_{1i}, a_{2i}, \ldots, a_{ni})$ và cho $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{K}^n$. Khi đó $\mathbf{b}|_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$ khi và chỉ khi

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

Do đó, bài toán tìm tọa độ của **b** đối với cơ sở $\mathscr{B}$ đưa về phép giải hệ phương trình tuyến tính trên. Hệ này là hệ Cramer và có nghiệm duy nhất vì

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \neq 0$$

(do $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ độc lập tuyến tính).

**4.2.5 Mệnh đề.** *Trong không gian vector $n$ chiều $V$ cho một cơ sở bất kì $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ và một hệ vector $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$. Đặt $\mathbf{b}_j|_{\mathscr{A}} = (b_{1j}, b_{2j}, \dots, b_{nj})$ và $B = (b_{ij})_{n\times m}$. Khi đó, hệ $\mathscr{B}$ độc lập tuyến tính khi và chỉ khi* $\operatorname{rank} B = m$.

*Chứng minh.* Theo giả thiết ta có biểu thị tuyến tính $\mathbf{b}_j = \sum_{i=1}^{n} b_{ij}\mathbf{a}_i$. Khi đó, xét tổ hợp tuyến tính

$$\mathbf{0} = \sum_{j=1}^{m} \lambda_j \mathbf{b}_j = \sum_{j=1}^{m} \lambda_j \sum_{i=1}^{n} b_{ij}\mathbf{a}_i = \sum_{i=1}^{n} \Big( \sum_{j=1}^{m} \lambda_j b_{ij} \Big) \mathbf{a}_i.$$

Do hệ $\mathscr{A}$ là cơ sở ta phải có $\sum_{j=1}^{m} \lambda_j b_{ij} = 0$ với mọi $i = 1, 2, \dots, n$. Vậy hệ $\mathscr{B}$ là độc lập tuyến tính khi và chỉ khi các $\lambda_j = 0$ với mọi $j = 1, \dots, m$. Điều này tương đương với hệ phương trình $\sum_{j=1}^{m} \lambda_j b_{ij} = 0$ với mọi $i = 1, 2, \dots, n$ có nghiệm duy nhất, nghĩa là $\operatorname{rank} B = m$. Do đó, hệ $\mathscr{B}$ độc lập tuyến tính khi và chỉ khi $\operatorname{rank} B = m$. □

**4.2.6 Mệnh đề.** *Trong không gian vector $n$ chiều $V$ cho một cơ sở bất kì $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ và một hệ vector $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$. Đặt $\mathbf{b}_j|_{\mathscr{A}} = (b_{1j}, b_{2j}, \dots, b_{nj})$ và $B = (b_{ij})_{n\times m}$. Khi đó,* $\operatorname{rank}\mathscr{B} = \operatorname{rank} B$.

*Chứng minh.* Nếu hệ $\mathscr{B}$ toàn vector không thì rõ ràng $\operatorname{rank}\mathscr{B} = 0 = \operatorname{rank} B$. Ngược lại, trong $\mathscr{B}$ có một hệ con độc lập tuyến tính tối đại $(\mathbf{b}_{k_1}, \dots, \mathbf{b}_{k_r})$. Do đó, $\operatorname{rank}\mathscr{B} = r$. Theo Mệnh đề 4.2.5 ta có hạng của ma trận

$$B' = \begin{bmatrix} b_{1k_1} & \dots & b_{1k_r} \\ b_{2k_1} & \dots & b_{2k_r} \\ \dots & \dots & \dots \\ b_{nk_1} & \dots & b_{nk_r} \end{bmatrix}$$

là $r$. Ta nhận thấy $B'$ là ma trận con của ma trận $B$, cho nên $\operatorname{rank} B \geq r$.

Giả sử $\operatorname{rank} B = s$. Đặt $\mathbf{x}_j = (b_{1j}, b_{2j}, \dots, b_{nj}) \in \mathbb{R}^n$ với mọi $j = 1, 2, \dots, m$. Theo Định lý 3.3.1 ta có $\operatorname{rank}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) = s$. Do đó trong hệ $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$ tồn tại hệ con độc lập tuyến tính tối đại $(\mathbf{x}_{l_1}, \dots, \mathbf{x}_{l_s})$. Cũng theo Định lý 3.3.1 ta có hạng của ma trận

$$B'' = \begin{bmatrix} b_{1l_1} & \dots & b_{1l_s} \\ b_{2l_1} & \dots & b_{2l_s} \\ \dots & \dots & \dots \\ b_{nl_1} & \dots & b_{nl_s} \end{bmatrix}$$

bằng $s$. Từ đó theo Mệnh đề 4.2.5 ta được hệ $(\mathbf{b}_{l_1}, \dots, \mathbf{b}_{l_s})$ độc lập tuyến tính. Hơn nữa, hệ ấy là hệ con của $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$, cho nên $\operatorname{rank} \mathscr{B} \geq s = \operatorname{rank} B$. Vậy ta phải có $\operatorname{rank} \mathscr{B} = \operatorname{rank} B$. □

## 4.3 Ma trận chuyển cơ sở và công thức đổi tọa độ

**4.3.1** Giả sử trong không gian vector $n$ chiều $V$ có hai cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ và $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n)$. Giả sử tọa độ của mỗi vector $\mathbf{e}'_i$ đối với cơ sở $\mathscr{B}$ được cho bởi $\mathbf{e}'_i|_{\mathscr{B}} = (a_{1i}, a_{2i}, \dots, a_{ni})$ với $i = 1, 2, \dots, n$. Ta thiết lập ma trận

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

đó chính là việc ghép các ma trận cột của các tọa độ $\mathbf{e}'_i$ đối với cơ sở $\mathscr{B}$. Ma trận này được gọi là **ma trận đổi (chuyển) cơ sở (tọa độ)** từ $\mathscr{B}$ sang

$\mathscr{B}'$. Mặt khác, ta thấy $\mathbf{e}'_j = \sum_{i=1}^n a_{ij}\mathbf{e}_i$ với $i = 1, \dots, n$; do đó, ta có thể viết gọn lại dưới dạng ma trận như sau

$$\begin{bmatrix}\mathbf{e}'_1 & \mathbf{e}'_2 & \dots & \mathbf{e}'_n\end{bmatrix} = \begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n\end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \hdotsfor{4} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Bây giờ giả sử $\mathbf{x} \in V$ bất kì có tọa độ đối với $\mathscr{B}$ và $\mathscr{B}'$ là $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \dots, x_n)$ và $\mathbf{x}|_{\mathscr{B}'} = (x'_1, x'_2, \dots, x'_n)$. Ta sẽ tìm mối liên hệ giữa $(x_1, x_2, \dots, x_n)$ và $(x'_1, x'_2, \dots, x'_n)$. Ta có

$$\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i = \sum_{j=1}^n x'_j\mathbf{e}'_j = \sum_{j=1}^n x'_j \sum_{i=1}^n a_{ij}\mathbf{e}_i = \sum_{i=1}^n \Big(\sum_{j=1}^n a_{ij}x'_j\Big)\mathbf{e}_i.$$

Do đó ta có $x_i = \sum_{j=1}^n a_{ij}x'_j$ với $i = 1, 2, \dots, n$, và ta có thể viết dưới dạng ma trận như sau

(4.3.2) $$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \hdotsfor{4} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix} \quad \text{hay} \quad \mathbf{x}|_{\mathscr{B}} = A\mathbf{x}|_{\mathscr{B}'}.$$

Công thức trên được gọi là **công thức đổi tọa độ**. Ta có biểu đồ giao hoán

$$\begin{tikzcd} & \mathbf{x} \arrow[dl, "\mathbf{x}|_{\mathscr{B}'}"'] \arrow[dr, "\mathbf{x}|_{\mathscr{B}}"] & \\ \mathscr{B}' \arrow[rr, "A"'] & & \mathscr{B} \end{tikzcd}$$

Như trên ta đã biết ma trận $A$ được gọi là ma trận chuyển cơ sở từ $\mathscr{B}$ sang $\mathscr{B}'$, nhưng để dễ nhớ và dựa vào phần tử của nó ta có thể gọi ma trận $A$ là ma trận tọa độ (theo cột) của cơ sở $\mathscr{B}'$ đối với $\mathscr{B}$.

**4.3.3 Thí dụ.** Trong không gian $\mathbb{R}^3$ xét cơ sở $\mathscr{B} = (\mathbf{u}_1 = (1, 1, 0), \mathbf{u}_2 = (0, 1, 1), \mathbf{u}_3 = (1, 0, 1))$. Khi đó ma trận chuyển từ cơ sở chính tắc $\mathscr{E}$ sang cơ sở $\mathscr{B}$ là

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

Khi đó công thức đổi tọa

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix}$$

Giả sử $\mathbf{x} \in \mathbb{R}^3$ có tọa độ đối với cơ sở $\mathscr{B}$ là $\mathbf{x}|_{\mathscr{B}} = (1, 2, 3)$. Khi đó ta có

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}$$

Vậy $\mathbf{x} = (4, 3, 5)$. □

**4.3.4 Mệnh đề.** *Giả sử trong không gian vector $n$ chiều $V$ có cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$. Khi đó ma trận đổi cơ sở từ $\mathscr{B}$ sang $\mathscr{B}$ là ma trận đơn vị $I_n$.*

*Chứng minh.* Theo cách thành lập ma trận đổi cơ sở ở trên, ta thấy ma trận đó là duy nhất. Ta cũng thấy được tính chất này là hiển nhiên. □

**4.3.5 Mệnh đề.** *Giả sử trong không gian vector $n$ chiều $V$ có ba cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$, $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n)$ và $\mathscr{B}'' = (\mathbf{e}''_1, \mathbf{e}''_2, \dots, \mathbf{e}''_n)$. Gọi $A$ là ma trận chuyển tọa độ từ $\mathscr{B}$ sang $\mathscr{B}'$, $B$ là ma trận chuyển tọa độ từ $\mathscr{B}'$ sang $\mathscr{B}''$ và $C$ là ma trận chuyển tọa độ từ $\mathscr{B}$ sang $\mathscr{B}''$. Khi đó $C = AB$. Ta có biểu đồ hình thức sau*

$$\begin{tikzcd} \mathscr{B}'' \arrow[rr, "B"] \arrow[dr, "C"'] & & \mathscr{B}' \arrow[dl, "A"] \\ & \mathscr{B} & \end{tikzcd}$$

*Ta chú ý đến hướng mũi tên và thứ tự trong phép nhân ma trận.*

*Chứng minh.* Theo giả thiết ta có

$$\begin{bmatrix} \mathbf{e}'_1 & \mathbf{e}'_2 & \dots & \mathbf{e}'_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{bmatrix} A$$

và

$$\begin{bmatrix}\mathbf{e}''_1 & \mathbf{e}''_2 & \dots & \mathbf{e}''_n\end{bmatrix} = \begin{bmatrix}\mathbf{e}'_1 & \mathbf{e}'_2 & \dots & \mathbf{e}'_n\end{bmatrix} B.$$

Suy ra

$$\begin{bmatrix}\mathbf{e}''_1 & \mathbf{e}''_2 & \dots & \mathbf{e}''_n\end{bmatrix} = \left(\begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n\end{bmatrix} A\right)B = \begin{bmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n\end{bmatrix}(AB).$$

Vậy ta có $C = AB$. □

**4.3.6 Mệnh đề.** *Giả sử trong không gian vector $n$ chiều $V$ có hai cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ và $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n)$. Gọi $A$ là ma trận chuyển tọa độ từ $\mathscr{B}$ sang $\mathscr{B}'$. Khi đó $A$ là ma trận không suy biến. Hơn nữa, $A^{-1}$ là ma trận chuyển tọa độ từ $\mathscr{B}'$ sang $\mathscr{B}$.*

*Chứng minh.* Gọi $B$ là ma trận chuyển cơ sở từ $\mathscr{B}'$ sang $\mathscr{B}$. Theo các tính chất trên ta được ma trận chuyển cơ sở từ $\mathscr{B}$ sang $\mathscr{B}$ qua $\mathscr{B}'$ là $AB$ và phải bằng $I_n$, và ma trận chuyển từ $\mathscr{B}'$ sang $\mathscr{B}'$ qua $\mathscr{B}$ là $BA$ và phải bằng $I_n$. Vậy $AB = BA = I_n$, suy ra $A$ khả nghịch và nghịch đảo của nó là $B$. □

**4.3.7 Thí dụ.** Trong không gian $\mathbb{R}^3$ cho hai cơ sở $\mathscr{A} = (\mathbf{a}_1 = (1,1,1);\ \mathbf{a}_2 = (-1,1,2); \mathbf{a}_3 = (1,2,3))$ và $\mathscr{B} = (\mathbf{b}_1 = (1,0,1); \mathbf{b}_2 = (0,1,1);\ \mathbf{b}_3 = (1,1,0))$. Đặt

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

Ta xét biểu đồ sau

$$\begin{tikzcd}
\mathscr{B} \arrow[rr, "C"] \arrow[dr, "B"'] & & \mathscr{A} \arrow[dl, "A"] \\
& \mathscr{E} &
\end{tikzcd}$$

Theo các tính chất trên thì ma trận chuyển tọa độ từ $\mathscr{A}$ sang $\mathscr{B}$ là $C =$

$A^{-1}B$. Ta tìm $A^{-1}$

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow[d_3 \to d_3 - d_1]{d_2 \to d_2 - d_1} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & -1 & 1 & 0 \\ 0 & 3 & 2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{d_2 \to \frac{1}{2} d_2} \left[\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 3 & 2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[d_3 \to d_3 - 3d_2]{d_1 \to d_1 + d_2} \left[\begin{array}{ccc|ccc} 1 & 0 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{3}{2} & 1 \end{array}\right]$$

$$\xrightarrow[d_3 \to 2d_3]{\substack{d_1 \to d_1 + 3d_3 \\ d_2 \to d_2 - d_3}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 5 & -3 \\ 0 & 1 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & 1 & -3 & 2 \end{array}\right]$$

Vậy

$$C = \begin{bmatrix} -1 & 5 & -3 \\ -1 & 2 & -1 \\ 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -4 & 2 & 4 \\ -2 & 1 & 1 \\ 3 & -1 & -2 \end{bmatrix}$$

Ma trận $C$ là ma trận chuyển cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$. □

**4.3.8 Nhận xét.** Qua thí dụ trên ta có kết quả tổng quát sau: Để tìm ma trận chuyển cơ sở trong $\mathbb{R}^n$ từ $\mathscr{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ sang $\mathscr{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ ta lập ma trận khối bổ sung $(A|B)$ trong đó $A$ và $B$ lần lượt là ma trận tọa độ của các cơ sở $\mathscr{A}$ và $\mathscr{B}$; sau đó dùng phép biến đổi sơ cấp theo dòng đưa ma trận khối ấy về dạng $(I_n|C)$; ma trận $C$ chính là ma trận chuyển cơ sở cần tìm.

**4.3.9 Thí dụ.** Trong $\mathbb{R}^3$ cho cơ sở $\mathscr{B} = (\mathbf{u}_1 = (1,1,0), \mathbf{u}_2 = (0,1,1), \mathbf{u}_3 = (1,0,1))$ và cơ sở $\mathscr{B}' = (\mathbf{u}'_1 = (2,1,1), \mathbf{u}'_2 = (1,2,1), \mathbf{u}'_3 = (1,1,2))$. Giả sử $\mathbf{x} \in \mathbb{R}^3$ có tọa độ đối với cơ sở $\mathscr{B}$ là $\mathbf{x}|_{\mathscr{B}} = (6,5,5)$. Hãy tìm $\mathbf{x}$ và tọa độ của nó đối với cơ sở $\mathscr{B}'$.

*Giải.* Từ công thức chuyển tọa độ ta tìm được tọa độ của $\mathbf{x}$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 11 \\ 11 \\ 10 \end{bmatrix}$$

Vậy $\mathbf{x} = (11, 11, 10)$. Để tìm tọa độ của $\mathbf{x}$ đối với cơ sở $\mathscr{B}'$ ta biến đổi

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 11 \\ 1 & 2 & 1 & 11 \\ 1 & 1 & 2 & 10 \end{array}\right] \xrightarrow[\substack{d_2 \to d_2 - d_1 \\ d_3 \to d_3 - 2d_1}]{d_1 \leftrightarrow d_3} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 10 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & -3 & -9 \end{array}\right]$$

$$\xrightarrow[d_3 \to d_3 + d_2]{d_1 \to d_1 - d_2} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -4 & -8 \end{array}\right]$$

$$\xrightarrow[\substack{d_2 \to d_2 + d_3 \\ d_1 \to d_1 - 3d_3}]{d_3 \to -\frac{1}{4} d_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Vậy $\mathbf{x}|_{\mathscr{B}'} = (3, 3, 2)$. Ta cũng có thể tìm tọa độ của $\mathbf{x}$ đối với cơ sở $\mathscr{B}'$ bằng cách tìm ma trận chuyển tọa độ từ $\mathscr{B}'$ sang $\mathscr{B}$. Ta biến đổi ma trận

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 & 1 & 1 \end{array}\right] \xrightarrow[\substack{d_2 \to d_2 - d_1 \\ d_3 \to d_3 - 2d_1}]{d_1 \leftrightarrow d_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & -1 & -3 & 1 & -2 & -1 \end{array}\right]$$

$$\xrightarrow[d_3 \to d_3 + d_2]{d_1 \to d_1 - d_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & -1 & 1 & 2 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & -4 & 2 & -2 & -2 \end{array}\right]$$

$$\xrightarrow[\substack{d_2 \to d_2 + d_3 \\ d_1 \to d_1 - 3d_3}]{d_3 \to -\frac{1}{4} d_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array}\right]$$

vậy ma trận chuyển cơ sở từ $\mathscr{B}'$ sang $\mathscr{B}$ là $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ nên ta tìm tọa độ của $\mathbf{x}$ đối với cơ sở $\mathscr{B}'$ như sau

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 6 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix} \qquad \text{suy ra} \qquad \mathbf{x}|_{\mathscr{B}'} = (3, 3, 2). \qquad \square$$

## Bài tập

**1** ) Hệ vector nào trong các hệ sau là cơ sở của $\mathbb{R}^3$.

**(a)** $\mathscr{B}_1 = ((1,2,3),(0,2,3))$

**(b)** $\mathscr{B}_2 = ((1,2,3),(0,2,3),(0,0,5))$

**(c)** $\mathscr{B}_3 = ((1,1,2),(1,2,5),(5,3,4))$

**(d)** $\mathscr{B}_4 = ((-1,0,0),(-1,1,0),(1,-1,1),(2,0,5))$

**2** ) Trong $\mathbb{R}^3$, chứng minh $\mathscr{B} = (\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3)$ là cơ sở và tìm tọa độ của $\mathbf{u}$ trong $\mathscr{B}$ với $\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3$ và $\mathbf{u}$ cho như dưới đây.

**(a)** $\mathbf{u}_1 = (1,1,1)$, $\mathbf{u}_2 = (1,1,2)$, $\mathbf{u}_3 = (1,2,3)$, $\mathbf{u} = (6,9,14)$

**(b)** $\mathbf{u}_1 = (2,1,-3)$, $\mathbf{u}_2 = (3,2,-5)$, $\mathbf{u}_3 = (1,-1,1)$, $\mathbf{u} = (6,2,-7)$

**(c)** $\mathbf{u}_1 = (1,-1,0)$, $\mathbf{u}_2 = (1,0,-1)$, $\mathbf{u}_3 = (2,0,0)$, $\mathbf{u} = (-3,1,-2)$

**3** ) Trong $\mathbb{R}^3$ cho hai hệ vector $\mathscr{A} = (\mathbf{a}_1 = (1,2,1), \mathbf{a}_2 = (-1,1,1), \mathbf{a}_3 = (1,3,2))$ và $\mathscr{B} = (\mathbf{b}_1 = (2,1,1), \mathbf{b}_2 = (-2,2,1), \mathbf{b}_3 = (2,3,2))$. Chứng minh rằng $\mathscr{A}$ và $\mathscr{B}$ là các cơ sở của $\mathbb{R}^3$. Tìm ma trận chuyển cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$.

**4** ) Trong $\mathbb{R}^3$ cho hai hệ vector $\mathscr{A} = (\mathbf{a}_1 = (1,0,1), \mathbf{a}_2 = (1,1,0), \mathbf{a}_3 = (0,1,1))$ và $\mathscr{B} = (\mathbf{b}_1 = (1,1,2), \mathbf{b}_2 = (-1,1,1), \mathbf{b}_3 = (1,2,3))$. Chứng minh rằng $\mathscr{A}$ và $\mathscr{B}$ là các cơ sở của $\mathbb{R}^3$. Tìm ma trận chuyển cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$.

**5** ) Trong $\mathbb{R}^3$ cho hai hệ vector $\mathscr{A} = (\mathbf{a}_1 = (1,1,1), \mathbf{a}_2 = (1,1,2), \mathbf{a}_3 = (1,2,3))$ và $\mathscr{B} = (\mathbf{b}_1 = (2,1,-1), \mathbf{b}_2 = (3,2,-5), \mathbf{b}_3 = (1,-1,m))$. Chứng minh rằng $\mathscr{A}$ là cơ sở của $\mathbb{R}^3$. Tìm $m$ để $\mathscr{B}$ là cơ sở của $\mathbb{R}^3$. Với $m$ tìm được, tìm ma trận chuyển cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$ và tìm tọa độ của vector $\mathbf{x} = (1,0,0)$ trong hai cơ sở đó.

**6** ) Trong $\mathbb{R}^3$ cho hai cơ sở $\mathscr{B} = ((1,1,0),(0,1,1),(1,0,1))$, $\mathscr{B}' = ((0,0,1),(1,-1,0),(1,1,1))$.

**(a)** Tìm ma trận đổi cơ sở từ $\mathscr{B}$ sang $\mathscr{B}'$ và từ $\mathscr{B}'$ sang $\mathscr{B}$.

**(b)** Tìm tọa độ của $\mathbf{x} = (1, -1, 1)$ trong hai cơ sở đó.

**7 )** Trong không gian $\mathbb{R}_n[x]$ cho hai hệ $\mathscr{B} = (1, x, x^2, \dots, x^n)$ và $\mathscr{B}' = (1, x-a, (x-a)^2, \dots, (x-a)^n)$ ở đó $a \in \mathbb{R}$.

**(a)** Chứng minh $\mathscr{B}$ và $\mathscr{B}'$ là các cơ sở của $\mathbb{R}_n[x]$. Tìm ma trận đổi cơ sở từ $\mathscr{B}$ sang $\mathscr{B}'$.

**(b)** Tìm tọa độ của $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ trong $\mathscr{B}$ và $\mathscr{B}'$.

# § 5 Không gian con

## 5.1 Khái niệm không gian con

**5.1.1 Định nghĩa.** Cho $V$ là một không gian vector trên trường $\mathbb{K}$. $U$ là tập con khác rỗng của $V$. Ta nói $U$ là **không gian vector con** (subspace) của $V$ nếu hai phép toán trên $V$ thu hẹp trên $U$ lại là các phép toán của $U$, đồng thời tập $U$ cùng với hai phép toán đó lập thành một không gian vector trên $\mathbb{K}$. Ta có thể dùng kí hiệu $U \Subset V$.

**5.1.2 *Chú ý.*** Từ định nghĩa ta có thể thấy với mọi không gian vector $V$ đều có hai không gian con $\{\mathbf{0}\}$ và $V$.

Trong định nghĩa trên có phần thừa vì nếu một tập khác rỗng $U$ mà phép cộng và phép nhân với vô hướng của $V$ thu hẹp trên $U$ cũng là phép cộng và phép nhân với vô hướng trên $U$ thì với hai phép toán đó $U$ là một không gian vector. Ta sẽ thấy điều này trong định lý sau.

**5.1.3 Định lý.** *Tập con khác rỗng $U$ của $V$ là không gian vector con của $V$ khi và chỉ khi*

**(i)** $\mathbf{x} + \mathbf{y} \in U$ *với mọi* $\mathbf{x}, \mathbf{y} \in U$

**(ii)** $\alpha\mathbf{x} \in U$ *với mọi* $\alpha \in \mathbb{K}$ *và mọi* $\mathbf{x} \in U$.

*Chứng minh.* Nếu $U$ là không gian con của $V$ thì rõ ràng ta có hai tính chất trên bởi vì phép cộng và nhân với vô hướng của $V$ bảo toàn trên $U$.

Ngược lại, giả sử trên tập $U$ ta có phép toán cộng và nhân vô hướng của $V$ thỏa hai tính chất trên. Nhờ hai tính chất đó mà trên $U$ có phép cộng hai vector và phép nhân vô hướng với vector, và chúng chính là các phép toán của $V$. Ta nhận thấy trừ tiên đề thứ ba và thứ tư các tiên đề còn lại hiển nhiên đúng trên $U$ bởi vì phép toán trên $U$ cũng chính là phép toán trên $V$. Phần tử không $\mathbf{0}$ của $V$ thuộc $U$ vì $U$ khác rỗng nên có $\mathbf{a} \in U$ và $\mathbf{0} = 0\mathbf{a} \in U$. Với $\mathbf{a} \in U$ và $-1 \in \mathbb{K}$ nên $-\mathbf{a} = (-1)\mathbf{a} \in U$, nghĩa là phần tử đối của $\mathbf{a}$ cũng thuộc $U$. □

**5.1.4 Định lý.** *Tập con khác rỗng $U$ của $V$ là không gian vector con của $V$ khi và chỉ khi $\alpha\mathbf{x} + \beta\mathbf{y} \in U$ với mọi $\mathbf{x}, \mathbf{y} \in U$ và mọi $\alpha, \beta \in \mathbb{K}$.*

**5.1.5 Thí dụ.** Trong $\mathbb{K}^3$ cho $U = \{(x_1, x_2, x_3) : x_1 + x_2 - x_3 = 0\}$. Rõ ràng $U \neq \varnothing$ bởi vì $(1, 0, 1) \in U$. Lấy $\mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3) \in U$. Ta có $x_1 + x_2 - x_3 = 0$ và $y_1 + y_2 - y_3 = 0$. Suy ra $(x_1 + y_1) + (x_2 + y_2) - (x_3 + y_3) = (x_1 + x_2 - x_3) + (y_1 + y_2 - y_3) = 0$, cho nên $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3) \in U$. Hơn nữa, với mọi $\alpha \in \mathbb{K}$ ta có $\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \alpha x_3)$ và $\alpha x_1 + \alpha x_2 - \alpha x_3 = \alpha(x_1 + x_2 - x_3) = 0$. Suy ra $\alpha\mathbf{x} \in U$. Vậy $U$ là không gian vector con của $\mathbb{K}^3$. □

**5.1.6 Thí dụ.** Xét không gian $\mathbb{K}[x]$ các đa thức trên $\mathbb{K}$. Đặt $\mathbb{K}_n[x]$ là tập các đa thức của $\mathbb{K}[x]$ có bậc không quá $n$. Ta có thể chứng minh được $\mathbb{K}_n[x]$ cùng với phép toán cộng hai đa thức và nhân đa thức với một số là không gian vector con của $\mathbb{K}[x]$ và gọi nó là *không gian các đa thức bậc không quá $n$.* □

**5.1.7 Định lý.** *Mọi không gian con của một không gian vector hữu hạn chiều $V$ đều là không gian vector hữu hạn chiều với số chiều không quá số chiều của không gian $V$.*

*Chứng minh.* Giả sử $\dim V = n$ và $U$ là một không gian vector con của $V$. Nếu $U = \{\mathbf{0}\}$ thì $\dim U = 0 \leq n$. Giả sử $U \neq \{\mathbf{0}\}$. Khi đó, mọi hệ độc lập tuyến tính trong $U$ đều là hệ độc lập tuyến tính trong $V$, cho nên số vector trong hệ ấy nhỏ hơn hoặc bằng $n$. Theo bổ đề Zorn trong $U$ tồn tại hệ độc lập tuyến tính tối đại, $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ với $m \leq n$. Khi đó, với mỗi $\mathbf{x} \in U$ tùy ý ta phải có hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m, \mathbf{x})$ phụ thuộc tuyến tính. Theo Định lý 2.2.6 ta có vector $\mathbf{x}$ biểu thị tuyến tính được qua hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$. Vậy hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ là hệ sinh của $U$; do đó, nó là cơ sở của $U$. Vậy $U$ là không gian hữu hạn chiều và $\dim U = m \leq n$. □

**5.1.8 Nhận xét.** Nếu $U \Subset V$ và $V$ hữu hạn chiều thì

**(i)** $\dim U = 0$ khi và chỉ khi $U = \{\mathbf{0}\}$, không gian con tầm thường.

**(ii)** $\dim U = \dim V$ khi và chỉ khi $U = V$, không gian con không thực sự.

**(iii)** $0 < \dim U < \dim V$ khi và chỉ khi $U$ khác $V$ và khác $\{\mathbf{0}\}$, không gian con thực sự không tầm thường.

## 5.2 Giao và tổng của các không gian con

**5.2.1 Định lý.** *Cho $V$ là một không gian vector và $\{U_i\}_{i \in I}$ là một họ các không gian con của $V$. Khi đó, $U = \bigcap\limits_{i \in I} U_i$ là không gian con của $V$, và nó được gọi là **không gian giao** của họ $\{U_i\}_{i \in I}$.*

*Chứng minh.* Ta có thể thấy $U \neq \varnothing$ vì $\mathbf{0} \in U$. Với mọi $\mathbf{x}, \mathbf{y} \in U$, ta có $\mathbf{x}, \mathbf{y} \in U_i$ với mọi $i \in I$. Do $U_i$ là không gian con của $V$ nên $\mathbf{x} + \mathbf{y} \in U_i$ và $\lambda \mathbf{x} \in U_i$ với mọi $i \in I$ với mọi $\lambda \in \mathbb{K}$. Suy ra $\mathbf{x} + \mathbf{y} \in U$ và $\lambda \mathbf{x} \in U$ với mọi $\lambda \in \mathbb{K}$. Vậy $U$ là không gian con của $V$. □

**5.2.2 Thí dụ.** Trong $\mathbb{R}^3$ xét hai tập con $U_1 = \{(x, y, 0) : x, y \in \mathbb{R}\}$ và $U_2 = \{(x, 0, z) : x, z \in \mathbb{R}\}$. Ta có thể chứng minh được $U_1$ và $U_2$ là hai không gian con của $\mathbb{R}^3$. Vậy $U_1 \cap U_2$ là không gian con của $\mathbb{R}^3$, và $U_1 \cap U_2 = \{(x, 0, 0) : x \in \mathbb{R}\}$. □

**5.2.3 Mệnh đề.** *Nếu $A$ và $B$ là các không gian vector con của không gian vector $V$, thì tập*

$$A + B = \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}$$

*là một không gian vector con của $V$, và được gọi là* **không gian tổng** *của $A$ và $B$. Hơn nữa, $A+B$ là không gian vector con nhỏ nhất chứa $A\cup B$.*

*Chứng minh.* Theo định nghĩa tập $A+B$ ta nhận thấy $A+B\neq\varnothing$. Với $\mathbf{x},\mathbf{y}\in A+B$ tùy ý, tồn tại $\mathbf{a}_1,\mathbf{a}_2\in A$ và $\mathbf{b}_1,\mathbf{b}_2\in B$ sao cho $\mathbf{x}=\mathbf{a}_1+\mathbf{b}_1$ và $\mathbf{y}=\mathbf{a}_2+\mathbf{b}_2$. Khi đó, ta có $\mathbf{a}_1+\mathbf{a}_2\in A$ và $\mathbf{b}_1+\mathbf{b}_2\in B$, và với mọi $\lambda\in\mathbb{K}$ ta có $\lambda\mathbf{a}_1\in A$ và $\lambda\mathbf{b}_1\in B$. Ta có

$$\mathbf{x}+\mathbf{y}=(\mathbf{a}_1+\mathbf{b}_1)+(\mathbf{a}_2+\mathbf{b}_2)=(\mathbf{a}_1+\mathbf{a}_2)+(\mathbf{b}_1+\mathbf{b}_2)\in A+B.$$

và

$$\lambda\mathbf{x}=\lambda(\mathbf{a}_1+\mathbf{b}_1)=\lambda\mathbf{a}_1+\lambda\mathbf{b}_1\in A+B.$$

Vậy $A+B$ là không gian vector con của $V$.

Với $\mathbf{x}\in A\cup B$ tùy ý ta có $\mathbf{x}\in A$ hoặc $\mathbf{x}\in B$, ta cũng có $\mathbf{0}\in A$ và $\mathbf{0}\in B$. Do đó $\mathbf{x}=\mathbf{x}+\mathbf{0}=\mathbf{0}+\mathbf{x}\in A+B$. Vậy $A\cup B\subseteq A+B$. Giả sử $U$ là một không gian vector con của $V$ chứa $A\cup B$. Lấy $\mathbf{x}\in A+B$ tùy ý. Khi đó tồn tại $\mathbf{a}\in A$ và $\mathbf{b}\in B$ sao cho $\mathbf{x}=\mathbf{a}+\mathbf{b}$. Do $A\cup B\subseteq U$ nên $\mathbf{a}\in U$ và $\mathbf{b}\in U$, suy ra $\mathbf{a}+\mathbf{b}\in U$ hay $\mathbf{x}\in U$. Vậy $A+B\subseteq U$. $\square$

**5.2.4** Khái niệm không gian tổng của hai không gian con có thể mở rộng cho một họ khác rỗng tùy ý các không gian con $(W_i)_{i\in I}$ của $V$. Tổng $\sum\limits_{i\in I}W_i$ được định nghĩa như sau.

$$\sum_{i\in I}W_i=\left\{\mathbf{x}:\exists J\subset I, J \text{ khác rỗng và hữu hạn}, \mathbf{x}_j\in W_j, \mathbf{x}=\sum_{j\in J}\mathbf{x}_j\right\}$$

Thực sự $\sum\limits_{i\in I}W_i$ là không gian con nhỏ nhất chứa $\bigcup\limits_{i\in I}W_i$.

**5.2.5 Định lý.** *Cho $A$ và $B$ là không gian vector con hữu hạn chiều của không gian vector $V$. Khi đó, không gian vector $A+B$ cũng hữu hạn chiều và*

$$\dim(A\cap B)=\dim A+\dim B-\dim(A+B).$$

*Chứng minh.* Ta có $A\cap B$ là không gian giao của $A$ và $B$, nó là không gian hữu hạn chiều. Chọn $(\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_r)$ là một cơ sở của $A\cap B$, suy ra

$\dim(A \cap B) = r$ (có thể hệ cơ sở đó bằng $\varnothing$ khi $A \cap B = \{\mathbf{0}\}$ và $r = 0$). Bổ sung thêm một số vector để được một cơ sở $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s)$ của $A$, suy ra $\dim A = r + s$; và bổ sung thêm một số vector để được một cơ sở $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r, \mathbf{c}_1, \dots, \mathbf{c}_t)$ của $B$, suy ra $\dim B = r + t$. Ta sẽ chứng tỏ $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s, \mathbf{c}_1, \dots, \mathbf{c}_t)$ là cơ sở của $A + B$.

Xét tổ hợp tuyến tính của hệ vector trên

$$\sum_{i=1}^{r} \alpha_i \mathbf{a}_i + \sum_{j=1}^{s} \beta_j \mathbf{b}_j + \sum_{k=1}^{t} \gamma_k \mathbf{c}_k = \mathbf{0}.$$

Từ đó ta đặt $\mathbf{x} = -\sum_{k=1}^{t} \gamma_k \mathbf{c}_k = \sum_{i=1}^{r} \alpha_i \mathbf{a}_i + \sum_{j=1}^{s} \beta_j \mathbf{b}_j$. Từ các cơ sở của không gian vector $A$ và $B$ ở trên ta suy ra được $\mathbf{x} \in A$ và $\mathbf{x} \in B$. Do đó $\mathbf{x} \in A \cap B$, cho nên ta có thể viết $\mathbf{x} = \sum_{i=1}^{r} \lambda_i \mathbf{a}_i$. Vậy

$$\sum_{i=1}^{r} \lambda_i \mathbf{a}_i = \sum_{i=1}^{r} \alpha_i \mathbf{a}_i + \sum_{j=1}^{s} \beta_j \mathbf{b}_j \ \text{ suy ra } \ \sum_{i=1}^{r} (\alpha_i - \lambda_i) \mathbf{a}_i + \sum_{j=1}^{s} \beta_j \mathbf{b}_j = \mathbf{0}.$$

Vì $(\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{b}_1, \dots, \mathbf{b}_s)$ là cơ sở của $A$ nên ta phải có $\alpha_i = \lambda_i$ với mọi $i = 1, \dots, r$ và $\beta_j = 0$ với mọi $j = 1, \dots, s$. Thay vào tổ hợp tuyến tính ban đầu ta có

$$\sum_{i=1}^{r} \alpha_i \mathbf{a}_i + \sum_{k=1}^{t} \gamma_k \mathbf{c}_k = \mathbf{0}.$$

Vì $(\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{c}_1, \dots, \mathbf{c}_t)$ là cơ sở của $B$ nên ta phải có $\alpha_i = 0$ với mọi $i = 1, \dots, r$ và $\gamma_k = 0$ với mọi $k = 1, \dots, t$. Do đó ta được hệ vector $(\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{b}_1, \dots, \mathbf{b}_s, \mathbf{c}_1, \dots, \mathbf{c}_t)$ độc lập tuyến tính.

Với mọi $\mathbf{x} \in A + B$, tồn tại $\mathbf{a} \in A$ và $\mathbf{b} \in B$ sao cho $\mathbf{x} = \mathbf{a} + \mathbf{b}$. Vì $(\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{b}_1, \dots, \mathbf{b}_s)$ là cơ sở của $A$ và $(\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{c}_1, \dots, \mathbf{c}_t)$ là cơ sở của $B$ nên ta có thể viết

$$\mathbf{a} = \sum_{i=1}^{r} \alpha_i \mathbf{a}_i + \sum_{j=1}^{s} \beta_j \mathbf{b}_j \qquad \text{và} \qquad \mathbf{b} = \sum_{i=1}^{r} \alpha'_i \mathbf{a}_i + \sum_{k=1}^{t} \gamma_k \mathbf{c}_k.$$

Khi đó

$$\mathbf{x} = \mathbf{a} + \mathbf{b} = \sum_{i=1}^{r} (\alpha_i + \alpha'_i) \mathbf{a}_i + \sum_{j=1}^{s} \beta_j \mathbf{b}_j + \sum_{k=1}^{t} \gamma_k \mathbf{c}_k.$$

Vậy $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_s, \mathbf{c}_1, \dots, \mathbf{c}_t)$ là một hệ sinh của $A+B$. Do đó nó là cơ sở của $A+B$ và $\dim(A+B) = r+s+t$. Vậy

$$\begin{aligned}\dim A + \dim B - \dim(A+B) &= r+s+r+t-(r+s+t)\\ &= r\\ &= \dim A \cap B. \qquad \square\end{aligned}$$

## Bài tập

**1**) Chứng minh rằng: Tập con khác rỗng $U$ của $V$ là không gian vector con của $V$ khi và chỉ khi $\alpha\mathbf{x} + \beta\mathbf{y} \in U$ với mọi $\mathbf{x}, \mathbf{y} \in U$ và mọi $\alpha, \beta \in \mathbb{K}$.

**2**) Xét xem các tập sau có là không gian vector hay không gian con hay không.

**(a)** Tập các đa thức hệ số thực bậc $n$.

**(b)** $D_n(\mathbb{R})$ các ma trận vuông cấp $n$ đối xứng hệ số thực.

**(c)** Tập $GL_n(\mathbb{R})$ các ma trận vuông thực cấp $n$ không suy biến.

# § 6 Không gian sinh bởi hệ vector và không gian nghiệm

## 6.1 Không gian con sinh bởi một hệ vector

**6.1.1 Định lý.** *Cho $V$ là một không gian vector, và hệ $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ trong $V$. Khi đó, tập hợp tất cả các tổ hợp tuyến tính của hệ $\mathscr{A}$ là một không gian vector con của $V$ và được gọi là **không gian con sinh bởi hệ vector** $\mathscr{A}$, hay **bao tuyến tính** của $\mathscr{A}$, kí hiệu $\langle\mathscr{A}\rangle$.*

$$\langle\mathscr{A}\rangle = \left\{ \sum_{i=1}^{m} \lambda_i \mathbf{a}_i : \lambda_i \in \mathbb{K}, i = 1, \dots, m \right\}.$$

*Hơn nữa, không gian $\langle\mathscr{A}\rangle$ hữu hạn chiều và $\dim\langle\mathscr{A}\rangle = \operatorname{rank}\mathscr{A}$.*

*Chứng minh.* Ta thấy $\langle\mathscr{A}\rangle \neq \varnothing$ vì $\mathbf{0} = \sum_{i=1}^{m} 0\mathbf{a}_i \in \langle\mathscr{A}\rangle$. Giả sử $\mathbf{x}, \mathbf{y} \in \langle\mathscr{A}\rangle$. Khi đó tồn tại các số $\lambda_1, \lambda_2, \ldots, \lambda_m$ và $k_1, k_2, \ldots, k_m$ sao cho $\mathbf{x} = \sum_{i=1}^{m} \lambda_i\mathbf{a}_i$ và $\mathbf{y} = \sum_{i=1}^{m} k_i\mathbf{a}_i$. Suy ra

$$\mathbf{x} + \mathbf{y} = \sum_{i=1}^{m} \lambda_i\mathbf{a}_i + \sum_{i=1}^{m} k_i\mathbf{a}_i = \sum_{i=1}^{m}(\lambda_i + k_i)\mathbf{a}_i \in \langle\mathscr{A}\rangle.$$

Hơn nữa, với $\alpha \in \mathbb{K}$ tùy ý, ta có

$$\alpha\mathbf{x} = \alpha\sum_{i=1}^{m} \lambda_i\mathbf{a}_i = \sum_{i=1}^{m}(\alpha\lambda_i)\mathbf{a}_i \in \langle\mathscr{A}\rangle.$$

Vậy $\langle\mathscr{A}\rangle$ là không gian con của $V$.

Nếu các vector thuộc $\mathscr{A}$ đều là vector không, thì $\langle\mathscr{A}\rangle = \{\mathbf{0}\}$; suy ra $\dim\langle\mathscr{A}\rangle = 0 = \operatorname{rank}\mathscr{A}$. Giả sử trong $\mathscr{A}$ có vector khác không. Gọi $(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r})$ là hệ con độc lập tuyến tính tối đại của $\mathscr{A}$. Khi đó $\operatorname{rank}\mathscr{A} = r$. Ta thấy hệ $\mathscr{A}$ và hệ $(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r})$ tương đương. Do đó, mọi vector trong $\langle\mathscr{A}\rangle$ đều biểu thị tuyến tính được qua hệ $\mathscr{A}$ nên cũng biểu thị tuyến tính được qua hệ $(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r})$. Suy ra hệ $(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r})$ là hệ sinh của $\langle\mathscr{A}\rangle$. Vậy hệ $(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r})$ là cơ sở của $\langle\mathscr{A}\rangle$ (vì nó độc lập tuyến tính). Do đó $\dim\langle\mathscr{A}\rangle = r = \operatorname{rank}\mathscr{A}$. □

**6.1.2 Nhận xét.** Qua chứng minh định lý trên chúng ta nhận thấy nếu $\dim\langle\mathscr{A}\rangle > 0$ thì một hệ con độc lập tuyến tính tối đại của $\mathscr{A}$ là một cơ sở của $\langle\mathscr{A}\rangle$. Hơn nữa, trong trường hợp $\mathscr{A}$ là hệ vector trong $\mathbb{K}^n$, ta lập ma trận $A$ bởi các vector trong $\mathscr{A}$ dưới dạng ma trận cột. Khi đó, các hàng khác không trong biến đổi sơ cấp theo dòng của ma trận $A^t$ sẽ cho ta một cơ sở khác của $\langle\mathscr{A}\rangle$.

**6.1.3 Thí dụ.** Trong $\mathbb{R}^3$ tìm bao tuyến tính của hệ $(\mathbf{a}_1 = (1,1,1), \mathbf{a}_2 = (2,3,4), \mathbf{a}_3 = (4,5,6))$.

Lập ma trận dòng của ba vector $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ và biến đổi nó

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 4 & 5 & 6 \end{bmatrix} \xrightarrow[d_2 \to d_2 - 2d_1]{d_3 \to d_3 - 3d_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow{d_3 \to d_3 - d_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Vậy $\dim(\langle\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\rangle) = 2$ và $((1,1,1), (0,1,2))$ là một cơ sở. □

**6.1.4 Thí dụ.** Tìm cơ sở và số chiều của không gian con sinh bởi hệ vector $((1,2,0,1,3),(2,3,1,-1,4),(2,5,-1,3,0),(3,5,1,-2,7),(4,8,0,2,4))$ trong $\mathbb{R}^5$.

Ta biến đổi sơ cấp ma trận sau

$$\begin{bmatrix}1&2&0&1&3\\2&3&1&-1&4\\2&5&-1&3&0\\3&5&1&-2&7\\4&8&0&2&4\end{bmatrix}\xrightarrow[\substack{d_4-3d_1\\d_5-4d_1}]{\substack{d_2-2d_1\\d_3-2d_1}}\begin{bmatrix}1&2&0&1&3\\0&-1&1&-3&-2\\0&1&-1&1&-6\\0&-1&1&-5&-2\\0&0&0&-2&-8\end{bmatrix}$$

$$\xrightarrow[d_4-d_2]{d_3+d_2}\begin{bmatrix}1&2&0&1&3\\0&-1&1&-3&-2\\0&0&0&-2&-8\\0&0&0&-2&0\\0&0&0&-2&-8\end{bmatrix}$$

$$\xrightarrow[d_5-d_3]{d_4-d_3}\begin{bmatrix}1&2&0&1&3\\0&-1&1&-3&-2\\0&0&0&-2&-8\\0&0&0&0&8\\0&0&0&0&0\end{bmatrix}$$

Vậy số chiều của không gian con sinh bởi hệ vector đã cho là $4$ và cơ sở của nó là $((1,2,0,1,3),(2,3,1,-1,4),(2,5,-1,3,0),(3,5,1,-2,7))$ hay một cơ sở khác là $((1,2,0,1,3),(0,-1,1,-3,-2),(0,0,0,-2,-8),(0,0,0,0,8))$. $\square$

**6.1.5 Định lý.** *Cho $A$ và $B$ là hai ma trận cùng cấp $m \times n$. Khi đó, ta có*

(6.1.6) $$\operatorname{rank}(A+B) \leq \operatorname{rank} A + \operatorname{rank} B.$$

*Chứng minh.* Gọi $\mathbf{a}_i \in \mathbb{R}^m$ là vector ứng với cột thứ $i$ của ma trận $A$ và $\mathbf{b}_j \in \mathbb{R}^m$ là vector ứng với cột thứ $j$ của ma trận $B$. Khi đó ta có $\operatorname{rank} A = \dim\langle \mathbf{a}_1,\ldots,\mathbf{a}_n\rangle$ và $\operatorname{rank} B = \dim\langle \mathbf{b}_1,\ldots,\mathbf{b}_n\rangle$. Mặt khác, rõ ràng $\mathbf{a}_i + \mathbf{b}_i \in \langle \mathbf{a}_1,\ldots,\mathbf{a}_n\rangle + \langle \mathbf{b}_1,\ldots,\mathbf{b}_n\rangle$. Do đó, theo Định lý 5.2.5

ta có

$$\begin{aligned}\operatorname{rank}(A+B) &= \operatorname{rank}(\mathbf{a}_1+\mathbf{b}_1,\dots,\mathbf{a}_n+\mathbf{b}_n)\\ &\le \dim(\langle \mathbf{a}_1,\dots,\mathbf{a}_n\rangle + \langle \mathbf{b}_1,\dots,\mathbf{b}_n\rangle)\\ &\le \dim\langle \mathbf{a}_1,\dots,\mathbf{a}_n\rangle + \dim\langle \mathbf{b}_1,\dots,\mathbf{b}_n\rangle\\ &= \operatorname{rank} A + \operatorname{rank} B.\end{aligned}$$

Vậy ta có được tính chất cho hạng của ma trận $\operatorname{rank}(A+B) \le \operatorname{rank} A + \operatorname{rank} B$. □

**6.1.7 Thí dụ. (OSV07)** Cho ma trận $A = (a_{ij})_{n\times n}$ trong đó phần tử $a_{ij} = i + j$ với $i, j = 1, \dots, n$. Tính $\operatorname{rank}(A)$.

*Giải.* Nếu $n = 1$ ta có ngay $\operatorname{rank}(A) = 1$. Nếu $n \ge 2$ thì ma trận $A$ được viết lại

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 & 1\\ 2 & 2 & \dots & 2 & 2\\ \dots & \dots & \dots & \dots & \dots\\ n-1 & n-1 & \dots & n-1 & n-1\\ n & n & \dots & n & n\end{bmatrix} + \begin{bmatrix} 1 & 2 & \dots & n-1 & n\\ 1 & 2 & \dots & n-1 & n\\ \dots & \dots & \dots & \dots & \dots\\ 1 & 2 & \dots & n-1 & n\\ 1 & 2 & \dots & n-1 & n\end{bmatrix}$$

$$= B_1 + B_2$$

Do đó, theo định lý trên ta có $\operatorname{rank}(A) \le \operatorname{rank}(B_1) + \operatorname{rank}(B_2) = 1 + 1 = 2$. Mặt khác, ma trận con cấp hai của $A$ ở phía trên và bên trái là $C = \begin{bmatrix} 2 & 3\\ 3 & 4\end{bmatrix}$ có định thức khác không, cho nên $\operatorname{rank}(A) \ge 2$. Vậy $\operatorname{rank}(A) = 2$. □

**6.1.8 Mệnh đề.** *Nếu* $A = \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\rangle$ *và* $B = \langle \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\rangle$, *thì* $A + B = \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\rangle$.

*Chứng minh.* Lấy $\mathbf{x} \in \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\rangle$ tùy ý. Khi đó tồn tại các số $\lambda_1, \lambda_2, \dots, \lambda_m, k_1, k_2, \dots, k_n$ sao cho $\mathbf{x} = \sum_{i=1}^m \lambda_i\mathbf{a}_i + \sum_{j=1}^n k_j\mathbf{b}_j$. Vì $\sum_{i=1}^m \lambda_i\mathbf{a}_i \in A$ và $\sum_{j=1}^n k_j\mathbf{b}_j \in B$ nên $\mathbf{x} \in A + B$.

Mặt khác, lấy $\mathbf{x} \in A+B$ tùy ý. Khi đó tồn tại $\mathbf{a} \in A$ và $\mathbf{b} \in B$ sao cho $\mathbf{x} = \mathbf{a} + \mathbf{b}$. Do $\mathbf{a} \in A$ nên tồn tại $\alpha_1, \alpha_2, \dots, \alpha_m$ sao cho $\mathbf{a} = \sum_{i=1}^m \alpha_i\mathbf{a}_i$;

và do $\mathbf{b} \in B$ nên tồn tại $\beta_1, \beta_2, \ldots, \beta_n$ sao cho $\mathbf{b} = \sum_{j=1}^{n} \beta_j \mathbf{b}_j$. Vậy $\mathbf{x} = \sum_{i=1}^{m} \alpha_i \mathbf{a}_i + \sum_{j=1}^{n} \beta_j \mathbf{b}_j$, suy ra $\mathbf{x} \in \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m, \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n \rangle$. Vậy phải có $A + B = \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m, \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n \rangle$. □

**6.1.9 Thí dụ.** Với $A$ và $B$ là không gian con của không gian vector $V$, ta đã biết $A+B$ là một không gian con của $V$ và mỗi phần tử của nó là tổng các phần tử của $A$ và $B$, tuy nhiên biểu diễn này nói chung là không duy nhất. Ta thấy điều này qua ví dụ cụ thể sau. Xét $\mathbb{R}^3$ với $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, $\mathbf{e}_3 = (0, 0, 1)$. Giả sử $A = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle$ và $B = \langle \mathbf{e}_2, \mathbf{e}_3 \rangle$. Khi đó $\mathbf{a} \in A$ có dạng $\mathbf{a} = (a_1, a_2, 0)$ và $\mathbf{b} \in B$ có dạng $\mathbf{b} = (0, b_2, b_3)$. Theo mệnh đề trên ta có $A + B = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_3 \rangle = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle = \mathbb{R}^3$. Vậy với mỗi $(x, y, z) \in A + B$ ta có nhiều cách biểu diễn

$$(x, y, z) = (x, y, 0) + (0, 0, z) = (x, 0, 0) + (0, y, z) = (x, \tfrac{y}{2}, 0) + (0, \tfrac{y}{2}, z).$$

□

**6.1.10 (Ứng dụng)** Xét hệ phương trình tuyến tính tổng quát

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{6.1.11}$$

Đặt $\mathbf{a}_i = (a_{1i}, a_{2i}, \ldots, a_{mi}) \in \mathbb{K}^m$ và $\mathbf{b} = (b_1, b_2, \ldots, b_m)$. Khi đó, hệ trên được viết lại

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

Như ta đã biết hệ phương trình (6.1.11) có nghiệm khi và chỉ khi $\mathbf{b}$ biểu thị tuyến tính được qua hệ vector $(\mathbf{a}_1, \mathbf{a}_2 \ldots, \mathbf{a}_n)$. Điều này cũng có thể được phát biểu lại trong định lý sau.

**6.1.12 Định lý.** *Hệ phương trình tuyến tính* (6.1.11) *có nghiệm khi và chỉ khi* $\mathbf{b} \in \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \rangle$.

## 6.2 Không gian con các nghiệm của hệ phương trình tuyến tính thuần nhất

**6.2.1 Định lý.** *Cho hệ phương trình tuyến tính thuần nhất trên $\mathbb{R}$ và dạng ma trận của nó*

$$(6.2.2)\qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \qquad \textit{hay} \qquad A\mathbf{x} = \mathbf{0}.$$

*Khi đó, tập hợp nghiệm $S$ của hệ là một không gian vector con của $\mathbb{R}^n$, ta gọi đó là **không gian nghiệm của hệ phương trình tuyến tính thuần nhất** (6.2.2). Hơn nữa, nếu $A$ là ma trận hệ số có hạng bằng $r$, thì* $\dim S = n - r$.

*Chứng minh.* Ta nhận thấy $\mathbf{0} \in S$ nên $S \neq \varnothing$. Lấy $\mathbf{x}, \mathbf{y} \in S$, nghĩa là $A\mathbf{x} = \mathbf{0}$ và $A\mathbf{y} = \mathbf{0}$. Ta có

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0}.$$

Vậy $\mathbf{x} + \mathbf{y}$ là nghiệm của hệ hay $\mathbf{x} + \mathbf{y} \in S$. Với mọi $\lambda \in \mathbb{R}$ tùy ý, ta có $A(\lambda\mathbf{x}) = \lambda(A\mathbf{x}) = \mathbf{0}$. Vậy $\lambda\mathbf{x} \in S$. Do đó $S$ là không gian vector con.

Việc chứng minh $\dim S = n - r$ được chuyển đến chương sau (xem trang 290). □

**6.2.3 *Chú ý.*** Không gian nghiệm của hệ phương trình (6.2.2) còn được gọi là **không gian không** của ma trận $A$.

**6.2.4 Thí dụ. (OSV03)** Cho $A$ là ma trận vuông thỏa mãn điều kiện $A^{2003} = O$. Chứng minh rằng với mọi số nguyên dương $n$ ta luôn có $\operatorname{rank}(A) = \operatorname{rank}(A + A^2 + \cdots + A^n)$.

*Giải.* Khẳng định là hiển nhiên với $n = 1$. Với $n \geq 2$, ta xét hai hệ phương trình thuần nhất $A\mathbf{x} = \mathbf{0}$ và $(A+A^2+\cdots+A^n)\mathbf{x} = \mathbf{0}$. Rõ ràng mọi nghiệm của $A\mathbf{x} = \mathbf{0}$ đều là nghiệm của $(A + A^2 + \cdots + A^n)\mathbf{x} = \mathbf{0}$. Giả sử $\mathbf{x}_0$ là nghiệm của hệ thứ hai. Khi đó ta có $A\mathbf{x}_0 = -(A^2 + \cdots + A^n)\mathbf{x}_0 = BA^2\mathbf{x}_0$ với $B = -I + \cdots - A^{n-2}$. Dễ dàng thấy rằng $AB = BA$, cho nên ta có

$$\begin{aligned} A\mathbf{x}_0 = BA^2\mathbf{x}_0 = BABA^2\mathbf{x}_0 = B^2A^3\mathbf{x}_0 = B^2A^2BA^2\mathbf{x}_0 = B^3A^4\mathbf{x}_0 \\ = \cdots = B^{2002}A^{2003}\mathbf{x}_0 = \mathbf{0}. \end{aligned}$$

Vậy $\mathbf{x}_0$ cũng là nghiệm của hệ $A\mathbf{x} = \mathbf{0}$. Như vậy hai hệ phương trình tuyến tính thuần nhất đang xét có cùng tập hợp nghiệm. Do đó theo Định lý 6.2.1 ta suy ra được $\operatorname{rank}(A) = \operatorname{rank}(A + A^2 + \cdots + A^n)$. □

**6.2.5 Định nghĩa.** Mỗi cơ sở của không gian nghiệm của hệ (6.2.2) được gọi là một **hệ nghiệm cơ bản**.

**6.2.6** Khi $r = n$ thì hệ (6.2.2) có nghiệm duy nhất nên không có hệ nghiệm cơ bản. Với $r < n$, hệ (6.2.2) có hệ nghiệm cơ bản. Để tìm hệ nghiệm cơ bản của hệ (6.2.2) ta làm như sau.

Do $r < n$ và không mất tính tổng quát ta có thể giả sử rằng định thức con cấp cao nhất khác 0 của ma trận $A$ là:

$$D = \begin{vmatrix} a_{11} & \ldots & a_{1r} \\ \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rr} \end{vmatrix} \neq 0.$$

Khi đó hệ (6.2.2) tương đương với

$$\begin{cases} a_{11}x_1 + \cdots + a_{1r}x_r = -a_{1,r+1}x_{r+1} - \cdots - a_{1n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + \cdots + a_{rr}x_r = -a_{r,r+1}x_{r+1} - \cdots - a_{rn}x_n. \end{cases}$$

Mỗi nghiệm của hệ phụ thuộc vào $n - r$ ẩn tự do: $x_{r+1}, x_{r+2}, \ldots, x_n$. Cho $x_{r+1} = 1$, $x_{r+2} = \cdots = x_n = 0$, ta được một nghiệm $\mathbf{c}_1 = (c_{11}, \ldots, c_{1r}, 1, 0, \ldots, 0)$. Tiếp tục cho $x_{r+1} = 0$, $x_{r+2} = 1$, $x_{r+3} = \cdots = x_n = 0$, ta được một nghiệm $\mathbf{c}_2 = (c_{21}, \ldots, c_{2r}, 0, 1, \ldots, 0)$. Lần lượt như thế ta được $n - r$ nghiệm riêng: $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{n-r}$. Đó là $n - r$ vector thuộc $S$. Dễ dàng kiểm chứng hạng của hệ $(\mathbf{c}_1, \mathbf{c}_2, \ldots \mathbf{c}_{n-r})$ bằng

$n-r$, để ý ma trận

$$C=\begin{bmatrix} c_{11} & \dots & c_{1r} & 1 & 0 & \dots & 0\\ c_{21} & \dots & c_{2r} & 0 & 1 & \dots & 0\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots\\ c_{n-r,1} & \dots & c_{n-r,r} & 0 & 0 & \dots & 1\end{bmatrix}.$$

Vậy hệ $(\mathbf{c}_1,\mathbf{c}_2,\dots,\mathbf{c}_{n-r})$ độc lập tuyến tính. Do $\dim S=n-r$ nên hệ $(\mathbf{c}_1,\mathbf{c}_2,\dots\mathbf{c}_{n-r})$ là cơ sở của $S$. Vậy hệ $(\mathbf{c}_1,\mathbf{c}_2,\dots\mathbf{c}_{n-r})$ là một hệ nghiệm cơ bản.

**6.2.7 Thí dụ.** Tìm cơ sở, số chiều của không gian nghiệm của hệ

$$\begin{cases} x_1-2x_2+x_3+x_4+x_5=0\\ 2x_1-4x_2-x_3+2x_4+2x_5=0\\ -x_1+2x_2+2x_3-x_4+3x_5=0.\end{cases}$$

Ta có $A=\begin{bmatrix} 1 & -2 & 1 & 1 & 1\\ 2 & -4 & -1 & 2 & 2\\ -1 & 2 & 2 & -1 & 3\end{bmatrix}$. Ta thực hiện các phép biến đổi sơ cấp trên dòng của ma trận $A$.

$$A\xrightarrow[d_3\to d_3+d_1]{d_2\to d_2-2d_1}\begin{bmatrix} 1 & -2 & 1 & 1 & 1\\ 0 & 0 & -3 & 0 & 0\\ 0 & 0 & 3 & 0 & 4\end{bmatrix}\xrightarrow[\substack{d_3\to d_3+d_2\\ d_2\to -\frac{1}{3}d_2}]{d_1\to d_1+\frac{1}{3}d_2}\begin{bmatrix} 1 & -2 & 1 & 1 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 4\end{bmatrix}$$

Vậy ta được $\operatorname{rank} A=3$. Do đó hệ có vô số nghiệm phụ thuộc vào 2 tham số $x_2,x_4$

$$x_1=2x_2-x_4 \qquad x_3=0 \qquad x_5=0.$$

hay

$$\mathbf{x}=\begin{bmatrix} 2x_2-x_4\\ x_2\\ 0\\ x_4\\ 0\end{bmatrix}=x_2\begin{bmatrix} 2\\ 1\\ 0\\ 0\\ 0\end{bmatrix}+x_4\begin{bmatrix} -1\\ 0\\ 0\\ 1\\ 0\end{bmatrix}$$

$\dim S=2$ và cơ sở của nó là $((2,1,0,0,0),(-1,0,0,1,0))$. □

**6.2.8 Thí dụ. (OSV06)** Xác định $n$ để hệ phương trình sau có 3 nghiệm độc lập tuyến tính

$$\begin{cases} 1^2x_1 + 2^2x_2 + \cdots + n^2x_n = 0 \\ 2^2x_1 + 3^2x_2 + \cdots + (n+1)^2x_n = 0 \\ 3^2x_1 + 4^2x_2 + \cdots + (n+2)^2x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ n^2x_1 + (n+1)^2x_2 + \cdots + (2n-1)^2x_n = 0 \end{cases}$$

*Giải* Gọi $A$ là ma trận hệ số của phương trình

$$A = \begin{bmatrix} 1^2 & 2^2 & \dots & n^2 \\ 2^2 & 3^2 & \dots & (n+1)^2 \\ 3^2 & 4^2 & \dots & (n+2)^2 \\ \dots & \dots & \dots & \dots \\ n^2 & (n+1)^2 & \dots & (2n-1)^2 \end{bmatrix}$$

Với $k = 1, \dots, k-1$ lấy dòng $k+1$ trừ dòng $k$ ta được

$$A_1 = \begin{bmatrix} 1^2 & 2^2 & \dots & n^2 \\ 3 & 5 & \dots & 2n+1 \\ 5 & 7 & \dots & 2n+3 \\ \dots & \dots & \dots & \dots \\ 2n-1 & 2n+1 & \dots & 4n-3 \end{bmatrix}$$

Tiếp tục lấy dòng $k+1$ trừ dòng $k$ với $k = 2, \dots, n-1$ ta được

$$A_2 = \begin{bmatrix} 1^2 & 2^2 & \dots & n^2 \\ 3 & 5 & \dots & 2n+1 \\ 2 & 2 & \dots & 2 \\ \dots & \dots & \dots & \dots \\ 2 & 2 & \dots & 2 \end{bmatrix}$$

Ta có $\operatorname{rank} A = \operatorname{rank} A_2 = 3$ với $n \geq 3$. Số chiều của không gian nghiệm là $\dim S = n - \operatorname{rank} A$. Ta muốn có $\dim S \geq 3$ cho nên $n \geq 6$. □

### 6.2.9 Bài toán ngược

Cho $U = \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m \rangle$. Tìm hệ phương trình tuyến tính thuần nhất sao cho không gian nghiệm của nó chính là $U = \langle \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m \rangle$. Ta có

$$\mathbf{x} \in U \iff \text{phương trình } t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \cdots + t_m\mathbf{a}_m = \mathbf{x} \text{ có nghiệm.}$$

Ta sử dụng Định lý 2.1.3 về điều kiện cần và đủ để hệ phương trình có nghiệm đưa đến hệ phương trình tuyến tính thuần nhất cho tọa độ của $\mathbf{x}$.

**6.2.10 Thí dụ.** Cho $U = \langle (\mathbf{a}_1 = (1, 0, 1, 1), \mathbf{a}_2 = (1, -1, 0, 1)) \rangle \subset \mathbb{R}^4$. Tìm hệ phương trình tuyến tính thuần nhất có không gian nghiệm là $U$.

Ta có $\mathbf{x} = (x_1, x_2, x_3, x_4) \in U$ khi và chỉ khi $\mathbf{x}$ biểu thị tuyến tính qua được $\mathbf{a}_1, \mathbf{a}_2$, nghĩa là tồn tại $t_1, t_2$ sao cho $\mathbf{x} = t_1\mathbf{a}_1 + t_2\mathbf{a}_2$. Tương đương với hệ

$$\begin{cases} t_1 + t_2 = x_1 \\ -t_2 = x_2 \\ t_1 = x_3 \\ t_1 + t_2 = x_4 \end{cases}$$

có nghiệm. Ta biến đổi sơ cấp ma trận bổ sung của hệ phương trình trên

$$\begin{bmatrix} 1 & 1 & | & x_1 \\ 0 & -1 & | & x_2 \\ 1 & 0 & | & x_3 \\ 1 & 1 & | & x_4 \end{bmatrix} \xrightarrow[d_4 \to d_4 - d_1]{d_3 \to d_3 - d_1} \begin{bmatrix} 1 & 1 & | & x_1 \\ 0 & -1 & | & x_2 \\ 0 & -1 & | & x_3 - x_1 \\ 0 & 0 & | & x_4 - x_1 \end{bmatrix}$$

$$\xrightarrow{d_3 \to d_3 - d_2} \begin{bmatrix} 1 & 1 & | & x_1 \\ 0 & -1 & | & x_2 \\ 0 & 0 & | & x_3 - x_1 - x_2 \\ 0 & 0 & | & x_4 - x_1 \end{bmatrix}$$

Do đó hệ trên có nghiệm khi và chỉ khi $\begin{cases} x_1 + x_2 - x_3 = 0 \\ x_1 - x_4 = 0. \end{cases}$ □

## Bài tập

**1)** Tìm một cơ sở và số chiều của không gian con $W$ của $\mathbb{R}^4$ sinh bởi hệ vector sau:

**(a)** $((1,4,-1,3),(2,1,-3,-1),(0,2,1,-5))$

**(b)** $((1,-4,-2,1),(1,-3,-1,2),(3,-8,-2,7))$

**2)** Tìm cơ sở số chiều của không gian con của $\mathbb{R}^5$ sinh bởi hệ vector $((2,0,1,3,-1),(1,1,0,-1,1),(0,-2,1,5,3),(1,-3,2,9,5))$.

**3)** Tìm một cơ sở và số chiều của không gian con $W$ của $\mathcal{M}_2(\mathbb{R})$ sinh bởi hệ sau

$$\left(\begin{bmatrix}1 & -5\\ -4 & 2\end{bmatrix},\begin{bmatrix}1 & 1\\ -1 & 5\end{bmatrix},\begin{bmatrix}2 & -4\\ -5 & 7\end{bmatrix},\begin{bmatrix}1 & -7\\ -5 & 1\end{bmatrix}\right)$$

**4)** Tìm một cơ sở và số chiều của không gian con $W$ của $\mathbb{R}_3[x]$ sinh bởi hệ $(x^3+2x^2-2x+1, x^3+3x^2-x+4, 2x^3+x^2-7x-7)$.

**5)** Trong $\mathbb{R}^4$ cho các không gian con sau đây

$$W_1 = \{(a,b,c,d) : b-2c+d=0\},\quad W_2 = \{(a,b,c,d) : a=d, b=2c\}$$

Tìm cơ sở, số chiều của $W_1$, $W_2$ và $W_1 \cap W_2$.

**6)** Trong không gian $\mathbb{R}^5$, cho $W_1$ là không gian con sinh bởi hệ vector $((1,3,-2,2,3),(1,4,-3,4,2),(2,3,-1,-2,9))$ và $W_2$ là không gian con sinh bởi hệ vector $((1,3,0,2,1),(1,5,-6,6,3),(2,5,3,2,1))$. Tìm một cơ sở và số chiều của các không gian con $W_1+W_2$ và $W_1 \cap W_2$.

**7)** Trong không gian $\mathbb{R}^4$, cho $W_1$ là không gian con sinh bởi hệ vector $((1,2,1,1),(3,6,5,7),(4,8,6,8),(8,16,12,20))$ và $W_2$ là không gian con sinh bởi hệ vector $((2,7,2,2),(1,3,1,1),(3,10,4,3),(6,21,7,6))$. Tìm một cơ sở và số chiều của $W_1$, $W_2$, $W_1+W_2$ và $W_1 \cap W_2$.

**8)** Tìm cơ sở, số chiều của các không gian vector $U$, $V$, $U+V$, $U \cap V$ với

$$\begin{cases} U = \langle (2,1,3,1),(1,2,0,1),(-1,1,-3,0)\rangle \\ V = \langle (2,1,3,-1),(-1,1,-3,1),(4,5,3,-1),(1,5,-3,1)\rangle. \end{cases}$$

**9)** Tìm cơ sở và số chiều của không gian nghiệm của các hệ phương trình tuyến tính thuần nhất sau:

**(a)** $\begin{cases} x + 3y + 2z = 0 \\ x + 5y + z = 0 \\ 3x + 5y + 8z = 0 \end{cases}$ **(b)** $\begin{cases} x - 2y + 7z = 0 \\ 2x + 3y - 2z = 0 \\ 2x - y + z = 0 \end{cases}$

**(c)** $\begin{cases} x + 4y + 2z = 0 \\ 2x + y + 5z = 0 \end{cases}$

**10)** Tìm cơ sở và số chiều của không gian nghiệm của các hệ phương trình

**(a)** $\begin{cases} x_1 + 2x_2 - 2x_3 + 2x_4 - x_5 = 0 \\ x_1 + 2x_2 - x_3 + 3x_4 - 2x_5 = 0 \\ 2x_1 + 4x_2 - 7x_3 + x_4 + x_5 = 0 \end{cases}$

**(b)** $\begin{cases} x_1 + 2x_2 - x_3 + 3x_4 - 4x_5 = 0 \\ 2x_1 + 4x_2 - 2x_3 - x_4 + 5x_5 = 0 \\ 2x_1 + 4x_2 - 2x_3 + 4x_4 - 2x_5 = 0 \end{cases}$

**11)** Hãy tìm hệ phương trình tuyến tính thuần nhất mà không gian nghiệm của nó là

**(a)** $\langle (2, -1, 0, 1), (1, 0, -1, 2), (1, -1, 1, -1), (3, -1, -1, 3) \rangle$ trong $\mathbb{R}^4$.

**(b)** $\langle (1, -2, 0, 3, -1), (2, -3, 2, 5, -3), (1, -2, 1, 2, -2) \rangle$ trong $\mathbb{R}^5$.

# § 7 Tổng trực tiếp và không gian thương

## 7.1 Tổng trực tiếp

**7.1.1 Định nghĩa.** Cho $U$ và $W$ là hai không gian con của không gian vector $V$. Nếu mỗi vector $\mathbf{x} \in U + W$ chỉ có một cách biểu diễn duy nhất thành tổng của các vector từ $U$ và $W$ thì không gian con $U + W$ được gọi là *tổng trực tiếp* của $U$ và $W$, ký hiệu $U \oplus W$. Nghĩa là nếu $\mathbf{x} = \mathbf{u} + \mathbf{w} = \mathbf{u}' + \mathbf{w}'$ với $\mathbf{u}, \mathbf{u}' \in U$ và $\mathbf{w}, \mathbf{w}' \in W$ thì $\mathbf{u} = \mathbf{u}'$ và $\mathbf{w} = \mathbf{w}'$.

**7.1.2 Thí dụ.** Trong $\mathbb{R}^4$ xét $U = \langle(1,0,0,0),(0,1,0,0)\rangle$ và $V = \langle(0,0,0,1)\rangle$. Khi đó $U + V = \langle(1,0,0,0),(0,1,0,0),(0,0,0,1)\rangle$. Mỗi vector của $U + V$ đều có dạng $(x, y, 0, z)$ và có biểu diễn duy nhất thành tổng các vector của các không gian con $U$ và $V$ như sau

$$(x, y, 0, z) = (x, y, 0, 0) + (0, 0, 0, z).$$

Vậy $U + V$ là một tổng trực tiếp, nên ta có thể viết $U \oplus V$. Ta nhận thấy $U \cap V = \{(0,0,0,0)\}$. Điều này vẫn đúng trong trường hợp tổng quát và ta có định lý sau. □

**7.1.3 Định lý.** *Cho hai không gian con $U_1$ và $U_2$ của không gian vector $V$. Khi đó các khẳng định sau là tương đương.*

**(i)** *$U_1 + U_2$ là tổng trực tiếp.*

**(ii)** $U_1 \cap U_2 = \{\mathbf{0}\}$

**(iii)** *Vector $\mathbf{0}$ biểu diễn duy nhất thành tổng của hai vector từ $U_1$ và $U_2$, tức là*

$$\begin{cases} \mathbf{0} = \mathbf{u}_1 + \mathbf{u}_2 \\ \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2 \end{cases} \qquad \Longleftrightarrow \qquad \mathbf{u}_1 = \mathbf{u}_2 = \mathbf{0}.$$

*Chứng minh.* (i) $\Rightarrow$ (ii) Giả sử $U_1 + U_2$ là tổng trực tiếp và $\mathbf{x}$ là vector bất kỳ của $U_1 \cap U_2$. Khi đó $\mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x}$ đều là biểu diễn của $\mathbf{x}$ thành tổng các vector từ $U_1$ và $U_2$. Do tính duy nhất của biểu diễn này, ta có $\mathbf{x} = \mathbf{0}$. Tức là $U_1 \cap U_2 = \{\mathbf{0}\}$.

(ii) $\Rightarrow$ (iii) Giả sử $U_1 \cap U_2 = \{\mathbf{0}\}$ và $\mathbf{0} = \mathbf{u}_1 + \mathbf{u}_2$ với $\mathbf{u}_1 \in U_1$ và $\mathbf{u}_2 \in U_2$. Từ đẳng thức trên suy ra $\mathbf{u}_2 = -\mathbf{u}_1 \in U_1$, cho nên $\mathbf{u}_2 \in U_1 \cap U_2$. Do đó $\mathbf{u}_2 = \mathbf{0}$ và suy ra $\mathbf{x}_1 = \mathbf{0}$. Từ đó ta dễ dàng kết luận được $\mathbf{0} = \mathbf{u}_1 + \mathbf{u}_2$ với $\mathbf{u}_1 \in U_1$ và $\mathbf{u}_2 \in U_2$ khi và chỉ khi $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{0}$.

(iii) $\Rightarrow$ (i) Giả sử $\mathbf{0}$ chỉ có một cách biểu diễn thành tổng $\mathbf{0} = \mathbf{0} + \mathbf{0}$ của các vector từ $U_1$ và $U_2$. Lấy tùy ý $\mathbf{x} \in U_1 + U_2$. Giả sử ta có biểu diễn của $\mathbf{x}$ thành tổng các vector từ $U_1$ và $U_2$ là $\mathbf{x} = \mathbf{u}_1 + \mathbf{u}_2 = \mathbf{u}_1' + \mathbf{u}_2'$ với $\mathbf{u}_1, \mathbf{u}_1' \in U_1$ và $\mathbf{u}_2, \mathbf{u}_2' \in U_2$. Khi đó ta có

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (\mathbf{u}_1 + \mathbf{u}_2) - (\mathbf{u}_1' + \mathbf{u}_2') = (\mathbf{u}_1 - \mathbf{u}_1') + (\mathbf{u}_2 - \mathbf{u}_2').$$

Do $\mathbf{u}_1 - \mathbf{u}_1' \in U_1$ và $\mathbf{u}_2 - \mathbf{u}_2' \in U_2$ nên ta suy ra $\mathbf{u}_1 - \mathbf{u}_1' = \mathbf{u}_2 - \mathbf{u}_2' = \mathbf{0}$ hay $\mathbf{u}_1 = \mathbf{u}_1'$ và $\mathbf{u}_2 = \mathbf{u}_2'$. Vì vậy ta kết luận được $U_1 + U_2$ là một tổng trực tiếp. □

**7.1.4 Hệ quả.** *Nếu tổng $U + W$ của không gian con hữu hạn chiều $U$ và $W$ của không gian vector $V$ là tổng trực tiếp thì*

$$\dim U + \dim W = \dim(U \oplus W).$$

*Hơn nữa nếu $(\mathbf{a}_1, \ldots, \mathbf{a}_m)$ và $(\mathbf{a}_{m+1}, \ldots, \mathbf{a}_n)$ lần lượt là cơ sở của $U$ và $W$ (với $0 \leq m \leq n$) thì $(\mathbf{a}_1, \ldots, \mathbf{a}_m, \mathbf{a}_{m+1}, \ldots, \mathbf{a}_n)$ là một cơ sở của $U \oplus W$.*

*Chứng minh.* Kết quả suy ra trực tiếp từ các Định lý 5.2.5 và 7.1.3. □

## 7.2 Không gian thương

**7.2.1** Giả sử $V$ là một không gian vector và $W$ là một không gian con của $V$. Trên $V$ ta xét quan hệ $\sim$ định nghĩa như sau: với mọi $\mathbf{x}, \mathbf{y} \in V$

$$\mathbf{x} \sim \mathbf{y} \iff \mathbf{x} - \mathbf{y} \in W$$

$\sim$ là một quan hệ tương đương. Thật vậy

- Với mọi $\mathbf{x} \in V$, ta có $\mathbf{x} - \mathbf{x} = \mathbf{0} \in W$ nên $\mathbf{x} \sim \mathbf{x}$, nghĩa là $\sim$ phản xạ
- Với mọi $\mathbf{x}, \mathbf{y} \in V$, nếu $\mathbf{x} \sim \mathbf{y}$ thì $\mathbf{x} - \mathbf{y} \in W$, suy ra $\mathbf{y} - \mathbf{x} \in W$ hay $\mathbf{y} \sim \mathbf{x}$, nghĩa là $\sim$ đối xứng.
- Với mọi $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, nếu $\mathbf{x} \sim \mathbf{y}$ và $\mathbf{y} \sim \mathbf{z}$ thì $\mathbf{x} - \mathbf{y}, \mathbf{y} - \mathbf{z} \in W$, suy ra $\mathbf{x} - \mathbf{z} = (\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z}) \in W$ hay $\mathbf{x} \sim \mathbf{z}$, nghĩa là $\sim$ bắc cầu.

Khi đó, tập các lớp tương đương trên $V$ theo quan hệ $\sim$ được ký hiệu $V/_W$. Lớp tương đương của $\mathbf{x} \in V$ được ký hiệu là $\tilde{\mathbf{x}}$.

**7.2.2 Nhận xét.** **(1)** Nếu $W = \{\mathbf{0}\}$, không gian con không, thì với mọi $\mathbf{x}, \mathbf{y} \in V$ ta có $\mathbf{x} \sim \mathbf{y}$ khi và chỉ khi $\mathbf{x} - \mathbf{y} = \mathbf{0}$ hay $\mathbf{x} = \mathbf{y}$. Do đó $\tilde{\mathbf{x}} = \{\mathbf{x}\}$ với mọi $\mathbf{x} \in V$, cho nên $V/_W = V/_{\{\mathbf{x}\}} \equiv V$.

**(2)** Nếu $W = V$, không gian con không thực sự, thì với mọi $\mathbf{x}, \mathbf{y} \in V$ ta đều có $\mathbf{x} \sim \mathbf{y}$. Bởi vậy $\tilde{\mathbf{0}} = \tilde{\mathbf{x}} = V$ với mọi $\mathbf{x} \in V$. Do đó $V/_W = V/_V = \{\tilde{\mathbf{0}}\}$.

**7.2.3 Mệnh đề.** *Cho $W$ là một không gian con của không gian vector $V$. Khi đó với mọi $\mathbf{x} \in V$ ta có $\tilde{\mathbf{x}} = \mathbf{x} + W = \{\mathbf{x} + \mathbf{w} : \mathbf{w} \in W\}$. Đặc biệt $\tilde{\mathbf{0}} = W$.*

*Chứng minh.* Với mọi $\mathbf{w} \in W$, ta có $\mathbf{x} + \mathbf{w} \sim \mathbf{x}$ nên $\mathbf{x} + \mathbf{w} \in \tilde{\mathbf{x}}$, do đó $\mathbf{x} + W \subseteq \tilde{\mathbf{x}}$. Hơn nữa, với mọi $\mathbf{y} \in \tilde{\mathbf{x}}$ ta có $\mathbf{y} \sim \mathbf{x}$ nên $\mathbf{y} - \mathbf{x} = \mathbf{w} \in W$; do đó $\mathbf{y} = \mathbf{x} + \mathbf{w} \in \mathbf{x} + W$. Vậy $\tilde{\mathbf{x}} \subseteq \mathbf{x} + W$. Do đó ta được $\tilde{\mathbf{x}} = \mathbf{x} + W$. Nói riêng $\tilde{\mathbf{0}} = \mathbf{0} + W = W$. □

**7.2.4 Định lý.** *Cho $W$ là một không gian con của không gian vector $V$. Tập hợp thương $V/_W$ cùng với hai phép toán cộng và nhân với vô hướng xác định bởi*

$$\tilde{\mathbf{x}} + \tilde{\mathbf{y}} = \widetilde{\mathbf{x} + \mathbf{y}}, \qquad \lambda\tilde{\mathbf{x}} = \widetilde{\lambda\mathbf{x}}$$

*với mọi $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in V/_W$ và mọi $\lambda \in \mathbb{K}$ lập thành một không gian vector và gọi là **không gian thương** của $V$ theo không gian con $W$.*

*Chứng minh.* Trước hết, cần kiểm chứng rằng các phép toán định nghĩa như trên là hợp lý, tức là không phụ thuộc phần tử đại diện. Thật vậy, nếu $\tilde{\mathbf{x}} = \widetilde{\mathbf{x}'}$ và $\tilde{\mathbf{y}} = \widetilde{\mathbf{y}'}$ thì $\mathbf{x} \sim \mathbf{x}'$ và $\mathbf{y} \sim \mathbf{y}'$ hay $\mathbf{x} - \mathbf{x}', \mathbf{y} - \mathbf{y}' \in W$, cho nên $(\mathbf{x} + \mathbf{y}) - (\mathbf{x}' + \mathbf{y}') \in W$ hay $\mathbf{x} + \mathbf{y} \sim \mathbf{x}' + \mathbf{y}'$ tức là $\widetilde{\mathbf{x} + \mathbf{y}} = \widetilde{\mathbf{x}' + \mathbf{y}'}$. Bên cạnh đó ta cũng có $\lambda\mathbf{x} - \lambda\mathbf{x}' = \lambda(\mathbf{x} - \mathbf{x}') \in W$ hay $\lambda\mathbf{x} \sim \lambda\mathbf{x}'$, tức là $\widetilde{\lambda\mathbf{x}} = \widetilde{\lambda\mathbf{x}'}$.

Dễ dàng kiểm chứng hai phép toán trên $V/_W$ thỏa 8 tiên đề của không gian vector, trong đó vector không chính là $\tilde{\mathbf{0}}$, vector đối của $\tilde{\mathbf{x}}$ là $\widetilde{-\mathbf{x}}$. Trình bày chi tiết cho việc kiểm chứng ấy xin dành cho bạn đọc xem như bài tập. Vậy $V/_W$ là một không gian vector. □

**7.2.5 Định lý. (Số chiều của không gian thương)** *Nếu $W$ là không gian con của không gian hữu hạn chiều $V$ thì không gian thương cũng hữu hạn chiều và*

$$\dim(V/_W) = \dim V - \dim W.$$

*Chứng minh.* Giả sử $\dim V = n$. Khi đó, đương nhiên $W$ hữu hạn chiều và $\dim W = m \leq n$ ($m = 0$ nếu $W = \{\mathbf{0}\}$). Lấy một cơ sở $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ của $W$ (cơ sở này là $\varnothing$ nếu $W = \{\mathbf{0}\}$) rồi bổ sung để được cơ sở $\mathscr{B} = (\mathbf{e}_1, \dots, \mathbf{e}_m, \mathbf{e}_{m+1}, \dots, \mathbf{e}_n)$ của $V$ (nếu $W = V$ thì $m = n$ và không cần phải bổ sung vector nào). Ta chứng tỏ hệ $\widetilde{\mathscr{B}} = (\widetilde{\mathbf{e}_{m+1}}, \dots, \widetilde{\mathbf{e}_n})$ là một cơ sở của $V/_W$.

Giả sử $\alpha_{m+1}\widetilde{\mathbf{e}_{m+1}} + \cdots + \alpha_n\widetilde{\mathbf{e}_n} = \tilde{\mathbf{0}}$ là một tổ hợp tuyến tính bất kỳ bằng không của $\widetilde{\mathscr{B}}$ trong $V/_W$. Khi đó $(\alpha_{m+1}\mathbf{e}_{m+1} + \cdots + \alpha_n\mathbf{e}_n) = \tilde{\mathbf{0}}$ cho nên $\alpha_{m+1}\mathbf{e}_{m+1} + \cdots + \alpha_n\mathbf{e}_n \sim \mathbf{0}$ hay $\alpha_{m+1}\mathbf{e}_{m+1} + \cdots + \alpha_n\mathbf{e}_n \in W$. Do đó nó phải biểu thị tuyến tính qua cơ sở $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ của $W$, tức là phải có $\alpha_{m+1}\mathbf{e}_{m+1} + \cdots + \alpha_n\mathbf{e}_n = \alpha_1\mathbf{e}_1 + \cdots + \alpha_m\mathbf{e}_m$ hay

$$\alpha_1\mathbf{e}_1 + \cdots + \alpha_m\mathbf{e}_m + \alpha_{m+1}\mathbf{e}_{m+1} + \cdots + \alpha_n\mathbf{e}_n = \mathbf{0}.$$

Vì $\mathscr{B}$ là một cơ sở của $V$, nên phải có $-\alpha_1 = \cdots = -\alpha_m = \alpha_{m+1} = \cdots = \alpha_n = 0$. Suy ra $\widetilde{\mathscr{B}}$ độc lập tuyến tính.

Mặt khác, với mọi $\tilde{\mathbf{x}} \in V/_W$, do $\mathbf{x} \in V$ nên ta có thể viết $\mathbf{x} = \sum_{i=1}^{n} \lambda_i\mathbf{e}_i$ nên

$$\mathbf{x} - \sum_{i=m+1}^{n} \lambda_i\mathbf{e}_i = \sum_{i=1}^{m} \lambda_i\mathbf{e}_i \in W \qquad \text{hay} \qquad \mathbf{x} \sim \sum_{i=m+1}^{n} \lambda_i\mathbf{e}_i.$$

Từ đó ta có $\tilde{\mathbf{x}} = \widetilde{\sum_{i=m+1}^{n} \lambda_i\mathbf{e}_i} = \sum_{i=m+1}^{n} \lambda_i\widetilde{\mathbf{e}_i}$. Tức là $\tilde{\mathbf{x}}$ biểu thị tuyến tính được qua $\tilde{B}$. Vậy $\widetilde{\mathscr{B}}$ là một cơ sở của $V/_W$. Do đó $\dim(V/_W) = n - m = \dim V - \dim W$. □

**7.2.6 Định nghĩa.** Cho $W$ là một không gian con của không gian hữu hạn chiều $V$. Hiệu $\dim V - \dim W$ được gọi là *số đối chiều* của $W$ trong $V$, ký hiệu $\operatorname{codim}_V(W)$. Không gian con có số đối chiều bằng 1 gọi là *siêu không gian con* của $V$. Như vậy ta có $\dim(V/_W) = \operatorname{codim}_V(W)$.

**7.2.7 Thí dụ.** Trong $\mathbb{R}^3$ xét hai không gian con $W = \langle(1,1,1),(0,1,1)\rangle$ và $Z = \langle(0,0,1)\rangle$. Hãy xác định số chiều và một cơ sở của các không gian thương $\mathbb{R}^3/_W$ và $\mathbb{R}^3/_Z$.

*Giải.* Ta thấy $\operatorname{rank}((1,1,1),(0,1,1)) = 2$ và $\operatorname{rank}((0,0,1)) = 1$. Do đó $\dim W = 2$ với $((1,1,1),(0,1,1))$ là một cơ sở của $W$, còn $\dim Z = 1$ và đương nhiên $((0,0,1))$ là cơ sở của $Z$.

Hơn nữa, vì $\begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0$ nên $((1,1,1),(0,1,1),(0,0,1))$ là một cơ sở của $\mathbb{R}^3$. Từ đó suy ra $\dim(\mathbb{R}^3/_W) = \dim \mathbb{R}^3 - \dim W = 3 - 2 = 1$ và $\mathbb{R}^3/_W = \langle(\widetilde{0,0,1})\rangle$. Còn $\dim(\mathbb{R}^3/_Z) = \dim \mathbb{R}^3 - \dim Z = 3 - 1 = 2$ và $\mathbb{R}^3/_Z = \langle(\widetilde{1,1,1}),(\widetilde{0,1,1})\rangle$. □

## Bài tập

**1**) Trong $\mathbb{R}^4$ hãy tìm bao tuyến tính $W$ của hệ vector sau đây

**(a)** $(\mathbf{u} = (1,1,0,0), \mathbf{v} = (0,1,1,0), \mathbf{w} = (0,0,1,1))$

**(b)** $(\mathbf{u} = (1,1,1,0), \mathbf{v} = (0,1,1,1))$

**(c)** $(\mathbf{u} = (1,2,1,2), \mathbf{v} = (2,1,2,1), \mathbf{w} = (3,3,3,3))$

Trong mỗi trường hợp hãy tính $\dim(\mathbb{R}^4/_W)$.

**2**) Hoàn thành việc chứng minh Định lý 7.2.4.

**3**) Cho $W_1, W_2, \ldots, W_m$ với $m \geq 2$ là các không gian con hữu hạn chiều của không gian vector $V$. Chứng minh rằng các không gian con $\bigcup\limits_{i=1}^{m} W_i$ và $\sum\limits_{i=1}^{m} W_i$ cũng hữu hạn chiều và

$$\sum_{i=1}^{m} \dim W_i = \dim\Big(\sum_{i=1}^{m} W_i\Big) + \sum_{i=1}^{m} \dim\Big[W_i \cap \Big(\sum_{j<i} W_j\Big)\Big].$$

**4**) Chứng minh rằng $\mathbb{R}$ và $\mathbb{C}$ đều là các không gian vector vô hạn chiều trên $\mathbb{Q}$.

# Chương V

# Ánh xạ tuyến tính

## § 1 Khái niệm và tính chất cơ bản ánh xạ tuyến tính

### 1.1 Khái niệm và ví dụ

**1.1.1 Định nghĩa.** Cho hai không gian vector $U$ và $V$ trên trường $\mathbb{K}$. Ánh xạ $f : V \to U$ gọi là **ánh xạ tuyến tính** nếu

**(i)** Với mọi $\mathbf{a}, \mathbf{b} \in V$, ta có $f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b})$

**(ii)** Với mọi $\lambda \in \mathbb{K}$, với mọi $\mathbf{a} \in V$, ta có $f(\lambda\mathbf{a}) = \lambda f(\mathbf{a})$.

Trường hợp $U = V$ thì ta gọi $f$ là **phép biến đổi tuyến tính** hay **toán tử tuyến tính** trên $V$.

**1.1.2 Định lý.** *Cho hai không gian vector $U$ và $V$ trên trường $\mathbb{K}$. Ánh xạ $f : V \to U$ là ánh xạ tuyến tính khi và chỉ khi với mọi $\alpha, \beta \in \mathbb{K}$ và mọi $\mathbf{a}, \mathbf{b} \in V$ ta có $f(\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha f(\mathbf{a}) + \beta f(\mathbf{b})$.*

*Chứng minh.* Giả sử $f$ là ánh xạ tuyến tính. Do $\alpha\mathbf{a}, \beta\mathbf{b} \in V$, ta suy ra được $f(\alpha\mathbf{a} + \beta\mathbf{b}) = f(\alpha\mathbf{a}) + f(\beta\mathbf{b}) = \alpha f(\mathbf{a}) + \beta f(\mathbf{b})$.

Ngược lại, với mọi $\mathbf{a}, \mathbf{b} \in V$ và $\alpha \in \mathbb{K}$, ta có $f(\mathbf{a}+\mathbf{b}) = f(1\mathbf{a}+1\mathbf{b}) = 1f(\mathbf{a})+1f(\mathbf{b}) = f(\mathbf{a})+f(\mathbf{b})$ và $f(\alpha\mathbf{a}) = f(\alpha\mathbf{a}+0\mathbf{b}) = \alpha f(\mathbf{a})+0f(\mathbf{b}) = \alpha f(\mathbf{a}) + \mathbf{0} = \alpha f(\mathbf{a})$. Vậy $f$ là ánh xạ tuyến tính. $\square$

**1.1.3 Thí dụ.** Cho hai không gian vector $U$ và $V$ trên trường $\mathbb{K}$. Ánh xạ $\theta : V \to U$ xác định như sau $\theta(\mathbf{a}) = \mathbf{0}$ với mọi $\mathbf{a} \in V$. Ta có $\theta$ là ánh xạ tuyến tính bởi vì với mọi $\mathbf{a}, \mathbf{b} \in V$ và $\alpha, \beta \in \mathbb{K}$, ta có $\theta(\alpha\mathbf{a} + \beta\mathbf{b}) = \mathbf{0} = \alpha\mathbf{0} + \beta\mathbf{0} = \alpha\theta(\mathbf{a}) + \beta\theta(\mathbf{b})$. □

**1.1.4 Thí dụ.** Cho không gian vector $V$ và ánh xạ đồng nhất $\mathrm{Id}_V : V \to V$, $\mathrm{Id}_V(\mathbf{a}) = \mathbf{a}$ với mọi $\mathbf{a} \in V$. $\mathrm{Id}_V$ là ánh xạ tuyến tính trên $V$, nó là phép biến đổi tuyến tính trên $V$. Thật vậy, với mọi $\mathbf{a}, \mathbf{b} \in V$ và mọi $\alpha, \beta \in \mathbb{K}$, ta có $\mathrm{Id}_V(\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha\mathbf{a} + \beta\mathbf{b} = \alpha\mathrm{Id}_V(\mathbf{a}) + \beta\mathrm{Id}_V(\mathbf{b})$. □

**1.1.5 Thí dụ.** Giả sử $V$ là một $\mathbb{K}$-không gian vector $n$ chiều và $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ là một cơ sở của nó. Khi đó, ánh xạ tọa độ $c_{\mathscr{B}} : V \to \mathbb{K}^n$ xác định bởi $c_{\mathscr{B}}(\mathbf{x}) = \mathbf{x}|_{\mathscr{B}}$ là một ánh xạ tuyến tính. □

**1.1.6 Thí dụ.** Ánh xạ $p_i : \mathbb{K}^n \to \mathbb{K}$ xác định bởi $p_i(x_1, \dots, x_n) = x_i$ trong đó $i = 1, \dots, n$ là một ánh xạ tuyến tính và gọi là *phép chiếu chính tắc* lên thành phần thứ $i$. □

**1.1.7 Thí dụ.** Cho ma trận $M$ có cấp $m \times n$ trên $\mathbb{K}$. Khi đó ánh xạ

$$\begin{aligned} f_M : \mathbb{K}^n &\to \mathbb{K}^m \\ \mathbf{x} &\mapsto M\mathbf{x} \end{aligned}$$

là một ánh xạ tuyến tính. (Nhắc lại ta đồng nhất mỗi phần tử của $\mathbb{K}^n$ với ma trận cột.) □

**1.1.8 Thí dụ.** Cho $W$ là một không gian con của $V$. Khi đó ánh xạ $p : V \to V/_W$ xác định bởi $p(\mathbf{x}) = \tilde{\mathbf{x}}$ là một ánh xạ tuyến tính và gọi là *phép chiếu chính tắc* $V$ lên $V/_W$. □

**1.1.9 Thí dụ.** Cho $W$ là một không gian con của $V$. Khi đó ánh xạ nhúng $i : W \to V$, xác định bởi $i(\mathbf{x}) = \mathbf{x}$ với mọi $\mathbf{x} \in W$ là một ánh xạ tuyến tính và gọi là *phép nhúng chính tắc* $W$ vào $V$. □

**1.1.10 Thí dụ.** Xét không gian vector $\mathbb{R}[x]$. Ánh xạ $\varphi : \mathbb{R}[x] \to \mathbb{R}[x]$ xác định bởi $\varphi(f(x)) = f'(x)$. Ta có thể thấy $\varphi$ là phép biến đổi tuyến tính trên $\mathbb{R}[x]$. Thật vậy, với mọi $f(x), g(x) \in \mathbb{R}[x]$ và mọi $\alpha, \beta \in \mathbb{R}$, ta có $\varphi(\alpha f(x) + \beta g(x)) = (\alpha f(x) + \beta g(x))' = \alpha f'(x) + \beta g'(x) = \alpha\varphi(f(x)) + \beta\varphi(g(x))$. □

**1.1.11 Thí dụ.** Ánh xạ $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi $f(x_1, x_2, x_3) = (x_2 - x_1, x_3 - x_2)$ là một ánh xạ tuyến tính. Thật vậy, với $\mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3) \in V$ và $\alpha, \beta \in \mathbb{R}$, ta có $f(\alpha\mathbf{x} + \beta\mathbf{y}) = f(\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \alpha x_3 + \beta y_3) = (\alpha x_2 + \beta y_2 - (\alpha x_1 + \beta y_1), \alpha x_3 + \beta y_3 - (\alpha x_2 + \beta y_2)) = \alpha(x_2 - x_1, x_3 - x_2) + \beta(y_2 - y_1, y_3 - y_2) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$. □

**1.1.12 Thí dụ.** Trong không gian vector $\mathbb{R}^3$, ánh xạ $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (x_1 + x_2, 0, x_1 + x_3)$ là một phép biến đổi tuyến tính trên $\mathbb{R}^3$. Thật vậy, với mọi $\mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ và mọi $\alpha, \beta \in \mathbb{R}$, ta có $f(\alpha\mathbf{x} + \beta\mathbf{y}) = (\alpha x_1 + \beta y_1 + \alpha x_2 + \beta y_2, 0, \alpha x_1 + \beta y_1 + \alpha x_3 + \beta y_3) = \alpha(x_1 + x_2, 0, x_1 + x_3) + \beta(y_1 + y_2, 0, y_1 + y_3) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$. □

**1.1.13 Thí dụ.** Trong không gian vector $\mathbb{R}^3$, ánh xạ $g : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $g(x_1, x_2, x_3) = (x_1 + x_2, 1, x_1 + x_3)$ không là ánh xạ tuyến tính. Thật vậy, lấy $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ tùy ý và $\lambda = 0$, ta có $g(\lambda\mathbf{x}) = g(\mathbf{0}) = (0, 1, 0)$, trong khi đó $\lambda g(\mathbf{x}) = 0g(\mathbf{x}) = (0, 0, 0)$. □

## 1.2 Các tính chất cơ bản của ánh xạ tuyến tính

**1.2.1 Mệnh đề.** *Cho ánh xạ tuyến tính $f : V \to U$. Ta có*

**(i)** $f(\mathbf{0}) = \mathbf{0}$

**(ii)** $f(-\mathbf{a}) = -f(\mathbf{a})$

**(iii)** $f(\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n) = \lambda_1 f(\mathbf{a}_1) + \lambda_2 f(\mathbf{a}_2) + \cdots + \lambda_n f(\mathbf{a}_n)$.

*Chứng minh.*

**(i)** Với $\mathbf{a} \in V$, ta có $f(\mathbf{a}) = f(\mathbf{a} + \mathbf{0}) = f(\mathbf{a}) + f(\mathbf{0})$. Suy ra $f(\mathbf{0}) = \mathbf{0}$.

**(ii)** Ta có $\mathbf{0} = f(\mathbf{0}) = f(\mathbf{a} + (-\mathbf{a})) = f(\mathbf{a}) + f(-\mathbf{a})$. Suy ra $f(-\mathbf{a}) = -f(\mathbf{a})$.

**(iii)** Dùng quy nạp và Định lý 1.1.2 ta sẽ có được kết quả. □

**1.2.2 Mệnh đề.** *Cho hai ánh xạ tuyến tính $f : V \to V'$ và $g : V' \to V''$. Khi đó ánh xạ hợp thành $g \circ f : V \to V''$ cũng là một ánh xạ tuyến tính. Nói riêng, ánh xạ thu hẹp $f|_W : W \to V'$ của ánh xạ tuyến tính $f : V \to V'$ trên không gian con $W$ của $V$ cũng là ánh xạ tuyến tính.*

*Chứng minh.* Với mọi $\mathbf{x}, \mathbf{y} \in V$ và mọi $\lambda, \mu \in \mathbb{K}$ ta có

$$\begin{aligned} g \circ f(\lambda\mathbf{x} + \mu\mathbf{y}) &= g(f(\lambda\mathbf{x} + \mu\mathbf{y}) \\ &= g(\lambda f(\mathbf{x}) + \mu f(\mathbf{y})) \\ &= \lambda g(f(\mathbf{x})) + \mu g(f(\mathbf{y})) \\ &= \lambda g \circ f(\mathbf{x}) + \mu g \circ f(\mathbf{x}). \end{aligned}$$

Do đó $g \circ f$ là một ánh xạ tuyến tính.

Đặc biệt với $i : W \to V$ là phép nhúng chính tắc ta có $f|_W = f \circ i : W \to V'$ là một ánh xạ tuyến tính. □

**1.2.3 Mệnh đề.** *Cho ánh xạ tuyến tính $f$ từ không gian vector $V$ đến không gian vector $U$ và $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một hệ vector của $V$. Khi đó, nếu hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ phụ thuộc tuyến tính trong $V$, thì $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ là một hệ vector phụ thuộc tuyến tính của $U$. Hay, nếu $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ là một hệ vector độc lập tuyến tính trong $U$, thì $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ độc lập tuyến tính trong $V$.*

*Chứng minh.* Do $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là hệ phụ thuộc tuyến tính của $V$, nên tồn tại $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ không đồng thời bằng 0 sao cho $\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n = \mathbf{0}$. Suy ra $f(\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n) = \mathbf{0}$, cho nên $\lambda_1 f(\mathbf{a}_1) + \lambda_2 f(\mathbf{a}_2) + \cdots + \lambda_n f(\mathbf{a}_n) = \mathbf{0}$. Vậy hệ $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ phụ thuộc tuyến tính. □

**1.2.4 Thí dụ.** Cho $f$ là một toán tử tuyến tính trên không gian vector $V$ và cho $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ là một hệ vector thỏa $f(\mathbf{x}_1) = \mathbf{x}_1$ và $f(\mathbf{x}_k) = \mathbf{x}_k - \mathbf{x}_{k-1}$ với mọi $k = 2, \ldots, n-1$. Khi đó, $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ là một hệ độc lập tuyến tính.

*Giải*. Rõ ràng khẳng định đúng với $n = 1$. Giả sử khẳng định đúng với $n-1$. Xét tổ hợp tuyến tính $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n = \mathbf{0}$. Lấy giá trị của $f$ hai vế ta được

$$a_1\mathbf{x}_1 + a_2 f(\mathbf{x}_2) + \cdots + a_n f(\mathbf{x}_n) = \mathbf{0}.$$

Từ đó suy ra $a_2(f(\mathbf{x}_2) - \mathbf{x}_2) + \cdots + a_n(f(\mathbf{x}_n) - \mathbf{x}_n) = \mathbf{0}$ hay $a_2\mathbf{x}_1 + \cdots + a_n\mathbf{x}_{n-1} = \mathbf{0}$. Theo giả thiết quy nạp $(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1})$ độc lập tuyến tính nên suy ra $a_2 = \cdots = a_n = 0$. Vậy ta phải có $a_1\mathbf{x}_1 = \mathbf{0}$, cho nên $a_1 = 0$. Do đó $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ là một hệ độc lập tuyến tính. Khẳng định đúng với $n$. $\square$

**1.2.5 Định lý.** *Cho ánh xạ tuyến tính $f$ từ không gian vector $V$ đến không gian vector $U$ và $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một hệ vector của $V$. Khi đó, ta có* $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \geq \operatorname{rank}(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$.

*Chứng minh.* Giả sử $\operatorname{rank}\{f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n)\} = k$. Nếu $k = 0$, thì rõ ràng $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \geq 0$. Nếu $k > 0$, thì tồn tại hệ con độc lập tuyến tính tối đại của hệ vector $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ và gọi là $(f(\mathbf{a}_{i_1}), f(\mathbf{a}_{i_2}), \ldots, f(\mathbf{a}_{i_k}))$. Theo Mệnh đề 1.2.3 ta phải có hệ vector $(\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \ldots, \mathbf{a}_{i_k})$ độc lập tuyến tính. Do đó $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \geq k$. $\square$

## 1.3 Tiêu chuẩn xác định ánh xạ tuyến tính

**1.3.1 Định lý.** *Cho $U$ và $V$ là hai không gian vector trên trường $\mathbb{K}$, $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một cơ sở của $V$ ($\dim V = n$) và $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ là một hệ vector bất kì của $U$. Khi đó, tồn tại duy nhất một ánh xạ tuyến tính $f : V \to U$ thỏa $f(\mathbf{a}_i) = \mathbf{b}_i$ với mọi $i = 1, 2, \ldots, n$.*

*Chứng minh.* Ta xác định một ánh xạ $f : V \to U$ như sau: với mọi $\mathbf{x} \in V$ tồn tại duy nhất tọa độ $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ của $\mathbf{x}$ đối với $\mathscr{A}$, và ta có thể viết $\mathbf{x} = \lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n$; khi đó,

$$f(\mathbf{x}) = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \cdots + \lambda_n\mathbf{b}_n \in U.$$

Từ đó ta có thể thấy $f(\mathbf{a}_i) = \mathbf{b}_i$ với mọi $i = 1, 2, \ldots, n$.

Với mọi $k \in \mathbb{K}$, ta có tọa độ của $k\mathbf{x}$ là $(k\lambda_1, k\lambda_2, \ldots, k\lambda_n)$, cho nên $f(k\mathbf{x}) = k\lambda_1\mathbf{b}_1 + k\lambda_2\mathbf{b}_2 + \cdots + k\lambda_n\mathbf{b}_n = k(\lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \cdots + \lambda_n\mathbf{b}_n) = kf(\mathbf{x})$.

Lấy thêm $\mathbf{y} \in V$ tùy ý và có biểu diễn $\mathbf{y} = \alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots + \alpha_n\mathbf{a}_n$. Ta có $f(\mathbf{y}) = \alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \cdots + \alpha_n\mathbf{b}_n$ và $\mathbf{x} + \mathbf{y} = (\lambda_1 + \alpha_1)\mathbf{a}_1 + (\lambda_2 + \alpha_2)\mathbf{a}_2 + \cdots + (\lambda_n + \alpha_n)\mathbf{a}_n$. Khi đó

$$
\begin{aligned}
f(\mathbf{x} + \mathbf{y}) &= (\lambda_1 + \alpha_1)\mathbf{b}_1 + (\lambda_2 + \alpha_2)\mathbf{b}_2 + \cdots + (\lambda_n + \alpha_n)\mathbf{b}_n \\
&= \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \cdots + \lambda_n\mathbf{b}_n + \alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \cdots + \alpha_n\mathbf{b}_n \\
&= f(\mathbf{x}) + f(\mathbf{y}).
\end{aligned}
$$

Vậy ánh xạ $f$ xác định như trên là một ánh xạ tuyến tính.

Giả sử có một ánh xạ tuyến tính $g : V \to U$ thỏa $g(\mathbf{a}_i) = \mathbf{b}_i$ với mọi $i = 1, \ldots, n$, ta chứng minh $f = g$. Với mọi $\mathbf{x} \in V$, ta có $\mathbf{x} = \lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n$. Khi đó, ta có

$$
\begin{aligned}
g(\mathbf{x}) &= g(\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n) \\
&= \lambda_1 g(\mathbf{a}_1) + \lambda_2 g(\mathbf{a}_2) + \cdots + \lambda_n g(\mathbf{a}_n) \\
&= \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \cdots + \lambda_n\mathbf{b}_n \\
&= f(\mathbf{x}).
\end{aligned}
$$

Vậy $f = g$. □

**1.3.2 Nhận xét. (a)** Trong phép chứng minh trên, ta còn chỉ ra công thức xác định $f$. Đó là công thưc

$$f(\mathbf{x}) = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2 + \cdots + \lambda_n\mathbf{b}_n \in U$$

trong đó $\mathbf{x}|_{\mathscr{A}} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

**(b)** Trong định lý trên không đòi hỏi $U$ là một không gian hữu hạn chiều; cũng không đòi hỏi hệ $(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ phải là một cơ sở của $U$.

**1.3.3 Thí dụ.** Trong $\mathbb{R}^3$ xét cơ sở chính tắc $\mathscr{E} = (\mathbf{e}_1 = (1, 0, 0), \mathbf{e}_2 = (0, 1, 0), \mathbf{e}_3 = (0, 0, 1))$ và trong $\mathbb{R}^2$ cho 3 vector $\mathbf{v}_1 = (1, 1)$, $\mathbf{v}_2 = (2, 3)$, $\mathbf{v}_3 = (4, 5)$. Hãy xác định ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^2$ sao cho $f(\mathbf{e}_i) = \mathbf{v}_i$ với $i = 1, 2, 3$.

*Giải.* Với mỗi $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$, vì $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ nên ta có

$$\begin{aligned} f(\mathbf{x}) &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 \\ &= (x_1, x_1) + (2x_2, 3x_2) + (4x_3, 5x_3) \\ &= (x_1 + 2x_2 + 4x_3, x_1 + 3x_2 + 5x_3). \end{aligned}$$

Đây chính là công thức xác định ánh xạ tuyến tính $f$. □

**1.3.4 Thí dụ.** Trong $\mathbb{R}^3$ cho hai hệ $(\mathbf{u}_1 = (1,1,0), \mathbf{u}_2 = (0,1,1), \mathbf{u}_3 = (1,0,1))$ và $(\mathbf{v}_1 = (0,2,2), \mathbf{v}_2 = (2,0,2), \mathbf{v}_3 = (2,4,6))$. Hỏi có tồn tại duy nhất hay không một toán tử tuyến tính $f$ trên $\mathbb{R}^3$ sao cho $f(\mathbf{u}_i) = \mathbf{v}_i$ với $i = 1,2,3$. Nếu có, hãy xác định $f$.

*Giải.* Rõ ràng $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ là một cơ sở của $\mathbb{R}^3$ vì $\begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix} = 2 \neq 0.$ Do đó, theo Định lý 1.3.1, có duy nhất một toán tử tuyến tính $f$ thỏa mãn yêu cầu $f(\mathbf{u}_i) = \mathbf{v}_i$ với $i = 1,2,3$.

Giả sử $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ là một vector tùy ý và giả sử $\mathbf{x} = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \lambda_3\mathbf{u}_3$. Khi đó

$$\begin{cases} \lambda_1 + \lambda_3 = x_1 \\ \lambda_1 + \lambda_2 = x_2 \\ \lambda_2 + \lambda_3 = x_3 \end{cases} \qquad \text{hay} \qquad \begin{cases} \lambda_1 = \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 \\ \lambda_2 = -\frac{1}{2}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 \\ \lambda_3 = \frac{1}{2}x_1 - \frac{1}{2}x_2 + \frac{1}{2}x_3 \end{cases}$$

Suy ra $f(\mathbf{x}) = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v}_3 = (2x_3, 3x_1 - x_2 + x_3, 3x_1 - x_2 + 3x_3)$. □

## Bài tập

**1)** Trong các ánh xạ sau đây, xét xem ánh xạ nào là ánh xạ tuyến tính

**(a)** $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x_1, x_2, x_3) = (2x_1 - x_2, x_3 - x_2, x_1)$

**(b)** $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x_1, x_2, x_3) = (x_1, x_2 + 3, x_3 - x_1)$

**(c)** $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x_1, x_2, x_3) = (x_2 + x_3, 2x_3 + x_1, 2x_1 + x_2)$

**(d)** $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x_1, x_2, x_3) = (x_1^2, x_2^2, x_3^2)$

**(e)** $f : \mathbb{R}^3 \to \mathbb{R}^3)$, $f(x_1, x_2, x_3) = (x_1, x_2, 4)$

**(f)** $f : \mathbb{R}^3 \to \mathbb{R}^2$, $f(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3)$

**(g)** $f : \mathbb{R}^3 \to \mathbb{R}^4$, $f(x_1, x_2, x_3) = (x_1 + x_2, x_2 - x_3, x_3 - x_1, x_1 - x_2)$

**(h)** $f : \mathbb{R}_n[x] \to \mathbb{R}_{n+1}[x]$, $f(p(x)) = x^{n+1} + p(x)$

**(i)** $f : \mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K})$, $f(M) = AM - MA$, trong đó $A$ là ma trận cố định thuộc $\mathcal{M}_n(\mathbb{K})$.

**2 )** Chứng minh rằng các ánh xạ sau là phép biến đổi tuyến tính trên $\mathbb{R}^3$

**(a)** $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x_1, x_2, x_3) = (x_1 + x_2 + x_3, -x_1 + 2x_2 + x_3, x_1 + 3x_2 + 2x_3)$.

**(b)** $g : \mathbb{R}^3 \to \mathbb{R}^3$, $g(x_1, x_2, x_3) = (-x_1 + 2x_2 + x_3, x_3, -x_2 + x_3)$.

**3 )** Cho các ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ trong đó $f(x_1, x_2, x_3) = (x_1 + x_2 + x_3, -x_1 + 2x_2 + x_3, x_1 + 3x_2 + 2x_3)$ và $g : \mathbb{R}^3 \to \mathbb{R}^3$ trong đó $g(x_1, x_2, x_3) = (-x_1 + 2x_2 + x_3, x_3, -x_2 + x_3)$.

**(a)** Tìm các ánh xạ $f + g$, $f \circ g$, và $g \circ f$.

**(b)** $f$ và $g$ có là song ánh không? Nếu có tìm ánh xạ ngược.

**4 )** Cho $V$ là một $\mathbb{K}$-không gian vector. Chứng minh rằng với mỗi $\lambda \in \mathbb{K}$ ánh xạ $V \to V$ xác định bởi $\mathbf{x} \mapsto \lambda\mathbf{x}$ là một toán tử tuyến tính trên $V$.

**5 )** Với $C[a, b]$ không gian các hàm thực liên tục trên $[a, b]$, chứng minh rằng phép lấy tích phân xác định

$$\begin{aligned} C[a,b] &\to \mathbb{R} \\ f(x) &\mapsto \int_a^b f(x)dx \end{aligned}$$

là một ánh xạ tuyến tính.

**6)** Chứng minh tính chất (iii) của Mệnh đề 1.2.1.

**7)** Cho ánh xạ tuyến tính $f : \mathbb{R}^2 \to \mathbb{R}^2$ xác định bởi $f(3,1) = (2,-4)$, $f(1,1) = (0,2)$. Xác định $f(x_1, x_2)$.

**8)** Cho ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi $f(1,2,3) = (1,0)$, $f(2,5,3) = (1,0)$, $f(1,0,10) = (0,1)$. Xác định $f(x_1, x_2, x_3)$.

**9)** Tìm ánh xạ tuyến tính $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ xác định bởi

$$f(1) = 1 + x, \qquad f(x) = 3 - x^2, \qquad f(x^2) = 4 + 2x - 3x^2.$$

# § 2 Ma trận và biểu thức tọa độ của ánh xạ tuyến tính

## 2.1 Ma trận của một ánh xạ tuyến tính

**2.1.1 Định nghĩa.** Cho $f : V \to U$ là một ánh xạ tuyến tính từ không gian vector $n$ chiều $V$ đến không gian vector $m$ chiều $U$. Giả sử $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ là cơ sở của $V$ và $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$ là cơ sở của $U$. Khi đó, mỗi vector $f(\mathbf{a}_j) \in U$ đều có dạng

$$f(\mathbf{a}_j) = a_{1j}\mathbf{b}_1 + a_{2j}\mathbf{b}_2 + \cdots + a_{mj}\mathbf{b}_m = \sum_{i=1}^{m} a_{ij}\mathbf{b}_i,$$

tức là $f(\mathbf{a}_j)|_{\mathscr{B}} = (a_{1j}, a_{2j}, \dots, a_{mj})$ với mọi $j = 1, \dots, n$. Do đó ta xác định được một ma trận $A = (a_{ij})_{m\times n}$. Ngược lại, do $f$ hoàn toàn xác định duy nhất bởi ảnh $(f(\mathbf{a}_1), f(\mathbf{a}_2), \dots, f(\mathbf{a}_n))$ của cơ sở $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ nên $f$ cũng hoàn toàn xác định nếu biết tất cả các hệ số $a_{ij}$, với $i = 1, \dots, m$ và $j = 1, \dots, n$; tức là $f$ được xác định duy nhất bởi ma trận $A = (a_{ij})_{m\times n}$. Ma trận $A = (a_{ij})_{m\times n}$ gọi là **ma trận của ánh xạ tuyến tính** $f$ trong cặp cơ sở $(\mathscr{A}, \mathscr{B})$. Ta kí hiệu $f|_{(\mathscr{A},\mathscr{B})} = A$ và biểu đồ thể hiện

$$(V, \mathscr{A}) \xrightarrow[A]{f} (U, \mathscr{B})$$

Ta viết lại ở dạng ma trận cột như sau

$$\begin{bmatrix} f(\mathbf{a}_1) & f(\mathbf{a}_2) & \dots & f(\mathbf{a}_n) \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_m \end{bmatrix} A.$$

Đặc biệt, khi $f$ là một toán tử tuyến tính trên $V$ (phép biến đổi tuyến tính trên $V$), $f : V \to V$, thì ma trận của $f$ trong cặp cơ sở $(\mathscr{A}, \mathscr{A})$ được gọi đơn giản là *ma trận của $f$ trong cơ sở $\mathscr{A}$*. Ma trận này là ma trận vuông cấp $n$ trên $\mathbb{K}$ lập nên từ các cột tọa độ theo thứ tự của các vector $f(\mathbf{a}_1), f(\mathbf{a}_2), \dots, f(\mathbf{a}_n)$ trong chính cơ sở $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$.

**2.1.2 Thí dụ.** Ma trận của ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi $f(x_1, x_2, x_3) = (x_2 - x_1, x_3 - x_2)$ đối với hai cơ sở chính tắc là $A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

**2.1.3 Thí dụ.** Ma trận của ánh xạ tuyến tính $\varphi : \mathbb{R}_n[x] \to \mathbb{R}_{n-1}[x]$ xác định bởi $\varphi(p(x)) = p'(x)$ đối với cơ sở $(1, x, \dots, x^n)$ và $(1, x, \dots, x^{n-1})$:

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n \end{bmatrix}. \qquad \square$$

**2.1.4 Thí dụ.** Ma trận của phép biến đổi tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (x_1 + x_2, 0, x_1 + x_3)$ đối với cơ sở chính tắc là $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. $\square$

## 2.2 Biểu thức tọa độ của ánh xạ tuyến tính

Cho $f : V \to U$ là một ánh xạ tuyến tính từ $\mathbb{K}$-không gian vector $n$ chiều $V$ đến $\mathbb{K}$-không gian vector $m$ chiều $U$, cho $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ là cơ sở của $V$ và $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$ là cơ sở của $U$. Gọi $A = (a_{ij})_{m\times n}$ là ma trận của $f$ đối với cặp cơ sở $(\mathscr{A}, \mathscr{B})$. Với mỗi $\mathbf{x} \in V$, ta muốn thiết lập một hệ thức liên hệ giữa các tọa độ của **x** trong $\mathscr{A}$ và tọa độ của ảnh $f(\mathbf{x}) \in U$ trong $\mathscr{B}$. Giả sử

$$\mathbf{x}|_{\mathscr{A}} = (x_1, x_2, \ldots, x_n) \qquad \text{và} \qquad f(\mathbf{x})|_{\mathscr{B}} = (y_1, y_2, \ldots, y_m),$$

tức là $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{a}_i$ và $f(\mathbf{x}) = \sum_{j=1}^{m} y_j \mathbf{b}_j$. Mặt khác, ta có

$$f(\mathbf{x}) = f\Big(\sum_{i=1}^{n} x_i \mathbf{a}_i\Big) = \sum_{i=1}^{n} x_i f(\mathbf{a}_i) = \sum_{i=1}^{n} x_i \sum_{j=1}^{m} a_{ji} \mathbf{b}_j = \sum_{j=1}^{m} \Big(\sum_{i=1}^{n} x_i a_{ji}\Big) \mathbf{b}_j.$$

Do đó, $y_j = \sum_{i=1}^{n} a_{ji} x_i$ với mọi $j = 1, \ldots, m$. Vậy ta có biểu thức ma trận sau

$$A\mathbf{x}|_{\mathscr{A}} = f(\mathbf{x})|_{\mathscr{B}} \tag{2.2.1}$$

trong công thức trên $\mathbf{x}|_{\mathscr{A}}$ và $f(\mathbf{x})|_{\mathscr{B}}$ được hiểu là ma trận cột của tọa độ của **x** và $f(\mathbf{x})$. Công thức (2.2.1) được gọi là **biểu thức tọa độ** của $f$ trong cặp cơ sở $(\mathscr{A}, \mathscr{B})$.

Đặc biệt, $U = V$, ta có phép biến đổi tuyến tính $f : V \to V$. Khi đó, $f(\mathbf{x})|_{\mathscr{A}} = A\mathbf{x}|_{\mathscr{A}}$ được viết gọn lại là

$$f(\mathbf{x}) = A\mathbf{x} \tag{2.2.2}$$

khi cơ sở $\mathscr{A}$ được ngầm hiểu và **x** và $f(\mathbf{x})$ được hiểu là các ma trận cột tọa độ đối với cơ sở $\mathscr{A}$.

**2.2.3 Thí dụ.** Biểu thức tọa độ của ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi $f(x_1, x_2, x_3) = (x_2 - x_1, x_3 - x_2)$ đối với hai cơ sở chính tắc là

$$f(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \qquad \square$$

**2.2.4 Thí dụ.** Biểu thức tọa độ của toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (x_1 + x_2, 0, x_1 + x_3)$ đối với cơ sở chính tắc là

$$f(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \qquad \square$$

**2.2.5 Định lý.** *Ma trận của hợp thành hai ánh xạ tuyến tính bằng tích các ma trận của chúng theo cùng thứ tự, tức là nếu $A$ và $B$ lần lượt là ma trận của các ánh xạ tuyến tính $f : V \to V'$ và $g : V' \to V''$ trong các cặp cơ sở $(\mathscr{B}, \mathscr{B}')$ và $(\mathscr{B}', \mathscr{B}'')$ tương ứng, thì ma trận của $g \circ f : V \to V''$ trong cặp cơ sở $(\mathscr{B}, \mathscr{B}'')$ chính là $BA$. Nghĩa là biểu đồ sau là giao hoán*

$$\begin{tikzcd}
(V,\mathscr{B}) \arrow[rr, "f", "A"'] \arrow[dr, "g\circ f", "BA"'] & & (V',\mathscr{B}') \arrow[dl, "B"', "g"] \\
& (V'',\mathscr{B}'') &
\end{tikzcd}$$

*Chứng minh.* Giả sử $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$, $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_m)$, $\mathscr{B}'' = (\mathbf{e}''_1, \mathbf{e}''_2, \ldots, \mathbf{e}''_p)$ và $A = (a_{ij})_{m\times n}$, $B = (b_{ki})_{p\times m}$. Khi đó ta có

$$f(\mathbf{e}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{e}'_i, \qquad g(\mathbf{e}'_j) = \sum_{k=1}^{p} b_{ki}\mathbf{e}''_k.$$

Do đó ta được

$$\begin{aligned} g\circ f(\mathbf{e}_j) = g(f(\mathbf{e}_j)) &= g\Big(\sum_{i=1}^{m} a_{ij}\mathbf{e}'_i\Big) = \sum_{i=1}^{m} a_{ij}g(\mathbf{e}'_i) \\ &= \sum_{i=1}^{m} a_{ij}\Big(\sum_{k=1}^{p} b_{ki}\mathbf{e}''_k\Big) = \sum_{k=1}^{p}\Big(\sum_{i=1}^{m} b_{ki}a_{ij}\Big)\mathbf{e}''_k \end{aligned}$$

Suy ra ma trận của $g \circ f$ trong cặp cơ sở $(\mathscr{B}, \mathscr{B}'')$ là $C = (c_{kj})_{p\times n}$ với $c_{kj} = \sum\limits_{i=1}^{m} b_{ki}a_{ij}$ trong đó $k = 1, \ldots, p$, $j = 1, \ldots, n$; tức là $C = BA$.

Ta có thể trình bày chứng minh ngắn gọn như sau

$$g\circ f(\mathbf{x})|_{\mathscr{B}''} = g(f(\mathbf{x}))|_{\mathscr{B}''} = Bf(\mathbf{x})|_{\mathscr{B}'} = B(A\mathbf{x}|_{\mathscr{B}}) = (BA)\mathbf{x}|_{\mathscr{B}}$$

Do đó ta phải có $C = BA$. □

**2.2.6 Thí dụ.** Cho hai ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^2$ và $g : \mathbb{R}^2 \to \mathbb{R}^3$ xác định như sau.

$$\begin{aligned} f(x, y, z) &= (2x + y - z, x + 2y + 3z) \quad &&\text{với mọi } (x, y, z) \in \mathbb{R}^3 \\ g(x', y') &= (x' - y', x' + 2y', x' + y') \quad &&\text{với mọi } (x', y') \in \mathbb{R}^2. \end{aligned}$$

Khi đó ma trận của $f$ và $g$ trong cặp cơ sở chính tắc lần lượt là

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 3 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$$

Do đó ma trận của ánh xạ hợp thành $g \circ f$ là

$$BA = \begin{bmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -4 \\ 4 & 5 & 5 \\ 3 & 3 & 2 \end{bmatrix}$$

Từ đó ánh xạ tuyến tính $g \circ f : \mathbb{R}^3 \to \mathbb{R}^3$ được xác định như sau:

$$g \circ f(x, y, z) = (x - y - 4z, 4x + 5y + 5z, 3x + 3y + 2z), \quad (x, y, z) \in \mathbb{R}^3.$$

□

## 2.3 Hai ma trận của ánh xạ tuyến tính trong hai cặp cơ sở khác nhau

**2.3.1 Định lý.** *Cho ánh xạ tuyến tính $f : V \to U$ có ma trận trong các cặp cơ sở $(\mathscr{A}, \mathscr{B})$ và $(\mathscr{A}_1, \mathscr{B}_1)$ lần lượt là $A$ và $A_1$, còn $B$ và $C$ là các ma trận đổi cơ sở từ $\mathscr{A}$ sang $\mathscr{A}_1$ và từ $\mathscr{B}$ sang $\mathscr{B}_1$ (một cách tương ứng). Khi đó,*

(2.3.2) $$A_1 = C^{-1}AB.$$

*Ta có biểu đồ sau:*

$$\begin{tikzcd} (V, \mathscr{A}) \arrow[r, "f", "A"'] & (U, \mathscr{B}) \\ (V, \mathscr{A}_1) \arrow[u, "B"] \arrow[r, "f", "A_1"'] & (U, \mathscr{B}_1) \arrow[u, "C"'] \end{tikzcd}$$

*Đặc biệt, khi $U = V$, nếu $A$ và $B$ lần lượt là ma trận của toán tử tuyến tính $f : V \to V$ trong hai cơ sở $\mathscr{A}$ và $\mathscr{B}$ và $C$ là ma trận đổi cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$, thì*

(2.3.3) $$B = C^{-1}AC.$$

*Ta có biểu đồ sau*

$$\begin{array}{ccc}
(V,\mathscr{A}) & \xrightarrow[A]{f} & (V,\mathscr{A}) \\
{\scriptstyle C}\uparrow & & {\scriptstyle C}\uparrow \\
(V,\mathscr{B}) & \xrightarrow[B]{f} & (V,\mathscr{B})
\end{array}$$

*Chứng minh.* Theo giả thiết của định lý ta có biểu thức tọa độ của ánh xạ tuyến tính $f$ đối với cặp cơ sở $(\mathscr{A},\mathscr{B})$ và cặp cơ sở $(\mathscr{A}_1,\mathscr{B}_1)$ lần lượt là

$$f(\mathbf{x})|_{\mathscr{B}} = A\mathbf{x}|_{\mathscr{A}} \qquad\qquad f(\mathbf{x})|_{\mathscr{B}_1} = A_1\mathbf{x}|_{\mathscr{A}_1}$$

Cũng theo giả thiết ta có công thức đổi tọa độ sau

$$\mathbf{x}|_{\mathscr{A}} = B\mathbf{x}|_{\mathscr{A}_1} \qquad\qquad f(\mathbf{x})|_{\mathscr{B}} = Cf(\mathbf{x})|_{\mathscr{B}_1}.$$

Từ đó ta tính được

$$f(\mathbf{x})|_{\mathscr{B}_1} = C^{-1}f(\mathbf{x})|_{\mathscr{B}} = C^{-1}(A\mathbf{x}|_{\mathscr{A}}) = (C^{-1}A)(B\mathbf{x}|_{\mathscr{A}_1}) = (C^{-1}AB)\mathbf{x}|_{\mathscr{A}_1}.$$

Do đó, ta có $A_1 = C^{-1}AB$. □

**2.3.4 Nhận xét.** Định lý trên cho ta thấy một ý nghĩa của quan hệ tương đương và đồng dạng trên các ma trận: các ma trận của cùng một ánh xạ tuyến tính (tương ứng, toán tử tuyến tính) đều tương đương (tương ứng, đồng dạng) với nhau. Hơn nữa, đều ngược lại cũng đúng: nếu $A \in \mathcal{M}_{m,n}(\mathbb{K})$ (tương ứng, $A \in \mathcal{M}_n(\mathbb{K})$) là ma trận của ánh xạ tuyến tính $f : V \to V'$ (tương ứng, toán tử tuyến tính $f : V \to V$) trong một cặp cơ sở (tương ứng, cơ sở) nào đó và $B \sim A$ (tương ứng, $B \backsim A$) thì $B$ cũng là ma trận của $f$ trong một cặp cơ sở (tương ứng, cơ sở) khác.

**2.3.5 Thí dụ.** Cho ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi

$$f(x,y,z) = (x+y-z, x-y+z) \qquad \text{với mọi } (x,y,z) \in \mathbb{R}^3$$

Tìm ma trận của $f$ trong cặp cơ sở $(\mathscr{B},\mathscr{B}')$ biết rằng $\mathscr{B} = (\mathbf{u}_1 = (1,1,0), \mathbf{u}_2 = (0,1,1), \mathbf{u}_2 = (1,0,1))$ và $\mathscr{B}' = (\mathbf{u}'_1 = (2,1), \mathbf{u}'_2 = (1,1))$.

*Giải.* Ma trận của $f$ trong cặp cơ sở chính tắc $(\mathscr{E}_3, \mathscr{E}_2)$ là

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

Ma trận đổi cơ sở từ $\mathscr{E}_3$ sang $\mathscr{B}$ là $B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. Ma trận đổi cơ sở từ $\mathscr{E}_2$ sang $\mathscr{B}'$ là $C = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$. Dễ dàng tính được $C^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$. Do đó ma trận của $f$ trong cặp cơ sở $(\mathscr{B}, \mathscr{B}')$ là

$$A_1 = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -2 \\ -2 & 0 & 4 \end{bmatrix}. \qquad \square$$

**2.3.6 Thí dụ.** Cho toán tử tuyến tính $f : \mathbb{R}^2 \to \mathbb{R}^2$ xác định như sau $f(x, y) = (x + 2y, 2x + y)$ với mọi $(x, y) \in \mathbb{R}^2$. Tìm ma trận của $f$ trong cơ sở $\mathscr{B} = (\mathbf{u}_1 = (2, 1), \mathbf{u}_2 = (3, 2))$.

*Giải.* Ma trận của $f$ trong cơ sở chính tắc là $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Ma trận đổi cơ sở từ $\mathscr{E}_2$ sang $\mathscr{B}$ là $C = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ với $C^{-1} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$. Vậy ma trận của $f$ trong cơ sở $\mathscr{B}$ là

$$B = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} -7 & -10 \\ 6 & 9 \end{bmatrix}. \qquad \square$$

**2.3.7 Thí dụ.** Tìm ma trận của toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (x_1 + x_2, 0, x_1 + x_3)$ đối với cơ sở $\mathscr{B} = ((1, 0, 0), (1, 1, 0), (1, 1, 1))$.

*Giải.* Ta biết ma trận của toán tử tuyến tính $f$ đối với cơ sở chính tắc là

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

Hơn nữa, ma trận chuyển cơ sở từ cơ sở chính tắc sang cơ sở $\mathscr{B}$ là

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

Do đó, theo công thức (2.3.3) ta được ma trận của toán tử tuyến tính $f$ đối với cơ sở $\mathscr{B}$ là

$$\begin{aligned} B &= C^{-1}AC \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 0 & 0 & 0 \\ 1 & 1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & 2 \\ -1 & -1 & -2 \\ 1 & 1 & 2 \end{bmatrix}. \end{aligned}$$

□

## Bài tập

**1)** Tìm công thức tọa độ của các ánh xạ tuyến tính $f$ được xác định như sau:

**(a)** $f : \mathbb{R}^3 \to \mathbb{R}^3$ với $f(1,1,2) = (1,0,0)$, $f(2,1,1) = (0,1,1)$, $f(2,2,3) = (0,-1,0)$.

**(b)** $f : \mathbb{R}^3 \to \mathbb{R}^3$ với $f(1,2,3) = (-1,0,1)$, $f(-1,1,1) = (0,1,0)$, $f(1,3,4) = (1,0,2)$.

**(c)** $f : \mathbb{R}^3 \to \mathbb{R}^2$ với $f(1,0,1) = (1,0)$, $f(0,1,1) = (1,-1)$, $f(1,1,0) = (0,-1)$.

**2)** Cho ánh xạ tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3, x_4) = (x_1 - x_2 + x_3, 2x_1 + x_4, 2x_2 + x_3 - x_4)$. Tìm ma trận của $f$ trong cặp cơ sở $(\mathscr{B}, \mathscr{B}')$ với $\mathscr{B} = ((1,1,1,1), (0,1,1,1), (0,0,1,1), (0,0,0,1))$ và $\mathscr{B}' = ((1,1,1), (1,1,0), (1,0,0))$.

**3)** Cho toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (2x_2 + x_3, x_1 - 4x_2, 3x_1)$. Tìm ma trận của $f$ đối với cơ sở $\mathscr{B} = ((1,1,1), (1,1,0), (1,0,0))$.

**4)** Toán tử tuyến tính $f$ trên $\mathbb{R}^3$ có ma trận trong cơ sở chính tắc $\mathscr{E}_3 = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ là $\begin{bmatrix} 0 & -2 & 1 \\ 3 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}$. Tìm ma trận của $f$ trong cơ sở $\mathscr{B} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, trong đó $\mathbf{u}_1 = 3\mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3$, $\mathbf{u}_2 = 2\mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3$, $\mathbf{u}_3 = -\mathbf{e}_1 + 2\mathbf{e}_2 + 5\mathbf{e}_3$.

**5)** Cho toán tử tuyến tính $f$ trên $\mathbb{R}^4$ có ma trận trong cơ sở chính tắc là $\begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 5 & 11 & 2 \\ 0 & -1 & 3 & 1 \\ 1 & 2 & 1 & 3 \end{bmatrix}$. Tìm ma trận của $f$ trong cơ sở $\mathscr{B} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4)$, trong đó $\mathbf{u}_1 = \mathbf{e}_1$, $\mathbf{u}_2 = \mathbf{e}_1 + \mathbf{e}_2$, $\mathbf{u}_3 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{u}_4 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4$.

**6)** Trong $\mathbb{R}^3$ cho hai cơ sở $\mathscr{A} = (\mathbf{a}_1 = (1,0,0), \mathbf{a}_2 = (0,1,1), \mathbf{a}_3 = (1,0,1))$ và $\mathscr{B} = (\mathbf{b}_1 = (1,-1,0), \mathbf{b}_2 = (0,1,-1), \mathbf{b}_3 = (1,0,1))$ và ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ sao cho $f(\mathbf{a}_i) = \mathbf{b}_i$ với $i = 1, 2, 3$.

**(a)** Tìm công thức tọa độ của ánh xạ tuyến tính $f$ trong cơ sở chính tắc.

**(b)** Tìm ma trận của $f$ trong các cơ sở hay cặp cơ sở sau

| | | |
|---|---|---|
| **(i)** $\mathscr{A}$ | **(iii)** $\mathscr{B}$ | **(v)** $(\mathscr{E}, \mathscr{A})$ |
| **(ii)** $(\mathscr{A}, \mathscr{B})$ | **(iv)** $(\mathscr{B}, \mathscr{A})$ | **(vi)** $(\mathscr{E}, \mathscr{B})$ |

**7)** Cho ánh xạ tuyến tính $f : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ xác định bởi $f(p(x)) = p'(x)$. Tìm ma trận của $f$ trong cơ sở

**(a)** $(\mathbf{u}_0 = 1, \mathbf{u}_1 = x, \mathbf{u}_2 = x^2, \dots, \mathbf{u}_n = x^n)$

**(b)** $\left(\mathbf{v}_0 = 1, \mathbf{v}_1 = x - a, \mathbf{v}_2 = \frac{(x-a)^2}{2!}, \dots, \mathbf{v}_n = \frac{(x-a)^n}{n!}\right)$ với $a \in \mathbb{R}$.

**8)** Trong không gian $\mathcal{M}_2(\mathbb{R})$ cho ma trận $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ và các ánh xạ

**(a)** $f : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ xác định bởi $f(M) = AM$

**(b)** $f : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ xác định bởi $f(M) = MA$

Chứng minh rằng $f$ và $g$ là các toán tử tuyến tính trên $\mathcal{M}_2(\mathbb{R})$. Tìm ma trận của $f$ và $g$ trong cơ sở sau

$$\left(\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right).$$

**9)** Ma trận của một toán tử tuyến tính trong cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ thay đổi thế nào đổi chỗ hai vector $\mathbf{e}_i$ và $\mathbf{e}_j$?

**10)** Giả sử $A$ và $B$ lần lượt là ma trận của toán tử tuyến tính $f : V \to V$ trong hai cơ sở $\mathscr{A}$ và $\mathscr{B}$. Chứng minh rằng $\operatorname{tr}(A) = \operatorname{tr}(B)$.

# § 3 Hạt nhân - ảnh

## 3.1 Ảnh và ảnh ngược của một không gian con

**3.1.1 Định lý.** *Cho ánh xạ tuyến tính $f : V \to U$. Nếu $W$ là không gian con của $V$ thì $f(W)$ là không gian con của $U$. Hơn nữa, nếu $W = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ thì $f(W) = \langle f(\mathbf{a}_1), \dots, f(\mathbf{a}_n) \rangle$*

*Chứng minh.* Ta thấy $\mathbf{0} \in f(W)$ vì $\mathbf{0} \in W$ và $f(\mathbf{0}) = \mathbf{0}$. Với mọi $\mathbf{x}, \mathbf{y} \in f(W)$ và $\alpha, \beta \in \mathbb{K}$, tồn tại $\mathbf{a}, \mathbf{b} \in W$ sao cho $f(\mathbf{a}) = \mathbf{x}$ và $f(\mathbf{b}) = \mathbf{y}$. Do $W$ là không gian con của $V$ nên $\alpha\mathbf{a} + \beta\mathbf{b} \in W$, suy ra $f(\alpha\mathbf{a} + \beta\mathbf{b}) \in f(W)$. Mặt khác, ta có $f(\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha f(\mathbf{a}) + \beta f(\mathbf{b})$. Vậy $\alpha\mathbf{x} + \beta\mathbf{y} \in f(W)$. Do đó, $f(W)$ là không gian con của $U$.

Giả sử $W = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$. Lấy $\mathbf{x} \in f(W)$ tùy ý. Khi đó, tồn tại $\mathbf{a} \in W$ sao cho $f(\mathbf{a}) = \mathbf{x}$. Mặt khác, tồn tại $\lambda_1, \dots, \lambda_n$ sao cho $\mathbf{a} = \sum_{i=1}^n \lambda_i \mathbf{a}_i$. Do đó,

$$\mathbf{x} = f\Big(\sum_{i=1}^n \lambda_i \mathbf{a}_i\Big) = \sum_{i=1}^n \lambda_i f(\mathbf{a}_i).$$

Từ đó ta suy ra mọi vector trong $f(W)$ đều biểu thị tuyến tính được qua hệ vector $(f(\mathbf{a}_1), \dots, f(\mathbf{a}_n))$. Do đó $f(W) = \langle f(\mathbf{a}_1), \dots, f(\mathbf{a}_n) \rangle$. $\square$

**3.1.2 Định lý.** *Cho ánh xạ tuyến tính $f : V \to U$. Nếu $W$ là không gian con hữu hạn chiều của $V$, thì $f(W)$ là không gian con hữu hạn chiều của $U$ và $\dim f(W) \leq \dim W$.*

*Chứng minh.* Do $W$ là không gian con hữu hạn chiều của $V$, nên $\dim W = k$. Nếu $k = 0$, thì $W = \{\mathbf{0}\}$, suy ra $f(W) = \{\mathbf{0}\}$, cho nên $\dim f(W) = 0$. Nếu $k > 0$, thì trong $W$ có một cơ sở $\mathscr{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_k)$. Ta có, $W = \langle \mathscr{A} \rangle$ suy ra (theo Định lý 3.1.1) $f(W) = \langle f(\mathbf{a}_1), \ldots, f(\mathbf{a}_k) \rangle$. Do đó, $\dim f(W) \leq k$. □

**3.1.3 Hệ quả.** *Cho $A$ là một ma trận $m \times n$ và $B$ là một ma trận $n \times p$. Khi đó, ta có*

(3.1.4) $$\operatorname{rank}(AB) \leq \min\{\operatorname{rank} A, \operatorname{rank} B\}.$$

*Chứng minh.* Xét ánh xạ tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^m$ xác định bởi ma trận $A$ (đối với cơ sở chính tắc). Gọi $\mathbf{b}_j$ là vector trong $\mathbb{R}^n$ ứng với cột $j$ của ma trận $B$. Khi đó $f(\mathbf{b}_j)$ ứng với cột $j$ của ma trận $AB$. Theo định lý trên ta có $\dim\langle f(\mathbf{b}_1), \ldots, f(\mathbf{b}_p)\rangle \leq \dim\langle \mathbf{b}_1, \ldots, \mathbf{b}_p\rangle$, cho nên $\operatorname{rank} AB \leq \operatorname{rank} B$. Mặt khác, ta có $\operatorname{rank} AB = \operatorname{rank}(AB)^t = \operatorname{rank} B^t A^t \leq \operatorname{rank} A^t = \operatorname{rank} A$. Như vậy ta được $\operatorname{rank}(AB) \leq \min\{\operatorname{rank} A, \operatorname{rank} B\}$. □

**3.1.5 Hệ quả.** *Cho $A$ là một ma trận cấp $m \times n$, $B$ là ma trận vuông cấp $n$ khả nghịch và $C$ là ma trận vuông cấp $m$ khả nghịch. Khi đó, ta có*

(3.1.6) $$\operatorname{rank} A = \operatorname{rank}(AB) = \operatorname{rank}(CA).$$

*Chứng minh.* Theo hệ quả trên ta có

$$\operatorname{rank} A \geq \operatorname{rank}(AB) \geq \operatorname{rank}(ABB^{-1}) = \operatorname{rank}(A).$$

Vậy $\operatorname{rank} A = \operatorname{rank}(AB)$. Tương tự ta cũng chứng minh được $\operatorname{rank} A = \operatorname{rank}(CA)$. □

**3.1.7 Thí dụ. (OSV02)** Cho $P$ và $Q$ là hai ma trận vuông cấp $n$ thỏa mãn các điều kiện sau:

$$P^2 = P, \qquad Q^2 = Q$$

và $I-(P+Q)$ là một ma trận khả nghịch ($I$ là ma trận đơn vị cấp $n$). Chứng minh rằng hạng của $P$ và $Q$ bằng nhau.

*Giải.* Từ giả thiết và hệ quả trên ta có

$$\operatorname{rank} P = \operatorname{rank}(P(I-(P+Q))) = \operatorname{rank}(P-P^2-PQ) = \operatorname{rank}(PQ)$$
$$\operatorname{rank} Q = \operatorname{rank}((I-(P+Q))Q) = \operatorname{rank}(Q-PQ-Q^2) = \operatorname{rank}(PQ)$$

Do đó ta được $\operatorname{rank} P = \operatorname{rank} Q$. □

**3.1.8 Thí dụ. (OSV99)** Cho $A$ là ma trận có 1999 dòng và 2000 cột. $A^t$ là ma trận chuyển vị của $A$ và $B$ là ma trận phụ hợp của ma trận $A^tA$. Biết rằng $\det(AA^t) \neq 0$ và $B \neq O$. Xác định hạng của ma trận $B$.

*Giải.* Do $\det(AA^t) \neq 0$ nên $\operatorname{rank}(AA^t) = 1999$. Mặt khác ta có $1999 = \operatorname{rank}(AA^t) \leq \operatorname{rank} A \leq 1999$. Vậy $\operatorname{rank} A = 1999$. Suy ra $\operatorname{rank}(A^tA) \leq \operatorname{rank} A = 1999$. Ta có $A^tA$ là ma trận vuông cấp 2000 nếu $\operatorname{rank}(A^tA) < 1999$ nên mọi định thức của ma trận con vuông cấp 1999 đều bằng không; điều này kéo theo mâu thuẫn với giả thiết $B \neq O$. Vậy $\operatorname{rank}(A^tA) = 1999$, cho nên $\det(A^tA) = 0$. Gọi $c_{ij}$ là các phần tử của $A^tA$ và $C_{ij}$ là phần bù đại số của $c_{ij}$. Do đó, ta có

$$B = \begin{bmatrix} C_{11} & C_{21} & \dots & C_{2000,1} \\ C_{12} & C_{22} & \dots & C_{2000,2} \\ \dots & \dots & \dots & \dots \\ C_{1,2000} & C_{2,2000} & \dots & C_{2000,2000} \end{bmatrix}$$

do $\det(A^tA) = 0$ nên $c_{i1}C_{j1} + c_{i2}C_{j2} + \cdots + c_{i,2000}C_{j,2000} = 0$ với mọi $i$ và $j$. Vậy các cột của ma trận $B$ là nghiệm của hệ phương trình tuyến tính thuần nhất $(A^tA)\mathbf{x} = \mathbf{0}$ với 2000 ẩn số và ma trận hệ số có hạng bằng 1999. Gọi $S$ là không gian nghiệm của hệ $A^tA\mathbf{x} = \mathbf{0}$ Do đó ta có

$$1 \leq \operatorname{rank}(B) \leq \dim(S) = 2000 - \operatorname{rank}(A^tA) = 1.$$

Vậy $\operatorname{rank}(B) = 1$. □

**3.1.9 Định lý.** *Cho ánh xạ tuyến tính $f : V \to U$. Nếu $W'$ là một không gian con của $U$, thì $f^{-1}(W')$ là một không gian con của $V$.*

*Chứng minh.* Ta thấy $\mathbf{0} \in f^{-1}(W')$ vì $f(\mathbf{0}) = \mathbf{0} \in W'$. Lấy $\mathbf{a}, \mathbf{b} \in f^{-1}(W')$ tùy ý và $\alpha, \beta \in \mathbb{K}$ tùy ý. Khi đó, $f(\mathbf{a}), f(\mathbf{b}) \in W'$. Do $W'$ là một không gian con nên $f(\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha f(\mathbf{a}) + \beta f(\mathbf{b}) \in W'$, suy ra $\alpha\mathbf{a} + \beta\mathbf{b} \in f^{-1}(W')$. Vậy $f^{-1}(W')$ là một không gian con của $V$. $\square$

## 3.2 Hạt nhân và ảnh của ánh xạ tuyến tính

Cho ánh xạ tuyến tính $f : V \to U$. Theo Định lý 3.1.1 ta có $f(V) = \{f(\mathbf{x}) : \mathbf{x} \in V\}$ là không gian con của $U$; và theo Định lý 3.1.9 ta có $f^{-1}(\mathbf{0}) = \{\mathbf{x} \in V : f(\mathbf{x}) = \mathbf{0}\}$ là không gian con của $V$.

**3.2.1 Định nghĩa.** Cho ánh xạ tuyến tính $f : V \to U$. Khi đó, không gian con $\{\mathbf{x} \in V : f(\mathbf{x}) = \mathbf{0}\}$ của $V$ được gọi là **hạt nhân** (kernel, null space) của $f$, kí hiệu $\ker f$, và không gian con $f(V) = \{f(\mathbf{x}) : \mathbf{x} \in V\}$ được gọi là **ảnh** (image of $V$) của $f$ (range of $f$), kí hiệu $\mathrm{Im} f$.

**3.2.2 Thí dụ.** Xét toán tử tuyến tính trên $\mathbb{R}^3$ như sau $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x, y, z) = (x, y, 0)$. Khi đó, dễ thấy rằng

$$\mathrm{Im} f = \{(x, y, 0) : x, y \in \mathbb{R}\} \qquad \mathrm{rank}(f) = \dim(\mathrm{Im}\, f) = 2$$

và

$$\ker f = \{(0, 0, z) : z \in \mathbb{R}\} \qquad \mathrm{def}(f) = \dim(\ker f) = 1.$$

Ta nhận thấy $\mathrm{rank}(f) + \dim(\ker f) = 3 = \dim(\mathbb{R}^3)$. Điều này không phải là trường hợp riêng mà nó đúng cho mọi không gian hữu hạn chiều, nó được chứng minh trong định lý sau. $\square$

**3.2.3 Định lý.** *Cho ánh xạ tuyến tính $f : V \to U$. Nếu $V$ là không gian hữu hạn chiều, thì $\ker f$ và $\mathrm{Im} f$ cũng là các không gian hữu hạn chiều và $\dim(\mathrm{Im} f) + \dim(\ker f) = \dim V$. Khi đó, $\dim(\mathrm{Im} f)$ được gọi là **hạng** của ánh xạ tuyến tính $f$, kí hiệu $\mathrm{rank}(f)$, và $\dim(\ker f)$ được gọi là **số khuyết** của $f$, kí hiệu $\mathrm{def}(f)$.*

*Chứng minh.* Vì $\ker f$ là không gian con của $V$, nên $\ker f$ là không gian hữu hạn chiều. Giả sử $\dim(\ker f) = r$ và $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r)$ là cơ sở của

$\ker f$. Ta có $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r)$ là họ độc lập tuyến tính của $V$, cho nên ta bổ sung thêm các vector để được $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r, \mathbf{a}_{r+1}, \dots, \mathbf{a}_n)$ là cơ sở của $V$ (với $\dim V = n$). Ta sẽ chứng minh $(f(\mathbf{a}_{r+1}), \dots, f(\mathbf{a}_n))$ là cơ sở của $\mathrm{Im} f$ (khi đó, $\dim(\mathrm{Im} f) = n - r$).

Với mọi $\mathbf{y} \in \mathrm{Im} f$, khi đó tồn tại $\mathbf{x} \in V$ để $f(\mathbf{x}) = \mathbf{y}$. Do $\mathbf{x} \in V$ nên ta có thể viết $\mathbf{x} = \lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_r\mathbf{a}_r + \lambda_{r+1}\mathbf{a}_{r+1} + \cdots + \lambda_n\mathbf{a}_n$, suy ra

$$\begin{aligned}\mathbf{y} &= \lambda_1 f(\mathbf{a}_1) + \lambda_2 f(\mathbf{a}_2) + \cdots + \lambda_r f(\mathbf{a}_r) + \lambda_{r+1} f(\mathbf{a}_{r+1}) + \cdots + a_n f(\mathbf{a}_n)\\ &= \lambda_{r+1} f(\mathbf{a}_{r+1}) + \cdots + \lambda_n f(\mathbf{a}_n)\end{aligned}$$

(vì $f(\mathbf{a}_i) = \mathbf{0}$ với mọi $i = 1, \dots, r$). Suy ra $(f(\mathbf{a}_{r+1}), \dots, f(\mathbf{a}_n))$ là hệ sinh của $\mathrm{Im} f$.

Giả sử $\lambda_{r+1} f(\mathbf{a}_{r+1}) + \cdots + \lambda_n f(\mathbf{a}_n) = \mathbf{0}$. Khi đó, ta có $f(\lambda_{r+1}\mathbf{a}_{r+1} + \cdots + \lambda_n\mathbf{a}_n) = \mathbf{0}$. Suy ra $\lambda_{r+1}\mathbf{a}_{r+1} + \cdots + \lambda_n\mathbf{a}_n \in \ker f$. Vì $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r)$ là cơ sở của $\ker f$ nên ta có thể viết

$$\lambda_{r+1}\mathbf{a}_{r+1} + \cdots + \lambda_n\mathbf{a}_n = \lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_r\mathbf{a}_r.$$

hay

$$-\lambda_1\mathbf{a}_1 - \lambda_2\mathbf{a}_2 - \cdots - \lambda_r\mathbf{a}_r + \lambda_{r+1}\mathbf{a}_{r+1} + \cdots + \lambda_n\mathbf{a}_n = \mathbf{0}.$$

Từ đó suy ra $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$ vì $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ là cơ sở của $V$; đặc biệt, $\lambda_{r+1} = \cdots = \lambda_n = 0$. Vậy hệ $(f(\mathbf{a}_{r+1}), \dots, f(\mathbf{a}_n))$ độc lập tuyến tính.

Do đó, $(f(\mathbf{a}_{r+1}), \dots, f(\mathbf{a}_n))$ là cơ sở của $\mathrm{Im} f$ và

$$\dim(\mathrm{Im} f) + \dim(\ker f) = \dim V. \qquad \square$$

**3.2.4 Mệnh đề.** *Nếu* $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ *là hệ sinh của* $V$ *và* $f : V \to U$ *là ánh xạ tuyến tính, thì* $(f(\mathbf{a}_1), \dots, f(\mathbf{a}_m))$ *là hệ sinh của* $\mathrm{Im} f$.

*Chứng minh.* Kết quả được suy ra trực tiếp từ Định lý 3.1.1. $\square$

**3.2.5 Thí dụ.** Xét phép lấy đạo hàm trên không gian $\mathbb{R}_n[x]$ các đa thức bậc không quá $n$ ($n \geq 1$)

$$\begin{aligned} d : \mathbb{R}_n[x] &\to \mathbb{R}_n[x]\\ p(x) &\mapsto p'(x)\end{aligned}$$

Ta biết rằng $\dim(\mathbb{R}_n[x]) = n+1$ với một cơ sở là $\mathscr{E} = (1, x, x^2, \dots, x^n)$. Ta cũng dễ dàng thấy rằng $\mathrm{Im}(d) = \mathbb{R}_{n-1}[x]$, $\mathrm{rank}(d) = \dim(\mathbb{R}_{n-1}[x]) = n$; $\ker(d) = \{p(x) = C : C \in \mathbb{R}\}$ (tập các đa thức hằng), $\mathrm{def}(d) = \dim(\ker(d)) = 1$. □

**3.2.6 Mệnh đề.** *Cho ánh xạ tuyến tính* $f : V \to U$, $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ *là cơ sở của* $V$, *và* $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$ *là cơ sở của* $U$. *Khi đó,* $\mathrm{rank}\, f = \mathrm{rank}\, A$, *trong đó* $A$ *là ma trận của* $f$ *đối với cặp cơ sở* $(\mathscr{A}, \mathscr{B})$.

*Chứng minh.* Vì $\mathscr{A}$ là cơ sở của $V$ nên nó cũng là hệ sinh, theo mệnh đề trên hệ vector $(f(\mathbf{a}_1), f(\mathbf{a}_2), \dots, f(\mathbf{a}_n))$ là hệ sinh của $\mathrm{Im} f$. Do đó, $\mathrm{rank}\, f = \dim(\mathrm{Im} f) = \mathrm{rank}(f(\mathbf{a}_1), f(\mathbf{a}_2), \dots, f(\mathbf{a}_n))$. Ta có $A$ là ma trận của $f$ đối với cặp cơ sở $(\mathscr{A}, \mathscr{B})$ nên $A$ chính là ma trận tọa độ của hệ $(f(\mathbf{a}_1), f(\mathbf{a}_2), \dots, f(\mathbf{a}_n))$ đối với cơ sở $\mathscr{B}$. Vậy $\mathrm{rank}\, f = \mathrm{rank}\, A$ (Mệnh đề 4.2.6, trang 235). □

**3.2.7 Thí dụ. (OSV07)** Cho $A$ là ma trận vuông cấp $n$ thỏa mãn các điều kiện $A^{2006} = A$ và $\mathrm{rank}\, A = 1$. Chứng minh rằng $I - A$ là ma trận suy biến.

*Giải.* Xét ánh xạ tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$ xác định bởi $f(\mathbf{x}) = A\mathbf{x}$. Ta có $\mathrm{rank}\, f = \mathrm{rank}\, A = 1$. Vậy tồn tại $\mathbf{a} \neq \mathbf{0}$ sao cho $f(\mathbb{R}^n) = \langle \mathbf{a} \rangle$. Vì vậy tồn tại $\lambda$ sao cho $A\mathbf{a} = \lambda\mathbf{a}$. Từ giả thiết ta biến đổi được

$$\lambda\mathbf{a} = A\mathbf{a} = A^{2006}\mathbf{a} = \lambda A^{2005}\mathbf{a} = \dots = \lambda^{2006}\mathbf{a}.$$

Từ đó suy ra $(\lambda - \lambda^{2006})\mathbf{a} = \mathbf{0}$. Vì $\mathbf{a} \neq \mathbf{0}$ nên $\lambda - \lambda^{2006} = 0$ suy ra $\lambda = 0$ hay $\lambda = 1$. Nếu $\lambda = 0$ và do $f(\mathbb{R}^n) = \langle \mathbf{a} \rangle$ nên suy ra

$$A^2\mathbf{x} = Af(\mathbf{x}) = \lambda_x A\mathbf{a} = \mathbf{0} \qquad \text{với mọi } \mathbf{x} \in \mathbb{R}^n.$$

Vậy $A^2 = O$, suy ra $A = A^{2006} = O$, mâu thuẫn với $\mathrm{rank}\, A = 1$. Do đó, ta phải có $\lambda = 1$. Khi đó, $A\mathbf{a} = \mathbf{a}$ suy ra $(A - I)\mathbf{a} = \mathbf{0}$. Vậy phương trình $(A - I)\mathbf{x} = \mathbf{0}$ có nghiệm không tầm thường, cho nên $A - I$ suy biến. □

**3.2.8 Định lý.** *Cho $A$ là một ma trận $m \times n$ và $B$ là một ma trận $n \times p$. Khi đó ta có*

$$\text{(3.2.9)} \qquad \operatorname{rank} A + \operatorname{rank} B \leq n + \operatorname{rank}(AB).$$

*Chứng minh.* Gọi $f : \mathbb{R}^n \to \mathbb{R}^m$ là ánh xạ tuyến tính xác định bởi ma trận $A$. Khi đó ta có $n = \dim(\operatorname{Im} f) + \dim(\ker f) = \operatorname{rank} A + \dim(\ker f)$. Gọi $\mathbf{b}_i \in \mathbb{R}^n$ là vector ứng với cột thứ $i$ của ma trận $B$, và đặt $U = \langle \mathbf{b}_1, \ldots, \mathbf{b}_p \rangle$. Xét ánh xạ tuyến tính thu hẹp $f|_U : U \to \mathbb{R}^m$. Rõ ràng ta có $\ker(f|_U) \subseteq \ker f$. Theo Định lý 3.2.3 ta có $\dim U = \operatorname{rank}(f|_U) + \dim(\ker f|_U)$. Do đó ta được

$$\begin{aligned} \operatorname{rank} B &= \operatorname{rank}(A\mathbf{b}_1, \ldots, A\mathbf{b}_p) + \dim(\ker f|_U) \\ &\leq \operatorname{rank} AB + \dim(\ker f) \\ &\leq \operatorname{rank} AB + n - \operatorname{rank} A \end{aligned}$$

Do đó ta được kết quả quan trọng về hạng ma trận $\operatorname{rank} A + \operatorname{rank} B \leq n + \operatorname{rank} AB$. □

**3.2.10 Thí dụ. (OSV94)** Chứng minh rằng nếu $A$ là ma trận vuông cấp $n$ với $A^2 = I$ thì $\operatorname{rank}(A + I) + \operatorname{rank}(I - A) = n$.

*Giải.* Sử dụng tính chất hạng của tích ma trận và hạng của tổng ma trận ta có

$$\begin{aligned} n &= n + \operatorname{rank}(I - A^2) \\ &= n + \operatorname{rank}(A + I)(I - A) \\ &\geq \operatorname{rank}(A + I) + \operatorname{rank}(I - A) \\ &\geq \operatorname{rank}(2I) \\ &= n \end{aligned}$$

Do đó ta phải có $\operatorname{rank}(A + I) + \operatorname{rank}(I - A) = n$. □

**3.2.11 Thí dụ. (OSV09)** Cho $A, B, C$ là các ma trận vuông cấp $n$ sao cho $C$ giao hoán với $A$ và $B$, $C^2 = I$ và $AB = 2(A + B)C$.

**(a)** Chứng minh $AB = BA$.

**(b)** Nếu có thêm điều kiện $A+B+C=0$ hãy chứng tỏ $\operatorname{rank}(A-C)+\operatorname{rank}(B-C)=n$.

*Giải.* **(a)** Từ giả thiết $AB = 2(A+B)C$ ta viết lại

$$AB - 2AC - 2BC = A(B-2C) - 2CB = O$$

tương đương

$$A(B-2C) - 2C(B-2C) - 4C^2 = 0$$

hay

$$(A-2C)(B-2C) = 4I.$$

Từ đó dễ dàng suy ra được $A-2C$ và $B-2C$ có tích giao hoán. Từ $(A-2C)(B-2C) = (B-2C)(A-2C)$ ta biến đổi được

$$AB - 2AC - 2CB + 4C^2 = BA - 2BC - 2CA + 4C^2.$$

Mặt khác, do $C$ giáo hoán với $A$ và $B$ nên suy ra được $AB = BA$.

**(b)** Do $A+B+C=O$ và hạng của tổng hai ma trận ta có

$$\begin{aligned}\operatorname{rank}(A-C)+\operatorname{rank}(B-C) &\geq \operatorname{rank}(A-C+B-C)\\ &= \operatorname{rank}(-3C)\\ &\geq n.\end{aligned}$$

Mặt khác, từ bất đẳng thức giữa hạng của tích hai ma trận và hạng của các ma trận thành phần ta có

$$\begin{aligned}&\operatorname{rank}(A-C)+\operatorname{rank}(B-C)\\ &\qquad\leq n + \operatorname{rank}(A-C)(B-C)\\ &\qquad\leq n + \operatorname{rank}(AB - AC - CB + C^2)\\ &\qquad= n + \operatorname{rank}(2(A+B)C - AC - BC + C^2)\\ &\qquad= n + \operatorname{rank}((A+B+C)C)\\ &\qquad= n + \operatorname{rank} O\\ &\qquad= n.\end{aligned}$$

Từ các kết quả thu được ta suy ra $\operatorname{rank}(A-C)+\operatorname{rank}(B-C)=n$. □

**3.2.12 Thí dụ. (OSV95)** Cho ma trận vuông $A = (a_{ij})$ cấp $n > 1$ có hạng $r$. Xét ma trận $A^* = (A_{ij})$, trong đó $A_{ij}$ là phần bù đại số của $a_{ij}$ trong $A$. Tìm hạng của $A^*$.

*Giải.* Ta nhận thấy $(A^*)^t$ là ma trận phụ hợp của $A$, nên nếu $\operatorname{rank}(A) = n$ thì $\operatorname{rank}(A^*) = n$. Nếu $\operatorname{rank}(A) < n-1$ thì mọi định thức con cấp $n-1$ đều bằng không, cho nên $\operatorname{rank}(A^*) = 0$.

Nếu $\operatorname{rank}(A) = n-1$ thì tồn tại $A_{ij}$ khác không và $A(A^*)^t = O$. Do đó, $\operatorname{rank}(A^*) \geq 0$. Mặt khác, do $\operatorname{rank}(A) + \operatorname{rank}(A^*)^t \leq n + \operatorname{rank}(A(A^*)^t)$ nên $\operatorname{rank}(A^*) \leq n+0-(n-1) = 1$. Vậy $\operatorname{rank}(A^*) = 1$. □

**3.2.13 (Chứng minh Định lý 6.2.1 trang 253)** Xét ánh xạ tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^m$ có biểu thức tọa độ đối với các cơ sở chính tắc là $f(\mathbf{x}) = A\mathbf{x}$. Khi đó, không gian nghiệm của hệ phương trình $A\mathbf{x} = \mathbf{0}$ chính là hạt nhân của $f$ hay $\ker f$. Vậy theo Định lý 3.2.3 và Mệnh đề 3.2.6 ta có

$$\dim S = \dim(\ker f) = \dim(\mathbb{R}^n) - \dim(\operatorname{Im} f) = n - \operatorname{rank} A = n - r. \square$$

**3.2.14 Thí dụ.** Cho ánh xạ tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^3$ cho bởi $f(x,y,z,t) = (x-2y+t, 3x-y+z, 4x-3y+z+t)$ với mọi $(x,y,z,t) \in \mathbb{R}^4$

Ma trận của $f$ trong cặp cơ sở chính tắc là $A = \begin{bmatrix} 1 & -2 & 0 & 1 \\ 3 & -1 & 1 & 0 \\ 4 & -3 & 1 & 1 \end{bmatrix}$. Do $\ker f$ chính là không gian nghiệm của hệ phương trình $A\mathbf{x} = \mathbf{0}$, cho nên ta biến đổi sơ cấp trên các dòng của $A$

$$A \xrightarrow[d_3 \to d_3 - 4d_1]{d_2 \to d_2 - 3d_1} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 5 & 1 & -3 \\ 0 & 5 & 1 & -3 \end{bmatrix} \xrightarrow{d_3 \to d_3 - d_2} \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 5 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ta chọn biến tự do là $y$ và $t$, ta được nghiệm của hệ phương trình viết dưới dạng ma trận

$$\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 2\alpha - \beta \\ \alpha \\ -5\alpha + 3\beta \\ \beta \end{bmatrix} = \alpha \begin{bmatrix} 2 \\ 1 \\ -5 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 3 \\ 1 \end{bmatrix} \qquad \alpha, \beta \in \mathbb{R}.$$

Vậy $\ker f = \langle (2,1,-5,0),(-1,0,3,1)\rangle$ và $\dim(\ker f) = 2$.

Vì $\dim(\operatorname{Im} f) = \dim(\mathbb{R}^4) - \dim(\ker f) = 4 - 2 = 2$ và $\operatorname{Im} f = \langle f(\mathbf{e}_1), f(\mathbf{e}_2), f(\mathbf{e}_3), f(\mathbf{e}_4)\rangle$ nên suy ra $\operatorname{Im} f = \langle (1,3,4),(-2,-1,-3)\rangle$ (do hệ hai vector $f(\mathbf{e}_1) = (1,3,4)$ và $f(\mathbf{e}_2) = (-2,-1,-3)$ độc lập tuyến tính). □

**3.2.15 Thí dụ.** Cho toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ cho bởi $f(x,y,z) = (x-y+z, 2x+y+3z, 4x-y+5z)$ với mọi $(x,y,z) \in \mathbb{R}^3$.

Ma trận của $f$ trong cơ sở chính tắc là $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & 3 \\ 4 & -1 & 5 \end{bmatrix}$. Để xác định $\operatorname{Im} f$ ta biến đổi sơ cấp ma trận $A^t$

$$A^t = \begin{bmatrix} 1 & 2 & 4 \\ -1 & 1 & -1 \\ 1 & 3 & 5 \end{bmatrix} \xrightarrow[d_3 \to d_3 - d_1]{d_2 \to d_2 + d_1} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 3 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow[d_3 \to d_3 - \frac{1}{3}d_2]{d_2 \to \frac{1}{3}d_2} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó $\operatorname{rank}(A) = 2$ cho nên $\dim(\operatorname{Im} f) = 2$. Hơn nữa, ta có $\operatorname{Im} f = \langle (1,2,4),(0,1,1)\rangle$.

Để xác định $\ker f$ ta xác định không gian nghiệm của hệ phương trình $A\mathbf{x} = \mathbf{0}$. Ta biến đổi sơ cấp ma trận $A$.

$$A \xrightarrow[d_3 \to d_3 - 4d_1]{d_2 \to d_2 - 2d_1} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 3 & 1 \\ 0 & 3 & 1 \end{bmatrix} \xrightarrow[\substack{d_2 \to \frac{1}{3}d_2 \\ d_3 \to d_3 - d_2}]{d_1 \to d_1 + \frac{1}{3}d_2} \begin{bmatrix} 1 & 0 & \frac{4}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

Vậy nghiệm của hệ phương trình $A\mathbf{x} = \mathbf{0}$ là

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{4}{3}\alpha \\ -\frac{1}{3}\alpha \\ \alpha \end{bmatrix} = \frac{\alpha}{3}\begin{bmatrix} -4 \\ -1 \\ 3 \end{bmatrix}.$$

Từ đó suy ra $\dim(\ker f) = 1$ và $\ker f = \langle (-4,-1,3)\rangle$. □

## 3.3 Không gian con bất biến

Những không gian con mà được toán tử tuyến tính ánh xạ vào chính nó thì có ý nghĩa quan trọng trong việc nghiên cứu về các toán tử tuyến tính.

**3.3.1 Định nghĩa.** Cho toán tử tuyến tính trên $\mathbb{K}$-không gian vector $V$. Không gian con $U$ của $V$ được gọi là **không gian con bất biến** đối với $f$ nếu $f(U) \subseteq U$; nghĩa là $f(\mathbf{x}) \in U$ với mọi $\mathbf{x} \in U$.

**3.3.2 Thí dụ.** Giả sử $f$ là một toán tử tuyến tính trên không gian vector $V$. Khi đó, các không gian con sau là $f$-bất biến: $\{\mathbf{0}\}$, $V$, $\mathrm{Im} f$, $\ker f$. $\square$

**3.3.3 Thí dụ.** Cho $f$ là một toán tử tuyến tính trên $\mathbb{R}^3$ xác định bởi

$$f(a, b, b) = (a + b, b + c, 0).$$

Do đó mặt phẳng $Oxy$, $\{(x, y, 0) : x, y \in \mathbb{R}\}$, và trục $Ox$, $\{(x, 0, 0) : x \in \mathbb{R}\}$, là những không gian con $f$-bất biến của $\mathbb{R}$. $\square$

**3.3.4 Mệnh đề.** *Cho $f$ là một toán tử tuyến tính trên không gian vector $V$. Nếu $U$ là không gian con bất biến đối với $f$ thì $f(U)$ và $f^{-1}(U)$ cũng là không gian con bất biến đối với $f$.*

*Chứng minh.* Theo giả thiết ta có $f(U) \subseteq U$. Do đó $f(f(U)) \subseteq f(U)$. Điều này có nghĩa là $f(U)$ là một không gian con bất biến đối với $f$.

Do $f(U) \subset U$ nên $U \subseteq f^{-1}(U)$. Mặt khác, ta luôn có $f(f^{-1}(U)) \subseteq U$. Do đó, $f(f^{-1}(U)) \subseteq f^{-1}(U)$. Vậy $f^{-1}(U)$ là không gian con bất biến đối với $f$. $\square$

**3.3.5 Định nghĩa.** Cho $f$ là một toán tử tuyến tính trên một không gian vector $V$ và $\mathbf{x}$ là một phần tử khác không bất kỳ của $V$. Không gian con

$$W = \langle \mathbf{x}, f(\mathbf{x}), f^2(\mathbf{x}), \ldots \rangle$$

được gọi là **không gian con $f$-cyclic** của $V$ *sinh bởi* $\mathbf{x}$.

Dễ dàng chứng minh được không gian con $W$ ở trên là $f$-bất biến. Thật sự, $W$ là không gian con $f$-bất biến nhỏ nhất chứa $\mathbf{x}$; nghĩa là bất kỳ không gian con $f$-bất biến chứa $\mathbf{x}$ phải chứa $W$.

**3.3.6 Thí dụ.** Cho $f$ là một toán tử tuyến tính trên $\mathbb{R}^3$ định nghĩa bởi

$$f(a, b, c) = (-b + c, a + c, 3c).$$

Ta xác định không gian con $f$-cyclic sinh bởi $\mathbf{e}_1 = (1, 0, 0)$. Do

$$f(\mathbf{e}_1) = f(1, 0, 0) = (0, 1, 0) = \mathbf{e}_2$$

và

$$f^2(\mathbf{e}_1) = f(f(\mathbf{e}_1)) = f(0, 1, 0) = (-1, 0, 0) = -\mathbf{e}_1,$$

$f^3(\mathbf{e}_1) = f(f^2(\mathbf{e}_1)) = f(-\mathbf{e}_1) = -\mathbf{e}_2$, $f^4(\mathbf{e}_1) = f(f^3(\mathbf{e}_1)) = f(-\mathbf{e}_2) = \mathbf{e}_1, \ldots$, nó kéo theo rằng $\langle \mathbf{e}_1, f(\mathbf{e}_1), f^2(\mathbf{e}_1), \ldots \rangle = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle = \{(a, b, 0) : a, b \in \mathbb{R}\}$. □

**3.3.7 Thí dụ.** Cho $T$ là toán tử tuyến tính trên $\mathbb{R}[x]$, không gian các đa thức, được định nghĩa bởi $T(p) = p'$ (đạo hàm của $p$). Khi đó không gian con $T$-cyclic sinh bởi $x^2$ là $\langle x^2, 2x, 1 \rangle = \mathbb{R}_2[x]$. □

## Bài tập

**1)** Cho ánh xạ tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3, x_4) = (x_1 - x_2 + x_3, 2x_1 + x_4, 2x_2 + x_3 - x_4)$. Tìm một cơ sở của $\ker f$ và một cơ sở của $\operatorname{Im} f$.

**2)** Cho toán tử tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^4$ xác định bởi $f(x_1, x_2, x_3, x_4) = (x_1 + 2x_2 + 4x_3 - 3x_4, 3x_1 + 5x_2 + 6x_3 - 4x_4, 4x_1 + 5x_2 - 2x_3 + 3x_4, 3x_1 + 8x_2 + 24x_3 - 19x_4)$. Tìm một cơ sở của $\ker f$ và $\operatorname{Im} f$.

**3)** Tìm toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ có ảnh sinh bởi hai vector $(1, 2, 3)$ và $(4, 5, 6)$. Hỏi $f$ có duy nhất không?

**4)** Tìm ánh xạ tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^3$ có nhân sinh bởi hai vector $(1, 2, 3, 4)$ và $(0, 1, 1, 1)$. Ánh xạ $f$ như vậy có duy nhất không?

**5)** Cho toán tử tuyến tính $f$ trên $\mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (ax_1 + x_2 + x_3, x_1 + ax_2 + x_3, x_1 + x_2 + ax_3)$, trong đó $a$ là một số thực nào đó. Tìm $a$ sao cho $\operatorname{rank} f = 3$. Khi $\operatorname{rank} f < 3$, tìm $\ker f$ và $\operatorname{Im} f$.

**6)** Cho $V$ là một $\mathbb{K}$-không gian vector và $W$ là một không gian con của $V$. Chứng minh rằng

**(a)** Tồn tại một toán tử tuyến tính $f$ trên $V$ mà $\operatorname{Im} f = W$.

**(b)** Tồn tại một toán tử tuyến tính $g$ trên $V$ mà $\ker g = W$.

**7)** Tìm toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ có ảnh sinh bởi hai vector $(1, 2, 3)$ và $(4, 5, 6)$. Hỏi $f$ có duy nhất không?

**8)** Tìm ánh xạ tuyến tính $g : \mathbb{R}^4 \to \mathbb{R}^3$ có hạt nhân sinh bởi hai vector $(1, 2, 3, 4)$ và $(0, 1, 1, 1)$. Hỏi ánh xạ $g$ như vậy có duy nhất không?

**9)** Cho toán tử tuyến tính $f$ trên $\mathbb{R}^4$ có ma trận trong cơ sở chính tắc là $A = \begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 5 & 11 & 2 \\ 0 & -1 & 3 & 1 \\ 1 & 2 & 1 & 3 \end{bmatrix}$. Tìm $\ker f$ và $\operatorname{Im} f$.

**10)** Cho ánh xạ tuyến tính $f : V \to V'$ và $\lambda \in \mathbb{K}$, $\lambda \neq 0$. Chứng minh rằng các ánh xạ $f$ và $\lambda f$ có cùng hạt nhân và ảnh.

**11)** **(OSV97)** Chứng minh rằng với một ma trận vuông cấp $n$ cho trước trên trường số thực đều tìm được số nguyên dương $N$ sao cho

$$\operatorname{rank}(A^k) = \operatorname{rank}(A^{k+1}) \qquad \text{với mọi } k \geq N.$$

**12)** Chứng minh rằng $\mathbb{K}_n[x]$ là một không gian $D$-cyclic với $D$ là toán tử tuyến tính $D(f) = f'$ (toán tử đạo hàm).

# § 4 Đơn cấu, toàn cấu, đẳng cấu

## 4.1 Các khái niệm và tính chất cơ bản

**4.1.1 Định nghĩa.** Cho $U$ và $V$ là các không gian vector trên trường $\mathbb{K}$ và $f : V \to U$ là ánh xạ tuyến tính. Khi đó, $f$ được gọi là **đơn cấu** (tương ứng, **toàn cấu**, **đẳng cấu**) nếu $f$ là đơn ánh (theo tương ứng là toàn ánh, song ánh).

Đặc biệt, nếu $V = U$, thì đơn cấu (toàn cấu, đẳng cấu) $f : V \to V$ gọi là tự đơn cấu (tự toàn cấu, tự đẳng cấu) của $V$.

**4.1.2 Thí dụ.**

**(i)** Cho không gian vector $V$ và $U$ là không gian con của $V$. Khi đó, phép nhúng chính tắc $i : U \to V$ (với mọi $\mathbf{x} \in U$ ta có $i(\mathbf{x}) = \mathbf{x}$) là đơn cấu.

**(ii)** Mọi phép chiếu chính tắc $p : V \to V/_W$ của không gian $V$ lên không gian thương $V/_W$ của nó đều là một toàn cấu.

**(iii)** Xét không gian vector $V$ và ánh xạ đồng nhất, $\mathrm{Id}_V$, trên $V$. Khi đó, $\mathrm{Id}_V$ là một tự đẳng cấu của $V$.

**(iv)** Với $\mathbb{K}$-không gian vector $n$ chiều $V$, ánh xạ tọa độ $c_{\mathscr{B}} : V \to \mathbb{K}^n$ ứng với cơ sở $\mathscr{B}$ của $V$ là một đẳng cấu. □

**4.1.3 Tính chất.** **(i)** *Cho $f : V \to U$ và $g : U \to W$ là các đơn cấu (tương ứng, toàn cấu, đẳng cấu). Khi đó, hợp thành $g \circ f$ là một đơn cấu (tương ứng, toàn cấu, đẳng cấu).*

**(ii)** *Nếu $f : V \to U$ là một đẳng cấu, thì $f^{-1} : U \to V$ cũng là một đẳng cấu.*

**4.1.4 Định nghĩa.** Cho $U$ và $V$ là hai không gian vector. Ta nói $V$ **đẳng cấu** với $U$ nếu tồn tại một đẳng cấu $f : V \to U$, kí hiệu, $V \cong U$.

**4.1.5 Mệnh đề.** *Quan hệ đẳng cấu giữa các không gian vector là một quan hệ tương đương.*

*Chứng minh.* Từ tính chất trên ta suy ra được mệnh đề này. □

## 4.2 Các dấu hiệu tương đương của đơn cấu và toàn cấu

**4.2.1 Định lý.** *Cho $f : V \to U$ là một ánh xạ tuyến tính trong đó $V$ là không gian hữu hạn chiều. Khi đó, các khẳng định sau là tương đương*

**(1)** *$f$ là một đơn cấu.*

**(2)** $\ker f = \{\mathbf{0}\}$.

**(3)** *Nếu hệ vector* $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ *độc lập tuyến tính trong* $V$*, thì hệ vector* $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ *độc lập tuyến tính trong* $U$*.*

**(4)** *Nếu* $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ *là một hệ vector trong* $V$ *thì* $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \operatorname{rank}(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$.

**(5)** *Nếu* $W$ *là không gian con của* $V$*, thì* $\dim f(W) = \dim W$.

**(6)** $\operatorname{rank} f = \dim V$.

*Chứng minh.*

(1)⇒(2) Giả sử $f$ là một đơn cấu. Lấy $\mathbf{x} \in \ker f$ tùy ý. Suy ra $f(\mathbf{x}) = \mathbf{0} = f(\mathbf{0})$. Do đó $\mathbf{x} = \mathbf{0}$ vì $f$ là đơn ánh. Vậy $\ker f = \{\mathbf{0}\}$.

(2)⇒(3) Giả sử $\ker f = \{\mathbf{0}\}$ và hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ độc lập tuyến tính trong $V$. Xét tổ hợp tuyến tính $\lambda_1 f(\mathbf{a}_1) + \lambda_2 f(\mathbf{a}_2) + \cdots + \lambda_n f(\mathbf{a}_n) = \mathbf{0}$. Suy ra $f(\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n) = \mathbf{0}$. Vậy $\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n \in \ker f$. Do đó, $\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n = \mathbf{0}$. Từ đó suy ra $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. Vậy hệ $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ độc lập tuyến tính.

(3)⇒(4) Nếu hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ toàn là vector không, thì hệ vector $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ toàn là vector không; suy ra $\operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = 0 = \operatorname{rank}(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$. Giả sử trong hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ có vector khác không. Khi đó, lấy $(\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_r})$ là một hệ con độc lập tuyến tính tối đại của hệ $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. Theo (3) ta được hệ $(f(\mathbf{a}_{i_1}), \ldots, f(\mathbf{a}_{i_r}))$ độc lập tuyến tính và là hệ con của $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$. Do đó, $\operatorname{rank}(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n)) \geq r = \operatorname{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$. Theo Định lý 1.2.5 ta có được kết quả.

(4)⇒(5) Khi $W = \{\mathbf{0}\}$ dễ dàng thấy được kết quả. Giả sử $(\mathbf{a}_1, \ldots, \mathbf{a}_r)$ là một cơ sở của $W$. Theo (4) thì $\operatorname{rank}(f(\mathbf{a}_1), \ldots, f(\mathbf{a}_r)) = \operatorname{rank}(\mathbf{a}_1, \ldots, \mathbf{a}_r) = r$. Suy ra hệ $(f(\mathbf{a}_1), \ldots, f(\mathbf{a}_r))$ độc lập tuyến tính trong $f(W)$. Theo Định lý 3.1.1 ta suy ra được hệ $(f(\mathbf{a}_1), \ldots, f(\mathbf{a}_r))$ cũng là một hệ sinh trong $f(W)$. Vậy hệ

$(f(\mathbf{a}_1), \ldots, f(\mathbf{a}_r))$ là cơ sở của $f(W)$. Suy ra $\dim f(W) = r = \dim W$.

(5)⇒(6) Do $V$ là không gian con của $V$, nên theo (5) ta có $\dim f(V) = \dim V$. Do đó, $\operatorname{rank} f = \dim V$.

(6)⇒(1) Khi $V = \{\mathbf{0}\}$ rõ ràng $f$ là một đơn cấu. Giả sử $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là cơ sở của $V$. Theo Mệnh đề 3.2.4 hệ $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ là hệ sinh của $\mathrm{Im} f$. Hơn nữa, theo (6) ta có $\operatorname{rank} f = \dim V = n$, cho nên hệ $(f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ là cơ sở của $\mathrm{Im} f$. Giả sử có $\mathbf{u}, \mathbf{u}' \in V$ sao cho $f(\mathbf{u}) = f(\mathbf{u}')$. Ta có biểu diễn $\mathbf{u} = \sum_{i=1}^n \lambda_i \mathbf{a}_i$ và $\mathbf{u}' = \sum_{i=1}^n \lambda_i' \mathbf{a}_i$. Khi đó, ta có

$$\mathbf{0} = f(\mathbf{u}) - f(\mathbf{u}') = f(\mathbf{u} - \mathbf{u}') = f\Big(\sum_{i=1}^n (\lambda_i - \lambda_i')\mathbf{a}_i\Big) = \sum_{i=1}^n (\lambda_i - \lambda_i') f(\mathbf{a}_i).$$

Từ đó suy ra $\lambda_i = \lambda_i'$ với mọi $i = 1, \ldots, n$ (vì $(f(\mathbf{a}_1), \ldots, f(\mathbf{a}_n))$ là cơ sở của $\mathrm{Im} f$). Do đó, ta có $\mathbf{u} = \mathbf{u}'$. Vậy ta được $f$ là đơn cấu. □

**4.2.2 Thí dụ.** Cho $f : \mathbb{R}^3 \to \mathbb{R}^4$ là ánh xạ tuyến tính xác định bởi ảnh của cơ sở chính tắc $\mathscr{E} = (\mathbf{e}_1 = (1,0,0), \mathbf{e}_2 = (0,1,0), \mathbf{e}_3 = (0,0,1))$ như sau: $f(\mathbf{e}_1) = \mathbf{u}_1 = (1,1,2,2)$, $f(\mathbf{e}_2) = \mathbf{u}_2 = (2,3,5,6)$, $f(\mathbf{e}_3) = (4,5,9,10)$. Hỏi $f$ có là đơn cấu không?

*Giải.* Ta thấy rằng ảnh của $f$ là không gian sinh bởi các vector $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$. Ta lập ma trận $A$ mà các dòng ứng với các vector $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ rồi biến đổi sơ cấp theo dòng ta được

$$A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 3 & 5 & 6 \\ 4 & 5 & 9 & 10 \end{bmatrix} \xrightarrow[d_3 \to d_3 - 4d_1]{d_2 \to d_2 - 2d_1} \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix} \xrightarrow{d_3 \to d_3 - d_2} \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Suy ra $\operatorname{rank}(f) = \dim\langle \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \rangle = \operatorname{rank} A = 2 < 3$. Vậy $f$ không là đơn cấu. □

**4.2.3 Định lý.** *Cho $f : V \to U$ là một ánh xạ tuyến tính trong đó $U$ là không gian hữu hạn chiều. Khi đó, các khẳng định sau là tương đương*

**(1)** *$f$ là một toàn cấu.*

**(2)** $\mathrm{Im} f = U$.

**(3)** $\operatorname{rank} f = \dim U$.

**(4)** *Nếu $\mathcal{S}$ là hệ sinh của $V$, thì $f(\mathcal{S})$ là hệ sinh của $U$.*

**(5)** *Tồn tại $\mathcal{S}$ là một hệ sinh của $V$ sao cho $f(\mathcal{S})$ là hệ sinh của $U$.*

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

**4.2.4 Thí dụ.** Xét ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi $f(\mathbf{e}_1) = \mathbf{u}_1 = (1, 1)$, $f(\mathbf{e}_2) = \mathbf{u}_2 = (1, 2)$, $f(\mathbf{e}_3) = \mathbf{u}_3 = (0, 0)$. Khi đó ta có

$$\operatorname{rank} f = \operatorname{rank}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \operatorname{rank}\left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}\right) = 2 = \dim(\mathbb{R}^2).$$

Vậy $f$ là một toàn cấu. □

Kết hợp kết quả của hai định lý trên ta có hệ quả sau.

**4.2.5 Hệ quả.** *Cho $f : V \to U$ là một ánh xạ tuyến tính từ $\mathbb{K}$-không gian vector $n$ chiều $V$ đến không gian vector $U$, và $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là một cơ sở của $V$. Khi đó,*

**(1)** *$f$ là đơn cấu khi và chỉ khi hệ $(f(\mathbf{e}_1), f(\mathbf{e}_2), \ldots, f(\mathbf{e}_n))$ độc lập tuyến tính trong $U$.*

**(2)** *$f$ là toàn cấu khi và chỉ khi hệ $(f(\mathbf{e}_1), f(\mathbf{e}_2), \ldots, f(\mathbf{e}_n))$ là một hệ sinh của $U$.*

**(3)** *$f$ là đẳng cấu khi và chỉ khi hệ $(f(\mathbf{e}_1), f(\mathbf{e}_2), \ldots, f(\mathbf{e}_n))$ là một cơ sở của $U$.*

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

**4.2.6 Nhận xét.** Từ Định lý 4.2.1, Định lý 4.2.3 và Hệ quả 4.2.5 ta có: nếu $f : V \to U$ là một ánh xạ tuyến tính thì

- $f$ là đơn cấu khi và chỉ khi $\operatorname{rank}(f) = \dim V \leq \dim U$.
- $f$ là toàn cấu khi và chỉ khi $\operatorname{rank}(f) = \dim U \leq \dim V$.
- $f$ là đẳng cấu khi và chỉ khi $\operatorname{rank}(f) = \dim V = \dim U$.

Đặc biệt, nếu $f : V \to U$ là ánh xạ tuyến tính và $\dim V = \dim U$ thì $f$ là đơn cấu khi và chỉ khi $f$ là toàn cấu khi và chỉ khi $f$ là đẳng cấu.

**4.2.7 Thí dụ.** Cho $f : \mathbb{R}^3 \to \mathbb{R}^3$ là ánh xạ tuyến tính xác định bởi $f(\mathbf{e}_1) = (1, 1, 1)$, $f(\mathbf{e}_2) = (1, 1, 0)$, $f(\mathbf{e}_3) = (1, 0, 0)$.

Do $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = -1 \neq 0$ nên $(f(\mathbf{e}_1), f(\mathbf{e}_2), f(\mathbf{e}_3))$ là một cơ sở của $\mathbb{R}^3$. Vậy $f$ là một đẳng cấu. □

**4.2.8 Định lý.** *Hai $\mathbb{K}$-không gian vector hữu hạn chiều đẳng cấu với nhau khi và chỉ khi chúng có cùng số chiều.*

*Chứng minh.* Giả sử không gian vector $V$ đẳng cấu với không gian vector $U$. Khi đó tồn tại đẳng cấu $f : V \to U$. Giả sử $\mathscr{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ là một cơ sở của $V$. Theo hệ quả trên ta có $(f(\mathbf{a}_1), \ldots, f(\mathbf{a}_n))$ là một cơ sở của $U$. Cho nên $\dim U = n = \dim V$.

Ngược lại, giả sử $\dim V = \dim U = n$. Gọi $\mathscr{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ là một cơ sở của $V$ và $\mathscr{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Khi đó, theo Định lý 1.3.1 thì tồn tại duy nhất một ánh xạ tuyến tính $f : V \to U$ sao cho $f(\mathbf{a}_i) = \mathbf{b}_i$ với mọi $i = 1, \ldots, n$. Một lần nữa theo hệ quả trên ta suy ra được $f$ là một đẳng cấu hay $V$ và $U$ đẳng cấu với nhau. □

**4.2.9 Nhận xét.** Cho $\mathbb{K}$-không gian vector $n$ chiều $V$ và một cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ trong nó. Khi đó, ta có đẳng cấu tọa độ ứng với $\mathscr{B}$ là $\mathbf{c}_B : V \to \mathbb{K}^n$ với $c_{\mathscr{B}}(\mathbf{x}) = (x_1, x_2, \ldots, x_n)$ ở đó $(x_1, x_2, \ldots, x_n)$ chính là tọa độ của vector $\mathbf{x} \in V$ đối với cơ sở $\mathscr{B}$. Nhờ đẳng cấu tọa độ này, ta có thể chuyển các kết quả đã biết trong $\mathbb{K}^n$ sang $V$. Chẳng hạn, xét hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ trong $V$ với $\mathbf{a}_j|_B = (a_{1j}, a_{nj}, \ldots, a_{nj})$. Khi đó, ta được

- $\mathrm{rank}(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \mathrm{rank}(A)$ với $A = (a_{ij})_{n\times m}$
- Hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ độc lập tuyến tính trong $V$ khi và chỉ khi $\mathrm{rank}(A) = m$
- Hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ phụ thuộc tuyến tính trong $V$ khi và chỉ khi $\mathrm{rank}(A) < m$.
- Đặc biệt, trường hợp $m = n$ ta có $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một cơ sở của $V$ khi và chỉ khi $\det(A) \neq 0$.

**4.2.10 Thí dụ.** Xét tính độc lập tuyến tính của hệ các đa thức $(p_1(x) = 1 + x + x^2 + x^3, p_2(x) = 2x + 3x^2 + 4x^3, p_3(x) = 4x^2 + 5x^3)$ trong không gian $\mathbb{R}_3[x]$ các đa thức bậc không quá 3.

*Giải.* Ta có $\dim(\mathbb{R}_3[x]) = 4$ với $\mathscr{B} = (1, x, x^2, x^3)$ là một cơ sở của nó. Rõ ràng $p_1(x)|_{\mathscr{B}} = (1, 1, 1, 1)$, $p_2(x)|_{\mathscr{B}} = (0, 2, 3, 4)$, $p_3(x)|_{\mathscr{B}} = (0, 0, 4, 5)$. Do đó

$$\mathrm{rank}(p_1(x), p_2(x), p_3(x)) = \mathrm{rank}\left(\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 4 \\ 1 & 4 & 5 \end{bmatrix}\right) = 3.$$

Vậy $(p_1(x), p_2(x), p_3(x))$ độc lập tuyến tính trong $\mathbb{R}_3[x]$. □

## 4.3 Định lý cơ bản về nhân tử hóa ánh xạ tuyến tính

**4.3.1 Định lý.** *Giả sử $V$ và $U$ là hai $\mathbb{K}$-không gian vector và $f : V \to U$ là một ánh xạ tuyến tính. Khi đó, tồn tại duy nhất một ánh xạ tuyến tính $\tilde{f} : V/_{\ker f} \to U$ sao cho $f = \tilde{f} \circ p$ ở đó $p : V \to V/_{\ker f}$ là phép chiếu chính tắc $V$ lên không gian thương $V/_{\ker f}$. Hơn nữa, $\tilde{f}$ là một đơn cấu (và thường gọi là đơn cấu cảm sinh từ $f$).*

*Chứng minh. Sự tồn tại của $\tilde{f}$.* Ta định nghĩa ánh xạ $\tilde{f} : V/_{\ker f} \to U$ như sau $\tilde{f}(\tilde{\mathbf{x}}) = f(\mathbf{x})$ với mọi $\tilde{\mathbf{x}} \in V/_{\ker f}$. Trước hết ta cần chứng tỏ rằng $\tilde{f}$ định nghĩa như vậy là hợp lý, nghĩa là không phụ thuộc đại diện $\mathbf{x} \in V$ của lớp $\tilde{\mathbf{x}} \in V/_{\ker f}$. Thật vậy, nếu $\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$ tức là $\mathbf{x} - \mathbf{y} \in \ker f$, do đó $f(\mathbf{x} - \mathbf{y}) = \mathbf{0}$ hay $f(\mathbf{x}) = f(\mathbf{y})$. Mặt khác, với mọi $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in V/_{\ker f}$ và mọi $\lambda, \mu \in \mathbb{K}$, ta có

$$\begin{aligned}
\tilde{f}(\lambda\tilde{x} + \mu\tilde{\mathbf{y}}) &= \tilde{f}(\widetilde{\lambda\mathbf{x} + \mu\mathbf{y}}) = f(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda f(\mathbf{x}) + \mu f(\mathbf{y}) \\
&= \lambda\tilde{f}(\tilde{\mathbf{x}}) + \mu\tilde{f}(\tilde{\mathbf{y}})
\end{aligned}$$

Do đó $\tilde{f}$ là một ánh xạ tuyến tính. Hơn nữa, theo định nghĩa của phép chiếu chính tắc $p : V \to V/_{\ker f}$ với $p(\mathbf{x}) = \tilde{\mathbf{x}}$ ta có

$$\tilde{f} \circ p(\mathbf{x}) = \tilde{f}(p(\mathbf{x})) = \tilde{f}(\tilde{\mathbf{x}}) = f(\mathbf{x}) \qquad \text{với mọi } \mathbf{x} \in V$$

Tức là $\tilde{f} \circ p = f$ hay biểu đồ sau giao hoán.

$$\begin{tikzcd}
V \arrow[rr, "f"] \arrow[dr, "p"'] & & U \\
& V/_{\ker f} \arrow[ur, dashed, "\tilde{f}"'] &
\end{tikzcd}$$

*Sự duy nhất của $\tilde{f}$.* Nếu $g : V/_{\ker f} \to U$ cũng là một ánh xạ tuyến tính thỏa mãn đẳng thức $f = g \circ p$ thì với mọi $\tilde{\mathbf{x}} \in V/_{\ker f}$ ta có

$$g(\tilde{\mathbf{x}}) = g(p(\mathbf{x})) = g \circ p(\mathbf{x}) = f(\mathbf{x}) = \tilde{f}(\tilde{\mathbf{x}})$$

Do đó, $g = \tilde{f}$ hay $\tilde{f}$ là duy nhất.

*Tính đơn cấu của $\tilde{f}$.* Rõ ràng với mọi $\mathbf{x} \in \ker \tilde{f}$, ta có $\tilde{f}(\tilde{\mathbf{x}}) = \mathbf{0}$ hay $f(\mathbf{x}) = \mathbf{0}$, suy ra $\mathbf{x} \in \ker f$ hay $\tilde{\mathbf{x}} = \tilde{\mathbf{0}} \in V/_{\ker f}$. Do đó ta được $\ker \tilde{f} = \{\tilde{\mathbf{0}}\}$. Vậy $\tilde{f}$ là đơn cấu. □

**4.3.2 Nhận xét.** Định lý này mang tên *định lý cơ bản về nhân tử hóa ánh xạ tuyến tính* (hay định lý cơ bản về ánh xạ tuyến tính) vì nó cho phép phân tích mỗi ánh xạ tuyến tính $f$ thành hợp thành của hai ánh xạ $\tilde{f}$ và $p$. Hơn nữa, $p$ là một toàn cấu và $\tilde{f}$ là một đơn cấu.

Vì mọi ánh xạ tuyến tính đều có thể xem là một toàn cấu lên ảnh của nó nên mệnh đề sau đây là hệ quả trực tiếp của định lý cơ bản về nhân tử hóa ánh xạ tuyến tính và Định lý 7.2.5 (trang 262)

**4.3.3 Hệ quả.** **(i)** *Đối với mọi ánh xạ tuyến tính $f : V \to U$ của các $\mathbb{K}$-không gian vector ta đều có $V/_{\ker f} \cong \operatorname{Im} f$.*

**(ii)** *Nếu $f : V \to U$ là một ánh xạ tuyến tính từ $\mathbb{K}$-không gian vector $n$ chiều $V$ đên $\mathbb{K}$-không gian vector $U$ thì*

$$\dim(\operatorname{Im} f) + \dim(\ker f) = \operatorname{rank}(f) + \operatorname{def}(f) = n.$$

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

## Bài tập

**1)** Chứng minh rằng hợp thành của các đẳng cấu là một đẳng cấu; nghịch đảo của một đẳng cấu cũng là một đẳng cấu.

**2)** Chứng minh rằng nếu $f$ là tự đẳng cấu của không gian vector $V$ và $U$ là một không gian con bất biến đối với $f$ thì $U$ cũng là một không gian con bất biến đối với $f^{-1}$.

**3)** Cho ánh xạ tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3, x_4) = (x_1 - x_2 + x_3, 2x_1 + x_4, 2x_2 + x_3 - x_4)$. Tìm một cơ sở của $\ker f$ và một cơ sở của $\operatorname{Im} f$. Hỏi $f$ có phải là đơn cấu, toàn cấu không?

**4)** Chứng minh Tính chất 4.1.3.

**5)** Chứng minh Định lý 4.2.3.

**6)** Chứng minh Hệ quả 4.2.5.

**7)** Cho toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (x_1 - x_2 + x_3, x_1 + x_3, x_1 + x_2 + \alpha x_3)$. Tìm $\alpha$ để $f$ không là đẳng cấu. Với $\alpha$ vừa tìm được, hãy tìm số chiều và cơ sở của $\ker f$ và $\operatorname{Im} f$.

**8 )** Cho toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (3x_1 + 5x_2 + 2x_3, 4x_1 + 7x_2 + (m+1)x_3, x_1 + x_2 - 4mx_3)$ Tìm giá trị của $m$ để $f$ không là một đẳng cấu. Với $m$ vừa tìm được hãy tìm cơ sở và số chiều của $\mathrm{Im} f$, $\ker f$.

**9 )** Cho toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x, y, z) = (-2x + y + z, x - 2y + z, x + y + \lambda z)$. Tìm $\lambda$ để $f$ không khả nghịch.

**10 ) (OSV93)** Cho hai số phức $\alpha$ và $\beta$ với $\alpha \neq \beta$ và cho ánh xạ

$$
\begin{aligned}
f : \mathbb{C}_3[x] &\to \mathbb{C}_2[x] \\
p(x) &\mapsto p(x+\alpha) - p(x+\beta)
\end{aligned}
$$

**(a)** Chứng minh $f$ là toàn cấu; **(b)** Tìm $\ker f$.

**11 )** Chứng minh Hệ quả 4.3.3.

# § 5 Không gian các ánh xạ tuyến tính và không gian đối ngẫu

## 5.1 Không gian các ánh xạ tuyến tính

**5.1.1** Cho hai $\mathbb{K}$-không gian vector $V$ và $W$. Xét tập hợp tất cả các ánh xạ tuyến tính từ $V$ đến $W$, ký hiệu là $\mathrm{Hom}_{\mathbb{K}}(V, W)$ hay chỉ đơn giản là $\mathrm{Hom}(V, W)$ khi không sợ nhầm lẫn. Hiển nhiên $\mathrm{Hom}(V, W) \neq \varnothing$ vì ít nhất có ánh xạ không $\theta : V \to W$ thuộc nó. Ta sẽ trang bị cho $\mathrm{Hom}(V, W)$ một cấu trúc $\mathbb{K}$-không gian vector bằng cách định nghĩa phép cộng và phép nhân với vô hướng như sau:

$$
\begin{aligned}
(f+g)(\mathbf{x}) &= f(\mathbf{x}) + g(\mathbf{x}), \qquad \text{với mọi } \mathbf{x} \in V \\
(\lambda f)(\mathbf{x}) &= \lambda f(\mathbf{x}), \qquad \text{với mọi } \mathbf{x} \in V
\end{aligned}
\tag{5.1.2}
$$

với mọi $f, g \in \mathrm{Hom}(V, W)$ mọi $\lambda \in \mathbb{K}$.

Dễ dàng thử lại rằng $f + g$ và $\lambda f$ đều là ánh xạ tuyến tính từ $V$ đến $W$. Nói cách khác, (5.1.2) cho ta phép cộng trên $\mathrm{Hom}(V, W)$ và phép nhân một vô hướng (thuộc $\mathbb{K}$ với một ánh xạ tuyến tính thuộc $\mathrm{Hom}(V, W)$.

Hơn nữa, ta cũng có thể kiểm chứng được rằng $\mathrm{Hom}(V, W)$ cùng với hai phép toán ấy là một $\mathbb{K}$-không gian vector và được gọi là **không gian các ánh xạ tuyến tính** từ $V$ đến $W$.

Trong không gian $\mathrm{Hom}(V, W)$, ánh xạ không $\theta : V \to W$ chính là vector không, còn vector đối của $f \in \mathrm{Hom}(V, W)$ chính là $(-1)f$ và được ký hiệu như thường lệ là $-f$.

**5.1.3** Khi $W = V$, không gian $\mathrm{Hom}(V, W)$ các toán tử tuyến tính trên $V$ được ký hiệu lại là $\mathrm{End}(V)$ (hay $\mathrm{End}_{\mathbb{K}}(V)$ khi cần phải chỉ rõ trường cơ sở $\mathbb{K}$). Trên $\mathrm{End}(V)$ đương nhiên còn có phép hợp thành các toán tử tuyến tính. Dễ dàng kiểm chứng được rằng phép nhân phân phối đối với phép cộng, tức là

$$(f + g) \circ h = f \circ h + g \circ h \qquad h \circ (f + g) = h \circ f + h \circ g,$$

với mọi $f, g, h \in \mathrm{End}(V)$. Bởi vậy, ngoài cấu trúc $\mathbb{K}$-không gian vector $\mathrm{End}(V)$ còn là một vành với phép cộng và phép hợp thành. Vành $\mathrm{End}(V)$ gọi là **vành các toán tử tuyến tính** trên $V$.

**5.1.4** Tập hợp con $\mathrm{GL}(V) = \{f \in \mathrm{End}(V) : f$ là tự đẳng cấu$\}$ rõ ràng là một nhóm (với phép toán hợp thành) và gọi là **nhóm các tự đẳng cấu** của $V$.

**5.1.5** Cho $V$ và $W$ lần lượt là $\mathbb{K}$-không gian vector $n$ và $m$ chiều. Giả sử trong $V$ và $W$ tương ứng đã cho cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ và $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_m)$. Khi đó, mỗi ánh xạ tuyến tính $f : V \to W$ đều có một ma trận duy nhất $A_f \in \mathcal{M}_{m,n}(\mathbb{K})$ trong cặp cơ sở $(\mathscr{B}, \mathscr{B}')$. Vậy ta thu được ánh xạ

$$\begin{aligned} \mathcal{A} : \mathrm{Hom}(V, W) &\to \mathcal{M}_{m,n}(\mathbb{K}) \\ f &\mapsto \mathcal{A}(f) = A_f. \end{aligned}$$

**5.1.6 Định lý.** *Với giả thiết và ký hiệu như trên,* $\mathrm{Hom}(V, W)$ *là* $\mathbb{K}$*-không gian vector* $nm$ *(*$= \dim(V)\dim(W)$*) chiều và ánh xạ* $\mathcal{A}$ *là một đẳng cấu. Nói riêng,* $\mathrm{End}(V)$ *là không gian* $n^2$ *chiều.*

*Chứng minh.* Hiển nhiên nếu $m = 0$ hay $n = 0$ thì $\mathrm{Hom}(V, W)$ chỉ gồm duy nhất ánh xạ không và do đó là không gian 0 chiều. Ta sẽ chứng minh định lý với $m, n > 0$.

(1) Xét hệ $nm$ phần tử $(f_{ij})_{i=1,\dots,m;j=1,\dots,n}$ với $f_{ij}$ ($1 \le i \le m$, $1 \le j \le n$) là ánh xạ tuyến tính từ $V$ đến $W$ được cho bởi điều kiện $f_{ij}(\mathbf{e}_k) = \delta_{jk}\mathbf{e}'_i$; $k = 1, \dots, n$ ở đó $\delta_{jk} = \begin{cases} 0 & \text{khi } j \neq k \\ 1 & \text{khi } j = k; \end{cases}$ $j, k = 1, \dots, n$. Theo định lý cơ bản về sự xác định ánh xạ tuyến tính, có một và chỉ một ánh xạ $f_{ij}$ như thế ($i = 1, \dots, m$; $j = 1, \dots, n$). Ta sẽ chỉ ra hệ $(f_{ij})_{i=1,\dots,m;j=1,\dots,n}$ là một cơ sở của $\mathrm{Hom}(V, W)$.

Giả sử $f \in \mathrm{Hom}(V, W)$ là một ánh xạ tuyến tính tùy ý từ $V$ đến $W$ và $f(\mathbf{e}_k) = \sum\limits_{i=1}^{m} a_{ik}\mathbf{e}'_i$ với $k = 1, \dots, n$. Khi đó ta có

$$\begin{aligned} f(\mathbf{e}_k) = \sum_{i=1}^{m} a_{ik}\mathbf{e}'_i &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}(\delta_{jk}\mathbf{e}'_i) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} f_{ij}(\mathbf{e}_k) \\ &= \Big(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} f_{ij}\Big)(\mathbf{e}_k) \end{aligned}$$

với mọi $k = 1, \dots, n$. Do đó, $f = \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} a_{ij} f_{ij}$, tức là $f$ biểu thị tuyến tính được qua hệ $(f_{ij})_{i=1,\dots,m;j=1,\dots,n}$. Nói cách khác, $(f_{ij})_{i=1,\dots,m;j=1,\dots,n}$ là một hệ sinh của $\mathrm{Hom}(V, W)$.

Mặt khác, nếu $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} a_{ij} f_{ij} = \theta$ (ánh xạ không) là một tổ hợp tuyến tính bằng không của hệ $(f_{ij})_{i=1,\dots,m;j=1,\dots,n}$ trong $\mathrm{Hom}(V, W)$ thì với mỗi $k = 1, \dots, n$ ta có

$$\begin{aligned} \mathbf{0} = \Big(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} f_{ij}\Big)(\mathbf{e}_k) &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} f_{ij}(\mathbf{e}_k) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}(\delta_{jk}\mathbf{e}'_i) \\ &= \sum_{i=1}^{m} a_{ik}\mathbf{e}'_i. \end{aligned}$$

Từ tính độc lập tuyến tính của $\mathscr{B}'$ ta suy ra $a_{ik} = 0$ với mỗi $i = 1, \dots, m$. Do đó ta được $a_{ij} = 0$ với mỗi $i = 1, \dots, m$ và $j = 1, \dots, n$. Vậy hệ

$(f_{ij})_{i=1,\ldots,m;j=1,\ldots,n}$ độc lập tuyến tính, cho nên nó chính là một cơ sở của $\mathrm{Hom}(V,W)$ và $\dim(\mathrm{Hom}(V,W)) = nm$.

(2) Với $\alpha, \beta \in \mathbb{K}$ tùy ý và $f, g$ tùy ý thuộc $\mathrm{Hom}(V,W)$ giả sử

$$\mathcal{A}(f) = A = (a_{ij})_{m\times n}, \qquad \mathcal{A}(g) = B = (b_{ij})_{m\times n}.$$

Khi đó theo định nghĩa của ánh xạ $\mathcal{A}$ ta có

$$f(\mathbf{e}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{e}'_i, \qquad g(\mathbf{e}_j) = \sum_{i=1}^{m} b_{ij}\mathbf{e}'_i \qquad j = 1,\ldots,n.$$

Do đó với mỗi $j = 1,\ldots,n$ ta có

$$\begin{aligned}(\alpha f + \beta g)(\mathbf{e}_j) &= \alpha f(\mathbf{e}_j) + \beta g(\mathbf{e}_j)\\ &= \alpha\sum_{i=1}^{m} a_{ij}\mathbf{e}'_i + \beta\sum_{i=1}^{m} b_{ij}\mathbf{e}'_i\\ &= \sum_{i=1}^{m}(\alpha a_{ij} + \beta b_{ij})\mathbf{e}'_j\end{aligned}$$

Vậy $\alpha f + \beta g$ có ma trận trong cặp cơ sở $(\mathscr{B}, \mathscr{B}')$ là

$$(\alpha a_{ij} + \beta b_{ij})_{m\times n} = \alpha A + \beta B.$$

Suy ra $\mathcal{A}(\alpha f + \beta g) = \alpha\mathcal{A}(f) + \beta\mathcal{A}(g)$, hay $\mathcal{A}$ là một ánh xạ tuyến tính.

Mặt khác, theo định lý cơ bản về sự xác định ánh xạ tuyến tính, với mỗi ma trận $A = (a_{ij})_{m\times n} \in \mathcal{M}_{m,n}(\mathbb{K})$, có một và chỉ một ánh xạ tuyến tính $f_A : V \to W$ sao cho $f_A(\mathbf{e}_j) = \sum\limits_{i=1}^{m} a_{ij}\mathbf{e}'_i$ với $j = 1,\ldots,n$. Do đó ta thu được ánh xạ

$$\begin{aligned}\mathcal{F} : \mathcal{M}_{m,n}(\mathbb{K}) &\to \mathrm{Hom}(V,W)\\ A &\mapsto \mathcal{F}(A) = f_A.\end{aligned}$$

Dễ thấy rằng

$$\mathcal{A}\circ\mathcal{F} = \mathrm{Id}_{\mathcal{M}_{m,n}(\mathbb{K})}, \qquad \mathcal{F}\circ\mathcal{A} = \mathrm{Id}_{\mathrm{Hom}(V,W)}$$

Vậy $\mathcal{A}$ là một đẳng cấu tuyến tính. □

**5.1.7 Hệ quả.** *$\mathcal{M}_{m,n}(\mathbb{K})$ là $\mathbb{K}$-không gian vector $mn$ chiều. Đặc biệt, $\mathcal{M}_n(\mathbb{K})$ là $\mathbb{K}$-không gian vector $n^2$ chiều.*

**5.1.8 Nhận xét.** Gọi $I_{ij}$ là ma trận cấp $m \times n$ trên $\mathbb{K}$ mà tất cả các phần tử đều bằng 0 ngoại trừ phần tử ở dòng thứ $i$ cột thứ $j$ bằng 1 với $i = 1, \dots, m$ và $j = 1, \dots, n$. Khi đó, dễ thấy $I_{ij}$ chính là ma trận của $f_{ij}$ trong cặp cơ sở $(\mathscr{B}, \mathscr{B}')$ tức là $\mathcal{A}(f_{ij}) = I_{ij}$ với $i = 1, \dots, m$ và $j = 1, \dots, n$. Suy ra $(I_{ij})_{i=1,\dots,m;j=1,\dots,n}$ là ảnh của cơ sở $(f_{ij})_{i=1,\dots,m;j=1,\dots,n}$ bởi đẳng cấu $\mathcal{A}$, do đó $(I_{ij})_{i=1,\dots,m;j=1,\dots,n}$ là một cơ sở của $\mathcal{M}_{m,n}(\mathbb{K})$.

## 5.2 Không gian đối ngẫu và không gian song đối ngẫu

**5.2.1 Định nghĩa.** Cho $V$ là một $\mathbb{K}$-không gian vector. Không gian vector $\mathrm{Hom}(V, \mathbb{K})$ các ánh xạ tuyến tính từ $V$ đến $\mathbb{K}$ (xem $\mathbb{K}$ là một không gian vector trên chính nó) được gọi là **không gian đối ngẫu** của $V$, ký hiệu là $V^*$. Mỗi $f \in V^*$, tức là một ánh xạ tuyến tính từ $V$ đến $\mathbb{K}$ được gọi là một **dạng tuyến tính** trên $V$.

**5.2.2** Giá trị $f(\mathbf{x}) \in \mathbb{K}$ của dạng tuyến tính $f \in V^*$ tại $\mathbf{x} \in V$ thường được viết lại là $\langle f, \mathbf{x} \rangle$. Theo cách viết này, tính tuyến tính và các phép toán tuyến tính trên $V^*$ được thể hiện như sau

$$\Big\langle \sum_{i=1}^{m} \alpha_i f_i, \sum_{j=1}^{n} \beta_j \mathbf{x}_j \Big\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j \langle f_i, \mathbf{x}_j \rangle$$

với mọi $f_1, f_2, \dots, f_m \in V^*$, mọi $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in V$, và mọi $\alpha_1, \dots, \alpha_m$, $\beta_1, \dots, \beta_n \in \mathbb{K}$.

Khi $V$ là một $\mathbb{K}$-không gian vector $n$ chiều thì đương nhiên không gian đối ngẫu $V^*$ cũng $n$ chiều, vì $\mathbb{K}$ là không gian 1 chiều và $\dim(V^*) = \dim(\mathrm{Hom}(V, \mathbb{K})) = n$.

**5.2.3 Định lý.** *Cho $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ là một cơ sở của $\mathbb{K}$-không gian vector $n$ chiều $V$. Khi đó tồn tại duy nhất một cơ sở $\mathscr{B}^* = (\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_n^*)$*

*của không gian đối ngẫu $V^*$ sao cho*

(5.2.4) $$\langle \mathbf{e}_i^*, \mathbf{e}_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{khi } i \neq j \\ 1 & \text{khi } i = j \end{cases} \qquad i, j = 1, \ldots, n.$$

*Cơ sở $\mathscr{B}^*$ gọi là **cơ sở đối ngẫu** của $\mathscr{B}$.*

*Chứng minh.* Rõ ràng hệ $\mathscr{B}^*$ của $V^*$ hoàn toàn được xác định và duy nhất bởi điều kiện (5.2.4) và theo định lý cơ bản về sự xác định ánh xạ tuyến tính. Ta còn phải chứng tỏ $\mathscr{B}^*$ là một cơ sở của $V^*$. Vì $\mathscr{B}^*$ gồm $n$ dạng tuyến tính nên thực chất chỉ còn phải kiểm tra tính độc lập tuyến tính của $\mathscr{B}^*$ trong $V^*$. Giả sử $\theta = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i^*$ là một tổ hợp tuyến tính bằng ánh xạ không của $\mathscr{B}^*$. Khi đó với mỗi $j = 1, \ldots, n$ ta có

$$0 = \Big\langle \sum_{i=1}^{n} \alpha_i \mathbf{e}_i^*, \mathbf{e}_j \Big\rangle = \sum_{i=1}^{n} \alpha_i \langle \mathbf{e}_i^*, \mathbf{e}_j \rangle = \sum_{i=1}^{n} \alpha_i \delta_{ij} = \alpha_j.$$

Vậy $\mathscr{B}^*$ độc lập tuyến tính, do đó nó là một cơ sở của $V$. □

**5.2.5 Định lý.** *Cho $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là một cơ sở của $\mathbb{K}$-không gian vector $n$ chiều $V$ và $\mathscr{B}^* = (\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_n^*)$ là cơ sở đối ngẫu của nó trong không gian đối ngẫu $V^*$. Khi đó với mọi vector $\mathbf{x} \in V$ và mọi dạng tuyến tính $f \in V^*$ ta luôn có*

**(i)** $\mathbf{x} = \sum_{i=1}^{n} \langle \mathbf{e}_i^*, \mathbf{x} \rangle \mathbf{e}_i$, *tức là* $x|_{\mathscr{B}} = (\langle \mathbf{e}_1^*, \mathbf{x} \rangle, \langle \mathbf{e}_2^*, \mathbf{x} \rangle, \ldots, \langle \mathbf{e}_n^*, \mathbf{x} \rangle)$

**(ii)** $f = \sum_{j=1}^{n} \langle f, \mathbf{e}_j \rangle \mathbf{e}_j^*$ *tức là* $f|_{\mathscr{B}^*} = (\langle f, \mathbf{e}_1 \rangle, \langle f, \mathbf{e}_2 \rangle, \ldots, \langle f, \mathbf{e}_n \rangle)$

**(iii)** $\langle f, \mathbf{x} \rangle = \sum_{j=1}^{n} \langle f, \mathbf{e}_j \rangle \langle \mathbf{e}_j^*, \mathbf{x} \rangle$.

*Chứng minh.* Giả sử $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$ và $f = \sum_{i=1}^{n} \xi_i \mathbf{e}_i^*$. Khi đó ta có

**(i)** $\langle \mathbf{e}_i^*, \mathbf{x} \rangle = \Big\langle \mathbf{e}_i^*, \sum_{j=1}^{n} \mathbf{x}_j \mathbf{e}_j \Big\rangle = \sum_{j=1}^{n} \mathbf{x}_j \langle \mathbf{e}_i^*, \mathbf{e}_j \rangle = \sum_{j=1}^{n} x_j \delta_{ij} = x_i$ với mọi $i = 1, \ldots, n$. Do đó $\mathbf{x} = \sum_{i=1}^{n} \langle \mathbf{e}_i^*, \mathbf{x} \rangle \mathbf{e}_i$.

**(ii)** $\langle f, \mathbf{e}_j\rangle = \left\langle \sum\limits_{i=1}^{n} \xi_i \mathbf{e}_i^*, \mathbf{e}_j \right\rangle = \sum_{i=1}^{n} \xi_i \langle \mathbf{e}_i^*, \mathbf{e}_j\rangle = \sum_{i=1}^{n} \xi \delta_{ij} = \xi_j$ với mọi $j = 1, \ldots, n$. Suy ra $f = \sum\limits_{j=1}^{n} \langle f, \mathbf{e}_j\rangle \mathbf{e}_j^*$.

**(iii)** $\langle f, \mathbf{x}\rangle = \left\langle \sum\limits_{i=1}^{n} \langle f, \mathbf{e}_j\rangle \mathbf{e}_i^*, \sum\limits_{j=1}^{n} \langle \mathbf{e}_i^*, \mathbf{x}\rangle \mathbf{e}_i \right\rangle = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \langle f, \mathbf{e}_j\rangle \langle \mathbf{e}_i^*, \mathbf{x}\rangle \langle \mathbf{e}_j^*, \mathbf{e}_i\rangle = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \langle f, \mathbf{e}_j\rangle \langle \mathbf{e}_i^*, \mathbf{x}\rangle \delta_{ji} = \sum\limits_{i=1}^{n} \langle f, \mathbf{e}_i\rangle \langle \mathbf{e}_i^*, \mathbf{x}\rangle$. □

**5.2.6 Nhận xét.** Vì $V$ là $V^*$ đều $n$ chiều nên chúng đẳng cấu với nhau. Khi đã cho trong $V$ một cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ thì một đẳng cấu $\varphi : V \to V^*$ có thể cho một cách tự nhiên bởi điều kiện $\varphi(\mathbf{e}_i) = \mathbf{e}_i^*$ với $i = 1, \ldots, n$ và $\mathscr{B}^* = (\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_n^*)$ là cơ sở đối ngẫu của $\mathscr{B}$. Đáng chú ý là đẳng cấu này *không chính tắc* theo nghĩa: nếu đổi cơ sở $\mathscr{B}$ thành cơ sở $\mathscr{B}'$ trong $V$ thì nói chung $f$ không biến $\mathscr{B}'$ thành cơ sở đối ngẫu $\mathscr{B}'^*$ của $\mathscr{B}'$. Tuy nhiên có thể khắc phục điều này khi thay $V^*$ bởi *không gian đối ngẫu* $(V^*)^*$ của nó.

## Bài tập

**1 )** Trình bày chứng minh chi tiết $\mathrm{Hom}(V, W)$ là $\mathbb{K}$-không gian vector.

**2 )** Chứng minh tường minh $\mathrm{End}(V)$ là một vành.

**3 )** Cho các ánh xạ $f : \mathbb{R}^3 \to \mathbb{R}^2$ và $g : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi

$$f(x, y, z) = (2x, y + z) \qquad g(x, y, z) = (x - y, z)$$

Hãy xác định các ánh xạ $f + g$, $3f$, $2f - 5g$.

**4 )** Cho các ánh xạ $f : \mathbb{R}^3 \to \mathbb{R}^2$ và $g : \mathbb{R}^2 \to \mathbb{R}^2$ xác định bởi

$$f(x, y, z) = (2x, y + z) \qquad g(x, y) = (y, x)$$

Hãy xác định $g \circ f$.

**5)** Cho $f$ và $g$ là các toán tử tuyến tính trên $\mathbb{R}^2$ xác định bởi

$$f(x,y) = (y,x) \qquad\qquad g(x,y) = (0,x)$$

Hãy xác định các ánh xạ $f \circ g$, $g \circ f$, $f^2$, $g^2$.

**6)** Chứng minh rằng toán tử tuyến tính $f$ trên $\mathbb{R}^3$ cho dưới đây là một tự đẳng cấu. Tìm $f^{-1}$

**(a)** $f(x,y,z) = (x - 3y - 2x, y - 4z, z)$

**(b)** $f(x,y,z) = (x + z, x - z, y)$

**7)** Cho $V$ là một $\mathbb{K}$-không gian vector 2 chiều và $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2)$ là một cơ sở của $V$. Giả sử $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2)$ và $\mathscr{B}'' = (\mathbf{e}''_1, \mathbf{e}''_2)$ là hai cơ sở khác của $V$ sao cho

$$\mathbf{e}'_1 = \mathbf{e}_1 + 2\mathbf{e}_2, \mathbf{e}'_2 = 2\mathbf{e}_1 + 3\mathbf{e}_2; \qquad \mathbf{e}''_1 = 3\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}''_2 = 4\mathbf{e}_1 + 2\mathbf{e}_2$$

Cho biết ma trận của $f$ trong $\mathscr{B}'$ là $\begin{bmatrix} 3 & 8 \\ 4 & 5 \end{bmatrix}$, ma trận của $g$ trong $\mathscr{B}''$ là $\begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}$ ở đó $f$, $g$ là các toán tử tuyến tính trên $V$. Hãy tìm ma trận của $f + g$ trong cơ sở $\mathscr{B}''$.

**8)** Giả sử toán tử $f$ trên $V$ có ma trận trong một cơ sở nào đó của $V$ là $A = \begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix}$. Cho đa thức $p(x) = 3x^2 - 2x + 5$. Hãy tìm ma trận của toán tử $p(f) = 3f^2 - 2f + 5\mathrm{Id}$ trong cơ sở đó.

**9)** Cho toán tử tuyến tính $f : V \to V$ và đa thức

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n \qquad\qquad \text{với } a_0 \neq 0$$

Chứng minh rằng nếu $p(f) = a_0 \mathrm{Id}_V + a_1 f + \cdots + a_n f^n$ là toán tử không trên $V$ thì $f$ khả nghịch. Tìm $f^{-1}$.

**10 )** Cho các toán tử tuyến tính $f, g, h$ trên $\mathbb{R}^2$ xác định bởi

$$f(x,y) = (x, 2y), \qquad g(x,y) = (y, x+y), \qquad h(x,y) = (0, x).$$

Chứng minh $f, g, h$ độc lập tuyến tính trong $\operatorname{End}(\mathbb{R}^2)$.

**11 )** Cho hai ánh xạ tuyến tính $f, g : V \to V'$. Chứng minh rằng $\operatorname{rank}(f+g) \leq \operatorname{rank}(f) + \operatorname{rank}(g)$.

# § 6 Không gian song đối ngẫu và chuyển vị của ánh xạ tuyến tính

## 6.1 Không gian song đối ngẫu

**6.1.1 Định nghĩa.** Không gian đối ngẫu $(V^*)^*$ của không gian $V^*$ gọi là **không gian song đối ngẫu** của $\mathbb{K}$-không gian vector $V$, ký hiệu $V^{**}$

Khi $\dim V = n$ thì đương nhiên cả $V^*$ lẫn $V^{**}$ đều $n$ chiều. Mỗi cơ sở $\mathscr{B}$ của $V$ sinh ra cơ sở đối ngẫu $\mathscr{B}^*$ trong $V^*$. Đến lượt mình, $\mathscr{B}^*$ lại sinh ra một cơ sở đối ngẫu $(\mathscr{B}^*)^*$ trong $V^{**}$. Cơ sở $(\mathscr{B}^*)^*$ gọi là **cơ sở song đối ngẫu** của cơ sở $\mathscr{B}$, ký hiệu là $\mathscr{B}^{**}$.

**6.1.2 Định lý.** *Cho $\mathbb{K}$-không gian vector $V$ với không gian đối ngẫu $V^*$ và không gian song đối ngẫu $V^{**}$. Khi đó ta có*

**(i)** *Với mỗi $\mathbf{x} \in V$, ánh xạ $\hat{\mathbf{x}} : V^* \to \mathbb{K}$ xác định bởi $\hat{\mathbf{x}}(f) = \langle f, \mathbf{x} \rangle$ là một dạng tuyến tính trên $V^*$, tức là $\hat{\mathbf{x}} \in V^{**}$.*

**(ii)** *Tương ứng $\mathbf{x} \mapsto \hat{\mathbf{x}}$ xác định một đồng cấu từ $V$ đến $V^{**}$.*

**(iii)** *Nếu $V$ là không gian $n$ chiều thì đồng cấu $V \to V^{**}$ xác định bởi $\mathbf{x} \mapsto \hat{\mathbf{x}}$ là một đẳng cấu và gọi là **đẳng cấu chính tắc** từ $V$ lên $V^{**}$.*

*Chứng minh.*

**(i)** Với mọi $f, g \in V^*$ mọi $\lambda, \mu \in \mathbb{K}$, ta có

$$\hat{\mathbf{x}}(\lambda f + \mu g) = \langle \lambda f + \mu g, \mathbf{x} \rangle = \lambda \langle f, \mathbf{x} \rangle + \mu \langle g, \mathbf{x} \rangle = \lambda \hat{\mathbf{x}}(f) + \mu \hat{\mathbf{x}}(g).$$

Do đó $\hat{\mathbf{x}}$ là dạng tuyến tính trên $V^*$ hay $\hat{\mathbf{x}} \in V^{**}$.

**(ii)** Với mọi $\mathbf{x}, \mathbf{y} \in V$ và mọi $\lambda, \mu \in \mathbb{K}$, ta có

$$\begin{aligned}(\widehat{\lambda\mathbf{x} + \mu\mathbf{y}})(f) &= \langle f, \lambda x + \mu\mathbf{y}\rangle \\ &= \lambda\langle f, \mathbf{x}\rangle + \mu\langle f, \mathbf{y}\rangle \\ &= \lambda\hat{\mathbf{x}}(f) + \mu\hat{\mathbf{y}}(f) \\ &= (\lambda\hat{\mathbf{x}} + \mu\hat{\mathbf{y}})(f)\end{aligned}$$

với mọi $f \in V^*$. Suy ra $\widehat{\lambda\mathbf{x} + \mu\mathbf{y}} = \lambda\hat{\mathbf{x}} + \mu\hat{\mathbf{y}}$, tức là tương ứng $\mathbf{x} \mapsto \hat{\mathbf{x}}$ là một đồng cấu từ $V$ đến $V^{**}$.

**(iii)** Giả sử $V$ là không gian $n$ chiều và $\varphi : V \to V^{**}$ chính là đồng cấu cho bởi $\varphi(\mathbf{x}) = \hat{\mathbf{x}}$ với mọi $\mathbf{x} \in V$. Cho trong $V$ một cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$. Gọi $\mathscr{B}^* = (\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_n^*)$ là cơ sở đối ngẫu của $\mathscr{B}$ trong $V^*$. Khi đó với mọi $\mathbf{x} \in V$ sao cho $\mathbf{x} \in \ker\varphi$ ta có $\hat{\mathbf{x}} = \varphi(\mathbf{x}) = \theta$. Suy ra với mọi $f \in V^*$ ta có $\hat{\mathbf{x}}(f) = \langle f, \mathbf{x}\rangle = 0$. Đặc biệt, với mỗi $i = 1, \ldots, n$ ta có

$$\langle \mathbf{e}_i^*, \mathbf{x}\rangle = \hat{\mathbf{x}}(\mathbf{e}_i^*) = 0.$$

Suy ra

$$\mathbf{x} = \sum_{i=1}^{n} \langle \mathbf{e}_i^*, \mathbf{x}\rangle \mathbf{e}_i = \mathbf{0}.$$

Vậy $\ker\varphi = \{\mathbf{0}\}$. Do đó $\varphi$ là một đơn cấu. Vì $\dim V = n = \dim V^{**}$ nên $\varphi$ cũng là đẳng cấu. $\square$

**6.1.3 Nhận xét.** **(a)** Trong trường hợp tổng quát, khi $V$ là không gian vô hạn chiều, đồng cấu $\varphi : V \to V^{**}$ xác định bởi $\varphi(\mathbf{x}) = \hat{\mathbf{x}}$ cũng là một đơn cấu, nhưng nói chung không phải là toàn cấu.

**(b)** Do $\hat{\mathbf{x}} \in V^{**}$ với mọi $\mathbf{x} \in V$ nên $\hat{\mathbf{x}}(f)$ còn viết là $\langle \hat{\mathbf{x}}, f\rangle$, tức là $\langle \hat{\mathbf{x}}, f\rangle = \langle f, \mathbf{x}\rangle$ với mọi $\mathbf{x} \in V$ mọi $f \in V^*$

**(c)** Khi $\dim V = n$ chiều, đẳng cấu $\varphi : V \to V^{**}$ xác định bởi $\varphi(\mathbf{x}) = \hat{\mathbf{x}}$ là đẳng cấu *chính tắc* theo nghĩa sau đây: nó chuyển mỗi cơ sở của $V$ thành cơ sở song đối ngẫu trong $V^{**}$. Bởi vậy, $\varphi$ có tên là **đẳng**

**cấu chính tắc** từ $V$ lên $V^{**}$. Ta thường đồng nhất $V$ với $V^{**}$ bởi đẳng cấu chính tắc này; tức là đồng nhất $\mathbf{x} \in V$ với $\hat{\mathbf{x}} \in V^{**}$. Khi đó, có thể viết $\langle \mathbf{x}, f\rangle$ thay cho $\langle \hat{\mathbf{x}}, f\rangle$; đồng thời luôn có $\langle \mathbf{x}, f\rangle = \langle f, \mathbf{x}\rangle$; với mọi $\mathbf{x} \in V$ mọi $f \in V^{**}$. Nói cách khác, các vector và các dạng tuyến tính là bình đẳng.

## 6.2 Chuyển vị của một ánh xạ tuyến tính

**6.2.1 Định lý.** *Cho $V$ và $W$ là hai $\mathbb{K}$-không gian vector. Khi đó,*

**(i)** *Với mỗi $\varphi \in \mathrm{Hom}(V, W)$, tồn tại duy nhất $\varphi^t \in \mathrm{Hom}(W^*, V^*)$ sao cho*

$$\langle \varphi^t(f), \mathbf{x}\rangle = \langle f, \varphi(\mathbf{x})\rangle \qquad \text{với mọi } \mathbf{x} \in V, f \in W^*. \tag{6.2.2}$$

*Ánh xạ tuyến tính $\varphi^t$ gọi là **chuyển vị** của ánh xạ tuyến tính $\varphi$.*

**(ii)** *Tương ứng $\varphi \mapsto \varphi^t$ là một ánh xạ tuyến tính từ $\mathrm{Hom}(V, W)$ đến $\mathrm{Hom}(W^*, V^*)$ và gọi là **phép chuyển vị**.*

*Chứng minh.*

**(i)** Cho $\varphi \in \mathrm{Hom}(V, W)$. Ta định nghĩa $\varphi^t : W^* \to V^*$ bởi qui tắc $\varphi^t(f) = f \circ \varphi \in V^*$ với mọi $f \in W^*$. Khi đó $\varphi^t$ thỏa mãn hệ thức (6.2.2) vì

$$\langle \varphi^t(f), \mathbf{x}\rangle = \langle f \circ \varphi, \mathbf{x}\rangle = f \circ \varphi(\mathbf{x}) = f(\varphi(\mathbf{x})) = \langle f, \varphi(\mathbf{x})\rangle$$

với mọi $\mathbf{x} \in V$ và $f \in W^*$. Hơn nữa, với mọi $f, g \in W^*$ và mọi $\lambda, \mu \in \mathbb{K}$, ta cũng có

$$\begin{aligned}\langle \varphi^t(\lambda f + \mu g), \mathbf{x}\rangle &= \langle \lambda f + \mu g, \varphi(\mathbf{x})\rangle \\ &= \lambda\langle f, \varphi(\mathbf{x})\rangle + \mu\langle g, \varphi(\mathbf{x})\rangle \\ &= \lambda\langle \varphi^t(f), \mathbf{x}\rangle + \mu\langle \varphi^t(g), \mathbf{x}\rangle \\ &= \langle \lambda\varphi^t(f) + \mu\varphi^t(g), \mathbf{x}\rangle\end{aligned}$$

với mọi $\mathbf{x} \in V$. Vậy $\varphi^t(\lambda f + \mu g) = \lambda\varphi^t(f) + \mu\varphi^t(g)$. Do đó $\varphi^t$ là ánh xạ tuyến tính từ $W^*$ đến $V^*$. Tức là $\varphi^t \in \mathrm{Hom}(W^*, V^*)$.

Nếu $\psi \in \mathrm{Hom}(W^*, V^*)$ cũng là ánh xạ tuyến tính thỏa mãn hệ thức (6.2.2), tức là $\langle \psi(f), \mathbf{x}\rangle = \langle f, \varphi(\mathbf{x})\rangle$ với mọi $\mathbf{x} \in V$ và mọi $f \in W^*$. Suy ra $\langle \varphi^t(f), \mathbf{x}\rangle = \langle \psi(f), \mathbf{x}\rangle$ với mọi $\mathbf{x} \in V$ và mọi $f \in W^*$. Do đó ta suy ra $\varphi^t(f) = \psi(f)$ với mọi $f \in W^*$. Do đó $\varphi^t = \psi$.

**(ii)** Với mọi $\varphi, \psi \in \mathrm{Hom}(V, W)$ mọi $\alpha, \beta \in \mathbb{K}$, ta đều có

$$\begin{aligned}
\langle (\alpha\varphi + \beta\psi)^t(f), \mathbf{x}\rangle &= \langle f, (\alpha\varphi + \beta\psi)(\mathbf{x})\rangle \\
&= \langle f, \alpha\varphi(\mathbf{x}) + \beta\psi(\mathbf{x})\rangle \\
&= \alpha\langle f, \varphi(\mathbf{x})\rangle + \beta\langle f, \psi(\mathbf{x})\rangle \\
&= \alpha\langle \varphi^t(f), \mathbf{x}\rangle + \beta\langle \psi^t(f), \mathbf{x}\rangle \\
&= \langle (\alpha\varphi^t + \beta\psi^t)(f), \mathbf{x}\rangle
\end{aligned}$$

với mọi $\mathbf{x} \in V$ và mọi $f \in W^*$. Suy ra

$$(\alpha\varphi + \beta\psi)^t(f) = (\alpha\varphi^t + \beta\psi^t)(f) \qquad \text{với mọi } f \in W^*$$

tức là $(\alpha\varphi + \beta\psi)^t = \alpha\varphi^t + \beta\psi^t$. Vậy tương ứng $\varphi \mapsto \varphi^t$ là một ánh xạ tuyến tính từ $\mathrm{Hom}(V, W)$ đến $\mathrm{Hom}(W^*, V^*)$. $\square$

**6.2.3 Nhận xét.** Từ địnhg nghĩa của phép chuyển vị dễ dàng suy ra được các tính chất sau

**(i)** $(\mathrm{Id}_V)^t = \mathrm{Id}_{V^*}$ phép chuyển vị bảo toàn đơn vị của $\mathrm{End}(V)$.

**(ii)** $(\varphi \circ \psi)^t = \psi^t \circ \varphi^t$ với mọi $\varphi \in \mathrm{Hom}(W, Z)$ và mọi $\psi \in \mathrm{Hom}(V, W)$.

**(iii)** $(\varphi^{-1})^t = (\varphi^t)^{-1}$

**6.2.4 Định lý.** *Cho $V$ và $W$ tương ứng là các $\mathbb{K}$-không gian vector $n$ và $m$ chiều và $\varphi \in \mathrm{Hom}(V, W)$ là một ánh xạ tuyến tính tùy ý từ $V$ đến $W$.*

**(i)** *Giả sử $\mathscr{B}, \mathscr{C}$ lần lượt là cơ sở của $V$, $W$ và $\mathscr{B}^*$, $\mathscr{C}^*$ tương ứng là cơ sở đối ngẫu của $\mathscr{B}$, $\mathscr{C}$. Khi đó nếu ma trận của $\varphi$ trong cặp cơ sở $(\mathscr{B}, \mathscr{C})$ là $A \in \mathcal{M}_{m,n}(\mathbb{K})$ thì ma trận của $\varphi^t \in \mathrm{Hom}(W^*, V^*)$ trong cặp cơ sở $(\mathscr{C}^*, \mathscr{B}^*)$ là $A^t \in \mathcal{M}_{n,m}(\mathbb{K})$.*

**(ii)** *Nếu đồng nhất $V$ với $V^{**}$ và $W$ với $W^{**}$ bởi các đẳng cấu chính tắc thì ta có $(\varphi^t)^t = \varphi$.*

*Chứng minh.*

**(i)** Giả sử $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$, $\mathscr{C} = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$, $\mathscr{B}^* = (\mathbf{e}_1^*, \ldots, \mathbf{e}_n^*)$, $\mathscr{C}^* = (\mathbf{u}_1^*, \ldots, \mathbf{u}_m^*)$, $A = (a_{ij})_{m\times n}$. Với mọi $j = 1, \ldots, n$, ta có

$$\varphi(\mathbf{e}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{u}_i$$

và mọi $j = 1, \ldots, m$

$$\begin{aligned}
\varphi^t(\mathbf{u}_j^*) &= \sum_{i=1}^{m} \langle \varphi^t(\mathbf{u}_j^*), \mathbf{e}_i \rangle \mathbf{e}_i^* \\
&= \sum_{i=1}^{n} \langle \mathbf{u}_j^*, \varphi(\mathbf{e}_i) \rangle \mathbf{e}_i^* \\
&= \sum_{i=1}^{n} \left\langle \mathbf{u}_j^*, \sum_{k=1}^{m} a_{ki}\mathbf{u}_k \right\rangle \mathbf{e}_i^* \\
&= \sum_{i=1}^{n} \left( \sum_{k=1}^{m} a_{ki} \langle \mathbf{u}_j^*, \mathbf{u}_k \rangle \right) \mathbf{e}_i^* \\
&= \sum_{i=1}^{n} \left( \sum_{k=1}^{m} a_{ki}\delta_{jk} \right) \mathbf{e}_i^* \\
&= \sum_{i=1}^{m} a_{ji}\mathbf{e}_i^* = \sum_{i=1}^{m} b_{ij}\mathbf{e}_i^*
\end{aligned}$$

Do đó $b_{ij} = a_{ji}$ với mọi $i = 1, \ldots, n$ và $j = 1, \ldots, m$. Suy ra ma trận của $\varphi^t$ trong cặp cơ sở $(\mathscr{C}^*, \mathscr{B}^*)$ là $B = (b_{ij})_{n\times m} = A^t$.

**(ii)** Khi đã đồng nhất $V$ với $V^{**}$ và $W$ với $W^{**}$ bởi các đẳng cấu chính tắc ta sẽ có

$$\langle (\varphi^t)^t(\mathbf{x}), f \rangle = \langle \mathbf{x}, \varphi^t(f) \rangle = \langle \varphi^t(f), \mathbf{x} \rangle = \langle f, \varphi(\mathbf{x}) \rangle = \langle \varphi(\mathbf{x}), f \rangle$$

với mọi $f \in W^*$ và mọi $\mathbf{x} \in V$. Do đó, $(\varphi^t)^t(\mathbf{x}) = \varphi(\mathbf{x})$ với mọi $\mathbf{x} \in V$ hay $(\varphi^t)^t = \varphi$. □

**6.2.5 Nhận xét.** Định lý trên cho ta một ý nghĩa của phép chuyển vị các ma trận. Phép chuyển vị các ma trận tương ứng với phép chuyển vị các ánh xạ tuyến tính (trên các không gian vector hữu hạn chiều).

## Bài tập

**1**) Chứng minh Nhận xét 6.2.3.

**2**) Với mỗi không gian vector $V$ và cơ sở $\mathscr{B}$ dưới đây, xác định công thức tường minh cho các phần tử trong cơ sở đối ngẫu $\mathscr{B}^*$ cho không gian $V^*$.

**(a)** $V = \mathbb{R}^3$, $\mathscr{B} = ((1,0,1),(1,2,1),(0,0,1))$

**(b)** $V = \mathbb{R}_2[x]$, $\mathscr{B} = (1, x, x^2)$.

**3**) Cho $V = \mathbb{R}^3$ và định nghĩa $f_1, f_2, f_3 \in V^*$ như sau

$$f_1(x,y,z) = x - 2y, \quad f_2(x,y,z) = x + y + z, \quad f_3(x,y,z) = y - 3z.$$

Chứng minh rằng $(f_1, f_2, f_3)$ là một cơ sở cho $V^*$, và tìm một cơ sở cho $V$ mà nó có cơ sở đối ngẫu là $(f_1, f_2, f_3)$.

# § 7 Giá trị riêng và Vector riêng

## 7.1 Các khái niệm

**7.1.1 Định nghĩa.** Cho $f$ là một toán tử tuyến tính trên $\mathbb{K}$-không gian vector $V$. Số $\lambda \in \mathbb{K}$ được gọi là một **giá trị riêng** (eigenvalue) của $f$ nếu tồn tại vector $\mathbf{u} \neq \mathbf{0}$ sao cho $f(\mathbf{u}) = \lambda\mathbf{u}$. Khi đó, vector $\mathbf{u}$ được gọi là **vector riêng** (eigenvector) của $f$ ứng với giá trị riêng $\lambda$.

**7.1.2 Mệnh đề.** *Cho $\lambda$ là giá trị riêng của toán tử tuyến tính $f : V \to V$. Khi đó, tập $E(\lambda) = \{\mathbf{u} \in V : f(\mathbf{u}) = \lambda\mathbf{u}\}$ là một không gian con của $V$. Không gian con $E(\lambda)$ được gọi là* ***không gian con riêng*** *(eigenspace) của $f$ ứng với giá trị riêng $\lambda$.*

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

**7.1.3 Định nghĩa.** Cho ma trận vuông $A$ cấp $n$ trên $\mathbb{K}$. Số $\lambda \in \mathbb{K}$ được gọi là **giá trị riêng** của $A$ nếu tồn tại vector $\mathbf{x} \in \mathbb{K}^n \setminus \{\mathbf{0}\}$ sao cho $A\mathbf{x} = \lambda\mathbf{x}$. Khi đó, vector $\mathbf{x}$ được gọi là **vector riêng** của $A$ ứng với giá trị riêng $\lambda$.

**7.1.4 Nhận xét.** Giả sử ta có toán tử tuyến tính $f : V \to V$ trong không gian vector $n$ chiều $V$, và $A$ là ma trận của $f$ đối với cơ sở $\mathscr{B}$ nào đó của $V$. Khi đó, ta có $f(\mathbf{x}) = A\mathbf{x}$, trong đó $\mathbf{x}$ và $f(\mathbf{x})$ được hiểu là ma trận cột tọa độ của $\mathbf{x}$ và $f(\mathbf{x})$ đối với cơ sở $\mathscr{B}$. Do đó, từ các định nghĩa trên ta thấy rằng

**(i)** $\lambda$ là giá trị riêng của $f$ khi và chỉ khi $\lambda$ là giá trị riêng của $A$.

**(ii)** Vector $\mathbf{u}$ là vector riêng của $f$ ứng với giá trị riêng $\lambda$ khi và chỉ khi $\mathbf{u}$ là vector riêng của ma trận $A$ ứng với giá trị riêng $\lambda$.

Như vậy bài toán tìm vector riêng và giá trị riêng của một toán tử tuyến tính tương đương với việc tìm vector riêng và giá trị riêng của ma trận của nó trong một cơ sở nào đó. Hơn nữa, trong các phát biểu ta có thể sử dụng song song ngôn ngữ ma trận và ngôn ngữ toán tử tuyến tính.

**7.1.5 Mệnh đề.** *Cho ma trận vuông $A$ cấp $n$. Khi đó, $\lambda$ là giá trị riêng của $A$ khi và chỉ khi ma trận $A - \lambda I_n$ là ma trận suy biến.*

*Chứng minh.* Giả sử $\lambda$ là giá trị riêng của $A$. Khi đó, tồn tại vector khác không $\mathbf{u}$ sao cho $A\mathbf{u} = \lambda\mathbf{u}$ hay $(A - \lambda I)\mathbf{u} = \mathbf{0}$. Vậy hệ phương trình tuyến tính thuần nhất $(A - \lambda I)\mathbf{x} = \mathbf{0}$ có nghiệm không tầm thường, cho nên $A - \lambda I$ là một ma trận suy biến.

Ngược lại, giả sử $A - \lambda I_n$ là ma trận suy biến. Điều đó kéo theo hệ phương trình tuyến tính thuần nhất $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$ có nghiệm không tầm thường là $\mathbf{u}$. Do đó, ta có $A\mathbf{u} = \lambda\mathbf{u}$. Vậy $\lambda$ là một giá trị riêng của ma trận $A$ và có vector riêng tương ứng là $\mathbf{u}$. □

**7.1.6 Nhận xét.** Từ mệnh đề trên ta nhận thấy nếu $A$ là ma trận chéo thì các phần tử trên đường chéo là các giá trị riêng của ma trận $A$.

**7.1.7 Thí dụ. (OSV11)** Hai sinh viên $A$ và $B$ chơi trò chơi như sau: Cho một bảng vuông $n \times n$ ô với $n \geq 2$. Mỗi lượt, $A$ chọn một số nguyên điền vào vị trí $(i, j)$ nào đó (tùy chọn không lặp lại). Sau đó $B$ được quyền chỉnh sửa giá trị đó bằng cách giữ nguyên hoặc thêm bớt 1 đơn vị. Trò chơi kết thúc sau khi điền xong bảng để nhận được ma trận $X$. $B$ khẳng định rằng luôn có cách để nhận được ma trận $X$ khả nghịch và không có điểm bất động (tức là không có vector $\mathbf{v} \neq \mathbf{0}$ để $X\mathbf{v} = \mathbf{v}$). Khẳng định của $B$ đúng hay sai? Hãy chứng minh nhận định của bạn.

*Giải.* $B$ chọn $x_{ii} \equiv -1 \pmod 3$ và $x_{ij} \equiv 0 \pmod 3$ với $i \neq j$. Khi đó, ta có $X \equiv -I \pmod 3$ và $\det(X) \equiv (-1)^n \pmod 3$. Suy ra $\det(X) \neq 0$ hay $X$ khả nghịch. Nếu $X\mathbf{v} = \mathbf{v}$ thì $\det(X - I) = 0$, suy ra $\det(X - I) \equiv 0 \pmod 3$. Trong khi đó $\det(X - I) \equiv \det(-2I) \equiv 1 \pmod 3$. Ta gặp phải mâu thuẫn. Vậy ma trận $X$ không có điểm bất động. Do đó, khẳng định của $B$ là đúng. □

**7.1.8 Mệnh đề.** *Cho toán tử tuyến tính $f : V \to V$. Nếu toán tử $f$ có giá trị riêng $\lambda \neq 0$, thì không gian con riêng $E(\lambda)$ là không gian con bất biến đối với $f$. Hơn nữa, ta có $f(E(\lambda)) = E(\lambda)$.*

*Chứng minh.* Lấy $\mathbf{u} \in E(\lambda)$ tùy ý. Ta thấy $\lambda\mathbf{u} \in E(\lambda)$ và $f(\mathbf{u}) = \lambda\mathbf{u}$. Do đó, $f(\mathbf{u}) \in E(\lambda)$. Vậy $f(E(\lambda)) \subseteq E(\lambda)$. Vậy $E(\lambda)$ là không gian con bất biến đối với $f$.

Ngược lại, lấy $\mathbf{u} \in E(\lambda)$ tùy ý. Ta cũng có $\frac{1}{\lambda}\mathbf{u} \in E(\lambda)$, suy ra $f(\frac{1}{\lambda}\mathbf{u}) \in f(E(\lambda))$. Mặt khác, ta có $f(\frac{1}{\lambda}\mathbf{u}) = \mathbf{u}$. Vậy $\mathbf{u} \in f(E(\lambda))$. Do đó, $E(\lambda) \subseteq f(E(\lambda))$. Từ đó ta có điều phải chứng minh, $f(E(\lambda)) = E(\lambda)$. □

## 7.2 Đa thức đặc trưng

Cho $A$ là ma trận vuông cấp $n$ trên $\mathbb{K}$. Ta biết rằng $\lambda \in \mathbb{K}$ là giá trị riêng của $A$ khi và chỉ khi ma trận $A - \lambda I_n$ suy biến. Điều này tương đương với $\det(A - \lambda I_n) = 0$. Nghĩa là giá trị riêng của $A$ là tất cả các nghiệm của đa thức $\det(A - \lambda I_n)$ và ngược lại. Vậy để tìm giá trị riêng của ma trận $A$ ta giải phương trình $\det(A - \lambda I_n) = 0$.

**7.2.1 Thí dụ. (OSV03)** Cho ma trận $A = \begin{bmatrix} a & b & 0 & 0 \\ b & a & -1 & 0 \\ 0 & -1 & a & b \\ 0 & -1 & a & b \\ 0 & 0 & b & a \end{bmatrix}$. Với $a, b$ là các số thực, $a > |b|$, hãy chỉ ra rằng có ít nhất một giá trị riêng của $A$ là một số dương.

*Giải.* Ta tính định thức

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} a-\lambda & b & 0 & 0 \\ b & a-\lambda & -1 & 0 \\ 0 & -1 & a-\lambda & b \\ 0 & -1 & a & b \\ 0 & 0 & b & a-\lambda \end{vmatrix} \\ &= (a-\lambda)^4 - (2b^2+1)(a-\lambda)^2 + b^4 \\ &= ((a-\lambda)^2 - (a-\lambda) - b^2)((a-\lambda)^2 + (a-\lambda) - b^2) \end{aligned}$$

Do đó $\det(A - \lambda I) = 0$ khi và chỉ khi $a - \lambda = \frac{1 \pm \sqrt{4b^2+1}}{2}$ hay $a - \lambda = \frac{-1 \pm \sqrt{4b^2+1}}{2}$. Vậy giá trị riêng của $A$ là $\lambda = a + \frac{1 \pm \sqrt{4b^2+1}}{2}$ và $\lambda = a + \frac{-1 \pm \sqrt{4b^2+1}}{2}$; rõ ràng có giá trị riêng dương. □

**7.2.2 Định nghĩa.** Cho ma trận vuông $A$ cấp $n$ trên $\mathbb{K}$. Đa thức $\det(A - \lambda I_n)$ được gọi là **đa thức đặc trưng** của ma trận $A$, kí hiệu $\chi_A(\lambda)$ hay $\chi(\lambda)$. Ma trận $A - \lambda I_n$ được gọi là **ma trận đặc trưng** của $A$. Phương trình $\chi_A(\lambda) = 0$ được gọi là **phương trình đặc trưng** của $A$.

**7.2.3 Thí dụ. (OSV01)** Cho $A$ là ma trận vuông cấp $n$ có tất cả các phần tử đều là những số nguyên chẵn. Chứng minh rằng ma trận $A$ không thể có giá trị riêng là một số nguyên lẻ.

*Giải.* Ta nhận thấy đa thức đặc trưng của $A$ có các hệ số chẵn trừ hệ số của $\lambda^n$ là $(-1)^n$ (do tất cả các phần tử của $A$ nguyên chẵn). Do đó đa thức đặc trưng không thể có nghiệm nguyên chẵn. Vậy ma trận $A$ không thể có giá trị riêng là một số nguyên lẻ. □

Ta thấy rằng nếu xét trên trường số phức $\mathbb{C}$ thì đa thức đặc trưng $\chi_A(\lambda)$ luôn có $n$ nghiệm kể cả bội, nghĩa là ta có thể phân tích được

$$\chi_A(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n).$$

Hai đẳng thức liên quan đến các giá trị riêng sau có được bằng cách chúng ta đồng nhất các hệ số tương ứng của lũy thừa $\lambda$ khi khai triển định thức ở vế trái và nhân ra ở vế phải (chứng minh chi tiết dành cho bạn đọc xem như bài tập).

$$\begin{aligned}\lambda_1 + \lambda_2 + \cdots + \lambda_n &= \operatorname{tr}(A)\\ \lambda_1\lambda_2\cdots\lambda_n &= \det(A).\end{aligned}$$

**7.2.4 Thí dụ. (OSV94)** Cho $a_{ij}$ (với $i, j = 1, 2, \ldots, n$) là các số nguyên. Hãy giải hệ phương trình

$$\begin{cases} \frac{1}{2}x_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\\ \frac{1}{2}x_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{1}{2}x_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{cases}$$

*Giải.* Xét ma trận $A = (a_{ij})_{n\times n}$. Do $A$ là ma trận có các phần tử là số nguyên nên đa thức đặc trưng của nó hệ số nguyên. Do đó, giá trị riêng hữu tỉ của $A$ phải là số nguyên, cho nên nó không thể có giá trị riêng là $\frac{1}{2}$. Do đó $\det(A - \frac{1}{2}I_n) \neq 0$. Điều đó có nghĩa là hệ phương trình chỉ có một nghiệm duy nhất là $x_1 = x_2 = \cdots = x_n = 0$. □

**7.2.5 Định lý.** *Các ma trận vuông đồng dạng có cùng đa thức đặc trưng.*

*Chứng minh.* Giả sử $A$ và $B$ là hai ma trận vuông cấp $n$ đồng dạng. Khi đó, tồn tại ma trận vuông $C$ cấp $n$ không suy biến sao cho $B = C^{-1}AC$. Ta có,

$$\begin{aligned}\chi_B(\lambda) &= \det(B - \lambda I_n)\\ &= \det(C^{-1}AC - \lambda C^{-1}I_nC)\\ &= \det(C^{-1}(A - \lambda I_n)C)\end{aligned}$$

$$\begin{aligned} &= \det(C^{-1})\det(A - \lambda I_n)\det C \\ &= \det(A - \lambda I_n) \\ &= \chi_A(\lambda). \end{aligned}$$
□

**7.2.6 Thí dụ. (OSV99)** Cho $A$, $B$ là các ma trận vuông cấp $n$. Chứng minh rằng tập các giá trị riêng của $AB$ và $BA$ trùng nhau.

*Giải*. Nếu $A$ là ma trận khả nghịch thì $AB$ và $A^{-1}ABA = BA$ có cùng đa thức đặc trưng.

Nếu $A$ là ma trận suy biến, thì nó có một giá trị riêng là 0. Hơn nữa, $A$ chỉ có hữu hạn giá trị riêng (không quá $n$). Do đó trong lân cận đủ nhỏ của 0 ma trận $A$ không có giá trị riêng nào ngoài 0. Đặt $A_k = A - \frac{1}{k}I$. Khi đó với $k$ đủ lớn $A_k$ không suy biến (vì $\frac{1}{k}$ không là giá trị riêng của $A$). Do đó theo kết quả trên $A_kB$ và $BA_k$ có cùng đa thức đặt trưng; nghĩa là

$$\det(A_kB - \lambda I) = \det(BA_k - \lambda I) \qquad \text{với mọi } k \text{ đủ lớn}$$

Lấy giới hạn hai vế khi $k \to \infty$ ta được $\det(AB - \lambda I) = \det(BA - \lambda I)$.□

**7.2.7 Thí dụ. (OSV03)** Biết rằng mọi ma trận thực đối xứng đều có các giá trị riêng là số thực. Chứng minh rằng nếu $\alpha, \beta, \gamma$ là các số thực khác không và $a, b, c, d, p, q$ là các số thực tùy ý, thì ma trận $B = \begin{bmatrix} a & b\frac{\alpha}{\beta} & c\frac{\alpha}{\gamma} \\ b\frac{\beta}{\alpha} & d & p\frac{\beta}{\gamma} \\ c\frac{\gamma}{\alpha} & p\frac{\gamma}{\beta} & q \end{bmatrix}$ cũng có các giá trị riêng là số thực thực.

*Giải*. Đặt $C = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix}$. Khi đó ta tính được

$$C^{-1}BC = \begin{bmatrix} a & b & c \\ b & d & p \\ c & p & q \end{bmatrix} = A.$$

Vậy hai ma trận $A$ và $B$ có cùng giá trị riêng mà $A$ là ma trận đối xứng, cho nên $B$ có các giá trị riêng là số thực. □

**7.2.8 Thí dụ. (OSV04)** Cho các ma trận

$$A = \begin{bmatrix} -1 & -3 & 0 \\ -3 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \qquad T = \begin{bmatrix} -1 & -3 & -3 \\ 0 & -2 & 5 \\ -3 & 1 & 1 \end{bmatrix}$$

**(a)** Tính $B = T^{-1}AT$.

**(b)** Tìm giá trị riêng và vector riêng của ma trận $A$.

*Giải.* **(a)** Tính toán trực tiếp ta được

$$T^{-1} = \frac{1}{70}\begin{bmatrix} -7 & 0 & -21 \\ -15 & -10 & 5 \\ -6 & 10 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

**(b)** Vì $B$ là ma trận chéo nên nó có các giá trị riêng là $-1$, $-3$, $4$. Theo định lý trên $A$ cũng nhận các giá trị ấy làm giá trị riêng. Để tìm không gian con riêng của ma trận $A$ ứng với giá trị riêng $\lambda$ ta tìm không gian nghiệm của phương trình $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Trình bày chi tiết việc tìm không gian con riêng ở phần cuối của mục này. Ta có thể kiểm tra được rằng một vector riêng ứng với giá trị riêng $-1$ là $(-1, 0, -3)$; một vector riêng ứng với giá trị riêng $-3$ là $(-3, -2, 1)$; một vector riêng ứng với giá trị riêng $4$ là $(-3, 5, 1)$. □

**7.2.9 Nhận xét.** Ta thấy rằng việc xác định đa thức đặc trưng của ma trận chéo thật sự là rất đơn giản. Nếu như trong chứng minh định lý trên ma trận $B$ là ma trận chéo thì ta nói rằng ma trận $A$ chéo hóa được bởi ma trận $C$. Việc xác định một ma trận vuông nào đó có chéo hóa được không và chéo hóa bởi ma trận nào là những vấn đề đặc biệt quan trọng của đại số tuyến tính.

Từ định lý trên ta nhận thấy việc tìm giá trị riêng của một toán tử tuyến tính trên không gian vector hữu hạn chiều đưa về việc tìm giá trị riêng của ma trận của nó trong một cơ sở bất kỳ nào. Định lý sau cho ta một ứng dụng của giá trị riêng và vector riêng.

**7.2.10 Định lý.** **(i)** *Nếu $V$ là một $\mathbb{C}$-không gian vector $n$ chiều thì mọi phép biến đổi tuyến tính của $V$ đều có ít nhất một không gian con bất biến một chiều.*

**(ii)** *Nếu $V$ là một $\mathbb{R}$-không gian vector $n$ chiều thì mọi phép biến đổi tuyến tính của $V$ đều có ít nhất một không gian con bất biến hoặc một chiều hoặc hai chiều.*

*Chứng minh.* Cho $\mathscr{B}$ là một cở sở của không gian vector $V$ và $f$ là một phép biến đổi tuyến tính trên $V$. Gọi $A$ là ma trận của $f$ đối với cơ sở $\mathscr{B}$.

**(i)** $A$ là một ma trận trên $\mathbb{C}$. Khi đó đa thức đặc trưng của $A$, $\chi_A(\lambda) = \det(A-\lambda I_n)$, là đa thức bậc $n$ trên $\mathbb{C}$. Khi đó phương trình $\chi_A(\lambda) = 0$ luôn có nghiệm trong $\mathbb{C}$, nghĩa là $A$ có một giá trị riêng $\lambda_0$ nào đó. Gọi $\mathbf{a} \in V$ là vector riêng ứng với giá trị riêng $\lambda_0$, nghĩa là $f(\mathbf{a}) = \lambda_0\mathbf{a}$. Với mọi $\mathbf{x} \in \langle\mathbf{a}\rangle$, tồn tại $k \in \mathbb{C}$ sao cho $\mathbf{x} = k\mathbf{a}$. Do đó $f(\mathbf{x}) = f(k\mathbf{a}) = k\lambda\mathbf{a} \in \langle\mathbf{a}\rangle$. Vậy $f(\langle\mathbf{a}\rangle) \subseteq \langle\mathbf{a}\rangle$ hay $\langle\mathbf{a}\rangle$ là một không gian con bất biến của $f$.

**(ii)** $A$ là một ma trận trên $\mathbb{R}$. Khi đó đa thức đặc trưng của $A$, $\chi_A(\lambda) = \det(A - \lambda I_n)$, là đa thức bậc $n$ trên $\mathbb{R}$. Khi đó phương trình đặc trưng $\chi_A(\lambda) = 0$ có thể có nghiệm thực hoặc không có nghiệm thực nào. Trường hợp $\chi_A(\lambda) = 0$ có một nghiệm thực $\lambda_0$. Khi đó $\lambda_0$ là một giá trị riêng của $f$ cho nên nó có một vector riêng $\mathbf{a} \in V$ nào đó. Cũng như trong trường hợp trên ta được $\langle\mathbf{a}\rangle$ là một không gian con bất biến của $f$.

Trường hợp $\chi_A(\lambda) = 0$ không có nghiệm thực. Tuy nhiên nó phải có một nghiệm phức $\lambda_0 = u + iv$ nào đó. Chúng ta xem $A$ là một ma trận trên $\mathbb{C}$. Khi đó $\lambda_0$ là một giá trị riêng của $A$, cho nên nó có một vector riêng $\mathbf{c} \in \mathbb{C}^n$ của $A$. Ta viết lại $\mathbf{c} = \mathbf{a} + i\mathbf{b}$ với $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. Khi đó ta có

$$A\mathbf{c} = \lambda_0\mathbf{c}$$

hay

$$A\mathbf{a} + iA\mathbf{b} = (u + iv)(\mathbf{a} + i\mathbf{b}) = u\mathbf{a} - v\mathbf{b} + i(v\mathbf{a} + u\mathbf{b}).$$

Do đó

$$A\mathbf{a} = u\mathbf{a} - v\mathbf{b} \qquad A\mathbf{b} = v\mathbf{a} + u\mathbf{b}.$$

Giả sử $\mathbf{a}$ và $\mathbf{b}$ phụ thuộc tuyến tính trong $\mathbb{R}^n$. Khi đó tồn tại $\alpha$ và $\beta$ không đồng thời bằng không sao cho $\alpha\mathbf{a} + \beta\mathbf{b} = \mathbf{0}$. Nếu $\alpha = 0$ thì $\mathbf{b} = \mathbf{0}$ (do $\beta \neq 0$), suy ra $A\mathbf{a} = u\mathbf{a}$. Đây là điều vô lý vì $u$ là giá trị riêng thực của $A$. Vậy $\alpha \neq 0$, cho nên ta có thể viết $\mathbf{a} = k\mathbf{b}$ với $k = -\beta/\alpha$. Khi đó, ta có

$$(uk-v)\mathbf{b} = u\mathbf{a}-v\mathbf{b} = A\mathbf{a} = A(k\mathbf{b}) = k(v\mathbf{a}+u\mathbf{b}) = (k^2v+ku)\mathbf{b}.$$

Do đó $uk - v = k^2v + ku$ hay $(k^2+1)v = 0$, cho nên $v = 0$ trái với giả thiết $\lambda_0 = u$ là một số phức. Vậy $\mathbf{a}$ và $\mathbf{b}$ là độc lập tuyến tính. Gọi $\mathbf{x}, \mathbf{y} \in V$ sao cho $\mathbf{x}|_{\mathscr{B}} = \mathbf{a}$ và $\mathbf{y}|_{\mathscr{B}} = \mathbf{b}$. Khi đó ta có

$$\begin{aligned} f(\mathbf{x})|_{\mathscr{B}} &= A\mathbf{a} = u\mathbf{a} - v\mathbf{b} = u\mathbf{x}|_{\mathscr{B}} - v\mathbf{y}|_{\mathscr{B}} = (u\mathbf{x} - v\mathbf{y})|_{\mathscr{B}} \\ f(\mathbf{y})|_{\mathscr{B}} &= A\mathbf{b} = v\mathbf{a} + u\mathbf{b} = v\mathbf{x}|_{\mathscr{B}} + u\mathbf{y}|_{\mathscr{B}} = (v\mathbf{x} + u\mathbf{y})|_{\mathscr{B}}. \end{aligned}$$

Do đó $f(\mathbf{x}) = u\mathbf{x} - v\mathbf{y}$ và $f(\mathbf{y}) = v\mathbf{x} + u\mathbf{y}$, cho nên $f(\mathbf{x}), f(\mathbf{y}) \in \langle \mathbf{x}, \mathbf{y} \rangle$. Vậy $f(\langle \mathbf{x}, \mathbf{y} \rangle) \subseteq \langle \mathbf{x}, \mathbf{y} \rangle$ hay $\langle \mathbf{x}, \mathbf{y} \rangle$ là một không gian con bất biến của $f$. $\square$

## 7.3 Thuật toán tìm giá trị riêng và vector riêng của ma trận

**7.3.1 Nội dung thuật toán** Cho $A$ là ma trận vuông cấp $n$ trên $\mathbb{K}$. Ta tìm giá trị riêng của $A$ và không gian con riêng tương ứng như sau:

**(i)** Lập đa thức đặc trưng $\chi_A(\lambda) = \det(A - \lambda I_n)$.

**(ii)** Giải phương trình đặc trưng $\chi_A(\lambda) = 0$ trên $\mathbb{K}$ (mỗi nghiệm của phương trình là một giá trị riêng của $A$).

**(iii)** Với mỗi giá trị riêng $\lambda$, tìm không gian không của ma trận $A - \lambda I_n$, nó cũng là không gian nghiệm của hệ phương trình $(A - \lambda I_n)\mathbf{x} = \mathbf{0}$. Đây chính là không gian con riêng ứng với giá trị riêng $\lambda$, $E(\lambda)$. Hơn nữa, mỗi vector khác không trong $E(\lambda)$ là vector riêng ứng với giá trị riêng $\lambda$.

**7.3.2 Thí dụ.** Tìm các giá trị riêng và không gian con riêng tương ứng của ma trận $A = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ trên $\mathbb{R}$.

Ta có đa thức đặc trưng của ma trận $A$ là

$$\begin{aligned}\chi_A(\lambda) &= \begin{vmatrix} 3-\lambda & -2 & 0 \\ -2 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix} \\ &= (5-\lambda)((3-\lambda)^2 - 4) \\ &= (5-\lambda)^2(1-\lambda).\end{aligned}$$

Do đó, ma trận $A$ có hai giá trị riêng $\lambda = 1$ và $\lambda = 5$. Với $\lambda = 1$, ta biến đổi ma trận

$$A - I_3 = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \xrightarrow[d_1 \to \frac{1}{2}d_1]{d_2 \to d_2 + d_1} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$
$$\xrightarrow[d_2 \leftrightarrow d_3]{d_3 \to \frac{1}{4}d_3} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó, $\mathbf{x} \in E(1)$ khi và chỉ khi $\mathbf{x} = \begin{bmatrix} x_2 \\ x_2 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. Vậy không gian con riêng ứng với giá trị riêng $\lambda = 1$ là $E(1) = \langle (1, 1, 0) \rangle$.

Với $\lambda = 5$, ta biến đổi ma trận

$$A - 5I_3 = \begin{bmatrix} -2 & -2 & 0 \\ -2 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow[d_1 \to -\frac{1}{2}d_1]{d_2 \to d_2 + d_1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó, $\mathbf{x} \in E(5)$ khi và chỉ khi $\mathbf{x} = \begin{bmatrix} -x_2 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ với $x_2, x_3 \in \mathbb{R}$ tùy ý. Vậy không gian con riêng ứng với giá trị riêng $\lambda = 5$ là $E(5) = \langle (-1, 1, 0), (0, 0, 1) \rangle$. □

**7.3.3 Thí dụ.** Cho $f$ là một toán tử tuyến tính trên $\mathbb{R}^3$ có ma trận trong cơ sở $\mathscr{B} = (\mathbf{b}_1 = (1,1,1), \mathbf{b}_2 = (1,1,0), \mathbf{b}_3 = (1,0,0))$ là

$$A = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

*Giải.* Trong Thí dụ 7.3.2 ta đã biết ma trận $A$ có giá trị riêng $\lambda = 1$ với không gian con riêng $E(1) = \langle (1,1,0) \rangle$ và giá trị riêng $\lambda = 5$ với không gian con riêng $E(5) = \langle (-1,1,0), (0,0,1) \rangle$. Do đó, $f$ có một giá trị riêng là $\lambda = 1$ và không gian con của nó sinh bởi vector $\mathbf{u}_1|_{\mathscr{B}} = (1,1,0)$; và $f$ có một giá trị riêng $\lambda = 5$ và không gian con của nó sinh bởi hai vector $\mathbf{u}_2|_{\mathscr{B}} = (-1,1,0)$ và $\mathbf{u}_3|_{\mathscr{B}} = (0,0,1)$.

Mặt khác, ta có ma trận đổi cơ sở từ $\mathscr{E}_3$ sang $\mathscr{B}$ là $B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.
Từ đó ta tìm được $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ như sau

$$\mathbf{u}_1 = B\mathbf{u}_1|_{\mathscr{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = B\mathbf{u}_2|_{\mathscr{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\mathbf{u}_3 = B\mathbf{u}_3|_{\mathscr{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Vậy $f$ có không gian con riêng ứng giá trị riêng $\lambda = 1$ là $\langle (2,2,1) \rangle$; và không gian con riêng ứng với giá trị riêng $\lambda = 5$ là $\langle (0,0,-1), (1,0,0) \rangle$. $\square$

**7.3.4 Thí dụ.** Tìm giá trị riêng và không gian con riêng tương ứng của ma trận $A = \begin{bmatrix} 3 & 3 & 2 \\ 1 & 1 & -2 \\ -3 & -1 & 0 \end{bmatrix}$ trên $\mathbb{C}$.

Đa thức đặc trưng của $A$ là

$$\begin{aligned}
\chi_A(\lambda) &= \begin{vmatrix} 3-\lambda & 3 & 2 \\ 1 & 1-\lambda & -2 \\ -3 & -1 & -\lambda \end{vmatrix} \\
&= -\begin{vmatrix} 1 & 1-\lambda & -2 \\ 3-\lambda & 3 & 2 \\ -3 & -1 & -\lambda \end{vmatrix} \\
&= -\begin{vmatrix} 1 & 1-\lambda & -2 \\ 0 & 4\lambda-\lambda^2 & 8-2\lambda \\ 0 & 2-3\lambda & -6-\lambda \end{vmatrix} \\
&= (4\lambda-\lambda^2)(6+\lambda)+(2-3\lambda)(8-2\lambda) \\
&= (4-\lambda)(\lambda^2+4)
\end{aligned}$$

Do đó ma trận $A$ có các giá trị riêng $\lambda = 4$, $\lambda = 2i$, và $\lambda = -2i$.

Với $\lambda = 4$, ta biến đổi ma trận

$$\begin{aligned}
A - 4I_3 = \begin{bmatrix} -1 & 3 & 2 \\ 1 & -3 & -2 \\ -3 & -1 & -4 \end{bmatrix} &\xrightarrow[d_3\to d_3-3d_1]{d_2\to d_2+d_1} \begin{bmatrix} 1 & -3 & -2 \\ 0 & 0 & 0 \\ 0 & -10 & -10 \end{bmatrix} \\
&\xrightarrow[\substack{d_2\leftrightarrow d_3 \\ d_1\to d_1+3d_2}]{d_3\to -\frac{1}{10}d_3} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned}$$

Do đó, $\mathbf{x} \in E(4)$ khi và chỉ khi $\mathbf{x} = \begin{bmatrix} -x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$. Vậy không gian con riêng ứng với giá trị riêng $\lambda = 4$ là $E(4) = \langle(-1,-1,1)\rangle$.

Với $\lambda = 2i$, ta biến đổi ma trận

$$A - 2iI_3 = \begin{bmatrix} 3-2i & 3 & 2 \\ 1 & 1-2i & -2 \\ -3 & -1 & -2i \end{bmatrix}$$

$$\xrightarrow[\substack{d_2\to d_2-(3-2i)d_1\\ d_3\to d_3+3d_1}]{d_2\leftrightarrow d_1}\begin{bmatrix}1 & 1-2i & -2\\ 0 & 4+8i & 8-4i\\ 0 & 2-6i & -6-2i\end{bmatrix}$$

$$\xrightarrow[\substack{d_1\to d_1-(1-2i)d_2\\ d_3\to d_3-(2-6i)d_3}]{d_2\to\frac{1}{4+8i}d_2}\begin{bmatrix}1 & 0 & i\\ 0 & 1 & -i\\ 0 & 0 & 0\end{bmatrix}$$

Do đó, $\mathbf{x}\in E(2i)$ khi và chỉ khi $\mathbf{x}=\begin{bmatrix}-ix_3\\ ix_3\\ x_3\end{bmatrix}=x_3\begin{bmatrix}-i\\ i\\ 1\end{bmatrix}$. Vậy không gian con riêng ứng với giá trị riêng $\lambda=2i$ là $E(2i)=\langle(-i,i,1)\rangle$.

Với $\lambda=-2i$, ta biến đổi ma trận

$$A+2iI_3=\begin{bmatrix}3+2i & 3 & 2\\ 1 & 1+2i & -2\\ -3 & -1 & 2i\end{bmatrix}$$

$$\xrightarrow[\substack{d_2\to d_2-(3+2i)d_1\\ d_3\to d_3+3d_1}]{d_2\leftrightarrow d_1}\begin{bmatrix}1 & 1+2i & -2\\ 0 & 4-8i & 8+4i\\ 0 & 2+6i & -6+2i\end{bmatrix}$$

$$\xrightarrow[\substack{d_1\to d_1-(1+2i)d_2\\ d_3\to d_3-(2+6i)d_3}]{d_2\to\frac{1}{4-8i}d_2}\begin{bmatrix}1 & 0 & -i\\ 0 & 1 & i\\ 0 & 0 & 0\end{bmatrix}$$

Do đó, $\mathbf{x}\in E(-2i)$ khi và chỉ khi $\mathbf{x}=\begin{bmatrix}ix_3\\ -ix_3\\ x_3\end{bmatrix}=x_3\begin{bmatrix}i\\ -i\\ 1\end{bmatrix}$. Vậy không gian con riêng ứng với giá trị riêng $\lambda=-2i$ là $E(-2i)=\langle(i,-i,1)\rangle$. $\square$

## Bài tập

**1**) Chứng minh Mệnh đề 7.1.2.

**2**) Chứng minh mỗi vector riêng chỉ ứng với một giá trị riêng duy nhất.

**3**) Chứng minh rằng một ma trận vuông có các phần tử nguyên thì không thể có giá trị riêng là $\frac{1}{2}$.

**4**) Chứng minh rằng nếu $\lambda$ là một giá trị riêng của ma trận vuông $A$ thì $\lambda^k$ là một giá trị riêng của $A^k$.

**5**) Chứng minh rằng nếu $\lambda \neq 0$ là một giá trị riêng của ma trận vuông khả nghịch $A$ thì $\frac{1}{\lambda}$ là một giá trị riêng của $A^{-1}$.

**6**) Tìm giá trị riêng và một cơ sở của không gian con riêng của toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$, biết ma trận của $f$ trong cơ sở chính tắc là

**(a)** $\begin{bmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{bmatrix}$ **(c)** $\begin{bmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{bmatrix}$

**(d)** $\begin{bmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{bmatrix}$ **(e)** $\begin{bmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 8 \end{bmatrix}$ **(f)** $\begin{bmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{bmatrix}$

**(g)** $\begin{bmatrix} 1 & -4 & -8 \\ -4 & 7 & -4 \\ -8 & -4 & 1 \end{bmatrix}$ **(h)** $\begin{bmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{bmatrix}$ **(i)** $\begin{bmatrix} 15 & -18 & -16 \\ 9 & -12 & -8 \\ 4 & -4 & -6 \end{bmatrix}$

**7**) Tìm các giá trị riêng và vector riêng độc lập tuyến tính của các ma trận phức sau

**(a)** $\begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & i \\ 0 & i \end{bmatrix}$ **(c)** $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & -3i \\ i & -1 \end{bmatrix}$

**8**) Cho toán tử tuyến tính $f : \mathbb{K} \to \mathbb{K}$ có ma trận trong cơ sở chính tắc là

$$\begin{bmatrix} 4 & -5 & 7 \\ 1 & -4 & 9 \\ -4 & 0 & 5 \end{bmatrix}$$

Tìm giá trị riêng và cơ sở của không gian con riêng trong trường hợp $\mathbb{K} = \mathbb{R}$ và $\mathbb{K} = \mathbb{C}$.

**9**) Giả sử toán tử tuyến tính $f$ trên $\mathbb{R}$-không gian vector $n$ chiều $V$ có ma trận đối với một cơ sở nào đó là $A$. Chứng minh rằng

**(a)** Đa thức đặc trưng của $f$ có dạng

$$\chi_A(\lambda) = (-\lambda)^n + c_1(-\lambda)^{n-1} + c_2(-\lambda)^{n-2} + \cdots + c_n$$

trong đó $c_k$ là tổng của các định thức con cấp $k$ của $A$ mà chỉ số dòng và chỉ số cột trùng nhau.

**(b)** Tổng tất cả các giá trị riêng của $f$ bằng vết của ma trận $A$ còn tích của tất cả các giá trị riêng của $f$ bằng định thức của ma trận $A$.

**10 )** Cho $A$ là ma trận vuông cấp $n$. Chứng minh rằng

**(a)** Tồn tại một và chỉ một đa thức với hệ số cao nhất bằng 1 nhận $A$ làm nghiệm và có bậc thấp nhất so với các đa thức nhận $A$ làm nghiệm (gọi là **đa thức tối tiểu** của ma trận $A$).

**(b)** Các ma trận đồng dạng có cùng một đa thức tối tiểu.

**11 )** Cho ma trận $A = \begin{bmatrix} b_1 & c_1 & 0 & 0 & \dots & 0 & 0 \\ a_1 & b_2 & c_2 & 0 & \dots & 0 & 0 \\ 0 & a_2 & b_3 & c_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & a_{n-1} & b_n \end{bmatrix}$ với $a_i, b_i, c_i \in \mathbb{R}$ và $a_i \neq 0$, $b_i \neq 0$, $c_i \neq 0$, $a_{i+1}c_i > 0$ với mọi $i$. Hãy chứng minh rằng các giá trị riêng của $A$ là số thực.

# § 8 Tính chất của đa thức đặc trưng

## 8.1 Đa thức đặc trưng và không gian con bất biến

**8.1.1 Định lý.** *Cho $f$ là một toán tử tuyến tính trên một không gian vector hữu hạn chiều $V$, và cho $W$ là một không gian con $f$-bất biến của $V$. Khi đó đa thức đặc trưng của toán tử tuyến tính $f|_W : W \to W$ chia hết đa thức đặc trưng của $f$.*

*Chứng minh.* Chọn một cơ sở $\mathscr{B}_W = (\mathbf{v}_1, \dots, \mathbf{v}_k)$ cho $W$ và mở rộng nó trên một cơ sở $\mathscr{B} = (\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n)$ cho $V$. Gọi $A$ là ma trận của $f$ đối với cơ sở $\mathscr{B}$ và $B_1$ là ma trận của $f|_W$ đối với cơ sở $\mathscr{B}_W$. Khi đó, ta có

$$A = \begin{bmatrix} B_1 & B_2 \\ O & B_3 \end{bmatrix},$$

ở đây $O$ là ma trận không cấp $(n-k)\times k$, $B_1$ và $B_2$ là các ma trận có cấp thích hợp. Gọi $\chi(\lambda)$ là đa thức đặc trưng của $f$ và $\chi_W(\lambda)$ là đa thức đặc trưng của $f|_W$. Khi đó

$$\begin{aligned} \chi(\lambda) = \det(A - \lambda I_n) &= \det \begin{bmatrix} B_1 - \lambda I_k & B_2 \\ O & B_3 - \lambda I_{n-k} \end{bmatrix} \\ &= \det(B_1 - \lambda I_k)\det(B_3 - \lambda I_{n-k}) \\ &= \chi_W(\lambda)\cdot \det(B_3 - tI_{n-k}) \end{aligned}$$

Vì vậy $\chi_W(\lambda)$ chia hết $\chi(\lambda)$. □

**8.1.2 Thí dụ.** Cho $f$ là toán tử tuyến tính trên $\mathbb{R}^4$ định nghĩa bởi

$$f(a,b,c,d) = (a+b+2c-d, b+d, 2c-d, c+d),$$

và cho $W = \{(t,s,0,0) : t,s \in \mathbb{R}\}$. Chú ý rằng $W$ là một không gian con $f$-bất biến của $\mathbb{R}^4$ bởi vì với bất kỳ vector $(a,b,0,0) \in \mathbb{R}^4$,

$$f(a,b,0,0) = (a+b,b,0,0) \in W.$$

Chọn $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2)$ là một cơ sở cho $W$. Mở rộng $\mathscr{B}$ thành cơ sở chuẩn $\mathscr{E}$ cho $\mathbb{R}^4$. Khi đó

$$B_1 = f|_W|_{\mathscr{B}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \text{và} \qquad A = f|_{\mathscr{E}} = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Cho $\chi(\lambda)$ là đa thức đặc trưng của $f$ và $\chi_W(\lambda)$ là đa thức đặc trưng của

$f_W$. Khi đó

$$\chi(\lambda) = \det(A - \lambda I_4) = \begin{vmatrix} 1-\lambda & 1 & 2 & -1 \\ 0 & 1-\lambda & 0 & 1 \\ 0 & 0 & 2-\lambda & -1 \\ 0 & 0 & 1 & 1-\lambda \end{vmatrix}$$

$$= \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} \cdot \begin{vmatrix} 2-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix}$$

$$= \chi_W(\lambda) \begin{vmatrix} 2-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix}.$$

□

Theo định lý 8.1.1 chúng ta có thể sử dụng đa thức đặc trưng của $f_W$ để biết thêm thông tin về đa thức đặc trưng của $f$. Lưu ý, không gian con cyclic thì rất tiện ích vì đa thức đặc trưng của sự thu hẹp một toán tử tuyến tính $f$ vào một không gian con cyclic thì không khó để tính.

**8.1.3 Định lý.** *Cho $f$ là một toán tử tuyến tính trên một không gian vector hữu hạn chiều $V$, và cho $W$ kí hiệu không gian con $f$-cyclic của $V$ tạo bởi một vector $\mathbf{v} \neq \mathbf{0}$, $\mathbf{v} \in V$. Cho $k = \dim(W)$. Khi đó*

**(a)** $(\mathbf{v}, f(\mathbf{v}), f^2(\mathbf{v}), \dots, f^{k-1}(\mathbf{v}))$ *là một cơ sở trong $W$.*

**(b)** *Nếu $a_0\mathbf{v} + a_1 f(\mathbf{v}) + \cdots + a_{k-1} f^{k-1}(\mathbf{v}) + f^k(\mathbf{v}) = 0$, thì đa thức đặc trưng của $f_W$ là $\chi_W(\lambda) = (-1)^k(a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k)$.*

*Chứng minh.*

**(a)** Do $\mathbf{v} \neq \mathbf{0}$, tập $(\mathbf{v})$ là độc lập tuyến tính. Cho $j$ là số nguyên dương lớn nhất mà

$$\mathscr{B} = (\mathbf{v}, f(\mathbf{v}), \dots, f^{j-1}(\mathbf{v}))$$

vẫn độc lập tuyến tính. Vậy $j$ bao giờ cũng tồn tại bởi vì $V$ hữu hạn chiều. Gọi $U$ là không gian con sinh bởi $\mathscr{B}$. Khi đó $\mathscr{B}$ là một cơ sở trong $U$. Ngoài ra, $f^j(\mathbf{v}) \in U$ do định nghĩa của $j$ và Định lý 2.2.6 (trang 216). Chúng ta dùng điều này để chỉ ra rằng $U$ là một không gian con $f$-bất biến của $V$. Cho $\mathbf{w} \in U$. Do $\mathbf{w}$ là một tổ hợp tuyến tính của những phần tử của $\mathscr{B}$ nên tồn tại những vô hướng $b_0, b_1, \dots, b_{j-1}$ sao cho

$$\mathbf{w} = b_0\mathbf{v} + b_1 f(\mathbf{v}) + \cdots + b_{j-1} f^{j-1}(\mathbf{v}),$$

và do đó

$$f(\mathbf{w}) = b_0 f(\mathbf{v}) + b_1 f^2(\mathbf{v}) + \cdots + b_{j-1} f^j(\mathbf{v}).$$

Vì vậy $f(\mathbf{w})$ là một tổ hợp tuyến tính của những phần tử của $U$, và do đó thuộc vào $U$. Vì vậy $U$ là một $f$-bất biến. Ngoài ra, $\mathbf{v} \in U$. Do $W$ là không gian con $f$-bất biến nhỏ nhất của $V$ mà chứa $\mathbf{v}$, và do đó $W \subseteq U$. Rõ ràng $U \subseteq W$ (vì $\mathscr{B} \subset W$), và do đó chúng ta kết luận rằng $U = W$. Từ đó suy ra $\mathscr{B}$ là một cơ sở trong $W$, và do đó $\dim(W) = j$. Vậy $j = k$. Điều đó chứng minh (a).

**(b)** Bây giờ xem $\mathscr{B}$ (từ (a)) như một cơ sở có thứ tự trong $W$. Cho $a_0, a_1, \dots, a_{k-1}$ là những vô hướng sao cho

$$a_0\mathbf{v} + a_1 f(\mathbf{v}) + \cdots + a_{k-1} f^{k-1}(\mathbf{v}) + f^k(\mathbf{v}) = 0.$$

Chú ý rằng

$$f_W|_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{k-1} \end{bmatrix},$$

nó có đa thức đặc trưng

$$\chi_W(\lambda) = (-1)^k (a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k)$$

Vì vậy $\chi_W(\lambda)$ là đa thức đặc trưng của $f_W$, chứng minh xong (b).□

**8.1.4 Thí dụ.** Toán tử tuyến tính trên $\mathbb{R}$ xác định bởi $f(a, b, c) = (-b + c, a + c, 3c)$ có $W = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle$ là không gian con $f$-cyclic tạo bởi $\mathbf{e}_1$. Chúng ta tính đa thức đặc trưng $\chi(\lambda)$ của $f_W$ bằng hai cách: bằng định lý 8.1.3 và bằng phương pháp tính định thức.

**(a)** *Bằng định lý 8.1.3*. Chúng ta có $(\mathbf{e}_1, \mathbf{e}_2)$ là cơ sở của $W$ và $f^2(\mathbf{e}_1) = -\mathbf{e}_1$. Do đó, $\chi(\lambda) = \lambda^2 + 1$ bởi định lý 8.1.3(b).

**(b)** *Bằng phương pháp tính định thức*. Cho $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2)$, nó là một cơ sở có thứ tự trong $W$. Do $f(\mathbf{e}_1) = \mathbf{e}_2$ và $f(\mathbf{e}_2) = -\mathbf{e}_1$, chúng ta có

$$f_W|_{\mathscr{B}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

và do đó,

$$\chi(\lambda) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1. \qquad \square$$

**8.1.5 Định lý.** *Cho $f$ là một toán tử tuyến tính trên một không gian hữu hạn chiều $V$, và giả sử rằng $V = W_1 \oplus \cdots \oplus W_k$, ở đây $W_i$ là các không gian con $f$-bất biến của $V$ với $i = 1, 2, \ldots, k$). Giả sử rằng $\chi(\lambda)$ là đa thức đặc trưng của $f$ và $\chi_i(\lambda)$ là đa thức đặc trưng của $f|_{W_i}$ với $1 \leq i \leq k$. Khi đó*

$$\chi(\lambda) = \chi_1(\lambda) \cdot \chi_2(\lambda) \cdots \chi_k(\lambda).$$

*Chứng minh.* Chứng minh bằng phép quy nạp theo $k$. Đầu tiên giả sử rằng $k = 2$. Cho $\mathscr{B}_1$ là một cơ sở cho $W_1$, $\mathscr{B}_2$ là cơ sở cho $W_2$. Khi đó $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2$ là một cơ sở cho $V$ (Hệ quả 7.1.4 trang 261). Đặt $A = f_{\mathscr{B}}$, $B_1 = f|_{W_1}\big|_{\mathscr{B}_1}$, và $B_2 = f|_{W_2}\big|_{\mathscr{B}_2}$. Từ đó suy ra rằng

$$A = \begin{pmatrix} B_1 & O \\ O' & B_2 \end{pmatrix}$$

ở đây $O$ và $O'$ là những ma trận không có cấp thích hợp. Vì vậy

$$\chi(\lambda) = \det(A - \lambda I) = \det(B_1 - \lambda I) \cdot \det(B_2 - \lambda I) = \chi_1(\lambda) \cdot \chi_2(\lambda);$$

kết quả đã chứng minh cho $k = 2$.

Bây giờ giả định rằng định lý đúng cho số hạng $k-1$, ở đây $k-1 \geq 2$, và giả sử rằng $V$ là tổng trực tiếp của $k$ không gian con

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

Đặt $W = W_1 \oplus W_2 \oplus \cdots \oplus W_{k-1}$. Nó thì dễ dàng để kiểm tra lại rằng $W$ là $f$-bất biến và $V = W \oplus W_k$. Vì vậy, bởi trường hợp cho $k = 2$, $\chi(\lambda) = \chi_W(\lambda) \cdot \chi_k(\lambda)$, ở đây $\chi_W(\lambda)$ là đa thức đặc trưng của $f|_W$. Rõ ràng $W = W_1 \oplus W_2 \oplus \cdots \oplus W_{k-1}$, và do đó $\chi_W(\lambda) = \chi_1(\lambda) \cdot \chi_2(\lambda) \cdots \chi_{k-1}(\lambda)$ bởi giả thuyết quy nạp. Chúng ta kết luận rằng $\chi(\lambda) = \chi_1(\lambda) \cdot \chi_2(\lambda) \cdots \chi_{k-1}(\lambda) \cdot \chi_k(\lambda)$. $\square$

**8.1.6 Thí dụ.** Cho $f$ là một toán tử tuyến tính trên $\mathbb{R}^4$ được định nghĩa bởi

$$f(a, b, c, d) = (2a - b, a + b, c - d, c + d),$$

và cho $W_1 = \{(s, t, 0, 0) : s, t \in \mathbb{R}\}$ và $W_2 = \{(0, 0, s, t) : s, t \in \mathbb{R}\}$. Chú ý rằng $W_1$ và $W_2$ là mỗi $f$-bất biến và $\mathbb{R}^4 = W_1 \oplus W_2$. Cho $\mathscr{B}_1 = \{\mathbf{e}_1, \mathbf{e}_2\}$, $\mathscr{B}_2 = \{\mathbf{e}_3, \mathbf{e}_4\}$, và $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$. Khi đó $\mathscr{B}_1$ là một cơ sở có thứ tự cho $W_1$, $\mathscr{B}_2$ là một cơ sở có thứ tự cho $W_2$, và $\mathscr{B}$ là một cơ sở có thứ tự cho $\mathbb{R}^4$. Cho $A = f|_{\mathscr{B}}$, $B_1 = f|_{W_1}\big|_{\mathscr{B}_1}$, và $B_2 = f|_{W_2}\big|_{\mathscr{B}_2}$. Khi đó

$$B_1 = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}, \qquad B_2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

và

$$A = \begin{bmatrix} B_1 & O \\ O & B_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Cho $\chi(\lambda)$, $\chi_1(\lambda)$, và $\chi_2(\lambda)$ ký hiệu những đa thức đặc trưng của $f$, $f|_{W_1}$, và $f|_{W_2}$, theo tương ứng. Khi đó,

$$\chi(\lambda) = \det(A - \lambda I) = \det(B_1 - \lambda I) \cdot \det(B_2 - \lambda I) = \chi_1(\lambda) \cdot \chi_2(\lambda). \ \square$$

Từ phép chứng minh của định lý trên và thí dụ minh họa trên ta dễ dàng nhận thấy kết quả trong định lý sau.

**8.1.7 Định lý.** *Cho $f$ là một toán tử tuyến tính trên một không gian vector hữu hạn chiều $V$, và cho $W_1, W_2, \ldots, W_k$ là các không gian con $f$-bất* *biến của $V$ sao cho $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Với mỗi $i$, cho $\mathscr{B}_i$ là một cơ sở có thứ tự cho $W_i$, và cho $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2 \cup \cdots \cup \mathscr{B}_k$. Cho $A$ là ma trận của $f$ đối với cơ sở $\mathscr{B}$ và $B_i$ là ma trận của $f|_{W_i}$ đối với cơ sở $\mathscr{B}_i$ với $i = 1, 2, \ldots, k$. Khi đó,*

$$A = \begin{bmatrix} B_1 & O & \ldots & O \\ O & B_2 & \ldots & O \\ \ldots & \ldots & \ldots & \ldots \\ O & O & \ldots & B_k \end{bmatrix} = B_1 \oplus B_2 \oplus \cdots \oplus B_k$$

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

## 8.2 Định lý Cayley-Hamilton

**8.2.1 Định lý. (Cayley-Hamilton)** *Cho $f$ là một toán tử tuyến tính trên một không gian vector hữu hạn chiều $V$, và cho $\chi(\lambda)$ là đa thức đặc trưng của $f$. Khi đó $\chi(f) = \theta$, ánh xạ không.*

*Chứng minh.* Chúng ta chứng minh $\chi(f)(\mathbf{v}) = \mathbf{0}$ với mọi $\mathbf{v} \in V$. Đây là điều hiển nhiên nếu $\mathbf{v} = \mathbf{0}$ bởi vì $\chi(f)$ là tuyến tính, cho nên giả sử rằng $\mathbf{v} \neq \mathbf{0}$. Cho $W$ là không gian con $f$-cyclic tạo bởi $\mathbf{v}$, và giả sử rằng $\dim(W) = k$. Bởi định lý 8.1.3 (b), tồn tại những vô hướng $a_0, a_1, \ldots, a_{k-1}$ sao cho

$$f^k(\mathbf{v}) = -a_0\mathbf{v} - a_1 f(\mathbf{v}) - \cdots - f^{k-1}(\mathbf{v}).$$

Do định lý 8.1.3 suy ra

$$\chi_W(\lambda) = (-1)^k(a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k)$$

là đa thức đặc trưng của $f_W$. Từ hai phương trình trên ta có

$$\chi_W(f)(\mathbf{v}) = (-1)^k(a_0 I + a_1 f + \cdots + a_{k-1} f^{k-1} + f^k)(\mathbf{v}) = \mathbf{0}.$$

Bởi định lý 8.1.1, $\chi_W(\lambda)$ chia hết $\chi(\lambda)$, và do đó tồn tại đa thức $q(\lambda)$ sao cho $\chi(\lambda) = q(\lambda)\chi_W(\lambda)$. Vì vậy

$$\chi(f)(\mathbf{v}) = (q(f) \circ \chi_W(f))(\mathbf{v}) = q(f)(\chi_W(f)(\mathbf{v})) = q(f)(\mathbf{0}) = \mathbf{0}. \quad \square$$

**8.2.2 Thí dụ.** Cho $f$ là một toán tử tuyến tính trên $\mathbb{R}^2$ định nghĩa bởi $f(a,b) = (a+2b, -2a+b)$, và cho $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2)$ là cơ sở chính tắc của $\mathbb{R}^2$. Khi đó

$$A = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

là ma trận của $f$ đối với cơ sở $\mathscr{B}$. Do đó đa thức đặc trưng của $f$ là

$$\chi(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 5.$$

Ta dễ dàng thử lại được $\theta = \chi(f) = f^2 - 2f + 5I$ là ánh xạ không trên $\mathbb{R}^2$. Tương tự,

$$\chi(A) = A^2 - 2A + 5I = \begin{pmatrix} -3 & 4 \\ -4 & -3 \end{pmatrix} + \begin{pmatrix} -2 & -4 \\ 4 & -2 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \qquad \square$$

Thí dụ trên gợi ý kết quả sau.

**8.2.3 Hệ quả. (Định lý Cayley-Hamilton cho ma trận)** *Cho $A$ là ma trận cấp $n \times n$, và cho $\chi(\lambda)$ là đa thức đặc trưng của $A$. Khi đó $\chi(A) = O$, ma trận không cấp $n \times n$.*

*Chứng minh.* Dành cho bạn đọc xem như bài tập. $\square$

**8.2.4 Thí dụ. (OSV11)** Cho các ma trận thực $A$, $B$ vuông cùng cấp $n$. Đặt $C = AB - BA$. Chứng minh rằng nếu ma trận $C$ giao hoán với cả hai ma trận $A$ và $B$ thì tồn tại số nguyên dương $m$ sao cho $C^m = O$.

*Giải.* Ta có $\operatorname{tr}(C) = \operatorname{tr}(AB) - \operatorname{tr}(BA) = 0$. Do $C$ giao hoán với $A, B$ nên với $k$ nguyên dương ta có $C^j = (AB - BA)C^{j-1} = ABC^{j-1} - BAC^{j-1} = AC^{j-1}B - BAC^{j-1}$. Do đó, ta có $\operatorname{tr}(C^j) = 0$. Giả sử $C$ có các giá trị riêng khác 0 phân biệt là $\lambda_1, \ldots, \lambda_k$ với bội số tương ứng $s_1, \dots, s_k$. Khi đó, $\lambda_1^j, \dots, \lambda_k^j$ là các giá trị riêng của $C^j$ với bội số $s_1, \dots, s_k$. Vì vậy ta có hệ

$$\begin{cases} s_1\lambda_1 + s_2\lambda_2 + \cdots + s_k\lambda_k = \operatorname{tr}(C) = 0 \\ s_1\lambda_1^2 + s_2\lambda_2^2 + \cdots + s_k\lambda_k^2 = \operatorname{tr}(C^2) = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ s_1\lambda_1^k + s_2\lambda_2^2 + \cdots + s_k\lambda_k^k = \operatorname{tr}(C^k) = 0 \end{cases}$$

Bởi vì các $\lambda_1, \dots, \lambda_k$ phân biệt nên

$$\begin{vmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_k \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_k^2 \\ \dots & \dots & \dots & \dots \\ \lambda_1^k & \lambda_2^k & \dots & \lambda_k^k \end{vmatrix} = \lambda_1\lambda_2\cdots\lambda_k \prod_{1\le i<j\le k} (\lambda_j - \lambda_i) \neq 0$$

Ta dẫn đến điều vô lý vì hệ trên không có nghiệm tầm thường đối với $(s_1, s_2, \dots, s_k)$. Do đó, $C$ không có giá trị riêng khác không, nghĩa là tất cả các giá trị riêng của $C$ đều bằng không. Vậy đa thức đặc trưng của $A$ là $\chi(\lambda) = (-\lambda)^n$. Theo định lý Cayley-Hamilton ta được $(-C)^n = O$ hay $C^n = O$. $\square$

**8.2.5 Thí dụ. (OSV11)** Cho 3 dãy số $\{x_n\}$, $\{y_n\}$, $\{z_n\}$ xác định như sau $x_0 = y_0 = z_0$ và $\begin{cases} x_{n+1} = 4x_n - y_n - 5z_n \\ y_{n+1} = 2x_n - 2z_n \\ z_{n+1} = x_n - 2z_n \end{cases}$ Tính $x_{2011}$.

*Giải.* Đặt $A = \begin{bmatrix} 4 & -1 & -5 \\ 2 & 0 & -2 \\ 1 & 0 & -2 \end{bmatrix}$ và $\mathbf{u}_n = (x_n, y_n, z_n)$ với $n \in \mathbb{N}$. Khi đó $\mathbf{u}_n = A\mathbf{u}_0$. Đa thức đặc trưng của $A$ là $\chi(\lambda) = -(\lambda-1)(\lambda+1)(\lambda-2)$. Ta có biểu diễn $\lambda^{2011} = \chi(\lambda)g(\lambda) + a\lambda^2 + b\lambda + c$. Gán các giá trị $-1, 1, 2$ cho $\lambda$ ta được hệ

$$\begin{cases} a - b + c = -1 \\ a + b + c = 1 \\ 4a + 2b + c = 2^{2011} \end{cases} \quad \text{suy ra} \quad \begin{cases} a = \frac{2^{2011}-2}{3} \\ b = 1 \\ c = \frac{2-2^{2011}}{3} \end{cases}$$

Từ định lý Cayley-Hamilton ta suy ra được

$$A^{2011} = \chi(A)g(A) + \frac{2^{2011}-2}{3}\begin{bmatrix}4 & -1 & -5\\ 2 & 0 & -2\\ 1 & 0 & -2\end{bmatrix}^2 + \begin{bmatrix}4 & -1 & -5\\ 2 & 0 & -2\\ 1 & 0 & -2\end{bmatrix} + \frac{2-2^{2011}}{3}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$= \frac{2^{2011}-2}{3}\begin{bmatrix}8 & -4 & -8\\ 6 & -3 & -6\\ 2 & -1 & -2\end{bmatrix} + \begin{bmatrix}4 & -1 & -5\\ 2 & 0 & -2\\ 1 & 0 & -2\end{bmatrix}$$

Do $\mathbf{u}_{2011} = A^{2011}\mathbf{u}_0$ nên $x_{2011} = \frac{2^{2011}-2}{3}(8-4-8)x_0 + (4-1-5)x_0 = \frac{2-2^{2013}}{3}x_0$. □

## 8.3 Ứng dụng của đa thức đặc trưng

Tiếp theo chúng ta có một số tính chất mở rộng của đa thức đặc trưng.

**8.3.1 Định lý.** *Cho $A$ là một ma trận vuông cấp $n$ trên trường số phức $\mathbb{C}$ và $p(t)$ là một đa thức bậc $m$. Nếu $\lambda_1, \lambda_2, \ldots, \lambda_n$ là các giá trị riêng của $A$, thì*

**(a)** $\det(p(A)) = p(\lambda_1)p(\lambda_2)\cdots p(\lambda_n)$.

**(b)** $p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)$ *là các giá trị riêng của* $p(A)$.

*Chứng minh.*

**(a)** Đa thức đặc trưng của $A$ được viết thành $\chi_A(\lambda) = (-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$. Gọi các nghiệm của đa thức $p(t)$ là $\alpha_1, \alpha_2, \ldots, \alpha_m$, cho nên $p(t) = c(t-\alpha_1)(t-\alpha_2)\cdots(t-\alpha_m)$ với $c$ là hệ số của bậc cao nhất. Do đó ta viết được

$$p(A) = c(A-\alpha_1 I)(A-\alpha_2 I)\cdots(A-\alpha_m I).$$

Theo tính chất của định thức, đặc biệt định thức của tích hai ma trận bằng tích hai định thức, ta tính được

$$\begin{aligned}
\det(p(A)) &= c^n|A-\alpha_1 I|\cdot|A-\alpha_2 I|\cdots|A-\alpha_m I| \\
&= c^n\chi_A(\alpha_1)\chi_A(\alpha_2)\cdots\chi_A(\alpha_m) \\
&= c^n\prod_{i=1}^{m}(-1)^n(\alpha_i-\lambda_1)(\alpha_i-\lambda_2)\cdots(\alpha_i-\lambda_n) \\
&= c^n\prod_{i=1}^{m}\prod_{j=1}^{n}(\lambda_j-\alpha_i) \\
&= \prod_{j=1}^{n}c^n(\lambda_j-\alpha_1)(\lambda_j-\alpha_2)\cdots(\lambda_j-\alpha_m) \\
&= \prod_{j=1}^{n}p(\lambda_j).
\end{aligned}$$

**(b)** Xét đa thức $f(t)=p(t)-\lambda$ và áp dụng kết quả trên ta có $\det(f(A)) = f(\lambda_1)f(\lambda_2)\cdots f(\lambda_n)$ hay

$$\det(p(A)-\lambda I) = (p(\lambda_1)-\lambda)(p(\lambda_2)-\lambda)\cdots(p(\lambda_n)-\lambda).$$

Vậy các giá trị riêng của $p(A)$ là $p(\lambda_1)$, $p(\lambda_2)$, ..., $p(\lambda_n)$. □

**8.3.2 Thí dụ.** Cho đa thức $p(t)=t^{2012}+1$ và ma trận $A=\begin{pmatrix}2&1&0\\1&2&0\\1&1&1\end{pmatrix}$.

Tính $\det(p(A))$ và tìm các giá trị riêng của $p(A)$.

*Giải*. Đa thức đặc trưng của $A$ là

$$\chi_A(\lambda)=\begin{vmatrix}2-\lambda&1&0\\1&2-\lambda&0\\1&1&1-\lambda\end{vmatrix}=(1-\lambda)(\lambda^2-4\lambda+3)$$

Vậy $A$ có hai giá trị riêng $\lambda=1$ (bội hai) và $\lambda=3$. Do đó theo định lý trên ta tính được

$$\det(p(A))=(1^{2012}+1)(1^{2012}+1)(3^{2012}+1)=4(3^{2012}+1),$$

và các giá trị riêng của $p(A) = A^{2012} + I$ là $p(1) = 2$ và $p(3) = 3^{2012} + 1$. □

**8.3.3 Thí dụ. (OSV99)** Cho $f(x) = x^{1999} + x^2 - 1$ và cho ma trận $C = \begin{bmatrix} 4 & 3 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 9 & -1 & 0 \\ 1 & 2 & 5 & 2 \end{bmatrix}$. Tính $\det(f(C))$.

*Giải.* Đa thức đặc trưng của ma trận $C$ là

$$\begin{aligned}\chi(\lambda) &= (-1-\lambda)(2-\lambda)((4-\lambda)(3-\lambda)-6) \\ &= (\lambda+1)(\lambda-2)(\lambda-1)(\lambda-6).\end{aligned}$$

Vậy ma trận $C$ có các giá trị riêng $1, -1, 2, 6$. Do đó theo Định lý 8.3.1 ta có

$$\det(f(C)) = f(1)f(-1)f(2)f(6) = -(2^{1999}+3)(6^{1999}+35). \quad \square$$

**8.3.4 Thí dụ. (OSV08)** Cho $A$ là ma trận vuông cấp $n$ khả nghịch. Mọi phần tử của các ma trận $A$, $A^{-1}$ là số nguyên. Chứng minh rằng nếu $A$ có $n$ giá trị riêng đều là các số thực thì $|\det(A + A^{-1})| \geq 2^n$.

*Giải.* Do các phần tử của $A$, $A^{-1}$ đều là các số nguyên nên $\det(A)$ và $\det(A^{-1})$ cũng là các số nguyên. Hơn nữa, ta có $\det(A)\det(A^{-1}) = 1$ cho nên $\det A = \det(A^{-1}) = \pm 1$. Từ đó ta suy ra được

$$|\det(A + A^{-1})| = |\det(A)\det(A + A^{-1})| = |\det(A^2 + I)|.$$

Gọi $\lambda_1, \lambda_2, \ldots, \lambda_n$ là các giá trị riêng của $A$. Theo định lý trên $\lambda_1^2 + 1, \lambda_2^2 + 1, \ldots, \lambda_n^2 + 1$ là các giá trị riêng của $I + A^2$ và

$$\begin{aligned}|\det(A^2 + I)| &= (\lambda_1^2+1)(\lambda_2^2+1)\cdots(\lambda_n^2+1) \\ &\geq 2^n|a_1a_2\cdots a_n| \qquad \text{(bất đẳng thức Cauchy)} \\ &= 2^n|\det A| = 2^n.\end{aligned}$$

□

**8.3.5 Hệ quả.** *Cho $A$ là ma trận vuông cấp $n$. Nếu $\lambda_0$ là giá trị riêng của $A$ thì $\lambda_0^n$ là giá trị riêng của $A^n$.*

*Chứng minh.* Với $p(t) = t^n$ và do $\lambda_0$ là giá trị riêng của $A$, theo định lý trên $p(\lambda_0) = \lambda_0^n$ là một giá trị riêng của $p(A) = A^n$. $\square$

**8.3.6 Thí dụ.** Cho ma trận $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Hãy tìm một giá trị riêng của ma trận $A^{2009}$.

*Giải.* Đa thức đặc trưng của ma trận $A$ là

$$\chi_A(\lambda) = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = -\lambda^3 + 3\lambda + 2 = -(\lambda+1)^2(\lambda-2)$$

Vậy $A$ có giá trị riêng $-1$ bội hai và 2. Do đó ma trận $A^{2009}$ có $(-1)^{2009} = -1$ là giá trị riêng bội hai và $2^{2009}$ là giá trị riêng. $\square$

**8.3.7 Thí dụ. (OSV06)** Cho ma trận $A = \begin{bmatrix} 2 & 4 & 4 \\ 8 & 6 & 8 \\ 0 & 4 & 6 \end{bmatrix}$. Chứng minh rằng $\det(I - A^{2006}) \neq 0$.

*Giải.* Đa thức đặc trưng của ma trận $A$ là $\chi_A(\lambda) = -\lambda^3 + 14\lambda^2 + 4\lambda - 56$, và giá trị riêng của nó là 2, $-2$, 14. Do đó ma trận $A^{2006}$ có các giá trị riêng là $2^{2006}$, $(-2)^{2006}$ và $14^{2006}$, cho nên nó không có giá trị riêng là 1. Vì vậy $\det(A^{2006} - I) \neq 0$ hay $\det(I - A^{2006}) \neq 0$. $\square$

**8.3.8 Thí dụ. (OSV12)** Cho $A, B \in \mathcal{M}_{100}(\mathbb{R})$ là hai ma trận thỏa mãn $A^{101} = O$ và $AB = 2A + 3B$. Chứng minh rằng $(A+B)^{100} = O$.

*Giải.* Nếu $\lambda$ là giá trị riêng của $A$ thì $\lambda^{101}$ là giá trị riêng của $A^{101}$; nhưng $A^{101} = O$ cho nên $\lambda = 0$. Vậy $A$ chỉ có các giá trị riêng bằng 0, cho nên đa thức đặc trưng của $A$ là $\chi(\lambda) = \lambda^{100}$. Theo định lý Cayley-Hamilton ta có $A^{100} = O$. Mặt khác, do $AB = 2A + 3B$ nên $(A - 3I)(B - 2I) = 6I$ hay $\frac{1}{6}(A - 3I)(B - 2I) = I$. Từ tính chất ma trận nghịch đảo của nhau ta có $\frac{1}{6}(B - 2I)(A - 3I) = I$, suy ra $BA = 2A + 3B$. Vậy $AB = BA$.

Bên cạnh đó, ta có $3(A+B) = AB + A = A(B+I) = (B+I)A$ (vì $AB = BA$), nghĩa là $A$ giao hoán với $B+I$. Do đó, ta được

$$(A+B)^{100} = \frac{1}{3^{100}}(A(B+I))^{100} = \frac{1}{3^{100}}A^{100}(B+I)^{100} = O. \quad \square$$

**8.3.9 Hệ quả.** *Cho $A$ là một ma trận vuông cấp $n$, trên trường số phức $\mathbb{C}$. Khi đó $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$ với $\lambda_i$, $i = 1, 2, \ldots, n$ là các giá trị riêng của $A$. Từ đó suy ra các giá trị riêng thu được đều là ước của* $\det(A)$.

*Chứng minh.* Cho $p(t) = t$, theo Định lý 8.3.1 ta có $|A| = \det(p(A)) = p(\lambda_1)p(\lambda_2)\cdots p(\lambda_n) = \lambda_1\lambda_2\cdots\lambda_n$. $\square$

**8.3.10 Thí dụ. (OSV08)** Cho $A$ là ma trận vuông thực cấp 3 có vết là 8. Tổng các phần tử trên mỗi hàng của $A$ bằng 4 và $\det(A) = 16$. Xác định các giá trị riêng của $A$.

*Giải.* Từ giả thiết đa thức đặc trưng của ma trận $A$ có dạng

$$\chi(\lambda) = (-1)^3(\lambda^3 - (\mathrm{tr}(A))\lambda^2 + a - \det(A) = -\lambda^3 + 8\lambda^2 - a\lambda + 16$$

Mặt khác, ta có

$$\begin{aligned}
\chi(\lambda) &= \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} \\
&= \begin{vmatrix} a_{11}+a_{12}+a_{13}-\lambda & a_{12} & a_{13} \\ a_{21}+a_{22}+a_{23}-\lambda & a_{22}-\lambda & a_{23} \\ a_{31}+a_{32}+a_{33}-\lambda & a_{32} & a_{33}-\lambda \end{vmatrix} \\
&= (4-\lambda)\begin{vmatrix} 1 & a_{12} & a_{13} \\ 1 & a_{22}-\lambda & a_{23} \\ 1 & a_{32} & a_{33}-\lambda \end{vmatrix}
\end{aligned}$$

Vậy ma trận $A$ có một giá trị riêng là $\lambda = 4$, cho nên $0 = -4^3 + 8\cdot 4^2 - a4 + 16$ suy ra $a = 20$. Lúc này đa thức đặc trưng trở thành $\chi(\lambda) = -\lambda^3 + 8\lambda^2 - 20\lambda + 16 = -(\lambda-4)(\lambda-2)^2$. Vậy $A$ có 4 là giá

trị riêng đơn và 2 là giá trị riêng kép. Rõ ràng có ma trận $A$ thỏa điều kiện đề bài, chẳng hạn
$$A = \begin{bmatrix} 2 & 2 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{bmatrix} \qquad \square$$

## Bài tập

**1)** Cho $\chi(\lambda)$ là đa thức đặc trưng của ma trận vuông $A$ cấp $n$. Chứng minh rằng $\chi(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) + q(\lambda)$, trong đó $q(\lambda)$ là một đa thức có bậc không quá $n - 2$.

**2)** Cho $\chi(\lambda)$ là đa thức đặc trưng của ma trận vuông $A$ cấp $n$. Chứng minh rằng nếu **a** là vector riêng ứng với giá trị riêng $\lambda$ thì $\chi(A)\mathbf{a} = \chi(\lambda)\mathbf{a}$; nghĩa là **x** là vector riêng ứng với giá trị riêng $\chi(\lambda)$ của ma trận $\chi(A)$.

# § 9 Tính chất các vector riêng và chéo hóa ma trận vuông

## 9.1 Tính chất giá trị riêng và vector riêng

**9.1.1 Định lý.** *Nếu $\lambda_0$ là giá trị riêng bội $k$ (nghĩa là $\lambda_0$ là nghiệm bội $k$ của đa thức đặc trưng $\chi(\lambda)$) của toán tử tuyến tính $f$ trên $V$, thì* $\dim(E(\lambda_0)) \leq k$.

*Chứng minh.* Giả sử $E(\lambda_0)$ có một cơ sở là $(\mathbf{u}_1, \ldots, \mathbf{u}_l)$. Bổ sung vào cơ sở này để thu được một cơ sở của $V$ là $\mathscr{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_l, \mathbf{u}_{l+1}, \ldots, \mathbf{u}_n)$. Vì $f(\mathbf{u}_1) = \lambda_0\mathbf{u}_1, \ldots, f(\mathbf{u}_l) = \lambda_0\mathbf{u}_l$ nên ma trận của $f$ trong cơ sở $\mathscr{B}$ có dạng
$$A = \begin{bmatrix} \lambda_0 & \ldots & 0 & a_{1,l+1} & \ldots & a_{1n} \\ \multicolumn{6}{c}{\dotfill} \\ 0 & \ldots & \lambda_0 & a_{l,l+1} & \ldots & a_{ln} \\ 0 & \ldots & 0 & a_{l+1,l+1} & \ldots & a_{l+1,n} \\ \multicolumn{6}{c}{\dotfill} \\ 0 & \ldots & 0 & a_{n,l+1} & \ldots & a_{nn} \end{bmatrix}$$

trong đó $f(\mathbf{u}_i)|_{\mathscr{B}} = (a_{1i}, \ldots, a_{ni})$ với mọi $i = l+1, \ldots, n$. Do đó, đa thức đặc trưng của $A$ có dạng

$$\chi_A(\lambda) = \begin{vmatrix} \lambda_0 - \lambda & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_0 - \lambda \end{vmatrix} \cdot \begin{vmatrix} a_{l+1,l+1} - \lambda & \ldots & a_{l+1,n} \\ \vdots & \ddots & \vdots \\ a_{n,l+1} & \ldots & a_{nn} - \lambda \end{vmatrix}$$
$$= (\lambda_0 - \lambda)^l \psi(\lambda)$$

trong đó $\psi(\lambda)$ là một đa thức bậc $n-l$. Vì $\lambda_0$ là nghiệm bội $k$ của $\chi_A(\lambda)$ nên $k \geq l$. Vậy $\dim E(\lambda) \leq k$. □

**9.1.2 Thí dụ. (OSV10)** Cho $A$ là ma trận thực, vuông cấp $n \geq 2$, có vết là 10 và $\operatorname{rank} A = 1$. Tìm đa thức đặc trưng và đa thức tối tiểu của $A$.

*Giải.* Vì $\operatorname{rank} A = 1 < n$ nên $\det A = 0$. Do đó $A$ có giá trị riêng là 0. Hơn nữa ta có $\dim(E(0)) = n - \operatorname{rank} A = n - 1$, cho nên giá trị riêng 0 có số bội lớn hơn hoặc bằng $n-1$. Theo giả thiết $\operatorname{tr}(A) = 10$, bằng tổng tất cả giá trị riêng (gồm cả phức) của $A$, cho nên $A$ có đúng một giá trị riêng là 10 và 0 là giá trị riêng với số bội là $n-1$. Vì vậy đa thức đặc trưng của $A$ là $\chi_A(\lambda) = (-1)^n \lambda^{n-1}(\lambda - 10)$. Ta có thể kiểm tra được không có một đa thức $p(\lambda)$ bậc nhỏ hơn 2 ước của $\chi_A$ sao cho $p(A) = 0$. Mặt khác, vì $\operatorname{rank}(A) = 1$ nên $A$ có dạng

$$A = \begin{bmatrix} \lambda_1 a_1 & \lambda_1 a_2 & \ldots & \lambda_1 a_n \\ \lambda_2 a_1 & \lambda_2 a_2 & \ldots & \lambda_2 a_n \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_n a_1 & \lambda_n a_2 & \ldots & \lambda_n a_n \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} = \mathbf{u}\mathbf{v}^t$$

với $(\lambda_1 a_1, \lambda_1 a_2, \ldots, \lambda_1 a_n) \neq \mathbf{0}$. Lưu ý rằng $\mathbf{v}^t\mathbf{u} = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n = 10$; từ đó ta có thể tính trực tiếp được $A^2 = \mathbf{u}\mathbf{v}^t\mathbf{u}\mathbf{v}^t = \mathbf{u}(\mathbf{v}^t\mathbf{u})\mathbf{v}^t = \mathbf{u}\operatorname{tr}(A)\mathbf{v}^t = 10A$. Vậy đa thức tối tiểu của $A$ là $p(\lambda) = \lambda^2 - 10\lambda$. □

**9.1.3 Định lý.** *Cho toán tử tuyến tính $f$ trên $V$. Giả sử $\lambda_1, \ldots, \lambda_m$ là các giá trị riêng khác nhau của $f$. Nếu $\mathbf{u}_i$ là vector riêng ứng với giá trị riêng $\lambda_i$ với $i = 1, \ldots, m$, thì hệ $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ độc lập tuyến tính.*

*Chứng minh.* Chúng ta chứng minh định lý bằng qui nạp theo $m$. Nếu $m = 1$ và $\mathbf{u}_1$ là vector riêng ứng với giá trị riêng $\lambda_1$, thì $\mathbf{u}_1 \neq 0$ nên hệ $(\mathbf{u}_1)$ độc lập tuyến tính. Giả sử định lý đúng với $m = k$, nghĩa là $k$ vector riêng của toán tử tuyến tính $f$ ứng với $k$ giá trị riêng khác nhau thì lập thành một hệ độc lập tuyến tính.

Xét $k+1$ vector riêng $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}$ ứng với $k+1$ giá trị riêng khác nhau $\lambda_1, \dots, \lambda_k, \lambda_{k+1}$. Ta có $f(\mathbf{u}_i) = \lambda_i \mathbf{u}_i$ với $i = 1, \dots, k+1$. Xét tổ hợp tuyến tính

(9.1.4) $$\sum_{i=1}^{k+1} \alpha_i \mathbf{u}_i = \mathbf{0}.$$

Vì $f$ là toán tử tuyến tính nên suy ra

(9.1.5) $$f\Big(\sum_{i=1}^{k+1} \alpha_i \mathbf{u}_i\Big) = \sum_{i=1}^{k+1} \alpha_i \lambda_i \mathbf{u}_i = \mathbf{0}.$$

Nhân cả hai vế biểu thức (9.1.4) cho $\lambda_{i+1}$ rồi trừ đi vế theo vế biểu thức (9.1.5) ta được

$$\sum_{i=1}^{k} (\lambda_{k+1} - \lambda_i)\alpha_i \mathbf{u}_i = \mathbf{0}.$$

Theo giả thiết qui nạp ta có hệ vector $(\mathbf{u}_1, \dots, \mathbf{u}_k)$ độc lập tuyến tính và các $\lambda_1, \dots, \lambda_k, \lambda_{k+1}$ đôi một khác nhau, nên từ đẳng thức trên ta suy ra được $\alpha_i = 0$ với mọi $i = 1, \dots, k$. Từ đó (9.1.4) trở thành $\alpha_{k+1}\mathbf{u}_{k+1} = \mathbf{0}$. Do $\mathbf{u}_{k+1} \neq \mathbf{0}$ ta phải có $\alpha_{k+1} = 0$. Vậy hệ $(\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1})$ độc lập tuyến tính. □

**9.1.6 Định lý.** *Giả sử $\lambda_1, \lambda_2, \dots, \lambda_m$ là các giá trị riêng khác nhau của toán tử tuyến tính $f$. Nếu với mỗi $i = 1, 2, \dots, m$ hệ $(\mathbf{u}_{i1}, \dots, \mathbf{u}_{il_i})$ độc lập tuyến tính trong $E(\lambda_i)$, thì hệ $(\mathbf{u}_{11}, \dots, \mathbf{u}_{1l_1}, \mathbf{u}_{21}, \dots, \mathbf{u}_{2l_2}, \dots, \mathbf{u}_{m1}, \dots, \mathbf{u}_{ml_m})$ độc lập tuyến tính.*

*Chứng minh.* Xét tổ hợp tuyến tính $\sum_{i=1}^{m}\sum_{j=1}^{l_i}\alpha_{ij}\mathbf{u}_{ij} = \mathbf{0}$. Đặt $\mathbf{u}_i = \sum_{j=1}^{l_i}\alpha_{ij}\mathbf{u}_{ij}$ với $i = 1, \ldots, m$. Ta nhận thấy $\mathbf{u}_i \in E(\lambda_i)$ và $\sum_{i=1}^{m}\mathbf{u}_i = \mathbf{0}$. Giả sử tồn tại $i_0$ sao cho $\mathbf{u}_{i_0} \neq \mathbf{0}$. Đặt $I = \{i : 1 \leq i \leq m, \mathbf{u}_i \neq \mathbf{0}\}$. Ta có $\sum_{i\in I}\mathbf{u}_i = \mathbf{0}$, và $\mathbf{u}_i$ là một vector riêng ứng với giá trị riêng $\lambda_i$ với mỗi $i \in I$. Theo Định lý 9.1.3 ta có hệ vector $(\mathbf{u}_i)_{i\in I}$ độc lập tuyến tính, điều này mâu thuẫn biểu thức $\sum_{i\in I}\mathbf{u}_i = \mathbf{0}$. Vậy $\mathbf{u}_i = \mathbf{0}$ với mọi $i = 1, \ldots, m$. Do đó, với mỗi $i = 1, \ldots, m$, ta có $\sum_{j=1}^{l_i}\alpha_{ij}\mathbf{u}_{ij} = \mathbf{0}$, suy ra $\alpha_{ij} = 0$ với mọi $j = 1, \ldots, l_i$ vì hệ $(\mathbf{u}_{i1}, \ldots, \mathbf{u}_{il_i})$ độc lập tuyến tính. Vậy ta được hệ $(\mathbf{u}_{11}, \ldots, \mathbf{u}_{1l_1}, \mathbf{u}_{21}, \ldots, \mathbf{u}_{2l_2}, \ldots, \mathbf{u}_{m1}, \ldots, \mathbf{u}_{ml_m})$ độc lập tuyến tính. □

## 9.2 Ma trận vuông chéo hóa được

Cho $V$ là một $\mathbb{K}$-không gian vector $n$ chiều và $f$ là một toán tử tuyến tính trên $V$. Ta biết rằng, ứng với mỗi cơ sở của $V$, $f$ có một ma trận duy nhất và các tính toán trên $f$ có thể quy về các tính toán trên ma trận của nó. Đương nhiên, nếu ma trận càng đơn giản thì việc tính toán càng dễ dàng. Do vậy, ta mong muốn tìm một cơ sở của $V$ sao cho ma trận của $f$ có dạng đơn giản, chẳng hạn như dạng ma trận chéo. Vấn đề đặt ra là, đối với mỗi toán tử tuyến tính $f$ trên $V$, liệu có tìm được một cơ sở của $V$ để ma trận của $f$ trong cơ sở đó có dạng chéo hay không? Rất tiếc, nói chung câu trả lời là phủ định. Phần này sẽ nghiên cứu về vấn đề nay.

**9.2.1 Định nghĩa.** Toán tử tuyến tính trên không gian vector hữu hạn chiều $V$ được gọi là **chéo hóa được** nếu tồn tại một cơ sở của $V$ sao cho ma trận của $f$ đối với cơ sở đó là ma trận chéo. Việc tìm một cơ sở như vậy gọi là **chéo hóa** $f$.

**9.2.2 Định nghĩa.** Ma trận vuông $A$ trên $\mathbb{K}$ được gọi là **chéo hóa được** nếu tồn tại một ma trận không suy biến $C$ sao cho $C^{-1}AC$ là ma trận đường chéo. Khi đó, ta nói ma trận *$C$ làm chéo hóa $A$* và *$A$ được chéo hóa* bởi $C$. Việc tìm ma trận $C$ như vậy được gọi là *chéo hóa $A$*.

**9.2.3 Định lý.** *Toán tử tuyến tính $f$ trên không gian vector hữu hạn chiều là chéo hóa được khi và chỉ khi với mọi cơ sở của $V$ thì ma trận của $f$ đối với cơ sở ấy là chéo hóa được.*

*Chứng minh.* Gọi $\mathscr{A}$ là một cơ sở bất kỳ của $V$ và ma trận của $f$ trong cơ sở $\mathscr{A}$ là $A$. Khi đó $A$ chéo hóa được khi và chỉ khi tồn tại ma trận không suy biến $C$ sao cho $C^{-1}AC$ là ma trận chéo. Khi đó ma trận của $f$ đối với cơ sở $\mathscr{B}$ là ma trận chéo với $\mathscr{B}$ là cơ sở sao cho $C$ là ma trận chuyển cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$. (Xem công thức (2.3.3)). □

**9.2.4 Định lý.** *Điều kiện cần và đủ để ma trận vuông cấp $n$ chéo hóa được là nó có $n$ vector riêng độc lập tuyến tính.*

*Chứng minh.* Giả sử ma trận vuông $A$ cấp $n$ chéo hóa được. Khi đó, tồn tại ma trận vuông không suy biến $C$ sao cho $C^{-1}AC = D$ là ma trận chéo. Đặt

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \qquad \text{và} \qquad \mathbf{c}_j = (c_{1j}, c_{2j}, \dots, c_{nj}).$$

Do $C^{-1}AC = D$ nên $AC = CD$. Ta viết lại

$$A \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

suy ra

$$\begin{bmatrix} A\mathbf{c}_1 & A\mathbf{c}_2 & \dots & A\mathbf{c}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{c}_1 & \lambda_2\mathbf{c}_2 & \dots & \lambda_n\mathbf{c}_n \end{bmatrix}$$

Vậy $A\mathbf{c}_i = \lambda_i\mathbf{c}_i$ với mọi $i = 1, 2, \dots, n$. Do $C$ không suy biến nên $\mathbf{c}_i \neq \mathbf{0}$. Vậy $\mathbf{c}_i$ là vector riêng của $A$ ứng với giá trị riêng $\lambda_i$. Hơn nữa, hệ $(\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n)$ độc lập tuyến tính vì ma trận $C$ không suy biến.

Ngược lại, giả sử ma trận $A$ có $n$ vector riêng độc lập tuyến tính $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ tương ứng với $n$ giá trị riêng $\lambda_1, \lambda_2, \ldots, \lambda_n$ (các $\lambda_i$ này có thể trùng nhau). Đặt

$$P = \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n\end{bmatrix} \qquad \text{và} \qquad D = \begin{bmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ \multicolumn{4}{c}{\dotfill}\\ 0 & 0 & \ldots & \lambda_n\end{bmatrix}.$$

Vì hệ vector $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ độc lập tuyến tính nên $P$ là ma trận không suy biến. Mặt khác, ta có $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$ với mọi $i = 1, 2, \ldots, n$, suy ra $\begin{bmatrix}A\mathbf{u}_1 & A\mathbf{u}_2 & \ldots & A\mathbf{u}_n\end{bmatrix} = \begin{bmatrix}\lambda_1\mathbf{u}_1 & \lambda_2\mathbf{u}_2 & \ldots & \lambda_n\mathbf{u}_n\end{bmatrix}$. Ta viết lại

$$A\begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n\end{bmatrix} = \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n\end{bmatrix}\begin{bmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ \multicolumn{4}{c}{\dotfill}\\ 0 & 0 & \ldots & \lambda_n\end{bmatrix}$$

hay $AP = PD$. Do đó, $P^{-1}AP = D$. Vậy $A$ chéo hóa được bởi $P$. □

**9.2.5 Nhận xét.** Trong chứng minh định lý trên ta thấy nếu trong ma trận $A$ ta biết có $n$ giá trị riêng (có thể trùng nhau) và $n$ vector riêng tương ứng độc lập tuyến tính thì ta tìm được ma trận chéo của $A$ và ma trận làm chéo hóa $A$.

**9.2.6 Hệ quả.** *Nếu ma trận vuông cấp $n$ có $n$ giá trị riêng đôi một khác nhau thì nó chéo hóa được.*

*Chứng minh.* Giả sử ma trận vuông cấp $n$ có $n$ giá trị riêng đôi một khác nhau. Theo Định lý 9.1.3 thì hệ $n$ vector riêng ứng với $n$ giá trị riêng ấy là độc lập tuyến tính. Khi đó, theo Định lý 9.2.4 thì ma trận đang xét chéo hóa được. □

**9.2.7 Định lý.** *Ma trận vuông cấp $n$ chéo hóa được khi và chỉ khi $A$ có $n$ giá trị riêng kể cả bội và số chiều của tất cả không gian con riêng bằng số bội của giá trị riêng tương ứng.*

*Chứng minh.* Giả sử ma trận $A$ vuông cấp $n$ chéo hóa được. Khi đó, theo Định lý 9.2.4 thì ma trận $A$ có $n$ vector riêng độc lập tuyến tính, cụ thể là $(\mathbf{u}_{11}, \dots, \mathbf{u}_{1l_1}, \mathbf{u}_{21}, \dots, \mathbf{u}_{2l_2}, \mathbf{u}_{m1}, \dots, \mathbf{u}_{ml_m})$, trong đó $\mathbf{u}_{ij}$ là vector riêng tương ứng với giá trị riêng $\lambda_i$ với $j = 1, \dots, l_i$ và $i = 1, \dots, m$ và $l_1 + l_2 + \cdots + l_m = n$.

Gọi $n_i$ là số bội của giá trị riêng $\lambda_i$. Theo Định lý 9.1.1 ta có $\dim(E(\lambda_i)) \leq n_i$. Do hệ vector $(\mathbf{u}_{i1}, \dots, \mathbf{u}_{il_i})$ độc lập tuyến tính và thuộc không gian con riêng $E(\lambda_i)$ nên $l_i \leq \dim(E(\lambda_i)) \leq n_i$. Suy ra

$$n = l_1 + l_2 + \cdots + l_m \leq n_1 + n_2 + \cdots + n_m \leq n.$$

Do đó, ta có $l_i = \dim(E(\lambda_i)) = n_i$ với mọi $i = 1, \dots, m$.

Ngược lại, giả sử ma trận $A$ có $m$ giá trị riêng đôi một khác nhau $\lambda_1, \dots, \lambda_m$ với số bội tương ứng là $l_1, \dots, l_m$ thỏa $l_1 + \cdots + l_m = n$. Với mỗi $i = 1, \dots, m$ ta có $\dim(E(\lambda_i)) = l_i$. Gọi $(\mathbf{u}_{i1}, \dots, \mathbf{u}_{il_i})$ là một cơ sở của không gian con riêng $E(\lambda_i)$. Theo Định lý 9.1.6 thì hệ $n$ vector $(\mathbf{u}_{11}, \dots, \mathbf{u}_{1l_1}, \dots, \mathbf{u}_{m1}, \dots, \mathbf{u}_{ml_m})$ độc lập tuyến tính. Hơn nữa, mỗi vector trong hệ là một vector riêng của ma trận $A$. Do đó, theo Định lý 9.2.4 thì ma trận $A$ chéo hóa được. $\square$

**9.2.8** Từ các kết quả trên, việc giải bài toán chéo hóa ma trận hay chéo hóa toán tử tuyến tính được thực hiện theo các bước sau. Đối với bài toán chéo hóa ma trận $A \in \mathcal{M}_n(\mathbb{K})$.

**Bước 1.** Lập đa thức đặc trưng của $A$ và giải phương trình đặc trưng trong $\mathbb{K}$ để tìm các giá trị riêng của $A$.

(a) Nếu $A$ không có giá trị riêng nào trong $\mathbb{K}$ thì $A$ không chéo hóa được. Bài toán kết thúc.

(b) Giả sử $A$ có $k$ giá trị riêng đôi một khác nhau $\lambda_1$, …, $\lambda_k$ với số bội tương ứng là $n_1, \dots, n_k$.

(i) Nếu $n_1 + \cdots + n_k < n$ thì $A$ không chéo hóa được. Bài toán kết thúc.

(ii) Nếu $n_1 + \cdots + n_k = n$ thì làm tiếp bước 2.

**Bước 2.** Với mỗi giá trị riêng $\lambda_i$ tìm $E(\lambda_i)$ và $\dim(E(\lambda_i))$ với $i = 1, \ldots, k$

(a) Nếu tồn tại ít nhất một $\lambda_i$ mà $\dim(E(\lambda_i)) < n_i$ thì $A$ không chéo hóa được. Bài toán kết thúc.

(b) Nếu $\dim(E(\lambda_i)) = n_i$ với mọi $i = 1, \ldots, k$ thì kết luận $A$ chéo hóa được. Với mỗi $\lambda_i$, tìm một cơ sở của không gian con riêng $E(\lambda_i)$ với $i = 1, \ldots, k$. Sau đó làm tiếp bước 3.

**Bước 3.** Lập ma trận $C$ mà các cột lần lượt là các vector cơ sở của các không gian con riêng $E(\lambda_i)$ với $i = 1, \ldots, k$.

Khi đó $C$ là ma trận làm chéo hóa $A$. Hơn nữa, $D = C^{-1}AC$ la ma trận chéo mà các phần tử trên đường chéo chính lần lượt là các giá trị riêng tương ứng với các vector riêng trong cột ma trận $C$. (Giá trị riêng $\lambda_i$ xuất hiện $n_i$ lần với $i = 1, \ldots, k$).

Trường hợp $f$ là toán tử tuyến tính của $\mathbb{K}$-không gian vector $V$ có ma trận là $A$ trong một cơ sở $\mathscr{B}$ nào đó của $V$. Ta giải bài toán chéo hóa ma trận $A$ ta cũng được lời giải cho bài toán chéo hóa $f$. Tuy nhiên, cần chú ý rằng ma trận $C$ thu được là ma trận chuyển từ cơ sở $\mathscr{B}$ sang cơ sở mới mà $f$ có ma trận chéo $D$ trong nó.

## 9.3 Các ví dụ

**9.3.1 Thí dụ.** Từ Thí dụ 7.3.4, ta thấy ma trận $A = \begin{bmatrix} 3 & 3 & 2 \\ 1 & 1 & -2 \\ 3 & -1 & 0 \end{bmatrix}$ chỉ có một giá trị riêng đơn trên $\mathbb{R}$ là $\lambda = 4$ nên ma trận $A$ không chéo hóa được trên $\mathbb{R}$. $\square$

**9.3.2 Thí dụ.** Theo Thí dụ 7.3.2, ta biết ma trận $A = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ có giá trị riêng $\lambda = 1$ với vector riêng tương ứng là $\mathbf{u}_1 = (1, 1, 0)$ và giá trị riêng $\lambda = 5$ bội hai với vector riêng tương ứng là $\mathbf{u}_2 = (-1, 1, 0)$ và

$\mathbf{u}_3 = (0, 0, 1)$. Vậy ma trận $A$ chéo hóa được và ma trận làm chéo hóa $A$ là

$$C = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \text{Hơn nữa, ta có} \qquad C^{-1}AC = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}. \quad \square$$

**9.3.3 Thí dụ.** Toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x, y, z) = (x - y, y - z, x + z)$ có chéo hóa được không?

*Giải.* Ta thấy ma trận của $f$ đối với cơ sở chính tắc là $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$.

Đa thức đặc trưng của $A$ là

$$\begin{aligned} \begin{vmatrix} 1-\lambda & -1 & 0 \\ 0 & 1-\lambda & -1 \\ 1 & 0 & 1-\lambda \end{vmatrix} &= (1-\lambda)\begin{vmatrix} 1-\lambda & -1 \\ 0 & 1-\lambda \end{vmatrix} + \begin{vmatrix} -1 & 0 \\ 1-\lambda & -1 \end{vmatrix} \\ &= (1-\lambda)^3 + 1. \end{aligned}$$

Vậy ma trận $A$ chỉ có một giá trị riêng đơn là $\lambda = 2$. Do đó, $A$ không chéo hóa được. Vậy $f$ không chéo hóa được. $\square$

**9.3.4 Thí dụ.** Ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi $f(x_1, x_2, x_3) = (x_1 + x_2 + x_3, 2x_2 + x_3, 2x_2 + 3x_3)$ có chéo hóa được không?

*Giải.* Ta thấy ma trận của $f$ đối với cơ sở chính tắc là $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 2 & 3 \end{bmatrix}$. Đa thức đặc trưng của $A$ là

$$\begin{aligned} \begin{vmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & 1 \\ 0 & 2 & 3-\lambda \end{vmatrix} &= (1-\lambda)((2-\lambda)(3-\lambda) - 2) \\ &= (1-\lambda)^2(4-\lambda). \end{aligned}$$

Vậy ma trận $A$ chỉ có giá trị riêng bội hai $\lambda = 1$ và giá trị riêng đơn $\lambda = 4$.

Với $\lambda = 1$, ta biến đổi ma trận

$$A - I_3 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\mathbf{x} \in E(1)$ khi và chỉ khi $\mathbf{x} = (x_1, -x_3, x_3) = x_1(1,0,0) + x_3(0,-1,1)$. Vậy $E(1) = \langle (1,0,0), (0,-1,1) \rangle$.

Với $\lambda = 4$, ta biến đổi ma trận

$$A - 4I_3 = \begin{bmatrix} -3 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 2 & -1 \end{bmatrix} \to \begin{bmatrix} -3 & 1 & 1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

$\mathbf{x} \in E(4)$ khi và chỉ khi $\mathbf{x} = (\frac{1}{2}x_3, \frac{1}{2}x_3, x_3) = \frac{1}{2}x_3(1,1,2)$. Vậy $E(4) = \langle (1,1,2) \rangle$.

Vậy ma trận $A$ chéo hóa được, và ma trận làm chéo nó là $C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$. Do đó, ánh xạ tuyến tính $f$ chéo hóa được và ma trận của nó đối với cơ sở $((1,0,0), (0,-1,1), (1,1,2))$ là ma trận $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. □

**9.3.5 Thí dụ. (OSV95)** Cho ma trận $M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Tính $M^n$ ứng với mọi $n$ nguyên dương cho trước.

*Giải.* Đa thức đặc trưng của ma trận $M$ là

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 = -(\lambda - 1)(\lambda - 3).$$

Vậy ma trận $M$ có hai giá trị riêng $\lambda = 1$ và $\lambda = 3$.

Với $\lambda = 1$, ta biến đổi ma trận

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Vậy một vector riêng ứng với giá trị riêng $\lambda = 1$ là $(1, -1)$.

Với $\lambda = 3$, ta biến đổi ma trận

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \to \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$$

Vậy một vector riêng ứng với giá trị riêng $\lambda = 3$ là $(1, 1)$.

Từ các kết quả trên ta có ma trận $M$ được chéo hóa bởi ma trận

$$C = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \qquad \text{Suy ra} \qquad C^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Từ đó ta biểu diễn được

$$M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = CDC^{-1}.$$

Do đó ta tính được

$$M^n = CD^nC^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3^n \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{3^n+1}{2} & \frac{3^n-1}{2} \\ \frac{3^n-1}{2} & \frac{3^n+1}{2} \end{bmatrix}.$$

□

**9.3.6 Thí dụ. (OSV98)** Cho $A = \begin{bmatrix} \frac{1}{2} & 1 & 1 \\ 0 & \frac{1}{3} & 1 \\ 0 & 0 & \frac{1}{6} \end{bmatrix}$. Đặt

$$A^n = \begin{bmatrix} a_{11}(n) & a_{12}(n) & a_{13}(n) \\ a_{21}(n) & a_{22}(n) & a_{23}(n) \\ a_{31}(n) & a_{32}(n) & a_{33}(n) \end{bmatrix}.$$

Tìm $\lim\limits_{n\to\infty} a_{ij}(n)$ với $i, j = 1, 2, 3$.

*Giải.* Đa thức đặc trưng của ma trận $A$ là $\chi(\lambda) = (\frac{1}{2} - \lambda)(\frac{1}{3} - \lambda)(\frac{1}{6} - \lambda)$. Vậy ma trận $A$ có các giá trị riêng $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{6}$. Với $\lambda = \frac{1}{2}$ ta biến đổi

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & -\frac{1}{6} & 1 \\ 0 & 0 & -\frac{1}{3} \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó $E(\frac{1}{2}) = \langle(1,0,0)\rangle$. Với $\lambda = \frac{1}{3}$ ta biến đổi

$$\begin{bmatrix} -\frac{1}{6} & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{6} \end{bmatrix} \to \begin{bmatrix} 1 & 6 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó ta được $E(\frac{1}{3}) = \langle(6,-1,0)\rangle$. Với $\lambda = \frac{1}{6}$ ta biến đổi

$$\begin{bmatrix} \frac{1}{3} & 1 & 1 \\ 0 & \frac{1}{6} & 1 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -15 \\ 0 & 1 & 6 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó ta được $E(\frac{1}{6}) = \langle(15,-6,1)\rangle$. Từ các kết quả thu được ta có biểu diễn $A = CDC^{-1}$ với

$$D = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{6} \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 6 & 15 \\ 0 & -1 & -6 \\ 0 & 0 & 1 \end{bmatrix}$$

Khi đó ta có $A^n = CD^nC^{-1}$. Rõ ràng

$$\lim_{n\to\infty} D^n = \lim_{n\to\infty} \begin{bmatrix} \frac{1}{2^n} & 0 & 0 \\ 0 & \frac{1}{3^n} & 0 \\ 0 & 0 & \frac{1}{6^n} \end{bmatrix} = O$$

cho nên $\lim\limits_{n\to\infty} A^n = O$. Do đó, $\lim\limits_{n\to\infty} a_{ij}(n) = 0$ với $i, j = 1, 2, 3$. □

**9.3.7 Thí dụ. (OSV02)** Cho $A = \begin{bmatrix} \frac{\sqrt{3}}{2} + 1 & -\frac{5}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} - 1 \end{bmatrix}$. Tính $A^{2002}$.

*Giải.* Đa thức đặc trưng của ma trận $A$ là

$$\chi(\lambda) = \Big(\frac{\sqrt{3}}{2}+1-\lambda\Big)\Big(\frac{\sqrt{3}}{2}-1-\lambda\Big)+\frac{5}{2}\cdot\frac{1}{2} = \Big(\lambda-\frac{\sqrt{3}}{2}-\frac{i}{2}\Big)\Big(\lambda-\frac{\sqrt{3}}{2}+\frac{i}{2}\Big)$$

Vậy có hai giá trị riêng $\frac{\sqrt{3}}{2} + \frac{i}{2}$ và $\frac{\sqrt{3}}{2} - \frac{i}{2}$. Với $\lambda = \frac{\sqrt{3}}{2} + \frac{i}{2}$, ta có

$$\begin{bmatrix} 1 - \frac{i}{2} & -\frac{5}{2} \\ \frac{1}{2} & -1 - \frac{i}{2} \end{bmatrix} \to \begin{bmatrix} 1 & -2 - i \\ 0 & 0 \end{bmatrix}$$

Do đó, ta được $E(\frac{\sqrt{3}}{2}+\frac{i}{2}) = \langle(2+i,1)\rangle$. Với $\lambda = \frac{\sqrt{3}}{2}-\frac{i}{2}$, ta biến đổi

$$\begin{bmatrix} 1+\frac{i}{2} & -\frac{5}{2} \\ \frac{1}{2} & -1+\frac{i}{2} \end{bmatrix} \to \begin{bmatrix} 1 & -2+i \\ 0 & 0 \end{bmatrix}$$

Do đó, ta được $E(\frac{\sqrt{3}}{2}-\frac{i}{2}) = \langle(2-i,1)\rangle$. Vậy ta có biểu diễn $A = CDC^{-1}$ với

$$D = \begin{bmatrix} \frac{\sqrt{3}}{2}+\frac{i}{2} & 0 \\ 0 & \frac{\sqrt{3}}{2}-\frac{i}{2} \end{bmatrix} \qquad C = \begin{bmatrix} 2+i & 2-i \\ 1 & 1 \end{bmatrix}$$

Do đó, ta tính được

$$\begin{aligned} A^{2002} &= CD^{2002}C^{-1} \\ &= \begin{bmatrix} 2+i & 2-i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{2002i\frac{\pi}{6}} & 0 \\ 0 & e^{-2002i\frac{\pi}{6}} \end{bmatrix} \begin{bmatrix} -\frac{i}{2} & \frac{1}{2}+i \\ \frac{i}{2} & \frac{1}{2}-i \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2}-\sqrt{3} & \frac{5}{2}\sqrt{3} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2}+\sqrt{3} \end{bmatrix} \end{aligned}$$

□

**9.3.8 Thí dụ. (OSV05)** Cho ma trận $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$. Tìm ma trận $B$ có các giá trị riêng dương sao cho $B^2 = A$.

*Giải.* Đa thức đặc trưng của ma trận $A$ là

$$\begin{vmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} = \lambda^2 - 5\lambda + 4$$

Vậy $A$ có hai giá trị riêng $\lambda = 1$ và $\lambda = 4$.

Với $\lambda = 1$, ta biến đổi ma trận

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

Vậy một vector riêng ứng với giá trị riêng $\lambda = 1$ là $(2,-1)$.

Với $\lambda = 4$, ta biến đổi ma trận

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

Vậy một vector riêng ứng với giá trị riêng $\lambda = 4$ là $(1, 1)$.

Từ các kết quả trên ta có ma trận $A$ được chéo hóa bởi ma trận $C = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$ với $C^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$. Khi đó, ta có $A = CDC^{-1}$ với $D = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. Từ đó ta lấy

$$B = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{5}{3} \end{bmatrix}.$$

Ta dễ dàng kiểm tra thấy rằng $B^2 = A$ và $B$ có hai giá trị riêng là $1$ và $2$. Ngược lại nếu $B$ là một ma trận có các giá trị riêng dương thỏa $B^2 = A$ thì các giá trị riêng đó phải là $1$ và $2$ cho nên $B$ có dạng $B = Q \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} Q^{-1}$ và thỏa $QDQ^{-1} = CDC^{-1}$. Từ đó sau những phép tính trực tiếp khá dài dòng ta suy ra được $B$ là ma trận chỉ ra như trên (nghĩa là chỉ có duy nhất một $B$ thỏa điều kiện bài toán. □

**9.3.9 Thí dụ. (OSV08)** Cho $A$ là ma trận thực vuông cấp 2 thỏa mãn điều kiện $\det A < 0$. Chứng minh rằng tồn tại hai số thực phân biệt $\lambda_1$, $\lambda_2$ và hai ma trận $A_1$, $A_2$ sao cho $A^n = \lambda_1^n A_1 + \lambda_2^n A_2$ với mọi $n = 1, 2, \ldots$.

*Giải.* Đa thức đặt trưng của $A$ là

$$\chi_A(\lambda) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A).$$

Do $\det A < 0$ nên $\Delta > 0$ vậy đa thức đặc trưng có hai nghiệm phân biệt $\lambda_1$ và $\lambda_2$. Do đó ma trận $A$ chéo hóa được, nghĩa là tồn tại ma trận $C$ khả nghịch sao cho $A = CDC^{-1}$ với

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \qquad \text{suy ra} \qquad D^n = \begin{bmatrix} \lambda_1^n & 0 & \\ 0 & \lambda_2^n & 0 \end{bmatrix}.$$

Từ đó ta viết được

$$\begin{aligned} A^n = CD^nC^{-1} &= C \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} C^{-1} \\ &= C \begin{bmatrix} \lambda_1^n & 0 \\ 0 & 0 \end{bmatrix} C^{-1} + C \begin{bmatrix} 0 & 0 \\ 0 & \lambda_2^n \end{bmatrix} C^{-1} \end{aligned}$$

$$= \lambda_1^n C \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} C^{-1} + \lambda_2^n C \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} C^{-1}.$$

Vậy $A_1 = C \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} C^{-1}$ và $A_2 = C \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} C^{-1}$. □

**9.3.10 Thí dụ. (OSV10)** Cho $\{u_n\}$, $\{v_n\}$, $\{w_n\}$ là các dãy số thực được xác định bởi $u_0 = v_0 = w_0 = 1$ và với mọi $n \in \mathbb{N}$ ta có

$$\begin{cases} u_{n+1} = -u_n - 7v_n + 5w_n \\ v_{n+1} = -2u_n - 8v_n + 6w_n \\ w_{n+1} = -4u_n - 16v_n + 12w_n \end{cases}$$

Chứng minh rằng $v_n - 2$ là số nguyên chia hết cho $2^n$.

*Giải*. Đặt $A = \begin{bmatrix} -1 & -7 & 5 \\ -2 & -8 & 6 \\ -4 & -16 & 12 \end{bmatrix}$, $\mathbf{x}_n = (u_n, v_n, w_n)$. Với mọi $n \in \mathbb{N}$ ta có $\mathbf{x}_{n+1} = A\mathbf{x}_n$ suy ra $\mathbf{x}_n = A^n\mathbf{x}_0$. Đa thức đặc trưng của $A$ là $\chi(\lambda) = -\lambda(\lambda - 1)(\lambda - 2)$. Do đó $A$ có các giá trị riêng là 0, 1, 2; và $A$ chéo hóa được. Với $\lambda = 0$, ta có

$$A = \begin{bmatrix} -1 & -7 & 5 \\ -2 & -8 & 6 \\ -4 & -16 & 12 \end{bmatrix} \to \begin{bmatrix} 3 & 0 & -1 \\ 0 & 3 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó $E(0) = \langle (1, 2, 3) \rangle$. Với $\lambda = 1$, ta có

$$A - I = \begin{bmatrix} -2 & -7 & 5 \\ -2 & -9 & 6 \\ -4 & -16 & 11 \end{bmatrix} \to \begin{bmatrix} 4 & 0 & -3 \\ 0 & 4 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó $E(1) = \langle (3, 2, 4) \rangle$. Với $\lambda = 2$, ta có

$$A - 2I = \begin{bmatrix} -3 & -7 & 5 \\ -2 & -10 & 6 \\ -4 & -16 & 10 \end{bmatrix} \to \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó, ta có $E(2) = \langle (1,1,2) \rangle$. Vậy với các kết quả đạt được ta có biểu diễn $A = CDC^{-1}$ với

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 4 & 2 \end{bmatrix}$$

Từ đó ta tính được

$$\mathbf{x}_n = CD^nC^{-1}\mathbf{x}_0 = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 4 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^n \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 2 & 2 & 1 \\ 3 & 4 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 3 - 3 \cdot 2^n \\ 2 - 3 \cdot 2^n \\ 4 - 6 \cdot 2^n \end{bmatrix}$$

Do đó ta được $v_n = 2 - 3 \cdot 2^n$, ta suy ra điều phải chứng minh. □

## Bài tập

**1**) Các ma trận (trên $\mathbb{R}$) sau có chéo hóa được hay không? Nếu chéo hóa được, hãy tìm ma trận làm chéo hóa nó và ma trận chéo tương ứng.

**(a)** $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{bmatrix}$ **(c)** $\begin{bmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{bmatrix}$

**(d)** $\begin{bmatrix} 1 & -4 & -8 \\ -4 & 7 & -4 \\ -8 & -4 & 1 \end{bmatrix}$ **(e)** $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ **(f)** $\begin{bmatrix} 15 & -18 & -16 \\ 9 & -12 & -8 \\ 4 & -4 & -6 \end{bmatrix}$

**2**) Các ma trận sau có chéo hóa được trong $\mathbb{C}$ không? Nếu chéo hóa được, hãy tìm ma trận làm chéo hóa nó.

**(a)** $\begin{bmatrix} 4 & -1 & -2 \\ 2 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$

**3)** Với giá trị nào của $m$ thì hai ma trận $A = \begin{bmatrix} 3 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ và $B = \begin{bmatrix} -2 & -2 & -2 \\ 2 & 3 & m \\ 4 & 2 & 4 \end{bmatrix}$ đồng dạng với cùng một ma trận chéo?

**4)** Cho $A = \dfrac{1}{4}\begin{bmatrix} -1 & 3 & 2 \\ 3 & -1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$. Tính $A^n$.

**5)** Các toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ sau có chéo hóa được không? Tìm cơ sở chéo hóa nếu có cho mỗi toán tử tuyến tính.

**(a)** $f(x, y, z) = (x + y, y + z, -2y - z)$

**(b)** $f(x, y, z) = (x - y + z, x + y - z, -x + y + z)$

**(c)** $f(x, y, z) = (x - y, y - z, x + z)$.

**6) (OSV96)** Cho $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 2 \end{bmatrix}^n = \begin{bmatrix} a_{11}(n) & a_{12}(n) & a_{13}(n) \\ a_{21}(n) & a_{22}(n) & a_{23}(n) \\ a_{31}(n) & a_{32}(n) & a_{33}(n) \end{bmatrix}$. Tính

$$\lim_{n\to\infty} \frac{a_{22}(n)}{a_{32}(n)}.$$

**7) (OSV97)** Giả sử $x_0, y_0, z_0$ là các số thực cho trước. Hãy xác định tất cả các số thực $x_n, y_n, z_n$ với $n = 0, 1, 2, \ldots$, thỏa hệ phương trình

$$\begin{cases} x_{n+1} = -x_n + y_n + z_n \\ y_{n+1} = x_n - y_n + z_n \\ z_{n+1} = x_n + y_n - z_n \end{cases} \qquad n = 0, 1, 2, \ldots$$

**8) (OSV05)** Cho ma trận $M = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Đặt $M^n = (b_{ij}(n))_{i,j=1,2,3}$, $n \in \mathbb{N}$, $n \geq 2$. Tính $S_n = \sum\limits_{i=1}^{3} \sum\limits_{j=1}^{3} b_{ij}(n)$.

# Chương VI

# Dạng song tuyến tính và dạng toàn phương

Chúng ta đã nghiên cứu những vấn đề "tuyến tính" trên không gian vector, nay chuyển sang các vấn đề về "bậc hai". Trong toán học, vật lý, cơ học, ... chúng ta gặp nhiều dạng toàn phương khác nhau. Như nhà toán học Dieudonné đã nói, không có lý thuyết toán học nào trong đó không có sự tham gia của một dạng toàn phương nào đấy. Đặc biệt hình học Euclid xây dựng trên một dạng toàn phương xác định dương. Đường bậc hai, mặt bậc hai chẳng qua là tập các điểm của một hàm đa thức bậc hai trên không gian nào đó; khái niệm này gắn chặt với khái niệm dạng toàn phương.

## § 1 Dạng song tuyến tính và dạng toàn phương

### 1.1 Các khái niệm cơ bản

**1.1.1 Định nghĩa.** Cho $V$ là không gian vector trên $\mathbb{K}$. Ánh xạ $\xi : V \times V \to \mathbb{K}$ được gọi là một **dạng song tuyến tính** trên $V$ nếu $\xi$ có các tính chất sau với mọi $\mathbf{x}, \mathbf{x}', \mathbf{y}, \mathbf{y}' \in V$ với mọi $\lambda \in \mathbb{K}$:

**(i)** $\xi(\mathbf{x} + \mathbf{x}', \mathbf{y}) = \xi(\mathbf{x}, \mathbf{y}) + \xi(\mathbf{x}', \mathbf{y})$.

**(ii)** $\xi(\lambda\mathbf{x}, \mathbf{y}) = \lambda\xi(\mathbf{x}, \mathbf{y})$.

**(iii)** $\xi(\mathbf{x}, \mathbf{y} + \mathbf{y}') = \xi(\mathbf{x}, \mathbf{y}) + \xi(\mathbf{x}, \mathbf{y}')$.

**(iv)** $\xi(\mathbf{x}, \lambda\mathbf{y}) = \lambda\xi(\mathbf{x}, \mathbf{y})$.

Nói cách khác, một dạng song tuyến tính trên $V$ là một hàm hai biến vector trên $V$ nhận giá trị trong $\mathbb{K}$ và tuyến tính theo từng biến.

**1.1.2 Định nghĩa.** Dạng song tuyến tính $\xi$ trên $V$ được gọi là **đối xứng** nếu $\xi$ có thêm tính chất $\xi(\mathbf{x}, \mathbf{y}) = \xi(\mathbf{y}, \mathbf{x})$ với mọi $\mathbf{x}, \mathbf{y} \in V$.

**1.1.3 Định nghĩa.** Dạng song tuyến tính $\xi$ trên $V$ được gọi là **phản xứng** nếu $\xi$ có thêm tính chất $\xi(\mathbf{x}, \mathbf{y}) = -\xi(\mathbf{y}, \mathbf{x})$ với mọi $\mathbf{x}, \mathbf{y} \in V$.

**1.1.4 Thí dụ.** Cho $V$ là một không gian vector trên $\mathbb{K}$ và $f, g$ là hai ánh xạ tuyến tính từ $V$ đến $\mathbb{K}$. Khi đó, ánh xạ $\xi : V \times V \to \mathbb{K}$ xác định bởi $\xi(\mathbf{x}, \mathbf{y}) = f(\mathbf{x})g(\mathbf{y})$ là một dạng song tuyến tính trên $V$. Đặc biệt, nếu $f = g$ thì $\xi$ là dạng song tuyến tính đối xứng. □

**1.1.5 Tính chất.** *Cho $\xi$ là một dạng song tuyến tính trên $V$. Khi đó, ta có*

**(i)** $\xi(\mathbf{0}, \mathbf{y}) = 0$ *và* $\xi(\mathbf{x}, \mathbf{0}) = 0$,

**(ii)** $\xi\Big(\sum\limits_{i=1}^{n} \lambda_i \mathbf{x}_i, \mathbf{y}\Big) = \sum\limits_{i=1}^{n} \lambda_i \xi(\mathbf{x}_i, \mathbf{y})$,

**(iii)** $\xi\Big(\mathbf{x}, \sum\limits_{j=1}^{m} \mu_j \mathbf{y}_j\Big) = \sum\limits_{j=1}^{m} \mu_j \xi(\mathbf{x}, \mathbf{y}_j)$,

**(iv)** $\xi\Big(\sum\limits_{i=1}^{n} \lambda_i \mathbf{x}_i, \sum\limits_{j=1}^{m} \mu_j \mathbf{y}_j\Big) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} \lambda_i \mu_j \xi(\mathbf{x}_i, \mathbf{y}_j)$,

*với mọi* $\mathbf{x}, \mathbf{y}, \mathbf{x}_i, \mathbf{y}_j \in V$ *với mọi* $\lambda_i, \mu_j \in \mathbb{K}$.

*Chứng minh.* Việc chứng minh các tính chất này dựa vào định nghĩa của dạng song tuyến tính và nhận xét sau: với mọi $\mathbf{x}_0, \mathbf{y}_0 \in V$ thì ánh xạ $f_{\mathbf{y}_0} : V \to \mathbb{K}$ xác định bởi $f_{\mathbf{y}_0}(\mathbf{x}) = \xi(\mathbf{x}, \mathbf{y}_0)$ và ánh xạ $g_{\mathbf{x}_0} : V \to \mathbb{K}$ xác định bởi $g_{\mathbf{x}_0}(\mathbf{y}) = \xi(\mathbf{x}_0, \mathbf{y})$ là hai ánh xạ tuyến tính. □

**1.1.6 Định nghĩa.** Cho không gian vector $V$ trên $\mathbb{K}$. Ánh xạ $p : V \to \mathbb{K}$ được gọi là **dạng toàn phương** trên $V$ nếu tồn tại một dạng song tuyến tính đối xứng $\xi$ trên $V$ sao cho $p(\mathbf{x}) = \xi(\mathbf{x}, \mathbf{x})$ với mọi $\mathbf{x} \in V$. Khi đó, ta nói $p$ là dạng toàn phương xác định bởi dạng song tuyến tính đối xứng $\xi$, còn $\xi$ gọi là *dạng cực* của dạng toàn phương $p$.

**1.1.7 Nhận xét.** Với mỗi dạng toàn phương $p$ trên $V$ và với mọi $\mathbf{x} \in V$, mọi $\lambda \in \mathbb{K}$, ta dễ dàng thấy rằng

$$p(\lambda\mathbf{x}) = \lambda^2 p(\mathbf{x}).$$

Nghĩa là $p$ có tính *thuần nhất bậc hai.*

**1.1.8 Thí dụ.** Cho $\xi : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ và $p : \mathbb{K}^n \to \mathbb{K}$ là các ánh xạ xác định bởi $\xi(\mathbf{x}, \mathbf{y}) = \sum\limits_{i=1}^{n} x_i y_i$ và $p(\mathbf{x}) = \sum\limits_{i=1}^{n} x_i^2$ với mọi $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{K}^n$. Khi đó, $\xi$ là một dạng song tuyến tính trên $\mathbb{K}^n$ và $p$ là dạng toàn phương trên $\mathbb{K}^n$. Hơn nữa, $\xi$ là dạng cực của $p$. $\square$

**1.1.9 Thí dụ.** Cho $V = C_{[a,b]}$ là không gian các hàm liên tục trên $[a, b]$ với $a, b \in \mathbb{R}$ và $a < b$. Khi đó ánh xạ $\xi : V \times V \to \mathbb{R}$ cho bởi công thức

$$\xi(\mathbf{x}, \mathbf{y}) = \int_a^b \mathbf{x}(t)\mathbf{y}(t)dt \qquad \text{với mọi } \mathbf{x}, \mathbf{y} \in V$$

là một dạng song tuyến tính đối xứng trên $V$. Nó xác định dạng toàn phương $q$ trên $V$ sau đây

$$q(\mathbf{x}) = \int_a^b \mathbf{x}^2(t)dt, \qquad \text{với mọi } \mathbf{x} \in V. \qquad \square$$

**1.1.10 Nhận xét.** Trong định nghĩa dạng toàn phương ta thấy: với mỗi dạng song tuyến tính đối xứng trên $V$ ta xác định được một dạng toàn phương. Ngược lại, với mỗi dạng toàn phương ta cũng xác định được một dạng song tuyến tính đối xứng mà nó là dạng cực của dạng toàn phương đã cho. Thật vậy, giả sử $p$ là dạng toàn phương trên $V$. Khi đó dạng cực của $p$ xác định bởi

$$\xi(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[p(\mathbf{x} + \mathbf{y}) - p(\mathbf{x}) - p(\mathbf{y})]. \tag{1.1.11}$$

**1.1.12** Ký hiệu $\mathcal{B}(V)$ là tập tất cả các dạng song tuyến tính trên $V$, $\mathcal{S}(V)$ là tập tất cả các dạng song tuyến tính đối xứng trên $V$, và $\mathcal{Q}(V)$ là tập tất cả các dạng toàn phương trên $V$. Trên $\mathcal{B}(V)$ và $\mathcal{Q}(V)$ xét hai phép toán cộng và nhân với vô hướng một cách tự nhiên như sau

$$\begin{aligned} (\xi+\eta)(\mathbf{x},\mathbf{y}) &= \xi(\mathbf{x},\mathbf{y})+\eta(\mathbf{x},\mathbf{y}) \qquad & (\lambda\xi)(\mathbf{x},\mathbf{y}) &= \lambda\xi(\mathbf{x},\mathbf{y}) \\ (p+p')(\mathbf{x}) &= p(\mathbf{x})+p'(\mathbf{x}) & (\lambda p)(\mathbf{x}) &= \lambda(p(\mathbf{x})) \end{aligned}$$

với mọi $\mathbf{x},\mathbf{y}\in V$ mọi $\lambda\in\mathbb{K}$ và mọi $\xi,\eta\in\mathcal{B}(V)$ mọi $p,p'\in\mathcal{Q}(V)$. Khi đó dễ dàng chứng tỏ rằng $\mathcal{B}(V)$ và $\mathcal{Q}(V)$ cùng với hai phép toán trên là các $\mathbb{K}$-không gian vector. Bên cạnh đó, ta cũng có $\mathcal{S}(V)$ là một không gian con của $\mathcal{B}(V)$.

## 1.2 Ma trận và biểu thức tọa độ

**1.2.1 Định nghĩa.** Cho $V$ là không gian $n$ chiều trên $\mathbb{K}$, cho dạng song tuyến tính $\xi$ trên $V$, và cho $\mathscr{A}=(\mathbf{a}_1,\dots,\mathbf{a}_n)$ là cơ sở của $V$. Đặt $a_{ij}=\xi(\mathbf{a}_i,\mathbf{a}_j)$. Khi đó, ma trận vuông $(a_{ij})_{n\times n}$ cấp $n$ được gọi là **ma trận của $\xi$ trong cơ sở $\mathscr{A}$**, kí hiệu $A|_{\mathscr{A}}$.

**1.2.2** Với giả thiết và kí hiệu như trong định nghĩa trên, cho $\mathbf{x}|_{\mathscr{A}}=(x_1,\dots,x_n)$ và $\mathbf{y}|_{\mathscr{A}}=(y_1,\dots,y_n)$. Ta có

$$\xi(\mathbf{x},\mathbf{y})=\xi\Big(\sum_{i=1}^n x_i\mathbf{a}_i,\sum_{j=1}^n y_j\mathbf{a}_j\Big)=\sum_{i=1}^n\sum_{j=1}^n x_iy_j\xi(\mathbf{a}_i,\mathbf{a}_j).$$

Vậy ta viết lại

$$\xi(\mathbf{x},\mathbf{y})=\sum_{i=1}^n\sum_{j=1}^n x_iy_ja_{ij}=\mathbf{x}|_{\mathscr{A}}^t A\mathbf{y}|_{\mathscr{A}}. \tag{1.2.3}$$

Công thức trên được gọi là **biểu thức tọa độ của $\xi$ trong cơ sở $\mathscr{A}$**.

**1.2.4 Thí dụ.** Xét không gian vector $\mathbb{R}^3$ và ánh xạ $\xi$ xác định bởi $\xi(\mathbf{x},\mathbf{y})=x_1y_1-2x_1y_2+x_2y_1+x_1y_3+3x_2y_3$ với $\mathbf{x}=(x_1,x_2,x_3)$ và $\mathbf{y}=$

$(y_1, y_2, y_3)$. Ta có thể chứng minh được $\xi$ là một dạng song tuyến tính trên $\mathbb{R}^3$. Ma trận của $\xi$ trong cơ sở chính tắc là

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}. \qquad \square$$

**1.2.5** Bây giờ ta có giả thiết thêm là $\xi$ là dạng song tuyến tính đối xứng. Khi đó $\xi(\mathbf{a}_i, \mathbf{a}_j) = \xi(\mathbf{a}_j, \mathbf{a}_i)$ với mọi $i, j = 1, \ldots, n$ hay $a_{ij} = a_{ji}$ với mọi $i, j = 1, \ldots, n$. Do đó, ma trận $A|_{\mathscr{A}}$ của $\xi$ trong cơ sở $\mathscr{A}$ là *ma trận đối xứng.*

Ngược lại, nếu dạng song tuyến tính $\xi$ trên $V$ có ma trận $A|_{\mathscr{A}}$ trong cơ sở $\mathscr{A}$ là ma trận đối xứng thì ta có thể chứng minh được $\xi$ là dạng song tuyến tính đối xứng. Vậy ta có mệnh đề sau.

**1.2.6 Mệnh đề.** *Dạng song tuyến tính $\xi$ trên $V$ là đối xứng khi và chỉ khi ma trận của nó trong một cơ sở nào đó của $V$ là ma trận đối xứng.*

**1.2.7 Thí dụ.** Xét không gian vector $\mathbb{R}^3$ và xét ánh xạ $\xi$ xác định bởi $\xi(\mathbf{x}, \mathbf{y}) = x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_1y_3 + x_3y_1 - x_2y_2 + 3x_3y_3$ với $\mathbf{x} = (x_1, x_2, x_3)$ và $\mathbf{y} = (y_1, y_2, y_3)$. Ta có $\xi$ là một dạng song tuyến tính đối xứng trên $\mathbb{R}^3$ và ma trận của nó trong cơ sở chính tắc là

$$A = \begin{bmatrix} 1 & -2 & 1 \\ -2 & -1 & 0 \\ 1 & 0 & 3 \end{bmatrix}.$$

Khi đó, dạng toàn phương xác định bởi $\xi$ là $p(\mathbf{x}) = x_1^2 - 4x_1x_2 + 2x_1x_3 - x_2^2 + 3x_3^2$ với $\mathbf{x} = (x_1, x_2, x_3)$. $\square$

**1.2.8** Giả sử $p$ là một dạng toàn phương trên $V$ và $\xi$ là dạng cực của $p$. Gọi $A = (a_{ij})_{n \times n}$ là ma trận của $\xi$ trong cơ sở $\mathscr{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Ta nhận thấy $A$ là ma trận đối xứng. Với mọi $\mathbf{x} \in V$ mà $\mathbf{x}|_{\mathscr{A}} = (x_1, \ldots, x_n)$, ta có

$$
\begin{aligned}
p(\mathbf{x}) &= \xi(\mathbf{x},\mathbf{x}) = \mathbf{x}|_{\mathscr{A}}^{t} A\mathbf{x}|_{\mathscr{A}} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \qquad (1.2.9)\\
&= \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{1\le i<j\le n} a_{ij}x_ix_j.
\end{aligned}
$$

Ma trận đối xứng $A$ cũng được gọi là **ma trận của $p$ trong cơ sở $\mathscr{A}$.** Công thức trên được gọi là **biểu thức tọa độ của $p$ trong cơ sở $\mathscr{A}$.**

Ngược lại, nếu ánh xạ $p : V \to \mathbb{K}$ xác định bởi $p(\mathbf{x}) = \sum_{1\le i\le j\le n} b_{ij}x_ix_j$ với mọi $\mathbf{x} \in V$ và $(x_1, \ldots, x_n)$ là tọa độ của $\mathbf{x}$ đối với cơ sở $\mathscr{A}$, thì $p$ là dạng toàn phương trên $V$ với ma trận trong cơ sở $\mathscr{A}$ là $A = (a_{ij})$ xác định bởi

$$
(1.2.10) \qquad \begin{cases} a_{ij} = a_{ji} = \frac{1}{2}b_{ij} & 1 \le i < j \le n \\ a_{ii} = b_{ii} & i = 1, \ldots, n. \end{cases}
$$

**1.2.11 Thí dụ.** Xét dạng toàn phương $p$ trên không gian vector $\mathbb{R}^3$ xác định bởi $p(\mathbf{x}) = x_1^2 - x_2^2 + x_3^2 - x_1x_2 - 2x_2x_3 + 3x_1x_3$ với $\mathbf{x} = (x_1, x_2, x_3)$. Khi đó, ma trận của $p$ trong cơ sở chính tắc là

$$
A = \begin{bmatrix} 1 & -\frac{1}{2} & \frac{3}{2} \\ -\frac{1}{2} & -1 & -1 \\ \frac{3}{2} & -1 & 1 \end{bmatrix},
$$

và dạng cực của $p$ được xác định bởi $\xi(\mathbf{x},\mathbf{y}) = x_1y_1 - \frac{1}{2}x_1y_2 + \frac{3}{2}x_1y_3 - \frac{1}{2}x_2y_1 - x_2y_2 - x_2y_3 + \frac{3}{2}x_3y_1 - x_3y_2 + x_3y_3$. □

**1.2.12 Mệnh đề.** *Nếu dạng song tuyến tính $\xi$ và $\xi'$ trên $V$ có ma trận trong cơ sở $\mathscr{A}$ lần lượt là $A$ và $A'$, thì $\xi + \xi'$ và $\lambda\xi$ là các dạng song tuyến tính trên $V$ có ma trận trong cơ sở $\mathscr{A}$ lần lượt là $A + A'$ và $\lambda A$ (với mọi $\lambda \in \mathbb{K}$).*

*Chứng minh.* Theo giả thiết ta có biểu thức

$$
\xi(\mathbf{x},\mathbf{y}) = \mathbf{x}|_{\mathscr{A}}^{t} A\mathbf{y}|_{\mathscr{A}} \qquad \text{và} \qquad \xi'(\mathbf{x},\mathbf{y}) = \mathbf{x}|_{\mathscr{A}}^{t} A'\mathbf{y}|_{\mathscr{A}}.
$$

Do đó, ta tính được

$$\begin{aligned}(\xi+\xi')(\mathbf{x},\mathbf{y}) &= \xi(\mathbf{x},\mathbf{y})+\xi'(\mathbf{x},\mathbf{y})\\ &= \mathbf{x}|_{\mathscr{A}}^t A\mathbf{y}|_{\mathscr{A}}+\mathbf{x}|_{\mathscr{A}}^t A'\mathbf{y}|_{\mathscr{A}}\\ &= \mathbf{x}|_{\mathscr{A}}^t (A+A')\mathbf{y}|_{\mathscr{A}}.\end{aligned}$$

Từ đó ta thấy được $\xi+\xi'$ là một dạng song tuyến tính và ma trận của nó trong cơ sở $\mathscr{A}$ là $A+A'$.

Mặt khác, ta có

$$(\lambda\xi)(\mathbf{x},\mathbf{y})=\lambda\xi(\mathbf{x},\mathbf{y})=\lambda(\mathbf{x}|_{\mathscr{A}}^t A\mathbf{y}|_{\mathscr{A}})=\mathbf{x}|_{\mathscr{A}}^t(\lambda A)\mathbf{y}|_{\mathscr{A}}.$$

Do đó, ta có thể thấy được $\lambda\xi$ là một dạng song tuyến tính và ma trận của nó trong cơ sở $\mathscr{A}$ là $\lambda A$. □

**1.2.13 Mệnh đề.** *Nếu dạng toàn phương $p$ và $p'$ trên $V$ có ma trận trong cơ sở $\mathscr{A}$ lần lượt là $A$ và $A'$, thì $p+p'$ và $\lambda p$ là các dạng toàn phương trên $V$ có ma trận trong cơ sở $\mathscr{A}$ lần lượt là $A+A'$ và $\lambda A$ (với mọi $\lambda\in\mathbb{K}$).*

*Chứng minh.* Gọi $\xi$ và $\xi'$ là dạng cực của các dạng toàn phương $p$ và $p'$, suy ra chúng có ma trận trong cơ sở $\mathscr{A}$ lần lượt là $A$ và $A'$ (là các ma trận đối xứng). Theo mệnh đề trên thì $\xi+\xi'$ và $\lambda\xi$ là các dạng song tuyến tính trên $V$ có ma trận trong cơ sở $\mathscr{A}$ lần lượt là $A+A'$ và $\lambda A$. Do $A$ và $A'$ là các ma trận đối xứng nên $A+A'$ và $\lambda A$ là các ma trận đối xứng. Do đó, $\xi+\xi'$ và $\lambda\xi$ là các dạng song tuyến tính đối xứng. Chúng ta có thể kiểm tra được các dạng toàn phương xác định bởi $\xi+\xi'$ và $\lambda\xi$ lần lượt là $p+p'$ và $\lambda p$. □

**1.2.14** Một dạng toàn phương trên $\mathbb{K}^n$ còn gọi là **dạng toàn phương $n$ biến** trên $\mathbb{K}$. Như vậy, ánh xạ $p:\mathbb{K}^n\to\mathbb{K}$ với $(x_1,x_2,\ldots,x_n)\mapsto p(x_1,x_2,\ldots,x_n)$ và $p\neq 0$ là một dạng toàn phương $n$ biến trên $\mathbb{K}$ khi và chỉ khi $p(x_1,x_2,\ldots,x_n)$ là một *đa thức đẳng cấp bậc hai của $n$ biến $x_1,x_2,\ldots,x_n$*. Đây là đặc trưng cơ bản của một dạng toàn phương (khác hằng không) $n$ biến trên $\mathbb{K}$.

**1.2.15 Thí dụ. (a)** Cho ánh xạ $p:\mathbb{R}^3\to\mathbb{R}$ xác định bởi

$$p(x,y,z)=3x^2+4xy-2xz+y^2+6yz-2z^2$$

với mọi $(x, y, z) \in \mathbb{R}^3$. Vì $p(x, y, z)$ là một đa thức đẳng cấp bậc hai của 3 biến thực $x, y, z$ nên $p$ là một dạng toàn phương 3 biến thực. Theo công thức (1.2.9) ma trận của $p$ trong cơ sở chính tắc của $\mathbb{R}^3$ là

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

**(b)** Ánh xạ $p : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto p(x, y, z) = xy - xz + yz$, cũng là một dạng toàn phương 3 biến thực với ma trận trong cơ sở chính tắc là

$$A = \begin{bmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}.$$

**(c)** Tuy nhiên ánh xạ $f : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto f(x, y, z) = x^2 + y - z^2$, không là một dạng toàn phương 3 biến thực vì $f(x, y, z)$ không phải là một đa thức đẳng cấp bậc hai của 3 biến $x, y, z$, cụ thể $f(x, y, z)$ chứa đơn thức $y$ bậc nhất. □

## Bài tập

**1)** Giả sử $\xi$ là một dạng song tuyến tính đối xứng trên không gian vector $V$. Với mọi $\mathbf{x}, \mathbf{y} \in V$, chứng minh rằng

$$\xi(\mathbf{x}, \mathbf{y}) = \frac{1}{4}[\xi(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) - \xi(\mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y})].$$

**2)** Xét dạng song tuyến tính $\xi$ trong Thí dụ 1.1.4. Chứng minh rằng nếu $\xi(\mathbf{x}, \mathbf{x}) = 0$ với mọi $\mathbf{x} \in V$ thì hoặc $f = 0$ hoặc $g = 0$.

**3)** Cho $f$ và $g$ là hai dạng tuyến tính trên $\mathbb{K}$-không gian vector $V$. Với điều kiện nào thì ánh xạ $p : V \to \mathbb{K}$ xác định bởi $p(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ là một dạng toàn phương? Khi đó, xác định dạng cực của $p$.

**4)** Cho $V$ là một $\mathbb{K}$-không gian vector. Chứng minh rằng $\mathcal{B}(V)$ và $\mathcal{Q}(V)$ là các $\mathbb{K}$-không gian vector, và $\mathcal{S}(V)$ là không gian con của $\mathcal{B}(V)$.

**5**) Chứng minh nhận xét 1.1.10.

**6**) Cho $V$ là một không gian vector $n$ chiều trên $\mathbb{K}$, cho $\mathscr{A}$ là một cơ sở trong $V$, và cho $A$ là một ma trận vuông cấp $n$ trên $\mathbb{K}$. Chứng minh rằng ánh xạ $\xi : V \times V \to \mathbb{K}$ xác định bởi $\xi(\mathbf{x}, \mathbf{y}) = \mathbf{x}|^t_{\mathscr{A}} A\mathbf{y}|_{\mathscr{A}}$ là một dạng song tuyến tính trên $V$.

**7**) Chứng minh rằng mọi dạng song tuyến tính đều là tổng của một dạng song tuyến tính đối xứng và một song tuyến tính phản xứng.

**8**) Cho $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$. Xét xem các ánh xạ $\xi$ cho dưới đây có phải là dạng song tuyến tính trên $\mathbb{R}^2$ hay không?

**(a)** $\xi(\mathbf{x}, \mathbf{y}) = 2x_1y_2 - 3x_2y_1$

**(b)** $\xi(\mathbf{x}, \mathbf{y}) = x_1 + y_2$

**(c)** $\xi(\mathbf{x}, \mathbf{y}) = 3x_2y_2$

**(d)** $\xi(\mathbf{x}, \mathbf{y}) = 1$

**9**) Cho $f$ là dạng song tuyến tính trên $V$. Với mọi tập con $N$ của $V$ ta đặt

$$N^\perp = \{\mathbf{v} \in V : f(\mathbf{u}, \mathbf{v}) = 0 \text{ với mọi } \mathbf{u} \in N\}$$
$$N^\top = \{\mathbf{v} \in V : f(\mathbf{v}, \mathbf{u}) = 0 \text{ với mọi } \mathbf{u} \in N\}$$

Chứng minh rằng

**(a)** $N^\perp$ và $N^\top$ là các không gian con của $V$.

**(b)** Nếu $N_1 \subseteq N_2$ thì $N_2^\perp \subseteq N_1^\perp$ và $N_2^\top \subseteq N_1^\top$

**(c)** $\{\mathbf{0}\}^\perp = \{\mathbf{0}\}^\top = V$.

**(d)** $N^\perp = \langle N\rangle^\perp$ và $N^\top = \langle N\rangle^\top$.

**10**) Trên $\mathbb{R}^3$ với biểu thức tọa độ dạng cực của dạng toàn phương $p$ là $\xi(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 - x_3y_3$. Với ký hiệu như bài tập trên, hãy tìm $\langle(1, 1, 2)\rangle^\perp$ và $\langle(1, 1, 2)\rangle^\top$.

**11**) Chứng minh Mệnh đề 1.2.6.

**12)** Tìm ma trận của các dạng toàn phương trên $\mathbb{R}^3$ sau trong cơ sở chính tắc

**(a)** $2x_1^2 + 3x_2^2 + x_3^2 + x_1x_2 - 2x_1x_3 + 3x_2x_3$

**(b)** $x_1^2 + 2x_2^2 + x_3^2 + x_1x_2 - 2x_1x_3 + 3x_2x_3$

**(c)** $x_1^2 + x_2^2 - 3x_3^2 + 2x_1x_2 - 4x_2x_3$

**(d)** $(x_1 + x_2 - x_3)(x_1 - x_2)$

**13)** Cho $\mathcal{M}_2(\mathbb{R})$ và $P = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$. Xét ánh xạ

$$\begin{aligned} f : \mathcal{M}_2(\mathbb{R}) \times \mathcal{M}_2(\mathbb{R}) &\to \mathbb{R} \\ (M, N) &\mapsto f(M, N) = \operatorname{tr}(M^t P N) \end{aligned}$$

**(a)** Chứng minh $f$ là dạng song tuyến tính trên $\mathcal{M}_2(\mathbb{R})$

**(b)** Tìm ma trận của $f$ trong cơ sở chính tắc của $\mathcal{M}_2(\mathbb{R})$.

**14)** Cho $\xi_1$ và $\xi_2$ là hai dạng tuyến tính trên không gian vector $n$ chiều $V$.

**(a)** Hỏi $\mathbf{v} \mapsto \xi_1(\mathbf{v})\xi_2(\mathbf{v})$ có phải là một dạng toàn phương trên $V$ không? Nếu có, dạng cực của nó là gì?

**(b)** Chứng minh rằng nếu $\xi_1(\mathbf{v})\xi_2(\mathbf{v}) = 0$ với mọi $\mathbf{v} \in V$ thì hoặc $\xi_1 = 0$ hoặc $\xi_2 = 0$.

**15)** Cho $\xi$ là dạng song tuyến tính đối xứng trên không gian vector $\mathbb{R}^3$ và có ma trận trong cơ sở chính tắc là $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & 1 \end{bmatrix}$.

**(a)** Tìm $\operatorname{Rad} \xi = \{\mathbf{x} : \xi(\mathbf{x}, \mathbf{y}) = 0 \text{ với mọi } \mathbf{y} \in \mathbb{R}^3\}$ (Tập $\operatorname{Rad} \xi$ được gọi là **hạch** của $\xi$).

**(b)** Tìm tất cả các vector $\mathbf{a}$ sao cho $\xi(\mathbf{a}, \mathbf{a}) = 0$ (gọi là các **vector $\xi$-đẳng hướng**).

**(c)** Tìm không gian con 2 chiều $W$ sao cho $\xi|_W = 0$. Khi đó xác định $W^\circ = \{\mathbf{x} : \xi(\mathbf{x}, \mathbf{y}) = 0$ với mọi $\mathbf{y} \in W\}$.

**16 )** Chứng minh rằng dạng song tuyến tính $\xi$ trên không gian vector hữu hạn chiều $V$ là phản xứng khi và chỉ khi $\xi(\mathbf{x}, \mathbf{x}) = 0$ với mọi $\mathbf{x} \in V$. Hơn nữa, chứng minh rằng nếu $\xi$ là phản xứng thì hạng của $\xi$ là một số chẵn.

# § 2 Biểu thức tọa độ khi thay đổi cơ sở và hạng của dạng toàn phương

## 2.1 Biểu thức tọa độ khi thay đổi cơ sở

**2.1.1** Cho $\xi$ là một dạng song tuyến tính trên $\mathbb{K}$-không gian vector $n$ chiều $V$. Giả sử trong $V$ có hai cơ sở $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ và $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$. Giả sử $C$ là ma trận chuyển cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$. Gọi $A$ và $B$ tương ứng là ma trận của $\xi$ trong cơ sở $\mathscr{A}$ và $\mathscr{B}$. Khi đó, ta có biểu thức tọa độ của $\xi$ trong các cơ sở $\mathscr{A}$ và $\mathscr{B}$ là

$$\xi(\mathbf{x}, \mathbf{y}) = \mathbf{x}|_{\mathscr{A}}^t A \mathbf{y}|_{\mathscr{A}} \qquad \text{và} \qquad \xi(\mathbf{x}, \mathbf{y}) = \mathbf{x}|_{\mathscr{B}}^t B \mathbf{y}|_{\mathscr{B}}.$$

Mặt khác, ta có công thức đổi tọa độ từ $\mathscr{A}$ sang $\mathscr{B}$ là

$$\mathbf{x}|_{\mathscr{A}} = C\mathbf{x}|_{\mathscr{B}} \qquad \text{và} \qquad \mathbf{y}|_{\mathscr{A}} = C\mathbf{y}|_{\mathscr{B}}.$$

Do đó, ta có

$$\xi(\mathbf{x}, \mathbf{y}) = \mathbf{x}|_{\mathscr{A}}^t A \mathbf{y}|_{\mathscr{A}} = (C\mathbf{x}_{\mathscr{B}})^t AC\mathbf{y}_{\mathscr{B}} = \mathbf{x}|_{\mathscr{B}}^t C^t AC\mathbf{y}|_{\mathscr{B}}.$$

Vậy ta có công thức mối liên hệ giữa các ma trận của $\xi$ trong các cơ sở $\mathscr{A}$ và $\mathscr{B}$ là

$$B = C^t AC. \tag{2.1.2}$$

**2.1.3 Nhận xét.** Ta nhận thấy $A$ và $B$ là hai ma trận tương đương nhau. Các kết quả trên vẫn đúng đối với dạng toàn phương.

**2.1.4 Thí dụ.** Trên không gian vector $\mathbb{R}^3$ cho dạng song tuyến tính $\xi(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_3 + 3x_3y_2$ với mọi $\mathbf{x} = (x_1, x_2, x_3)$, $\mathbf{y} = (y_1, y_2, y_3)$. Tìm ma trận của $\xi$ trong cơ sở $\mathscr{A} = ((1, 0, 0), (1, 1, 0), (1, 1, 1))$.

Ta có ma trận của $\xi$ trong cơ sở chính tắc là

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 3 & 0 \end{bmatrix}.$$

Ma trận chuyển cơ sở từ cơ sở chính tắc sang cơ sở $\mathscr{A}$ là

$$C = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Vậy ma trận của $\xi$ trong cơ sở $\mathscr{A}$ là

$$B = C^t A C = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 4 & 5 \end{bmatrix}. \quad \square$$

## 2.2 Hạng của dạng toàn phương

**2.2.1 Mệnh đề.** *Giả sử $p$ là dạng toàn phương trên không gian vector $n$ chiều $V$ và $\mathscr{A}$, $\mathscr{B}$ là hai cơ sở của $V$. Nếu $A$ và $B$ là ma trận của $p$ trong cơ sở $\mathscr{A}$ và $\mathscr{B}$ theo tương ứng, thì* $\operatorname{rank} A = \operatorname{rank} B$.

*Chứng minh.* Theo công thức mối liên hệ giữa các ma trận của $p$, công thức (2.1.2), ta có

$$B = C^t A C,$$

trong đó $C$ là ma trận đổi cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$, cho nên $C$ là ma trận không suy biến. Từ đó ta có được $\operatorname{rank} B = \operatorname{rank} A$. $\square$

**2.2.2 Định nghĩa.** Cho dạng toàn phương $p$ trên không gian vector $n$ chiều $V$, và $A$ là ma trận của $p$ trong cơ sở nào đó của $V$. Khi đó, hạng của ma trận $A$ được gọi là **hạng** của dạng toàn phương $p$, kí hiệu $\operatorname{rank} p$. Nếu $\operatorname{rank} p = n$, thì ta nói dạng toàn phương $p$ *không suy biến*. Nếu $\operatorname{rank} p < n$, thì ta nói dạng toàn phương $p$ *suy biến*.

**2.2.3 *Ghi chú.*** Tương tự ta cũng có thể nêu khái niệm hạng của một dạng song tuyến tính và tính suy biến của một dạng song tuyến tính.

**2.2.4 Thí dụ.** Trong không gian vector $\mathbb{R}^3$, xét dạng toàn phương

$$p(\mathbf{x}) = x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$$

với $\mathbf{x} = (x_1, x_2, x_3)$. Ta có ma trận của $p$ trong cơ sở chính tắc là

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

Thực hiện các phép biến đổi sơ cấp trên dòng của ma trận $A$ để tìm hạng của nó ta được

$$A \to \begin{bmatrix} 1 & 2 & 1 \\ 0 & -3 & -1 \\ 0 & -1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 \\ 0 & -3 & -1 \\ 0 & 0 & \frac{7}{3} \end{bmatrix}.$$

Vậy $\operatorname{rank} A = 3$. Do đó, hạng của dạng toàn phương $p$ bằng 3, hay $\operatorname{rank} p = 3$. Hơn nữa, $p$ là dạng toàn phương không suy biến. □

## Bài tập

**1)** Cho $\xi$ là dạng song tuyến tính trên $\mathbb{R}^2$ xác định bởi

$$\xi(\mathbf{x}, \mathbf{y}) = 3x_1y_1 - 2x_1y_2 + 4x_2y_1 - x_2y_2$$

với $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$.

**(a)** Tìm ma trận $A$ của $f$ trong cơ sở $\mathscr{B} = (\mathbf{b}_1 = (1, 1), \mathbf{b}_2 = (1, 2))$.

**(b)** Tìm ma trận $B$ của $f$ trong cơ sở $\mathscr{B}' = (\mathbf{b}'_1 = (1, -1), \mathbf{b}'_2 = (3, 1))$.

**(c)** Tìm ma trận chuyển cơ sở từ $\mathscr{B}$ sang $\mathscr{B}'$ và kiểm chứng $B = C^tAC$.

**2)** Trên không gian $\mathbb{R}^3$ cho dạng song tuyến tính $\xi$. Tìm ma trận của $\xi$ trong cơ sở $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, nếu với mọi $\mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$, $\xi$ được cho bởi

**(a)** $\xi(\mathbf{x},\mathbf{y}) = x_1y_1 + 2x_2y_2 + 3x_3y_3$; với

$$\mathbf{u}_1 = (1,1,1), \mathbf{u}_2 = (1,1,-1), \mathbf{u}_3 = (1,-1,-1)$$

**(b)** $\xi(\mathbf{x},\mathbf{y}) = x_1y_1 + x_2y_3 + x_3y_1$; với

$$\mathbf{u}_1 = (1,0,0), \mathbf{u}_2 = (1,1,0), \mathbf{u}_3 = (1,1,1).$$

**3)** Trên không gian vector $\mathbb{R}^3$ cho dạng song tuyến tính $\xi(\mathbf{x},\mathbf{y}) = x_1y_1 + 2x_2y_2 + 3x_3y_3$ với mọi $\mathbf{x} = (x_1,x_2,x_3)$ và $\mathbf{y} = (y_1,y_2,y_3)$. Tìm ma trận của $\xi$ trong cơ sở $(\mathbf{u}_1 = (1,1,1), \mathbf{u}_2 = (1,1,-1), \mathbf{u}_3 = (1,-1,-1))$.

**4)** Cho dạng toàn phương 3 biến thực $p : \mathbb{R}^3 \to \mathbb{R}$ xác định bởi

$$p(x,y,z) = 3x^2 + 12xy - 6xz + 8y^2 - 28yz - 12z^2.$$

**(a)** Lập ma trận và biểu thức tọa độ của $p$ trong cơ sở

$$\mathscr{B} = ((1,1,0),(0,1,1),(1,0,1))$$

**(b)** Lập ma trận và biểu thức tọa độ của $p$ trong cơ sở

$$\mathscr{B}' = ((1,0,0),(-2,1,0),(5,-2,1))$$

**5)** Biết dạng toàn phương $p$ trên không gian $\mathbb{R}^3$ có ma trận trong cơ sở $\mathscr{B} = ((0,0,1),(0,1,1),(1,1,1))$ là

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

Viết biểu thức tọa độ của $p$ trong cơ sở chính tắc của $\mathbb{R}^3$.

**6)** Trên không gian $n$ chiều $V$ cho dạng song tuyến tính $\xi$ và hai cơ sở $\mathscr{B}, \mathscr{B}'$ với $C$ là ma trận đổi cơ sở từ $\mathscr{B}$ sang $\mathscr{B}'$. Tìm ma trận của $\xi$ trong cơ sở $\mathscr{B}'$ biết $C$ và biểu thức tọa độ của $\xi$ trong $\mathscr{B}$ cho như sau:

**(a)** $\xi(\mathbf{x},\mathbf{y}) = x_1y_1 + x_1y_2 + x_2y_1 - x_2y_2$, $C = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$; $\mathbf{x}|_{\mathscr{B}} = (x_1,x_2)$, $\mathbf{y}|_{\mathscr{B}} = (y_1,y_2)$.

**(b)** $\xi(\mathbf{x},\mathbf{y}) = x_1y_2 + x_2y_3 + x_3y_4$, $C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$; $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, x_3, x_4)$, $\mathbf{y}|_{\mathscr{B}} = (y_1, y_2, y_3, y_4)$.

**7 )** Tìm ma trận, hạng và tính suy biến của dạng toàn phương trên $\mathbb{R}^3$ có biểu thức tọa độ

$$p(\mathbf{x}) = 2x_1^2 + 4x_2^2 + 9x_3^2 + 4x_1x_2 - 6x_1x_3 - 2x_2x_3$$

với mọi $\mathbf{x} = (x_1, x_2, x_3)$.

**8 )** Xác định hạng và tính suy biến của cac dạng toàn phương 3 biến thực sau

**(a)** $p_1(x,y,z) = 2x^2 - 3y^2 - 5z^2$

**(b)** $p_2(x,y,z) = 2xy + 2xz + 2\lambda yz$

**(c)** $p_3(x,y,z) = x^2 + 2xy + 6xz + 2y^2 + 8yz + \lambda z^2$

**9 )** Cho dạng song tuyến tính $\xi$ trên $\mathbb{R}^3$ có ma trận trong cơ sở $\mathscr{B} = ((1,0,0),(-1,1,0),(0,-1,1))$ là $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & a \end{bmatrix}$

**(a)** Tìm ma trận $A$ của dạng song tuyến tính $\xi$ trong cơ sở chính tắc của $\mathbb{R}^3$.

**(b)** Với giá trị nào của $a$ thì ma trận $A$ có hai giá trị riêng?

**(c)** Dạng song tuyến tính trên có đối xứng không? Nếu có, hãy tìm dạng toàn phương tương ứng trong cơ sở chính tắc.

**10 )** Biết dạng toàn phương 3 biến thực $p$ có ma trận trong cơ sở $\mathscr{B} = ((0,0,1),(0,1,1),(1,1,1))$ là $\begin{bmatrix} 2 & -3 & 1 \\ -3 & 2 & 4 \\ 1 & 4 & -5 \end{bmatrix}$. Viết biểu thức tọa độ của $p$ trong cơ sở chính tắc của $\mathbb{R}^3$.

**11)** Trên không gian $\mathbb{R}^3$ với cơ sở chính tắc cho dạng toàn phương $p(x, y, z) = x^2 + yz + xz$.

**(a)** Viết ma trận $A$ của $p$ trong cơ sở chính tắc của $\mathbb{R}^3$.

**(b)** Viết ma trận $B$ của $p$ trong cơ sở $\mathscr{B} = ((1, 1, 1), (1, 1, -1), (1, -1, -1))$

**(c)** Tính hạng của $p$.

# § 3 Dạng chính tắc của dạng toàn phương

Giống như các toán tử tuyến tính, mỗi dạng toàn phương $p$ trên $\mathbb{K}$-không gian vector $n$ chiều $V$ (với một cơ sở đã chọn) đều được biểu diễn bởi một ma trận (đối xứng) và cùng với nó là một biểu thức tọa độ dạng đa thức đẳng cấp bậc hai của các tọa độ (nếu $p \neq 0$). Dạng toàn phương $p$ sẽ trở nên đơn giản, dễ khảo sát nếu ta khéo chọn cơ sở trong $V$ để ma trận (và do đó cả biểu thức tọa độ) của $p$ có dạng đơn giản (ta sẽ xét sau này là ma trận chéo).

## 3.1 Khái niệm và sự tồn tại dạng chính tắc của dạng toàn phương

**3.1.1 Định nghĩa.** Cho $V$ là một không gian vector $n$ chiều trên $\mathbb{K}$ và $p$ là một dạng toàn phương trên $V$. Nếu có một cơ sở $\mathscr{B}$ của $V$ sao cho ma trận của $p$ trong $\mathscr{B}$ là ma trận chéo

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

thì biểu thức tọa độ của $p$ trong $\mathscr{B}$ có dạng

(3.1.2)
$$p(\mathbf{x}) = a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2 \qquad \text{với } \mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \dots, x_n),$$

được gọi là **dạng chính tắc** của $p$. Khi đó, các hệ số $a_1, a_2, \dots, a_n$ ở trên được gọi là **hệ số chính tắc**, và cơ sở $\mathscr{B}$ của $V$ được gọi là *cơ sở chính tắc đối với $p$* hay **cơ sở $p$-chính tắc**.

**3.1.3 Nhận xét.** Nếu dạng toàn phương $p$ có dạng chính tắc (3.1.2), thì

**(i)** $\operatorname{rank}(p)$ là số các hệ số chính tắc khác không.

**(ii)** $p$ suy biến khi và chỉ khi $a_1 a_2 \cdots a_n = 0$

**(iii)** $p$ không suy biến khi và chỉ khi $a_1 a_2 \cdots a_n \neq 0$.

**3.1.4 Định lý.** *Đối với mọi dạng toàn phương $p$ trên $\mathbb{K}$-không gian vector $n$ chiều $V$, luôn tồn tại cơ sở $p$-chính tắc trong $V$.*

*Chứng minh.* Ta chứng minh bằng phương pháp quy nạp theo số chiều $n$ của không gian vector $V$. Khi $n = 1$, khẳng định của định lý hiển nhiên đúng, mọi cơ sở đều là cơ sở $p$-chính tắc.

Giả sử khẳng định của định lý đúng với mọi dạng toàn phương trên mọi $\mathbb{K}$-không gian vector $n-1$ chiều. Giả sử $p$ là một dạng toàn phương trên $\mathbb{K}$-không gian vector $n$ chiều $V$. Nếu $p = 0$ thì mọi cơ sở của $V$ đều là $p$-chính tắc. Ta xét trường hợp $p \neq 0$. Gọi $\xi$ là dạng cực của $p$. Do $p \neq 0$ nên tồn tại $\mathbf{a}_1 \in V$ sao cho $\mathbf{a}_1 \neq \mathbf{0}$ và $p(\mathbf{a}_1) = \xi(\mathbf{a}_1, \mathbf{a}_1) = a_1 \neq 0$. Bổ sung vào hệ $(\mathbf{a}_1)$ để được cơ sở $(\mathbf{a}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ của $V$. Đặt

$$\mathbf{u}_i = \mathbf{b}_i - \frac{1}{a_1}\xi(\mathbf{a}_1, \mathbf{b}_i)\mathbf{a}_1 \qquad \text{với } i = 2, \ldots, n$$

Xét tổ hợp tuyến tính $\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}$. Theo định nghĩa các $\mathbf{u}_i$ ta được

$$\Big(\alpha_1 - \frac{\alpha_2}{a_1}\xi(\mathbf{a}_1, \mathbf{b}_2) - \cdots - \frac{\alpha_n}{a_1}\xi(\mathbf{a}_1, \mathbf{b}_n)\Big)\mathbf{a}_1 + \alpha_2\mathbf{b}_2 + \cdots + \alpha_n\mathbf{b}_n = \mathbf{0}.$$

Do $(\mathbf{a}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ độc lập tuyến tính suy ra $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Vậy hệ $(\mathbf{a}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ độc lập tuyến tính nên nó cũng là cơ sở của $V$. Hơn nữa, ta có

$$\begin{aligned}\xi(\mathbf{a}_1, \mathbf{u}_i) &= \xi\Big(\mathbf{a}_1, \mathbf{b}_i - \frac{1}{a_1}\xi(\mathbf{a}_1, \mathbf{b}_i)\mathbf{a}_1\Big)\\ &= \xi(\mathbf{a}_1, \mathbf{b}_i) - \frac{1}{a_1}\xi(\mathbf{a}_1, \mathbf{b}_i)\xi(\mathbf{a}_1, \mathbf{a}_1)\\ &= 0\end{aligned}$$

với $i = 2, \ldots, n$. Đặt $U = \langle \mathbf{u}_2, \ldots, \mathbf{u}_n \rangle$ ta có $\dim U = n - 1$ và $\mathbf{x} \in U$ nên ta có thể biểu diễn $\mathbf{x} = \lambda_2\mathbf{u}_2 + \cdots + \lambda_n\mathbf{u}_n$. Do đó,

$$\begin{aligned}\xi(\mathbf{a}_1, \mathbf{x}) &= \xi(\mathbf{a}_1, \lambda_2\mathbf{u}_2 + \cdots + \lambda_n\mathbf{u}_n)\\ &= \lambda_2\xi(\mathbf{a}_1, \mathbf{u}_2) + \cdots + \lambda_n\xi(\mathbf{a}_1, \mathbf{u}_n)\\ &= 0.\end{aligned}$$

Đặt $p_1 = p|_U$ là thu hẹp của $p$ trên $U$. Ta có thể thấy $p_1$ là dạng toàn phương trên không gian vector $n-1$ chiều $U$. Theo giả thiết quy nạp trong $U$ tồn tại cơ sở $p_1$-chính tắc $\mathscr{A}_1 = (\mathbf{a}_2, \ldots, \mathbf{a}_n)$ sao cho $\mathbf{x} \in U$ và $\mathbf{x}|_{\mathscr{A}_1} = (x_2, \ldots, x_n)$ thì $p_1(\mathbf{x}) = a_2x_2^2 + \cdots + a_nx_n^2$, trong đó $a_i = p_1(\mathbf{a}_i) = p(\mathbf{a}_i) = \xi(\mathbf{a}_i, \mathbf{a}_i)$ với $i = 2, \ldots, n$.

Ta có thể thấy hệ $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là cơ sở của $V$. Hơn nữa, $\xi(\mathbf{a}_1, \mathbf{a}_i) = 0$ với mọi $i = 2, \ldots, n$ vì $\mathbf{a}_i \in U$. Do đó, ma trận của $\xi$ hay của $p$ trong cơ sở $\mathscr{A}$ là

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

Vậy $\mathscr{A}$ là cơ sở $p$-chính tắc trong $V$. $\square$

**3.1.5 Nhận xét.** Theo chứng minh trên, từ một vector $\mathbf{a}_1 \neq \mathbf{0}$ tùy ý trong $V$ mà $p(\mathbf{a}_1) \neq 0$ ta đều có thể xây dựng được một cơ sở $p$-chính tắc $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ trong $V$. Nói riêng $p$ chấp nhận vô số cơ sở $p$-chính tắc.

## 3.2 Phương pháp Lagrange đưa dạng toàn phương về dạng chính tắc

Ta đã biết mọi dạng toàn phương $p$ trên $\mathbb{K}$-không gian vector $n$ chiều $V$ đều có cơ sở $p$-chính tắc trong $V$. Việc tìm một cơ sở $p$-chính tắc và từ đó ta có biểu thức tọa độ của $p$ có dạng chính tắc gọi là *phép đưa dạng toàn phương về dạng chính tắc*.

**3.2.1** Giả sử đã cho dạng toàn phương $p$ trên $\mathbb{K}$-không gian vector $n$ chiều $V$ và $A = (a_{ij})_n \in \mathcal{M}_n(\mathbb{K})$ là ma trận của $p$ trong một cơ sở

$\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ đã cho của $V$. Nếu $\mathscr{B}$ chưa phải là cơ sở chính tắc đối với $p$ thì $A$ không là ma trận chéo.

Việc đưa $p$ về dạng chính tắc thực chất là việc chọn một cơ sở $p$-chính tắc $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n)$ của $V$ để biểu thức tọa độ của $p$ trong $\mathscr{B}'$ có dạng chính tắc

$$p(\mathbf{x}) = a_1 x_1'^2 + a_2 x_2'^2 + \cdots + a_n x_n'^2$$

với mọi $\mathbf{x} \in V$ mà $\mathbf{x}|_{\mathscr{B}'} = (x'_1, x'_2, \ldots, x'_n)$; $a_i = p(\mathbf{e}'_i)$ với $i = 1, \ldots, n$. Hơn nữa, bằng cách đánh số lại $\mathscr{B}'$ (nếu cần) ta luôn làm cho dạng chính tắc của biểu thức tọa độ của $p$ có dạng

(3.2.2) $$p(\mathbf{x}) = a_1 x_1'^2 + a_2 x_2'^2 + \cdots + a_r x_r'^2,$$

với $r = \operatorname{rank}(p)$, $a_1 a_2 \cdots a_r \neq 0$.

Về phương diện ma trận, việc đưa $p$ về dạng chính tắc (3.2.2) tương đương với việc tìm ma trận khả nghịch $C \in \mathrm{GL}_n(\mathbb{K})$ ($C$ chính là ma trận đổi cơ sở từ $\mathscr{B}$ sang $\mathscr{B}'$) sao cho $A' = C^t A C$ có dạng chéo như sau:

(3.2.3) $$A' = \begin{bmatrix} a_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & 0 & 0 & \ldots & 0 \\ \multicolumn{7}{c}{\dotfill} \\ 0 & 0 & \ldots & a_r & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \multicolumn{7}{c}{\dotfill} \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{bmatrix}$$

($A'$ chính là ma trận của $p$ trong cơ sở $\mathscr{B}'$) với $a_1 a_2 \cdots a_r \neq 0$.

Trong thực hành, để đưa $p$ về dạng chính tắc, ta thường không bắt đầu từ việc chọn cơ sở $p$-chính tắc $\mathscr{B}'$ mà trước hết biến đổi trực tiếp biểu thức tọa độ của $q$ để đưa nó về dạng chính tắc (3.2.2), hoặc biến đổi sơ cấp ma trận $A$ để tìm ra ma trận $C$ làm chéo hóa $A$ thành $A'$ như trong (3.2.3); sau đó mới chỉ ra cơ sở $p$-chính tắc $\mathscr{B}'$.

**3.2.4 Thuật toán Lagrange.** Cho dạng toàn phương $p$ trên $\mathbb{K}$-không gian vector $n$ chiều $V$ với biểu thức tọa độ trong cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là

$$p(\mathbf{x}) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

trong đó $\mathbf{x} \in V$ và $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \dots, x_n)$.

- **Bước 1**: Biến đổi để đưa $p(\mathbf{x})$ về dạng

$$p(\mathbf{x}) = a_1 x_1'^2 + p_1(x_2, \dots, x_n)$$

  trong đó $a_1 \in \mathbb{K}$ ($a_1 \neq 0$ nếu $x_1$ không vắng mặt trong biểu thức tọa độ của $p$), $p_1$ chỉ chứa $x_2, \dots, x_n$.

- **Bước 2**: Biến đổi để đưa $p_1$ về dạng

$$p_1(x_2, \dots, x_n) = a_2 x_2'^2 + p_2(x_3, \dots, x_n)$$

  với $a_2 \in \mathbb{K}$ ($a_2 \neq 0$ nếu $x_2$ không vắng mặt trong biểu thức của $p_1$), $p_2$ là biểu thức chỉ chứa $x_3, \dots, x_r$.

- Cứ tiếp tụ như thế, nhiều nhất sau $n-1$ bước, $p$ sẽ là tổng các bình phương, tức là $p$ có dạng chính tắc.

Vì các bước hoàn toàn tương tự nhau nên sau đây ta chỉ cần trình bày rõ **bước 1** đưa $p(\mathbf{x})$ về dạng $a_1 x_1'^2 + p_2(x_2, \dots, x_n)$. Có hai trường hợp:

*Trường hợp 1*: $a_{11}, a_{22}, \dots, a_{nn}$ không đồng thời triệt tiêu (tức là $p$ chứa ít nhất một bình phương của một biến nào đó). Nếu cần có thể đánh số lại $\mathscr{B}$, nên ngay từ đầu có thể giả thiết rằng $a_{11} \neq 0$. Khi đó $p(\mathbf{x})$ được viết lại như sau

$$\begin{aligned}
p(\mathbf{x}) &= a_{11}x_1^2 + 2\sum_{j=2}^{n} a_{1j}x_1x_j + \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{2\le i<j\le n} a_{ij}x_ix_j \\
&= a_{11}\Big(x_1^2 + 2x_1\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j + \Big(\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\Big)^2\Big) \\
&\qquad - a_{11}\Big(\sum_{j=2}^{n}\frac{a_{1j}}{a_{11}}x_j\Big)^2 + \sum_{j=2}^{n} a_{ii}x_i^2 + 2\sum_{2\le i<j\le n} a_{ij}x_ix_j \\
&= a_1 x_1'^2 + p_1(x_2, \dots, x_n),
\end{aligned}$$

với $a_1 = a_{11} \neq 0$, $x'_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n$ và

$$p_1(x_2, \ldots, x_n) = -a_{11}\Big(\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}}x_j\Big)^2 + \sum_{j=2}^{n} a_{ii}x_i^2 + 2\sum_{2\leq i<j\leq n} a_{ij}x_ix_j$$

Rõ ràng $p_1$ chỉ chứa $x_2, \ldots, x_n$, hơn nữa $p_1$ là một đa thức đẳng cấp bậc hai đối với $x_2, \ldots, x_n$ (nếu $q_1 \neq 0$).

Quá trình vừa thực hiện trên tương ứng với phép đổi tọa độ (hay đổi biến) như sau

$$\begin{cases} x'_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n \\ x'_2 = x_2 \\ \ldots\ldots \\ x'_n = x_n \end{cases}$$

hay

$$\begin{cases} x_1 = x'_1 - \frac{a_{12}}{a_{11}}x'_2 - \cdots - \frac{a_{1n}}{a_{11}}x'_n \\ x_2 = x'_2 \\ \ldots\ldots \\ x_2 = x'_n \end{cases}$$

Tức là thay cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ bởi cơ sở $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ mà ma trận đổi cơ sở từ $\mathscr{B}$ sang $\mathscr{A}$ là

$$B = \begin{bmatrix} 1 & -\frac{a_{12}}{a_{11}} & \ldots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix} \in \mathrm{GL}_n(\mathbb{K})$$

Trong cơ sở $\mathscr{A}$, biểu thức tọa độ của của $p$ chính là $p(\mathbf{x}) = a_1x_1'^2 + p_1(x'_2, \ldots, x'_n)$.

*Trường hợp 2*: $a_{11} = a_{22} = \cdots = a_{nn} = 0$ (tức là $p$ không chứa bình phương của một biến bất kỳ nào). Nếu $a_{1j} = 0$ với mọi $j = 2, \ldots, n$ thì

$$p(\mathbf{x}) = 2\sum_{2\leq i<j\leq n} a_{ij}x_ix_j,$$

và chỉ còn chứa $x_2, \ldots, x_n$. Lúc đó chỉ cần đặt $p(\mathbf{x}) = p_1(x_2, \ldots, x_n) = 0x_1'^2 + p_1(x_2, \ldots, x_n)$.

Giả sử $a_{12}, \ldots, a_{1n}$ không đồng thời triệt tiêu. Ta có thể cho rằng $a_{12} \neq 0$ (điều này luôn đạt được sau phép đánh số lại cơ sở $\mathscr{B}$ một cách thích hợp). Dùng phép đổi tọa độ sau

$$\begin{cases} x_1 = x_1' + x_2' \\ x_2 = x_1' - x_2' \\ x_3 = x_3' \\ \ldots\ldots\ldots\ldots \\ x_n = x_n' \end{cases}$$

tức là thay cơ sở $\mathscr{B}$ bởi cơ sở $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ với ma trận đổi cơ sở từ $\mathscr{B}$ sang $\mathscr{A}$ là

$$B = \begin{bmatrix} 1 & 1 & 0 & \ldots & 0 \\ 1 & -1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

để ý rằng $\det(B) = -2$. Khi đó

$$\begin{aligned} p(\mathbf{x}) &= 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + 2\sum_{2\leq i<j\leq n} a_{ij}x_ix_j \\ &= 2a_{12}(x_1' + x_2')(x_1' - x_2') + 2\sum_{j=3}^{n} a_{ij}(x_1' + x_2')x_j' \\ &\quad + 2\sum_{j=3}^{n} a_{2j}(x_1' - x_2')x_j' + 2\sum_{3\leq i<j\leq n} a_{ij}x_i'x_j' \\ &= a_{11}'x_1'^2 + p'(x_1', \ldots, x_n') \end{aligned}$$

với $a_{11}' = 2a_{12} \neq 0$ và $p'(x_1', \ldots, x_n') = p(x_1, \ldots, x_n) - a_{11}'x_1'^2$ không chứa $x_1'^2$. Nói cách khác biểu thức tọa độ của $p$ trong $\mathscr{A}$ lại có dạng thuộc **trường hợp 1** đã xét (vì nó chứa ít nhất một bình phương, $a_{11}'x_1'^2$, của biến $x'_1$). Thực hiện tiếp như trường hợp 1 ta đưa được $p(\mathbf{x})$ về dạng cần thiết $p(\mathbf{x}) = a_1 x''^2_1 + p_1$.

**3.2.5 Thí dụ.** Xét dạng toàn phương trong $\mathbb{R}^3$ có biểu thức tọa độ trong cơ sở chính tắc: $p(x_1, x_2, x_3) = 2x_1^2 - 2x_2^2 + 3x_3^2 - 2x_1x_2 + 4x_1x_3 - x_2x_3$.

Theo thuật toán Lagrange ta biến đổi như sau

$$\begin{aligned}
p(\mathbf{x}) &= 2x_1^2 - 2x_1x_2 + 4x_1x_3 - 2x_2^2 + 3x_3^2 - x_2x_3 \\
&= 2(x_1 - \tfrac{1}{2}x_2 + x_3)^2 - \frac{1}{2}x_2^2 - 2x_3^2 + 2x_2x_3 - 2x_2^2 + 3x_3^2 - x_2x_3 \\
&= 2x'^2_1 - \frac{5}{2}(x_2^2 - \tfrac{2}{5}x_2x_3) + x_3^2 \\
&= 2x'^2_1 - \frac{5}{2}(x_2 - \tfrac{1}{5}x_3)^2 + \frac{1}{10}x_3^2 + x_3^2 \\
&= 2x'^2_1 - \frac{5}{2}x'^2_2 + \frac{11}{10}x'^2_3
\end{aligned}$$

và ta có công thức chuyển tọa độ từ cơ sở $p$-chính tắc sang cơ sơ chính tắc

$$\begin{cases} x'_1 = x_1 - \frac{1}{2}x_2 + x_3 \\ x'_2 = x_2 - \frac{1}{5}x_3 \\ x'_3 = x_3 \end{cases} \qquad \text{hay} \qquad \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} & 1 \\ 0 & 1 & -\frac{1}{5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Do đó, ma trận chuyển tọa độ từ cơ sở chính tắc sang cơ sở $p$-chính tắc là

$$A = \begin{bmatrix} 1 & -\frac{1}{2} & 1 \\ 0 & 1 & -\frac{1}{5} \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \frac{1}{2} & -\frac{9}{10} \\ 0 & 1 & \frac{1}{5} \\ 0 & 0 & 1 \end{bmatrix}$$

Vậy cơ sở $p$-chính tắc là $\mathscr{A} = ((1, 0, 0), (\frac{1}{2}, 1, 0), (-\frac{9}{10}, \frac{1}{5}, 1))$. □

**3.2.6 Thí dụ.** Xét dạng toàn phương $p(x, y, z) = 2xy + 4xz - 2yz$. Đưa dạng toàn phương $p$ về dạng chính tắc và tìm cơ sở $p$-chính tắc tương ứng.

Do $p$ không chứa một bình phương nào nên đặt $x = x' + y'$ và $y =$

$x' - y'$. Khi đó, ta có

$$\begin{aligned} p(\mathbf{x}) = p(x,y,z) &= 2(x'^2 - y'^2) + 4(x'+y')z - 2(x'-y')z \\ &= 2(x'^2 + x'z) - 2y'^2 + 6y'z \\ &= 2(x' + \tfrac{1}{2}z)^2 - 2(y'^2 - 3y'z) - \frac{1}{2}z^2 \\ &= 2x''^2 - 2(y' - \tfrac{3}{2}z)^2 + 4z^2 \\ &= 2x''^2 - 2y''^2 + 4z''^2 \end{aligned}$$

và ta có công thức chuyển tọa độ từ cơ sở $p$-chính tắc sang cơ sơ chính tắc

$$\begin{cases} x'' = x' + \frac{1}{2}z = \frac{1}{2}x + \frac{1}{2}y + \frac{1}{2}z \\ y'' = y' - \frac{3}{2}z = \frac{1}{2}x - \frac{1}{2}y - \frac{3}{2}z \\ z'' = z \end{cases} \quad \text{hay} \quad \begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Do đó, ma trận chuyển tọa độ từ cơ sở chính tắc sang cơ sở $p$-chính tắc là

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

Vậy cơ sở $p$-chính tắc là $\mathscr{A} = ((1,1,0),(1,-1,0),(1,-2,1))$. Hơn nữa, $p$ là dạng toàn phương có hạng là 3 nên nó không suy biến. $\square$

## 3.3 Phương pháp đưa dạng toàn phương về dạng chính tắc nhờ các phép biến đổi sơ cấp trên ma trận của nó

Cho dạng toàn phương $p$ trên $\mathbb{K}$-không gian vector $n$ chiều $V$ với ma trận trong cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ đã cho là $A = (a_{ij})_n \in \mathcal{M}_n(\mathbb{K})$. Hiển nhiên $A$ là ma trận đối xứng. Việc đưa $p$ về dạng chính tắc chính là việc tìm ma trận không suy biến $C$ sao cho $C^tAC$ là ma trận chéo.

**3.3.1** Vì mỗi ma trận khả nghịch đều là tích của các ma trận sơ cấp và vì $A$ là ma trận đối xứng, nên ta có thể thu được dạng chéo $C^tAC$ của $A$ nhờ một dãy các phép biến đổi sơ cấp dòng và dãy tương tự các phép biến đổi sơ cấp cột trên $A$. Khi đó, chính dãy các biến đổi sơ cấp dòng (tương ứng cột) sẽ biến ma trận đơn vị $I_n$ thành $C^t$ (tương ứng $C$). Từ đó ta có thuật toán sau.

**3.3.2 Nội dung thuật toán.** Lập ma trận chia khối $\begin{bmatrix} A \mid I_n \end{bmatrix}$ bằng cách ghép thêm ma trận đơn vị $I_n$ vào bên phải $A$. Thực hiện các phép biến đổi sơ cấp trên các dòng, đồng thời lập lại các biến đổi cùng kiểu trên các cột của $\begin{bmatrix} A \mid I_n \end{bmatrix}$ để đưa ma trận $A$ về dạng chéo. Khi đó $I_n$ sẽ biến thành $C^t$.

**3.3.3 Thí dụ.** Cho dạng toàn phương 3 biến thực

$p(x, y, z) = x^2 + 4xy + 6xz + 5y^2 + 8yz + 8z^2$, với mọi $(x, y, z) \in \mathbb{R}^3$.

Hãy đưa $p$ về dạng chính tắc, xác định hạng và tính suy biến của $p$.

*Giải*. Theo biểu thức định nghĩa, ma trận của $p$ trong cơ sở chính tắc của $\mathbb{R}^3$ là

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 8 \end{bmatrix}.$$

Lập ma trận $\begin{bmatrix} A \mid I_3 \end{bmatrix}$ rồi biến đổi sơ cấp để chéo hóa như sau

$$\begin{bmatrix} A \mid I_3 \end{bmatrix} = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 4 & 0 & 1 & 0 \\ 3 & 4 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[d_3 \to d_3 - 3d_1]{d_2 \to d_2 - 2d_1} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & -2 & -1 & -3 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[c_3 \to c_3 - 3c_1]{c_2 \to c_2 - 2c_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & -2 & -1 & -3 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{d_3 \to d_3 + 2d_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & -5 & -7 & 2 & 1 \end{array}\right]$$

$$\xrightarrow{c_3 \to c_3 + 2c_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -5 & -7 & 2 & 1 \end{array}\right]$$

Khối bên trái của ma trận sau cùng đã có dạng chéo. Đặt

$$C^t = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -7 & 2 & 1 \end{bmatrix}$$

Khi đó $C = \begin{bmatrix} 1 & -2 & -7 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ và $C^tAC = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}$. Thay cơ sở chính tắc $\mathscr{E}_3$ bởi cơ sở $\mathscr{B}$ sao cho $C$ là ma trận đổi cơ sở từ $\mathscr{E}_3$ sang $\mathscr{B}$, tức là dùng phép đổi biến

$$\begin{cases} x = x' - 2y' - 7z' \\ y = y' + 2z' \\ z = z' \end{cases}$$

Lúc đó $\operatorname{rank}(p) = 3$ và $p$ không suy biến. □

**3.3.4 Nhận xét.** Qua thí dụ trên ta nhận thấy sau mỗi bước thực hiện phép biến đổi sơ cấp dòng cùng với phép biến đổi cột tương ứng, phần tử được đánh dấu giữ nguyên giá trị, trong khi đó các phần tử còn lại trong cột và dòng của khối bên trái của phần tử đánh dấu bằng không. Do đó ta chỉ cần thực hiện phép biến đổi sơ cấp theo dòng và ghi các kết quả ở các dòng biến đổi còn dòng của phần tử đánh dấu của khối bên trái ta có các phần tử bằng không trừ phần tử đánh dấu.

**3.3.5 Thí dụ.** Hãy đưa dạng toàn phương 3 biến thực sau đây về dạng chính tắc

$$p(x, y, z) = 2xy - 4xz + 6yz \qquad \text{với mọi } (x, y, z) \in \mathbb{R}^3.$$

Xác định hạng, tính suy biến của $p$.

*Giải.* Ma trận của $p$ trong cơ sở chính tắc của $\mathbb{R}^3$ là

$$A = \begin{bmatrix} 0 & 1 & -2 \\ 1 & 0 & 3 \\ -2 & 3 & 0 \end{bmatrix}.$$

Lập ma trận $\left[A \mid I_3\right]$ và biến đổi sơ cấp để chéo hóa $A$ như sau:

$$\left[A \mid I_3\right] = \left[\begin{array}{ccc|ccc} 0 & 1 & -2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[c_1 \to c_1 + c_2]{d_1 \to d_1 + d_2} \left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[\substack{c_2 \to c_2 - \frac{1}{2}c_1 \\ c_3 \to c_3 - \frac{1}{2}c_1}]{\substack{d_2 \to d_2 - \frac{1}{2}d_1 \\ d_3 \to d_3 - \frac{1}{2}d_1}} \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & \frac{5}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{5}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array}\right]$$

$$\xrightarrow[c_3 \to c_3 + 5c_2]{d_3 \to d_3 + 5d_2} \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 12 & -3 & 2 & 1 \end{array}\right].$$

Đặt $C^t = \begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -3 & 2 & 1 \end{bmatrix}$. Khi đó

$$C = \begin{bmatrix} 1 & -\frac{1}{2} & -3 \\ 1 & \frac{1}{2} & 2 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{và} \qquad C^t A C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 12 \end{bmatrix}.$$

Do đó, nếu dùng phép đổi biến

$$\begin{cases} x = x' - \frac{1}{2}y' - 3z' \\ y = x' + \frac{1}{2}y' + 2z' \\ z = z' \end{cases}$$

thì $p$ nhận dạng chính tắc $p(x, y, z) = 2x'^2 - \frac{1}{2}y'^2 + 12z'^2$. Hơn nữa, ta có $\operatorname{rank}(p) = 3$ và $p$ không suy biến. □

### 3.4 Phương pháp Jacobi đưa dạng toàn phương về dạng chính tắc

**3.4.1 Định lý. (Thuật toán Jacobi)** *Cho dạng toàn phương $p$ trên không gian vector $n$ chiều $V$ với ma trận là $A = (a_{ij})_n$ trong một cơ sở*

*$\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ nào đó của $V$. Giả sử các định thức con chính $D_1, D_2, \ldots, D_{n-1}$ của $A$ khác không, trong đó*

$$D_k = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix}.$$

*Khi đó, tồn tại duy nhất cơ sở p-chính tắc $\mathscr{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ của $V$ sao cho*

$$\begin{aligned} &\mathbf{u}_1 = \mathbf{b}_1 \\ &\mathbf{u}_2 = \alpha_{21}\mathbf{b}_1 + \mathbf{b}_2 \\ &\mathbf{u}_3 = \alpha_{31}\mathbf{b}_1 + \alpha_{32}\mathbf{b}_2 + \mathbf{b}_3 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\mathbf{u}_n = \alpha_{n1}\mathbf{b}_1 + \alpha_{n2}\mathbf{b}_2 + \cdots + \alpha_{n,n-1}\mathbf{b}_{n-1} + \mathbf{b}_n \end{aligned} \tag{3.4.2}$$

*Hơn nữa, với mọi $\mathbf{y} \in V$ mà $\mathbf{y}|_{\mathscr{U}} = (y_1, y_2, \ldots, y_n)$, biểu thức tọa độ của $p$ trong $\mathscr{U}$ là*

$$p(\mathbf{y}) = D_1 y_1^2 + \frac{D_2}{D_1} y_2^2 + \cdots + \frac{D_n}{D_{n-1}} y_n^2. \tag{3.4.3}$$

*Chứng minh. Sự tồn tại:* Với mỗi $i \in \{2, \ldots, n\}$ xét hệ phương trình tuyến tính

$$\begin{cases} a_{11}t_1 + a_{12}t_2 + \cdots + a_{1,i-1}t_{i-1} = -a_{1i} \\ a_{21}t_1 + a_{22}t_2 + \cdots + a_{2,i-1}t_{i-1} = -a_{2i} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{i-1,1}t_1 + a_{i-1,2}t_2 + \cdots + a_{i-1,i-1}t_{i-1} = -a_{i-1,i} \end{cases} \tag{3.4.4}$$

gồm $i-1$ phương trình $i-1$ ẩn $t_1, t_2, \ldots, t_{i-1}$. Đây là hệ Cramer vì ma trận hệ số của hệ có định thức là $D_{i-1} \neq 0$. Do đó, hệ (3.4.4) có nghiệm duy nhất, và gọi nghiệm đó là

$$(t_1, t_2, \ldots, t_{i-1}) = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{i,i-1}).$$

Đặt $\mathbf{u}_i = \alpha_{i1}\mathbf{b}_1 + \alpha_{i2}\mathbf{b}_2 + \cdots + \alpha_{i,i-1}\mathbf{b}_{i-1} + \mathbf{b}_i$ với $i = 2, \ldots, n$ và $\mathbf{u}_1 = \mathbf{b}_1$. Khi đó, hiển nhiên ta nhận được cơ sở $\mathscr{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ thỏa mãn (3.4.2). Tính toán trực tiếp ta có (chú ý đến nghiệm của hệ (3.4.4))

$$\xi(\mathbf{u}_1, \mathbf{u}_i) = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix} A \begin{bmatrix} \alpha_{i,1} \\ \vdots \\ \alpha_{i,i-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0$$

với mọi $2 \leq i \leq n$ trong đó $\xi$ là dạng song tuyến tính cực của $p$; và với mọi $1 < j < i \leq n$ ta có

$$\xi(\mathbf{u}_j, \mathbf{u}_i) = \begin{bmatrix} \alpha_{j1} & \ldots & \alpha_{j,j-1} & 1 & 0 & \ldots & 0 \end{bmatrix} A \begin{bmatrix} \alpha_{i1} \\ \vdots \\ \alpha_{i,i-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0$$

Vậy ma trận của $\xi$ hay của $p$ trong cơ sở $\mathscr{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ có dạng chéo, tức là $p$ có dạng chính tắc trong $\mathscr{U}$ hay $\mathscr{U}$ là một cơ sở $p$-chính tắc.

*Sự duy nhất:* Giả sử còn có một cơ sở $p$-chính tắc $\mathscr{V} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ sao cho

$$\begin{aligned}
&\mathbf{v}_1 = \mathbf{b}_1 \\
&\mathbf{v}_2 = \beta_{21}\mathbf{b}_1 + \mathbf{b}_2 \\
&\mathbf{v}_3 = \beta_{31}\mathbf{b}_1 + \beta_{32}\mathbf{b}_2 + \mathbf{b}_3 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\mathbf{v}_n = \beta_{n1}\mathbf{b}_1 + \beta_{n2}\mathbf{b}_2 + \cdots + \beta_{n,n-1}\mathbf{b}_{n-1} + \mathbf{b}_n
\end{aligned}$$

Khi đó, $\xi(\mathbf{v}_j, \mathbf{v}_i) = 0$ với mọi $1 \leq j < i \leq n$. Từ đó suy ra

$$\xi(\mathbf{b}_1, \mathbf{v}_i) = \xi(\mathbf{b}_2, \mathbf{v}_i) = \cdots = \xi(\mathbf{b}_{i-1}, \mathbf{v}_i) = 0, \quad \text{với mọi } i = 2, \ldots, n.$$

Mặt khác, ta có $\xi(\mathbf{b}_k, \mathbf{v}_i) = \mathbf{b}_k|_{\mathscr{B}} A\mathbf{v}_i|_{\mathscr{B}} = a_{k1}\beta_{i1} + \cdots + a_{k,i-1}\beta_{i,i-1} + a_{ki}$. Do đó, ta được $(\beta_{i1}, \beta_{i2}, \ldots, b_{i,i-1})$ là nghiệm của hệ (3.4.4). Suy ra

$$\beta_{i1} = \alpha_{i1}, \quad \beta_{i2} = \alpha_{i2}, \quad \ldots, \quad \beta_{i,i-1} = \alpha_{i,i-1} \quad \text{với } i = 2, \ldots, n$$

hay $\alpha_{ij} = \beta_{ij}$ với mọi $1 \leq j < i \leq n$; tức là $\mathbf{v}_i = \mathbf{u}_i$ với mọi $i = 1, \ldots, n$ hay $\mathscr{U} = \mathscr{V}$, nói cách khác $\mathscr{U}$ là duy nhất.

Giải hệ (3.4.4) theo công thức Cramer ta được

$$\alpha_{ij} = (-1)^{i+j} \frac{D_{i-1,j}}{D_{i-1}},$$

ở đó $(-1)^{i+j} D_{i-1,j}$ là giá trị định thức nhận được từ $D_{i-1}$ bằng cách thay cột thứ $j$ bởi cột tự do của hệ (3.4.4). Hơn nữa, $D_{i-1,j}$ là định thức con cấp $i-1$ của $A$ mà các phần tử của nó nằm trên giao của các dòng $1, 2, \ldots, i-1$ và các cột thứ $1, \ldots, j-1, j+1, \ldots, i$ với $1 \leq j < i \leq n$.

Ta có $\xi(\mathbf{u}_1, \mathbf{u}_1) = \xi(\mathbf{b}_1, \mathbf{b}_1) = a_{11} = D_1$. Với $2 \leq i \leq n$, ta tính được

$$\begin{aligned}
\xi(\mathbf{u}_i, \mathbf{u}_i) &= \begin{bmatrix} \alpha_{i1} & \ldots & \alpha_{i,i-1} & 1 & 0 & \ldots & 0 \end{bmatrix} A \begin{bmatrix} \alpha_{i1} \\ \vdots \\ \alpha_{i,i-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
&= a_{i1}\alpha_{i1} + \cdots + a_{i,i-1}\alpha_{i,i-1} + a_{ii} \\
&= \frac{1}{D_{i-1}} \big( a_{i1}(-1)^{i+1} D_{i-1,1} + \cdots \\
&\qquad\qquad + a_{i,i-1}(-1)^{i+i-1} D_{i-1,i-1} + a_{ii} D_{i-1} \big) \\
&= \frac{D_i}{D_{i-1}}.
\end{aligned}$$

Như vậy biểu thức tọa độ của $p$ trong cơ sở $\mathscr{U}$ chính là (3.4.3). $\square$

**3.4.5 Thí dụ.** Hãy đưa dạng toàn phương trên không gian vector $\mathbb{R}^3$ sau đây về dạng chính tắc bằng thuật toán Jacobi.

$$p(x_1, x_2, x_3) = 2x_1^2 + 3x_1x_2 + 4x_1x_3 + x_2^2 + x_3^2$$

Ta có ma trận đối xứng của $p$ trong cơ sở chính tắc là

$$A = \begin{bmatrix} 2 & \frac{3}{2} & 2 \\ \frac{3}{2} & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

Các định thức con chính của $A$ là

$$D_1 = 2 \qquad D_2 = \begin{vmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{vmatrix} = -\frac{1}{4} \qquad D_3 = \begin{vmatrix} 2 & \frac{3}{2} & 2 \\ \frac{3}{2} & 1 & 0 \\ 2 & 0 & 1 \end{vmatrix} = -\frac{17}{4}.$$

Vậy ta có thể dùng định lý Jacobi để tìm dạng chính tắc của $p$. Ta tính các hệ số $\alpha_{ij}$ với $1 \leq j < i \leq 3$. Ta có

$$\alpha_{21} = (-1)^{2+1}\frac{D_{11}}{D_1} = -\frac{\frac{3}{2}}{2} = -\frac{3}{4}$$

$$\alpha_{31} = (-1)^{3+1}\frac{D_{21}}{D_2} = \frac{\begin{vmatrix} \frac{3}{2} & 2 \\ 1 & 0 \end{vmatrix}}{-\frac{1}{4}} = 8$$

$$\alpha_{32} = (-1)^{3+2}\frac{D_{22}}{D_2} = -\frac{\begin{vmatrix} 2 & 2 \\ \frac{3}{2} & 0 \end{vmatrix}}{-\frac{1}{4}} = -12.$$

Vậy $p$ có dạng chính tắc

$$p(\mathbf{x}) = 2y_1^2 - \frac{1}{8}y_2^2 + 17y_3^2$$

trong đó $(y_1, y_2, y_3)$ là tọa độ của vector $\mathbf{x} \in \mathbb{R}^3$ trong cơ sở $p$-chính tắc $\mathscr{B} = ((1,0,0), (-\frac{3}{4}, 1, 0), (8, -12, 1))$. Hơn nữa, ta thấy $p$ là một dạng toàn phương không suy biến. □

## Bài tập

**1)** Đưa các dạng toàn phương sau trên không gian vector $\mathbb{R}^3$ về dạng chính tắc và tìm cơ sở chính tắc đối với dạng toàn phương ấy.

**(a)** $x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$.

**(b)** $x_1^2 - 2x_2^2 + x_3^2 + 2x_1x_2 + 4x_1x_3 + 2x_2x_3$.

**(c)** $2x_1^2 + 3x_2^2 + 4x_3^2 - 2x_1x_2 + 4x_1x_3 - 3x_2x_3$.

**(d)** $3x_1^2 - 2x_2^2 + 2x_3^2 + 4x_1x_2 - 3x_1x_3 - x_2x_3$.

**(e)** $x_1^2 + 2x_2^2 - 7x_3^2 - 4x_1x_2 + 8x_1x_3$.

**2)** Đưa các dạng toàn phương sau trên không gian vector $\mathbb{R}^3$ về dạng chính tắc và tìm cơ sở chính tắc đối với dạng toàn phương ấy.

**(a)** $x^2 + y^2 + 2(z^2 + t^2) + xz + xt + zt$

**(b)** $4x^2 + 3y^2 + 9z^2 + 8xz + 4xy + 4yt + 8zt + \lambda t^2$.

**3)** Đưa dạng toàn phương sau trên không gian vector $\mathbb{R}^n$ về dạng chính tắc và tìm cơ sở chính tắc tương ứng. $p(\mathbf{x}) = \sum\limits_{i=1}^{n} x_i^2 + \sum\limits_{1\leq i<j\leq n} x_ix_j$ với mọi $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

**4)** Trên $\mathbb{R}^3$ với cơ sở chính tắc cho dạng toàn phương $p(x, y, z) = 3x^2 + 4xy + 8xz - y^2 - 6yz + z^2$.

**(a)** Viết ma trận của $p$ trong cơ sở chính tắc, xác định dạng song tuyến tính cực $p$.

**(b)** Viết ma trận của $p$ trong cơ sở $((1, 1, 0), (0, 1, 1), (1, 0, 1))$.

**(c)** Đưa $p$ về dạng chính tắc, xác định hạng và tính chất của $p$.

**5)** Hãy đưa dạng toàn phương 4 biến thực sau về dạng chính tắc

$q(x, y, z, t) = 2x^2 + 8xy + 4xz - 4xt + 9y^2 + 6yz - 6yt + z^2 - 10zt + t^2$.

Từ đó dãy tìm hạng và xác định tính suy biến hay không suy biến của $p$.

**6**) Trên $\mathbb{R}^3$ với cơ sở chính tắc cho các dạng toàn phương sau

**(a)** $p(x, y, z) = 3x^2 + 4xy + 8xz - y^2 - 6yz + z^2$

**(b)** $p(x, y, z) = x^2 + xz - 2yz$

**(c)** $p(x, y, z) = 4x^2 - 6xy - 7y^2$

**(d)** $p(x, y, z) = xy + y^2$.

Đối với mỗi $p$ hãy

**(i)** Viết ma trận của $p$ trong cơ sở chính tắc, xác định dạng song tuyến tính cực của $p$

**(ii)** Viết ma trận của $p$ trong cơ sở $((1, 1, 0), (0, 1, 1), (1, 0, 1))$.

**(iii)** Đưa $p$ về dạng chính tắc, xác định hạng và các tính chất của $p$.

**7**) Đối với mỗi ma trận đối xứng $A$ sau đây hãy tìm ma trận khả nghịch $C$ để $C^t AC$ là ma trận chéo.

**(a)** $\begin{bmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & 1 & 1 \\ 1 & -2 & 2 \\ 1 & 2 & -1 \end{bmatrix}$

**(c)** $\begin{bmatrix} 1 & 2 & 4 & 1 \\ 2 & 2 & 0 & 2 \\ 4 & 0 & 1 & 2 \\ 1 & 2 & 2 & 1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix}$

## § 4 Dạng chuẩn tắc và định lý quán tính của dạng toàn phương thực

Trong mục này, ta chỉ xét các dạng toàn phương thực, tức là các dạng toàn phương trên các không gian vector thực ($\mathbb{K} = \mathbb{R}$).

## 4.1 Dạng chuẩn tắc của dạng toàn phương thực

**4.1.1 Định nghĩa.** Cho $p$ là một dạng toàn phương trên không gian vector thực $n$ chiều $V$. Giả sử $\operatorname{rank} p = r$ và $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ là một cơ sở $p$-chính tắc trong $V$ sao cho biểu thức tọa độ dạng của $p$ trong $\mathscr{B}$ là

$$p(\mathbf{x}) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2,$$

với mọi $\mathbf{x} \in V$ mà $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \dots, x_n)$ và $s + t = r$. Khi đó, $p$ gọi là có **biểu thức tọa độ dạng chuẩn tắc** (hay nói gọn hơn **dạng chuẩn tắc**). Cơ sở $\mathscr{B}$ được gọi là **cơ sở chuẩn tắc đối với** $p$ hay cơ sở $p$**-chuẩn tắc**.

**4.1.2 Định lý.** *Mọi dạng toàn phương thực $p$ trong không gian vector $n$ chiều $V$ luôn tồn tại cơ sở $p$-chuẩn tắc trong $V$.*

*Chứng minh.* Nếu $p = 0$, thì mọi cơ sở trong $V$ đều là cơ sở $p$-chuẩn tắc. Giả sử $p \neq 0$ và $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ là một cơ sở $p$-chính tắc và $p$ có dạng chính tắc trong $\mathscr{A}$ là

$$p(\mathbf{x}) = a_1x_1^2 + a_2x_2^2 + \cdots + a_rx_r^2, \qquad a_1a_2\cdots a_r \neq 0,$$

ở đó $0 < r = \operatorname{rank} p \leq n$, với mọi $\mathbf{x} \in V$ mà $\mathbf{x}|_{\mathscr{A}} = (x_1, x_2, \dots, x_n)$.

Sau khi đánh số lại các vector trong $\mathscr{A}$ một cách thích hợp (nếu cần), ta luôn có thể giả thiết rằng $a_1, \dots, a_s > 0$ và $a_{s+1}, \dots, a_{s+t} < 0$ với $0 \leq s, t \leq r$. Xét cơ sở $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ xác định như sau

$$\mathbf{b}_j = \begin{cases} \frac{1}{\sqrt{a_j}}\mathbf{a}_j & \text{khi } 1 \leq j \leq s \\ \frac{1}{\sqrt{-a_j}}\mathbf{a}_j & \text{khi } s+1 \leq j \leq s+t \\ \mathbf{a}_j & \text{khi } j \geq r = s+t. \end{cases}$$

Với mọi $\mathbf{x} \in V$ mà $\mathbf{x}|_{\mathscr{A}} = (x_1, x_2, \dots, x_n)$ và $\mathbf{x}|_{\mathscr{B}} = (x'_1, x'_2, \dots, x'_n)$, ta có công thức đổi tọa độ

$$x_i = \begin{cases} \frac{1}{\sqrt{a_i}}x'_i & \text{khi } 1 \leq i \leq s \\ \frac{1}{\sqrt{-a_i}}x'_i & \text{khi } s+1 \leq i \leq s+t \\ x'_i & \text{khi } i \geq s+t. \end{cases}$$

Khi đó, trong cơ sở $\mathscr{B}$, $p$ có biểu thức tọa độ dạng chuẩn tắc

$$p(\mathbf{x}) = x_1'^2 + \cdots + x_s'^2 - x_{s+1}'^2 - \cdots - x_{s+t}'^2$$

với mọi $\mathbf{x} \in V$ mà $\mathbf{x}|_{\mathscr{B}} = (x_1', x_2', \ldots, x_n')$; nói cách khác $\mathscr{B}$ chính là cơ sở $p$-chuẩn tắc cần tìm. $\square$

**4.1.3 Thí dụ.** Hãy đưa dạng toàn phương trên $\mathbb{R}^3$ sau đây về dạng chuẩn tắc và tìm cơ sở $p$-chuẩn tắc tương ứng.

$$p(x, y, z) = 2x^2 + 8xy + 4xz + 9y^2 + 12yz + 9z^2$$

Theo thuật toán Lagrange ta biến đổi được

$$\begin{aligned} p(x, y, z) &= 2x^2 + 4x(2y + z) + 9y^2 + 12yz + 9z^2 \\ &= 2(x + 2y + z)^2 + y^2 + 4yz + 7z^2 \\ &= 2(x + 2y + z)^2 + (y + 2z)^2 + 3z^2 \\ &= x'^2 + y'^2 + z'^2 \end{aligned}$$

và ta có công thức chuyển tọa độ từ cơ sở $p$-chuẩn tắc sang cơ sở chính tắc của $\mathbb{R}^3$

$$\begin{cases} x' = \sqrt{2}(x + 2y + z) \\ y' = y + 2z \\ z' = \sqrt{3}z \end{cases} \qquad \text{hay} \qquad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 2\sqrt{2} & \sqrt{2} \\ 0 & 1 & 2 \\ 0 & 0 & \sqrt{3} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Do đó, dạng toàn phương $p$ là không suy biến, và ma trận chuyển tọa độ từ cơ sở chính tắc của $\mathbb{R}^3$ sang cơ sở $p$-chuẩn tắc là

$$A = \begin{bmatrix} \sqrt{2} & 2\sqrt{2} & \sqrt{2} \\ 0 & 1 & 2 \\ 0 & 0 & \sqrt{3} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -2 & \sqrt{3} \\ 0 & 1 & -\frac{2}{\sqrt{3}} \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{bmatrix}$$

và cơ sở $p$-chuẩn tắc là $\mathscr{B} = ((\frac{1}{\sqrt{2}}, 0, 0), (-2, 1, 0), (\sqrt{3}, -\frac{2}{\sqrt{3}}, \frac{1}{\sqrt{3}}))$. $\square$

## 4.2 Định lý quán tính

Vì đối với dạng toàn phương bất kỳ nói chung có vô số cơ sở $p$-chính tắc, nên (theo chứng minh Định lý 4.1.2) nói chung cũng có vô số cơ sở $p$-chuẩn tắc. Giả sử $\mathscr{B}$ và $\mathscr{B}'$ là hai cơ sở $p$-chuẩn tắc và trong chúng $p$ có biểu thức tọa độ dạng chuẩn tắc lần lượt là

$$p(\mathbf{x}) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2$$
$$p(\mathbf{x}) = x_1'^2 + \cdots + x_{s'}'^2 - x_{s'+1}'^2 - \cdots - x_{s'+t'}'^2$$

trong đó $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$ $\mathbf{x}|_{\mathscr{B}'} = (x_1', x_2', \ldots, x_n')$ và $s + t = s' + t' = \operatorname{rank}(p)$. Một cách tự nhiên, ta có câu hỏi: phải chăng $s = s'$ và $t = t'$? Nói cách khác, liệu dạng chuẩn tắc của $p$ có duy nhất không? Câu trả lời là khẳng định. Đó là nội dung của *luật quán tính Sylvester đối với dạng toàn phương thực* trong định lý sau.

**4.2.1 Định lý.** *Cho $p$ là một dạng toàn phương trên một không gian vector thực $n$ chiều $V$. Khi đó, số $s$ các số hạng mang dấu "$+$" và số $t$ các số hạng mang dấu "$-$" trong mỗi dạng chuẩn tắc của $p$ là các hằng số chỉ phụ thuộc vào $p$, không phụ thuộc vào cơ sở $p$-chuẩn tắc được chọn.*

*Chứng minh.* Giả sử $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ và $\mathscr{B}' = (\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_n')$ là hai cơ sở $p$-chuẩn tắc, và biểu thức tọa độ của $p$ trong $\mathscr{B}$ và $\mathscr{B}'$ lần lượt là

$$p(\mathbf{x}) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2$$
$$p(\mathbf{x}) = x_1'^2 + \cdots + x_{s'}'^2 - x_{s'+1}'^2 - \cdots - x_{s'+t'}'^2$$

với mọi $\mathbf{x} \in V$ mà $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$ $\mathbf{x}|_{\mathscr{B}'} = (x_1', x_2', \ldots, x_n')$ và $s + t = s' + t' = r = \operatorname{rank}(p)$ thỏa $0 \leq s, s', t, t' \leq r \leq n$. Ta cần chứng tỏ $s = s'$ và $t = t'$.

Trước hết, ta chứng tỏ $s \leq s'$. Thật vậy, nếu $s = 0$, thì đương nhiên $s \leq s'$. Giả sử $s > 0$. Đặt $W = \langle \mathbf{e}_1, \ldots, \mathbf{e}_s \rangle$ và $Z = \langle \mathbf{e}_{s'+1}', \ldots, \mathbf{e}_n' \rangle$. Khi đó, với mọi $\mathbf{x} \in W \subset V$, ta có $\mathbf{x}$ biểu thị tuyến tính qua $(\mathbf{e}_1, \ldots, \mathbf{e}_s)$ tức là

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_s\mathbf{e}_s = x_1\mathbf{e}_1 + \cdots + x_s\mathbf{e}_s + 0\mathbf{e}_{s+1} + \cdots + 0\mathbf{e}_n.$$

Do đó, $\mathbf{x}|_{\mathscr{B}} = (x_1, \ldots, x_s, 0, \ldots, 0)$ và $p(\mathbf{x}) = x_1^2 + \cdots + x_s^2 \geq 0$. Hơn nữa, $p(\mathbf{x}) = 0$ (với $\mathbf{x} \in W$) khi và chỉ khi $\mathbf{x} = \mathbf{0}$.

Tương tự, mọi $\mathbf{x} \in Z$ đều có tọa độ đối với cơ sở $\mathscr{B}'$ dạng $\mathbf{x}|_{\mathscr{B}'} = (0, \ldots, 0, x'_{s'+1}, \ldots, x'_n)$. Do đó, $p(\mathbf{x}) = -x'^2_{s'+1} - \cdots - x'^2_{s'+t'} \leq 0$. Suy ra, nếu $\mathbf{x} \in W \cap Z$ thì $p(\mathbf{x}) \geq 0$ và $p(\mathbf{x}) \leq 0$, cho nên $p(\mathbf{x}) = 0$ hay $\mathbf{x} = \mathbf{0}$. Vậy $W \cap Z = \{\mathbf{0}\}$. Ta có,

$$\begin{aligned} s + (n - s') &= \dim(W) + \dim(Z) \\ &= \dim(W + Z) - \dim(W \cap Z) \\ &= \dim(W + Z) \leq n. \end{aligned}$$

Từ đó ta được $s \leq s'$.

Vì vai trò $s$ và $s'$ bình đẳng nên bằng lập luận tương tự ta cũng có $s' \leq s$; tức là $s = s'$. Suy ra $t = r - s = r - s' = t'$. □

**4.2.2 Định nghĩa.** Giả sử dạng chuẩn tắc của dạng toàn phương $p$ trên không gian vector thực $n$ chiều $V$ là

$$p(\mathbf{x}) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2,$$

với $0 \leq s, t \leq r = \operatorname{rank} p$, $s + t = r$. Khi đó, số $s$ (các số hạng mang dấu "$+$") gọi là **chỉ số dương quán tính** của $p$; còn số $t$ (các số hạng mang dấu "$-$") gọi là **chỉ số âm quán tính**. Cặp $(s, t)$ được gọi là **cặp chỉ số quán tính của** $p$.

**4.2.3 Thí dụ.** Xét dạng toàn phương trên không gian $\mathbb{R}^3$ sau

$$p(x_1, x_2, x_3) = x_1^2 - 3x_3^2 - 2x_1x_2 + 2x_1x_3 - 6x_2x_3.$$

Theo thuật toán Lagrange ta biến đổi như sau:

$$\begin{aligned} p(x_1, x_2, x_3) &= (x_1 - x_2 + x_3)^2 - x_2^2 - 4x_3^2 - 4x_2x_3 \\ &= (x_1 - x_2 + x_3)^2 - (x_2 + 2x_3)^2 \\ &= x_1'^2 - x_2'^2 \end{aligned}$$

Như vậy, dạng toàn phương $p$ là suy biến với chỉ số dương quán tính là $s = 1$ và chỉ số âm quán tính là $t = 1$, và

$$\begin{cases} x'_1 = x_1 - x_2 + x_3 \\ x'_2 = x_2 + 2x_3 \\ x'_3 = x_3 \end{cases}$$

Ma trận chuyển cơ sở từ cơ sở chính tắc trong $\mathbb{R}^3$ sang cơ sở $p$-chuẩn tắc là

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

Cơ sở $p$-chuẩn tắc là $\mathscr{B} = ((1,0,0),(1,1,0),(-3,-2,1))$. $\square$

### Bài tập

**1**) Đưa các dạng toàn phương sau (trên không gian vector $\mathbb{R}^3$) về dạng chuẩn tắc và xác định chỉ số quán tính của chúng.

**(a)** $2x_1^2 + 3x_2^2 + 4x_3^2 - 2x_1x_2 + 4x_1x_3 - 3x_2x_3$

**(b)** $x_1^2 - 3x_2^2 + x_3^2 - 3x_1x_2 + 5x_1x_3 - 4x_2x_3$

**(c)** $x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$

**(d)** $x_1^2 - 2x_2^2 + x_3^2 + 2x_1x_2 + 4x_1x_3 + 2x_2x_3$

# § 5 Phân loại dạng toàn phương thực

## 5.1 Phân loại dạng toàn phương thực

**5.1.1 Định nghĩa.** Cho $p$ là một dạng toàn phương trên không gian vector thực $V$. Ta bảo $p$ là **không âm** (tương ứng **không dương**) nếu $p(\mathbf{x}) \geq 0$ (tương ứng $p(\mathbf{x}) \leq 0$) với mọi $\mathbf{x} \in V$. Các dạng toàn phương không âm hay không dương được gọi chung là các dạng toàn phương **có dấu xác định**.

Ngược lại, $p$ được gọi là **đổi dấu** nếu tồn tại cặp vector khác không $\mathbf{x}, \mathbf{y} \in V$ sao cho

$$p(\mathbf{x}) > 0 > p(\mathbf{y}).$$

Ta bảo $p$ là **xác định dương** (tương ứng **xác định âm**) nếu $p$ không âm (tương ứng, không dương) và

$$p(\mathbf{x}) > 0 \qquad\qquad \text{(tương ứng } p(\mathbf{x}) < 0\text{)}$$

với mọi $\mathbf{x} \in V$ và $\mathbf{x} \neq \mathbf{0}$.

**5.1.2 Nhận xét.** Giả sử $p$ là một dạng toàn phương trên không gian vector thực $n$ chiều với hạng $r$ và có dạng chính tắc trên cơ sở $p$-chính tắc nào đó là

$$p(\mathbf{x}) = a_1x_1^2 + a_2x_2^2 + \cdots + a_rx_r^2$$

với $(x_1, x_2, \ldots, x_n)$ là tọa độ của $\mathbf{x}$ đối với cơ sở $p$-chính tắc đang xét. Khi đó

**(1)** Chỉ số dương quán tính của $p$ là số các hệ số chính tắc $a_i > 0$; và chỉ số âm quán tính của $p$ là số các hệ số chính tắc $a_i < 0$.

**(2)** $p$ không âm khi và chỉ khi $a_i > 0$ với mọi $i = 1, \ldots, r$; và $p$ không dương khi và chỉ khi $a_i < 0$ với mọi $i = 1, \ldots, r$.

**(3)** $p$ xác định dương khi và chỉ khi $r = n$ và $a_i > 0$ với mọi $i = 1, \ldots, n$; và $p$ xác định âm khi và chỉ khi $r = n$ và $a_i < 0$ với mọi $i = 1, \ldots, n$.

**(4)** $p$ xác định dương khi và chỉ khi $-p$ xác định âm.

**5.1.3 Thí dụ.** Phân loại dạng toàn phương trên $\mathbb{R}^3$ sau.

$$p(x_1, x_2, x_3) = 3x_1^2 - 2x_2^2 + 2x_3^2 + 4x_1x_2 - 3x_1x_3 - x_2x_3.$$

Dùng thuật toán Lagrange đưa $p$ về dạng chính tắc:

$$\begin{aligned} p(x_1,x_2,x_3) &= 3(x_1+\tfrac{2}{3}x_2-\tfrac{1}{2}x_3)^2-\tfrac{4}{3}x_2^2-\tfrac{3}{4}x_3^2+2x_2x_3 \\ &\qquad -2x_2^2+2x_3^2-x_2x_3 \\ &= 3(x_1+\tfrac{2}{3}x_2-\tfrac{1}{2}x_3)^2-\tfrac{10}{3}(x_2^2-\tfrac{3}{10}x_2x_3)+\tfrac{5}{4}x_3^2 \\ &= 3(x_1+\tfrac{2}{3}x_2-\tfrac{1}{2}x_3)^2-\tfrac{10}{3}(x_2-\tfrac{3}{20}x_3)^2+\tfrac{53}{40}x_3^2 \end{aligned}$$

Vậy dạng toàn phương $p$ không suy biến, đổi dấu, có chỉ số dương quán tính $s=2$ và chỉ số âm quán tính $t=1$. □

**5.1.4 Định lý.** *Cho $p$ là một dạng toàn phương trên không gian vector thực $V$ xác định bởi dạng song tuyến tính đối xứng $\xi$ trên $V$. Khi đó, các khẳng định sau đây là tương đương:*

**(1)** *$p$ có dấu xác định.*

**(2)** *$p$ và $\xi$ thỏa mãn bất đẳng thức Schwarz*

$$[\xi(\mathbf{x},\mathbf{y})]^2 \le p(\mathbf{x})p(\mathbf{y}) \qquad \text{với mọi } \mathbf{x},\mathbf{y}\in V. \tag{5.1.5}$$

*Hơn nữa, nếu $p$ là xác định dương hay xác định âm, thì dấu bằng trong bất đẳng thức trên xảy ra khi và chỉ khi hệ vector $(\mathbf{x},\mathbf{y})$ phụ thuộc tuyến tính.*

*Chứng minh.* (1) ⇒ (2) Giả sử $p$ có dấu xác định. Không mất tính tổng quát ta giả sử $p$ không âm (trường hợp $p$ không dương được lập luận tương tự). Lấy cặp vector $\mathbf{x},\mathbf{y}\in V$ tùy ý. Nếu $p(\mathbf{x})=p(\mathbf{y})=0$, thì ta có

$$\xi(\mathbf{x},\mathbf{y}) = \frac{1}{2}[p(\mathbf{x}+\mathbf{y})-p(\mathbf{x})-p(\mathbf{y})] = \frac{1}{2}p(\mathbf{x}+\mathbf{y}) \ge 0,$$
$$-\xi(\mathbf{x},\mathbf{y}) = \xi(\mathbf{x},-\mathbf{y}) = \frac{1}{2}[p(\mathbf{x}-\mathbf{y})-p(\mathbf{x})-p(\mathbf{y})] = \frac{1}{2}p(\mathbf{x}-\mathbf{y}) \ge 0.$$

Do đó, $\xi(\mathbf{x},\mathbf{y}) \ge 0 \ge \xi(\mathbf{x},\mathbf{y})$ hay $\xi(\mathbf{x},\mathbf{y})=0$. Vậy $0=[\xi(\mathbf{x},\mathbf{y})]^2 \le p(\mathbf{x})p(\mathbf{y})$, tức là bất đẳng thức Schwarz thỏa mãn.

Giả sử $p(\mathbf{x})\neq 0$, do đó $p(\mathbf{x})>0$ (trường hợp $p(\mathbf{y})>0$ làm tương tự). Khi đó, với mỗi $\lambda\in\mathbb{R}$, ta có $p(\lambda\mathbf{x}+\mathbf{y})\ge 0$ (vì $p$ không âm). Mặt khác ta có

$$p(\lambda\mathbf{x}+\mathbf{y}) = \xi(\lambda\mathbf{x}+\mathbf{y},\lambda\mathbf{x}+\mathbf{y}) = \lambda^2 p(\mathbf{x})+2\lambda\xi(\mathbf{x},\mathbf{y})+p(\mathbf{y}).$$

Vậy ta có $\lambda^2 p(\mathbf{x}) + 2\lambda\xi(\mathbf{x},\mathbf{y}) + p(\mathbf{y}) \geq 0$ với mọi $\lambda \in \mathbb{R}$ và $q(\mathbf{x}) > 0$, cho nên $[\xi(\mathbf{x},\mathbf{y})]^2 - p(\mathbf{x})p(\mathbf{y}) \geq 0$ hay $[\xi(\mathbf{x},\mathbf{y})]^2 \leq p(\mathbf{x})p(\mathbf{y})$. Như vậy, bất đẳng thức Schwarz đúng với mọi $\mathbf{x},\mathbf{y} \in V$.

(2) $\Rightarrow$ (1) Giả sử $[\xi(\mathbf{x},\mathbf{y})]^2 \leq p(\mathbf{x})p(\mathbf{y})$ với mọi $\mathbf{x},\mathbf{y} \in V$. Khi đó, $p(\mathbf{x})p(\mathbf{y}) \geq 0$ với mọi $\mathbf{x},\mathbf{y} \in V$. Do đó, $p$ không đổi dấu hay $p$ có dấu xác định.

Khi $p$ là xác định dương hay âm rõ ràng nếu $p(\mathbf{x}) = 0$ thì $\mathbf{x} = \mathbf{0}$, suy ra $(\mathbf{x},\mathbf{y})$ phụ thuộc tuyến tính. Nếu $p(\mathbf{x}) \neq 0$, thì theo chứng minh trên ta thấy dấu bằng xảy ra trong bất đẳng thức (5.1.5) khi và chỉ khi tồn tại $\lambda_0$ sao cho $p(\lambda_0\mathbf{x} + \mathbf{y}) = 0$ hay $\lambda_0\mathbf{x} + \mathbf{y} = \mathbf{0}$; điều này tương đương với $(\mathbf{x},\mathbf{y})$ phụ thuộc tuyến tính. □

## 5.2 Định lý Sylvester

**5.2.1 Định lý.** *Cho dạng toàn phương $p$ trên không gian vector thực $n$ chiều $V$ và $A = (a_{ij})_n$ là ma trận của $p$ trong một cơ sở $\mathscr{B}$ nào đó của $V$. Gọi $D_1$, $D_2$, ..., $D_n$ là $n$ định thức con chính của $A$. Khi đó,*

**(1)** *$p$ xác định dương khi và chỉ khi $D_i > 0$ với mọi $i = 1,\dots,n$.*

**(2)** *$p$ xác định âm khi và chỉ khi $D_1, D_2, \dots, D_n$ đan dấu và $D_1 < 0$.*

*Chứng minh.* (1) ($\Rightarrow$) Giả sử $p$ xác định dương. Trước hết, ta chứng tỏ $D_i \neq 0$ với mọi $i = 1,\dots,n$. Giả sử $D_k = 0$ với $k$ nào đó thuộc $\{1,\dots,n\}$. Xét phương trình tuyến tính thuần nhất

$$(5.2.2)\qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1k}x_k = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kk}x_k = 0 \end{cases}$$

Vì $D_k = 0$ nên hệ phương trình có ít nhất một nghiệm không tầm thường $(a_1, a_2, \dots, a_k) \neq \mathbf{0}$. Xét $\mathbf{a} \in V$ sao cho $\mathbf{a}|_{\mathscr{B}} = (a_1,\dots,a_k,0,\dots,0)$, nói riêng $\mathbf{a} \neq \mathbf{0}$. Khi đó,

$$p(\mathbf{a}) = \mathbf{a}|_{\mathscr{B}}^t A\mathbf{a}|_{\mathscr{B}} = 0$$

trái với giả thiết $p$ xác định dương. Như vậy $D_i \neq 0$ với mọi $i = 1, \ldots, n$.

Dùng thuật toán Jacobi đưa $p$ về dạng chính tắc ta được

$$p(\mathbf{x}) = D_1 y_1^2 + \frac{D_2}{D_1} y_2^2 + \cdots + \frac{D_n}{D_{n-1}} y_n^2$$

với $(y_1, y_2, \ldots, y_n)$ là tọa độ của **x** đối với cơ sở $p$-chính tắc trong thuật toán Jacobi. Vì $p$ xác định dương nên $D_1 > 0$, $\dfrac{D_2}{D_1} > 0, \ldots, \dfrac{D_n}{D_{n-1}} > 0$ hay $D_i > 0$ với mọi $i = 1, \ldots, n$.

($\Leftarrow$) Giả sử $D_i > 0$ với mọi $i = 1, \ldots, n$. Khi đó thuật toán Jacobi cho ta dạng chính tắc của $p$

$$p(\mathbf{x}) = D_1 y_1^2 + \frac{D_2}{D_1} y_2^2 + \cdots + \frac{D_n}{D_{n-1}} y_n^2$$

trong đó $(y_1, y_2, \ldots, y_n)$ là tọa độ của của **x** đối với cơ sở $p$-chính tắc trong thuật toán Jacobi. Do $D_i > 0$ với mọi $i = 1, \ldots, n$ nên $D_1 > 0$, $\dfrac{D_2}{D_1} > 0, \ldots, \dfrac{D_n}{D_{n-1}} > 0$, cho nên $p$ xác định dương.

(2) Ta biết rằng $p$ xác định âm khi và chỉ khi $-p$ xác định dương. Ma trận của $-p$ trong cơ sở $\mathscr{B}$ là $-A = (-a_{ij})_n$. Do đó, áp dụng kết quả (1) ta sẽ có được điều phải chứng minh. $\square$

**5.2.3 Thí dụ.** Tìm $\lambda$ để dạng toàn phương trên $\mathbb{R}^3$ sau xác định dương, xác định âm.

$$p(x_1, x_2, x_3) = 5x_1^2 + x_2^2 + \lambda x_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3.$$

Ma trận của $p$ trong cơ sở chính tắc là $A = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & \lambda \end{bmatrix}$. Các định thức con chính của $A$ là

$$D_1 = 5, \qquad D_2 = 1, \qquad D_3 = \lambda - 2.$$

Vậy $p$ xác định dương khi và chỉ khi $\lambda > 2$, và không có giá trị $\lambda$ nào để $p$ xác định âm. $\square$

**5.2.4 Thí dụ.** Dùng tiêu chuẩn Sylvester hãy biện luận về tính xác định âm của dạng toàn phương trên $\mathbb{R}^3$ sau.

$$p(x_1, x_2, x_3) = -2x_1^2 - 2x_2^2 - x_3^2 + 2x_1x_2 + 2ax_1x_3.$$

Ma trận của $p$ trong cơ sở chính tắc là $A = \begin{bmatrix} -2 & 1 & a \\ 1 & -2 & 0 \\ a & 0 & -1 \end{bmatrix}$. Các định thức con chính của $A$ là

$$D_1 = -2, \qquad D_2 = 3, \qquad D_3 = 2a^2 - 3.$$

Ta có, $p$ xác định âm khi và chỉ khi $2a^2 - 3 < 0$ hay $-\frac{\sqrt{6}}{2} < a < \frac{\sqrt{6}}{2}$. $\square$

## Bài tập

**1**) Tìm $\lambda$ sao cho các dạng toàn phương trên $\mathbb{R}^3$ sau đây xác định dương.

**(a)** $p(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + 3x_3^2 + 2\lambda x_1x_2 + 2x_1x_3$

**(b)** $p(x_1, x_2, x_3) = x_1^2 + x_2^2 + 5x_3^2 + 2\lambda x_1x_2 - 2x_1x_3 + 4x_2x_3$

**(c)** $p(x_1, x_2, x_3) = x_1^2 + 4x_2^2 + x_3^2 + 2\lambda x_1x_2 + 10x_1x_3 + 6x_2x_3$

**2**) Dùng tiêu chuẩn Sylvester hãy xét xem dạng toàn phương trên $\mathbb{R}^2$ nào là xác định dương, xác định âm?

**(a)** $p(x_1, x_2) = x_1^2 + 26x_2^2 + 10x_1x_2$

**(b)** $p(x_1, x_2) = -x_1^2 + 2x_1x_2 - 4x_2^2$

**3**) Dùng tiêu chuẩn Sylvester hãy xét xem dạng toàn phương trên $\mathbb{R}^3$ nào là xác định dương, xác định âm?

**(a)** $p(x_1, x_2, x_3) = -11x_1^2 - 6x_2^2 - 6x_3^2 + 12x_1x_2 - 12x_1x_3 + 6x_2x_3$

**(b)** $p(x_1, x_2, x_3) = 9x_1^2 + 6x_2^2 + 6x_3^2 + 12x_1x_2 - 10x_1x_3 - 2x_2x_3$

**4**) Cho dạng toàn phương ba biến thực:

$$q(x_1, x_2, x_3) = 2x_1^2 + 8x_1x_2 + 4x_1x_3 + 9x_2^2 + 12x_2x_3 + mx_3^2.$$

Hãy đưa $q$ về dạng chính tắc. Biện luận theo tham số $m$ về tính suy biến, không suy biến, tính xác định dương của $q$.

**5**) Chứng minh rằng dạng toàn phương là xác định dương khi và chỉ khi ma trận $A$ của nó có thể biểu diễn dưới dạng $A = C^tC$, trong đó $C$ là ma trận thực không suy biến.

**6**) Cho dạng toàn phương $p$ có ma trận tương ứng là $A$. Giả sử $p$ xác định dương. Chứng minh rằng dạng toàn phương có ma trận $A^{-1}$ cũng xác định dương.

**7**) Chứng minh rằng nếu dạng toàn phương $p$ trên không gian vector $V$ là xác định dương (t.ư âm) thì thu hẹp của $p$ lên không gian con bất kỳ của $V$ cũng xác định dương (t.ư âm).

**8**) Chứng minh rằng dạng toàn phương $p$ có ma trận trong một cơ sở nào đó như sau là xác định dương

$$A = \begin{bmatrix} 0.5 & 1 & 1 & 1 & \dots & 1 \\ 1 & 2.5 & 3 & 3 & \dots & 3 \\ 1 & 3 & 4.5 & 5 & \dots & 5 \\ 1 & 3 & 5 & 6.5 & \dots & 7 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 3 & 5 & 7 & \dots & \frac{4k-3}{2} \end{bmatrix}$$

# § 6 Dạng song tuyến tính và dạng toàn phương Hermite trên không gian phức

## 6.1 Dạng chuẩn tắc của dạng toàn phương phức

**6.1.1 Định nghĩa.** Cho $p$ là một dạng toàn phương trên không gian vector phức $n$ chiều $V$. Nếu trong một cơ sở $\mathscr{B}$ nào đó của $V$, $p$ có biểu thức tọa độ

$$p(\mathbf{x}) = x_1^2 + x_2^2 + \cdots + x_r^2 \tag{6.1.2}$$

với mọi $\mathbf{x} \in V$ mà $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$, $r = \operatorname{rank}(p)$ thì $p$ gọi là có biểu thức tọa độ **dạng chuẩn tắc**. Lúc đó, $\mathscr{B}$ gọi là **cơ sở $p$-chuẩn tắc**.

**6.1.3 Định lý.** *Mọi dạng toàn phương p trên không gian vector phức đều có cơ sở p-chuẩn tắc, tức là luôn đưa được p về dạng chuẩn tắc.*

*Chứng minh.* Trước hết, ta chọn cơ sở $p$-chính tắc $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ để $p$ có dạng chính tắc

$$p(\mathbf{x}) = a_1x_1^2 + a_2x_2^2 + \cdots + a_rx_2^2, \quad a_1a_2\cdots a_r \neq 0, \quad r = \operatorname{rank}(p)$$

Sau đó, dùng phép đổi tọa độ $x_j = \frac{1}{\sqrt{a_j}}x'_j$ với $j = 1, \ldots, r$ và $x_j = x'_j$ với $j = r+1, \ldots, n$, tức là thay cơ sở $\mathscr{B}$ bởi cơ sở $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n)$ với $\mathbf{e}'_j = \frac{1}{\sqrt{a_j}}\mathbf{e}_j$ với $j = 1, \ldots, r$ và $\mathbf{e}'_j = \mathbf{e}_j$ với $j = r+1, \ldots, n$ (ở đó $\sqrt{a_j}$ là một căn bậc hai phức của $a_j$ với $j = 1, \ldots, r$). Dễ thấy rằng $\mathscr{B}'$ là một cơ sở $p$-chuẩn tắc và trong $\mathscr{B}'$, $p$ có dạng chuẩn tắc (6.1.2). □

## 6.2 Dạng song tuyến tính và dạng toàn phương Hermite

**6.2.1 Định nghĩa.** Cho $V$ là một không gian vector phức. Ánh xạ $\xi : V \times V \to \mathbb{C}$ gọi là một **dạng song tuyến tính Hermite** (hay **dạng Hermite**) trên $V$ nếu $\xi$ thỏa mãn các tính chất sau:

**(i)** $\xi(\mathbf{x}, \mathbf{y}) = \overline{\xi(\mathbf{y}, \mathbf{x})}$

**(ii)** $\xi(\lambda\mathbf{x} + \mu\mathbf{y}, \mathbf{z}) = \lambda\xi(\mathbf{x}, \mathbf{z}) + \mu\xi(\mathbf{y}, \mathbf{z})$

với mọi $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ với mọi $\lambda, \mu \in \mathbb{C}$.

Ánh xạ $p : V \to \mathbb{C}$ gọi là một **dạng toàn phương Hermite** trên $V$ nếu có một dạng song tuyến tính $\xi$ trên $V$ để $p(\mathbf{x}) = \xi(\mathbf{x}, \mathbf{x})$, với mọi $\mathbf{x} \in V$. Khi đó $p$ được gọi là **dạng toàn phương Hermite xác định bởi** $\xi$, còn $\xi$ gọi là **dạng cực** của $p$.

**6.2.2 Thí dụ.** Xét không gian vector phức các hàm số $\varphi$ giá trị phức liên tục trên đoạn $[a, b]$. Khi đó ánh xạ

$$(\varphi, \psi) \mapsto \int_a^b \varphi(t)\overline{\psi(t)}dt$$

là một dạng song tuyến tính Hermite mà dạng toàn phương Hermite ứng với nó là

$$\varphi \mapsto \int_a^b \varphi(t)\overline{\varphi(t)}dt = \int_a^b |\varphi(t)|^2 dt. \qquad \square$$

**6.2.3 Mệnh đề.** *Nếu $\xi$ là một dạng song tuyến tính Hermite trên $V$ thì ta có*

**(i)** $\xi(\mathbf{x}, \lambda\mathbf{y} + \mu\mathbf{z}) = \bar{\lambda}\xi(\mathbf{x}, \mathbf{y}) + \bar{\mu}\xi(\mathbf{x}, \mathbf{z})$

**(ii)** $\xi(\mathbf{x}, \mathbf{x}) \in \mathbb{R}$

**(iii)** $\xi(\mathbf{x}, \mathbf{0}) = \xi(\mathbf{0}, \mathbf{x}) = 0$

**(iv)** $\xi\Big(\sum_{i=1}^{n} \lambda_i \mathbf{x}_i, \sum_{j=1}^{m} \mu_j \mathbf{y}_j\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} \lambda_i \bar{\mu}_j \xi(\mathbf{x}_i, \mathbf{y}_j)$

*với mọi* $\mathbf{x}, \mathbf{y}, \mathbf{z}; \mathbf{x}_1, \ldots, \mathbf{x}_n; \mathbf{y}_1, \ldots, \mathbf{y}_m \in V$, *mọi* $\lambda, \mu, \lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in \mathbb{C}$

*Chứng minh.* Dành cho bạn đọc xem như bài tập. $\square$

**6.2.4 Mệnh đề.** *Nếu $p$ là một dạng toàn phương Hermite thì $p(\mathbf{x}) \in \mathbb{R}$ với mọi $\mathbf{x} \in V$, tức là $p : V \to \mathbb{R}$.*

*Chứng minh.* Đây chính là tính chất (ii) của mệnh đề trên. $\square$

**6.2.5 Mệnh đề.** *Ký hiệu $\mathcal{Q}_H(V)$ và $\mathcal{S}_H(V)$ tương ứng là tập hợp các dạng toàn phương và các dạng song tuyến tính Hermite trên $V$ thì ta có một song ánh $\mathcal{S}_H(V) \to \mathcal{Q}_H(V)$ cho ứng mỗi $\xi \in \mathcal{S}_H(V)$ với dạng toàn* *phương $p \in \mathcal{Q}_H(V)$ xác định bởi $\xi$. Cụ thể $p$ và $\xi$ được liên hệ với nhau bởi hệ thức thuận nghịch sau đây*

$$\begin{aligned} p(\mathbf{x}) &= \xi(\mathbf{x}, \mathbf{x}) \\ \xi(\mathbf{x}, \mathbf{y}) &= \frac{1}{4}[p(\mathbf{x}+\mathbf{y}) - p(\mathbf{x}-\mathbf{y}) + ip(\mathbf{x}+i\mathbf{y}) - ip(\mathbf{x}-i\mathbf{y})] \end{aligned}$$

*với mọi $\mathbf{x}, \mathbf{y} \in V$.*

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

**6.2.6** Khi trong $V$ đã cho cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$, mỗi $\xi \in \mathcal{S}_H(V)$ hay $p \in \mathcal{Q}_H(V)$ cũng được biểu thị bởi một ma trận $A = (a_{ij})_n \in \mathcal{M}_n(\mathbb{C})$ có tính đối xứng liên hợp (tức là $a_{ij} = \overline{a_{ji}}$ với $i, j = 1, \dots, n$) và biểu thức tọa độ như sau

$$\xi(\mathbf{x}, \mathbf{y}) = \mathbf{x}|_{\mathscr{B}}^t A \overline{\mathbf{y}|_{\mathscr{B}}}$$

hay

$$p(\mathbf{x}) = \mathbf{x}|_{\mathscr{B}}^t A \overline{\mathbf{x}|_{\mathscr{B}}}$$

với mọi $\mathbf{x}, \mathbf{y} \in V$. Hơn nữa, có thể chọn cơ sở $\mathscr{B}$ để $p$ có dạng chuẩn tắc, tức là biểu thức tọa độ của $p$ trong $\mathscr{B}$ dạng

$$p(\mathbf{x}) = x_1\bar{x}_1 + x_2\bar{x}_2 + \cdots + x_s\bar{x}_s - x_{s+1}\bar{x}_{s+1} - \cdots - x_{s+t}\bar{x}_{s+t},$$

với $0 \le s, t \le r = \operatorname{rank}(p) \le n$, $s + t = r$. Các số $s, t$ cũng bất biến đối với $p$ đã cho và cũng gọi là **chỉ số dương, âm quán tính** của $p$.

**6.2.7** Cũng giống như dạng toàn phương thực, ta cũng phân loại dạng toàn phương Hermite trên không gian vector phức gồm dạng toàn phương Hermite không âm, không dương, có dấu xác định, đổi dấu, xác định dương, xác định âm.

## Bài tập

**1)** Cho $V$ làm một không gian vector phức. Một ánh xạ $\xi : V \times V \to \mathbb{C}$ được gọi là một **dạng tuyến tính rưỡi** trên $V$ nếu thỏa hai điều kiện

**(i)** $\xi(\lambda\mathbf{x} + \mu\mathbf{x}', \mathbf{y}) = \lambda\xi(\mathbf{x}, \mathbf{y}) + \mu\xi(\mathbf{x}', \mathbf{y})$

**(ii)** $\xi(\mathbf{x}, \lambda\mathbf{y} + \mu\mathbf{y}') = \overline{\lambda}\xi(\mathbf{x}, \mathbf{y}) + \overline{\mu}\xi(\mathbf{x}, \mathbf{y}')$

Khi đó,

**(a)** hãy xác định biểu thức tọa độ của dạng tuyến tính rưỡi $\xi$ đối với cơ sở $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ của $V$.

**(b)** Cho $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ là một cơ sở khác của $V$. Giả sử $C$ la ma trận chuyển cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$. Hãy xác định ma trận của $\xi$ đối với cơ sở $\mathscr{B}$.

**2)** Hãy đưa dạng toàn phương Hermite trên $\mathbb{C}^3$ sau đây về dạng chính tắc

**(a)** $p(x, y, z) = 2x\bar{x} + (2 + 3i)x\bar{y} + (4 - 5i)x\bar{z} + (2 - 3i)y\bar{x} + 5y\bar{y} + (6 + 2i)y\bar{z} + (4 + 5i)z\bar{x} + (6 - 2i)z\bar{y} - 7z\bar{z}$

**(b)** $p(x, y, z) = 3x\bar{x} + (2+i)x\bar{y} + (4+i)x\bar{z} + (2-i)y\bar{x} + 6y\bar{y} + iy\bar{z} + (4+i)z\bar{x} + iz\bar{y} + 3z\bar{z}$

**(c)** $p(x, y, z) = 4x\bar{x} - 3x\bar{y} + 5x\bar{z} - 3y\bar{x} + 2y\bar{y} + y\bar{z} + 5z\bar{x} + z\bar{y} - 6z\bar{z}$

**3)** Với ma trận $H \in \mathcal{M}_3(\mathbb{C})$ sau đây, hãy tìm ma trận khả nghịch $C$ sao cho $C^t HC$ là ma trận chéo.

**(a)** $\begin{bmatrix} 1 & 1+i & 2i \\ 1-i & 4 & 2-3i \\ -2i & 2+3i & 7 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 1+2i & 1-i \\ 1-2i & 4 & 2+i \\ 1+i & 2-i & 2 \end{bmatrix}$

**(c)** $\begin{bmatrix} 2 & 2-i & 1+2i \\ 2+i & 4 & 1-3i \\ 1-2i & 1+3i & 3 \end{bmatrix}$ **(d)** $\begin{bmatrix} 2 & 1-i & 2+3i \\ 1+i & 4 & 4-i \\ 2-3i & 4+i & 6 \end{bmatrix}$

**4)** Chứng minh Mệnh đề 6.2.3.

**5)** Chứng minh Mệnh đề 6.2.5.

# Chương VII

# Không gian Euclid

## § 1 Khái niệm về không gian Euclid

### 1.1 Không gian Euclid

**1.1.1 Định nghĩa.** Một dạng song tuyến tính đối xứng xác định dương $\xi$ trên một không gian vector thực $V$ được gọi là **tích vô hướng** trên $V$. Khi đó, giá trị $\xi(\mathbf{x}, \mathbf{y}) \in \mathbb{R}$ được gọi là tích vô hướng của **x** với **y** và thường được ký hiệu là $\langle \mathbf{x}, \mathbf{y} \rangle$. Như vậy, một ánh xạ từ $V \times V$ đến $\mathbb{R}$ với $(\mathbf{x}, \mathbf{y}) \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$ là một tích vô hướng trên $V$ khi và chỉ khi nó thỏa các điều kiện sau:

**(1)** $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ với mọi $\mathbf{x}, \mathbf{y} \in V$

**(2)** $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$ với mọi $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$

**(3)** $\langle \lambda\mathbf{x}, \mathbf{y} \rangle = \lambda\langle \mathbf{x}, \mathbf{y} \rangle$ với mọi $\mathbf{x}, \mathbf{y} \in V$ và mọi $\lambda \in \mathbb{R}$

**(4)** $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ với mọi $\mathbf{x} \in V$ và $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ khi và chỉ khi $\mathbf{x} = \mathbf{0}$.

**1.1.2 Định nghĩa.** Mỗi không gian vector thực $V$ cùng với tích vô hướng đã cho trên nó gọi là một **không gian Euclid**, kí hiệu là $V_E$. Nếu $V$ là không gian $n$ chiều thì $V_E$ cũng gọi là *không gian Euclid $n$ chiều*.

**1.1.3 Thí dụ.** Không gian $\mathbb{R}^n$ trở thành một không gian Euclid $n$ chiều với *tích vô hướng chính tắc*, tức là tích vô hướng $\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ với

$(\mathbf{x}, \mathbf{y}) \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$ xác định như sau.

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i \quad \text{với } \mathbf{x} = (x_1, x_2, \ldots, x_n), \mathbf{y} = (y_1, y_2, \ldots, y_n).$$

Không gian Euclid thu được gọi là *không gian Euclid chính tắc $n$ chiều.* □

**1.1.4 Thí dụ.** Xét không gian vector $\mathcal{C}_{[a,b]}$ trên $\mathbb{R}$, gồm các hàm thực liên tục trên $[a, b]$. Một ánh xạ $\langle \cdot, \cdot \rangle$ từ $\mathcal{C}_{[a,b]} \times \mathcal{C}_{[a,b]}$ đến $\mathbb{R}$ xác định bởi

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx,$$

với mọi $f, g \in \mathcal{C}_{[a,b]}$, là một tích vô hướng. Thật vậy

$$\begin{aligned}
\langle f, g \rangle &= \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = \langle g, f \rangle \\
\langle f + g, h \rangle &= \int_a^b (f(x) + g(x))h(x)dx = \int_a^b f(x)h(x)dx + \int_a^b g(x)h(x)dx \\
&= \langle f, h \rangle + \langle g, h \rangle \\
\langle \lambda f, g \rangle &= \int_a^b \lambda f(x)g(x)dx = \lambda \int_a^b f(x)g(x)dx = \lambda \langle f, g \rangle \\
\langle f, f \rangle &= \int_a^b f^2(x)dx \geq 0.
\end{aligned}$$

Hơn nữa, do hàm $f$ liên tục trên $[a, b]$ nên ta thấy rằng $\int_a^b f^2(x)dx = 0$ khi và chỉ khi $f(x) = 0$ với mọi $x \in [a, b]$, hay $\langle f, f \rangle = 0$ khi và chỉ khi $f = 0$.

Vậy không gian vector $\mathcal{C}_{[a,b]}$ cùng với tích vô hướng trên là một không gian Euclid. □

**1.1.5 Nhận xét.** Theo định nghĩa không gian Euclid ta có trong các không gian Euclid có đầy đủ mọi khái niệm và tính chất của các không gian vector thực, và các khái niệm tính chất liên quan đến tích vô hướng (đó chính là dạng song tuyến tính đối xứng xác định dương). Chúng tôi nêu ra các tính chất đối với tích vô hướng.

**1.1.6 Tính chất.** *Với mọi* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V_E$, *ta có*

**(i)** $\langle \mathbf{0}, \mathbf{x} \rangle = 0$

**(ii)** $\langle \mathbf{x}, -\mathbf{y} \rangle = \langle -\mathbf{x}, \mathbf{y} \rangle = -\langle \mathbf{x}, \mathbf{y} \rangle$

**(iii)** $\langle \mathbf{x} - \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle - \langle \mathbf{y}, \mathbf{z} \rangle$

**(iv)** *nếu* $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle$ *với mọi* $\mathbf{x} \in V_E$ *thì* $\mathbf{y} = \mathbf{z}$.

**(v)** *(Bất đẳng thức Cauchy-Schwarz)* $\langle \mathbf{x}, \mathbf{y} \rangle^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle$; *dấu* " $=$ " *xảy ra khi và chỉ khi* $(\mathbf{x}, \mathbf{y})$ *phụ thuộc tuyến tính.*

*Chứng minh.* Ta có các tính chất (i), (ii), (iii) do tích vô hướng là một dạng song tuyến tính trên $V$. Ta có bất đẳng thức Schwarz (v) do tích vô hướng là một dạng song tuyến tính xác định dương và (5.1.5) (trang 400). Để chứng minh (iv) ta chỉ việc chọn $\mathbf{x} = \mathbf{y} - \mathbf{x}$ và biến đổi được $\langle \mathbf{y} - \mathbf{z}, \mathbf{y} - \mathbf{z} \rangle = \langle \mathbf{y} - \mathbf{z}, \mathbf{y} \rangle - \langle \mathbf{y} - \mathbf{z}, \mathbf{z} \rangle = 0$, suy ra $\mathbf{y} = \mathbf{z}$. □

**1.1.7** Ta biết rằng một dạng song tuyến tính trên một không gian vector $n$ chiều sẽ có ma trận và biểu thức tọa độ trong một cơ sở của không gian vector đã cho. Ta cũng có điều này đối với tích vô hướng trong không gian Euclid $n$ chiều. Xét không gian Euclid $n$ chiều $V_E$ với một cơ sở của nó là $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$. Khi đó, ma trận của tích vô hướng trong cơ sở $\mathscr{B}$ là $A = (a_{ij})_n$ với $a_{ij} = \langle \mathbf{b}_i, \mathbf{b}_j \rangle$ trong đó $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$. Hơn nữa, công thức tọa độ của tích vô hướng trong cơ sở $\mathscr{B}$ là

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}|_{\mathscr{B}}^t A \mathbf{y}|_{\mathscr{B}} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j,$$

trong đó $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$ và $\mathbf{y}|_{\mathscr{B}} = (y_1, y_2, \ldots, y_n)$.

**1.1.8 Định nghĩa.** Cho $V_E$ là một không gian Euclid. Nếu $W$ là một không gian con của $V_E$, thì tích vô hướng trên $V_E$ sẽ cảm sinh một tích vô hướng trên $W$ và làm cho $W$ trở thành một không gian Euclid $W_E$. Ta cũng gọi $W_E$ là **không gian con** của $V_E$.

## 1.2 Độ dài và góc của các vector

**1.2.1 Định nghĩa.** Cho $V_E$ là không gian Euclid. Với mỗi $\mathbf{x} \in V_E$, do $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ nên nó xác định số thực $\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ mà ta kí hiệu là $\|\mathbf{x}\|$ và gọi là **độ dài** của vector $\mathbf{x}$.

**1.2.2 Mệnh đề. (Tính chất của độ dài)** *Cho $V_E$ là không gian Euclid và $\mathbf{x}, \mathbf{y} \in V_E$, và $\lambda \in \mathbb{R}$. Ta có*

**(i)** $\|\mathbf{x}\| \geq \mathbf{0}$, *và* $\|\mathbf{x}\| = 0$ *khi và chỉ khi* $\mathbf{x} = \mathbf{0}$.

**(ii)** $\|\lambda \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$.

**(iii)** *(Bất đẳng thức Cauchy-Schwarz-Bunhiakowsky)* $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$*; dấu "$=$" xảy ra khi và chỉ khi* $(\mathbf{x}, \mathbf{y})$ *phụ thuộc tuyến tính.*

**(iv)** *(Bất đẳng thức tam giác)* $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

**(v)** *(Đẳng thức hình bình hành)* $2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) = \|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2$.

**(vi)** $\big| \|\mathbf{x}\| - \|\mathbf{y}\| \big| \leq \|\mathbf{x} + \mathbf{y}\|$.

**(vii)** $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$.

**(viii)** $\left\| \dfrac{1}{\|\mathbf{x}\|} \mathbf{x} \right\| = 1$ *với* $\mathbf{x} \neq \mathbf{0}$.

*Chứng minh.* Các tính chất (i), (ii), (vii), (viii) được suy ra dễ dàng từ định nghĩa của độ dài và tích vô hướng. (iii) chính là một dạng khác của bất đẳng thức Schwarz

**(iv)** Ta có $\|\mathbf{x} + \mathbf{y}\|^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$. Do đó, $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

**(v)** Ta có $\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = (\|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2) + (\|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2) = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$.

**(vi)** Ta có $\|\mathbf{x}\| = \|(\mathbf{x} - \mathbf{y}) + \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\|$ và $\|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{x}\|$. Từ đó ta có được kết quả. □

**1.2.3 Thí dụ.** Trong không gian Euclid chính tắc $\mathbb{R}^n$ độ dài của vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ là

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

và bất đẳng thức Cauchy-Schwarz-Bunhiakowski được viết lại là

$$\Big|\sum_{i=1}^n x_i y_i\Big| \leq \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2}$$

Trong không gian Euclid $C_{[a,b]}$ với tích vô hướng được xác định ở ví dụ trước, độ dài của $f \in C_{[a,b]}$ là

$$\|f\| = \sqrt{\int_a^b [f(x)]^2 dx}.$$

Với $f, g \in C_{[a,b]}$, từ bất đẳng thức Cauchy-Schwarz-Bunhiakowski ta có được bất đẳng thức

$$\Big|\int_a^b f(x)g(x)dx\Big| \leq \sqrt{\int_a^b [f(x)]^2 dx}\sqrt{\int_a^b [g(x)]^2 dx}. \qquad \square$$

**1.2.4 Thí dụ.** Trong $\mathcal{C}_{[0,1]}$ xét tích vô hướng xác định trong Thí dụ 1.1.4. Ta có

$$\langle t, e^t\rangle = \int_0^1 te^t dt = 1$$

$$\|t\| = \sqrt{\langle t, t\rangle} = \sqrt{\int_0^1 t^2 dt} = \frac{1}{\sqrt{3}}$$

$$\|e^t\| = \sqrt{\langle e^t, e^t\rangle} = \sqrt{\int_0^1 e^{2t} dt} = \sqrt{\frac{e^2 - 1}{2}}$$

$$\|t + e^t\| = \sqrt{\langle t + e^t, t + e^t\rangle} = \sqrt{\int_0^1 (t + e^t) dt} = \sqrt{\frac{3e^2 + 11}{6}}$$

Ta nhận thấy bất đẳng thức Cauchy-Schwarz-Bunhiakowski và bất đẳng thức tam giác đúng cho trường hợp cụ thể hàm $t$ và $e^t$. $\square$

**1.2.5 Định lý.** *Cho không gian vector thực $V$ và một hàm $\|\cdot\| : V \to \mathbb{R}$ với $\mathbf{x} \mapsto \|\mathbf{x}\|$ thỏa mãn các tính chất*

**(i)** $\|\mathbf{x}\| \geq \mathbf{0}$, *và* $\|\mathbf{x}\| = 0$ *khi và chỉ khi* $\mathbf{x} = \mathbf{0}$,

**(ii)** $\|\lambda\mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$,

**(iii)** $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$,

**(iv)** $2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) = \|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2$,

*với mọi* $\mathbf{x}, \mathbf{y} \in V$. *Khi đó, công thức*

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2) \qquad \text{với mọi } \mathbf{x}, \mathbf{y} \in V$$

*xác định trên $V$ một tích vô hướng; hơn nữa, $\|\cdot\|$ chính là độ dài trên $V$ sinh bởi tích vô hướng này.*

*Chứng minh.* Đầu tiên ta có $\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) = \frac{1}{4}(\|\mathbf{y} + \mathbf{x}\|^2 - \|\mathbf{y} - \mathbf{x}\|^2) = \langle \mathbf{y}, \mathbf{x} \rangle$.

Với $\lambda = 0$ ta có $\langle \lambda\mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}(\|\mathbf{0} + \mathbf{y}\|^2 - \|\mathbf{0} - \mathbf{y}\|^2) = 0 = 0\langle \mathbf{x}, \mathbf{y} \rangle = \lambda\langle \mathbf{x}, \mathbf{y} \rangle$. Rõ ràng $\langle \lambda\mathbf{x}, \mathbf{y} \rangle = \lambda\langle \mathbf{x}, \mathbf{y} \rangle$ khi $\lambda = 1$. Giả sử $\langle k\mathbf{x}, \mathbf{y} \rangle = k\langle \mathbf{x}, \mathbf{y} \rangle$ với mọi $1 \leq k \leq n$ là một số nguyên dương và mọi phần tử $\mathbf{x}, \mathbf{y} \in V$. Ta có

$$\begin{aligned}
\langle (n+1)\mathbf{x}, \mathbf{y} \rangle &= \frac{1}{4}(\|(n+1)\mathbf{x} + \mathbf{y}\|^2 - \|(n+1)\mathbf{x} - \mathbf{y}\|^2) \\
&= \frac{1}{4}\big(2(\|n\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x}\|^2) - \|n\mathbf{x} + \mathbf{y} - \mathbf{x}\|^2 \\
&\qquad - 2(\|n\mathbf{x} - \mathbf{y}\|^2 + \|\mathbf{x}\|^2) + \|n\mathbf{x} - \mathbf{y} - \mathbf{x}\|^2\big) \\
&= \frac{1}{2}(\|n\mathbf{x} + \mathbf{y}\|^2 - \|n\mathbf{x} - \mathbf{y}\|^2) \\
&\qquad - \frac{1}{4}(\|(n-1)\mathbf{x} + \mathbf{y}\|^2 - \|(n-1)\mathbf{x} - \mathbf{y}\|^2) \\
&= 2\langle n\mathbf{x}, \mathbf{y} \rangle - \langle (n-1)\mathbf{x}, \mathbf{y} \rangle \\
&= 2n\langle \mathbf{x}, \mathbf{y} \rangle - (n-1)\langle \mathbf{x}, \mathbf{y} \rangle \\
&= (n+1)\langle \mathbf{x}, \mathbf{y} \rangle.
\end{aligned}$$

Vậy $\langle \alpha\mathbf{x}, \mathbf{y}\rangle = \alpha\langle \mathbf{x}, \mathbf{y}\rangle$ với mọi $\alpha \in \mathbb{N}$. Với $n$ là một số nguyên dương bất kỳ ta có

$$\begin{aligned}\langle -n\mathbf{x}, \mathbf{y}\rangle &= \frac{1}{4}(\| -n\mathbf{x} + \mathbf{y}\|^2 - \| -n\mathbf{x} - \mathbf{y}\|^2)\\ &= -\frac{1}{4}(\|n\mathbf{x} + \mathbf{y}\|^2 - \|n\mathbf{x} - \mathbf{y}\|^2)\\ &= -\langle n\mathbf{x}, \mathbf{y}\rangle = -n\langle \mathbf{x}, \mathbf{y}\rangle.\end{aligned}$$

Vậy $\langle \alpha\mathbf{x}, \mathbf{y}\rangle = \alpha\langle \mathbf{x}, \mathbf{y}\rangle$ với mọi $\alpha \in \mathbb{Z}$. Với $m$ là một số nguyên dương, ta có

$$\begin{aligned}\left\langle \tfrac{1}{m}\mathbf{x}, \mathbf{y}\right\rangle &= \frac{1}{4}(\|\tfrac{1}{m}\mathbf{x} + \mathbf{y}\|^2 - \|\tfrac{1}{m}\mathbf{x} - \mathbf{y}\|^2)\\ &= \frac{1}{4m^2}(\|m\mathbf{y} + \mathbf{x}\|^2 - \|m\mathbf{y} - \mathbf{x}\|^2)\\ &= \frac{1}{m^2}\langle m\mathbf{y}, \mathbf{x}\rangle\\ &= \frac{1}{m}\langle \mathbf{y}, \mathbf{x}\rangle = \frac{1}{m}\langle \mathbf{x}, \mathbf{y}\rangle.\end{aligned}$$

Với $p$ là một số hữu tỉ tùy ý. Khi đó tồn tại số nguyên dương $m$ và số nguyên $n$ sao cho $p = \frac{n}{m}$. Ta có

$$\langle p\mathbf{x}, \mathbf{y}\rangle = \langle \tfrac{n}{m}\mathbf{x}, \mathbf{y}\rangle = n\langle \tfrac{1}{m}\mathbf{x}, \mathbf{y}\rangle = \tfrac{n}{m}\langle \mathbf{x}, \mathbf{y}\rangle = p\langle \mathbf{x}, \mathbf{y}\rangle.$$

Vậy $\langle \alpha\mathbf{x}, \mathbf{y}\rangle = \alpha\langle \mathbf{x}, \mathbf{y}\rangle$ với mọi $\alpha \in \mathbb{Q}$.

Với mọi $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, ta có

$$\begin{aligned}\langle \mathbf{x} + \mathbf{y}, \mathbf{z}\rangle &= \frac{1}{2}\langle 2\mathbf{z}, \mathbf{x} + \mathbf{y}\rangle = \frac{1}{2}\langle \mathbf{x} + \mathbf{y}, 2\mathbf{z}\rangle\\ &= \frac{1}{8}(\|\mathbf{x} + \mathbf{y} + 2\mathbf{z}\|^2 - \|\mathbf{x} + \mathbf{y} - 2\mathbf{z}|^2)\\ &= \frac{1}{8}(2(\|\mathbf{x} + \mathbf{z}\|^2 + \|\mathbf{y} + \mathbf{z}\|^2) - \|\mathbf{x} - \mathbf{y}\|^2\\ &\qquad - 2(\|\mathbf{x} - \mathbf{z}\|^2 + \|\mathbf{y} - \mathbf{z}\|^2) + \|\mathbf{x} - \mathbf{y}\|^2)\\ &= \frac{1}{4}(\|\mathbf{x} + \mathbf{z}\|^2 - \|\mathbf{x} - \mathbf{z}\|^2) + \frac{1}{4}(\|\mathbf{y} + \mathbf{z}\|^2 - \|\mathbf{y} - \mathbf{z}\|^2)\\ &= \langle \mathbf{x}, \mathbf{z}\rangle + \langle \mathbf{y}, \mathbf{z}\rangle.\end{aligned}$$

Với $\alpha \in \mathbb{R}$ và $p \in \mathbb{Q}$, ta có

$$\begin{aligned}|\langle \alpha\mathbf{x},\mathbf{y}\rangle - \alpha\langle \mathbf{x},\mathbf{y}\rangle| &\leq |\langle \alpha\mathbf{x},\mathbf{y}\rangle - \langle p\mathbf{x},\mathbf{y}\rangle| + |p\langle \mathbf{x},\mathbf{y}\rangle - \alpha\langle \mathbf{x},\mathbf{y}\rangle \\ &= |\langle (\alpha-p)\mathbf{x},\mathbf{y}\rangle| + |\alpha-p|\cdot|\langle \mathbf{x},\mathbf{y}\rangle| \\ &= \frac{1}{4}\big|\|(\alpha-p)\mathbf{x}+\mathbf{y}\|^2 - \|(\alpha-p)\mathbf{x}-\mathbf{y}\|^2\big| \\ &\qquad\qquad + |\alpha-p|\cdot|\langle \mathbf{x},\mathbf{y}\rangle|\end{aligned}$$

Mặt khác, ta biến đổi được $\big|\|(\alpha-p)\mathbf{x}+\mathbf{y}\|^2 - \|(\alpha-p)\mathbf{x}-\mathbf{y}\|^2\big| = \big|\|(\alpha-p)\mathbf{x}+\mathbf{y}\| - \|(\alpha-p)\mathbf{x}-\mathbf{y}\|\big|\big(\|(\alpha-p)\mathbf{x}+\mathbf{y}\| + \|(\alpha-p)\mathbf{x}-\mathbf{y}\|\big) \leq \|2(\alpha-p)\mathbf{x}\|(2\|(\alpha-p)\mathbf{x}\|+2\|\mathbf{y}\|) = 4|\alpha-p|\|\mathbf{x}\|(|\alpha-p|\|\mathbf{x}\|+\|\mathbf{y}\|)$. Từ đó với mọi $p \in \mathbb{Q}$ ta có được

$$|\langle \alpha\mathbf{x},\mathbf{y}\rangle - \alpha\langle \mathbf{x},\mathbf{y}\rangle| \leq |\alpha-p|\big(\|\mathbf{x}\|(|\alpha-p|\|\mathbf{x}\|+\|\mathbf{y}\|) + |\langle \mathbf{x},\mathbf{y}\rangle|\big)$$

Ta nhận thấy vế phải dần về 0 khi $p \to \alpha$. Vậy do tính trù mật của $\mathbb{Q}$ ta suy ra được $\langle \alpha\mathbf{x},\mathbf{y}\rangle = \alpha\langle \mathbf{x},\mathbf{y}\rangle$ với mọi $\alpha \in \mathbb{R}$.

Cuối cùng ta có $\langle \mathbf{x},\mathbf{x}\rangle = \frac{1}{4}(\|\mathbf{x}+\mathbf{x}\|^2 - \|\mathbf{x}-\mathbf{x}\|^2) = \|\mathbf{x}\|^2 \geq 0$ với mọi $\mathbf{x} \in V$. Hơn nữa, $\langle \mathbf{x},\mathbf{x}\rangle = 0$ khi và chỉ khi $\|\mathbf{x}\| = 0$ hay $\mathbf{x} = \mathbf{0}$.

Vậy $\langle , \rangle$ là một tích vô hướng trên $V$, và ta cũng thấy độ dài của $\mathbf{x}$ sinh bởi tích vô hướng này là $\|\mathbf{x}\|$. $\square$

**1.2.6 Định nghĩa.** Cho $V_E$ là một không gian Euclid. **Góc** giữa hai vector khác không $\mathbf{x},\mathbf{y} \in V_E$, ký hiệu $(\widehat{\mathbf{x},\mathbf{y}})$, là góc có số đo (radian) thuộc $[0,\pi]$ sao cho

(1.2.7) $$\cos(\widehat{\mathbf{x},\mathbf{y}}) = \frac{\langle \mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|}.$$

**1.2.8 Mệnh đề.** *Cho không gian Euclid $V_E$ và các vector khác không $\mathbf{x},\mathbf{y} \in V_E$. Ta có*

**(i)** $(\widehat{\mathbf{x},\mathbf{y}}) = (\widehat{\mathbf{y},\mathbf{x}})$.

**(ii)** $(\widehat{\lambda\mathbf{x},\mu\mathbf{y}}) = \begin{cases} (\widehat{\mathbf{x},\mathbf{y}}) & \textit{nếu } \lambda\mu > 0 \\ \pi - (\widehat{\mathbf{x},\mathbf{y}}) & \textit{nếu } \lambda\mu < 0 \end{cases}$

**(iii)** $(\widehat{\mathbf{x},\mathbf{y}}) = 0$ *khi và chỉ khi* $\mathbf{y} = \lambda\mathbf{x}$ *với* $\lambda > 0$.

**(iv)** $(\widehat{\mathbf{x},\mathbf{y}}) = \pi$ *khi và chỉ khi* $\mathbf{y} = \lambda\mathbf{x}$ *với* $\lambda < 0$.

*Chứng minh.* Các tính chất trên được suy ra trực tiếp từ định nghĩa góc giữa hai vector và tính chất hàm cosine. □

**1.2.9 Định nghĩa.** Cho $V_E$ là một không gian Euclid hữu hạn chiều. Toán tử tuyến tính $f : V_E \to V_E$ được gọi là **bảo toàn góc** nếu $f$ là song ánh và với mọi $\mathbf{x}, \mathbf{y}$ khác $\mathbf{0}$ ta có $(\widehat{f(\mathbf{x}), f(\mathbf{y})}) = (\widehat{\mathbf{x},\mathbf{y}})$.

## Bài tập

**1)** Xét các ánh xạ $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ với $(\mathbf{x},\mathbf{y}) \mapsto \langle \mathbf{x},\mathbf{y}\rangle$ như sau

**(a)** $\langle \mathbf{x},\mathbf{y}\rangle = x_1y_1 + x_2y_2$

**(b)** $\langle \mathbf{x},\mathbf{y}\rangle = 3x_1y_2 + 3x_2y_1$

**(c)** $\langle \mathbf{x},\mathbf{y}\rangle = 3x_1y_2 + 5x_2y_2$

**(d)** $\langle \mathbf{x},\mathbf{y}\rangle = 2x_1^2 + y_1^2 + x_2^2 + 2y_2^2$

trong đó $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$. Ánh xạ nào trong chúng là tích vô hướng trên $\mathbb{R}^2$.

**2)** Chứng minh ánh xạ $\mathbb{R}_2[x] \times \mathbb{R}_2[x] \to \mathbb{R}$ xác định bởi

$$(a_0 + a_1x + a_2x^2, b_0 + b_1x + b_2x^2) \mapsto a_0b_0 + a_1b_1 + a_2b_2$$

là một tích vô hướng trên $\mathbb{R}_2[x]$.

**3)** Xác định $m$ để công thức sau đây là một tích vô hướng trên $\mathbb{R}^2$

$$\mathbf{x}\cdot\mathbf{y} = x_1y_1 + 2x_1y_2 + 2x_2y_1 + mx_2y_2,$$

trong đó $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$.

**4)** Cho $\langle\cdot,\cdot\rangle$ là một tích vô hướng trên $V$. Chứng minh rằng $\langle\cdot,\cdot\rangle'$ xác định bởi $\langle \mathbf{x},\mathbf{y}\rangle' = r\langle \mathbf{x},\mathbf{y}\rangle$ là một tích vô hướng trên $V$ với $r > 0$.

**5)** Cho $\langle \cdot, \cdot \rangle_1$ và $\langle \cdot, \cdot \rangle_2$ là hai tích vô hướng trên không gian vector $V$. Chứng minh rằng $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_1 + \langle \cdot, \cdot \rangle_2$ là một tích vô hướng trên $V$.

**6)** Chứng minh rằng $\langle A, B \rangle = \mathrm{tr}(B^t A)$ với $A, B \in \mathcal{M}_n(\mathbb{R})$ là một tích vô hướng trên $\mathcal{M}_n(\mathbb{R})$.

**7)** Chứng tỏ rằng các công thức sau không phải là một tích vô hướng trên không gian vector đã cho

**(a)** $\langle (a, b), (c, d) \rangle = ac - bd$ trên $\mathbb{R}^2$.

**(b)** $\langle A, B \rangle = \mathrm{tr}(A + B)$ trên $\mathcal{M}_2(\mathbb{R})$.

**(c)** $\langle f, g \rangle = \displaystyle\int_0^1 f'(t)g(t)dt$ trên $\mathcal{P}(\mathbb{R})$ các đa thức trên $\mathbb{R}$, trong đó $f'(t)$ là đạo hàm của $f(t)$.

**8)** Cho $f : \mathbb{R}^n \to \mathbb{R}^n$ là một ánh xạ tuyến tính. Chứng minh rằng $\|f(\mathbf{x})\| = \|\mathbf{x}\|$ với mọi $\mathbf{x} \in \mathbb{R}^n$ khi và chỉ khi $\langle f(\mathbf{x}), f(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$. Khi đó ta có $f$ là song ánh và $f^{-1}$ cũng có tính chất như thế.

**9)** Cho $V$ là một không gian vector và cho $W$ là một không gian Euclid với tích vô hương $\langle \cdot, \cdot \rangle$. Nêu $T : V \to W$ là ánh xạ tuyến tính, chứng minh rằng $\langle \mathbf{x}, \mathbf{y} \rangle' = \langle T(\mathbf{x}), T(\mathbf{y}) \rangle$ định nghĩa một tích vô hướng trên $V$ nếu và chỉ nếu $T$ là đơn ánh.

**10)** Cho $\mathscr{B}$ là một cơ sở của một không gian Euclid hữu hạn chiều. Chứng minh rằng nếu $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ với mọi $\mathbf{x} \in \mathscr{B}$ thì $\mathbf{y} = \mathbf{0}$.

**11)** Trong không gian Eulid $\mathbb{R}^n$ với tích vô hướng chính tắc, tìm độ dài của các vector **x**, **y**; tích vô hướng của **x**, **y**; và góc giữa chúng.

**(a)** $\mathbf{x} = (1, -2, 3), \quad \mathbf{y} = (3, 0, -4); \quad n = 3$

**(b)** $\mathbf{x} = (1, -3, 0, 2), \quad \mathbf{y} = (0, 5, -12, 0); \quad n = 4$

**(c)** $\mathbf{x} = (4, -2, 1, 2), \quad \mathbf{y} = (1, 3, -5, -1); \quad n = 4$

**(d)** $\mathbf{x} = (1, 1, \ldots, 1)$, $\mathbf{y} = (\alpha, \alpha, \ldots, \alpha)$, $\alpha \in \mathbb{R} \setminus \{0\}$; $n$ tùy ý.

**12)** Nếu $0 \leq \theta < \pi$ và ánh xạ tuyến tính $f : \mathbb{R}^2 \to \mathbb{R}^2$ có ma trận $\begin{bmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{bmatrix}$ đối với cơ sở chính tắc. Chứng minh rằng $f$ là bảo toàn góc, và nếu $\mathbf{x} \neq \mathbf{0}$ hãy tính $(\widehat{\mathbf{x}, f(\mathbf{x})})$.

**13)** Giả sử $f : \mathbb{R}^n \to \mathbb{R}^n$ là một ánh xạ tuyến tính sao cho tồn tại cơ sở $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ của $\mathbb{R}^n$ thỏa $f(\mathbf{b}_i) = \lambda_i \mathbf{b}_i$ với mọi $i = 1, 2, \ldots, n$. Chứng minh rằng $f$ là bảo toàn góc khi và chỉ khi tất cả các $\lambda_i$ bằng nhau.

**14)** Tìm điều kiện để ánh xạ tuyến tính $f : \mathbb{R}^2 \to \mathbb{R}^2$ bảo toàn góc.

**15)** Chứng minh rằng trong $\mathbb{R}^2$, $\|(a, b)\| = \max\{|a|, |b|\}$ thỏa 3 điều kiện đầu của Định lý 1.2.5; tuy nhiên không có một tích vô hướng $\langle \cdot, \cdot \rangle$ nào trên $\mathbb{R}^n$ thỏa $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle$ với mọi $\mathbf{x} \in \mathbb{R}^2$.

# § 2 Trực giao

## 2.1 Khái niệm trực giao và tính chất

**2.1.1 Định nghĩa.** Cho không gian Euclid $V_E$. Hai vector $\mathbf{x}, \mathbf{y} \in V_E$ được gọi là **trực giao với nhau**, kí hiệu $\mathbf{x} \perp \mathbf{y}$, nếu $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Ta nói vector $\mathbf{a}$ trực giao với tập $U \subseteq V_E$ nếu $\mathbf{a} \perp \mathbf{x}$ với mọi $\mathbf{x} \in U$.

**2.1.2 Mệnh đề. (Tính chất cơ bản)** *Cho không gian Euclid $V_E$ và $\mathbf{x}, \mathbf{y} \in V_E$. Ta có*

**(i)** $\mathbf{0} \perp \mathbf{x}$.

**(ii)** $\mathbf{x} \perp \mathbf{x}$ *khi và chỉ khi* $\mathbf{x} = \mathbf{0}$.

**(iii)** *(Định lý Pythagore)* $\mathbf{x} \perp \mathbf{y}$ *khi và chỉ khi* $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

**(iv)** *Với* $\mathbf{x}, \mathbf{y}$ *khác vector không,* $(\widehat{\mathbf{x}, \mathbf{y}}) = \dfrac{\pi}{2}$ *khi và chỉ khi* $\mathbf{x} \perp \mathbf{y}$.

*Chứng minh.* Các tính chất trên được suy ra trực tiếp từ định nghĩa, ta sẽ chứng minh tính chất (iii). Ta có

$$\|\mathbf{x}+\mathbf{y}\|^2 = \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{x}\rangle + 2\langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle.$$

Vậy $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ khi và chỉ khi $\langle \mathbf{x}, \mathbf{y}\rangle = 0$ hay $\mathbf{x} \perp \mathbf{y}$. □

**2.1.3 Định lý.** *Trong không gian Euclid $V_E$, cho $U = \langle \mathbf{u}_1, \ldots, \mathbf{u}_n\rangle$. Khi đó, $\mathbf{x} \perp U$ khi và chỉ khi $\mathbf{x} \perp \mathbf{u}_i$ với mọi $i = 1, \ldots, n$.*

*Chứng minh.* Nếu $\mathbf{x} \perp U$, thì rõ ràng $\mathbf{x} \perp \mathbf{u}_i$ với mọi $i = 1, \ldots, n$. Ngược lại, giả sử $\mathbf{x} \perp \mathbf{u}_i$ với mọi $i = 1, \ldots, n$. Khi đó, với mọi $\mathbf{u} \in U$ ta có biểu diễn $\mathbf{u} = \sum_{i=1}^n \alpha_i \mathbf{u}_i$. Suy ra

$$\langle \mathbf{x}, \mathbf{u}\rangle = \langle \mathbf{x}, \sum_{i=1}^n \alpha_i \mathbf{u}_i\rangle = \sum_{i=1}^n \alpha_i \langle \mathbf{x}, \mathbf{u}_i\rangle = 0,$$

hay $\mathbf{x} \perp \mathbf{u}$. Do đó $\mathbf{x} \perp U$. □

## 2.2 Không gian con trực giao

**2.2.1 Định nghĩa.** Cho $V_E$ là một không gian Euclid và $U$, $W$ là hai không gian con của $V_E$. $U$ và $W$ gọi là **trực giao với nhau**, ký hiệu $U \perp W$, nếu $\mathbf{u} \perp \mathbf{w}$ với mọi $\mathbf{u} \in U$ và $\mathbf{w} \in W$ (nghĩa là với mọi $\mathbf{u} \in U$ ta có $\mathbf{u} \perp W$ và với mọi $\mathbf{w} \in W$ ta có $\mathbf{w} \perp U$).

**2.2.2 Thí dụ.** Xét không gian Euclid chính tắc $\mathbb{R}^3$. Đặt

$$W = \langle (1,0,0)\rangle = \{(x,0,0) : x \in \mathbb{R}\};$$
$$Z = \langle (0,1,2)\rangle = \{(0,y,2y) : y \in \mathbb{R}\}.$$

Ta thấy $W \perp Z$ vì $\langle (x,0,0), (0,y,2y)\rangle = 0$. □

**2.2.3 Mệnh đề.** *Với các không gian con $U, W$ tùy ý trong $V_E$ ta có*

**(i)** $\{\mathbf{0}\} \perp U$.

**(ii)** *Nếu $U \perp W$, thì $U \cap W = \{\mathbf{0}\}$.*

**(iii)** *Nếu $V_E$ là không gian Euclid hữu hạn chiều và $U \perp W$, thì* $\dim U + \dim W \leq \dim V_E$.

*Chứng minh.*

**(i)** Hiển nhiên theo định nghĩa các không gian con trực giao.

**(ii)** Với $\mathbf{x} \in U \cap W$, ta có $\mathbf{x} \perp \mathbf{x}$ vì $\mathbf{x} \in U$, $\mathbf{x} \in W$, và $U \perp W$. Do đó, $\mathbf{x} = \mathbf{0}$. Vậy $U \cap W = \{\mathbf{0}\}$.

**(iii)** Do $U$ và $W$ là các không gian con của $V_E$ nên $U + W$ cũng là không gian con của $V_E$. Hơn nữa, do $U \cap W = \{\mathbf{0}\}$, ta có $\dim U + \dim W = \dim(U + W) \leq \dim V_E$. □

**2.2.4 Định lý.** *Giả sử $U$ và $V$ là hai không gian con của không gian Euclid hữu hạn chiều, $\mathscr{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_m)$ và $\mathscr{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ lần lượt là hai hệ sinh của $U$ và $V$. Khi đó, $U$ và $V$ trực giao với nhau khi và chỉ khi $\mathbf{a}_i \perp \mathbf{b}_j$ với mọi $i = 1, \ldots, m$ và $j = 1, \ldots, n$.*

*Chứng minh.* Nếu $U \perp V$, thì theo định nghĩa của các không gian con trực giao rõ ràng ta có $\mathbf{a}_i \perp \mathbf{b}_j$ với mọi $i = 1, \ldots, m$ và $j = 1, \ldots, n$.

Ngược lại, giả sử $\mathbf{a}_i \perp \mathbf{b}_j$ với mọi $i = 1, \ldots, m$ và $j = 1, \ldots, n$. Lấy $\mathbf{u} \in U$ và $\mathbf{v} \in V$ tùy ý. Ta có biểu diễn

$$\mathbf{u} = \sum_{i=1}^{m} \alpha_i \mathbf{a}_i \qquad \text{và} \qquad \mathbf{v} = \sum_{j=1}^{n} \beta_j \mathbf{b}_j.$$

Do đó

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j \langle \mathbf{a}_i, \mathbf{b}_j \rangle = 0.$$

Vậy $U \perp V$. □

## 2.3 Hệ vector trực giao và trực chuẩn

**2.3.1 Định nghĩa.** Giả sử $V_E$ là một không gian Euclid. Hệ vector $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ trong $V_E$ gọi là **hệ trực giao** nếu các vector của hệ khác không và đôi một trực giao nhau.

**2.3.2 Thí dụ.** Hệ vector $((1,1,1),(-2,1,1),(0,-1,1))$ là hệ trực giao trong $\mathbb{R}^3$. $\square$

**2.3.3 Định nghĩa.** Hệ vector $(\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n)$ trong không gian Euclid $V_E$ gọi là **hệ trực chuẩn** nếu nó là hệ trực giao và mọi vector của nó đều là vector đơn vị (có chiều dài bằng 1).

**2.3.4 Thí dụ.** Hệ vector $((\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}),(-\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}}),(0,-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}))$ là hệ trực chuẩn trong $\mathbb{R}^3$. $\square$

**2.3.5 Thí dụ.** Họ $(\cos x,\cos 2x,\ldots,\cos nx)$ là trực giao trong $C_{[0,\pi]}$.

Thật vậy, với $k\neq l$ thuộc $\{1,2,\ldots,n\}$ ta có

$$\int_0^\pi \cos(kx)\cos(lx)=\frac{1}{2}\int_0^\pi[\cos(k+l)x-\cos(k-l)x]dx=0. \quad \square$$

**2.3.6 Định lý.** *Mọi hệ trực giao (nói riêng, mọi hệ trực chuẩn) trong một không gian Euclid đều độc lập tuyến tính.*

*Chứng minh.* Giả sử $(\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n)$ là một hệ trực giao trong không gian Euclid. Xét tổ hợp tuyến tính

$$\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2+\cdots+\alpha_n\mathbf{u}_n=\mathbf{0}.$$

Với mỗi $i=1,2,\ldots,n$, ta có

$$0=\langle\mathbf{0},\mathbf{u}_i\rangle=\langle\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2+\cdots+\alpha_n\mathbf{u}_n,\mathbf{u}_i\rangle=\alpha_i\langle\mathbf{u}_i,\mathbf{u}_i\rangle,$$

suy ra $\alpha_i=0$. Vậy hệ vector $(\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n)$ độc lập tuyến tính. $\square$

## Bài tập

**1**) Trong không gian Euclid $\mathbb{R}^3$ cho $U$ và $V$ lần lượt là không gian nghiệm của các hệ phương trình tuyến tính thuần nhất

$$\begin{cases}x_1+x_2-x_3=0\\3x_1+x_2-x_3=0\end{cases}\qquad\begin{cases}3x_1+2x_2-2x_3=0\\-6x_1-4x_2+4x_3=0\end{cases}$$

Hai không gian con $U$ và $V$ có trực giao nhau không?

**2)** Trong không gian Euclid $\mathbb{R}^4$ cho $U$ và $V$ lần lượt là không gian nghiệm của các hệ phương trình tuyến tính thuần nhất

$$\begin{cases} x_1 - 2x_2 - 2x_3 - x_4 = 0 \\ 2x_1 + 3x_2 + x_3 + 2x_4 = 0 \end{cases} \qquad \begin{cases} x_1 + x_2 + x_3 + 3x_4 \\ x_1 + 4x_2 + 7x_4 = 0 \end{cases}$$

Hai không gian con $U$ và $V$ có trực giao nhau không?

**3)** Trong không gian Euclid $\mathbb{R}^4$ cho các không gian con

$$U = \langle (1,1,1,-2), (3,-1,5,0) \rangle \qquad V = \langle (1,3,0,2), (0,5,1,m) \rangle.$$

Xác định $m$ để $U \perp V$.

**4)** Chứng minh rằng $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots\}$ là một hệ trực giao trong $\mathcal{C}_{[-\pi,\pi]}$

**5)** Cho $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ là một hệ trực giao trong một không gian Euclid và cho $a_1, a_2, \ldots, a_k$ là các vô hướng. Chứng minh rằng

$$\Big\| \sum_{i=1}^{k} a_i \mathbf{v}_i \Big\| = \sum_{i=1}^{k} |a_i| \cdot \|\mathbf{v}_i\|^2.$$

**6)** Cho $V$ là một không gian vector (có thể vô hạn chiều) và cho $\mathscr{B}$ là một cơ sở của $V$. Với $\mathbf{x}, \mathbf{y} \in V$ tồn tại $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n \in \mathscr{B}$ sao cho

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{b}_i \qquad \text{và} \qquad \mathbf{y} = \sum_{i=1}^{n} y_i \mathbf{b}_i.$$

Định nghĩa

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i b_i$$

Chứng minh rằng $\langle \cdot, \cdot \rangle$ là một tích vô hướng trên $V$ và $\mathscr{B}$ là một cơ sở trực chuẩn của $V$. Do đó, mỗi không gian vector có thể được xem như là một không gian vector Euclid.

# § 3 Cơ sở trực giao - Cơ sở trực chuẩn

## 3.1 Khái niệm và tính chất

**3.1.1 Định nghĩa.** Cho $V_E$ là một không gian Euclid $n$ chiều. Cơ sở $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ của $V_E$ gọi là **cơ sở trực giao** (tương ứng, **cơ sở trực chuẩn**) nếu bản thân hệ $\mathscr{B}$ là một hệ trực giao (tương ứng, trực chuẩn).

**3.1.2 Thí dụ.** Trong không gian Euclid $\mathbb{R}^n$ cơ sở chính tắc $\mathscr{E}_n = (\mathbf{e}_1 = (1, 0, \ldots, 0), \mathbf{e}_2 = (0, 1, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, \ldots, 1))$ là một cơ sở trực chuẩn, và nó được gọi là cơ sở trực chuẩn chính tắc của $\mathbb{R}^n$. □

**3.1.3 Định lý.** *Mọi không gian Euclid $n$ chiều đều có cơ sở trực chuẩn.*

*Chứng minh.* Đó chính là thuật toán Gram-Schmidt ở phân mục sau. □

**3.1.4 Mệnh đề.** *Nếu $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là một cơ sở trực chuẩn của không gian Euclid $n$ chiều $V_E$, thì với mọi cặp vector $\mathbf{x}, \mathbf{y} \in V_E$ mà $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$ và $\mathbf{y}|_{\mathscr{B}} = (y_1, y_2, \ldots, y_n)$ ta đều có*

**(i)** $x_i = \langle \mathbf{x}, \mathbf{e}_i \rangle$ *với mọi* $i = 1, 2, \ldots, n$.

**(ii)** $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$.

**(iii)** $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}$

**(iv)** *Với* $\mathbf{x} \neq \mathbf{0}$ *và* $\mathbf{y} \neq \mathbf{0}$, *ta có* $\cos(\widehat{\mathbf{x}, \mathbf{y}}) = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}}$.

*Chứng minh.*

**(i)** Theo giả thiết ta có $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$. Do đó, với mỗi $i = 1, 2, \ldots, n$, ta có

$$\langle \mathbf{x}, \mathbf{e}_i \rangle = \langle x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n, \mathbf{e}_i \rangle = x_i \langle \mathbf{e}_i, \mathbf{e}_i \rangle = x_i$$

**(ii)** Tương tự như trên ta có

$$\langle \mathbf{x}, \mathbf{y} \rangle = \Big\langle \sum_{i=1}^{n} x_i \mathbf{e}_i, \sum_{j=1}^{n} y_j \mathbf{e}_j \Big\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \sum_{i=1}^{n} x_i y_i.$$

**(iii)** Suy ra từ định nghĩa độ dài và (ii).

**(iv)** Suy ra trực tiếp từ định nghĩa của cosine và tính chất (ii), (iii). □

**3.1.5 Định lý.** *Giả sử $\mathscr{A}$ và $\mathscr{B}$ là hai cơ sở trực chuẩn của không gian Euclid $n$ chiều $V_E$ và $A$ là ma trận đổi cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$. Khi đó, $A^t = A^{-1}$, tức là*

$$A^t A = A A^t = I_n.$$

*Ma trận vuông cấp $n$ thực $A$ có tính chất như trên gọi là* ***ma trận trực giao****.*

*Chứng minh.* Giả sử $\mathscr{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ và $\mathscr{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$, và $A = (a_{ij})_n$. Theo giả thiết ta có $\mathbf{b}_j = \sum\limits_{i=1}^{n} a_{ij} \mathbf{a}_i$ và $a_{ij} = \langle \mathbf{a}_i, \mathbf{b}_j \rangle$. Ta có

$$(A^t A)_{ij} = \sum_{l=1}^{k} a_{li} a_{lj} = \langle \mathbf{b}_i, \mathbf{b}_j \rangle = \delta_{ij}.$$

Vậy $A^t A = I_n$. Mặt khác, ta có $A$ là ma trận không suy biến nên tồn tại ma trận nghịch đảo $A^{-1}$. Suy ra $(A^t A) A^{-1} = I_n A^{-1}$, cho nên $A^t = A^{-1}$. □

**3.1.6 Định lý.** *Cho $\mathscr{A}$ là cơ sở trực chuẩn và $\mathscr{B}$ là một cơ sở tùy ý trong không gian Euclid $V_E$ và $A$ là ma trận chuyển cơ sở từ $\mathscr{A}$ sang $\mathscr{B}$. Khi đó, $\mathscr{B}$ là cơ sở trực chuẩn khi và chỉ khi $A$ là ma trận trực giao.*

*Chứng minh.* Theo định lý trên nếu $\mathscr{B}$ là một cơ sở trực chuẩn thì $A$ là ma trận trực giao.

Giả sử $A = (a_{ij})_n$ là ma trận trực giao, và giả sử $\mathscr{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ và $\mathscr{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Khi đó, ta có $\mathbf{b}_i = \sum\limits_{l=1}^{n} a_{li} \mathbf{a}_l$. Do đó,

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = \sum_{l=1}^{n} a_{li} a_{lj} = (A^t A)_{ij} = \delta_{ij}.$$

Vậy hệ $\mathscr{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ là một cơ sở trực giao. $\square$

## 3.2 Thuật toán Gram-Schmidt

**3.2.1 Nội dung thuật toán** Cho $(\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_n)$ là một hệ vector độc lập tuyến tính tùy ý $(n \geq 1)$ trong không gian Euclid $V_E$.

**(1)** Đặt $\mathbf{v}_1 = \mathbf{u}_1$ và $\mathbf{w}_1 = \dfrac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1$.

**(2)** Nếu $n > 1$, thì tiếp tục đặt

$$\mathbf{v}_j = \mathbf{u}_j - \sum_{i=1}^{j-1} \langle \mathbf{u}_j, \mathbf{w}_i \rangle \mathbf{w}_i, \qquad \mathbf{w}_j = \frac{1}{\|\mathbf{v}_j\|}\mathbf{v}_j$$

với $j = 2, \ldots, n$.

Khi đó, $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n)$ là một hệ trực chuẩn; hơn nữa, $\mathbf{w}_j$ thuộc không gian con $\langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_j \rangle$ với $j = 1, \ldots, n$.

*Chứng minh.* Rõ ràng hệ vector $(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n)$ là hệ vector đơn vị. Ta sẽ chứng minh bằng quy nạp hệ vector $(\mathbf{w}_1, \ldots, \mathbf{w}_k)$ là trực chuẩn và $\mathbf{w}_k \in \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \rangle$ với mọi $k = 1, \ldots, n$. Rõ ràng hệ vector $(\mathbf{w}_1)$ là trực chuẩn và $\mathbf{w}_1 \in \langle \mathbf{u}_1 \rangle$. Giả sử hệ vector $(\mathbf{w}_1, \ldots, \mathbf{w}_{k-1})$ là trực chuẩn với $1 < k \leq n$ và $\mathbf{w}_j \in \langle \mathbf{u}_1, \ldots, \mathbf{u}_j \rangle$ với mọi $j = 1, \ldots, k-1$. Với mỗi $i = 1, \ldots, k-1$, ta có

$$\begin{aligned}
\langle \mathbf{w}_i, \mathbf{w}_k \rangle &= \left\langle \mathbf{w}_i, \frac{1}{\|\mathbf{v}_k\|}\Big(\mathbf{u}_k - \sum_{j=1}^{k-1} \langle \mathbf{u}_k, \mathbf{w}_j \rangle \mathbf{w}_j\Big)\right\rangle \\
&= \frac{1}{\|\mathbf{v}_k\|}\Big(\langle \mathbf{w}_i, \mathbf{u}_k \rangle - \sum_{j=1}^{k-1} \langle \mathbf{u}_k, \mathbf{w}_j \rangle \langle \mathbf{w}_i, \mathbf{w}_j \rangle\Big) \\
&= \frac{1}{\|\mathbf{v}_k\|}(\langle \mathbf{w}_i, \mathbf{u}_k \rangle - \langle \mathbf{u}_k, \mathbf{w}_i \rangle) = 0.
\end{aligned}$$

Vậy $(\mathbf{w}_1, \ldots, \mathbf{w}_k)$ là hệ trực chuẩn. Mặt khác, từ định nghĩa $\mathbf{w}_k$ ta thấy $\mathbf{w}_k \in \langle \mathbf{w}_1, \ldots, \mathbf{w}_{k-1}, \mathbf{u}_k \rangle$. Do $\mathbf{w}_j \in \langle \mathbf{u}_1, \ldots, \mathbf{u}_j \rangle$ với mọi $j = 1, \ldots, k-1$ nên suy ra được $\mathbf{w}_k \in \langle \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \rangle$. $\square$

**3.2.2 Thí dụ.** Trực chuẩn hóa hệ độc lập tuyến tính $((1,1,0),(0,1,1))$ trong $\mathbb{R}^3$. Ta có

$$\mathbf{w}_1 = (\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0) \quad \mathbf{v}_2 = (0,1,1) - \tfrac{1}{\sqrt{2}}(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0) = (-\tfrac{1}{2}, \tfrac{1}{2}, 1).$$

Suy ra $\mathbf{w}_2 = (-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}})$. Do đó, hệ trực chuẩn cần tìm là $((\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}))$. □

**3.2.3 Thí dụ.** Trực chuẩn hóa cơ sở $((0,0,1),(1,1,1),(0,1,1))$ trong $\mathbb{R}^3$. Ta có

$$\mathbf{w}_1 = (0,0,1) \qquad \mathbf{v}_2 = (1,1,1) - 1(0,0,1) = (1,1,0).$$

Do đó, $\mathbf{w}_2 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$. Ta tiếp tục

$$\mathbf{v}_3 = (0,1,1) - 1(0,0,1) - \tfrac{1}{\sqrt{2}}(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0) = (-\tfrac{1}{2}, \tfrac{1}{2}, 0).$$

Suy ra $\mathbf{w}_3 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$. Vậy cơ sở trực chuẩn cần tìm là $((0,1,1)$, $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$, $(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0))$. □

**3.2.4 Mệnh đề.** *Cho $V_E$ là một không gian Euclid $n$ chiều. Mọi hệ trực chuẩn trong $V_E$ đều có thể bổ sung thành một cơ sở trực chuẩn của $V_E$.*

*Chứng minh.* Giả sử $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ là một hệ trực chuẩn trong $V_E$. Do đó, hệ ấy là độc lập tuyến tính. Vậy ta có thể bổ sung vào hệ ấy $n-m$ vector để được một cơ sở của $V_E$ là $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m, \mathbf{b}_{m+1}, \ldots, \mathbf{b}_n)$. Ta áp dụng thuật toán Gram-Schmidt bắt đầu từ hệ trực chuẩn $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ để thu được hệ trực chuẩn $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m, \mathbf{a}_{m+1}, \ldots, \mathbf{a}_n)$, và hệ này là một cơ sở của $V_E$. □

## Bài tập

**1**) Cho $U_1$ và $U_2$ là hai không gian con của không gian Euclid $n$ chiều $V_E$ với $\dim U_1 < \dim U_2$. Chứng minh rằng trong $U_2$ luôn tìm được một vector khác không trực giao với $U_1$.

**2**) Trong không gian $\mathbb{E}^4$ với tích vô hướng chính tắc chứng tỏ các hệ vector sau đây trực giao và bổ sung vào chúng để được các cơ sở trực giao của $\mathbb{E}^4$.

**(a)** $((1,-2,0,1),(0,1,0,2))$ **(b)** $((2,2,1,0),(2,-3,2,4))$

**(c)** $((3,-1,1,1),(-1,-1,1,1)$ **(d)** $((1,2,3,4),(0,1,-2,1))$

**(e)** $((1,1,1,2),(1,2,3,-3))$ **(f)** $((1,1,1,1),(1,-1,1,-1))$

**3**) Cho $V_E$ là không Euclid $n$ chiều với cơ sở $\mathscr{B}$. Chứng minh rằng $\mathscr{B}$ là cơ sở trực chuẩn của $V_E$ khi và chỉ khi với mọi $\mathbf{x},\mathbf{y}\in V_E$ ta đều có $\langle \mathbf{x},\mathbf{y}\rangle = \sum\limits_{i=1}^{n} x_i y_i$ với $\mathbf{x}|_{\mathscr{B}} = (x_1, x_2, \ldots, x_n)$ và $\mathbf{y}|_{\mathscr{B}} = (y_1, y_2, \ldots, y_n)$.

**4**) Trong các ma trận sau đây, ma trận nào là ma trận trực giao? Tìm ma trận nghịch đảo trong trường hợp ma trận là ma trận trực giao.

**(a)** $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 2 & -1 & 3 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & 0 \\ -2 & 1 & 1 \end{bmatrix}$ **(c)** $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{11}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{11}} & 0 \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

**(d)** $\begin{bmatrix} \frac{12}{13} & \frac{-4}{5} & \frac{12}{13} \\ \frac{3}{13} & 0 & \frac{-5}{13} \\ \frac{4}{13} & \frac{3}{5} & 0 \end{bmatrix}$ **(e)** $\begin{bmatrix} \frac{12}{13} & \frac{4}{13} & \frac{-3}{13} \\ \frac{-3}{13} & \frac{12}{13} & \frac{4}{13} \\ \frac{4}{13} & \frac{-3}{13} & \frac{12}{13} \end{bmatrix}$ **(f)** $\begin{bmatrix} \frac{1}{\sqrt{15}} & \frac{2}{\sqrt{6}} & \frac{2}{\sqrt{30}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{30}} \\ 0 & \frac{1}{\sqrt{6}} & \frac{-5}{\sqrt{30}} \end{bmatrix}$

**5**) Chứng minh mọi ma trận trực giao đều có định thức bằng $1$ hay $-1$.

**6**) Dùng thuật toán Gram-Schmidt trực chuẩn hóa hệ vector sau đây

**(a)** $((1,2,3),(0,2,0),(0,0,3))$

**(b)** $((1,0,0),(0,1,-1),(1,1,1))$

**(c)** $((1,2,2,-1),(1,1,-5,3),(3,2,8,-7))$

**(d)** $((1,1,-1,-2),(5,8,-2,-3),(3,9,3,8))$

**(e)** $((1,1,0,0),(0,0,1,1),(1,0,1,1),(0,1,0,-1))$

**(f)** $((1,0,-1,0),(0,1,0,0),(1,-1,0,0),(-1,0,0,1))$

**(g)** $((2,1,3,-1),(7,4,3,-3),(1,1,-6,0),(5,7,7,8))$

**7)** Xét không gian vector $C_{[-1,1]}$ (tập các hàm liên tục trên $[-1,1]$) với tích vô hướng xác định bởi

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx.$$

Tìm một cơ sở trực chuẩn của không gian con sinh bởi các hàm 1, $x$ và $x^2$.

**8)** Chứng minh rằng cơ sở $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ trong không gian Euclid $V$ là trực chuẩn nếu với mọi $\mathbf{x}, \mathbf{y} \in V$ có biểu diễn $\mathbf{x} = \sum\limits_{i=1}^{n} x_i\mathbf{a}_i$, $\mathbf{y} = \sum\limits_{i=1}^{n} y_i\mathbf{a}_i$ ta luôn có $\langle \mathbf{x}, \mathbf{y}\rangle = \sum\limits_{i=1}^{n} x_iy_i$.

# § 4 Phần bù trực giao và phép chiếu trực giao

## 4.1 Phần bù trực giao

**4.1.1 Định lý.** *Cho $U$ là không gian con của không gian Euclid $V_E$. Khi đó tập $W = \{\mathbf{x} \in V_E : \mathbf{x} \perp U\}$ là không gian con của $V_E$ thỏa $U \cap W = \{\mathbf{0}\}$.*

*Chứng minh.* Rõ ràng $\mathbf{0} \in W$ nên $W \neq \varnothing$. Lấy $\mathbf{x}, \mathbf{y} \in W$ tùy ý và $\alpha, \beta \in \mathbb{K}$. Với mỗi $\mathbf{u} \in U$, ta có

$$\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{u}\rangle = \alpha\langle \mathbf{x}, \mathbf{u}\rangle + \beta\langle \mathbf{y}, \mathbf{u}\rangle = 0.$$

Vậy $\alpha\mathbf{x} + \beta\mathbf{y} \perp U$ hay $\alpha\mathbf{x} + \beta\mathbf{y} \in W$. Do đó, $W$ là một không gian con của $V_E$. Hơn nữa, từ định nghĩa tập $W$ ta thấy rằng $W \perp U$, cho nên $U \cap W = \{\mathbf{0}\}$. □

**4.1.2 Định nghĩa.** Trong định lý trên không gian con $W$ được gọi là **phần bù trực giao** của không gian con $U$, kí hiệu $U^\perp$. Do $U \cap U^\perp = \{\mathbf{0}\}$ nên tổng $U + U^\perp$ được gọi là **tổng trực tiếp** và kí hiệu $U \oplus U^\perp$. Để nhấn mạnh tổng trực tiếp của các không gian con trực giao nhau ta thường dùng ký hiệu $U \overset{\perp}{\oplus} U^\perp$.

**4.1.3 Mệnh đề.** *Cho $V_E$ là không gian Euclid có các không gian con $U$, $W$, $U_1$, $U_2$. Khi đó, ta có*

**(i)** *Nếu $W \perp U$, thì $W \subseteq U^\perp$.*

**(ii)** *Nếu $U_1 \subseteq U_2$, thì $U_2^\perp \subseteq U_1^\perp$.*

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

**4.1.4 Mệnh đề.** *Cho $U$ là không gian con của không gian Euclid $V_E$ thỏa $U = \langle \mathbf{u}_1, \dots, \mathbf{u}_k \rangle$. Khi đó, $\mathbf{x} \in U^\perp$ khi và chỉ khi $\mathbf{x} \perp \mathbf{u}_i$ với mọi $i = 1, \dots, k$.*

*Chứng minh.* Nếu $\mathbf{x} \in U^\perp$, thì rõ ràng $\mathbf{x} \perp \mathbf{u}_i$ với mọi $i = 1, \dots, k$ (vì $\mathbf{u}_i \in U$).

Ngược lại, giả sử $\mathbf{x} \perp \mathbf{u}_i$ với mọi $i = 1, \dots, k$. Lấy $\mathbf{u} \in U$ bất kỳ, ta có biểu diễn $\mathbf{u} = \sum_{i=1}^k \alpha_i \mathbf{u}_i$. Do đó,

$$\langle \mathbf{x}, \mathbf{u} \rangle = \langle \mathbf{x}, \sum_{i=1}^k \alpha_i \mathbf{u}_i \rangle = \sum_{i=1}^k \alpha_i \langle \mathbf{x}, \mathbf{u}_i \rangle = 0.$$

Vậy $\mathbf{x} \perp \mathbf{u}$. Suy ra $\mathbf{x} \perp U$. □

**4.1.5 Thí dụ.** Xét không gian Euclid $\mathbb{R}^3$ và $U = \langle (1, 1, 1), (1, 2, 3) \rangle$. Theo mệnh đề trên ta có $(x_1, x_2, x_3) \in U^\perp$ khi và chỉ khi $x_1 + x_2 + x_3 = 0$ và $x_1 + 2x_2 + 3x_3 = 0$. Vậy $U^\perp = \langle (1, -2, 1) \rangle$. □

**4.1.6 Định lý.** *Nếu $U$ là không gian con hữu hạn chiều của không gian Euclid $V_E$, thì $V_E = U \oplus U^\perp$.*

*Chứng minh.* Do $U \cap U^\perp = \{\mathbf{0}\}$ nên $U + U^\perp = U \oplus U^\perp$. Rõ ràng $U + U^\perp \subseteq V_E$.

Vì $U$ là không gian con hữu hạn chiều trong $V_E$ nên nó có một cơ sở trực chuẩn $\mathscr{B} = (\mathbf{b}_1, \dots, \mathbf{b}_k)$ nào đó. Với $\mathbf{x} \in V_E$ bất kỳ, đặt $\mathbf{u} = \langle \mathbf{x}, \mathbf{b}_1 \rangle \mathbf{b}_1 + \cdots + \langle \mathbf{x}, \mathbf{b}_k \rangle \mathbf{b}_k$ và $\mathbf{v} = \mathbf{x} - \mathbf{u}$. Rõ ràng $\mathbf{x} = \mathbf{u} + \mathbf{v}$. Hơn nữa, với mỗi $i = 1, \dots, k$, ta có

$$\begin{aligned} \langle \mathbf{v}, \mathbf{b}_i \rangle &= \langle \mathbf{x}, \mathbf{b}_i \rangle - \langle \mathbf{u}, \mathbf{b}_i \rangle \\ &= \langle \mathbf{x}, \mathbf{b}_i \rangle - \langle \mathbf{x}, \mathbf{b}_1 \rangle \langle \mathbf{b}_1, \mathbf{b}_i \rangle - \cdots - \langle \mathbf{x}, \mathbf{b}_k \rangle \langle \mathbf{b}_k, \mathbf{b}_i \rangle. \end{aligned}$$

Bởi vì $(\mathbf{b}_1, \ldots, \mathbf{b}_k)$ là cơ sở trực chuẩn nên $\langle \mathbf{b}_j, \mathbf{b}_i \rangle = \delta_{ij}$. Do đó, ta có được $\langle \mathbf{v}, \mathbf{b}_i \rangle = \langle \mathbf{x}, \mathbf{b}_i \rangle - \langle \mathbf{x}, \mathbf{b}_i \rangle = 0$ hay $\mathbf{v} \perp \mathbf{b}_i$. Vì $U = \langle \mathbf{b}_1, \ldots, \mathbf{b}_k \rangle$ nên $\mathbf{v} \perp U$ hay $\mathbf{v} \in U^\perp$. Vậy $V_E \subseteq U + U^\perp$. Do đó $V_E = U \oplus U^\perp$. □

**4.1.7 Hệ quả.** *Nếu $U$ là không gian con của không gian Euclid hữu hạn chiều $V_E$, thì* $\dim V = \dim U + \dim U^\perp$.

*Chứng minh.* Do $V_E$ là không gian hữu hạn chiều, nên $U$ cũng hữu hạn chiều. Theo định lý trên ta có $V_E = U \oplus U^\perp$. Do đó, $\dim V_E = \dim U + \dim U^\perp$. □

**4.1.8 Mệnh đề.** *Cho $U$, $U_1$, $U_2$ là các không gian con trong không gian Euclid hữu hạn chiều $V_E$. Khi đó, ta có*

- **(i)** $(U^\perp)^\perp = U$.
- **(ii)** $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$.
- **(iii)** $U_1^\perp + U_2^\perp = (U_1 \cap U_2)^\perp$.

*Chứng minh.* Chúng tôi trình bày chứng minh tính chất (iii) còn hai tính chất đầu dành cho bạn đọc ở phần bài tập.

Do $V_E$ là không gian hữu hạn chiều nên các không gian con $U_1^\perp$ và $U_2^\perp$ cũng vậy, suy ra chúng có cơ sở. Gọi $\mathscr{A}$ là một cơ sở của $U_1^\perp$ và $\mathscr{B}$ là một cơ sở của $U_2^\perp$. Khi đó $\mathscr{A} \cup \mathscr{B}$ là một hệ sinh của $U_1^\perp + U_2^\perp$. Với mọi $\mathbf{x} \in U_1^\perp + U_2^\perp$, ta có $\mathbf{x}$ trực giao với $\mathscr{A} \cup \mathscr{B}$, suy ra $\mathbf{x}$ trực giao với $\mathscr{A}$ và $\mathbf{x}$ trực giao với $\mathscr{B}$. Do đó, $\mathbf{x} \perp U_1$ và $\mathbf{x} \perp U_2$, suy ra $\mathbf{x} \perp \mathbf{y}$ với mọi $\mathbf{y} \in U_1 \cap U_2$. Vậy $\mathbf{x} \in (U_1 \cap U_2)^\perp$, cho nên $U_1^\perp + U_2^\perp \subseteq (U_1 \cap U_2)^\perp$.

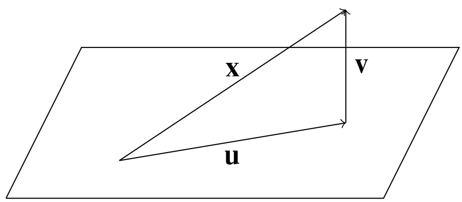
Ngược lại, với mọi $\mathbf{x} \in (U_1 \cap U_2)^\perp$ ta có $\mathbf{x} \perp U_1$ và $\mathbf{x} \perp U_2$. Suy ra $\mathbf{x} \in U_1^\perp$ và $\mathbf{x} \in U_2^\perp$. Vậy $\mathbf{x} \in U_1^\perp + U_2^\perp$ vì $U_1^\perp \cup U_2^\perp \subseteq U_1^\perp + U_2^\perp$. Do đó $(U_1 \cap U_2)^\perp \subseteq U_1^\perp + U_2^\perp$. Ta được điều phải chứng minh. □

## 4.2 Phép chiếu trực giao

**4.2.1 Định lý.** *Cho không gian Euclid $V_E$ và $U$ là không gian con của $V_E$ sao cho $V_E = U \oplus U^\perp$. Khi đó, với mọi $\mathbf{x} \in V_E$ ta đều viết được một cách duy nhất dưới dạng $\mathbf{x} = \mathbf{u} + \mathbf{v}$ trong đó $\mathbf{u} \in U$ và $\mathbf{v} \in U^\perp$.*

*Chứng minh.* Vì $V_E = U \oplus U^{\perp}$ nên với mọi $\mathbf{x} \in V_E$ ta đều có thể viết dưới dạng $\mathbf{x} = \mathbf{u} + \mathbf{v}$ với $\mathbf{u} \in U$ và $\mathbf{v} \in U^{\perp}$. Giả sử ta có thể viết một cách nữa $\mathbf{x} = \mathbf{u}' + \mathbf{v}'$ với $\mathbf{u}' \in U$ và $\mathbf{v}' \in U^{\perp}$. Khi đó, ta có $\mathbf{u} + \mathbf{v} = \mathbf{u}' + \mathbf{v}'$ suy ra $\mathbf{u} - \mathbf{u}' = \mathbf{v} - \mathbf{v}'$. Do $\mathbf{u} - \mathbf{u}' \in U$ và $\mathbf{v} - \mathbf{v}' \in U^{\perp}$ nên $\mathbf{u} - \mathbf{u}' = \mathbf{v} - \mathbf{v}' = \mathbf{0}$ vì $U \cap U^{\perp} = \{\mathbf{0}\}$. Vậy $\mathbf{u} = \mathbf{u}'$ và $\mathbf{v} = \mathbf{v}'$. $\square$

**4.2.2 Định nghĩa.** Cho không gian Euclid $V_E$ và $U$ là không gian con của $V_E$ sao cho $V_E = U \oplus U^{\perp}$. Khi đó, ánh xạ $P : V_E \to V_E$ xác định bởi $P(\mathbf{x}) = \mathbf{u}$ trong đó $\mathbf{x} = \mathbf{u} + \mathbf{v}$ với $\mathbf{u} \in U$ và $\mathbf{v} \in U^{\perp}$ là một toán tử tuyến tính trên $V_E$, nó được gọi là **phép chiếu trực giao** lên không gian con $U$ và $\mathbf{u}$ được gọi là **hình chiếu trực giao** của $\mathbf{x}$ lên $U$, $\mathbf{v}$ được gọi là **thành phần trực giao** của $\mathbf{x}$ đối với $U$.




Ta kiểm chứng trong định nghĩa trên $P$ là một ánh xạ tuyến tính. Với $\mathbf{x}, \mathbf{y} \in V_E$ ta có biểu diễn $\mathbf{x} = \mathbf{u} + \mathbf{v}$ và $\mathbf{y} = \mathbf{u}' + \mathbf{v}'$ với $\mathbf{u}, \mathbf{u}' \in U$ và $\mathbf{v}, \mathbf{v}' \in U^{\perp}$. Với $\alpha, \beta \in \mathbb{K}$, ta có $\alpha\mathbf{u} + \beta\mathbf{u}' \in U$, cho nên

$$P(\alpha\mathbf{x} + \beta\mathbf{y}) = P(\alpha\mathbf{u} + \beta\mathbf{u}' + \alpha\mathbf{v} + \beta\mathbf{v}') = \alpha\mathbf{u} + \beta\mathbf{u}' = \alpha P(\mathbf{x}) + \beta P(\mathbf{y}).$$

**4.2.3 Mệnh đề.** *Nếu $P$ là phép chiếu trực giao lên không gian con $U$ trong không gian Euclid $V_E$, thì $P(\mathbf{x}) = \mathbf{u}$ là vector trong $U$ gần vector $\mathbf{x} \in V_E$ nhất (nghĩa là $\|\mathbf{x} - \mathbf{u}\| = \min\{\|\mathbf{x} - \mathbf{u}'\| : \mathbf{u}' \in U\}$ và $\|\mathbf{x} - \mathbf{u}\|$ được gọi là khoảng cách từ $\mathbf{x}$ đến $U$).*

*Chứng minh.* Lấy $\mathbf{u}' \in U$ và $\mathbf{x} = \mathbf{u} + \mathbf{v}$ trong đó $\mathbf{v} \in U^{\perp}$. Khi đó,

$$\begin{aligned}
\|\mathbf{x} - \mathbf{u}'\|^2 &= \langle (\mathbf{u} - \mathbf{u}') + \mathbf{v}, (\mathbf{u} - \mathbf{u}') + \mathbf{v} \rangle \\
&= \langle \mathbf{u} - \mathbf{u}', \mathbf{u} - \mathbf{u}' \rangle + 2\langle \mathbf{u} - \mathbf{u}', \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \\
&= \|\mathbf{u} - \mathbf{u}'\|^2 + \|\mathbf{v}\|^2 \\
&\geq \|\mathbf{v}\|^2.
\end{aligned}$$

Dấu bằng xảy ra khi và chỉ khi $\|\mathbf{u} - \mathbf{u}'\|^2 = 0$ hay $\mathbf{u} = \mathbf{u}'$. □

**4.2.4** Cho $U$ là không gian con của không gian Euclid $V_E$ và $U$ có cơ sở trực chuẩn là $\mathscr{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_k)$. Khi đó, theo chứng minh Định lý 4.1.6 ta có công thức của phép chiếu trực giao lên $U$ là

$$P(\mathbf{x}) = \langle \mathbf{x}, \mathbf{b}_1 \rangle \mathbf{b}_1 + \cdots + \langle \mathbf{x}, \mathbf{b}_k \rangle \mathbf{b}_k. \tag{4.2.5}$$

**4.2.6 Thí dụ.** Tìm hình chiếu trực giao và thành phần trực giao của vector $\mathbf{x} = (-1, 2, 1)$ lên không gian con $U = \langle (2, 0, 2), (1, -1, -2) \rangle$ của không gian Eulid $\mathbb{R}^3$.

Trước hết ta tìm một cơ sở trực chuẩn của không gian con $U$. Ta có $\mathbf{w}_1 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ và

$$\mathbf{v}_2 = (1, -1, -2) - \frac{1}{2}(-1)(1, 0, 1) = \left(\frac{3}{2}, -1, -\frac{3}{2}\right)$$

Vậy $\mathbf{w}_2 = (\frac{3}{\sqrt{22}}, \frac{-2}{\sqrt{22}}, \frac{-3}{\sqrt{22}})$. Ta được một cơ sở trực chuẩn của $U$ là

$$((\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}), (\tfrac{3}{\sqrt{22}}, \tfrac{-2}{\sqrt{22}}, \tfrac{-3}{\sqrt{22}})).$$

Do đó, hình chiếu trực giao của $\mathbf{x}$ lên $U$ là

$$\mathbf{u} = \frac{1}{2}(0)(1, 0, 1) + \frac{1}{22}(-10)(3, -2, -3) = \frac{1}{11}(-15, 10, 15)$$

và thành phần trực giao của $\mathbf{x}$ là $\mathbf{v} = (-1, 2, 1) - \frac{1}{11}(-15, 10, 15) = (\frac{4}{11}, \frac{12}{11}, \frac{-4}{11})$. □

**4.2.7** Giả sử $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$ là một cơ sở của không gian con $U$ của không gian Eulid $V_E$. Với $\mathbf{x} \in V_E$ bất kỳ, ta có hình chiếu vuông góc của $\mathbf{x}$ lên $U$ có biểu diễn duy nhất $P(\mathbf{x}) = \sum\limits_{i=1}^{k} x_i\mathbf{a}_i$. Khi đó thành phần trực giao của $\mathbf{x}$ đối với $U$ trực giao với mọi $\mathbf{a}_i$ với $i = 1, \ldots, k$. Do đó

$$0 = \langle \mathbf{x} - P(\mathbf{x}), \mathbf{a}_i\rangle = \langle \mathbf{x}, \mathbf{a}_i\rangle - \sum_{j=1}^{n} x_j\langle \mathbf{a}_j, \mathbf{a}_i\rangle \qquad i = 1, \ldots, k$$

Vậy ta có hệ phương trình

$$\begin{cases} \langle \mathbf{a}_1, \mathbf{a}_1\rangle x_1 + \langle \mathbf{a}_1, \mathbf{a}_2\rangle x_2 + \cdots + \langle \mathbf{a}_1, \mathbf{a}_k\rangle x_k = \langle \mathbf{x}, \mathbf{a}_1\rangle \\ \langle \mathbf{a}_2, \mathbf{a}_1\rangle x_1 + \langle \mathbf{a}_2, \mathbf{a}_2\rangle x_2 + \cdots + \langle \mathbf{a}_2, \mathbf{a}_k\rangle x_k = \langle \mathbf{x}, \mathbf{a}_2\rangle \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \langle \mathbf{a}_k, \mathbf{a}_1\rangle x_1 + \langle \mathbf{a}_k, \mathbf{a}_2\rangle x_2 + \cdots + \langle \mathbf{a}_k, \mathbf{a}_k\rangle x_k = \langle \mathbf{x}, \mathbf{a}_k\rangle \end{cases}$$

có nghiệm duy nhất, hay đó là một hệ Cramer, cho nên định thức hệ số của nó khác không:

$$D = \begin{vmatrix} \langle \mathbf{a}_1, \mathbf{a}_1\rangle & \langle \mathbf{a}_1, \mathbf{a}_2\rangle & \ldots & \langle \mathbf{a}_1, \mathbf{a}_k\rangle \\ \langle \mathbf{a}_2, \mathbf{a}_1\rangle & \langle \mathbf{a}_2, \mathbf{a}_2\rangle & \ldots & \langle \mathbf{a}_2, \mathbf{a}_k\rangle \\ \dots & \dots & \dots & \dots \\ \langle \mathbf{a}_k, \mathbf{a}_1\rangle & \langle \mathbf{a}_k, \mathbf{a}_2\rangle & \ldots & \langle \mathbf{a}_k, \mathbf{a}_k\rangle \end{vmatrix} \neq 0$$

Định thức $D$ gọi là định thức Gram của hệ vector $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k)$

## Bài tập

**1)** Tìm cơ sở của phần bù trực giao $W^\perp$ của không gian con $W$ sau đây

**(a)** $W = \langle (1, -2, 0, 1), (0, 1, 0, 2)\rangle$

**(b)** $W = \langle (2, 2, 1, 0), (2, -3, 2, 4)\rangle$

**(c)** $W = \langle (1, 0, 2, 1), (2, 1, 2, 3), (0, 1, -2, 3), (0, 1, -2, 1)\rangle$

**(d)** $W$ là không gian nghiệm của hệ phương trình $\begin{cases} x_2 - 2x_3 = 0 \\ x_1 + 2x_2 - x_3 + x_4 = 0 \end{cases}$

**2**) Trong $\mathbb{E}^3$ cho $W_1$ và $W_2$ lần lượt là không gian nghiệm của các hệ phương trình tuyến tính thuần nhất

$$\begin{cases} x_1 + x_2 - x_3 = 0 \\ 3x_1 + x_2 - x_3 = 0 \end{cases} \qquad \begin{cases} 3x_1 + 2x_2 - 2x_3 = 0 \\ -6x_1 - 4x_2 + 4x_3 = 0 \end{cases}$$

Chứng minh $W_1 \perp W_2$.

**3**) Trong $\mathbb{E}^3$ cho $W_1$ và $W_2$ lần lượt là không gian nghiệm của ccs hệ phương trình tuyến tính thuần nhất

$$\begin{cases} x_1 - 2x_2 - 2x_3 - x_4 = 0 \\ 2x_1 + 3x_2 + x_3 + 2x_4 = 0 \end{cases} \qquad \begin{cases} x_1 + x_2 + x_3 + 3x_4 = 0 \\ x_1 + 4x_2 + 7x_4 = 0 \end{cases}$$

Chứng minh $W_1 \perp W_2$.

**4**) Trong $\mathbb{E}^4$ cho các không gian con

$$W_1 = \langle (1,1,1,-2), (3,-1,5,0) \rangle, \quad W_2 = \langle (1,3,0,2), (0,5,1,m) \rangle$$

Xác định $m$ để $W_1 \perp W_2$.

**5**) Tìm hình chiếu vuông góc và thành phần trực giao của **x** lên không gian con $W$:

**(a)** $\mathbf{x} = (4,-1,-3,4)$ và $W = \langle (1,1,1,1), (1,2,2,-1), (1,0,0,3) \rangle$

**(b)** $\mathbf{x} = (5,2,-2,2)$ và $W = \langle (2,1,1,-1), (1,1,3,0), (1,2,8,1) \rangle$

**(c)** $\mathbf{x} = (7,-4,-1,2)$ và $W$ là không gian nghiệm của hệ

$$\begin{cases} 2x_1 + x_2 + x_3 + 3x_4 = 0 \\ 3x_1 + 2x_2 + 2x_3 + x_4 = 0 \\ x_1 + 2x_2 + 2x_3 - 9x_4 = 0 \end{cases}$$

**6**) Chứng minh Mệnh đề 4.1.3.

**7**) Chứng minh Mệnh đề 4.1.8.

**8**) Cho $U$ là một không gian con của không gian Euclid $n$ chiều $V_E$ và vector $\mathbf{x}_0 \in V_E$. Ta gọi tập $P = U + \mathbf{x}_0 = \{\mathbf{x} + \mathbf{x}_0 : \mathbf{x} \in U\}$ là một *đa tạp tuyến tính* của $V_E$. Ta gọi khoảng cách từ một vector **a** thuộc $V_E$ đến đa tạp $P$ là số $\min\{\|\mathbf{a} - \mathbf{u}\| : \mathbf{u} \in P\}$. Chứng minh rằng khoảng cách từ **a** tới đa tạp $P$ bằng độ dài của đường trực giao hạ từ vector $\mathbf{a} - \mathbf{x}_0$ xuống $U$.

**9**) Tìm khoảng cách từ vector **a** thuộc $\mathbb{R}^4$ đến đa tạp $P$

**(a)** $\mathbf{a} = (4, 2, -5, 1)$ và $P$ xác định bởi hệ phương trình tuyến tính

$$\begin{cases} 2x_1 - 2x_2 + x_3 + 2x_4 = 9 \\ 2x_1 - 4x_2 + x_3 + 3x_4 = 12. \end{cases}$$

**(b)** $\mathbf{a} = (2, 4, -4, 12)$ và $P$ xác định bởi hệ phương trình tuyến tính

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

**10**) Ta gọi khoảng cách giữa hai đa tạp tuyến tính $P_1$ và $P_2$ là số nhỏ nhất trong tất cả các khoảng cách giữa hai điểm bất kỳ, một điểm thuộc $P_1$ còn một điểm thuộc $P_2$. Chứng minh rằng khoảng cách giữa hai đa tạp tuyến tính $P_1 = U_1 + \mathbf{x}_1$ và $P_2 = U_2 + \mathbf{x}_2$ bằng độ dài đường trực giao hạ từ vector $\mathbf{x}_1 - \mathbf{x}_2$ xuống không gian con $U = U_1 + U_2$.

**11**) Hãy tìm khoảng cách giữa hai đa tạp $P_1 = U_1 + \mathbf{x}_1$ và $P_2 = U_2 + \mathbf{x}_2$ với $\mathbf{x}_1 = (4, 5, 3, 2)$, $\mathbf{x}_2 = (1, -2, 1, -3)$, $U_1 = \langle (1, 2, 2, 2), (2, -2, 1, 2) \rangle$ và $U_2 = \langle (2, 0, 2, 1), (1, -2, 0, -1) \rangle$.

**12**) Cho **e** là một vector có độ dài bằng 1 của không gian Euclid $n$ chiều $V_E$. Chứng minh rằng một vector bất kỳ **x** thuộc $V_E$ luôn phân tích được một cách duy nhất dưới dạng $\mathbf{x} = \alpha\mathbf{e} + \mathbf{z}$, trong đó $\alpha \in \mathbb{R}$ còn $\mathbf{z} \perp \mathbf{e}$. Số thực $\alpha$ trong sự phân tích trên gọi là *hình chiếu của vector* **x** *trên phương* **e** và được ký hiệu $\mathrm{pr}_{\mathbf{e}}\mathbf{x}$. Chứng minh rằng

**(a)** $\mathrm{pr}_{\mathbf{e}}(\mathbf{x} + \mathbf{y}) = \mathrm{pr}_{\mathbf{e}}\mathbf{x} + \mathrm{pr}_{\mathbf{e}}\mathbf{y}$

**(b)** $\text{pr}_{\mathbf{e}}(\lambda\mathbf{x}) = \lambda\text{pr}_{\mathbf{e}}\mathbf{x}$

**(c)** Nếu $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ là một cơ sở trực chuẩn của $V_E$ thì với mọi $\mathbf{x} \in V_E$ ta có $\mathbf{x} = \sum\limits_{i=1}^{n}(\text{pr}_{\mathbf{e}_i}\mathbf{x})\mathbf{e}_i$.

**(d)** Nếu $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k)$ là một hệ trực chuẩn trong $V_E$ thì với mọi $\mathbf{x} \in V_E$ ta có $\sum\limits_{i=1}^{k}(\text{pr}_{\mathbf{e}_i}\mathbf{x})^2 \leq \|\mathbf{x}\|$. Dấu đẳng thức xảy ra khi và chỉ khi $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k)$ là một cơ sở trực chuẩn của $V_E$

**13 )** Chứng minh rằng trong tất cả các vector thuộc không gian con $U$, hình chiếu vuông góc $\mathbf{u}$ lên $U$ của vector $\mathbf{v}$ tạo với $\mathbf{v}$ một góc nhỏ nhất. Hơn nữa nếu có $\mathbf{u}' \in U$ sao cho $\cos(\widehat{\mathbf{v}, \mathbf{u}}) = \cos(\widehat{\mathbf{v}, \mathbf{u}'})$ thì $\mathbf{u}' = k\mathbf{u}$ với $k$ là số dương. Ta gọi góc nhỏ nhất này là góc giữa vector $\mathbf{v}$ và không gian con $U$.

**14 )** Hãy tìm góc giữa $\mathbf{v}$ và $U$ với

**(a)** $\mathbf{v} = (2, 2, 1, 1)$ và $U = \langle(3, 4, -4, -1), (0, 1, -1, 2)\rangle$

**(b)** $\mathbf{v} = (1, 0, 3, 0)$ và $U = \langle(5, 3, 4, -3), (1, 1, 4, 5), (2, -1, 1, 2)\rangle$.

**15 )** Chứng minh rằng định thức Gram

$$g(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k) = \begin{vmatrix} \langle\mathbf{e}_1, \mathbf{e}_1\rangle & \langle\mathbf{e}_1, \mathbf{e}_2\rangle & \dots & \langle\mathbf{e}_1, \mathbf{e}_k\rangle \\ \langle\mathbf{e}_2, \mathbf{e}_1\rangle & \langle\mathbf{e}_2, \mathbf{e}_2\rangle & \dots & \langle\mathbf{e}_2, \mathbf{e}_k\rangle \\ \dots & \dots & \dots & \dots \\ \langle\mathbf{e}_k, \mathbf{e}_1\rangle & \langle\mathbf{e}_k, \mathbf{e}_2\rangle & \dots & \langle\mathbf{e}_k, \mathbf{e}_k\rangle \end{vmatrix}$$

của hệ vector $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k)$ trong không gian Euclid $n$ chiều $V_E$ là không đổi qua quá trình trực giao hóa. Nghĩa là, nếu quá trình trực giao hóa biến hệ vector $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k)$ thành hệ trực giao $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$ thì $g(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k) = g(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$.

**16 )** Chứng minh hệ vector $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k)$ của không gian Euclid $E$ là phụ thuộc tuyến tính khi và chỉ khi $g(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k) = 0$.

# § 5 Đồng cấu trực giao

## 5.1 Đồng cấu trực giao - Toán tử trực giao

**5.1.1 Định nghĩa.** Cho hai không gian Euclid $V_E$ và $V'_E$ với tích vô hướng lần lượt là $\langle , \rangle$ và $\langle , \rangle'$. Ánh xạ $f : V_E \to V'_E$ được gọi là **đồng cấu trực giao** nếu $f$ là một ánh xạ tuyến tính và $f$ bảo toàn tích vô hướng, nghĩa là với mọi $\mathbf{x}, \mathbf{y} \in V_E$ ta có

$$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle' = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Đồng cấu trực giao $f : V_E \to V'_E$ gọi là một **đẳng cấu trực giao** nếu $f$ là một song ánh. Khi đó, ta bảo $V_E$ và $V'_E$ đẳng cấu trực giao với nhau, kí hiệu $V_E \cong V'_E$.

Đồng cấu trực giao $f : V_E \to V_E$ được gọi là **toán tử trực giao** trên $V_E$.

**5.1.2 Mệnh đề. (1)** *Mọi đồng cấu trực giao đều là đơn cấu.*

**(2)** *Hợp thành của hai đồng cấu trực giao lại là một đồng cấu trực giao.*

*Chứng minh.*

**(1)** Cho $f : V_E \to V'_E$ là một đồng cấu trực giao. Giả sử $\mathbf{x} \in \ker f$. Ta có $f(\mathbf{x}) = \mathbf{0}$ và

$$\langle \mathbf{x}, \mathbf{x} \rangle = \langle f(\mathbf{x}), f(\mathbf{x}) \rangle' = \langle \mathbf{0}, \mathbf{0} \rangle' = 0.$$

Do đó, $\mathbf{x} = \mathbf{0}$, suy ra $\ker f = \{\mathbf{0}\}$. Vậy $f$ là một đơn ánh.

**(2)** Giả sử $f : V_E \to V'_E$ và $g : V'_E \to V''_E$ là các đồng cấu trực giao. Ta thấy rằng $g \circ f$ là một ánh xạ tuyến tính. Với mọi $\mathbf{x}, \mathbf{y} \in V_E$ ta có

$$\langle g \circ f(\mathbf{x}), g \circ f(\mathbf{y}) \rangle'' = \langle g(f(\mathbf{x})), g(f(\mathbf{y})) \rangle'' = \langle f(\mathbf{x}), f(\mathbf{y}) \rangle' = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Vậy $g \circ f$ là một đồng cấu trực giao. □

**5.1.3 Định lý.** *Hai không gian Euclid hữu hạn chiều đẳng cấu trực giao với nhau khi và chỉ khi chúng có cùng số chiều.*

*Chứng minh.* Giả sử hai không gian Euclid hữu hạn chiều đẳng cấu trực giao nhau. Khi đó, rõ ràng chúng có cùng số chiều vì một đẳng cấu biến một cơ sở của không gian này thành một cơ sở của không gian kia.

Giả sử $V_E$ và $V'_E$ là hai không gian Euclid $n$ chiều. Khi đó, tồn tại hai cơ sở trực chuẩn tương ứng cho hai không gian ấy là $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ và $\mathscr{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$. Ta thiết lập một đẳng cấu $f : V_E \to V'_E$ thỏa $f(\mathbf{a}_i) = \mathbf{b}_i$. Với mọi $\mathbf{x}, \mathbf{y} \in V_E$ ta có biểu diễn $\mathbf{x} = \sum\limits_{i=1}^{n} \alpha_i \mathbf{a}_i$ và $\mathbf{y} = \sum\limits_{j=1}^{n} \beta_j \mathbf{a}_j$. Ta được

$$\begin{aligned}\langle f(\mathbf{x}), f(\mathbf{y})\rangle' &= \Big\langle \sum_{i=1}^{n} \alpha_i f(\mathbf{a}_i), \sum_{j=1}^{n} \beta_j f(\mathbf{a}_j) \Big\rangle' \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \langle \mathbf{b}_i, \mathbf{b}_j\rangle' \\ &= \sum_{i=1}^{n} \alpha_i \beta_i = \langle \mathbf{x}, \mathbf{y}\rangle.\end{aligned}$$

Vậy $f$ là một đẳng cấu trực giao hay $V_E$, $V'_E$ đẳng cấu trực giao nhau. □

**5.1.4 Định lý.** *Cho không gian Euclid hữu hạn chiều $V_E$ và ánh xạ $f : V_E \to V_E$. Khi đó, các khẳng định sau là tương đương*

**(1)** *$f$ là một toán tử trực giao.*

**(2)** *$f$ là một toán tử tuyến tính và biến cơ sở trực chuẩn thành cơ sở trực chuẩn.*

**(3)** *$f$ là một toán tử tuyến tính và giữ nguyên độ dài của các vector, tức là $\|f(\mathbf{x})\| = \|\mathbf{x}\|$ với mọi $\mathbf{x} \in V_E$.*

**(4)** *$\|f(\mathbf{x}) + f(\mathbf{y})\| = \|\mathbf{x} + \mathbf{y}\|$ với mọi $\mathbf{x}, \mathbf{y} \in V_E$.*

**(5)** *$\langle f(\mathbf{x}), f(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ với mọi $\mathbf{x}, \mathbf{y} \in V_E$.*

**(6)** *$f$ là một toán tử tuyến tính và có ma trận trong một cơ sở trực chuẩn nào đó của $V_E$ là ma trận trực giao.*

*Chứng minh.*

(1)⇒(2) Giả sử $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ là một cơ sở trực chuẩn của $V_E$. Vì $f$ là một toán tử trực giao nên $(f(\mathbf{a}_1), f(\mathbf{a}_2), \dots, f(\mathbf{a}_n))$ là một cơ sở của $V_E$. Hơn nữa, ta có

$$\langle f(\mathbf{a}_i), f(\mathbf{a}_j)\rangle = \langle \mathbf{a}_i, \mathbf{a}_j\rangle = \delta_{ij}.$$

Vậy $(f(\mathbf{a}_1), f(\mathbf{a}_2), \dots, f(\mathbf{a}_n))$ là một cơ sở trực chuẩn của $V_E$.

(2)⇒(3) Giả sử $f$ là một toán tử tuyến tính trên $V_E$ biến cơ sở trực chuẩn $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ thành cơ sở trực chuẩn $\mathscr{B} = (f(\mathbf{a}_1), f(\mathbf{a}_2), \dots, f(\mathbf{a}_n))$. Với $\mathbf{x} \in V_E$ tùy ý, giả sử $\mathbf{x}|_{\mathscr{A}} = (x_1, x_2, \dots, x_n)$. Tức là $\mathbf{x} = \sum\limits_{i=1}^{n} x_i\mathbf{a}_i$, suy ra $f(\mathbf{x}) = \sum\limits_{i=1}^{n} x_i f(\mathbf{a}_i)$ hay $f(\mathbf{x})|_{\mathscr{B}} = (x_1, x_2, \dots, x_n)$. Theo Mệnh đề 3.1.4 ta có

$$\|f(\mathbf{x})\| = \sqrt{\sum_{i=1}^{n} x_i^2} = \|\mathbf{x}\|.$$

(3)⇒(4) Giả sử $f$ là một toán tử tuyến tính trên $V_E$ sao cho $\|f(\mathbf{x})\| = \|\mathbf{x}\|$ với mọi $\mathbf{x} \in V_E$. Khi đó, với mọi $\mathbf{x}, \mathbf{y} \in V_E$ ta có ngay

$$\|f(\mathbf{x}) + f(\mathbf{y})\| = \|f(\mathbf{x} + \mathbf{y})\| = \|\mathbf{x} + \mathbf{y}\|.$$

(4)⇒(5) Giả sử với mọi $\mathbf{x}, \mathbf{y} \in V_E$ ta có $\|f(\mathbf{x}) + f(\mathbf{y})\| = \|\mathbf{x} + \mathbf{y}\|$. Với mọi $\mathbf{x} \in V_E$ ta thấy rằng $\|f(\mathbf{x}) + f(\mathbf{x})\| = \|\mathbf{x} + \mathbf{x}\|$ cho nên $\|f(\mathbf{x})\| = \|\mathbf{x}\|$. Nhờ vào mối liên hệ giữa tích vô hướng và độ dài ta có

$$\begin{aligned}\langle f(\mathbf{x}), f(\mathbf{y})\rangle &= \frac{1}{2}(\|f(\mathbf{x}) + f(\mathbf{y})\|^2 - \|f(\mathbf{x})\|^2 - \|f(\mathbf{y})\|^2)\\ &= \frac{1}{2}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2) = \langle \mathbf{x}, \mathbf{y}\rangle,\end{aligned}$$

với $\mathbf{x}, \mathbf{y} \in V_E$ tùy ý.

(5)⇒(1) Giả sử ánh xạ $f$ bảo toàn tích vô hướng. Với mọi $\mathbf{x}, \mathbf{y} \in V_E$ và mọi $\alpha, \beta \in \mathbb{R}$, ta có

$$\begin{aligned}
&\langle f(\alpha\mathbf{x}+\beta\mathbf{y}) - \alpha f(\mathbf{x}) - \beta f(\mathbf{y}), f(\alpha\mathbf{x}+\beta\mathbf{y}) - \alpha f(\mathbf{x}) - \beta f(\mathbf{y})\rangle \\
&\quad = \langle f(\alpha\mathbf{x}+\beta\mathbf{y}), f(\alpha\mathbf{x}+\beta\mathbf{y})\rangle + \alpha^2\langle f(\mathbf{x}), f(\mathbf{x})\rangle + \beta^2\langle f(\mathbf{y}), f(\mathbf{y})\rangle \\
&\qquad - 2\alpha\langle f(\alpha\mathbf{x}+\beta\mathbf{y}), f(\mathbf{x})\rangle - 2\beta\langle f(\alpha\mathbf{x}+\beta\mathbf{y}), f(\mathbf{y})\rangle + 2\alpha\beta\langle f(\mathbf{x}), f(\mathbf{y})\rangle \\
&\quad = \langle \alpha\mathbf{x}+\beta\mathbf{y}, \alpha\mathbf{x}+\beta\mathbf{y}\rangle + \alpha^2\langle\mathbf{x},\mathbf{x}\rangle + \beta^2\langle\mathbf{y},\mathbf{y}\rangle \\
&\qquad - 2\alpha\langle\alpha\mathbf{x}+\beta\mathbf{y},\mathbf{x}\rangle - 2\beta\langle\alpha\mathbf{x}+\beta\mathbf{y},\mathbf{y}\rangle + 2\alpha\beta\langle\mathbf{x},\mathbf{y}\rangle \\
&\quad = \alpha^2\langle\mathbf{x},\mathbf{x}\rangle + \beta^2\langle\mathbf{y},\mathbf{y}\rangle + 2\alpha\beta\langle\mathbf{x},\mathbf{y}\rangle - \langle\alpha\mathbf{x}+\beta\mathbf{y},\alpha\mathbf{x}+\beta\mathbf{y}\rangle \\
&\quad = 0.
\end{aligned}$$

Vậy $f(\alpha\mathbf{x}+\beta\mathbf{y}) - \alpha f(\mathbf{x}) - \beta f(\mathbf{y}) = 0$ hay $f(\alpha\mathbf{x}+\beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$. Do đó $f$ là một toán tử tuyến tính trên $V_E$, suy ra $f$ là một toán tử trực giao.

(2)⇒(6) Giả sử $f$ là một toán tử tuyến tính biến mọi cơ sở trực chuẩn thành cơ sở trực chuẩn. Giả sử $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một cơ sở trực chuẩn bất kỳ. Ta có $\mathscr{B} = (f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ là một cơ sở trực chuẩn. Theo Định lý 3.1.6 ta được ma trận của $f$ đối với cơ sở $\mathscr{A}$ là một ma trận trực giao.

(6)⇒(2) Giả sử $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một cơ sở trực chuẩn tùy ý. Khi đó ma trận của toán tử tuyến tính $f$ đối với $\mathscr{A}$ là một ma trận trực giao. Điều đó có nghĩa là ma trận cột tọa độ của hệ vector $\mathscr{B} = (f(\mathbf{a}_1), f(\mathbf{a}_2), \ldots, f(\mathbf{a}_n))$ đối với cở sở $\mathscr{A}$ là ma trận trực giao. Vì ma trận trực giao không suy biến nên $\mathscr{B}$ là một cơ sở của $V_E$. Hơn nữa theo Định lý 3.1.6 ta suy ra $\mathscr{B}$ là một cơ sở trực chuẩn của $V_E$. □

**5.1.5** Tập hợp các toán tử trực giao trên $V_E$ là một nhóm với phép hợp thành và gọi là **nhóm trực giao** trên $V_E$ ký hiệu là $\mathcal{O}(V_E)$.

## 5.2 Dạng chính tắc của một toán tử trực giao

**5.2.1 Định lý. (Dạng chính tắc của một toán tử trực giao)** *Cho $f$ là một toán tử tuyến tính trên không gian Euclid $n$ chiều. Để $f$ là một toán tử trực giao trên $V_E$ cần và đủ là tồn tại một cơ sở trực chuẩn của $V_E$ sao cho ma trận của $f$ trong cơ sở trực chuẩn đó là ma trận chia khối $M$ có dạng sau:*

$$M = \begin{bmatrix} M_1 & O & \dots & O & O & O & \dots & O \\ O & M_2 & \dots & O & O & O & \dots & O \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ O & O & \dots & M_k & O & O & \dots & O \\ O & O & \dots & O & \varepsilon_1 & 0 & \dots & 0 \\ O & O & \dots & O & 0 & \varepsilon_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ O & O & \dots & O & 0 & 0 & \dots & \varepsilon_{n-2k} \end{bmatrix} \in \mathcal{M}_n(\mathbb{R})$$

*ở đó $M_1, \dots, M_k$ là các khối $2 \times 2$*

$$M_i = \begin{bmatrix} \cos\varphi_i & -\sin\varphi_i \\ \sin\varphi_i & \cos\varphi_i \end{bmatrix} \qquad 0 < \varphi_i < 2\pi,$$

*$\varphi_i \neq \pi$, $i = 1, \dots, k$ $(0 \leq 2k \leq n)$; $\varphi_j = \pm 1$; $j = 1, \dots, n-2k$. Khi đó ta bảo $f$ có biểu thức tọa độ dạng chính tắc trong cơ sở trực chuẩn nói trên.*

**5.2.2** Trong không gian Euclid $n$ chiều $V_E$ cho hai cơ sở trực chuẩn $\mathscr{B}$, $\mathscr{B}'$ và $C$ là ma trận đổi cơ sở từ $\mathscr{B}$ sang $\mathscr{B}'$. Khi đó $C$ là ma trận trực giao cấp $n$, tức là $C^t = C^{-1}$, nói riêng $(\det C)^2 = \det(C^tC) = \det(I_n) = 1$ hay $\det C = \pm 1$. Ta bảo $\mathscr{B}$ và $\mathscr{B}'$ **cùng hướng** (hay **ngược hướng**) nếu $\det C = 1$ (hay $\det C = -1$). Quan hệ cùng hướng là một quan hệ tương đương trên tập hợp các cơ sở trực chuẩn của $V_E$. Hơn nữa chỉ có đúng hai lớp tương đương. Mỗi lớp gọi là một hướng của $V_E$. Nếu chọn một hướng là thuận (dương) thì hướng kia là nghịch (âm).

**5.2.3 Định nghĩa.** Hướng của $\mathbb{E}^n$ có đại diện là cơ sở chính tắc gọi là **hướng dương chính tắc** trên $\mathbb{E}^n$. Hướng còn lại là **hướng âm chính tắc** trên $\mathbb{E}^n$

**5.2.4 Định nghĩa.** Toán tử trực giao $f$ trên $V_E$ (hữu hạn chiều) gọi là **bảo toàn hướng** hay **đổi hướng** tùy vào việc $f$ biến một cơ sở trực chuẩn thành cơ sở cùng hướng hay ngược hướng với cơ sở đó.

**5.2.5 Mệnh đề.** *Nếu toán tử trực giao $f$ có ma trận trong một cơ sở trực chuẩn nào đó của $V_E$ là $M$ thì*

**(i)** *$f$ bảo toàn hướng khi và chỉ khi* $\det M = 1$.

**(ii)** *$f$ đổi hướng khi và chỉ khi* $\det M = -1$.

**5.2.6 Thí dụ.** **(a)** Cơ sở $((1,0),(0,-1))$ xác định hướng âm chính tắc trên $\mathbb{E}^2$ vì ma trận đổi cơ sở từ cơ sở chính tắc sang nó là $C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ có định thức bằng $-1$.

**(b)** Toán tử trực giao trên $\mathbb{E}^3$ cho bởi ma trận

$$M = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad 0 \le \varphi \le 2\pi$$

là bảo toàn hướng vì $\det M = 1$.

**(c)** Toán tử trực giao $g$ trên $\mathbb{E}^3$ cho bởi ma trận

$$N = \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ \sin\varphi & -\cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad 0 \le \varphi \le 2\pi$$

là đổi hướng vì $\det N = -1$. □

## Bài tập

**1)** Chứng minh rằng một đẳng cấu trực giao là bảo toàn góc.

**2)** Chứng minh rằng quan hệ "đẳng cấu" giữa các không gian Euclid là một quan hệ tương đương.

# § 6 Toán tử đối xứng

## 6.1 Khái niệm và tính chất

**6.1.1 Định nghĩa.** Cho $V_E$ là một không gian Euclid. Toán tử tuyến tính $f$ trên $V_E$ gọi là **đối xứng** nếu

$$\langle f(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, f(\mathbf{y})\rangle \qquad \text{với mọi } \mathbf{x}, \mathbf{y} \in V_E.$$

**6.1.2 Mệnh đề.** *Phép chiếu trực giao lên không gian con $U$ trong không gian Euclid $V_E$ là một toán tử đối xứng.*

*Chứng minh.* Theo giả thiết với $\mathbf{x}, \mathbf{y} \in V_E$ ta có biểu diễn duy nhất $\mathbf{x} = \mathbf{u} + \mathbf{v}$ và $\mathbf{y} = \mathbf{u}' + \mathbf{v}'$ với $\mathbf{u}, \mathbf{u}' \in U$ và $\mathbf{v}, \mathbf{v}' \in U^\perp$. Ta có, $P(\mathbf{x}) = \mathbf{u}$ và $P(\mathbf{y}) = \mathbf{u}'$, suy ra

$$\langle P(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{u}, \mathbf{u}' + \mathbf{v}'\rangle = \langle \mathbf{u}, \mathbf{u}'\rangle = \langle \mathbf{u} + \mathbf{v}, \mathbf{u}'\rangle = \langle \mathbf{x}, P(\mathbf{y})\rangle. \qquad \square$$

**6.1.3 Định lý.** *Để toán tử tuyến tính $f$ trên không gian Euclid $n$ chiều $V_E$ là đối xứng, cần và đủ là ma trận của $f$ trong một cơ sở trực chuẩn nào đó của $V_E$ là ma trận đối xứng.*

*Chứng minh.* Giả sử $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một cơ sở trực chuẩn của $V_E$ và $A = (a_{ij})_n$ là ma trận của $f$ trong cơ sở $\mathscr{A}$. Khi đó, ta có $f(\mathbf{a}_j) = \sum\limits_{i=1}^{n} a_{ij}\mathbf{a}_i$ và

$$\langle \mathbf{a}_i, f(\mathbf{a}_j)\rangle = \left\langle \mathbf{a}_i, \sum_{k=1}^{n} a_{kj}\mathbf{a}_k \right\rangle = \sum_{k=1}^{n} a_{kj}\langle \mathbf{a}_i, \mathbf{a}_k\rangle = a_{ij}$$

và

$$\langle f(\mathbf{a}_i), \mathbf{a}_j\rangle = \left\langle \sum_{k=1}^{n} a_{ki}\mathbf{a}_k, \mathbf{a}_j \right\rangle = \sum_{k=1}^{n} a_{ki}\langle \mathbf{a}_k, \mathbf{a}_j\rangle = a_{ji}.$$

Nếu $f$ là một toán tử đối xứng thì $\langle f(\mathbf{a}_i), \mathbf{a}_j\rangle = \langle \mathbf{a}_i, f(\mathbf{a}_j)\rangle$ hay $a_{ji} = a_{ij}$ với mọi $i, j = 1, \ldots, n$. Do đó $A$ là ma trận đối xứng.

Ngược lại, giả sử $A$ là một ma trận đối xứng. Khi đó, ta có $\langle f(\mathbf{a}_i), \mathbf{a}_j\rangle = \langle \mathbf{a}_i, f(\mathbf{a}_j)\rangle$. Với mọi $\mathbf{x}, \mathbf{y} \in V_E$ ta có biểu diễn $\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{a}_i$ và $\mathbf{y} = \sum_{j=1}^{n} \beta_j \mathbf{a}_j$. Ta tính được

$$\begin{aligned}
\langle f(\mathbf{x}), \mathbf{y}\rangle &= \Big\langle \sum_{i=1}^{n} \alpha_i f(\mathbf{a}_i), \sum_{j=1}^{n} \beta_j \mathbf{a}_j \Big\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\beta_j \langle f(\mathbf{a}_i), \mathbf{a}_j\rangle \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\beta_j \langle \mathbf{a}_i, f(\mathbf{a}_j)\rangle \\
&= \Big\langle \sum_{i=1}^{n} \alpha_i \mathbf{a}_i, \sum_{j=1}^{n} \beta_j f(\mathbf{a}_j) \Big\rangle \\
&= \langle \mathbf{x}, f(\mathbf{y})\rangle.
\end{aligned}$$

Vậy $f$ là một toán tử đối xứng. □

**6.1.4 Thí dụ.** Xét không gian Euclid 3 chiều chính tắc $\mathbb{E}^3$ với cơ sở trực chuẩn chính tắc $((1,0,0),(0,1,0),(0,0,1))$.

**(a)** Toán tử tuyến tính $f$ trên $\mathbb{E}^3$ cho bởi

$$f(x,y,z) = (x - 2y + 3z, -2x + 3y - 4z, 3x - 4y - z),$$

với mọi $(x,y,z) \in \mathbb{E}^3$, là một toán tử đối xứng vì ma trận của $f$ trong cơ sở trực chuẩn chính là ma trận đối xứng sau đây

$$A = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 3 & -4 \\ 3 & -4 & -1 \end{bmatrix}.$$

**(b)** Tương tự, toán tử tuyến tính $g$ sau đây cũng là một toán tử đối xứng trên $\mathbb{E}^3$

$$g(x,y,z) = (2x + y - z, x + 2y + 3z, -x + 3y + 4z). \qquad \square$$

**6.1.5 Nhận xét.** Đương nhiên trên mỗi không gian Euclid toán tử đồng nhất Id và bội tùy ý của nó $\lambda\mathrm{Id}$ ($\lambda \in \mathbb{R}$) là các toán tử đối xứng.

## 6.2 Chéo hóa ma trận đối xứng

**6.2.1 Định lý.** *Nếu $f$ là một toán tử đối xứng trên không gian Euclid $n$ chiều, thì $f$ có đủ $n$ giá trị riêng kể cả bội, các vector riêng ứng với giá trị riêng khác nhau đều trực giao với nhau, và tồn tại một cơ sở trực chuẩn của $V_E$ để $f$ có ma trận chéo trong cơ sở đó.*

*Chứng minh.* Gọi $A$ là ma trận của toán tử đối xứng $f$ trên không gian Euclid $V_E$ đối với một cơ sở nào đó. Ta biết rằng $A$ sẽ có $n$ giá trị riêng (kể cả bội) trên $\mathbb{C}$. Xét trên trường số phức $\mathbb{C}$. Giả sử $\lambda$ là một giá trị riêng của $A$ và $\mathbf{x} \in \mathbb{C}^n$ là một vector riêng của $A$ ứng với giá trị riêng $\lambda$, nghĩa là $A\mathbf{x} = \lambda\mathbf{x}$. Ta có (chú ý $A$ là ma trận thực đối xứng)

$$\overline{\overline{\mathbf{x}}^t A\mathbf{x}} = \overline{\overline{\mathbf{x}}}^t\overline{A}\overline{\mathbf{x}} = \mathbf{x}^t A\overline{\mathbf{x}} = (\mathbf{x}^t A\overline{\mathbf{x}})^t = \overline{\mathbf{x}}^t A\mathbf{x}.$$

Vậy $\overline{\mathbf{x}}^t A\mathbf{x}$ là một số thực. Do $A\mathbf{x} = \lambda\mathbf{x}$ ta có $\overline{\mathbf{x}}^t A\mathbf{x} = \lambda\overline{\mathbf{x}}^t\mathbf{x}$. Do đó $\lambda = \dfrac{\overline{\mathbf{x}}^t A\mathbf{x}}{\overline{\mathbf{x}}^t\mathbf{x}}$ là số thực. Vậy mọi giá trị riêng của $A$ đều là thực, cho nên toán tử đối xứng $f$ có $n$ giá trị riêng (kể cả bội)

Giả sử $\lambda_1 \neq \lambda_2$ là hai giá trị riêng phân biệt của toán tử đối xứng $f$. Gọi $\mathbf{x}$ là một vector riêng ứng với giá trị riêng $\lambda_1$ và $\mathbf{y}$ là một vector riêng ứng với giá trị riêng $\lambda_2$, nghĩa là $A\mathbf{x} = \lambda_1\mathbf{x}$ và $A\mathbf{y} = \lambda_2\mathbf{y}$. Ta có

$$\lambda_1\langle\mathbf{x}, \mathbf{y}\rangle = \langle\lambda_1\mathbf{x}, \mathbf{y}\rangle = \langle A\mathbf{x}, \mathbf{y}\rangle = \langle\mathbf{x}, A\mathbf{y}\rangle = \langle\mathbf{x}, \lambda_2\mathbf{y}\rangle = \lambda_2\langle\mathbf{x}, \mathbf{y}\rangle,$$

suy ra $(\lambda_1 - \lambda_2)\langle\mathbf{x}, \mathbf{y}\rangle = 0$. Do đó $\langle\mathbf{x}, \mathbf{y}\rangle = 0$ vì $\lambda_1 \neq \lambda_2$.

Ta chứng minh sự tồn tại của một cơ sở trực chuẩn của $V_E$ để toán tử đối xứng $f$ có ma trận chéo trong cơ sở đó theo số chiều của $V_E$. Khi $n = 1$ khẳng định hiển nhiên đúng với cơ sở trực chuẩn duy nhất. Giả sử luôn tồn tại cơ sở trực chuẩn của không gian Euclid $n-1$ chiều bất kỳ để toán tử đối xứng trên không gian ấy có ma trận chéo trong cơ sở ấy. Xét không gian Euclid $n$ chiều $V_E$ và một toán tử $f$ trên $V_E$. Gọi $\lambda$ là một giá trị riêng của $f$, và gọi $\mathbf{a}$ là một vector riêng đơn vị ứng với giá trị riêng

$\lambda$, tức là $f(\mathbf{a}) = \lambda\mathbf{a}$ và $\|\mathbf{a}\| = 1$. Đặt $U = \langle \mathbf{a} \rangle$. Ta có $V_E = U \oplus U^\perp$ và $\dim(U^\perp) = n - 1$. Với mọi $\mathbf{w} \in U^\perp$, ta có

$$\langle f(\mathbf{w}), \mathbf{a} \rangle = \langle \mathbf{w}, f(\mathbf{a}) \rangle = \langle \mathbf{w}, \lambda\mathbf{a} \rangle = \lambda \langle \mathbf{w}, \mathbf{a} \rangle = 0.$$

Vậy $f(\mathbf{w}) \in U^\perp$, cho nên $f(U^\perp) \subseteq U^\perp$. Do đó ánh xạ thu hẹp $f|_{U^\perp}$ xác định một ánh xạ $f|_{U^\perp} : U^\perp \to U^\perp$ và nó là một toán tử đối xứng trên không gian $U^\perp$. Theo giả thiết qui nạp tồn tại một cơ sở trực chuẩn $(\mathbf{a}_2, \ldots, \mathbf{a}_n)$ của $U^\perp$ sao cho ma trận của $f|_{U^\perp}$ đối với cơ sở này là ma trận chéo. Ta có thể thấy $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ là một cơ trực chuẩn và ma trận của $f$ đối với cơ sở này là ma trận chéo. □

**6.2.2 Định nghĩa.** Cho $f$ là một toán tử đối xứng trên không gian Euclid $n$ chiều $V_E$. Việc đi tìm một cơ sở trực chuẩn của $V_E$ để $f$ có ma trận chéo trong cơ sở đó gọi là phép **chéo hóa trực giao** toán tử đối xứng $f$.

Với mỗi ma trận thực đối xứng $A$ cấp $n$, ta cũng gọi việc đi tìm ma trận trực giao $C$ cấp $n$ sao cho $C^tAC$ có dạng chéo là **phép chéo hóa trực giao** $A$.

**6.2.3** Giả sử đã biết ma trận đối xứng $A \in \mathcal{M}_n(\mathbb{R})$ của toán tử đối xứng trong một cơ sở trực chuẩn $\mathscr{B}$ nào đó của không gian Euclid $n$ chiều $V_E$. Khi đó phép chéo hóa trực giao $f$ thực chất chính là phép chéo hóa trực giao ma trận $A$. Cụ thể nếu $C$ là ma trận chéo hóa trực giao $A$ thì $C$ chính là ma trận đổi cơ sở từ $\mathscr{B}$ sang cơ sở trực chuẩn $\mathscr{B}'$ để $f$ có ma trận chéo trong $\mathscr{B}'$.

**6.2.4 Thuật toán chéo hóa trực giao ma trận thực đối xứng**

Cho $A$ là một ma trận thực đối xứng tùy ý cấp $n$. Để chéo hóa trực giao $A$ ta thực hiện các bước sau đây.

**Bước 1.** Lập đa thức đặc trưng của $A$: $\chi(\lambda) = \det(A - \lambda I_n)$.

**Bước 2.** Giải phương trình đặc trưng $\chi(\lambda) = 0$ để tìm các giá trị riêng của $A$.

**Bước 3.** Với mỗi giá trị riêng $\lambda_i$ tìm cơ sở của không gian con riêng ứng với giá trị riêng $\lambda_i$. Dùng thuật toán Gram-Schmidt tìm cơ sở trực chuẩn của không gian con riêng ấy. Lập ma trận $C$ là ma trận các cột tọa độ của các vector trong các cơ sở trực chuẩn của các không gian con riêng. Ma trận $C$ chính là ma trận trực giao chéo hóa $A$ và $C^tAC$ là ma trận chéo với các phần tử trên đường chéo là các giá trị riêng tương ứng.

**6.2.5 Thí dụ.** Tìm ma trận trực giao chéo hóa ma trận $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

Đa thức đặc trưng của $A$ là

$$\chi(\lambda) = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = (\lambda-1)^2(4-\lambda).$$

Vậy $A$ có các giá trị riêng là $\lambda = 1$ (bội hai) và $\lambda = 4$.

Với $\lambda = 1$, ta biến đổi ma trận

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Vậy cơ sở của không gian con riêng $E(1)$ là $((-1,1,0),(-1,0,1))$. Dùng thuật toán Gram-Schmidt tìm cơ sở trực chuẩn của $E(1)$

$$\mathbf{u}_1 = (-\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0) \quad \mathbf{v}_2 = (-1,0,1) - \tfrac{1}{\sqrt{2}}(-\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0) = (-\tfrac{1}{2}, -\tfrac{1}{2}, 1).$$

Vậy một cơ sở trực chuẩn của $E(1)$ là $((-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}))$.

Với $\lambda = 4$, ta biến đổi ma trận

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -2 \\ 0 & 3 & -3 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Vậy $E(4) = \langle (1,1,1) \rangle$ và cơ sở trực chuẩn của $E(4)$ là $((\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}))$.

Ma trận trực giao chéo hóa $A$ là

$$C = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

và ma trận chéo tương ứng của $A$ là $C^tAC = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. □

**6.2.6 Thí dụ.** Cho $f$ là toán tử đối xứng trên $\mathbb{E}^3$ xác định bởi

$$f(x, y, z) = (5x - y + 2z, -x + 5y + 2z, 2x + 2y + 2z).$$

Tìm một cơ sở trực chuẩn của $\mathbb{E}^3$ sao cho trong cơ sở đó $f$ có ma trận chéo.

*Giải.* Ma trận của $f$ trong cơ sở chính tắc là $A = \begin{bmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix}$. Ta sẽ thực hiện chéo hóa ma trận $A$. Đa thức đặc trưng của ma trận là

$$\chi(\lambda) = \begin{vmatrix} 5-\lambda & -1 & 2 \\ -1 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} = -\lambda(\lambda - 6)^2$$

nên $A$ có hai giá trị riêng là 0 và 6. Với $\lambda = 0$, ta biến đổi sơ cấp ma trận

$$\begin{bmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 6 & 3 \\ 0 & -6 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Vậy không gian con riêng $E(0) = \langle (1, 1, -2) \rangle$ và cơ sở trực chuẩn của $E(0)$ là $(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}})$.

Với $\lambda = 6$, ta biến đổi ma trận

$$\begin{bmatrix} -1 & -1 & 2 \\ -1 & -1 & 2 \\ 2 & 2 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Vậy không gian con riêng $E(6) = \langle(-1,1,0),(2,0,1)\rangle$. Trực chuẩn hóa cơ sở của $E(6)$: $\mathbf{w}_1 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$,

$$\mathbf{v}_1 = (2,0,1) - \frac{1}{2}\langle(-1,1,0),(2,0,1)\rangle(-1,1,0) = (1,1,1)$$

suy ra $\mathbf{w}_2 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. Vậy $E(6) = \langle(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\rangle$.

Do đó ma trận chéo hóa trực giao là $\begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{bmatrix}$ và ma trận chéo tương ứng là $D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}$.

Vậy $\mathscr{B} = ((\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}), (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}))$ là một cơ sở trực chuẩn của $\mathbb{E}^3$ mà $f$ có ma trận trong $\mathscr{B}$ là $D$. □

## Bài tập

**1)** Tìm ma trận chéo hóa trực giao các ma trận sau đây và ma trận chéo tương ứng

**(a)** $\begin{bmatrix} 5 & 6 \\ 6 & 0 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$ **(c)** $\begin{bmatrix} 7 & 3\sqrt{5} \\ 3\sqrt{5} & 3 \end{bmatrix}$

**(d)** $\begin{bmatrix} 5 & -6 & -3 \\ -6 & 9 & 0 \\ -3 & 0 & 9 \end{bmatrix}$ **(e)** $\begin{bmatrix} 4 & -2 & 6 \\ -2 & 1 & -3 \\ 6 & -3 & 9 \end{bmatrix}$ **(f)** $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & -2\sqrt{2} \\ 0 & -2\sqrt{2} & 3 \end{bmatrix}$

**(g)** $\begin{bmatrix} 3 & -1 & 2 \\ -1 & 3 & 2 \\ 2 & 2 & 0 \end{bmatrix}$. **(h)** $\begin{bmatrix} 5 & -1 & 1 \\ -1 & 2 & -2 \\ 1 & -2 & 2 \end{bmatrix}$ **(i)** $\begin{bmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{bmatrix}$

**2)** Tìm cơ sở $p$-chính tắc của các dạng toàn phương trên $\mathbb{R}^3$ xác định bởi

**(a)** $p(x_1, x_2, x_3) = 5x_1^2 + 9x_2^2 + 9x_3^2 - 12x_1x_2 - 6x_1x_3$

**(b)** $p(x_1, x_2, x_3) = 4x_1^2 + x_2^2 + 9x_3^2 + 4x_1x_2 - 6x_2x_3 + 12x_1x_2$

# § 7 Dạng toàn phương trên không gian Euclid

## 7.1 Dạng chính tắc trực giao của dạng toàn phương

**7.1.1 Định lý.** *Cho $p$ là một dạng toàn phương trên không gian Euclid $n$ chiều $V_E$. Khi đó ta có*

**(i)** *Tồn tại một cơ sở trực chuẩn của $V_E$ để trong nó $p$ có dạng chính tắc*

$$p(\mathbf{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

**(ii)** *Nếu trong một cơ sở trực chuẩn khác của $V_E$, $p$ lại có dạng chính tắc*

$$p(\mathbf{x}) = \beta_1 x_1'^2 + \beta_2 x_2'^2 + \cdots + \beta_n x_n'^2$$

*thì $\{\lambda_1, \lambda_2, \ldots, \lambda_n\} = \{\beta_1, \beta_2, \ldots, \beta_n\}$ (cho phép sự lặp lại các phần tử).*

**7.1.2 Định nghĩa.** Dạng chính tắc trong định lý trên của dạng toàn phương $p$ trong một cơ sở trực chuẩn nào đó gọi là **dạng chính tắc trực giao** của $p$. Dạng chính tắc trực giao của $p$ là duy nhất không kể đến thứ tự của các hệ số chính tắc.

## 7.2 Đưa dạng toàn phương về dạng chính tắc trực giao

**7.2.1 Định nghĩa.** Cho $p$ là một dạng toàn phương trên không gian Euclid $n$ chiều $V_E$. Việc đi tìm một cơ sở trực chuẩn của $V_E$ để $p$ có dạng chính tắc trực giao gọi là **phép đưa $p$ về dạng chính tắc trực giao** hay **phép chính tắc hóa trực giao $p$**.

**7.2.2 Thuật toán chính tắc hóa trực giao.** Cho dạng toàn phương $p$ trên không gian Euclid $n$ chiều $V_E$ và $A \in \mathcal{M}_n(\mathbb{R})$ là ma trận đối xứng của $p$ trong một cơ sở trực chuẩn nào đó của $V_E$. Phép đưa $p$ về dạng chính tắc trực giao thực chất chính là phép chéo hóa trực giao ma trận đối xứng $A$ của $f$. Do đó ta tiến hành các bước như trong Thuật toán 6.2.4 với lưu ý rằng cần dựa vào ma trận trực giao $C$ ở bước 3 để chỉ ra phép đổi biến hay đổi cơ sở mà $p$ có dạng chính tắc trực giao.

**7.2.3 Thí dụ.** Trên $\mathbb{E}^3$ cho dạng toàn phương $p(x,y,z) = 2xy + 2xz + 2yz$. Hãy đưa $p$ về dạng chính tắc trực giao.

*Giải.* Trong cơ sở trực chuẩn chính tắc của $\mathbb{E}^3$ ma trận của $p$ là

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

Đa thức đặc trưng của $A$ là

$$\chi(\lambda) = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = (\lambda+1)^2(2-\lambda).$$

Do đó $A$ có giá trị riêng là $\lambda = 2$ và $\lambda = -1$. Khi $\lambda = 2$, ta biến đổi ma trận

$$\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 1 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó không gian con riêng $E(2) = \langle (1,1,1) \rangle$ và cơ sở trực chuẩn của nó là $((\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$.

Khi $\lambda = -1$, ta biến đổi ma trận

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Do đó không gian conn riêng $E(-1) = \langle (-1,1,0), (-1,0,1) \rangle$. Trực chuẩn hóa cơ sở của $E(-1)$ ta được $\mathbf{w}_1 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$ và

$$\mathbf{v}_1 = (-1,0,1) - \frac{1}{2}\langle (-1,0,1), (-1,1,0) \rangle (-1,1,0) = (-\frac{1}{2}, -\frac{1}{2}, 1)$$

nên $\mathbf{w}_2 = (-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}})$.

Vậy với cơ sở $\mathscr{B} = ((\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}), (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}))$ dạng song tuyến tính $p$ có dạng chính tắc trực giao

$$p(\mathbf{x}) = 2x_1'^2 - x_2'^2 - x_3'^2$$

trong đó $\mathbf{x}|_{\mathscr{B}} = (x_1', x_2', x_3')$. □

**7.2.4 Nhận xét.** **(a)** So sánh với thuật toán Lagrange, thuật toán dùng các phép biến đổi sơ cấp hay thuật toán Jacobi để đưa một dạng toàn phương trong không gian Euclid về dạng *chính tắc*, ta thấy phép đưa dạng toàn phương về dạng *chính tắc* trực giao nói chung phức tạp hơn. Điều này cũng dễ hiểu vì chúng ta đòi hỏi cơ sở được chọn (để $p$ có dạng chính tắc) phải là *cơ sở trực chuẩn*.

**(b)** Không gian Euclid khác với không gian vector bất kỳ ở chỗ trong nó có một tích vô hướng, do đó có thêm nhiều khái niệm, tính chất liên quan đến tích vô hướng như độ dài, góc, .... Việc tính toán độ dài, góc trở nên đơn giản và thuận tiện khi cơ sở được chọn là cơ sở trực chuẩn. Đó là lý do tại sao trong không gian Euclid (hữu hạn chiều) chúng ta chỉ quan tâm đến cơ sở trực chuẩn. Dạng chính tắc của mỗi dạng toàn phương cho phép ta tính toán về nó đơn giản nhưng sẽ là vô ý nghĩa nếu cơ sở được chọn cho dạng chính tắc đó không phải là cơ sở trực chuẩn. Bởi vậy, phép đưa dạng toàn phương về dạng chính tắc trực giao tuy phức tạp hơn các thuật toán chính tắc hóa khác nhưng là cần thiết và có ý nghĩa trong không gian Euclid hữu hạn chiều.

## Bài tập

**1**) Tìm phép biến đổi trực giao đưa các dạng toàn phương sau đây về dạng chính tắc trực giao và viết các dạng đó.

**(a)** $x_1^2 + x_2^2 + 4x_1x_2$

**(b)** $5x_1^2 + 12x_1x_2$

**(c)** $7x_1^2 + 3x_2^2 + 6\sqrt{5}x_1x_2$

**(d)** $2x_1^2 + x_2^2 + 3x_3^2 - 4\sqrt{2}x_2x_3$

**(e)** $5x_1^2 + 9x_2^2 + 9x_3^2 - 12x_1x_2 - 6x_1x_3$

**(f)** $4x_1^2 + x_2^2 + 9x_3^2 + 4x_1x_2 - 6x_2x_3 + 12x_1x_3$

**(g)** $5x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 4x_2x_3 + 2x_1x_3$

**2**) Tìm phép biến đổi trực giao đưa các dạng toàn phương sau đây về dạng chính tắc trực giao và viết các dạng đó.

**(a)** $\sum_{i=1}^{n} x_i^2 + \sum_{i<j}^{n} x_i x_j$ **(b)** $\sum_{i<j}^{n} x_i x_j$

# § 8 Đường bậc hai trên mặt phẳng và mặt bậc hai trong không gian

Không gian Euclid chính tắc $\mathbb{E}^2$ (tương ứng $\mathbb{E}^3$) có thể đồng nhất với mặt phẳng tọa độ (tương ứng, không gian tọa độ) Descartes vuông góc trong hình học giải tích sơ cấp. Mục này trình bày về lý thuyết các đường bậc hai tổng quát trên mặt phẳng và lý thuyết các mặt bậc hai tổng quát trong không gian trên cơ sở lý thuyết các dạng toàn phương trong $\mathbb{E}^2$ và $\mathbb{E}^3$.

## 8.1 Tích vô hướng của hai vector

**8.1.1** Khác với không gian vector tổng quát, không gian các vector tự do trên mặt phẳng hay trong không gian của hình học sơ cấp đã có khái niệm "độ dài" như một khái niệm nguyên thủy. Bởi vậy cần định nghĩa tích vô hướng sao cho nó sinh ra độ dài đã có đó.

**8.1.2 Định nghĩa.** Cho $\overrightarrow{a}$, $\overrightarrow{b}$ là hai vector tự do tùy ý trong mặt phẳng hay trong không gian hình học sơ cấp. **Tích vô hướng** của $\overrightarrow{a}$ và $\overrightarrow{b}$ là một số thực, ký hiệu $\overrightarrow{a}\overrightarrow{b}$, được xác định như sau:

$$\overrightarrow{a}\overrightarrow{b} = \frac{1}{4}(|\overrightarrow{a} + \overrightarrow{b}|^2 - |\overrightarrow{a} - \overrightarrow{b}|^2) \tag{8.1.3}$$

Theo Định lý 1.2.5 trang 414, dễ thấy tích vô hướng này sinh ra độ dài đã biết.

**8.1.4 Mệnh đề.** **(i)** *Tích vô hướng định nghĩa như trên thỏa 4 tiên đề của một tích vô hướng tổng quát.*

**(ii)** $\overrightarrow{a}\overrightarrow{b} = \frac{1}{2}(|\overrightarrow{a} + \overrightarrow{b}|^2 - |\overrightarrow{a}|^2 - |\overrightarrow{b}|^2)$ *với mọi* $\overrightarrow{a}, \overrightarrow{b}$

$\overrightarrow{a}\overrightarrow{b} = |\overrightarrow{a}| \cdot |\overrightarrow{b}| \cos(\widehat{\overrightarrow{a}, \overrightarrow{b}})$ *với mọi* $\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{a} \neq \overrightarrow{0} \neq \overrightarrow{b}$.

**(iii)** $\sqrt{\overrightarrow{a}^2} = |\overrightarrow{a}|$, *với mọi vector* $\overrightarrow{a}$.

**(iv)** $\overrightarrow{a} \perp \overrightarrow{b}$ *khi và chỉ khi* $\overrightarrow{a}\overrightarrow{b} = 0$.

**(v)** *Cặp vector đơn vị chỉ phương hai trục tọa độ trên mặt phẳng tọa độ Descartes vuông góc là một cơ sở trực chuẩn của không gian Euclid các vector tự do trên mặt phẳng. (Một khẳng định tương tự cũng có đối với bộ ba vector đơn vị chỉ phương các trục tọa độ trong không gian tọa độ Descartes vuông góc.*

## 8.2 Đường bậc hai trên mặt phẳng tọa độ Descartes vuông góc

**8.2.1 Định nghĩa.** Trên mặt phẳng tọa độ Descartes vuông góc $Oxy$, đường bậc hai là tập hợp tất cả các điểm $M(x, y)$ có tọa độ $(x, y)$ thỏa mãn một phương trình đại số bậc hai đối với $x, y$:

(8.2.2) $$f(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

trong đó hệ số $A, B, C, D, E, F$ là các số thực; $A, B, C$ không đồng thời bằng không. Phương trình (8.2.2) được gọi là **phương trình** (tổng quát) của đường bậc hai nói trên trong hệ tọa độ đã chọn. Hai đường bậc hai được xem là *trùng nhau* nếu các hệ số tương ứng trong phương trình của chúng tỷ lệ với nhau.

**8.2.3 Nhận xét.** Định nghĩa trên là hợp lý tức là không phụ thuộc vào hệ tọa độ đã chọn. Có thể có những đường bậc hai không chứa một điểm nào, chẳng hạn đường bậc hai $\mathscr{C}$ mà phương trình tổng quát là

$$x^2 + y^2 + 1 = 0.$$

Tuy nhiên phương trình trên xác định một tập hợp nghiệm không rỗng trên trường số phức. Bởi vậy, ta cũng gọi $\mathscr{C}$ là một *đường bậc hai ảo*.

**8.2.4** Xét các ma trận đối xứng $Q = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$ và $\overline{Q} = \begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix}$.
Chúng được gọi tương ứng là **ma trận nhỏ và lớn** của đường bậc hai cho bởi phương trình (8.2.2). Vì $A, B, C$ không đồng thời bằng không nên $0 < \operatorname{rank}(Q) \leq 2$. Đường bậc hai đang xét gọi là **suy biến** hay **không suy biến** tùy vào ma trận lớn $\overline{Q}$ suy biến hay không suy biến> Tính suy biến hay không suy biến cũng không phụ thuộc vào tọa độ.

**8.2.5 Định lý.** *Đối với mọi đường bậc hai $\mathscr{C}$ trên mặt phẳng, luôn tìm được một hệ tọa độ Descates vuông góc thích hợp sao cho phương trình của nó trong hệ tọa độ đó có một (và chỉ một) trong các dạng sau:*

**(1)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ *đường ellipse*

**(2)** $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ *đường hyperbola*

**(3)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$ *đường ellipse ảo*

**(4)** $x^2 + y^2 = 0$ *cặp đường thẳng ảo cắt nhau (một điểm)*

**(5)** $x^2 - y^2 = 0$ *cặp đường thẳng cắt nhau*

**(6)** $y^2 = 2px$ *(*$p > 0$*) đường parabola*

**(7)** $y^2 = a$ *(*$a > 0$*) cặp đường thẳng song song*

**(8)** $y^2 = -a$ *(*$a > 0$ *cặp đường thẳng ảo song song*

**(9)** $y^2 = 0$ *cặp đường thẳng trùng nhau*

*Các dạng phương trình trên gọi là phương trình chính tắc của đường bậc hai $\mathscr{C}$. Hệ tọa độ Descartes vuông góc mà $\mathscr{C}$ có phương trình chính tắc trong nó gọi là hệ tọa độ chính tắc đối với $\mathscr{C}$. Việc đi tìm hệ tọa độ chính tắc đối với $\mathscr{C}$ gọi là phép lập phương trình chính tắc của $\mathscr{C}$.*

**8.2.6 Nhận xét.** **(a)** Trên mặt phẳng chỉ có những đường bậc hai sau đây: đường ellipse, đường hyperbolic, đường parabolic, cặp đường thẳng cắt nhau, cặp đường thẳng song song, cặp đường thẳng trùng nhau, một điểm hoặc tập hợp rỗng.

**(b)** Đường bậc hai $\mathscr{C}$ không suy biến khi và chỉ khi phương trình chính tắc của nó thuộc một trong các dạng (1), (2), (3), (6). Nói riêng, nếu $\mathscr{C} \neq \varnothing$ và không suy biến thì $\mathscr{C}$ chính là một đường conic.

Các cặp đường thẳng và chỉ chúng là những đường bậc hai suy biến khác rỗng.

**8.2.7 Định lý. (Nhận biết các đường conic)** *Cho $\mathscr{C}$ là một đường bậc hai không suy biến trên mặt phẳng tọa độ Descartes vuông góc với phương trình* (8.2.2). *Khi đó*

**(i)** *$\mathscr{C}$ là đường ellipse hay đường ellipse ảo khi và chỉ khi $\Delta = AC - B^2 > 0$.*

**(ii)** *$\mathscr{C}$ là đường hyperbola khi và chỉ khi $\Delta = AC - B^2 < 0$*

**(iii)** *$\mathscr{C}$ là đường parabola khi và chỉ khi $\Delta = AC - B^2 = 0$.*

**(iv)** *Nói riêng, $\mathscr{C}$ là đường tròn hay đường tròn ảo khi và chỉ khi $A = C \neq 0$, $B = 0$.*

**8.2.8 Thí dụ.** Hãy xác định dạng của đường bậc hai

$$\mathscr{C} : x^2 - 4xy + 4y^2 + 4x - 3y - 7 = 0.$$

*Giải.* Ma trận nhỏ và lớn của $\mathscr{C}$ lần lượt là

$$Q = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \qquad \overline{Q} = \begin{bmatrix} 1 & -2 & 2 \\ -2 & 4 & -\frac{3}{2} \\ 2 & -\frac{3}{2} & 7 \end{bmatrix}.$$

Do $\det(\overline{Q}) = -\frac{25}{4}$ nên $\operatorname{rank}(\overline{Q}) = 3$. Do đó $\mathscr{C}$ không suy biến. Mặt khác, ta có $\Delta = \det(Q) = 0$, cho nên $\mathscr{C}$ là đường parabola. □

**8.2.9 Phương pháp lập phương trình chính tắc của đường bậc hai.** Giả sử trên mặt phẳng tọa độ Descartes vuông góc đã cho đường bậc hai $\mathscr{C}$ với phương trình tổng quát

$$f(x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0.$$

Xét dạng toàn phương hai biến thực

$$p(x, y) = Ax^2 + 2Bxy + Cy^2$$

xác định bởi phần đẳng cấp bậc hai của $f(x, y)$. Để lập phương trình chính tắc của $\mathscr{C}$ ta thực hiện các bước sau:

**Bước 1.** Chính tắc hóa trực giao $p$ nhờ một phép quay thích hợp hệ tọa độ đang xét.

**Bước 2.** Tính tiến hệ tọa độ một cách thích hợp để phương trình của $\mathscr{C}$ có dạng chính tắc.

**8.2.10 Thí dụ.** Lập phương trình chính tắc của đường bậc hai $\mathscr{C}$ cho bởi phương trình $5x^2 + 4xy + 8y^2 - 32x - 56y + 80 = 0$ trong hệ tọa độ $Oxy$.

*Giải.* Xét dạng toàn phương $p(x, y) = 5x^2 + 4xy + 8y^2$. Ma trận đối xứng của $p$ trong cơ sở chính tắc là $Q = \begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix}$ nó có đa thức đặc trưng:

$$\chi_Q(\lambda) = \begin{vmatrix} 5-\lambda & 2 \\ 2 & 8-\lambda \end{vmatrix} = (\lambda - 4)(\lambda - 9).$$

Vậy $Q$ có hai giá trị riêng là $\lambda = 4$ và $\lambda = 9$. Với $\lambda = 4$ ta biến đổi ma trận

$$Q - 4I = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

Từ đó ta được không gian con riêng $E(4) = \langle(-2, 1)\rangle$ với cơ sở trực chuẩn là $((-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}))$. Với $\lambda = 9$ ta biến đổi ma trận

$$Q - 9I = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \to \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}.$$

Từ đó ta được không gian con riêng $E(9) = \langle(1,2)\rangle$ với cơ sở trực chuẩn là $((\frac{1}{\sqrt5}, \frac{2}{\sqrt5}))$. Vậy ma trận chéo hóa trực giao $Q$ là

$$C = \begin{bmatrix} -\frac{2}{\sqrt5} & \frac{1}{\sqrt5} \\ \frac{1}{\sqrt5} & \frac{2}{\sqrt5} \end{bmatrix}.$$

Khi đó ta có $C^tQC = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$. Do đó với phép biến đổi tọa độ

$$\begin{cases} x = -\frac{2}{\sqrt5}x' + \frac{1}{\sqrt5}y' \\ y = \frac{1}{\sqrt5}x' + \frac{2}{\sqrt5}y' \end{cases}$$

dạng toàn phương $p$ được viết lại $p(x,y) = 4x'^2 + 9y'^2$, cho nên phương trình của $\mathscr{C}$ được viết lại là

$$4x'^2 + 9y'^2 + \frac{8}{\sqrt5}x' - \frac{144}{\sqrt5}y' + 80 = 0$$

Biến đổi tương đương phương trình này ta được các phương trình sau

$$4\Big(x'^2 + \frac{2}{\sqrt5}x' + \frac{1}{5}\Big) + 9\Big(y'^2 - \frac{16}{\sqrt5} + \frac{64}{5}\Big) + 80 - \frac{4}{5} - \frac{9\cdot 64}{5} = 0$$

$$4\Big(x' + \frac{1}{\sqrt5}\Big)^2 + 9\Big(y' - \frac{8}{\sqrt5}\Big)^2 = 36$$

$$\frac{\big(x' + \frac{1}{\sqrt5}\big)^2}{9} + \frac{\big(y' - \frac{8}{\sqrt5}\big)^2}{4} = 1$$

Do đó dùng phép tịnh tiến hệ tọa độ

$$\begin{cases} X = x' + \frac{1}{\sqrt5} \\ Y = y' - \frac{8}{\sqrt5} \end{cases}$$

thì $\mathscr{C}$ có phương trình chính tắc $\dfrac{X^2}{9} + \dfrac{Y^2}{4} = 1$. Như vậy $\mathscr{C}$ là một ellipse.

Tóm lại, nếu ta dùng phép đổi tọa độ

$$\begin{cases} x = -\frac{2}{\sqrt5}(X - \frac{1}{\sqrt5}) + \frac{1}{\sqrt5}(Y + \frac{8}{\sqrt5}) = -\frac{2}{\sqrt5}X + \frac{1}{\sqrt5}Y + 2 \\ y = \frac{1}{\sqrt5}(X - \frac{1}{\sqrt5}) + \frac{2}{\sqrt5}(Y + \frac{8}{\sqrt5}) = \frac{1}{\sqrt5}X + \frac{2}{\sqrt5}Y + 3 \end{cases}$$

thì $\mathscr{C}$ có phương trình chính tắc $\frac{X^2}{9} + \frac{Y^2}{4} = 1$ và $\mathscr{C}$ là ellipse có bán kính trục lớn $a = 3$ và bán kính trục nhỏ $b = 2$; và tâm đối xứng là $I(2, 3)$ (trong hệ tọa độ $Oxy$). $\square$

## 8.3 Mặt bậc hai trong không gian tọa độ Descartes vuông góc

**8.3.1 Định nghĩa.** Trong không gian tọa độ Descartes vuông góc $Oxyz$, **mặt bậc hai** là tập hợp các điểm $M(x, y, z)$ có tọa độ $(x, y, z)$ thỏa mãn phương trình đại số bậc hai đối với $x, y, z$:

$$\begin{aligned} f(x,y,z) = Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2 \\ +2Gx + 2Hy + 2Kz + L = 0 \end{aligned} \tag{8.3.2}$$

ở đó $A, B, C, D, E, F, G, H, K, L$ là các số thực và $A, B, C, D, E, F$ không đồng thời triệt tiêu. Phương trình (8.3.2) gọi là **phương trình (tổng quát)** của mặt bậc hai nói trên trong hệ tọa độ đã chọn. Hai mặt bậc hai được xem là **trùng nhau** nếu các hệ số tương ứng trong phương trình của chúng tỷ lệ với nhau.

**8.3.3 Nhận xét.** Tương tự như các đường bậc hai, ta cũng nhận thấy định nghĩa trên là hợp lý, tức là không phụ thuộc vào hệ tọa độ đã chọn. Hơn nữa, có thể có những mặt bậc hai không chứa điểm nào (rỗng). Chúng cũng được gọi là các *mặt bậc hai ảo*.

**8.3.4 Định nghĩa.** Các ma trận đối xứng thực

$$Q = \begin{bmatrix} A & B & C \\ B & D & E \\ C & E & F \end{bmatrix} \qquad \overline{Q} = \begin{bmatrix} A & B & C & G \\ B & D & E & H \\ C & E & F & K \\ G & H & K & L \end{bmatrix}$$

tương ứng được gọi là **ma trận nhỏ và lớn** của mặt bậc hai cho bởi phương trình (8.3.2). Vì $A, B, C, D, E, F$ không đồng thời bằng không nên $0 < \operatorname{rank}(Q) \leq 3$.

Mặt bậc hai đang xét gọi là **suy biến** (tương ứng **không suy biến**) nếu ma trận lớn $\overline{Q}$ của nó suy biến (tương ứng không suy biến).

**8.3.5 Định lý.** *Đối với mặt bậc hai $\mathscr{S}$ trong không gian, luôn tìm được một tọa độ Descartes vuông góc thích hợp để phương trình của $\mathscr{S}$ trong hệ tọa độ đó có một (và chỉ một) trong các dạng sau đây:*

**(1)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ *mặt ellipsoid*

**(2)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = -1$ *mặt ellipsoid ảo ($\varnothing$)*

**(3)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 0$ *mặt nón ảo (một điểm)*

**(4)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$ *mặt hyperboloid 1 tầng*

**(5)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = -1$ *mặt hyperboloid 2 tầng*

**(6)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 0$ *mặt nón elliptic*

**(7)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 2z$ *mặt paraboloid elliptic*

**(8)** $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 2z$ *mặt paraboloid hyperbolic (mặt yên ngựa)*

**(9)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ *mặt trụ elliptic*

**(10)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$ *mặt trụ elliptic ảo ($\varnothing$)*

**(11)** $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$ *cặp mặt phẳng ảo liên hợp (một đường thẳng)*

**(12)** $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ *mặt trụ hyperbolic*

**(13)** $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0$ *cặp mặt phẳng cắt nhau*

**(14)** $y^2 = 2px$ *mặt trụ parabolic*

**(15)** $x^2 = a$ *(*$a > 0$*) cặp mặt phẳng song song*

**(16)** $x^2 = -a$ *(*$a > 0$*) cặp mặt phẳng ảo song song (*$\varnothing$*)*

**(17)** $x^2 = 0$ *cặp mặt phẳng trùng nhau*

**8.3.6 Định nghĩa.** Các dạng phương trình (1),…, (17) trong định lý trên gọi là **phương trình chính tắc của mặt bậc hai** $\mathscr{S}$. Hệ tọa độ Descartes vuông góc mà $\mathscr{S}$ có phương trình chính tắc trong nó gọi là **hệ tọa độ chính tắc** đối với $\mathscr{S}$. Việc đi tìm hệ tọa độ chính tắc đối với $\mathscr{S}$ gọi là **phép lập phương trình chính tắc** của $\mathscr{S}$.

**8.3.7 Nhận xét.** **(i)** Trong không gian chỉ có các mặt bậc hai sau đây: mặt ellipsoid, mặt hyperboloid (một tầng hay hai tầng), mặt paraboloid (elliptic hay hyperbolic), mặt nón elliptic, mặt trụ (elliptic, hyperbolic, hay parabolic) cặp mặt phẳng (căt nhau, song song hay trùng nhau), một đường thẳng hoặc tập hợp rỗng.

**(ii)** Mặt bậc hai $\mathscr{S}$ không suy biến khi và chỉ khi nó có phương trình chính tắc thuộc một trong các dạng (1), (2), (4), (5), (7), (8). Nói riêng, nếu $\mathscr{S} \neq \varnothing$ và không suy biến thì $\mathscr{S}$ chính là một trong các mặt ellipsoid, hyperboloid, paraboloid. Các mặt nón, mặt trụ, cặp mặt phẳng, một đường thẳng (và chỉ chúng) là những mặt bậc hai suy biến không rỗng.

**8.3.8 Định lý. (Nhận biết các mặt bậc hai không suy biến)** *Cho* $\mathscr{S}$ *là một mặt bậc hai không suy biến trong không gian tọa độ Descartes vuông góc với phương trình* (8.3.2). *Gọi*

$$p(x, y, z) = Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2$$

*là dạng toàn phương ba biến thực cho bởi phần đẳng cấp bậc hai trong vế trái của phương trình* (8.3.2). *Khi đó*

**(i)** $\mathscr{S}$ *là mặt ellipsoid (kể cả mặt ellipsoid ảo) khi và chỉ khi* $p$ *xác định dương hoặc âm.*

**(ii)** *$\mathscr{S}$ là mặt hyperboloid (1 tầng hay 2 tầng) khi và chỉ khi hiệu chỉ chỉ số dương quán tính và âm quán tính là $\pm 1$.*

**(iii)** *$\mathscr{S}$ là mặt paraboloid (elliptic hay hyperbolic) khi và chỉ khi $p$ suy biến.*

**(iv)** *Nói riêng, $\mathscr{S}$ là mặt cầu (kể cả mặt cầu ảo) khi và chỉ khi $A = D = F \neq 0$ và $B = C = E = 0$.*

**8.3.9 Phương pháp lập phương trình chính tắc của mặt bậc hai.** Giả sử trong không gian tọa độ Descartes vuông góc đã cho mặt bậc hai $\mathscr{S}$ với phương trình tổng quát

$$\begin{aligned} f(x,y,z) = Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2 \\ +2Gx + 2Hy + 2Kz + L = 0. \end{aligned}$$

Xét dạng toàn phương ba biến thực

$$p(x,y,z) = Ax^2 + 2Bxy + 2Cxz + Dy^2 + 2Eyz + Fz^2$$

xác định bởi phần đẳng cấp bậc hai của $f(x,y,z)$. Để lập phương trình chính tắc của $\mathscr{S}$ ta thực hiện các bước sau

**Bước 1.** Chính tắc hóa trực giao dạng toàn phương $p$

**Bước 2.** Tịnh tiến hệ tọa độ một cách thích hợp để phương trình của $\mathscr{S}$ có dạng chính tắc.

**8.3.10 Thí dụ.** Hãy xác định dạng của mặt bậc hai $\mathscr{S}$ sau đây rồi lập phương trình chính tắc của nó

$$22x^2 + 8xy + 28y^2 + 15z^2 - 112x - 184y - 30z + 343 = 0.$$

*Giải.* Ma trận nhỏ và lớn của $\mathscr{S}$ lần lượt là

$$Q = \begin{bmatrix} 22 & 4 & 0 \\ 4 & 14 & 0 \\ 0 & 0 & 15 \end{bmatrix} \qquad \overline{Q} = \begin{bmatrix} 22 & 4 & 0 & -56 \\ 4 & 14 & 0 & -92 \\ 0 & 0 & 15 & -15 \\ -56 & -92 & -15 & 343 \end{bmatrix}$$

Ta tính được $\det(\overline{Q}) = -540000$, cho nên $\mathscr{S}$ không suy biến. Xét dạng toàn phương ba biến

$$p(x,y,z) = 22x^2 + 8xy + 28y^2 + 15z^2$$

xác định bởi phần đẳng cấp bậc hai trong vế trái của phương trình của $\mathscr{S}$.

$Q$ chính là ma trận của $p$ trong cơ sở trực giao chính tắc. Do

$$D_1 = 22, \quad D_2 = \begin{vmatrix} 22 & 4 \\ 4 & 28 \end{vmatrix} = 600, \quad D_3 = \begin{vmatrix} 22 & 4 & 0 \\ 4 & 28 & 0 \\ 0 & 0 & 15 \end{vmatrix} = 9000$$

nên $p$ xác định dương theo tiêu chuẩn Sylvester. Do đó $\mathscr{S}$ là một mặt ellipsoid.

Đa thức đặc trưng của ma trận $Q$ là

$$\chi_Q(\lambda) = \begin{vmatrix} 22-\lambda & 4 & 0 \\ 4 & 28-\lambda & 0 \\ 0 & 0 & 15-\lambda \end{vmatrix} = -(\lambda-30)(\lambda-20)(\lambda-15).$$

Vậy $Q$ có các giá trị riêng $\lambda = 30$, $\lambda = 20$, $\lambda = 15$. Với $\lambda = 30$ ta biến đổi ma trận

$$Q - 30I_3 = \begin{bmatrix} -8 & 4 & 0 \\ 4 & -2 & 0 \\ 0 & 0 & -15 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

suy ra không gian con $E(30) = \langle(1,2,0)\rangle$, nó có cơ sở trực chuẩn $((\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0))$. Với $\lambda = 20$, ta biến đổi ma trận

$$Q - 20I_3 = \begin{bmatrix} 2 & 4 & 0 \\ 4 & 8 & 0 \\ 0 & 0 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

suy ra không gian con $E(20) = \langle(-2,1,0)\rangle$, nó có cơ sở trực chuẩn $((-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0))$. Với $\lambda = 15$, ta biến đổi ma trận

$$Q - 15I_3 = \begin{bmatrix} 7 & 4 & 0 \\ 4 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

suy ra không gian con $E(15) = \langle(0,0,1)\rangle$, nó có cơ sở trực chuẩn $((0,0,1))$. Vậy ma trận trực giao chéo hóa $Q$ là

$$C = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} & 0 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Khi đó, $C^tQC = \begin{bmatrix} 30 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 15 \end{bmatrix}$. Do đó, phương trình của $p$ trong cơ sở $\mathscr{B} = ((\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0), (-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0), (0,0,1))$ là

$$p(\mathbf{x}) = 30x'^2 + 20y'^2 + 15z'^2 \qquad \text{với } \mathbf{x}|_{\mathscr{B}} = (x', y', z').$$

Dùng phép đổi tọa độ

$$\begin{cases} x = \frac{1}{\sqrt{5}}x' - \frac{2}{\sqrt{5}}y' \\ y = \frac{2}{\sqrt{5}}x' + \frac{1}{\sqrt{5}}y' \\ z = z' \end{cases}$$

phương trình của $\mathscr{S}$ trong hệ tọa độ mới là

$$30x'^2 + 20y'^2 + 15z'^2 - \frac{480}{\sqrt{5}}x' + \frac{40}{\sqrt{5}}y' - 30z' + 343 = 0$$

Biến đổi phương trình trên $\mathscr{S}$ trong hệ tọa độ mới

$$30\Big(x' - \frac{8}{\sqrt{5}}\Big)^2 + 20\Big(y' + \frac{1}{\sqrt{5}}\Big)^2 + 15(z'-1)^2 = 60$$

$$\frac{(x' - \frac{8}{\sqrt{5}})^2}{2} + \frac{(y' + \frac{1}{\sqrt{5}})^2}{3} + \frac{(z'-1)^2}{4} = 1$$

Đổi sang tọa độ khác

$$\begin{cases} X = x' - \frac{8}{\sqrt{5}} \\ Y = y' + \frac{1}{\sqrt{5}} \\ Z = z' - 1 \end{cases} \qquad \text{hay} \qquad \begin{cases} x = \frac{1}{\sqrt{5}}X - \frac{2}{\sqrt{5}}Y + 2 \\ y = \frac{2}{\sqrt{5}}X + \frac{1}{\sqrt{5}}Y + 3 \\ z = Z + 1 \end{cases}$$

Lúc đó ta nhận được phương trình chính tắc của $\mathscr{S}$ là

$$\frac{X^2}{2} + \frac{Y^2}{3} + \frac{Z^2}{4} = 1.$$

Đây là phương trình chính tắc của một mặt ellipsoid. Tâm của $\mathscr{S}$ là $I(2, 3, 1)$ (trong hệ tọa độ $Oxyz$), các bán trục là $\sqrt{2}, \sqrt{3}, 2$. $\square$

## Bài tập

**1**) Nhận dạng và lập phương trình chính tắc của các đường bậc hai sau:

**(a)** $2x^2 - 4xy - y^2 + 8 = 0$

**(b)** $x^2 + 2xy + y^2 + 8x + y = 0$

**(c)** $5x^2 + 4xy + 5y^2 - 9x = 0$

**(d)** $11x^2 + 24xy + 4y^2 - 15 = 0$

**(e)** $2x^2 + 4xy + 5y^2 - 24 = 0$

**(f)** $x^2 + xy + y^2 - 18 = 0$

**(g)** $5x^2 - 4xy + 8y^2 - 36$

**(h)** $25x^2 - 14xy + 25y^2 + 64x - 64y - 224 = 0$

**(i)** $9x^2 - 24xy + 16y^2 - 20x + 110y - 50 = 0$

**2**) Nhận dạng và lập phương trình chính tắc của các mặt bậc hai sau đây

**(a)** $2x^2 - 2xz + 2y^2 - 2yz + 3z^2 - 16 = 0$

**(b)** $2xy + 2xz + 2yz - 6x - 6y - 4z = 0$

**(c)** $7x^2 + 7y^2 + 10z^2 - 2xy - 4xz + 4yz - 12x + 12y + 60z - 24 = 0$

**(d)** $2xy - 6x + 10y + z - 31 = 0$

**(e)** $2x^2 + 2y^2 + 5z^2 - 4xy - 2xz + 2yz + 10x - 26y - 2z = 0$

**(f)** $11x^2 + 5y^2 + 2z^2 + 16xy + 4xz - 20yz + 2x + 2y + 2z + 1 = 0$

**(g)** $3y^2 + 3z^2 + 4xy + 4xz - 2yz + 4x + 1 = 0.$

# Chương VIII

# Dạng chuẩn tắc Jordan của ma trận vuông

Ta đã biết hai ma trận $A$ và $B$ đối với hai cơ sở khác nhau của cùng một phép biến đổi tuyến tính $f$ là đồng dạng, nghĩa là tồn tại ma trận không suy biến $C$ sao cho $B = C^{-1}AC$. Hơn nữa, quan hệ đồng dạng này là một quan hệ tương đương trên $\mathcal{M}(\mathbb{K})$. Mục đích cơ bản của chương này là xác định điều kiện cần và đủ để hai ma trận vuông cùng cấp đồng dạng nhau và trong mỗi lớp tương đương như vậy sẽ xác định một đại diện có dạng "đơn giản nhất". Để giải quyết bài toán này trước tiên ta cần khái niệm và tính chất của $\lambda$-ma trận.

## § 1 $\lambda$-ma trận

### 1.1 $\lambda$-ma trận và phép biến đổi sơ cấp

**1.1.1 Định nghĩa.** Ta gọi một **$\lambda$-ma trận** cấp $m \times n$ là ma trận có dạng

$$A(\lambda) = \begin{bmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \dots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \dots & a_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \dots & a_{mn}(\lambda) \end{bmatrix}$$

trong đó $a_{ij}(\lambda) \in \mathbb{K}[\lambda]$ với mọi $i = 1, \dots, m$ và $j = 1, \dots, n$. Như vậy $\lambda$-ma trận là một ma trận trên $\mathbb{K}[\lambda]$ hay thuộc $\mathcal{M}_{m,n}(\mathbb{K}[\lambda])$, nghĩa là mỗi phần tử của $\lambda$-ma trận là một đa thức trên $\mathbb{K}$.

**1.1.2 Nhận xét.** Với $A(\lambda)$ như trong định nghĩa, đặt

$$d = \max\left\{\deg(a_{ij}(\lambda)) : i = 1, \ldots, m; j = 1, \ldots, n\right\}.$$

Ta viết lại ma trận $A(\lambda)$ như sau

$$A(\lambda) = A_0 + \lambda A_1 + \cdots + \lambda^d A_d,$$

trong đó $A_i \in \mathcal{M}_{m,n}(\mathbb{K})$. Như vậy, mỗi $\lambda$-ma trận là một đa thức trên $\mathcal{M}_{m,n}(\mathbb{K})$ hay nó là phần tử của $\mathcal{M}_{m,n}(\mathbb{K})[\lambda]$

**1.1.3 Thí dụ.** $\begin{bmatrix} \lambda^2 + 2\lambda + 3 & \lambda - 8 \\ 3\lambda + 1 & 2\lambda^2 + 3 \end{bmatrix} = \begin{bmatrix} 3 & -8 \\ 1 & 3 \end{bmatrix} + \lambda \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}$ $+ \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. □

**1.1.4 Định nghĩa.** Cũng giống như ma trận trên trường số $\mathbb{K}$, ta có ba *phép biến đổi sơ cấp trên dòng* (hay cột) đối với $\lambda$-ma trận; cụ thể như sau

- Đổi chỗ hai dòng (cột) cho nhau
- Nhân một dòng (cột) cho một số khác không
- Nhân một dòng (cột) với một đa thức rồi cộng vào dòng (cột) khác

**1.1.5 Nhận xét.**
- Mỗi phép biến đổi sơ cấp đều có phép biến đổi ngược (nghĩa là nếu thực hiện liên tiếp hai phép toán này thì $\lambda$-ma trận không thay đổi).
- Phép biến đổi đổi chỗ hai dòng (cột) cho nhau có thể suy ra được từ hai phép biến đổi sơ cấp còn lại. Đổi chỗ hai dòng $i$ và $j$ ta thực hiện như sau: (a) cộng dòng $j$ vào dòng $i$; (b) lấy dòng $j$ trừ dòng $i$; (c) nhân dòng $j$ với $-1$.

$$\begin{bmatrix} i \\ j \end{bmatrix} \to \begin{bmatrix} i+j \\ j \end{bmatrix} \to \begin{bmatrix} i+j \\ -i \end{bmatrix} \to \begin{bmatrix} j \\ -i \end{bmatrix} \to \begin{bmatrix} j \\ i \end{bmatrix}$$

- Cũng như đối với ma trận thông thường các phép biến đổi sơ cấp chính là việc nhân trái hay phải của $\lambda$-ma trận đang xét với một ma trận sơ cấp thích hợp. Với kí hiệu ma trận sơ cấp đã được dùng ta có: (a) Đổi chỗ hai dòng (cột) $i$ và $j$ của ma trận $A(\lambda)$ ta thực hiện $E_{I;i,j}A(\lambda)$ $(A(\lambda)E_{I;i,j})$; (b) Nhân dòng (cột) $i$ của ma trận $A(\lambda)$ cho một số khác không $\alpha$ ta thực hiện $E_{II;i,\alpha}A(\lambda)$ $(A(\lambda)E_{II;i,\alpha})$; (c) Nhân dòng (cột) $j$ cho đa thức $f(\lambda)$ rồi cộng vào dòng (cột) $i$ của ma trận $A(\lambda)$ ta thực hiện $E_{III;i,j,f(\lambda)}A(\lambda)$ $(A(\lambda)E_{III;j,i,f(\lambda)})$. (Các ma trận sơ cấp phải có cấp thích hợp để thực hiện các phép nhân).

- Định thức của $E_{I;i,j}$ bằng $-1$, định thức của $E_{II;i,\alpha}$ bằng $\alpha$, định thức của $E_{III;i,j,f(\lambda)}$ bằng 1.

**1.1.6 Định nghĩa.** Cho $A(\lambda)$ và $B(\lambda)$ là hai $\lambda$-ma trận vuông cấp $n$. Ta nói $A(\lambda)$ **tương đương*** với $B(\lambda)$, kí hiệu $A(\lambda) \sim B(\lambda)$, nếu $B(\lambda)$ suy ra được từ $A(\lambda)$ sau một số hữu hạn phép biến đổi sơ cấp.

**1.1.7** Cho $A(\lambda)$ là một $\lambda$-ma trận vuông cấp $n$. Với mỗi số tự nhiên $1 \le k \le n$, ta kí hiệu $D_k(A)$ là ước chung lớn nhất (với hệ số bậc cao nhất của $\lambda$ bằng 1) của tất cả các định thức con cấp $k$ của $A(\lambda)$.

**1.1.8 Mệnh đề.** *Nếu hai $\lambda$-ma trận $A(\lambda)$ và $B(\lambda)$ tương đương, thì $D_k(A) = D_k(B)$ với mọi $k = 1, \ldots, n$.*

*Chứng minh.* Giả sử $C(\lambda)$ là ma trận thu được sau khi chúng ta áp dụng một phép biến đổi sơ cấp theo dòng vào ma trận $A(\lambda)$. Ta có ba loại phép biến đổi sơ cấp.

- Ma trận $C(\lambda)$ thu được bằng cách đổi chỗ hai dòng của ma trận $A(\lambda)$. Do đó mọi định thức con cấp $k$ của $C(\lambda)$ là định thức con của cấp $k$ của $A(\lambda)$ (và ngược lại), có thể thay đổi thứ tự dòng, cho nên chỉ khác dấu; do đó, ước chung lớn nhất của chúng là giống nhau. Nghĩa là $D_k(C) = D_k(A)$.

*Thực sự quan hệ này là một quan hệ tương đương

- Ma trận $C(\lambda)$ thu được bằng cách nhân một số khác không vào một hàng nào đó của $A(\lambda)$. Do đó mọi định thức con cấp $k$ của $C(\lambda)$ hoặc là một định thức con cấp $k$ của $A(\lambda)$ hoặc là hằng số trên nhân với định thức con cấp $k$ của $A(\lambda)$. Từ đó ta suy ra được ước chung lớn nhất của chúng không có gì thay đổi, nghĩa là $D_k(C) = D_k(A)$.

- Ma trận $C(\lambda)$ thu được bằng cách nhân dòng $j$ của $A(\lambda)$ với $f(\lambda)$ rồi cộng vào dòng $i$ của $A(\lambda)$. Khi đó các định thức con cấp $k$ của $C(\lambda)$ được chia thành ba loại.

  **(a)** Định thức con cấp $k$ của $C(\lambda)$ không chứa phần tử dòng $i$. Đây cũng là một định thức con cấp $k$ của $A(\lambda)$. Do đó, định thức con đang xét chia hết cho $D_k(A)$.

  **(b)** Định thức con cấp $k$ của $C(\lambda)$ có chứa phần tử dòng thứ $i$ và dòng thứ $j$. Định thức này đúng bằng định thức con cấp $k$ của $A(\lambda)$ tương ứng với các dòng và các cột được chọn. Do đó định thức con đang xét này cũng chia hết cho $D_k(A)$.

  **(c)** Định thức con cấp $k$ của $C(\lambda)$ có chứa phần từ dòng $i$ nhưng không chứa phần tử dòng $j$. Khi đó ta có thể thấy được định thức con cấp $k$ đang xét bằng $M + f(\lambda)N$ trong đó $M$ là một định thức con cấp $k$ của $A(\lambda)$ với các dòng và cột được lấy tương ứng còn $N$ là định thức cấp $k$ từ $M$ thay dòng tương ứng với $i$ bởi các phần tử tương ứng dòng $j$, cho nên $N$ là một định thức con cấp $k$ hay là trái dấu với một định thức con cấp $k$ của $A(\lambda)$. Vậy định thức con cấp $k$ của $C(\lambda)$ đang xét chia hết cho $D_k(A)$.

  Cả ba trường hợp định thức con cấp $k$ của $C(\lambda)$ chia hết cho $D_k(A)$. Vậy ta phải có $D_k(C)$ chia hết cho $D_k(A)$. Mặt khác, ta thấy $A(\lambda)$ thu được từ $C(\lambda)$ bằng cách nhân $-f(\lambda)$ với dòng $j$ của $C(\lambda)$ rồi cộng vào dòng $i$ của $C(\lambda)$. Bằng cách lập luận như trên ta có được $D_k(A)$ chia hết cho $D_k(C)$. Vậy $D_k(A) = D_k(C)$.

Đối với phép biến đổi sơ cấp theo cột ta cũng được kết quả như trên.

Do $A(\lambda)$ và $B(\lambda)$ tương đương nhau nên $B(\lambda)$ thu được từ $A(\lambda)$ bằng cách áp dụng một dãy các phép biến đổi sơ cấp. Dùng qui nạp và kết quả trên ta chứng minh được $D_k(A) = D_k(B)$. □

## 1.2 $\lambda$-ma trận chính tắc

**1.2.1 Định nghĩa.** $\lambda$-ma trận vuông $E(\lambda)$ gọi là có dạng chính tắc nếu $E(\lambda)$ có dạng

$$E(\lambda) = \begin{bmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & e_2(\lambda) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e_n(\lambda) \end{bmatrix}$$

trong đó $e_i(\lambda) \mid e_{i+1}(\lambda)$ (chia hết) với $i = 1, \dots, n-1$ và nếu $e_i(\lambda) \neq 0$ thì hệ số của bậc cao nhất trong $e_i(\lambda)$ là 1.

**1.2.2 Thí dụ.** Ma trận

$$A(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda - 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda^2 - 1 & 0 & 0 \\ 0 & 0 & 0 & \lambda^3 - \lambda & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

là một $\lambda$-ma trận chính tắc. □

**1.2.3 Mệnh đề.** *Nếu $E(\lambda)$ là $\lambda$-ma trận cấp $n$ chính tắc như trong định nghĩa trên, thì $D_k(E) = e_1(\lambda) \cdots e_k(\lambda)$ với $k = 1, \dots, n$.*

*Chứng minh.* Chúng tôi minh họa chứng minh mệnh đề này trong trường hợp $n = 4$. Trong trường hợp tổng quát, chúng ta làm tương tự hay dùng qui nạp. Giả sử

$$E(\lambda) = \begin{bmatrix} e_1(\lambda) & 0 & 0 & 0 \\ 0 & e_2(\lambda) & 0 & 0 \\ 0 & 0 & e_3(\lambda) & 0 \\ 0 & 0 & 0 & e_4(\lambda) \end{bmatrix}.$$

Ta có

$$
\begin{aligned}
D_1 &= \gcd(e_1(\lambda), e_2(\lambda), e_3(\lambda), e_4(\lambda)) = e_1(\lambda),\\
D_2 &= \gcd\big(e_1(\lambda)e_2(\lambda), e_1(\lambda)e_3(\lambda), e_1(\lambda)e_4(\lambda),\\
&\qquad\qquad e_2(\lambda)e_3(\lambda), e_2(\lambda)e_4(\lambda), e_3(\lambda)e_4(\lambda)\big)\\
&= e_1(\lambda)e_2(\lambda),\\
D_3 &= \gcd(e_1(\lambda)e_2(\lambda)e_3(\lambda), e_1(\lambda)e_2(\lambda)e_4(\lambda),\\
&\qquad\qquad e_1(\lambda)e_3(\lambda)e_4(\lambda), e_2(\lambda)e_3(\lambda)e_4(\lambda))\\
&= e_1(\lambda)e_2(\lambda)e_3(\lambda),\\
D_4 &= \gcd(e_1(\lambda)e_2(\lambda)e_3(\lambda)e_4(\lambda)) = e_1(\lambda)e_2(\lambda)e_3(\lambda)e_4(\lambda).
\end{aligned}
$$
□

**1.2.4 Định lý.** *Một $\lambda$-ma trận vuông bất kỳ đều có thể đưa được về dạng chính tắc và duy nhất bằng các phép biến đổi sơ cấp (theo dòng và theo cột). Nghĩa là một $\lambda$-ma trận vuông bất kỳ đều tương đương với một $\lambda$-ma trận chính tắc.*

*Chứng minh.* Dễ dàng thấy rằng $\lambda$-ma trận $O$ là một dạng chính tắc. Cho nên ta chỉ chứng minh định lý cho trường hợp $\lambda$-ma trận khác $O$. Ta sẽ chứng minh bằng qui nạp theo cấp của $\lambda$-ma trận.

Nếu cấp của $\lambda$-ma trận $A(\lambda)$ là $k = 1$, nghĩa là $A(\lambda) = \big[a_{11}(\lambda)\big]$. Ta nhân $A(\lambda)$ với số nghịch đảo của hệ số cao nhất của $a_{11}(\lambda)$ sẽ thu được dạng chính tắc.

Giả sử định lý đúng với mọi $\lambda$-ma trận có cấp $k = n - 1$. Xét $\lambda$-ma trận $A(\lambda)$ có cấp $k = n$. Khi đó, trong tất cả các $\lambda$-ma trận tương đương với $A(\lambda)$ có thể tìm được một ma trận

$$
B(\lambda) = \begin{bmatrix}
b_{11}(\lambda) & b_{12}(\lambda) & \dots & b_{1n}(\lambda)\\
b_{21}(\lambda) & b_{22}(\lambda) & \dots & b_{2n}(\lambda)\\
\vdots & \vdots & \ddots & \vdots\\
b_{n1}(\lambda) & b_{n2}(\lambda) & \dots & b_{nn}(\lambda)
\end{bmatrix}
$$

sao cho $b_{11}(\lambda) \neq 0$ và có bậc nhỏ nhất*. Ma trận $B(\lambda)$ này có tính chất là $b_{11}(\lambda)$ chia hết mọi phần tử cùng dòng 1 và cột 1. Thật vậy, giả sử tồn tại

---

*$b_{11}(\lambda)$ chính là ước chung lớn nhất với hệ số của bậc cao nhất $\lambda$ bằng 1 của các số hạng $a_{ij}(\lambda)$

$2 \leq k \leq n$ sao cho $b_{k1}(\lambda)$ không chia hết cho $b_{11}(\lambda)$. Khi đó, ta có thể viết $b_{k1}(\lambda) = b_{11}(\lambda)q_k(\lambda) + r_k(\lambda)$ với $q_k(\lambda), r_k(\lambda) \in \mathbb{K}[\lambda]$, $r_k(\lambda) \neq 0$ và $\deg(r_k(\lambda)) < \deg(b_{11}(\lambda))$. Tiến hành biến đổi sơ cấp trên $B(\lambda)$ như sau: (1) nhân dòng 1 với $-q_k(\lambda)$ rồi cộng vào dòng $k$, (2) sau đó đổi chỗ dòng 1 và dòng $k$ cho nhau. Ta thấy

$$A(\lambda) \sim B(\lambda) \sim \begin{bmatrix} r_k(\lambda) & \dots \\ \vdots & \ddots \end{bmatrix}.$$

Ta gặp phải mâu thuẫn với cách chọn $B(\lambda)$ và $\deg(r_k(\lambda)) < \deg(b_{11}(\lambda))$. Do đó, $b_{11}(\lambda)$ chia hết các đa thức trên cột 1. Tương tự ta cũng có $b_{11}(\lambda)$ chia hết các đa thức trên dòng 1.

Nhờ sự chia hết của $b_{11}(\lambda)$ với các đa thức trong cột 1 và dòng 1 và thực hiện các phép biến đổi sơ cấp ta đưa $\lambda$-ma trận $B(\lambda)$ về dạng

$$B(\lambda) \sim \begin{bmatrix} b_{11}(\lambda) & 0 & \dots & 0 \\ 0 & c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{bmatrix}.$$

Ta thấy $\lambda$-ma trận

$$C(\lambda) = \begin{bmatrix} c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \vdots & \ddots & \vdots \\ c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{bmatrix}$$

có cấp $n-1$, nên theo giả thiết qui nạp có thể đưa được về dạng chính tắc. Do đó ta có

$$A(\lambda) \sim B(\lambda) \sim \begin{bmatrix} b_{11}(\lambda) & 0 & \dots & 0 \\ 0 & e_2(\lambda) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e_n(\lambda) \end{bmatrix}.$$

Để $\lambda$-ma trận trên là $\lambda$-ma trận chính tắc ta phải có $b_{11}(\lambda)$ chia hết $e_2(\lambda)$. Thật vậy, giả sử $e_2(\lambda)$ không chia hết cho $b_{11}(\lambda)$. Khi đó, ta có thể viết

$e_2(\lambda) = b_{11}(\lambda)q(\lambda) + r(\lambda)$ với $r(\lambda) \neq 0$ và $\deg(r(\lambda)) < \deg(b_{11}(\lambda))$. Thực hiện phép biến đổi sơ cấp ta được

$$A(\lambda) \sim \begin{bmatrix} b_{11}(\lambda) & 0 & \dots & 0 \\ 0 & b_{11}(\lambda)q(\lambda) + r(\lambda) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e_n(\lambda) \end{bmatrix}$$

$$\sim \begin{bmatrix} r(\lambda) & 0 & \dots & 0 \\ 0 & b_{11}(\lambda) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e_n(\lambda) \end{bmatrix}.$$

Ta lại gặp mâu thuẫn với cách chọn $B(\lambda)$. Vậy $e_2(\lambda)$ chia hết cho $b_{11}(\lambda)$.

Giả sử $E(\lambda)$ là dạng chính của $\lambda$-ma trận $A(\lambda)$ và

$$E(\lambda) = \begin{bmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & e_2(\lambda) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e_n(\lambda) \end{bmatrix}$$

Theo Mệnh đề 1.1.8 ta có $D_k(A) = D_k(E) = e_1(\lambda) \cdots e_k(\lambda)$. Do đó, $e_1(\lambda) = D_1(A)$ và

$$e_k(\lambda) = \frac{D_k(A)}{D_{k-1}(A)} \qquad \text{với } k = 2, \dots, n$$

Cách xác định các đa thức $e_i(\lambda)$ trong này nói lên tính duy nhất của dạng chính tắc của $A(\lambda)$. □

**1.2.5 Định nghĩa.** Nếu $\lambda$-ma trận $A(\lambda)$ tương đương với ma trận chính tắc

$$E(\lambda) = \begin{bmatrix} e_1(\lambda) & 0 & \dots & 0 \\ 0 & e_2(\lambda) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e_n(\lambda) \end{bmatrix}$$

thì các đa thức $e_i(\lambda)$ gọi là các *nhân tử bất biến* của $\lambda$-ma trận $A(\lambda)$.

**1.2.6 Thí dụ.** Đưa $\lambda$-ma trận $A(\lambda) = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda+3 \end{bmatrix}$ về dạng chính tắc. Ta có

$$\begin{aligned} A(\lambda) &\xrightarrow{d_1 \to d_1 + d_2} \begin{bmatrix} \lambda & \lambda+3 \\ 0 & \lambda+3 \end{bmatrix} \\ &\xrightarrow{c_1 \to c_2 - c_1} \begin{bmatrix} 3 & \lambda+3 \\ \lambda+3 & \lambda+3 \end{bmatrix} \\ &\xrightarrow[d_2 \to d_2 - (\lambda+3)d_1]{d_1 \to d_1/3} \begin{bmatrix} 1 & \lambda/3+1 \\ 0 & -\lambda(\lambda+3)/3 \end{bmatrix} \\ &\xrightarrow[d_2 \to -3d_2]{c_2 \to c_2 - (\lambda/3+1)c_1} \begin{bmatrix} 1 & 0 \\ 0 & \lambda(\lambda+3) \end{bmatrix} \end{aligned}$$ □

**1.2.7 Nhận xét.** Giả sử $A(\lambda)$ là một $\lambda$-ma trận vuông cấp $n$ và $E(\lambda)$ là dạng chính tắc của $A(\lambda)$. Nếu $D_k(A) = 0$, thì $e_k(\lambda) = \cdots = e_n(\lambda) = 0$.

**1.2.8 Thí dụ.** Tìm dạng chính tắc của $\lambda$-ma trận

$$\begin{bmatrix} \lambda(\lambda+1) & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & (\lambda+1)^2 \end{bmatrix}.$$

Ta có $D_1 = \gcd(\lambda(\lambda+1), \lambda, (\lambda+1)^2) = 1$, $D_2 = \gcd(\lambda^2(\lambda+1), \lambda(\lambda+1)^3, \lambda(\lambda+1)^2) = \lambda(\lambda+1)$, và $D_3 = \lambda^2(\lambda+1)^3$. Vậy ta được dạng chính tắc

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda(\lambda+1) & 0 \\ 0 & 0 & \lambda(\lambda+1)^2 \end{bmatrix}.$$ □

**1.2.9 Định lý.** *Hai $\lambda$-ma trận $A(\lambda)$ và $B(\lambda)$ tương đương nhau khi và chỉ khi tồn tại hai $\lambda$-ma trận $P(\lambda)$ và $Q(\lambda)$ có định thức là hằng số khác* 0 *sao cho $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$.*

*Chứng minh.* Giả sử $A(\lambda)$ và $B(\lambda)$ tương đương nhau. Khi đó $A(\lambda)$ biến được thành $B(\lambda)$ sau một số hữu hạn phép biến đổi sơ cấp. Do đó ta có các ma trận sơ cấp $P_1, P_2, \ldots, P_r, Q_1, Q_2, \ldots, Q_s$ sao cho

$$B(\lambda) = P_r \cdots P_2 P_1 A(\lambda) Q_1 Q_2 \cdots Q_s.$$

Đặt $P(\lambda) = P_r \cdots P_2 P_1$ và $Q(\lambda) = Q_1 Q_2 \cdots Q_s$. Khi đó định thức của $P(\lambda)$ và $Q(\lambda)$ là những hằng số khác không (xem Nhận xét 1.1.5).

Giả sử $A(\lambda)$ và $B(\lambda)$ có cấp $n$ và thỏa $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$ với các định thức $|P(\lambda)|$ và $|Q(\lambda)|$ là các hằng số khác không. Do $|P(\lambda)|$ là hằng số khác không nên $D_n(P) = 1$. Do đó, ta được $D_1(P) = \cdots = D_n(P) = 1$. Vậy các nhân tử bất biến của $P(\lambda)$ là $e_i(\lambda) = 1$ với mọi $i = 1, \ldots, n$, hay $P(\lambda) \sim I$. Theo chứng minh phần trên ta có $P(\lambda) = P_r \cdots P_1 I P_1' \cdots P_{r'}'$ với $P_i, P_j'$ là các ma trận sơ cấp. Tương tự, ta cũng có $Q(\lambda) \sim I$ và $Q(\lambda) = Q_s \cdots Q_1 I Q_1' \cdots Q_{s'}'$ với $Q_i, Q_j'$ là các ma trận sơ cấp. Khi đó, ta có

$$B(\lambda) = P_r \cdots P_1 P_1' \cdots P_{r'}' A(\lambda) Q_s \cdots Q_1 Q_1' \cdots Q_{s'}'.$$

Vậy $A(\lambda)$ và $B(\lambda)$ tương đương nhau. □

**1.2.10 Bổ đề.** *Với mọi ma trận vuông $A$ cấp $n$ và mọi $\lambda$-ma trận $A(\lambda)$ cấp $n$, tồn tại $\lambda$-ma trận $Q(\lambda)$ và ma trận $R$ sao cho $A(\lambda) = (A - \lambda I)Q(\lambda) + R$. (Ta cũng có kết quả tương tự $A(\lambda) = Q'(\lambda)(A - \lambda I) + R'$.)*

*Chứng minh.* Bổ đề này được chứng minh bằng qui nạp theo bậc của $A(\lambda)$. Khi bậc của $A(\lambda) = 0$ ta có $A(\lambda) = (A - \lambda I)O + A(\lambda)$, suy ra bổ đề đúng.

Giả sử bổ đề đúng với mọi $\lambda$-ma trận có bậc $m - 1$. Xét một $\lambda$-ma trận $A(\lambda)$ có bậc $m$. Khi đó, ta có thể viết lại

$$\begin{aligned} A(\lambda) &= A_0 + \lambda A_1 + \cdots + \lambda^{m-1} A_{m-1} + \lambda^m A_m \\ &= A_0 + \lambda A_1 + \cdots + \lambda^{m-1}(A_{m-1} + AA_m) - (A - \lambda I)(\lambda^{m-1} A_m). \end{aligned}$$

Theo giả thiết qui nạp tồn tại $Q_1(\lambda)$ và ma trận $R$ sao cho

$$A_0 + \lambda A_1 + \cdots + \lambda^{m-1}(A_{m-1} + AA_m) = (A - \lambda I)Q_1(\lambda) + R.$$

Do đó, ta có

$$A(\lambda) = (A - \lambda I)(Q_1(\lambda) - \lambda^{m-1} A_m) + R.$$

Vậy bổ đề đúng với bậc của $A(\lambda)$ là $m$. Đó là điều cần chứng minh. □

**1.2.11 Định lý.** *Hai ma trận vuông $A$ và $B$ cấp $n$ trên $\mathbb{K}$ đồng dạng nhau đương khi và chỉ khi hai $\lambda$-ma trận $A - \lambda I_n$ và $B - \lambda I_n$ tương đương.*

*Chứng minh.* Giả sử $A$ và $B$ là hai ma trận đồng dạng nhau. Khi đó tồn tại ma trận không suy biến $C$ sao cho $B = C^{-1}AC$. Vậy

$$B - \lambda I = C^{-1}AC - \lambda C^{-1}IC = C^{-1}(A - \lambda I)C.$$

Theo Định lý 1.2.9 ta suy ra được $A - \lambda I$ và $B - \lambda I$ tương đương nhau.

Ngược lại, giả sử $A - \lambda I$ và $B - \lambda I$ tương đương nhau. Theo Định lý 1.2.9 tồn tại hai $\lambda$-ma trận $P(\lambda)$ và $Q(\lambda)$ có định thức là các số khác không sao cho $B - \lambda I = P(\lambda)(A - \lambda I)Q(\lambda)$. Vì $P(\lambda)$ và $Q(\lambda)$ có định thức là các số khác không nên chúng khả đảo, và ta có các đẳng thức

$$\begin{aligned} P(\lambda)^{-1}(B - \lambda I) &= (A - \lambda I)Q(\lambda); \\ (B - \lambda I)Q(\lambda)^{-1} &= P(\lambda)(A - \lambda I). \end{aligned}$$

Mặt khác, theo bổ đề trên tồn tại hai $\lambda$-ma trận $P_1(\lambda)$ và $Q_1(\lambda)$ và hai ma trận $P_0$ và $Q_0$ sao cho

$$P(\lambda) = (B - \lambda I)P_1(\lambda) + P_0; \quad Q(\lambda) = Q_1(\lambda)(B - \lambda I) + Q_0.$$

Từ các đẳng thức trên ta tính được

$$\begin{aligned} &B - \lambda I \\ &\quad = P(\lambda)(A - \lambda I)Q(\lambda) \\ &\quad = ((B - \lambda I)P_1(\lambda) + P_0)(A - \lambda I)(Q_1(\lambda)(B - \lambda I) + Q_0) \\ &\quad = (B - \lambda I)P_1(\lambda)(A - \lambda I)Q_1(\lambda)(B - \lambda I) + P_0(A - \lambda I)Q_0 \\ &\qquad\quad + (B - \lambda I)P_1(\lambda)(A - \lambda I)Q_0 + P_0(A - \lambda I)Q_1(\lambda)(B - \lambda I) \\ &\quad = P(\lambda)(A - \lambda I)Q_1(\lambda)(B - \lambda I) + (B - \lambda I)P_1(\lambda)(A - \lambda I)Q(\lambda) \\ &\qquad\quad - (B - \lambda I)P_1(\lambda)(A - \lambda I)Q_1(\lambda)(B - \lambda I) + P_0(A - \lambda I)Q_0. \end{aligned}$$

Suy ra

$$\begin{aligned}B - \lambda I - P_0(A - \lambda I)Q_0 &= (B - \lambda I)Q(\lambda)^{-1}Q_1(\lambda)(B - \lambda I)\\ &\quad + (B - \lambda I)P_1(\lambda)P(\lambda)^{-1}(B - \lambda I)\\ &\quad - (B - \lambda I)P_1(\lambda)(A - \lambda I)Q_1(\lambda)(B - \lambda I)\\ &= (B - \lambda I)[Q(\lambda)^{-1}Q_1(\lambda) + P_1(\lambda)P(\lambda)^{-1}\\ &\quad - P_1(\lambda)(A - \lambda I)Q_1(\lambda)](B - \lambda I).\end{aligned}$$

Đây là đẳng thức giữa hai $\lambda$-ma trận, do đó là đẳng thức giữa hai đa thức (với hệ số là các ma trận) theo $\lambda$. Nếu đa thức ở hai vế khác đa thức không, thì đa thức vế trái có bậc tối đa là 1 còn đa thức vế phải có bậc ít nhất là 2 nên không thể xảy ra. Do đó, đa thức ở hai vế phải là đa thức không, nghĩa là $B - \lambda I - P_0(A - \lambda I)Q_0 = O$. Ta phải có

$$B = P_0AQ_0 \qquad \text{và} \qquad I = P_0Q_0.$$

Suy ra $P_0 = Q_0^{-1}$, cho nên $B = Q_0^{-1}AQ_0$. Nghĩa là hai ma trận $A$ và $B$ đồng dạng. □

**1.2.12 Thí dụ.** Xét hai ma trận $A = \begin{bmatrix} -2 & 1 \\ 0 & 3 \end{bmatrix}$ và $B = \begin{bmatrix} -10 & -4 \\ 26 & 11 \end{bmatrix}$.

Các đa thức đặc trưng của chúng là $\chi_A(\lambda) = (\lambda + 2)(\lambda - 3)$ và $\chi_B(\lambda) = (\lambda + 10)(\lambda - 11) + 104 = \lambda^2 - \lambda - 6$. Ta cũng nhận thấy

$$D_1(A - \lambda I) = D_1(B - \lambda I) = 1.$$

Do đó $A - \lambda I = \begin{bmatrix} -2 - \lambda & 1 \\ 0 & 3 - \lambda \end{bmatrix}$ có dạng chính tắc là $\begin{bmatrix} 1 & 0 \\ 0 & \lambda^2 - \lambda - 6 \end{bmatrix}$ và $B - \lambda I = \begin{bmatrix} -10 - \lambda & -4 \\ 26 & 11 - \lambda \end{bmatrix}$ có dạng chính tắc là $\begin{bmatrix} 1 & 0 \\ 0 & \lambda^2 - \lambda - 6 \end{bmatrix}$. Vậy $A - \lambda I \sim B - \lambda I$, cho nên $A$ đồng dạng $B$. □

## Bài tập

**1**) Tìm các phép biến đổi sở cấp ngược của các phép biến đổi sơ cấp.

**2**) Chứng minh $D_k(C) = D_k(A)$ chi tiết cho trường hợp $C(\lambda)$ thu được từ $A(\lambda)$ bằng cách đổi chỗ hai dòng (hay cột) trong chứng minh Mệnh đề 1.1.8.

**3**) Tìm dạng chính tắc của các $\lambda$-ma trận sau:

**(a)** $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda+1 \end{bmatrix}$ **(b)** $\begin{bmatrix} \lambda^2-1 & 0 \\ 0 & \lambda^3-1 \end{bmatrix}$

**(c)** $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ **(d)** $\begin{bmatrix} \lambda^2-1 & 0 \\ 0 & (\lambda-1)^3 \end{bmatrix}$

**(e)** $\begin{bmatrix} \lambda & 0 \\ 0 & \lambda+5 \end{bmatrix}$ **(f)** $\begin{bmatrix} \lambda^2-1 & \lambda+1 \\ \lambda+1 & \lambda^2+2\lambda+1 \end{bmatrix}$

**(g)** $\begin{bmatrix} \lambda(\lambda+1) & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & (\lambda+1)^2 \end{bmatrix}$ **(h)** $\begin{bmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ \lambda^2+1 & \lambda^2 & -\lambda^2 \end{bmatrix}$

**4**) Tìm dạng chính tắc của các $\lambda$-ma trận sau:

**(a)** $\begin{bmatrix} \lambda-2 & -1 & 0 \\ 0 & \lambda-2 & -1 \\ 0 & 0 & \lambda-2 \end{bmatrix}$

**(b)** $\begin{bmatrix} \lambda(\lambda+1) & 0 & 0 \\ 0 & \lambda(\lambda-2) & 0 \\ 0 & 0 & \lambda(\lambda-3) \end{bmatrix}$

**(c)** $\begin{bmatrix} \lambda(\lambda-1) & 0 & 0 \\ 0 & \lambda(\lambda-2) & 0 \\ 0 & 0 & (\lambda-1)(\lambda-2) \end{bmatrix}$

**(d)** $\begin{bmatrix} \lambda^2 & \lambda^2-\lambda & 3\lambda^2 \\ \lambda^2-\lambda & 3\lambda^2-\lambda & \lambda^3+4\lambda^2-3\lambda \\ \lambda^2+\lambda & \lambda^2+\lambda & 3\lambda^2+3\lambda \end{bmatrix}$

**(e)** $\begin{bmatrix} \lambda & -1 & 0 & 0 & 0 \\ 0 & \lambda & -1 & 0 & 0 \\ 0 & 0 & \lambda & -1 & 0 \\ 0 & 0 & 0 & \lambda & -1 \\ 1 & 2 & 3 & 4 & \lambda+5 \end{bmatrix}$

**(f)** $\begin{bmatrix} (\lambda-1)(\lambda-2)(\lambda-3) & 0 & 0 & 0 \\ 0 & (\lambda-1)(\lambda-2)(\lambda-4) & 0 & 0 \\ 0 & 0 & (\lambda-1)(\lambda-3)(\lambda-4) & 0 \\ 0 & 0 & 0 & (\lambda-2)(\lambda-3)(\lambda-4) \end{bmatrix}$

**5)** Các cặp $\lambda$-ma trận sau có tương đương nhau không?

**(a)** $\begin{bmatrix} 3\lambda+1 & \lambda & 4\lambda-1 \\ 1-\lambda^2 & \lambda-1 & \lambda-\lambda^2 \\ \lambda^2+\lambda+2 & \lambda & \lambda^2+2\lambda \end{bmatrix}$ và $\begin{bmatrix} \lambda+1 & \lambda-2 & \lambda^2-2\lambda \\ 2\lambda & 2\lambda-3 & \lambda^2-2\lambda \\ -2 & 1 & 1 \end{bmatrix}$

**(b)** $\begin{bmatrix} \lambda^2+\lambda+1 & 3\lambda-\lambda^2 & 2\lambda^2+\lambda & \lambda^2 \\ \lambda^2+\lambda & 3\lambda-\lambda^2 & 2\lambda^2+\lambda & \lambda^2 \\ \lambda^2-\lambda & 2\lambda-\lambda^2 & 2\lambda^2+\lambda & \lambda^2 \\ \lambda^2 & -\lambda^2 & 2\lambda^2 & \lambda^2 \end{bmatrix}$ và $\begin{bmatrix} 3 & \lambda^2+1 & 3\lambda^2 & \lambda^2 \\ 2 & \lambda^2+1 & 3\lambda^2 & \lambda^2 \\ 0 & \lambda^2 & 3\lambda^2 & \lambda^2 \\ \lambda^2 & -\lambda^2 & 2\lambda^2 & \lambda^2 \end{bmatrix}$

**6)** Hãy tìm các $\lambda$-ma trận $P$ và $Q$ có định thức hằng số sao cho $B = PAQ$ với $A$ và $B$ là

**(a)** $A = \begin{bmatrix} 2\lambda^2-\lambda+1 & 3\lambda^2-2\lambda+1 \\ 2\lambda^2+\lambda-1 & 3\lambda^2+\lambda-2 \end{bmatrix}$

và $B = \begin{bmatrix} 3\lambda^3+7\lambda+2 & 3\lambda^3+4\lambda-1 \\ 2\lambda^3+5\lambda+1 & 2\lambda^3+3\lambda-1 \end{bmatrix}$

**(b)** $A = \begin{bmatrix} 2\lambda^2-\lambda-1 & 2\lambda^3+2\lambda^2-3\lambda \\ \lambda^2-\lambda & \lambda^3-\lambda \end{bmatrix}$

và $B = \begin{bmatrix} \lambda^3-\lambda^2+\lambda-1 & 2\lambda^3-\lambda^2+\lambda-2 \\ \lambda^3-\lambda^2 & 2\lambda^3-\lambda^2-\lambda \end{bmatrix}$

**7)** Hãy chia bên trái $\lambda$-ma trận $A$ cho $B-\lambda I$ với $B = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & 2 \\ 2 & -1 & 3 \end{bmatrix}$

và $A = \begin{bmatrix} -2\lambda^2+5\lambda+3 & -\lambda^2+3\lambda+2 & 6-\lambda \\ -3\lambda^2+7\lambda+11 & -3\lambda^2+9\lambda+1 & 8-2\lambda \\ -\lambda^2+2\lambda+8 & -2\lambda^2+5\lambda+3 & -\lambda+4 \end{bmatrix}$.

**8)** Chứng minh rằng các cặp ma trận sau đồng dạng.

**(a)** $A = \begin{bmatrix} 3 & 2 & -5 \\ 2 & 6 & -10 \\ 1 & 2 & -3 \end{bmatrix}$ và $B = \begin{bmatrix} 6 & 20 & -34 \\ 6 & 32 & -51 \\ 4 & 20 & -32 \end{bmatrix}$

**(b)** $A = \begin{bmatrix} 6 & 6 & -15 \\ 1 & 5 & -5 \\ 1 & 2 & -2 \end{bmatrix}$ và $B = \begin{bmatrix} 37 & -20 & -4 \\ 34 & -17 & -4 \\ 119 & -70 & -11 \end{bmatrix}$

**9)** Hãy tìm các giá trị riêng của ma trận

$$A = \begin{bmatrix} 0 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & -1 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

# § 2 Định lý Cayley-Hamilton

## 2.1 Nghiệm ma trận của đa thức và Định lý Cayley-Hamilton

**2.1.1 Định nghĩa.** Cho đa thức $f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \cdots + a_{m-1}\lambda + a_m \in \mathbb{K}[\lambda]$ và ma trận $A = (a_{ij})_{n\times n} \in \mathcal{M}_n(\mathbb{K})$. Khi đó ma trận vuông cấp $n$

$$a_0A^m + a_1A^{m-1} + \cdots + a_{m-1}A + a_mI_n$$

được gọi là **giá trị của đa thức** $f(\lambda)$ tại ma trận $A$, ký hiệu $f(A)$. Đặc biệt, nếu $f(A) = O = (0)_{n\times n}$ thì $A$ được gọi là **nghiệm** của đa thức $f(\lambda)$.

**2.1.2 Định lý.** *Mỗi ma trận vuông cấp $n$ trên trường $\mathbb{K}$ đều là nghiệm của ít nhất một đa thức khác không trong $\mathbb{K}[\lambda]$.*

*Chứng minh.* Ta biết rằng hệ $(E_{kl})_{k=1,\dots,n;l=1,\dots,n}$ với $E_{kl}$ là ma trận mà các phần tử của nó đều bằng không trừ phần tử ở hàng $k$ cột $l$ bằng 1 là một cơ sở của không gian vector $\mathcal{M}_n(\mathbb{K})$, cho nên $\dim(\mathcal{M}_n(\mathbb{K})) = n^2$. Do đó hệ $n^2+1$ ma trận $A^{n^2}, A^{n^2-1}, \dots, A^2, A, I_n$ phụ thuộc tuyến tính trong $\mathcal{M}_n(\mathbb{K})$. Suy ra tồn tại các số $a_0, a_1, \dots, a_{n^2} \in \mathbb{K}$ không đồng thời bằng 0 sao cho $a_0A^{n^2} + a_1A^{n^2-1} + \cdots + a_{n^2-1}A + a_{n^2}I_n = O$. Vậy $A$ là nghiệm ma trận của đa thức khác không $f(\lambda) = a_0\lambda^{n^2} + a_1\lambda^{n^2-1} + \cdots + a_{n^2-1}\lambda + a_{n^2} \in \mathbb{K}[\lambda]$. □

Định lý trên chỉ cho biết sự tồn tại của đa thức nhận $A$ làm nghiệm mà không cung cấp thuật toán để tìm đa thức này, hơn nữa đa thức được chỉ ra trong chứng minh định lý có bậc "quá cao". Thật sự ta có một đa thức có bậc $n$ thỏa định lý trên.

**2.1.3 Định lý. (Cayley-Hamilton)** *Mọi ma trận vuông đều là nghiệm của đa thức đặc trưng của nó, nghĩa là $\chi_A(A) = O$ (ma trận không).*

*Chứng minh.* Giả sử $A = (a_{ij})_n \in \mathcal{M}_n(\mathbb{K})$ và đa thức đặc trưng của $A$ là

$$\chi_A(\lambda) = |A - \lambda I| = (-1)^n(\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n)$$

Gọi $B$ là ma trận phụ hợp của ma trận $A - \lambda I$. Khi đó phần tử dòng $i$ cột $j$ của $B$ là phần bù đại số của phần tử dòng $j$ cột $i$ trong ma trận $A - \lambda I$ nên là một đa thức $b_{ij}(\lambda)$ với bậc không vượt quá $n-1$. Do đó ta có thể viết $B$ dưới dạng

$$B = B_1\lambda^{n-1} + B_2\lambda^{n-2} + \cdots + B_{n-1}\lambda + B_n$$

với $B_1, B_2, \ldots, B_n$ là các hằng ma trận vuông cấp $n$. Theo định nghĩa của ma trận phụ hợp ta có $(A - \lambda I)B = |A - \lambda I|I$. Ta viết lại

$$\begin{aligned}(A - \lambda I)&(B_1\lambda^{n-1} + B_2\lambda^{n-2} + \cdots + B_{n-1}\lambda + B_n)\\ &= (-1)^n(\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n)I\end{aligned}$$

Khai triển rút gọn và đồng nhất các hệ số của các lũy thừa của $\lambda$ ta được

$$\begin{aligned}(-1)^n a_n I &= AB_n\\ (-1)^n a_{n-1} I &= AB_{n-1} - B_n I\\ (-1)^n a_{n-2} I &= AB_{n-2} - B_{n-1} I\\ &\vdots\\ (-1)^n a_1 I &= AB_1 - B_2 I\\ (-1)^n I &= -B_1 I\end{aligned}$$

Từ đó suy ra được

$$\begin{aligned}
(-1)^n a_n I &= AB_n \\
(-1)^n a_{n-1} A &= A^2 B_{n-1} - AB_n I \\
(-1)^n a_{n-2} A^2 &= A^3 B_{n-2} - A^2 B_{n-1} I \\
&\vdots \\
(-1)^n a_1 A^{n-1} &= A^n B_1 - A^{n-1} B_2 I \\
(-1)^n A^n &= -A^n B_1 I
\end{aligned}$$

Cộng vế theo vế ta được $(-1)^n(A^n + a_1 A^{n-1} + \cdots + a_{n-1}A + a_n I) = O$ hay $\chi_A(A) = O$. □

**2.1.4 Thí dụ.** Ta kiểm chứng định lý Hamilton-Cayley bằng một ma trận đơn giản $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$. Ta có đa thức đặc trưng của $A$ là

$$\chi_A(\lambda) = \begin{vmatrix} 1-\lambda & 1 \\ 2 & -\lambda \end{vmatrix} = \lambda^2 - \lambda - 2$$

do đó ta có

$$\chi_A(A) = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}^2 - \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$ □

## 2.2 Tính giá trị đa thức tại ma trận và đa thức tối tiểu

**2.2.1** Cho $A \in \mathcal{M}_n(\mathbb{K})$ và $f$ là đa thức trên $\mathbb{K}$. Bài toán đặt ra là hãy thu gọn $f(A)$ và từ đó tính $f(A)$. Gọi $\chi_A(\lambda)$ là đa thức đặc trưng của $A$. Thực hiện phép chia $f(\lambda)$ cho $\chi_A(\lambda)$ ta được

$$f(\lambda) = q(\lambda)\chi_A(\lambda) + r(\lambda)$$

ở đó nếu $r(\lambda)$ khác hằng không thì nó là đa thức có bậc nhỏ hơn $n$. Do vậy theo định lý Hamilton-Cayley ta có

$$f(A) = q(A)\chi_A(A) + r(A) = r(A).$$

**2.2.2 Thí dụ.** Tính $f(A) = 2A^8 - 3A^5 + A^4 + A^2 - 4I_3$ với

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

*Giải.* Đa thức đặc trưng của ma trận $A$ là

$$\chi_A(\lambda) = \begin{vmatrix} 1-\lambda & 0 & 2 \\ 0 & -1-\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = -\lambda^3 + 2\lambda - 1.$$

Chia đa thức $2\lambda^8 - 3\lambda^5 + \lambda^4 + \lambda^2 - 4$ cho đa thức $-\lambda^3 + 2\lambda - 1$ ta được đa thức dư là $r(\lambda) = 24\lambda^2 - 37\lambda + 10$. Do đó $r(A) = 24A^2 - 37A + 10I_3$. Vì $A^2 = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$ nên

$$\begin{aligned} f(A) = r(A) &= 24 \begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} - 37 \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix} + 10 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -3 & 48 & -26 \\ 0 & 95 & -61 \\ 0 & -61 & 34 \end{bmatrix}. \end{aligned}$$

□

**2.2.3 Định lý.** *Cho $A \in \mathcal{M}_n(\mathbb{K})$. Khi đó tập hợp $J = \{f(\lambda) \in \mathbb{K}[\lambda] : f(A) = O\}$ là một ideal chính khác $0$ của $\mathbb{K}[\lambda]$. Phần tử sinh của $J$ với hệ số cao nhất bằng $1$ được gọi là* ***đa thức tối tiểu*** *của ma trận $A$, ký hiệu $p_A(\lambda)$.*

*Chứng minh.* Theo Định lý Hamilton-Cayley ta nhận thấy $J \neq \{0\}$. Với mọi $f(\lambda), g(\lambda) \in J$ và $h(\lambda) \in \mathbb{K}[\lambda]$, ta có $f(A) = O, g(A) = O$, cho nên

$$f(A) + g(A) = O \qquad\qquad f(A)h(A) = O.$$

Do đó $f(\lambda) + g(\lambda), f(\lambda)h(\lambda) \in J$. Vậy $J$ là một ideal của $\mathbb{K}[\lambda]$. Hơn nữa, do $\mathbb{K}[\lambda]$ là vành Euclid nên $J$ là ideal chính. □

Từ định nghĩa của đa thức tối tiểu và tính chất ideal chính ta có $p_A(\lambda)$ là ước trong $\mathbb{K}[\lambda]$ của mọi đa thức $g(\lambda) \in \mathbb{K}[\lambda]$ nhận $A$ là nghiệm.

Trong định lý 2.2.3 ta đã định nghĩa đa thức tối tiểu của một ma trận $A$. Từ định lý đó ta cũng có diễn giải khác về đa thức tối tiểu $p(\lambda)$: *đó là đa thức có hệ số cao nhất là* $1$ *và có bậc dương bé nhất sao cho* $p(A) = O$. Tương tự ta cũng có khái niệm đa thức tối tiểu cho một toán tử tuyến tính.

Cho $f$ là một toán tử tuyến tính trong không gian vector hữu hạn chiều. Một đa thức $p(t)$ được gọi là *đa thức tối tiểu* của $f$ nếu $p(t)$ là một đa thức có hệ số cao nhất là 1 và có bậc dương bé nhất với $p(f) = \theta$, ánh xạ không. Theo Định lý 2.2.3 ta có $\langle p(t) \rangle = \{g(\lambda) \in \mathbb{K}[\lambda] : g(f) = \theta\}$.

Sau đây chúng ta chỉ đề cập đến đa thức tối tiểu của ma trận, kết quả của đa thức tối tiểu của toán tử tuyến tính được suy ra tương tự bằng cách áp dụng cho ma trận của nó đối với một cơ sở nào đó.

## 2.3 Đặc trưng của đa thức tối tiểu

**2.3.1 Định lý.** *Cho $p(t)$ là đa thức tối tiểu của ma trận vuông $A$. Đại lượng vô hướng $\lambda$ là giá trị riêng của $A$ nếu và chỉ nếu $p(\lambda) = 0$. Do đó, đa thức đặc trưng và đa thức cực tiểu của $T$ có cùng tập nghiệm.*

*Chứng minh.* Cho $\chi(t)$ là đa thức đặc trưng của $A$. Khi đó $p(t)$ chia hết $\chi(t)$; tồn tại đa thức $q(t)$ sao cho $\chi(t) = q(t)p(t)$. Cho $\lambda$ là giá trị làm cho $p(t)$ bằng 0. Khi đó

$$\chi(\lambda) = q(\lambda)p(\lambda) = q(\lambda) \cdot 0 = 0.$$

Vậy $\lambda$ là giá trị làm cho $\chi(t)$ bằng 0, nghĩa là $\lambda$ là giá trị riêng của $A$.

Ngược lại, giả sử $\lambda$ là giá trị riêng của $A$ và $\mathbf{x}$ là vector riêng tương ứng với $\lambda$. Khi đó, ta có

$$\mathbf{0} = O\mathbf{x} = p(A)\mathbf{x} = p(\lambda)\mathbf{x}, \qquad \text{(bởi vì } A^m\mathbf{x} = \lambda^m\mathbf{x}\text{)}$$

Khi đó, do $\mathbf{x} \neq \mathbf{0}$ nên $p(\lambda) = 0$, suy ra $\lambda$ là một nghiệm của $p(t)$. □

**2.3.2 Hệ quả.** *Cho ma trận vuông $A$ có đa thức tối tiểu $p(t)$ và đa thức đặc trưng $\chi(t)$. Giả sử $\chi(t)$ phân tích thành thừa số như sau*

$$\chi(t) = (\lambda_1 - t)^{n_1}(\lambda_2 - t)^{n_2} \cdots (\lambda_k - t)^{n_k}$$

*tong đó $\lambda_1, \lambda_2, \ldots, \lambda_k$ là các giá trị riêng khác nhau của $A$. Khi đó, tồn tại các số nguyên $m_1, m_2, \ldots, m_k$ sao cho $1 \le m_i \le n_i$ với mọi $i$ và*

$$p(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k}.$$

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

**2.3.3 Thí dụ.** Chúng ta tìm đa thức tối tiểu của ma trận

$$A = \begin{bmatrix} 3 & -1 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 2 \end{bmatrix}.$$

Ma trận $A$ có đa thức đặc trưng

$$\chi(t) = \begin{vmatrix} 3-t & -1 & 0 \\ 0 & 2-t & 0 \\ 1 & -1 & 2-t \end{vmatrix} = -(t-2)^2(t-3),$$

đa thức tối tiểu của $A$ phải là $(t-2)(t-3)$ hoặc $(t-2)^2(t-3)$ bởi Hệ quả 2.3.2. Thế $A$ vào $p(t) = (t-2)(t-3)$ cho thấy

$$p(A) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 0 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix} = O,$$

và khi đó $p(t) = t^2 - 5t + 6$ là đa thức tối tiểu đối với $A$. □

**2.3.4 Thí dụ.** Cho $f$ là toán tử tuyến tính trong $\mathbb{R}^2$ định nghĩa bởi

$$f(a, b) = (2a + 5b, 6a + b).$$

Cho $\mathscr{E}$ là cơ sở chuẩn tắc $\mathbb{R}^2$. Khi đó, ma trận của $f$ đối với cơ sở $\mathscr{E}$ là

$$f|_{\mathscr{E}} = \begin{bmatrix} 2 & 5 \\ 6 & 1 \end{bmatrix}$$

và đa thức đặc trưng của $f$ là

$$\chi(t) = \begin{vmatrix} 2-t & 5 \\ 6 & 1-t \end{vmatrix} = (t-7)(t+4),$$

vậy đa thức tối tiểu của $f$ cũng là $(t-7)(t+4)$. □

**2.3.5 Thí dụ.** Cho $D$ là toán tử tuyến tính trong $\mathbb{R}_2[x]$ định nghĩa bởi $D(f) = f'$, là đạo hàm của $f$. Ta tìm đa thức tối tiểu của $T$. Cho $\mathscr{B}$ cơ sở chuẩn tắc của $\mathbb{R}_2[x]$. Khi đó, ma trận của $D$ đối với cơ sở $\mathscr{B}$ là

$$D|_{\mathscr{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

và khi đó đa thức đặc trưng của $D$ là $-t^3$. Như vậy theo Hệ quả 2.3.2 đa thức tối tiểu của $D$ là $t, t^2$, hoặc $t^3$. Do $D^2(x^2) = 2 \neq 0$, nên $D^2 \neq 0$, và do đó đa thức tối tiểu của $D$ phải là $t^3$. $\square$

Trong thí dụ trên dễ dàng kiểm tra được $\mathbb{R}_2[x]$ là không gian $D$-cyclic, và ta có đa thức tối tiểu và đa thức đặc trưng cùng bậc. Điều này không phải sự trùng hợp ngẫu nhiên; thật sự điều này luôn đúng cho mọi không gian có tính chất đó, được phát biểu chính thức trong định lý sau.

**2.3.6 Định lý.** *Cho $f$ là một toán tử tuyến tính trong không gian vector $n$-chiều $V$ sao cho $V$ là một không gian $f$-cyclic. Khi đó đa thức đặc trưng $\chi(t)$ và đa thức tối tiểu $p(t)$ có cùng bậc và $\chi(t) = (-1)^n p(t)$.*

*Chứng minh.* Giả sử $V$ là một không gian $f$-cyclic; khi đó tồn tại một phần tử $\mathbf{x} \in V$ sao cho

$$\mathscr{B} = (\mathbf{x}, f(\mathbf{x}), \ldots, f^{n-1}(\mathbf{x}))$$

là cơ sở của $V$. Cho

$$g(t) = a_0 + a_1 t + \cdots + a_k t^k,$$

là đa thức có bậc $k < n$. Khi đó $a_k \neq 0$ và

$$g(f)(\mathbf{x}) = a_0\mathbf{x} + a_1 f(\mathbf{x}) + \cdots + a_k f^k(\mathbf{x})$$

là tổ hợp tuyến tính của các phần tử của $\mathscr{B}$ có ít nhất một hệ số khác không, cụ thể là $a_k$. Do $\mathscr{B}$ là độc lập tuyến tính, $g(f)(\mathbf{x}) \neq 0$, và khi đó $g(f) \neq 0$. Vì vậy, đa thức tối tiểu của $f$ có bậc $n$, nó cũng là bậc của đa thức đặc trưng của $f$. $\square$

Định lí 2.3.6 cho điều kiện mà theo nó bậc của đa thức tối tiểu của một toán tử là lớn nhất có thể. Bây giờ ta nghiên cứu chiều ngược lại. Theo Định lí 2.3.1 bậc của đa thức tối tiểu của toán tử phải lớn hơn hoặc bằng số giá trị riêng phân biệt của toán tử. Kết quả tiếp theo chứng tỏ rằng toán tử mà với bậc của đa thức tối tiểu là nhỏ nhất có thể thì chính là toán tử chéo hóa được.

**2.3.7 Định lý.** *Cho $f$ là toán tử tuyến tính trong không gian vector hữu hạn chiều $V$. Khi đó $f$ là chéo hóa được nếu và chỉ nếu đa thức tối tiểu của $f$ có dạng*

$$p(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_k)$$

*với $\lambda_1, \lambda_2, \ldots, \lambda_k$ là các đại lượng vô hướng khác nhau. (Chú ý rằng $\lambda_i$ nhất thiết phải là các giá trị riêng khác nhau của $f$.)*

*Chứng minh.* Giả sử rằng $f$ là chéo hóa được. Cho $\lambda_1, \lambda_2, \ldots, \lambda_k$ là các giá trị riêng khác nhau của $f$, và định nghĩa

$$p(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_k).$$

Theo định lí 2.3.1, $p(t)$ chia hết đa thức cực tiểu của $f$. Đặt $\mathscr{B} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ là một cơ sở của $V$ gồm các vector riêng của $f$, và xét với mọi $\mathbf{v}_i \in \mathscr{B}$. Khi đó $(f - \lambda_j Id_V)(\mathbf{v}_i) = \mathbf{0}$ với giá trị riêng $\lambda_j$ nào đó. Do $(t - \lambda_j)$ chia hết $p(t)$, tồn tại đa thức $q_j(t)$ sao cho $p(t) = q_j(t)(t - \lambda_j)$. Do đó

$$p(f)(\mathbf{v}_i) = q_j(f)(f - \lambda_j Id_V)(\mathbf{v}_i) = 0.$$

Nó kéo theo rằng $p(f) = 0$, do $p(f)$ lấy giá trị mỗi phần tử của cơ sở $V$ là 0. Do đó, $p(t)$ là đa thức tối tiểu của $f$.

Ngược lại, giả sử tồn tại các đại lượng vô hướng phân biệt $\lambda_1, \lambda_2, \ldots, \lambda_k$ sao cho đa thức tối tiểu $p(t)$ của $f$ phân tích thành thừa số như sau

$$p(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_k).$$

Theo Định lí 2.3.1 các $\lambda_i$ là giá trị riêng của $f$. Chúng ta chứng minh bằng qui nạp toán học theo $n = \dim(V)$. Rõ ràng $f$ là chéo hóa được với $n = 1$. Bây giờ giả sử rằng mọi toán tử có tính chất trên là chéo hóa được với mọi không gian có số chiều nhỏ hơn $n$ với mọi $n > 1$, và giả sử rằng $\dim(V) = n$. Đặt $W = \mathrm{Im}(f - \lambda_k Id_V)$. Rõ ràng $W \neq V$ bởi vì $\lambda_k$ là một giá trị riêng của $f$. Nếu $W = \{\mathbf{0}\}$, thì $f = \lambda_k Id_V$, nó thì rõ ràng chéo hóa được. Vậy giả sử rằng $0 < \dim(W) < n$. Khi đó $W$ là $f$-bất biến, và với mọi $\mathbf{x} \in W$ do $p(f) = 0$ ta có

$$(f - \lambda_1 Id_V) \cdots (f - \lambda_{k-1} Id_V)(\mathbf{x}) = \mathbf{0}.$$

Nó suy ra rằng đa thức tối tiểu với $f|_W$ chia hết đa thức $(t - \lambda_1) \cdots (t - \lambda_{k-1})$. Khi đó bởi giả thuyết quy nạp, $f|_W$ là chéo hóa được. Hơn thế nữa, $\lambda_k$ không là giá trị riêng của $f|_W$ theo Định lí 2.3.1. Do đó $W \cap \ker(f - \lambda_k Id_V) = \{\mathbf{0}\}$. Bây giờ cho $\mathscr{B}_1 = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ là cơ sở của $W$ gồm các vector riêng của $f|_W$ (và khi đó là của $f$), và cho $\mathscr{B}_2 = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_p)$ là cơ sở của $\ker(f - \lambda_k Id_V)$, không gian riêng của $f$ tương ứng với $\lambda_k$. Khi đó $\mathscr{B}_1$ và $\mathscr{B}_2$ là rời nhau. Cũng quan sát thấy rằng $m + p = n$ (Định lý 3.2.3, trang 285). Ta chứng minh rằng $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2$ là độc lập tuyến tính. Xét đại lượng vô hướng $a_1, a_2, \ldots, a_m$ và $b_1, b_2, \ldots, b_p$ sao cho

$$a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m + b_1\mathbf{w}_1 + \cdots + b_p\mathbf{w}_p = 0.$$

Đặt

$$\mathbf{x} = \sum_{i=1}^{m} a_i\mathbf{v}_i \qquad \text{và} \qquad \mathbf{y} = \sum_{i=1}^{p} b_i\mathbf{w}_i.$$

Khi đó $\mathbf{x} \in W, \mathbf{y} \in \ker(f - \lambda_k Id_V)$, và $\mathbf{x} + \mathbf{y} = \mathbf{0}$. Nó suy ra rằng $\mathbf{x} = -\mathbf{y} \in W \cap \ker(f - \lambda_k Id_V)$, và do đó $\mathbf{x} = \mathbf{0}$. Do $\mathscr{B}_1$ là độc lập tuyến tính, ta có $a_1 = \cdots = a_m = 0$. Một cách tương tự $b_1 = \cdots = b_p = 0$, và ta kết luận rằng $\mathscr{B}$ là tập con độc lập tuyến tính của $V$ gồm $n$ vector riêng. Nó suy ra rằng $\mathscr{B}$ là cơ sở của $V$ gồm các vector riêng của $f$, và do đó $f$ là chéo hóa được. □

Ngoài các toán tử chéo hóa được, còn có các phương pháp khác để xác định đa thức tối tiểu của một toán tử bất kỳ trên không gian vector hữu hạn chiều. Chúng ta sẽ nghiên cứu một vài phương pháp ấy ở những bài sau.

**2.3.8 Thí dụ.** Ta xác định tất cả ma trận $A \in \mathcal{M}_{2\times 2}(\mathbb{R})$ mà $A^2 - 3A + 2I = 0$. Đặt $g(t) = t^2 - 3t + 2 = (t-1)(t-2)$. Khi đó $g(A) = 0$, đa thức tối tiểu $p(t)$ của $A$ chia hết $g(t)$. Do đó $p(t)$ chỉ có thể là $t-1, t-2$, và $(t-1)(t-2)$. Nếu $p(t) = t-1$ hoặc $p(t) = t-2$, thì $A = I$ hoặc $A = 2I$, tương ứng. Nếu $p(t) = (t-1)(t-2)$, thì $A$ là chéo hóa được với các giá trị riêng 1 và 2, và do đó $A$ là đồng dạng với ma trận

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}. \qquad \square$$

**2.3.9 Thí dụ.** Cho $A \in \mathcal{M}_{n\times n}(R)$ thỏa $A^3 = A$. Ta chứng minh rằng $A$ là chéo hóa được. Đặt $g(t) = t^3 - t = t(t+1)(t-1)$. Khi đó $g(A) = 0$, và do đó đa thức tối tiểu $p(t)$ của $A$ chia hết $g(t)$. Do $g(t)$ không có nhân tử lặp, nên $p(t)$ cũng không có. Vậy $A$ là chéo hóa được bởi Định lí 2.3.7. $\square$

## Bài tập

**1)** Giả sử $f(t) \in \mathbb{K}[t]$ và $A \in \mathcal{M}_n(\mathbb{K})$. Chứng minh rằng nếu $A\mathbf{x} = \lambda\mathbf{x}$ thì $f(A)\mathbf{x} = f(\lambda)\mathbf{x}$.

**2)** Nếu chúng ta dùng định nghĩa: *đa thức tối tiểu $p(\lambda)$ của ma trận $A$ là đa thức có hệ số cao nhất là 1 và có bậc dương bé nhất sao cho* $p(A) = O$, chứng minh rằng

**(a)** với mọi đa thức $f(t)$, nếu $f(A) = O$, thì $p(t)$ chia hết $f(t)$.

**(b)** đa thức cực tiểu của $A$ là duy nhất.

**3)** Chứng minh Hệ quả 2.3.2.

**4)** Tìm đa thức tối tiểu với mỗi ma trận sau. Từ đó xét xem ma trận ấy có chéo hóa được không?

**(a)** $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

**(c)** $\begin{bmatrix} 4 & -14 & 5 \\ 1 & -4 & 2 \\ 1 & -6 & 4 \end{bmatrix}$ **(d)** $\begin{bmatrix} 3 & 0 & 1 \\ 2 & 2 & 2 \\ -1 & 0 & 1 \end{bmatrix}$

**5**) Tìm đa thức tối tiểu với mỗi toán tử tuyến tính sau. Từ đó xét xem toán tử tuyến tính ấy có chéo hóa được không?

**(a)** $f$ trong $\mathbb{R}^2$, với $f(a,b) = (a+b, a-b)$

**(b)** $f$ trong $\mathbb{R}_2[x]$, với $f(p) = f' + 2f$

**(c)** $f$ trong $\mathbb{R}_2[x]$, với $f(p)(x) = -xp''(x) + p'(x) + 2p(x)$

**(d)** $f$ trong $\mathcal{M}_n(\mathbb{R})$, với $f(A) = A^t$. *Hướng dẫn*: chú ý rằng $f^2 = Id$.

**6**) Mô tả tất cả toán tử tuyến tính $f$ trong $\mathbb{R}^2$ sao cho $f$ là chéo hóa được và $f^3 - 2f^2 + f = 0$.

**7**) Cho $f$ là một toán tử tuyến tính trong không gian vector hữu hạn chiều, và cho $p(t)$ là đa thức cực tiểu của $f$. Chứng minh rằng

**(a)** $f$ là khả nghịch nếu và chỉ nếu $p(0) \neq 0$.

**(b)** Nếu $f$ là khả nghịch và $p(t) = t^n + \cdots + a_1 t + a_0$, thì

$$f^{-1} = -\frac{1}{a_0}(f^{n-1} + a_{n-1}f^{n-2} + \cdots + a_1 Id_V).$$

**8**) Cho $f$ là một toán tử tuyến tính trong không gian vector hữu hạn chiều $V$. Chứng minh rằng $V$ là một không gian con $f$-cyclic nếu và chỉ nếu mỗi không gian riêng là một chiều.

**9**) Cho $f$ là một toán tử tuyến tính trong không gian vector hữu hạn chiều $V$, và giả sử rằng $W$ là một không gian con $f$-bất biến của $V$. Chứng minh rằng đa thức tối tiểu của $f|_W$ chi hết đa thức tối tiểu của $f$.

**10)** Cho $D$ được lấy vi phân trong $P(R)$, không gian đa thức trên $R$. Chứng minh rằng không tồn tại đa thức $g(t)$ mà $g(D) = T_0$. Khi đó $D$ không có đa thức cực tiểu.

# § 3 Ma trận Jordan và khối Jordan

## 3.1 Khối Jorddan

**3.1.1 Định nghĩa.** Ma trận vuông cấp $k$ có dạng

$$A = \begin{bmatrix} \lambda_0 & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_0 & \dots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \dots & \lambda_0 & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_0 \end{bmatrix}$$

với $\lambda_0 \in \mathbb{K}$ được gọi là **khối Jordan** cấp $k$ thuộc giá trị $\lambda_0$.

**3.1.2** Với $A$ là khối Jordan như trong định nghĩa trên, ta tìm dạng chính tắc của $\lambda$-ma trận $A - \lambda I_k$

$$A - \lambda I_k = \begin{bmatrix} \lambda_0 - \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_0 - \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_0 - \lambda & \dots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \dots & \lambda_0 - \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_0 - \lambda \end{bmatrix}$$

Do $\det(A - \lambda I_k) = (\lambda_0 - \lambda)^k$ nên ta được $D_k(A - \lambda I_k) = (\lambda - \lambda_0)^k$. Mặt khác, do trong $A$ có một định lý con cấp $k-1$ là

$$D_{k-1} = \begin{vmatrix} 1 & 0 & \dots & 0 & 0 \\ \lambda_0 - \lambda & 1 & \dots & 0 & 0 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & \lambda_0 - \lambda & 1 \end{vmatrix} = 1$$

nên $D_{k-1}(A-\lambda I_k) = 1$. Vì $D_i(A-\lambda I_k)$ chia hết $D_{i+1}(A-\lambda I_k)$ với mọi $i = 1, \dots, k-1$, suy ra $D_{k-2}(A-\lambda I_k) = \cdots = D_1(A-\lambda I_k) = 1$. Ta có $e_1(\lambda) = D_1(A-\lambda I_k) = 1$, $e_i(\lambda) = \dfrac{D_i(A-\lambda I_k)}{D_{i-1}(A-\lambda I_k)} = 1$ với $i = 2\dots, k-1$, và $e_k(\lambda) = \dfrac{D_k(A-\lambda I_k)}{D_{k-1}(A-\lambda I_k)} = (\lambda-\lambda_0)^k$. Vậy

$$A-\lambda I_k \sim \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & (\lambda-\lambda_0)^k \end{bmatrix}$$

## 3.2 Phép biến đổi tuyến tính có ma trận biểu diễn là khối Jordan

**3.2.1 Mệnh đề.** *Giả sử phép biến đổi tuyến tính $f$ của không gian vector $V$ trên trường $\mathbb{K}$ có ma trận đối với cơ sở $\mathscr{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ là một ma trận khối Jordan $J$ thuộc giá trị $\lambda_1$. Khi đó*

$$(f-\lambda_1 \mathrm{Id}_V)^k(\mathbf{e}_j) = \begin{cases} \mathbf{0} & \textit{khi } k \geq j \\ \mathbf{e}_{j-k} & \textit{khi } 0 \leq k < j \leq n \end{cases}$$

*Chứng minh.* Giả sử

$$J = \begin{bmatrix} \lambda_1 & 1 & \dots & 0 & 0 \\ 0 & \lambda_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_1 & 1 \\ 0 & 0 & \dots & 0 & \lambda_1 \end{bmatrix}$$

Do đó ta có $f(\mathbf{e}_1) = \lambda_1\mathbf{e}_1$ và $f(\mathbf{e}_j) = \mathbf{e}_{j-1} + \lambda_1\mathbf{e}_j$. Ta chứng minh mệnh đề quy nạp theo $j$. Rõ ràng $(f-\lambda_1\mathrm{Id}_V)(\mathbf{e}_1) = f(\mathbf{e}_1) - \lambda_1\mathbf{e}_1 = \mathbf{0}$. Do đó

$$(f-\lambda_1 \mathrm{Id}_V)^k(\mathbf{e}_1) = \begin{cases} \mathbf{0} & \text{khi } k \geq 1 \\ \mathbf{e}_1 & \text{khi } k = 0 \end{cases}$$

nghĩa là mệnh đề đúng với $j = 1$. Giả sử mệnh đề đúng với $j - 1$. Khi đó ta có

$$\begin{aligned}
(f - \lambda_1 \mathrm{Id}_V)^k(\mathbf{e}_j) &= (f - \lambda_1 \mathrm{Id}_V)^{k-1}((f - \lambda_1 \mathrm{Id}_V)(\mathbf{e}_j)) \\
&= (f - \lambda_1 \mathrm{Id}_V)^{k-1}(\mathbf{e}_{j-1}) \\
&= \begin{cases} \mathbf{0} & \text{khi } k-1 \geq j-1 \\ \mathbf{e}_{j-1-(k-1)} & \text{khi } 0 \leq k-1 < j-1 \end{cases} \\
&= \begin{cases} \mathbf{0} & \text{khi } k \geq j \\ \mathbf{e}_{j-k} & \text{khi } 1 \leq k < j \end{cases}
\end{aligned}$$

Hơn nữa, hiển nhiên ta có $(f - \lambda_1 \mathrm{Id}_V)^0(\mathbf{e}_j) = \mathbf{e}_j$. Do đo mệnh đề đúng với $j$. □

**3.2.2 Định lý.** *Giả sử $f$ là phép biến đổi tuyến tính của không gian vector $n$ chiều $V$ trên trường $\mathbb{K}$ mà ma trận của $f$ đối với cơ sở $\mathbf{e}_1, \ldots, \mathbf{e}_n$ nào đó của $V$ là một khối Jordan $J$ thuộc giá trị $\lambda_1$. Khi đó các không gian $M_0 = \{\mathbf{0}\}$, $M_1 = \langle \mathbf{e}_1 \rangle$, $M_2 = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle, \ldots, M_n = \langle \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \rangle$ đều là các không gian con của $V$ bất biến qua $f$. Hơn thế nữa, mọi không gian con của $V$ bất biến qua $f$ đều có dạng $M_i$.*

*Chứng minh.* Rõ ràng $M_0$ là không gian con bất biến qua $f$. Mặt khác, ta có

$$J = \begin{bmatrix} \lambda_1 & 1 & \ldots & 0 & 0 \\ 0 & \lambda_1 & \ldots & 0 & 0 \\ \multicolumn{5}{c}{\ldots\ldots\ldots\ldots\ldots\ldots} \\ 0 & 0 & \ldots & \lambda_1 & 1 \\ 0 & 0 & \ldots & 0 & \lambda_1 \end{bmatrix}$$

và $f(\mathbf{e}_1) = \lambda_1 \mathbf{e}_1$, $f(\mathbf{e}_j) = \mathbf{e}_{j-1} + \lambda_1 \mathbf{e}_j$ với $2 \leq j \leq n$. Do đó, dễ dàng nhận thấy $M_1$ là không gian con bất biến qua $f$. Hơn nữa, với mỗi $k = 2, \ldots, n$ ta có $f(\mathbf{e}_1) \in M_k$ và $f(\mathbf{e}_i) \in M_k$ với mọi $i = 2, \ldots, k$. Vậy $M_k$ là không gian con bất biến qua $f$.

Ngược lại giả sử $M$ là một không gian con của $V$ bất biến qua $f$. Nếu $M = \{\mathbf{0}\}$ thì $M = M_0$ và nếu $M = V$ thì $M = M_n$. Giả sử $M \neq \{\mathbf{0}\}$

và $M \neq V$. Khi đó gọi **x** là một vector khác **0** thuộc $M$. Ta có thể giả sử $\mathbf{x} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_j\mathbf{e}_j$ với $\alpha_j \neq 0$ và $j \in \{1, \ldots, n\}$. Ta thấy

$$f(\mathbf{x}) \in M, \lambda_1\mathbf{x} \in M \qquad \text{nên} \qquad \mathbf{y} = (f - \lambda_1\mathrm{Id}_V)(\mathbf{x}) \in M.$$

Tương tự, ta cũng có

$$f(\mathbf{y}) \in M, \lambda_1\mathbf{y} \in M \qquad \text{nên} \qquad (f - \lambda_1\mathrm{Id}_V)(\mathbf{y}) \in M.$$

Suy ra

$$(f - \lambda_1\mathrm{Id}_V)^2(\mathbf{x}) = (f - \lambda_1\mathrm{Id}_V)((f - \lambda_1\mathrm{Id}_V)(\mathbf{x})) = (f - \lambda_1\mathrm{Id}_V)(\mathbf{y}) \in M.$$

Tổng quát ta được

$$(f - \lambda_1\mathrm{Id}_V)^k(\mathbf{x}) \in M, \qquad \text{với mọi } k \geq 0.$$

Nói riêng $(f - \lambda_1\mathrm{Id}_V)^{j-1}(\mathbf{x}) \in M$, trong khi đó

$$\begin{aligned}(f - \lambda_1\mathrm{Id}_V)^{j-1}(\mathbf{x}) &= (f - \lambda_1\mathrm{Id}_V)^{j-1}\Big(\sum_{l=1}^{j}\alpha_l\mathbf{e}_l\Big)\\ &= \sum_{l=1}^{j}\alpha_l(f - \lambda_1\mathrm{Id}_V)^{j-1}(\mathbf{e}_l)\\ &= \alpha_j(f - \lambda_1\mathrm{Id}_V)^{j-1}(\mathbf{e}_j)\\ &= \alpha_j\mathbf{e}_1\end{aligned}$$

cho nên $\alpha_j\mathbf{e}_1 \in M$, suy ra $\mathbf{e}_1 \in M$ và do đó $M_1 \subseteq M$. Hơn nữa, do $M \neq V$ nên tồn tại $i \in \{1, \ldots, n-1\}$ sao cho $M_i \subseteq M$ và $M_{i+1} \not\subseteq M$. Khi đó, ta khẳng định rằng $M_i = M$. Thật vậy, với mọi $x \in M \setminus \{\mathbf{0}\}$, giả sử $\mathbf{x} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_j\mathbf{e}_j$ với $\alpha_j \neq 0$ và $j \in \{1, \ldots, n\}$. Nếu $j > i$ thì $j - i - 1 \geq 0$ và $(f - \lambda_1\mathrm{Id}_V)^{j-i-1}(\mathbf{x}) \in M$ (như chứng minh ở trên). Mặt khác ta có

$$\begin{aligned}(f - \lambda_1\mathrm{Id}_V)^{j-i-1}(\mathbf{x}) &= (f - \lambda_1\mathrm{Id}_V)^{j-i-1}\Big(\sum_{k=1}^{j}\alpha_k\mathbf{e}_k\Big)\\ &= \sum_{k=1}^{j}\alpha_k(f - \lambda_1\mathrm{Id}_V)^{j-i-1}(\mathbf{e}_k)\\ &= \alpha_{j-i}\mathbf{e}_1 + \alpha_{j-i+1}\mathbf{e}_2 + \cdots + \alpha_{j-1}\mathbf{e}_i + \alpha_j\mathbf{e}_{i+1}\end{aligned}$$

Ta thấy rằng $\alpha_{j-i}\mathbf{e}_1 + \alpha_{j-i+1}\mathbf{e}_2 + \cdots + \alpha_{j-1}\mathbf{e}_i \in M_i \subseteq M$, cho nên

$$\mathbf{e}_{i+1} = \frac{1}{\alpha_j}\Big((f - \lambda_1 \mathrm{Id}_V)^{j-i-1}(\mathbf{x}) - (\alpha_{j-i}\mathbf{e}_1 + \alpha_{j-i+1}\mathbf{e}_2 + \cdots + \alpha_{j-1}\mathbf{e}_i)\Big)$$
$$\in M.$$

Suy ra $M_{i+1} \subseteq M$, đây là điều vô lý. Do đó ta phải có $j \leq i$, cho nên $\mathbf{x} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_j\mathbf{e}_j \in M_i$. Vậy $M_i = M$. □

## 3.3 Ma trận Jordan

**3.3.1 Định nghĩa.** Ma trận vuông cấp $n$, $J$, gọi là **ma trận Jordan** nếu nó có dạng

$$J = \begin{bmatrix} \lambda_1 & \varepsilon_1 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_{n-1} & \varepsilon_{n-1} \\ 0 & 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

trong đó $\lambda_i \in \mathbb{K}$, $\varepsilon_i \in \{0, 1\}$ với $i = 1, \dots, n-1$, và nếu $\varepsilon_i = 1$ thì $\lambda_i = \lambda_{i+1}$.

**3.3.2 Thí dụ.** Các ma trận chéo (nói riêng: ma trận không, ma trận đơn vị) đều là ma trận Jordan. □

**3.3.3 Nhận xét.** Theo định nghĩa ta thấy ma trận Jordan là ma trận được tạo thành từ các khối Jordan ghép lại trên đường chéo chính và các phần tử ngoài các khối Jordan ấy đều bằng 0.

**3.3.4 Thí dụ.** Ma trận sau là ma trận Jordan

$$\begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

nó được tạo thành từ bốn khối Jordan theo thứ tự là

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 \end{bmatrix} \qquad \begin{bmatrix} 5 \end{bmatrix}$$

Khối thứ nhất có cấp 3 ứng với $\lambda_0 = 3$, khối thứ hai có cấp 2 ứng với $\lambda_0 = 2$, khối thứ ba có cấp 1 ứng với $\lambda_0 = 1$, và khối thứ tư có cấp 1 ứng với $\lambda_0 = 5$. □

Vì lý do trên thay vì viết một cách tường minh ma trận Jordan ra ta có thể mô tả nó bằng cách lập bảng. Chẳng hạn, ta có bảng sau mô tả về các khối Jordan của một ma trận Jordan

| $\lambda_0$ | Số khối | Cấp của các khối | | |
|---|---|---|---|---|
| 1 | 3 | 3 | 2 | 1 |
| 0 | 3 | 2 | 1 | 1 |
| 2 | 2 | 3 | 1 | |

ma trận Jordan đó là

$$J = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2
\end{bmatrix}$$

## Bài tập

**1**) Tìm dạng chính tắc của các $\lambda$-ma trận sau:

**(a)** $\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$ **(b)** $\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$ **(c)** $\begin{bmatrix} \lambda & 1 & 1 & \dots & 1 \\ 0 & \lambda & 1 & \dots & 1 \\ 0 & 0 & \lambda & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix}$

**2**) Tìm dạng chính tắc của các $\lambda$-ma trận $A - \lambda I$ với ma trận $A$ sau

**(a)** $\begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & 1 & 1 & \dots & 1 & 1 \\ 0 & 0 & 1 & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$

# § 4 Dạng chuẩn tắc Jordan của ma trận vuông

## 4.1 Dạng chính tắc của $\lambda$-ma trận $J - \lambda I$

**4.1.1** Từ bảng mô tả các khối Jordan trong ma trận Jordan, $J$, ta có thể dễ dàng tìm dạng chính tắc của $\lambda$-ma trận $J - \lambda I_n$ (chú ý đến các cột *cấp của các khối*). Chẳng hạn, ma trận Jordan $J$ cấp 14 ở mục trước có bảng mô tả các khối Jordan

| $\lambda_0$ | Số khối | Cấp của các khối | | |
|---|---|---|---|---|
| 1 | 3 | 3 | 2 | 1 |
| 0 | 3 | 2 | 1 | 1 |
| 2 | 2 | 3 | 1 | |

Dạng chính tắc của $\lambda$-ma trận $J - \lambda I_{14}$ được xác định như sau. Ta nhận thấy chỉ có *ba* cột trong *cấp của các khối* nên $e_1(\lambda) = \cdots = e_{11}(\lambda) = 1$. Từ cột thứ ba trong *cấp của các khối* ta xác định được $e_{12}(\lambda) = (\lambda - 1)\lambda$. Từ cột thứ hai trong *cấp của các khối* ta xác định được $e_{13}(\lambda) = (\lambda - 1)^2\lambda(\lambda - 2)$, cuối cùng từ cột thứ nhất trong *cấp của các khối* ta xác định được $e_{14}(\lambda) = (\lambda - 1)^3\lambda^2(\lambda - 2)^3$. Vậy dạng chính tắc của ma trận

$J - \lambda I_{14}$ là ma trận vuông cấp 14 sau

$$\begin{bmatrix} 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 1 & 0 & 0 & 0 \\ 0 & \dots & 0 & \lambda(\lambda-1) & 0 & 0 \\ 0 & \dots & 0 & 0 & \lambda(\lambda-1)^2(\lambda-2) & 0 \\ 0 & \dots & 0 & 0 & 0 & \lambda^2(\lambda-1)^3(\lambda-2)^3 \end{bmatrix}$$

Ta có được cách tìm dạng chính tắc của $\lambda$-ma trận $J - \lambda I$ trên nhờ định lý sau.

**4.1.2 Định lý.** *Giả sử $J$ là ma trận Jordan, trong đó $k_i$ khối Jordan thuộc giá trị $\lambda_i$ với cấp là $m_{i1} \geq m_{i2} \geq \cdots \geq m_{ik_i}$ với $i = 1, \dots, s$. Bằng cách thêm số $0$ vào phía sau dãy trên, ta có thể giả sử tất cả các dãy này đều có cùng độ dài $k = \max\{k_1, k_2, \dots, k_s\}$,*

$$m_{i1} \geq m_{i2} \geq \cdots \geq m_{ik}, \qquad i = 1, 2, \dots, s.$$

*Khi đó, $\lambda$-ma trận $J - \lambda I$ có dạng chính tắc xác định bởi các nhân tử bất biến sau:*

$$\begin{aligned} e_n(\lambda) &= (\lambda - \lambda_1)^{m_{11}}(\lambda - \lambda_2)^{m_{21}} \cdots (\lambda - \lambda_s)^{m_{s1}} \\ e_{n-1}(\lambda) &= (\lambda - \lambda_1)^{m_{12}}(\lambda - \lambda_2)^{m_{22}} \cdots (\lambda - \lambda_s)^{m_{s2}} \\ &\vdots \\ e_{n-k+1}(\lambda) &= (\lambda - \lambda_1)^{m_{1k}}(\lambda - \lambda_2)^{m_{2k}} \cdots (\lambda - \lambda_s)^{m_{sk}} \\ e_{n-k}(\lambda) &= e_{n-k-1}(\lambda) = \cdots = e_1(\lambda) = 1. \end{aligned}$$

*Chứng minh.* Ta thay mỗi khối Jordan bởi dạng chính tắc của nó để được một $\lambda$-ma trận mới tương đương. Từ đó ta tính được dễ dàng hơn các biểu thức $D_k(J)$:

$$\begin{aligned} D_n(J) &= (\lambda - \lambda_1)^{m_{11}+m_{12}+\cdots+m_{1k}} \cdots (\lambda - \lambda_s)^{m_{s1}+m_{s2}+\cdots+m_{sk}} \\ D_{n-1}(J) &= (\lambda - \lambda_1)^{m_{12}+\cdots+m_{1k}} \cdots (\lambda - \lambda_s)^{m_{s2}+\cdots+m_{sk}} \\ &\vdots \\ D_{n-k+1}(J) &= (\lambda - \lambda_1)^{m_{1k}} \cdots (\lambda - \lambda_s)^{m_{sk}} \end{aligned}$$

và $D_{n-k}(J) = \cdots = D_1(J) = 1$. Theo công thức tính $e_k(\lambda)$ qua các $D_k(J)$ ta có được kết quả của định lý. □

**4.1.3 Thí dụ.** Tìm dạng chính tắc của $\lambda$-ma trận $J - \lambda I_{15}$ trong đó $J$ là ma trận Jordan có bảng mô tả các khối Jordan trong nó

| $\lambda_0$ | Số khối | Cấp của các khối | | |
|---|---|---|---|---|
| $-1$ | 2 | 3 | 2 | |
| 0 | 3 | 3 | 1 | 1 |
| 2 | 2 | 2 | 2 | |
| 3 | 1 | 1 | | |

Từ bảng trên ta suy ra được $e_1(\lambda) = \cdots = e_{12}(\lambda) = 1$, $e_{13}(\lambda) = \lambda$, $e_{14}(\lambda) = (\lambda+1)^2\lambda(\lambda-2)^2$, và $e_{15}(\lambda) = (\lambda+1)^3\lambda^3(\lambda-2)^2(\lambda-3)$. Do đó, $\lambda$-ma trận chính tắc cấp 15 cần tìm là

$$\begin{bmatrix} 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 1 & 0 & 0 & 0 \\ 0 & \dots & 0 & \lambda & 0 & 0 \\ 0 & \dots & 0 & 0 & \lambda(\lambda+1)^2(\lambda-2)^2 & 0 \\ 0 & \dots & 0 & 0 & 0 & \lambda^3(\lambda+1)^3(\lambda-2)^2(\lambda-3) \end{bmatrix}$$ □

**4.1.4** Nếu biết dạng chính tắc của ma trận $J - \lambda I$ trong đó $J$ là ma trận Jordan, thì ta có thể lập bảng mô tả các khối Jordan trong ma trận $J$ từ đó ta lập được ma trận $J$ (có thể khác nhau về thứ tự các khối). Chẳng hạn ta có dạng chính tắc của ma trận $J - \lambda I$ cấp 15

$$\begin{bmatrix} 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 1 & 0 & 0 & 0 \\ 0 & \dots & 0 & \lambda(\lambda+1) & 0 & 0 \\ 0 & \dots & 0 & 0 & \lambda^2(\lambda+1)(\lambda-2)^2 & 0 \\ 0 & \dots & 0 & 0 & 0 & \lambda^3(\lambda+1)^2(\lambda-2)^3 \end{bmatrix}$$

Các $\lambda_0$ chỉ nhận 3 giá trị là $-1$, 0 và 2. Ta nhận thấy số khối của một giá trị $\lambda_0$ nhiều nhất là 3. Ta có bảng mô tả các khối Jordan trong ma trận $J$ là

| $\lambda_0$ | Số khối | Cấp của các khối | | |
|---|---|---|---|---|
| $-1$ | 3 | 2 | 1 | 1 |
| 0 | 3 | 3 | 2 | 1 |
| 2 | 2 | 3 | 2 | |

## 4.2 Dạng chuẩn tắc Jordan của ma trận vuông

**4.2.1 Định nghĩa.** Cho ma trận $A$ vuông cấp $n$. Nếu $A$ đồng dạng với ma trận Jordan $J$, thì $J$ là duy nhất (không kể thứ tự các khối Jordan). Khi đó, ma trận $J$ được gọi là **dạng chuẩn tắc Jordan** của $A$.

Phép đưa ma trận vuông $A$ về dạng chuẩn tắc Jordan là việc đi tìm ma trận Jordan $J$ đồng dạng với ma trận $A$.

**4.2.2 Định lý.** *Cho $A$ là một ma trận vuông cấp $n$ trên trường $\mathbb{K}$. Khi đó, $A$ đồng dạng với một ma trận Jordan khi và chỉ khi đa thức đặc trưng của $A$ phân tích được thành tích của $n$ đa thức bậc nhất trên trường $\mathbb{K}$.*

*Chứng minh.* Giả sử đa thức đặc trưng của $A$ phân tích được thành tích của $n$ đa thức bậc nhất. Nghĩa là $|A-\lambda I| = (-1)^n(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_s)^{k_s}$ với $\lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{K}$ và $k_1+k_2+\cdots+k_s = n$. Khi đó trong $\lambda$-ma trận $A-\lambda I$ ta có: $D_n(A) = (\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_s)^{k_s}$. Dạng chính tắc của $A-\lambda I$ là

$$E(\lambda) = \begin{bmatrix} e_1(\lambda) & 0 & \ldots & 0 \\ 0 & e_2(\lambda) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & e_n(\lambda) \end{bmatrix}$$

trong đó các $e_i(\lambda)$ đều phân tích được thành tích của các đa thức bậc nhất trên trường $\mathbb{K}$. Gọi $J$ là ma trận Jordan mà dạng chính tắc của $J-\lambda I$ là $E(\lambda)$ thì $A$ và $J$ đồng dạng.

Ngược lại, giả sử $A$ đồng dạng với một ma trận Jordan $J$. Khi đó $A-\lambda I$ và $J-\lambda I$ là hai $\lambda$-ma trận tương đương nên có cùng dạng chính tắc, do đó $D_n(A) = D_n(J)$. Vì $D_n(J)$ phân tích được thành tích các đa thức bậc nhất nên $D_n(A)$ cũng vậy. Do đó, $|A-\lambda I|$ phân tích được thành tích các đa thức bậc nhất. □

### 4.2.3 Cách tìm dạng chuẩn tắc Jordan của $A$

- Lập ma trận $A - \lambda I$ và đưa nó về dạng chính tắc $E(\lambda)$.
- Tìm ma trận Jordan $J$ sao cho $J - \lambda I$ có dạng chính tắc cũng là $E(\lambda)$.

Giả sử $E(\lambda) = \begin{bmatrix} e_1(\lambda) & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & e_n(\lambda) \end{bmatrix}$. Có hai trường hợp xảy ra

**(i)** Tồn tại $e_i(\lambda)$ khác 1 mà không phân tích được thành tích các đa thức bậc 1. Khi đó, ta kết luận $A$ không đồng dạng với ma trận Jordan hay $A$ không có dạng chuẩn Jordan.

**(ii)** Mọi $e_i(\lambda)$ khác 1 đều phân tích được thành tích các đa thức bậc 1. Khi đó $A$ có dạng chuẩn tắc Jordan.

**4.2.4 Thí dụ.** Xét ma trận $A = \begin{bmatrix} 0 & -4 & 0 \\ 1 & -4 & 0 \\ 1 & -2 & -2 \end{bmatrix}$. Ta có $\lambda$-ma trận $A - \lambda I$

$$A - \lambda I = \begin{bmatrix} -\lambda & -4 & 0 \\ 1 & -4-\lambda & 0 \\ 1 & -2 & -2-\lambda \end{bmatrix}$$

Ta tính được $D_1 = 1$, $D_2 = \gcd((\lambda+2)^2, 2(\lambda+2), \lambda(\lambda+2), 4(\lambda+2), \lambda+2, -(\lambda+2), (\lambda+2)(\lambda+4)) = \lambda+2$, $D_3 = (\lambda+2)^3$. Do đó, dạng chính tắc của $A - \lambda I$ là

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda+2 & 0 \\ 0 & 0 & (\lambda+2)^2 \end{bmatrix}.$$

Vậy $A$ đồng dạng với một ma trận Jordan, và dạng chuẩn tắc Jordan của $A$ là $J = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$. □

## Bài tập

**1)** Hãy tìm dạng chuẩn tắc Jordan $J$ của các ma trận $A$ sau.

**(a)** $\begin{bmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{bmatrix}$ **(c)** $\begin{bmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{bmatrix}$

**(d)** $\begin{bmatrix} 9 & -6 & -2 \\ 18 & -12 & -3 \\ 18 & -9 & -6 \end{bmatrix}$ **(e)** $\begin{bmatrix} 4 & 6 & -15 \\ 1 & 3 & -5 \\ 1 & 2 & -4 \end{bmatrix}$ **(f)** $\begin{bmatrix} 12 & -6 & -2 \\ 18 & -9 & -3 \\ 18 & -9 & -3 \end{bmatrix}$

**(g)** $\begin{bmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{bmatrix}$ **(h)** $\begin{bmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{bmatrix}$ **(i)** $\begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ \alpha & 0 & \alpha \end{bmatrix}$

**2**) Hãy tìm dạng chuẩn tắc Jordan $J$ của các ma trận $A$ sau.

**(a)** $\begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 2 & 0 & \dots & 0 & 0 \\ 1 & 2 & 3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 2 & 3 & \dots & n-1 & 0 \\ 1 & 2 & 3 & \dots & n-1 & n \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ 0 & 1 & 2 & \dots & n-2 & n-1 \\ 0 & 0 & 1 & \dots & n-3 & n-2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 2 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}$

**(c)** $\begin{bmatrix} n & n-1 & n-2 & \dots & 2 & 1 \\ 0 & n & n-1 & \dots & 3 & 2 \\ 0 & 0 & n & \dots & 4 & 3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n & n-1 \\ 0 & 0 & 0 & \dots & 0 & n \end{bmatrix}$

**(d)** $\begin{bmatrix} \alpha & a_{12} & a_{13} & \dots & a_{1,n-1} & a_{1n} \\ 0 & \alpha & a_{23} & \dots & a_{2,n-1} & a_{2n} \\ 0 & 0 & \alpha & \dots & a_{3,n-1} & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \alpha & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & \alpha \end{bmatrix}$

# § 5 Chuẩn tắc hóa Jordan và ma trận chuẩn tắc hóa Jordan

## 5.1 Phép biến đổi lũy linh

**5.1.1 Định nghĩa.** Ta nói một phép biến đổi tuyến tính $f$ trên không gian vector $V$ là **lũy linh** nếu $f^k = 0$ với số nguyên dương $k$ nào đó.

**5.1.2 Thí dụ.** Nếu phép biến đổi tuyến tính $f$ có ma trận đối với cơ sở $\mathscr{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ là một khối Jordan thuộc giá trị $\lambda$ thì $f - \lambda\mathrm{Id}$ là một phép biến đổi lũy linh. Thật vậy theo Mệnh đề 3.2.1, ta có $(f - \lambda\mathrm{Id})^n(\mathbf{e}_j) = \mathbf{0}$ với mọi $j = 1, \dots, n$, cho nên $(f - \lambda\mathrm{Id})^n = 0$. Trường hợp đặc biệt $\lambda = 0$, ta có $f$ là lũy linh. Hơn nữa, ta có một phần của điều ngược lại, được nêu trong mệnh đề sau. □

**5.1.3 Mệnh đề.** *Nếu $f$ là phép biến đổi tuyến tính lũy linh thì nó chỉ có giá trị riêng là* $0$.

*Chứng minh.* Vì $f$ lũy linh nên tồn tại $k$ nguyên dương sao cho $f^k = 0$. Giả sử $\lambda$ là giá trị riêng của $f$ và một vector riêng $\mathbf{v}$ tương ứng. Khi đó ta có $f(\mathbf{v}) = \lambda\mathbf{v}$ và $f^k(\mathbf{v}) = \lambda^k\mathbf{v}$. Do $f^k = 0$ nên ta phải có $\lambda^k = 0$ hay $\lambda = 0$. □

Từ kết quả của mệnh đề trên ta nhận thấy một phép biến đổi lũy linh có giá trị riêng $0$ với số bội đại số bằng $n$. Ta xét một phép biến đổi lũy linh $f$ trên $V$. Khi đó gọi $k$ là nguyên dương nhỏ nhất sao cho $f^k = 0$, nghĩa là $f^{k-1} \neq 0$. Do đó trong $V$ tồn tại một vector $\mathbf{v}_k$ sao cho $f^{k-1}(\mathbf{v}_k) \neq \mathbf{0}$. Đặt $\mathbf{v}_{k-1} = f(\mathbf{v}_k)$, $\mathbf{v}_{k-2} = f(\mathbf{v}_{k-1}) = f^2(\mathbf{v}_k)$, ..., $\mathbf{v}_1 = f(\mathbf{v}_2) = f^{k-1}(\mathbf{v}_k)$. Tính chất của các vector $\mathbf{v}_1, \dots, \mathbf{v}_k$ được nêu trong mệnh đề sau.

**5.1.4 Mệnh đề.** *Với các ký hiệu như trên, hệ vector $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ là độc lập tuyến tính.*

*Chứng minh.* Giả sử ta có tổ hợp tuyến tính $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_k\mathbf{v}_k = \mathbf{0}$. Lần lượt tính giá trị phép biến đổi tuyến tính $f$, $f^2$, ..., $f^{k-1}$ tại tổ hợp tuyến tính này ta được

$$\begin{aligned}
\lambda_2\mathbf{v}_1 + \lambda_3\mathbf{v}_2 + \cdots + \lambda_k\mathbf{v}_{k-1} &= \mathbf{0} \\
\lambda_3\mathbf{v}_1 + \cdots + \lambda_k\mathbf{v}_{k-2} &= \mathbf{0} \\
&\vdots \\
\lambda_k\mathbf{v}_1 &= \mathbf{0}
\end{aligned}$$

Từ đó suy ngược ra được $\lambda_k = 0, \ldots, \lambda_2 = 0, \lambda_1 = 0$. Vậy hệ vector $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ là độc lập tuyến tính. □

**5.1.5 Định nghĩa.** Các vector $\mathbf{v}_1, \ldots, \mathbf{v}_k$ trong mệnh đề trên được gọi là **xâu Jordan** (hay dây Jordan) độ dài $k$, vector $\mathbf{v}_k$ được gọi là **vector sinh** của xâu và $\mathbf{v}_1$ được gọi là **vector cuối** của xâu.

Ta nhận thấy $\mathbf{v}_1$ là một vector riêng của $f$. Mặt khác, ta có $k \leq \dim V = n$, nên có thể bổ sung vào hệ độc lập tuyến tính $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ một cơ sở $(\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n)$ của $\langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle^\perp$ để được một cơ sở $(\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_n)$. Khi đó, ma trận của $f$ đối với cơ sở này có dạng

$$\begin{bmatrix}
0 & 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\
\multicolumn{8}{c}{\dotfill} \\
0 & 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\
0 & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & 0 & 0 & \ldots & 0 & \times & \ldots & \times \\
\multicolumn{8}{c}{\dotfill} \\
0 & 0 & 0 & \ldots & 0 & \times & \ldots & \times
\end{bmatrix}$$

trong đó ma trận con với $k$ dòng đầu và $k$ cột đầu là một khối Jordan thuộc giá trị 0. Khi thu hẹp $f$ trên không gian $\langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle^\perp$ ta cũng được phép biến đổi lũy linh. Tiến hành quá trình như trên ta tìm được một xâu Jordan có chiều dài ít nhất bằng 1 trong không gian $\langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle^\perp$. Và tiếp tục như thế không quá $n - 1$ lần ta tìm được các xâu Jordan, các vector trong các xâu Jordan lập thành một cơ sở của $V$ và ma trận của $f$ đối với cơ sở này là một ma trận Jordan với phần tử trên đường chéo chính bằng 0. Như vậy vấn đề trong việc tìm cơ sở mà ma trận của $f$ là một ma trận Jordan trở thành việc tìm các vector sinh của các xâu Jordan.

Tiếp theo, giả sử phép biến đổi tuyến tính $f$ có dạng $\lambda \mathrm{Id} + L$ trong đó $L$ là phép biến đổi lũy linh, nghĩa là $f$ chỉ có một giá trị riêng duy nhất là $\lambda$ với số bội đại số là $n$. Nếu chúng ta có thể tìm một cơ sở như đã tiến hành ở trên sao cho ma trận của $L$ đối với cơ sở này là ma trận Jordan $J$ với các phần tử nằm trên đường chéo là 0, thì ma trận của $f$ đối với cơ sở này sẽ là $\lambda I + J$, đây cũng là ma trận Jordan với các phần tử nằm trên đường chéo đều bằng $\lambda$.

**5.1.6 *Chú ý.*** Ta biết rằng có sự tương quan giữa phép biến đổi tuyến tính trên không gian vector hữu hạn chiều và ma trận, cho nên chúng ta cũng có khái niệm *ma trận lũy linh*. Cụ thể là: ma trận vuông $A$ được gọi là lũy linh nếu tồn tại số nguyên $k > 0$ sao cho $A^k = O$.

**5.1.7 Thí dụ. (OSV12)** **(a)** Chứng tỏ rằng ma trận tam giác trên có đường chéo chính toàn 0 là ma trận lũy linh và các ma trận này lập thành một không gian con $V_0$ của không gian $\mathcal{M}_n(\mathbb{R})$ các ma trận vuông cấp $n$ trên trường số thực. Tính $\dim V_0$.

**(b)** Giả sử $V$ là một không gian con nào đó của $\mathcal{M}_n(\mathbb{R})$ mà các phần tử của nó đều là ma trận lũy linh. Chứng minh rằng $\dim V \leq \dfrac{n^2 - n}{2}$.

*Giải.* **(a)** Nhận thấy rằng mọi ma trận tam giác trên $A$ cấp $n$ với đường chéo toàn 0 có đa thức đặc trưng là $\chi(\lambda) = (-\lambda)^n$. Do đó theo định lý Cayley-Hamilton ta được $(-A)^n = O$ hay $A^n = O$. Vậy $A$ là ma trận lũy linh. Rõ ràng tổ hợp tuyến tính của hai ma trận tam giác trên có đường chéo toàn 0 là một ma trận tam giác trên có đường chéo toàn 0. Vậy $V_0$ là một không gian con của $\mathcal{M}_n(\mathbb{R})$. Ta cũng nhận thấy số các vị trí mà các phần tử trong ma trận tam giác trên có đường chéo toàn 0 là $\frac{(n-1)n}{2}$. Do đó $\dim(V_0) = \frac{n^2-n}{2}$.

**(b)** Nếu $A \neq O$ là ma trận đối xứng thực, thì $A$ chéo hóa được. Từ đó suy ra các giá trị riêng của $A^k$ là $\lambda_1^k, \ldots, \lambda_n^k$, trong đó $\lambda_1, \ldots, \lambda_n$ là các giá trị riêng của $A$ (cho phép lặp lại). Suy ra $A$ không thể lũy linh bởi vì nếu ngược lại thì từ tính chất ma trận lũy linh chỉ có duy nhất giá trị riêng bằng 0 suy ra các giá trị $\lambda_1, \ldots, \lambda_n$ đều bằng 0 nên do tính chéo hóa $A$ là ma trận không (vô lý). Như vậy tập các ma trận đối xứng thực khác ma trận không $S \setminus \{O\}$ không giao với $V$ và $\dim S = \frac{n^2+n}{2}$. Do đó, ta có

$$\dim V = \dim(V+S) + \dim(V \cap S) - \dim S \leq n^2 - \frac{n^2+n}{2} = \frac{n^2-n}{2}.$$

□

## 5.2 Chuẩn tắc hóa Jordan

Cho trường hợp tổng quát với phép biến đổi tuyến tính $f$ có các giá trị riêng phân biệt nhưng có tổng bội đại số bằng $n$, có thể chứng minh được rằng với mỗi giá trị riêng ta tìm được một dãy các vector mà $f$ bất biến đối với không gian con sinh bởi các vector này và ma trận của $f$ đối với các vector này trong không gian sinh ấy là ma trận Jordan ứng với giá trị riêng này. Nó được chứng minh rằng gôm tất cả các vector như thế với nhau cho các giá trị riêng phân biệt cho ta một cơ sở của $V$ và ma trận của $f$ đối với cơ sở này là một ma trận Jordan. (Sự tồn tại của cơ sở này có thể suy ra được từ Định lý 4.2.2.)

Chú ý rằng các vector thành lập cơ sở cho không gian riêng con tương ứng giá trị riêng $\lambda_i$ là các vector cuối trong các xâu Jordan ứng với giá trị riêng $\lambda_i$. Khi tất các các xâu Jordan đều có chiều dài 1 thì ma trận Jordan trở thành ma trận chéo. Giả sử một phép biến đổi tuyến tính $f$ có ma trận đối với cơ sở $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6, \mathbf{v}_7, \mathbf{v}_8, \mathbf{v}_9, \mathbf{v}_{10})$ có dạng chuẩn tắc Jordan

$$\begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4
\end{bmatrix}$$

Vector $\mathbf{v}_3$ là vector sinh của xâu Jordan độ dài 3 ứng với giá trị riêng 1. Vector $\mathbf{v}_5$ là vector sinh của xâu Jordan độ dài 2 ứng với giá trị riêng 1. Vector $\mathbf{v}_9$ là vector sinh của xâu Jordan độ dài 2 ứng với giá trị riêng 4. Các vector $\mathbf{v}_1$ và $\mathbf{v}_4$ lập thành cơ sở cho không gian con riêng của giá trị riêng 1. Các vector $\mathbf{v}_6$ và $\mathbf{v}_7$ lập thành cơ sở cho không gian con riêng của giá trị riêng 2. Các vector $\mathbf{v}_8$ và $\mathbf{v}_{10}$ lập thành cơ sở cho không gian con riêng của giá trị riêng 4.

Bây giờ chúng ta chứng minh lại điều kiện đủ cho sự tồn tại của một cơ sở để phép biến đổi tuyến tính có ma trận dạng chuẩn tắc Jordan đối với cơ sở ấy.

**5.2.1 Định lý.** *Cho $V$ là một không gian vector $n$ chiều trên trường $\mathbb{K}$ và $f : V \to V$ là một phép biến đổi tuyến tính. Nếu $f$ có $n$ giá trị riêng (kể cả bội) thì tồn tại một cơ sở cho $V$ sao cho ma trận của $f$ đối với cơ sở này có dạng chuẩn tắc Jordan.*

*Chứng minh.* Đầu tiên chúng ta giả sử rằng 0 là một trong các giá trị riêng của $f$. Chúng ta chứng minh định lý bằng qui nạp theo $n$. Khi $n = 1$, ta có được kết luận của định lý bằng cách chọn một vector khác không bất kỳ làm cơ sở. Giả sử định lý đúng với mọi không gian có số chiều nhỏ hơn $n$. Ta xét không gian vector $V$ với số chiều $n$ và phép biến đổi tuyến tính. Do $f$ có một giá trị riêng là 0 nên $f$ không là đơn ánh, cho nên $\mathrm{Im} f \neq V$ và $r = \dim(\mathrm{Im} f) < n$. Rõ ràng thu hẹp $f|_{\mathrm{Im} f}$ là một phép biến đổi tuyến tính trên $\mathrm{Im} f$. Do đó theo giả thiết qui nạp tồn tại một cơ sở $(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ cho $\mathrm{Im} f$ sao cho ma trận của $f|_{\mathrm{Im} f}$ đối với cơ sở này có dạng chuẩn tắc Jordan.

Có $p$ vector trong cơ sở Jordan $(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ mà chúng là các vector cuối của $p$ xâu Jordan ứng với giá trị riêng 0. Chú ý rằng có thể có $p = 0$ nếu 0 không là giá trị riêng của $f|_{\mathrm{Im} f}$. Chúng ta gọi $p$ vector này là $\mathbf{z}_1, \ldots, \mathbf{z}_p$ và các vector sinh của các xâu này là $\mathbf{w}_1, \ldots, \mathbf{w}_p$; chúng ta gọi độ dài của xâu thứ $i$ là $k(i)$. Theo định nghĩa của xâu Jordan ta nhận thấy các vector trong $p$ xâu Jordan đang nói đến đều có mặt trong cơ sở $(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ của $\mathrm{Im} f$. Do $\mathbf{w}_j \in \mathrm{Im} f$ nên tồn tại $\mathbf{u}_j$ sao cho $f(\mathbf{u}_j) = \mathbf{w}_j$. Các vector $\mathbf{u}_j$ sẽ là các vector sinh cho các xâu Jordan độ dài $k(j)+1$ với các vector cuối $\mathbf{z}_j$. Số chiều của không gian con $\ker f$ là $n-r$ và chúng ta có $p$ vector độc lập tuyến tính $\mathbf{z}_1,\dots,\mathbf{z}_p$ trong $\ker f$. Do đó ta có thể tìm $n-r-p$ vector $\mathbf{x}_1,\dots,\mathbf{x}_{n-r-p}$ sao cho chúng cùng với các vector $\mathbf{z}_1,\dots,\mathbf{z}_p$ tạo thành một cơ sở của $\ker f$.

Chúng ta cho rằng khi bổ sung các vector $\mathbf{x}_1,\dots,\mathbf{x}_{n-r-p}$ và $\mathbf{u}_1,\dots,\mathbf{u}_p$ vào hệ vector $(\mathbf{v}_1,\dots,\mathbf{v}_r)$ sẽ tạo thành một cơ sở Jordan của $V$ cho $f$, nghĩa là ma trận của $f$ đối với cơ sở này là một ma trận chuẩn tắc Jordan. Đầu tiên chúng ta thấy hệ vector mới này có $r+(n-r-p)+p=n$ vector. Do đó ta chỉ cần chứng minh $(\mathbf{v}_1,\dots,\mathbf{v}_r,\mathbf{x}_1,\dots,\mathbf{x}_{n-r-p},\mathbf{u}_1,\dots,\mathbf{u}_p)$ là một hệ sinh của $V$. Lấy $\mathbf{v}\in V$ tùy ý. Khi đó $f(\mathbf{v})\in \mathrm{Im} f$ và do đó chúng ta có thể biểu diễn $f(\mathbf{v})=\sum\limits_{i=1}^{r} a_i\mathbf{v}_i$.

Mặt khác, ta có $p$ xâu Jordan trong $(\mathbf{v}_1,\dots,\mathbf{v}_r)$ ứng với giá trị riêng $0$ là $\mathbf{z}_1,\dots,\mathbf{w}_1;\ \dots;\ \mathbf{z}_p,\dots,\mathbf{w}_p$. Ta đã biết $f(\mathbf{u}_j)=\mathbf{w}_j$ với $j=1,\dots,p$. Khi $\mathbf{v}_i\in\{\mathbf{z}_1,\dots,\mathbf{w}_1,\dots,\mathbf{z}_p,\dots,\mathbf{w}_p\}\setminus\{\mathbf{w}_1,\dots,\mathbf{w}_p\}$ theo sự định nghĩa của xâu Jordan tồn tại $\mathbf{v}_j\in\{\mathbf{z}_1,\dots,\mathbf{w}_1,\dots,\mathbf{z}_p,\dots,\mathbf{w}_p\}$ sao cho $f(\mathbf{v}_j)=\mathbf{v}_i$. Với $\mathbf{v}_i\in\{\mathbf{v}_1,\dots,\mathbf{v}_r\}\setminus\{\mathbf{z}_1,\dots,\mathbf{w}_1,\dots,\mathbf{z}_p,\dots,\mathbf{w}_p\}$. Khi đó $\mathbf{v}_i$ phải nằm trong một xâu Jordan của $f|_{\mathrm{Im} f}$ ứng với giá trị riêng khác không, gọi xâu này là $\mathbf{v}'_1,\dots,\mathbf{v}'_k$ (các vector nằm trong cơ sở $(\mathbf{v}_1,\dots,\mathbf{v}_r)$). Đặt $S=\langle \mathbf{v}'_1,\dots,\mathbf{v}'_k\rangle$. Theo Định lý 3.2.2 ta có $S$ bất biến qua $f$. Hơn nữa $\dim S=k$ vì ma trận khối Jordan không suy biến. Do đó $f|_S$ là đẳng cấu tuyến tính trên $S$, cho nên tồn tại $\mathbf{y}\in S$ sao cho $f(\mathbf{y})=\mathbf{v}_i$. Ta có $\mathbf{y}$ là tổ hợp tuyến tính của các vector $\mathbf{v}'_1,\dots,\mathbf{v}'_k$.

Như vậy ta đã chứng minh được với mỗi $\mathbf{v}_i\in\{\mathbf{v}_1,\dots,\mathbf{v}_r\}$, ta có thể biểu diễn $\mathbf{v}_i$ như là một tổ hợp tuyến tính của các vector $f(\mathbf{v}_1),\dots,f(\mathbf{v}_r)$, $f(\mathbf{u}_1),\dots,f(\mathbf{u}_r)$. Từ đó ta có thể biểu diễn

$$f(\mathbf{v})=\sum_{i=1}^{r}a_i\mathbf{v}_i=\sum_{i=1}^{r}b_if(\mathbf{v}_i)+\sum_{i=1}^{p}c_jf(\mathbf{u}_j)=f\Big(\sum_{i=1}^{r}b_i\mathbf{v}_i+\sum_{i=1}^{p}c_j\mathbf{u}_j\Big).$$

Do đó ta có $f\Big(\mathbf{v}-\sum\limits_{i=1}^{r}b_i\mathbf{v}_i-\sum\limits_{i=1}^{p}c_j\mathbf{u}_j\Big)=\mathbf{0}$. Vậy $\mathbf{v}-\sum\limits_{i=1}^{r}b_i\mathbf{v}_i-\sum\limits_{i=1}^{p}c_j\mathbf{u}_j$ biểu diễn được bởi tổ hợp tuyến tính của các vector $\mathbf{z}_1,\dots,\mathbf{z}_p,\mathbf{x}_1,\dots,\mathbf{x}_{n-r-p}$. Hơn nữa, $\mathbf{z}_1,\dots,\mathbf{z}_p\in\{\mathbf{v}_1,\dots,\mathbf{v}_r\}$, cho nên ta có thể thấy $\mathbf{v}$ biểu diễn tuyến tính được qua các vector $\mathbf{v}_1,\dots,\mathbf{v}_r,\mathbf{x}_1,\dots,\mathbf{x}_{n-r-p},\mathbf{u}_1,\dots,\mathbf{u}_p$.

Theo cách xây dựng chứng minh trên, ma trận của $f$ đối với cơ sở gồm các vector $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{x}_1, \ldots, \mathbf{x}_{n-r-p}, \mathbf{u}_1, \ldots, \mathbf{u}_p$ sẽ có dạng chuẩn tắc Jordan. Nó sẽ có tất cả các khối Jordan như trong $f|_{\mathrm{Im} f}$ đối với các giá trị riêng khác 0. Với giá trị riêng 0 sẽ có $n-r-p$ xâu Jordan với độ dài 1 tương ứng với các $\mathbf{x}_i$ và $p$ xâu Jordan được nói ban đầu trong $(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ với độ dài $k(i)$ trở thành các xâu có độ dài $k(i)+1$ với các vector sinh $\mathbf{v}_i$. Điều này hoàn thành việc chứng minh khi $f$ nhận 0 là giá trị riêng.

Trong trường hợp tổng quát, giả sử một trong các giá trị riêng của $f$ là $\lambda$. Khi đó $f - \lambda \mathrm{Id}$ sẽ có một giá trị riêng là 0, cho nên theo kết quả trên tồn tại một cơ sở sao cho ma trận của $f - \lambda \mathrm{Id}$ đối với cơ sở này có dạng chuẩn tắc Jordan $J$. Do đó ma trận của $f$ đối với cơ sở này cũng có dạng chuẩn tắc Jordan (thu được bằng cách cộng $\lambda$ vào tất cả các phần tử trên đường chéo), $J + cI$. □

Chìa khóa để tìm một cơ sở Jordan cho phép biến đổi tuyến tính là tìm các vector sinh của các xâu Jordan và chúng ta tìm các vector sinh ấy một cách độc lập theo các giá trị riêng phân biệt. Cách để tìm vector sinh ứng giá trị riêng $\lambda$ là chúng ta xét liên tiếp các không gian không của $(f - \lambda \mathrm{Id})^k$ với số mũ $k$ lớn dần. Một cách khác, chúng ta có thể tìm vector cuối của xâu Jordan và tiến hành ngược lại để tìm đến vector sinh.

**5.2.2 Định nghĩa.** Giả sử ma trận vuông $A$ có dạng chuẩn tắc Jordan. Khi đó việc tìm các ma trận $P$ và $J$ trong mệnh đề trên gọi là **chuẩn tắc hóa Jordan** ma trận $A$, và ma trận $P$ gọi là **ma trận chuẩn tắc hóa Jordan** $A$.

**5.2.3 Mệnh đề.** *Nếu ma trận vuông $A$ cấp $n$ có dạng chuẩn tắc Jordan $J$, thì tồn tại một ma trận không suy biến $P$ sao cho $P^{-1}AP = J$.*

*Chứng minh.* Điều này là hiển nhiên theo định nghĩa dạng chuẩn tắc Jordan của một ma trận. Ta có $A$ đồng dạng với $J$ nên tồn tại ma trận $P$ sao cho $J = P^{-1}AP$. □

**5.2.4 Thí dụ.** Ta biết rằng ma trận $A = \begin{bmatrix} 0 & -4 & 0 \\ 1 & -4 & 0 \\ 1 & -2 & -2 \end{bmatrix}$ có dạng chuẩn tắc $J = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$. Ta sẽ tìm ma trận $P$ sao cho $P^{-1}AP = J$. Ta có thể thấy ma trận $A$ có một giá trị riêng bội 3 là $\lambda = -2$. Ta tìm không gian con riêng ứng với giá trị riêng $\lambda = -2$. Ta biến đổi

$$A + 2I = \begin{bmatrix} 2 & -4 & 0 \\ 1 & -2 & 0 \\ 1 & -2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Vậy $E(-2) = \langle (2, 1, 0), (0, 0, 1) \rangle$. Do đó chúng ta phải có hai xâu Jordan. Do $(A + 2I)^2 = O$ và $(1, 1, 1) \notin E(-2)$, ta tính

$$\begin{bmatrix} 2 & -4 & 0 \\ 1 & -2 & 0 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ -1 \end{bmatrix}$$

Ma trận $P$ được lập bởi các vector $(-2, -1, -1)$, $(1, 1, 1)$, $(0, 0, 1)$. Vậy $P = \begin{bmatrix} -2 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$ thỏa điều kiện $P^{-1}AP = J$. □

**5.2.5 Thí dụ.** Cho ma trận $A = \begin{bmatrix} 3 & -1 & 1 & 1 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 2 \\ -1 & 1 & 1 & 3 \end{bmatrix}$. Khi đó chúng ta có đa thức đặc trưng của $A$ là $\chi_A(\lambda) = (\lambda - 4)(\lambda - 2)^3$. Ta tính được

$$A - 4I = \begin{bmatrix} -1 & -1 & 1 & 1 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & -2 & 2 \\ -1 & 1 & 1 & -1 \end{bmatrix}$$

Từ đó suy ra được $E(4) = \langle(1, 1, 1, 1)\rangle$. Với $\lambda = 2$ ta tính

$$A - 2I = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ -1 & 1 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Do đó $E(2) = \langle(1, 1, 0, 0)\rangle$, cho nên chúng ta chỉ có một xâu Jordan ứng với giá trị riêng $\lambda = 2$. Vậy ta phải tính đến $(A - 2I)^3$.

$$(A - 2I)^2 = \begin{bmatrix} 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 4 \\ -2 & 2 & 2 & 2 \\ -2 & 2 & 2 & 2 \end{bmatrix}, \quad (A - 2I)^3 = \begin{bmatrix} -4 & 4 & 4 & 4 \\ -4 & 4 & 4 & 4 \\ -4 & 4 & 4 & 4 \\ -4 & 4 & 4 & 4 \end{bmatrix}$$

Từ đó ta nhận thấy không gian không của $(A - 2I)^3$ là

$$\langle(1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1)\rangle.$$

Ta phải chọn một vector $\mathbf{v}_3$ trong cơ sở này sao cho $(A-2I)^2\mathbf{v}_3 \neq \mathbf{0}$. Do đó ta chọn $\mathbf{v}_3 = (1, 0, 0, 1)$. Từ đó ta có

$$\mathbf{v}_2 = (A - 2I)\mathbf{v}_3 = (2, 0, 2, 0) \qquad \mathbf{v}_1 = (A - 2I)\mathbf{v}_2 = (4, 4, 0, 0).$$

Vì vậy ta có

$$P = \begin{bmatrix} 1 & 4 & 2 & 1 \\ 1 & 4 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \qquad \text{và} \qquad J = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

sao cho $P^{-1}AP = J$. □

**5.2.6 Thí dụ. (OSV09)** Tính $A^{2009}$, trong đó $A = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & -7 & 5 & 3 & 0 \\ 0 & -5 & 4 & 2 & 0 \\ 0 & -9 & 6 & 4 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$

*Giải.* Ta tính được đa thức đặc trưng của $A$ là $\chi(\lambda) = \det(A - \lambda I) = \lambda^4 - \lambda^5 + \lambda^2 - \lambda^3 = -\lambda^2(\lambda - 1)(\lambda^2 + 1)$. Do đó $A$ có các giá trị riêng thực 0 (bội 2) và 1, cho nên nó không có dạng chuẩn tắc Jordan thực. Ta sẽ tìm dạng chuẩn tắc Jordan phức. $A$ có các giá trị riêng phức $0$, $1$, $-i$, $i$. Với $\lambda = 0$ ta có

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & -7 & 5 & 3 & 0 \\ 0 & -5 & 4 & 2 & 0 \\ 0 & -9 & 6 & 4 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -2 & 0 \\ 0 & 0 & 3 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Từ đó suy ra không gian con riêng $E(0) = \langle (0, 2, 1, 3, 0) \rangle$, cho nên $A$ có một xâu Jordan có độ dài 2 ứng với giá trị riêng $\lambda = 0$. Hơn nữa, ta tính được

$$A^2 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -3 & 3 & 1 & 0 \\ 0 & -3 & 3 & 1 & 0 \\ 0 & -3 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

và không gian không của nó là $\langle (0, 1, 1, 0, 0), (0, 1, 0, 3, 0) \rangle$. Chọn vector $\mathbf{v}_2 = (0, 1, 1, 0, 0)$ có tính chất $A\mathbf{v}_2 \neq 0$. Khi đó, $\mathbf{v}_1 = A\mathbf{v}_2 = (0, -2, -1, -3, 0)$. Với $\lambda = 1$ ta có

$$A - I = \begin{bmatrix} -1 & 0 & 0 & 0 & -1 \\ 0 & -8 & 5 & 3 & 0 \\ 0 & -5 & 3 & 2 & 0 \\ 0 & -9 & 6 & 3 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Do đó, không gian con riêng $E(1) = \langle (0, 1, 1, 1, 0) \rangle$, cho nên ứng với giá trị riêng $\lambda = 1$ có một xâu Jordan có độ dài 1. Với $\lambda = -i$, ta có

$$A + iI = \begin{bmatrix} i & 0 & 0 & 0 & -1 \\ 0 & -7+i & 5 & 3 & 0 \\ 0 & -5 & 4+i & 2 & 0 \\ 0 & -9 & 6 & 5+i & 0 \\ 1 & 0 & 0 & 0 & i \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 & i \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Do đó ta có không gian con riêng $E(-i) = \langle(1,0,0,0,i)\rangle$, cho nên có một xâu Jordan ứng với giá trị riêng $\lambda = -i$. Với $\lambda = i$, ta có

$$A - iI = \begin{bmatrix} -i & 0 & 0 & 0 & -1 \\ 0 & -7-i & 5 & 3 & 0 \\ 0 & -5 & 4-i & 2 & 0 \\ 0 & -9 & 6 & 5-i & 0 \\ 1 & 0 & 0 & 0 & -i \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 & -i \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Do đó ta có không gian con riêng $E(i) = \langle(1,0,0,0,-i)\rangle$, cho nên có một xâu Jordan ứng với giá trị riêng $\lambda = i$. Như vậy ma trận $A$ có ma trận Jordan tương ứng là $J$ và ma trận $P$ sao cho $A = PJP^{-1}$ như sau

$$J = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & i \end{bmatrix} \qquad P = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ -2 & 1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & i & -i \end{bmatrix}$$

Từ đó ta có $A^{2009} = PJ^{2009}P^{-1}$. Dễ dàng tính được

$$J^{2009} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & (-i)^{2009} & 0 \\ 0 & 0 & 0 & 0 & i^{2009} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & i \end{bmatrix}$$

Từ đó ta tính được

$$A^{2009} = PJ^{2009}P^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & -3 & 3 & 1 & 0 \\ 0 & -3 & 3 & 1 & 0 \\ 0 & -3 & 3 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thí dụ này sẽ giải ngắn và đơn giản hơn nếu ta phát hiện quy luật lũy thừa của $A$; cụ thể là $A^{4n} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -3 & 3 & 1 & 0 \\ 0 & -3 & 3 & 1 & 0 \\ 0 & -3 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$. Khi đó, $A^{2009} = A^{502\cdot 4}A^3$. □

## Bài tập

**1)** Cho $f$ là một toán tử tuyến tính trên không gian vector 5 chiều. Nếu $f$ là lũy linh với $f^3 = 0$ nhưng $f^2 \neq 0$ thì $f$ có thể có ma trận biểu diễn có dạng chuẩn tắc Jordan nào?

**2)** Hãy tìm dạng chuẩn tắc Jordan $J$ của các ma trận $A$ sau và tìm ma trận $P$ sao cho $P^{-1}AP = J$.

**(a)** $\begin{bmatrix} 1 & -1 & 0 \\ -3 & 3 & 4 \\ 2 & -2 & -2 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 2 \\ 1 & -1 & 0 \end{bmatrix}$

**3)** Hãy tìm dạng chuẩn tắc Jordan $J$ của các ma trận $A$ sau và tìm ma trận $P$ sao cho $P^{-1}AP = J$.

**(a)** $\begin{bmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{bmatrix}$ **(b)** $\begin{bmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{bmatrix}$

**(c)** $\begin{bmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{bmatrix}$

**(e)** $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ **(f)** $\begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$

**4)** Cho $A$ là ma trận vuông cấp 3 mà nó chỉ có một giá trị riêng $\lambda$ bội 3. Khi đó $A$ có thể có những dạng chuẩn tắc Jordan nào?

**5)** Cho $A$ là ma trận vuông cấp 4 mà nó chỉ có một giá trị riêng $\lambda$ bội 3. Khi đó $A$ có thể có những dạng chuẩn tắc Jordan nào?

**6)** Cho $A$ là ma trận vuông cấp 5. Nếu $A^2 \neq O$ và $A^3 = O$ thì $A$ có thể có dạng chuẩn tắc Jordan nào?

**7)** Cho $f$ là một toán tử tuyến tính trên một không gian vector 6 chiều $V$. Nếu $f$ chỉ có một giá trị riêng $\lambda$ bội 6 và không gian con riêng $E(\lambda)$ có số chiều là 3 thì dạng chuẩn tắc Jordan của ma trận biểu diễn của $f$ có thể có ở những dạng nào?

**8)** Cho $f$ là một phép biến đổi tuyến tính trên một không gian vector hữu hạn chiều

**(a)** Chứng minh rằng $\mathrm{Im}(f^i) \subseteq \mathrm{Im}(f^j)$ với bất kỳ $i > j$.

**(b)** Nếu với $k$ nào đó mà $\mathrm{Im}(f^k) = \mathrm{Im}(f^{k+1})$ thì $\mathrm{Im}(f^k) = \mathrm{Im}(f^{k+j})$ với mọi $j \geq 1$.

**(c)** Chứng minh rằng tồn tại số nguyên dương nhỏ nhất $k_0$ sao cho $\mathrm{Im}(f^{k_0}) = \mathrm{Im}(f^{k_0+1})$

**(d)** Cho $k_1$ là số nguyên dương nhỏ nhất sao cho $\ker(f^{k_1}) = \ker(f^{k_1+1})$. Chứng minh rằng $k_1 = k_0$.

**9) (OSV03)** Cho $P$ và $Q$ là hai ma trận vuông cấp $n$ thỏa mãn các điều kiện sau: $PQ = QP$ và tồn tại các số nguyên dương $s, r$ sao cho $P^s = Q^r = O$. Chứng minh rằng các ma trận $I + (P + Q)$ và $I - (P + Q)$ là các ma trận khả nghịch.

# Chương IX

# Không gian Unita

Mục đích của chúng ta là trang bị tích vô hướng cho các không gian vector trên trường số phức. Nếu định nghĩa tích vô hướng như trong trường hợp $\mathbb{K}$ là trường số thực thì ta thấy có những trở ngại chẳng hạn như

- $\langle \mathbf{x}, \mathbf{x} \rangle$ không chắc thuộc $\mathbb{R}$, do đó không thỏa $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ với mọi $\mathbf{x}$.
- có thể xảy ra trường hợp $\mathbf{x} \neq \mathbf{0}$ nhưng $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ (ví dụ với tích vô hướng Euclid thông thường trong $\mathbb{C}^2$ ta có $(1, i) \neq 0$ nhưng $\langle (1, i), (1, i) \rangle = 1 + i^2 = 0$).

Vì vậy, khi $\mathbb{K} = \mathbb{C}$, chúng ta cần đưa ra một định nghĩa khác thích hợp hơn cho tích vô hướng.

## § 1 Khái niệm và tính chất cơ bản về không gian unita

### 1.1 Tích vô hướng phức

**1.1.1 Định nghĩa.** Cho $V$ là một không gian vector trên $\mathbb{C}$. Ta gọi một **tích vô hướng** trên $V$ là một ánh xạ

$$\begin{aligned} \langle\ ,\ \rangle : V \times V &\to \mathbb{C} \\ (\mathbf{x}, \mathbf{y}) &\mapsto \langle \mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

thỏa các điều kiện sau:

**(i)** $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$

**(ii)** $\langle \mathbf{x} + \mathbf{x}', \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}', \mathbf{y} \rangle$

**(iii)** $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$

**(iv)** nếu $\mathbf{x} \neq \mathbf{0}$ thì $\langle \mathbf{x}, \mathbf{x} \rangle > 0$

với mọi $\mathbf{x}, \mathbf{y} \in V$ và mọi $\alpha \in \mathbb{C}$. Khi đó không gian vector $V$ cùng với tích vô hướng $\langle\ ,\ \rangle$ được gọi là một **không gian unita**.

**1.1.2 Nhận xét.**

- Điều kiện (i) bảo đảm rằng $\langle \mathbf{x}, \mathbf{x} \rangle \in \mathbb{R}$.
- $\langle \mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{y}' \rangle = \overline{\langle \alpha\mathbf{y} + \beta\mathbf{y}', \mathbf{x} \rangle} = \overline{\langle \alpha\mathbf{y}, \mathbf{x} \rangle + \langle \beta\mathbf{y}', \mathbf{x} \rangle} =$
  $= \overline{\alpha\langle \mathbf{y}, \mathbf{x} \rangle + \beta\langle \mathbf{y}', \mathbf{x} \rangle} = \overline{\alpha}\langle \mathbf{x}, \mathbf{y} \rangle + \overline{\beta}\langle \mathbf{x}, \mathbf{y}' \rangle$.
- Ánh xạ $\langle\ ,\ \rangle : V \times V \to \mathbb{C}$ định nghĩa như trên còn được gọi là *dạng song tuyến tính đối xứng liên hợp* xác định trên $V$.

**1.1.3 Thí dụ.** Xét $V = \mathbb{C}^n$. Với mọi $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ và $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in V$, ta có

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j \overline{y}_j.$$

Đây thực sự là một tích vô hướng trên $\mathbb{C}^n$. Việc kiểm tra các điều kiện của tích vô hướng phức này xin dành cho bạn đọc xem như bài tập. Tích vô hướng phức này thường được ký hiệu $\mathbf{y}^H\mathbf{x}$. Khi đó $(V, \langle\ ,\ \rangle)$ là một không gian unita. □

## 1.2 Tính chất không gian unita

Ta có một số khái niệm và tính chất tương tự như không gian Euclid.

**1.2.1 Định lý.** *Cho $U$ là một không gian unita. Khi đó, với mọi $\mathbf{x}, \mathbf{y}, \mathbf{z} \in U$ và $\alpha \in \mathbb{C}$, ta có*

**(i)** $\langle \mathbf{x}, \mathbf{y}+\mathbf{z}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{x}, \mathbf{z}\rangle$

**(ii)** $\langle \mathbf{x}, \alpha\mathbf{y}\rangle = \overline{\alpha}\langle \mathbf{x}, \mathbf{y}\rangle$

**(iii)** $\langle \mathbf{x}, \mathbf{x}\rangle = 0$ *khi và chỉ khi* $\mathbf{x} = \mathbf{0}$

**(iv)** *nếu* $\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{z}\rangle$ *với mọi* $\mathbf{x} \in V$ *thì* $\mathbf{y} = \mathbf{z}$.

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

**1.2.2 Định nghĩa.** Cho không gian unita $U$. Ta có các định nghĩa sau:

- Ta gọi **độ dài** của vector $\mathbf{x} \in U$, ký hiệu $\|\mathbf{x}\|$ là số thực không âm $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}$.
- Hai vector $\mathbf{x}, \mathbf{y} \in U$ gọi là **trực giao** nhau nên $\langle \mathbf{x}, \mathbf{y}\rangle = 0$, ký hiệu $\mathbf{x} \perp \mathbf{y}$.
- Một hệ vector của $U$ gọi là **hệ trực giao** nếu chúng đôi một trực giao.
- **Hệ trực chuẩn** là một hệ trực giao mà độ dài mỗi vector bằng 1.

**1.2.3 Thí dụ.** Trong không gian $\mathbb{C}^2$ xét tích vô hướng phức. Với $\mathbf{x} = (5+i, 1-3i)$ và $\mathbf{y} = (2+i, -2+3i)$. Khi đó $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^H\mathbf{x} = (5+i)(2-i)+(1-3i)(-2-3i) = 11-3i+(-11+3i) = 0$, $\|\mathbf{x}\| = \sqrt{26+10} = 6$ và $\|\mathbf{y}\| = \sqrt{5+13} = 3\sqrt{2}$. □

**1.2.4 Định lý.** *Trong không gian unita, một họ vector trực giao không chứa vector* $\mathbf{0}$ *là họ độc lập tuyến tính.*

*Chứng minh.* Giả sử $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ là một hệ trực giao trong không gian unita. Xét tổ hợp tuyến tính

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}.$$

Với mỗi $i = 1, 2, \ldots, n$, ta có

$$0 = \langle \mathbf{0}, \mathbf{u}_i\rangle = \langle \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n, \mathbf{u}_i\rangle = \alpha_i\langle \mathbf{u}_i, \mathbf{u}_i\rangle,$$

suy ra $\alpha_i = 0$. Vậy hệ vector $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ độc lập tuyến tính. □

**1.2.5 Định lý.** *Nếu* $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *là một cơ sở trực chuẩn của không gian unita* $U$*, thì với mọi* $\mathbf{x}, \mathbf{y} \in U$ *và có biểu diễn* $\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j$ *và* $\mathbf{y} = \sum_{j=1}^{n} y_j \mathbf{e}_j$ *ta có* $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j \overline{y}_j$.

*Chứng minh.* Ta chỉ việc tính trực tiếp

$$\langle \mathbf{x}, \mathbf{y} \rangle = \Big\langle \sum_{j=1}^{n} x_j \mathbf{e}_j, \sum_{k=1}^{n} y_k \mathbf{e}_k \Big\rangle = \sum_{j=1}^{n} x_j \Big\langle \mathbf{e}_j, \sum_{k=1}^{n} y_k \mathbf{e}_k \Big\rangle = \sum_{j=1}^{n} x_j \overline{y}_j. \quad \square$$

**1.2.6 Định lý.** *Trong không gian unita luôn tồn tại cơ sở trực chuẩn.*

*Chứng minh.* Do trong không gian unita hữu hạn chiều luôn tồn tại một cơ sở, nên áp dụng thuật toán Gram-Schmidt thu được cơ sở trực chuẩn. □

**1.2.7 Định lý.** *Trong không gian unita ta có bất đẳng thức* $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$.

*Chứng minh.* Nếu $\mathbf{y} = \mathbf{0}$ thì bất đẳng thức đúng. Xét $\mathbf{y} \neq \mathbf{0}$. Ta có $\langle \mathbf{x} - \lambda \mathbf{y}, \mathbf{x} - \lambda \mathbf{y} \rangle \geq 0$ với mọi $\lambda \in \mathbb{C}$. Suy ra

$$\langle \mathbf{x}, \mathbf{x} \rangle - \lambda \langle \mathbf{y}, \mathbf{x} \rangle - \overline{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle + \lambda \overline{\lambda} \langle \mathbf{y}, \mathbf{y} \rangle \geq 0 \qquad \text{với mọi } \lambda \in \mathbb{C}.$$

Lấy $\lambda = \dfrac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}$, ta được

$$\langle \mathbf{x}, \mathbf{x} \rangle - \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} \langle \mathbf{y}, \mathbf{x} \rangle - \frac{\overline{\langle \mathbf{x}, \mathbf{y} \rangle}}{\langle \mathbf{y}, \mathbf{y} \rangle} \langle \mathbf{x}, \mathbf{y} \rangle + \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} \frac{\overline{\langle \mathbf{x}, \mathbf{y} \rangle}}{\langle \mathbf{y}, \mathbf{y} \rangle} \langle \mathbf{y}, \mathbf{y} \rangle \geq 0$$

suy ra

$$\langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle \overline{\langle \mathbf{x}, \mathbf{y} \rangle} - \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \geq 0$$

cho nên

$$\langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle \geq \langle \mathbf{x}, \mathbf{y} \rangle \overline{\langle \mathbf{x}, \mathbf{y} \rangle}.$$

Do đó ta được $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|$. □

## Bài tập

**1**) Kiểm tra các điều kiện của tích vô hướng đối với tích vô hướng phức $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^H \mathbf{x}$ trong $\mathbb{C}^n$ như ở Thí dụ 1.1.3.

**2**) Cho $H$ là không gian gồm tất cả các hàm giá trị phức biến thực liên tục trên $[0, 2\pi]$. Chứng minh rằng biểu thức sau xác định một tích vô hướng trên $H$

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)\overline{g(t)}dt \qquad f, g \in H.$$

Khi đó, hệ các hàm $S = (f_j)_{j \in \mathbb{N}}$ với $f_j(t) = e^{ijt}$ là trực chuẩn.

**3**) Trong mỗi cặp vector $\mathbf{z}$ và $\mathbf{w}$ trong $\mathbb{C}^n$ sau, hãy tính $\|\mathbf{z}\|$, $\|\mathbf{w}\|$, $\|\mathbf{z}+\mathbf{w}\|$, $\langle \mathbf{z}, \mathbf{w} \rangle$ và $\langle \mathbf{w}, \mathbf{z} \rangle$.

**(a)** $\mathbf{z} = (4 + 2i, 4i)$, $\quad \mathbf{w} = (-2, 2 + i)$

**(b)** $\mathbf{z} = (1 + i, 2i, 3 - i)$, $\quad \mathbf{w} = (2 - 4i, 5, 2i)$

**(c)** $\mathbf{z} = (2, 1 + i, i)$, $\quad \mathbf{w} = (2 - i, 2, 1 + 2i)$

**4**) Cho $\mathbf{z}_1 = (\frac{1+i}{2}, \frac{1-i}{2})$ và $\mathbf{z}_2 = (\frac{i}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$

**(a)** Chứng minh rằng $(\mathbf{z}_1, \mathbf{z}_2)$ là một hệ trực chuẩn trong $\mathbb{C}^2$.

**(b)** Viết vector $\mathbf{z} = (2+4i, -2i)$ như là một tổ hợp tuyến tính của $\mathbf{z}_1$ và $\mathbf{z}_2$.

**5**) Cho $(\mathbf{u}_1, \mathbf{u}_2)$ là một cơ sở trực chuẩn của $\mathbb{C}^2$ và cho $\mathbf{z} = (4 + 2i)\mathbf{u}_1 + (6 - 5i)\mathbf{u}_2$.

**(a)** Tính các giá trị $\mathbf{u}_1^H\mathbf{z}$, $\mathbf{z}^H\mathbf{u}_1$, $\mathbf{u}_2^H\mathbf{z}$, $\mathbf{z}^H\mathbf{u}_2$

**(b)** Xác định giá trị $\|\mathbf{z}\|$.

**6**) Cho $U$ là một không gian unita. Chứng minh rằng với mọi $\mathbf{x}, \mathbf{y} \in U$ ta có

**(a)** $\langle \mathbf{x}, \mathbf{y} \rangle = \dfrac{1}{4}\displaystyle\sum_{k=1}^{4} i^k \|\mathbf{x} + i^k \mathbf{y}\|^2.$

**(b)** $\displaystyle\sum_{k=1}^{4} \|\mathbf{x} + i^k \mathbf{y}\|^2 = 4(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2).$

**7)** Cho $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ là một cơ sở trực chuẩn của không gian unita $U$ và cho $\mathbf{z}$ và $\mathbf{w}$ là các phần tử của $U$. Chứng minh rằng

$$\langle \mathbf{z}, \mathbf{w} \rangle = \sum_{i=1}^{n} \langle \mathbf{z}, \mathbf{u}_i \rangle \langle \mathbf{u}_i, \mathbf{w} \rangle.$$

**8)** Trong $\mathbb{C}$, chứng minh rằng $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^H A\mathbf{x}$ là một tích vô hướng, trong đó $A = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$. Tính $\langle \mathbf{x}, \mathbf{y} \rangle$ với $\mathbf{x} = (1-i, 2+3i)$ và $\mathbf{y} = (2+i, 3-2i)$.

**9)** Cho không gian vector phức $V$ và một hàm $\|\cdot\| : V \to \mathbb{R}$ với $\mathbf{x} \mapsto \|\mathbf{x}\|$ thỏa mãn các tính chất

**(i)** $\|\mathbf{x}\| \geq \mathbf{0}$, và $\|\mathbf{x}\| = 0$ khi và chỉ khi $\mathbf{x} = \mathbf{0}$,

**(ii)** $\|\lambda \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$,

**(iii)** $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$,

**(iv)** $4(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2) = \sum_{k=1}^{4} \|\mathbf{x} + i^k \mathbf{y}\|^2$,

với mọi $\mathbf{x}, \mathbf{y} \in V$. Khi đó, công thức

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}\sum_{k=1}^{4} i^k \|\mathbf{x} + i^k \mathbf{y}\|^2 \qquad \text{với mọi } \mathbf{x}, \mathbf{y} \in V$$

xác định trên $V$ một tích vô hướng phức; hơn nữa, $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$ với mọi $\mathbf{x} \in V$.

# § 2 Ma trận trên $\mathbb{C}$

## 2.1 Các khái niệm về ma trận trên $\mathbb{C}$

**2.1.1** Xét một ma trận $A = (a_{ij})_{m\times n}$ trên $\mathbb{C}$. Chúng ta ký hiệu $\overline{A}$ là ma trận được thành lập từ $A$ bằng cách lấy liên hợp mọi phần tử của $A$, nghĩa là $\overline{A} = (\overline{a}_{ij})_{m\times n}$. Chuyển vị của ma trận $\overline{A}$ được ký hiệu $A^H$, nghĩa là $A^H = \overline{A}^t$. Như vậy, trong trường hợp $A$ là ma trận thực thì $A^H$ chính là ma trận chuyển vị $A^t$ của $A$. Cũng tương tự như tính chất của chuyển vị ma trận ta có mệnh đề sau.

**2.1.2 Mệnh đề.** *Với điều kiện các phép toán trên các ma trận sau thực hiện được, ta luôn có*

**(i)** $(A^H)^H = A$

**(ii)** $(\alpha A + \beta B)^H = \overline{\alpha}A^H + \overline{\beta}B^H$

**(iii)** $(AC)^H = C^H A^H$

*Chứng minh.* Dành cho bạn đọc xem như bài tập. □

**2.1.3 Định nghĩa.** Một ma trận $A$ được gọi là **ma trận Hermite** nếu $A = A^H$.

Trong trường hợp ma trận thực, ma trận Hermite chính là ma trận đối xứng.

**2.1.4 Thí dụ.** Ma trận $A = \begin{bmatrix} 3 & 2-i \\ 2+i & 4 \end{bmatrix}$ là ma trận Hermite do $A^H = \begin{bmatrix} \overline{3} & \overline{2-i} \\ \overline{2+i} & \overline{4} \end{bmatrix} = \begin{bmatrix} 3 & 2+i \\ 2-i & 4 \end{bmatrix} = A.$ □

**2.1.5 Định lý.** *Giá trị riêng của ma trận Hermite là thực. Hơn nữa, các vector riêng ứng với các giá trị riêng phân biệt là trực giao.*

*Chứng minh.* Cho $A$ là một ma trận Hermite. Cho $\lambda$ là một giá trị riêng của $A$ và $\mathbf{x}$ là một vector riêng ứng với giá trị riêng $\lambda$. Nếu $\alpha = \mathbf{x}^H A\mathbf{x}$, thì

$$\overline{\alpha} = \alpha^H = (\mathbf{x}^H A\mathbf{x})^H = \mathbf{x}^H A^H (\mathbf{x}^H)^H = \mathbf{x}^H A\mathbf{x} = \alpha.$$

Do đó $\alpha$ là thực. Nó kéo theo rằng

$$\alpha = \mathbf{x}^H A\mathbf{x} = \mathbf{x}^H \lambda\mathbf{x} = \lambda\|\mathbf{x}\|^2$$

và do đó

$$\lambda = \frac{\alpha}{\|\mathbf{x}\|^2}$$

là thực. Mặt khác, nếu $\mathbf{x}_1$ và $\mathbf{x}_2$ là hai vector riêng ứng với hai giá trị riêng phân biệt $\lambda_1$ và $\lambda_2$ thì

$$(A\mathbf{x}_1)^H\mathbf{x}_2 = \mathbf{x}_1^H A^H\mathbf{x}_2 = \mathbf{x}_1^H A\mathbf{x}_2 = \lambda_2\mathbf{x}_1^H\mathbf{x}_2$$

và

$$(A\mathbf{x}_1)^H\mathbf{x}_2 = (\mathbf{x}_2^H A\mathbf{x}_1)^H = (\lambda_1\mathbf{x}_2^H\mathbf{x}_1)^H = \lambda_1\mathbf{x}_1^H\mathbf{x}_2.$$

Do đó $\lambda_1\mathbf{x}_1^H\mathbf{x}_2 = \lambda_2\mathbf{x}_1^H\mathbf{x}_2$. Do $\lambda_1 \neq \lambda_2$ nó kéo theo rằng $\langle \mathbf{x}_2, \mathbf{x}_1\rangle = \mathbf{x}_1^H\mathbf{x}_2 = 0$. □

**2.1.6 Định nghĩa.** Một ma trận vuông không suy biến $A$ gọi là **ma trận unita** nếu $A^{-1} = A^H$.

Trong trường hợp ma trận thực, ma trận unita trở thành ma trận trực giao.

**2.1.7 Định lý.** *Ma trận chuyển giữa các cơ sở trực chuẩn trong không gian unita là một ma trận unita.*

*Chứng minh.* Giả sử $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ và $(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ là hai cơ sở trực chuẩn của không gian unita $U$. Nếu

$$\mathbf{b}_j = \sum_{l=1}^{n} \alpha_{lj}\mathbf{a}_l, \qquad \mathbf{a}_j = \sum_{l=1}^{n} \beta_{lj}\mathbf{b}_l \qquad j = 1, \ldots, n$$

thì

$$\beta_{kj} = \langle \mathbf{a}_j, \mathbf{b}_k \rangle = \left\langle \mathbf{a}_j, \sum_{l=1}^{n} \alpha_{lk} \mathbf{a}_l \right\rangle = \overline{\alpha}_{jk}.$$

Suy ra $(\beta_{kj}) = \overline{(\alpha_{kj})^t}$. Do đó, nếu $T$ là ma trận chuyển từ cơ sở trực chuẩn $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ sang cơ sở trực chuẩn $(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ thì ta có

$$T^{-1} = \overline{T}^t = T^H. \qquad \square$$

**2.1.8 Định nghĩa.** Ma trận $A$ được gọi là **chuẩn tắc** nếu $AA^H = A^H A$.

Bạn đọc có thể xác minh rằng $A = \begin{bmatrix} 2+2i & 3+i & 0 \\ 1-3i & 4+i & 0 \\ 0 & 0 & 3 \end{bmatrix}$ là một ma trận chuẩn tắc.

## 2.2 Chéo hóa ma trận trên $\mathbb{C}$

**2.2.1 Định lý.** *Nếu các giá trị riêng của một ma trận Hermite $A$ là phân biệt thì tồn tại một ma trận unita $U$ mà nó chéo hóa $A$.*

*Chứng minh.* Gọi $\mathbf{x}_i$ là vector riêng ứng với giá trị riêng $\lambda_i$ cho mỗi giá trị riêng $\lambda_i$ của $A$. Đặt $\mathbf{u}_i = (1/\|\mathbf{x}_i\|)\mathbf{x}_i$. Khi đó $\mathbf{u}_i$ là vector riêng đơn vị ứng với giá trị riêng $\lambda_i$ với mỗi $i$. Từ Định lý 2.1.5 kéo theo rằng $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ là hệ trực chuẩn trong $\mathbb{C}^n$. Đặt $U$ là ma trận mà cột thứ $i$ của nó là vector cột $\mathbf{u}_i$ với mỗi $i$. Khi đó $U$ là ma trận unita và $U$ chéo hóa $A$. $\square$

**2.2.2 Thí dụ.** Cho $A = \begin{bmatrix} 2 & 1-i \\ 1+i & 1 \end{bmatrix}$. $A$ là ma trận Hermite. Tìm ma trận unita $U$ chéo hóa $A$.

*Giải.* Đa thức đặt trưng của ma trận $A$ là $\chi_A(\lambda) = (2-\lambda)(1-\lambda) - (1+i)(1-i) = \lambda^2 - 3\lambda$. Vậy $A$ có hai giá trị riêng là $\lambda_1 = 3$ và $\lambda_2 = 0$. Khi $\lambda_1 = 3$ ta cần tìm không gian không của ma trận

$$\begin{bmatrix} -1 & 1-i \\ 1+i & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -1+i \\ 0 & 0 \end{bmatrix}$$

đó là $E(3) = \langle(1-i,1)\rangle$. Cho nên $(1-i,1)$ là vector riêng ứng với $\lambda_1 = 3$. Tương tự ta biến đổi ma trận

$$\begin{bmatrix} 2 & 1-i \\ 1+i & 1 \end{bmatrix} \to \begin{bmatrix} 2 & 1-i \\ 0 & 0 \end{bmatrix}$$

cho nên vector riêng ứng với giá trị riêng $\lambda_2 = 0$ là $(1-i,-2)$. Đặt

$$\mathbf{u}_1 = \frac{1}{\sqrt{2+1}}(1-i,1) \qquad \text{và} \qquad \mathbf{u}_2 = \frac{1}{\sqrt{2+4}}(1-i,-2)$$

Do đó $U = \dfrac{1}{\sqrt{3}}\begin{bmatrix} 1-i & \frac{1-i}{\sqrt{2}} \\ 1 & -\sqrt{2} \end{bmatrix}$ và

$$U^H AU = \frac{1}{3}\begin{bmatrix} 1+i & 1 \\ \frac{1+i}{\sqrt{2}} & -\sqrt{2} \end{bmatrix}\begin{bmatrix} 2 & 1-i \\ 1+i & 1 \end{bmatrix}\begin{bmatrix} 1-i & \frac{1-i}{\sqrt{2}} \\ 1 & -\sqrt{2} \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}.$$

$\square$

Thật sự Định lý trên vẫn còn đúng khi các giá trị riêng của $A$ không phân biệt. Để chứng minh điều này đầu tiên ta chứng minh định lý sau.

**2.2.3 Định lý. (Định lý Schur)** *Với mỗi ma trận $A$ cấp $n \times n$ tồn tại một ma trận unita $U$ sao cho $U^H AU$ là ma trận tam giác trên.*

*Chứng minh.* Ta chứng minh bằng qui nạp theo $n$. Kết quả rõ ràng khi $n = 1$. Giả sử định lý đúng với mọi ma trận vuông cấp $k \times k$. Giả sử $A$ là một ma trận cấp $(k+1) \times (k+1)$. Gọi $\lambda_1$ là một giá trị riêng của ma trận $A$ và $\mathbf{u}_1$ là một vector riêng đơn vị ứng với giá trị riêng $\lambda_1$. Dùng thuật toán Gram-Schmit xây dựng các vector $\mathbf{u}_2, \ldots, \mathbf{u}_n$ sao cho hệ $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$ là một cơ sở trực chuẩn cho $\mathbb{C}^{k+1}$. Đặt $W$ là ma trận mà cột thứ $i$ của nó là vector cột $\mathbf{u}_i$ với $i = 1, \ldots, k+1$. Do đó $W$ là unita. Cột thứ nhất của ma trận $W^H AW$ sẽ là $W^H A\mathbf{w}_1$. Ta tính

$$W^H A\mathbf{w}_1 = \lambda W^H \mathbf{w}_1 = \lambda_1 \mathbf{e}_1.$$

Do đó $W^H AW$ là ma trận có dạng

$$\left[\begin{array}{c|ccc} \lambda_1 & \times & \ldots & \times \\ \hline 0 & & & \\ \vdots & & M & \\ 0 & & & \end{array}\right]$$

ở đây $M$ là một ma trận cấp $k \times k$. Theo giả thiết qui nạp tồn tại một ma trận $V_1$ cấp $k \times k$ sao cho $V_1^H AV_=T_1$ ở đây $T_1$ là ma trận tam giác trên. Đặt

$$V = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & V_1 & \\ 0 & & & \end{array}\right]$$

Ta nhận thấy $V$ là ma trận unita và

$$V^H W^H AWV = \left[\begin{array}{c|ccc} \lambda_1 & \times & \dots & \times \\ \hline 0 & & & \\ \vdots & & V_1^H MV_1 & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} \lambda_1 & \times & \dots & \times \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right] = T.$$

Đặt $U = WV$, nó là ma trận unita, bởi vì

$$U^H U = (WV)^H WV = V^H W^H WV = I$$

và $U^H AU = T$. □

Cách viết thành nhân tử $A = UTU^H$ thường được gọi là *phân tích Schur* của $A$. Trong trường hợp $A$ là ma trận Hermite thì ma trận $T$ trở thành ma trận chéo, và ta có lại kết quả đã biết trước đây.

**2.2.4 Định lý.** *Nếu $A$ là ma trận Hermite thì tồn tại một ma trận unita $U$ mà nó chéo hóa $A$.*

*Chứng minh.* Theo Định lý 2.2.3, tồn tại một ma trận unita $U$ sao cho $U^H AU = T$, ở đây $T$ là ma trận tam giác trên. Mặt khác ta có

$$T^H = (U^H AU)^H = U^H A^H U = U^H AU = T.$$

Do đó, $T$ là ma trận Hermite và $T$ phải là ma trận chéo. □

**2.2.5 Nhận xét.** Trong trường hợp $A$ là ma trận thực đối xứng, giá trị riêng và vector riêng của nó phải là thực. Do đó ma trận chéo hóa $U$ của $A$ phải là trực giao. Vậy ta phát biểu kết quả trong hệ quả sau.

**2.2.6 Hệ quả.** *Nếu $A$ là một ma trận thực đối xứng thì tồn tại một ma trận trực giao $U$ mà nó chéo hóa $A$, nghĩa là $U^TAU = D$ là ma trận chéo.*

Định lý 2.2.4 kéo theo rằng mọi ma trận Hermite $A$ có thể phân tích thành tích $UDU^H$ trong đó $U$ là ma trận unita và $D$ là ma trận chéo. Do $U$ chéo hóa $A$, nên kéo theo rằng mọi phần tử của trên đường chéo của $D$ là giá trị riêng của $A$ và các cột của $U$ ứng với các vector riêng tương ứng với giá trị riêng trong cột tương ứng của $D$. Các vector riêng ứng với các cột của ma trận $U$ lập thành một cơ sở trực chuẩn cho $\mathbb{C}^n$, gọi nó là $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$. Cho $\mathbf{x} \in \mathbb{C}^n$ ta có biểu diễn $\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$, trong đó $c_i = \langle \mathbf{x}, \mathbf{u}_i\rangle = \mathbf{u}_i^H\mathbf{x}$. Khi đó

$$\begin{aligned}
A\mathbf{x} &= c_1A\mathbf{u}_1 + \cdots + c_nA\mathbf{u}_n \\
&= c_1\lambda_1\mathbf{u}_1 + \cdots c_n\lambda_n\mathbf{u}_n \\
&= \lambda_1(\mathbf{u}_1^H\mathbf{x})\mathbf{u}_1 + \ldots + \lambda_n(\mathbf{u}_n^H\mathbf{x})\mathbf{u}_n.
\end{aligned}$$

Có những ma trận không là Hermite mà vẫn có hệ các vector riêng trực chuẩn. Chẳng hạn, các ma trận phản đối xứng và ma trận phản Hermite (ma trận Hermite lệch), là ma trận $A^H = -A$, có tính chất này. Thật sự có một lớp ma trận lớn hơn có tính chất này. Kết quả sau đây sẽ làm rõ điều này nhưng ở dạng tổng quát.

**2.2.7 Định lý.** *Một ma trận $A$ là chuẩn tắc nếu và chỉ nếu $A$ có một hệ đầy đủ trực chuẩn các vector riêng.*

*Chứng minh.* Giả sử $A$ là một ma trận có một hệ đầy đủ trực chuẩn các vector riêng. Khi đó ta có thể biểu diễn $A = UDU^H$ ở đây $U$ là ma trận unita và $D$ là một ma trận chéo. Ta tính

$$AA^H = (UDU^H)(UDU^H)^H = UDU^HUD^HU^H = UDD^HU^H$$

và

$$A^HA = (UDU^H)^HUDU^H = UD^HU^HUDU^H = UD^HDU^H$$

Mặt khác, do

$$DD^H = D^H D = \begin{bmatrix} |\lambda_1|^2 & 0 & \dots & 0 \\ 0 & |\lambda_2|^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & |\lambda_n|^2 \end{bmatrix}$$

nó kéo theo rằng $AA^H = A^H A$, hay $A$ là ma trận chuẩn tắc.

Ngược lại, giả sử $A$ là một ma trận chuẩn tắc. Theo Định lý 2.2.3 tồn tại một ma trận unita $U$ và ma trận tam giác $T$ sao cho $T = U^H AU$. Chúng ta cũng có $T$ là ma trận chuẩn tắc. Thật vậy

$$T^H T = (U^H AU)^H U^H AU = U^H A^H UU^H AU = U^H A^H AU$$

và

$$TT^H = U^H AU(U^H AU)^H = U^H AUU^H A^H U = U^H AA^H U$$

Do $A^H A = AA^H$ nên kéo theo rằng $T^H T = TT^H$. So sánh các phần tử trên đường chéo của $TT^H$ và $T^H T$, chúng ta thấy rằng

$$\begin{aligned} |t_{11}|^2 + |t_{12}|^2 + \dots + |t_{1n}|^2 &= |t_{11}|^2 \\ |t_{22}|^2 + \dots + |t_{2n}|^2 &= |t_{12}|^2 + |t_{22}|^2 \\ &\vdots \\ |t_{nn}|^2 &= |t_{1n}|^2 + |t_{2n}|^2 + \dots + |t_{nn}|^2 \end{aligned}$$

Nó kéo theo rằng $t_{ij} = 0$ mỗi khi $i \neq j$. Do đó $U$ chéo hóa ma trận $A$ và các vector cột của $U$ là các vector riêng của $A$. □

## Bài tập

**1)** Chứng minh Mệnh đề 2.1.2.

**2)** Cho $A$ là một ma trận cấp $n \times n$. Định nghĩa

$$A_1 = \frac{1}{2}(A + A^H) \qquad A_2 = \frac{1}{2i}(A - A^H).$$

**(a)** Chứng minh rằng $A_1^H = A_1$, $A_2^H = A_2$ và $A = A_1 + iA_2$. Có hợp lý không khi định nghĩa $A_1$ và $A_2$ lần lượt là phần thực và phần ảo của ma trận $A$?

**(b)** Chứng minh rằng nếu $A = B_1 + iB_2$ thỏa $B_1^H = B_1$, $B_2^H = B_2$ thì $B_1 = A_1$ và $B_2 = A_2$.

**3)** Chứng minh rằng các phần tử trên đường chéo của ma trận Hermite phải là thực.

**4)** Những ma trận nào sau đây là Hermite? chuẩn tắc

**(a)** $\begin{bmatrix} 1-i & 2 \\ 2 & 3 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 2-i \\ 2+i & -1 \end{bmatrix}$

**(c)** $\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ **(d)** $\begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{bmatrix}$

**(e)** $\begin{bmatrix} 0 & i & 1 \\ i & 0 & -2+i \\ -1 & 2+i & 0 \end{bmatrix}$ **(f)** $\begin{bmatrix} 3 & 1+i & i \\ 1-i & 1 & 3 \\ -i & 3 & 1 \end{bmatrix}$

**5)** Chứng minh rằng nếu $A$ là ma trận Hermite thì $C^t A \overline{C}$ cũng là ma trận Hermite.

**6)** Chứng minh rằng nếu $A$ là ma trận Hermite thì $\det(A)$ là một số thực.

**7)** Chứng minh rằng $\langle A, B\rangle = \mathrm{tr}(B^H A)$ với mọi $A, B \in \mathcal{M}_n(\mathbb{C})$ là một tích vô hướng phức trên $\mathcal{M}_n(\mathbb{C})$. Nếu $n = 2$ và

$$A = \begin{bmatrix} 1 & 2+i \\ 3 & i \end{bmatrix} \qquad \text{và} \qquad B = \begin{bmatrix} 1+i & 0 \\ i & -i \end{bmatrix},$$

hãy tính $\|A\|$, $\|B\|$ và $\langle A, B\rangle$.

**8)** Tìm ma trận unita chéo hóa các ma trận sau

**(a)** $\begin{bmatrix} 1 & 3+i \\ 3-i & 4 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & i & 0 \\ -i & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

**9 )** Cho $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{bmatrix}$, tìm ma trận $B$ sao cho $B^H B = A$.

**10 )** Cho $U$ là ma trận unita. Chứng minh rằng

**(a)** $U$ là chuẩn tắc

**(b)** $\|U\mathbf{x}\| = \|\mathbf{x}\|$ với mọi $\mathbf{x} \in \mathbb{C}^n$

**(c)** Nếu $\lambda$ là giá trị riêng của $U$ thì $|\lambda| = 1$.

**11 )** Cho $\mathbf{u}$ là một vector đơn vị trong $\mathbb{C}^n$ và đặt $U = I - 2\mathbf{u}\mathbf{u}^H$. Chứng minh rằng ma trận $U$ vừa là unita vừa là Hermite và do đó nó là nghịch đảo của chính nó.

**12 )** Chứng minh rằng ma trận $U$ vừa là unita vừa là Hermite thì giá trị riêng của nó hoặc là 1 hoặc là $-1$.

**13 )** Cho $A$ là một ma trận cấp $2 \times 2$ với phân tích Schur là $UTU^H$ và giả sử rằng $t_{12} \neq 0$. Chứng minh rằng

**(a)** Giá trị riêng của $A$ là $\lambda_1 = t_{11}$ và $\lambda_2 = t_{22}$

**(b)** $\mathbf{u}_1$ là một vector riêng của $A$ ứng với $\lambda_1 = t_{11}$.

**(c)** $\mathbf{u}_2$ không là vector riêng của $A$ ứng với $\lambda_2 = t_{22}$.

**14 )** Chứng minh rằng $M = A + iB$ ($A$ và $B$ là các ma trận thực) là ma trận Hermite lệch khi và chỉ khi $A$ là ma trận phản đối xứng và $B$ là ma trận đối xứng.

**15 )** Chứng minh rằng nếu $A$ là ma trận Hermite lệch và $\lambda$ là một giá trị riêng của $A$ thì $\lambda$ là số thuần ảo.

**16 )** Cho $A$ là một ma trận thực cấp $2 \times 2$ với tính chất $a_{21}a_{12} > 0$ và đặt

$$r = \sqrt{a_{21}/a_{12}} \qquad \text{và} \qquad S = \begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix}$$

Tính $B = SAS^{-1}$. Bạn có kết luận gì về giá trị riêng và vector riêng của $B$? Bạn có kết luận gì về giá trị riêng và vector riêng của $A$?

**17)** Cho $p(x) = -x^3 + cx^2 + (c+3)x + 1$ ở đây $c$ là một số thực. Ký hiệu $C$ là ma trận ứng với $p(x)$,

$$C = \begin{bmatrix} c & c+3 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

và đặt

$$A = \begin{bmatrix} -1 & 2 & -c-3 \\ 1 & -1 & c+2 \\ -1 & 1 & -c-1 \end{bmatrix}$$

**(a)** Tính $A^{-1}CA$.

**(b)** Dùng kết quả trên chứng minh rằng $p(x)$ sẽ luôn có nghiệm thực với bất kỳ giá trị của $c$.

**18)** Cho $A$ là một ma trận Hermite với các giá trị riêng $\lambda_1, \ldots, \lambda_n$ và các vector trực chuẩn tương ứng $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Chứng minh rằng

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^H + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^H + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^H.$$

**19)** Đặt $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Viết $A$ dưới dạng tổng $\lambda_1 \mathbf{u}_1 \mathbf{u}_1^t + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^t$ ở đây $\lambda_1$ và $\lambda_2$ là các giá trị riêng và $\mathbf{u}_1$ và $\mathbf{u}_2$ là các vector riêng trực chuẩn tương ứng.

**20)** Cho $A$ là một ma trận Hermite với các giá trị riêng $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ và các vector riêng trực chuẩn tương ứng $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Với vector khác không $\mathbf{x}$ bất kỳ trong $\mathbb{R}^n$, *tỷ số Rayleigh* $\rho(\mathbf{x})$ được định nghĩa bởi

$$\rho(\mathbf{x}) = \frac{\langle A\mathbf{x}, \mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} = \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H \mathbf{x}}.$$

**(a)** Nếu $\mathbf{x} = c_1 \mathbf{u}_1 + \cdots + c_n \mathbf{u}_n$, chứng minh rằng

$$\rho(\mathbf{x}) = \frac{|c_1|^2 \lambda_1 + |c_2|^2 \lambda_2 + \cdots + |c_n|^2 \lambda_n}{\|\mathbf{c}\|^2}$$

**(b)** Chứng minh rằng $\lambda_n \leq \rho(\mathbf{x}) \leq \lambda_1$.

**(c)** Chứng minh rằng

$$\max_{\mathbf{x}\neq\mathbf{0}} \rho(\mathbf{x}) = \lambda_1 \qquad \text{và} \qquad \min_{\mathbf{x}\neq\mathbf{0}} \rho(\mathbf{x}) = \lambda_n.$$

# § 3 Các phép biến đổi tuyến tính trên không gian unita

## 3.1 Phép biến đổi tuyến tính và dạng song tuyến tính

**3.1.1 Định nghĩa.** Hai không gian unita $U_1$ và $U_2$ với tích vô hướng lần lượt là $\langle\ ,\ \rangle_1$ và $\langle\ ,\ \rangle_2$, gọi là **đẳng cấu nhau** nếu có một ánh xạ tuyến tính $f : U_1 \to U_2$ sao cho $f$ là đẳng cấu không gian vector và với mọi $\mathbf{x}, \mathbf{y} \in U_1$ ta có

$$\langle \mathbf{x}, \mathbf{y} \rangle_1 = \langle f(\mathbf{x}), f(\mathbf{y}) \rangle_2.$$

**3.1.2 Định lý.** *Hai không gian unita có cùng số chiều thì đẳng cấu nhau.*

**3.1.3 Định nghĩa.** Cho không gian unita $U$ có số chiều $n$ và $f$ là phép biến đổi tuyến tính của $U$. Khi đó ánh xạ

$$\begin{aligned} \varphi : U \times U &\to \mathbb{C} \\ (\mathbf{x}, \mathbf{y}) &\mapsto \langle f(\mathbf{x}), \mathbf{y} \rangle \end{aligned}$$

thỏa với mọi $\mathbf{x}, \mathbf{x}', \mathbf{y}, \mathbf{y}' \in U$ với mọi $\alpha \in \mathbb{C}$ các điều kiện sau

**(i)** $\varphi(\mathbf{x} + \mathbf{x}', \mathbf{y}) = \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{x}', \mathbf{y})$

**(ii)** $\varphi(\alpha\mathbf{x}, \mathbf{y}) = \alpha\varphi(\mathbf{x}, \mathbf{y})$

**(iii)** $\varphi(\mathbf{x}, \mathbf{y} + \mathbf{y}') = \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{x}, \mathbf{y}')$

**(iv)** $\varphi(\mathbf{x}, \alpha\mathbf{y}) = \overline{\alpha}\varphi(\mathbf{x}, \mathbf{y})$

được gọi là một **dạng song tuyến tính liên hợp** trên $U$, và nó được ký hiệu $\varphi_f$.

**3.1.4 Định lý.** *Ký hiệu $\mathscr{S}$ là tập tất cả các dạng song tuyến tính liên hợp trên $U$ và $\mathrm{Hom}(U, U)$ là tập tất cả các phép biến đổi tuyến tính trên $U$ thì ta có ánh xạ*

$$\begin{aligned} \Phi : \mathrm{Hom}(U, U) &\to \mathscr{S} \\ f &\mapsto \varphi_f \end{aligned}$$

*là song ánh ($\varphi_f$ xác định như trong định nghĩa trên).*

*Chứng minh.* Với mọi $f, g \in \mathrm{Hom}(U, U)$, nếu $\Phi(f) = \Phi(g)$ thì $\varphi_f = \varphi_g$. Khi đó với mọi $\mathbf{x}, \mathbf{y} \in U$ ta có

$$\langle f(\mathbf{x}), \mathbf{y}\rangle = \langle g(\mathbf{x}), \mathbf{y}\rangle \qquad \text{hay} \qquad \langle f(\mathbf{x}) - g(\mathbf{x}), \mathbf{y}\rangle = 0.$$

Do đó ta phải có $f(\mathbf{x}) - g(\mathbf{x}) = \mathbf{0}$ với mọi $\mathbf{x} \in U$, hay $f(\mathbf{x}) = g(\mathbf{x})$ với mọi $\mathbf{x} \in U$, cho nên $f = g$. Vậy $\Phi$ là đơn ánh.

Với mọi $\varphi \in X$, ta đặt

$$\begin{aligned} f : U &\to U \\ \mathbf{x} &\mapsto \sum_{j=1}^{n} \varphi(\mathbf{x}, \mathbf{e}_j)\mathbf{e}_j \end{aligned}$$

trong đó $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là cơ sở trực chuẩn nào đó của $U$. Khi đó với mọi $\mathbf{y} = \sum\limits_{j=1}^{n} y_j\mathbf{e}_j \in U$, ta có

$$\begin{aligned} \langle f(\mathbf{x}), \mathbf{y}\rangle &= \Big\langle \sum_{j=1}^{n} \varphi(\mathbf{x}, \mathbf{e}_j)\mathbf{e}_j, \sum_{j=1}^{n} y_j\mathbf{e}_j \Big\rangle \\ &= \sum_{j=1}^{n} \overline{y}_j \varphi(\mathbf{x}, \mathbf{e}_j) \\ &= \varphi\Big(\mathbf{x}, \sum_{j=1}^{n} y_j\mathbf{e}_j\Big) \\ &= \varphi(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Vì vậy ta có $\varphi_f(\mathbf{x}, \mathbf{y}) = \langle f(\mathbf{x}), \mathbf{y}\rangle = \varphi(\mathbf{x}, \mathbf{y})$ với mọi $\mathbf{x}, \mathbf{y} \in U$, cho nên $\Phi(f) = \varphi_f = \varphi$. Do đó $\Phi$ là toàn ánh. □

## 3.2 Phép biến đổi tuyến tính liên hợp

**3.2.1 Định lý.** *Cho $U$ là một không gian unita, $f \in \mathrm{Hom}(U,U)$. Khi đó tồn tại duy nhất phép biến đổi tuyến tính $f^* \in \mathrm{Hom}(U,U)$ sao cho*

$$\langle f(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, f^*(\mathbf{y})\rangle \qquad \text{với mọi } \mathbf{x},\mathbf{y} \in U$$

*Khi đó $f^*$ được gọi là* ***phép biến đổi tuyến tính liên hợp*** *của $f$.*

*Chứng minh.* Với mọi $\mathbf{t},\mathbf{z} \in U$, đặt $\phi(\mathbf{t},\mathbf{z}) = \overline{\varphi_f(\mathbf{z},\mathbf{t})} = \overline{\langle f(\mathbf{z}), \mathbf{t}\rangle}$. Với mọi $\mathbf{t},\mathbf{t}',\mathbf{z},\mathbf{z}' \in U$ với mọi $\alpha,\beta \in \mathbb{C}$, ta thấy

$$\begin{aligned}
\phi(\alpha\mathbf{t}+\beta\mathbf{t}',\mathbf{z}) &= \overline{\varphi_f(\mathbf{z},\alpha\mathbf{t}+\beta\mathbf{t}')}\\
&= \overline{\overline{\alpha}\varphi_f(\mathbf{z},\mathbf{t})+\overline{\beta}\varphi_f(\mathbf{z},\mathbf{t}')}\\
&= \alpha\overline{\varphi_f(\mathbf{z},\mathbf{t})}+\beta\overline{\varphi_f(\mathbf{z},\mathbf{t}')}\\
&= \alpha\phi(\mathbf{t},\mathbf{z})+\beta\phi(\mathbf{t}',\mathbf{z})
\end{aligned}$$

và

$$\begin{aligned}
\phi(\mathbf{t},\alpha\mathbf{z}+\beta\mathbf{z}') &= \overline{\varphi_f(\alpha\mathbf{z}+\beta\mathbf{z}',\mathbf{t})}\\
&= \overline{\alpha\varphi_f(\mathbf{z},\mathbf{t})+\beta\varphi_f(\mathbf{z}',\mathbf{t})}\\
&= \overline{\alpha}\cdot\overline{\varphi_f(\mathbf{z},\mathbf{t})}+\overline{\beta}\cdot\overline{\varphi_f(\mathbf{z}',\mathbf{t})}\\
&= \overline{\alpha}\phi(\mathbf{t},\mathbf{z})+\overline{\beta}\phi(\mathbf{t},\mathbf{z}').
\end{aligned}$$

Suy ra $\phi \in \mathscr{S}$. Do đó theo Định lý 3.1.4 tồn tại $f^* \in \mathrm{Hom}(U,U)$ sao cho $\phi = \varphi_{f^*}$. Khi đó với mọi $\mathbf{x},\mathbf{y} \in U$, ta có

$$\langle f(\mathbf{x}),\mathbf{y}\rangle = \varphi_f(\mathbf{x},\mathbf{y}) = \overline{\phi(\mathbf{y},\mathbf{x})} = \overline{\varphi_{f^*}(\mathbf{y},\mathbf{x})} = \overline{\langle f^*(\mathbf{y}),\mathbf{x}\rangle} = \langle \mathbf{x}, f^*(\mathbf{y})\rangle.$$

Mặt khác, nếu có $g \in \mathrm{Hom}(U,U)$ thỏa $\langle f(\mathbf{x}),\mathbf{y}\rangle = \langle \mathbf{x}, g(\mathbf{y})\rangle$ với mọi $\mathbf{x},\mathbf{y} \in U$ thì

$$\begin{aligned}
\varphi_g(\mathbf{x},\mathbf{y}) = \langle g(\mathbf{x}),\mathbf{y}\rangle = \overline{\langle \mathbf{y}, g(\mathbf{x})\rangle} &= \overline{\langle f(\mathbf{y}),\mathbf{x}\rangle}\\
&= \overline{\varphi_f(\mathbf{y},\mathbf{x})} = \phi(\mathbf{x},\mathbf{y}) = \varphi_{f^*}(\mathbf{x},\mathbf{y})
\end{aligned}$$

với mọi $\mathbf{x},\mathbf{y} \in U$. Suy ra $g = f^*$. □

**3.2.2 Định lý.** *Cho $U$ là một không gian unita. Khi đó với mọi $f, g \in \mathrm{Hom}(U, U)$ và mọi $\alpha \in \mathbb{C}$ ta có*

**(i)** $(f^*)^* = f$

**(ii)** $(f + g)^* = f^* + g^*$

**(iii)** $(f \circ g)^* = g^* \circ f^*$

**(iv)** $(\alpha f)^* = \overline{\alpha} f^*$

**(v)** $\mathrm{Id}_U^* = \mathrm{Id}_U$

*Chứng minh.* Với mọi $\mathbf{x}, \mathbf{y} \in U$, ta có

**(i)** $\langle (f^*)^*(\mathbf{x}), \mathbf{y}\rangle = \overline{\langle \mathbf{y}, (f^*)^*(\mathbf{x})\rangle} = \overline{\langle f^*(\mathbf{y}), \mathbf{x}\rangle} = \langle \mathbf{x}, f^*(\mathbf{y})\rangle = \langle f(\mathbf{x}), \mathbf{y}\rangle$. Suy ra $(f^*)^* = f$.

**(ii)** $\langle (f+g)^*(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, (f+g)(\mathbf{y})\rangle = \langle x, f(\mathbf{y})\rangle + \langle \mathbf{x}, g(\mathbf{y})\rangle = \langle f^*(\mathbf{x}), \mathbf{y}\rangle + \langle g^*(\mathbf{x}), \mathbf{y}\rangle = \langle (f^*+g^*)(\mathbf{x}), \mathbf{y}\rangle$. Suy $(f+g)^* = f^* + g^*$.

**(iii)** $\langle (f \circ g)^*(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, (f \circ g)(\mathbf{y})\rangle = \langle \mathbf{x}, f(g(\mathbf{y}))\rangle = \langle f^*(\mathbf{x}), g(\mathbf{y})\rangle = \langle g^*(f^*(\mathbf{x})), \mathbf{y}\rangle = \langle (g^* \circ f^*)(\mathbf{x}), \mathbf{y}\rangle$. Suy ra $(f \circ g)^* = g^* \circ f^*$.

**(iv)** $\langle (\alpha f)^*(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, (\alpha f)(\mathbf{y})\rangle = \langle \mathbf{x}, \alpha f(\mathbf{y})\rangle = \overline{\alpha}\langle \mathbf{x}, f(\mathbf{y})\rangle = \langle \overline{\alpha} f^*(\mathbf{x}), \mathbf{y}\rangle = \langle (\overline{\alpha} f^*)(\mathbf{x}), \mathbf{y}\rangle$. Suy ra $(\alpha f)^* = \overline{\alpha} f^*$.

**(v)** $\langle \mathrm{Id}_U^*(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, \mathrm{Id}_U(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$. Suy ra $\mathrm{Id}_U^* = \mathrm{Id}_U$. □

**3.2.3 Định lý.** *Đối với cùng một cơ sở trực chuẩn, ma trận của phép biến đổi tuyến tính liên hợp của $f$ là ma trận chuyển vị liên hợp của ma trận của $f$.*

*Chứng minh.* Giả sử $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là một cơ sở trực chuẩn nào đó của không gian unita $U$, và $f$ là một phép biến đổi tuyến tính của $U$ mà ma trận đối với cơ sở $\mathscr{B}$ là $A = (a_{ij})$. Khi đó $f(\mathbf{e}_j) = \sum\limits_{i=1}^{n} a_{ij}\mathbf{e}_i$ suy ra

$$a_{ij} = \langle f(\mathbf{e}_j), \mathbf{e}_i\rangle \qquad \text{với mọi } i, j = 1, 2, \ldots, n$$

Tương tự nếu $A^* = (b_{ij})$ là ma trận của $f^*$ đối với $\mathscr{B}$ thì với mọi $i, j = 1, 2, \ldots, n$ ta có

$$b_{ij} = \langle f^*(\mathbf{e}_j), \mathbf{e}_i \rangle = \overline{\langle \mathbf{e}_i, f^*(\mathbf{e}_j) \rangle} = \overline{\langle f(\mathbf{e}_i), \mathbf{e}_j \rangle} = \overline{a}_{ji}.$$

Vậy $A^* = \overline{A}^t$. □

**3.2.4 Nhận xét.** Xét một cơ sở trực chuẩn $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ bất kỳ và đặt $A$ và $B$ lần lượt là ma trận của phép biến đổi tuyến tính $f$ và $g$ đối với cơ sở $\mathscr{B}$. Khi đó ma trận đối với cơ sở $\mathscr{B}$ của $f + g$, $f \circ g$, $\alpha f$ và $\mathrm{Id}_U$ lần lượt là $A + B$, $AB$, $\alpha A$ và $I$. Trong khi đó, ma trận của $f^*$ đối với cơ sở $\mathscr{B}$ là $\overline{A}^t$; của $g^*$ là $\overline{B}^t$; của $(f+g)^*$ là $\overline{A+B}^t$; của $f^* + g^*$ là $\overline{A}^t + \overline{B}^t$; của $(f \circ g)^*$ là $\overline{AB}^t$; của $g^* \circ f^*$ là $\overline{B}^t\overline{A}^t$; của $(\alpha f)^*$ là $\overline{\alpha A}^c = \overline{\alpha}\overline{A}^t$.

## Bài tập

**1**) Cho $L$ là không gian con của không gian unita $U$ và $L$ bất biến qua phép biến đổi tuyến tính $f$. Chứng minh rằng phần bù trực giao $L^\perp$ bất biến qua phép biến đổi tuyến tính liên hợp $f^*$ của $f$.

**2**) Chứng minh rằng nếu $f$ là phép biến đổi tuyến tính của không gian unita $n$ chiều $U$ thì với mọi $k \in \{1, 2, \ldots, n\}$ đều tồn tại ít nhất một không gian con $k$ chiều của $U$ bất biến qua $f$.

**3**) Chứng minh rằng nếu vector **x** của không gian unita $U$ vừa là vector riêng của phép biến đổi tuyến tính $f$ ứng với giá trị riêng $\alpha$, vừa là vector riêng của phép biến đổi tuyến tính liên hợp $f^*$ ứng với giá trị riêng $\beta$ thì $\alpha = \overline{\beta}$.

**4**) Chứng minh rằng nếu phép biến đổi tuyến tính $f$ của không gian unita $n$ chiều $U$ có các giá trị riêng $\lambda_1, \lambda_2, \ldots, \lambda_n$ thì phép biến đổi tuyến tính liên hợp $f^*$ có các giá trị riêng $\overline{\lambda}_2, \overline{\lambda}_2, \ldots, \overline{\lambda}_n$.

# § 4 Phép biến đổi unita

## 4.1 Khái niệm và tính chất phép biến đổi unita

**4.1.1 Định nghĩa.** Phép biến đổi tuyến tính $f$ của không gian unita $U$ được gọi là **phép biến đổi unita** nếu $\langle \mathbf{x}, \mathbf{y} \rangle = \langle f(\mathbf{x}), f(\mathbf{y}) \rangle$ với mọi $\mathbf{x}, \mathbf{y} \in U$.

**4.1.2 Định lý.** *Cho $U$ là một không gian unita và $f$ là phép biến đổi tuyến tính của nó. Khi đó các mệnh đề sau tương đương*

**(1)** *$f$ là biến đổi unita.*

**(2)** *$f$ biến một cơ sở trực chuẩn thành một cơ sở trực chuẩn.*

**(3)** *Ma trận của $f$ đối với cơ sở trực chuẩn là ma trận unita.*

**(4)** $f^* = f^{-1}$.

**(5)** *$f$ bảo toàn độ dài của vector.*

*Chứng minh.*

(1)⇒(2) Giả sử $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là một cơ sở trực chuẩn. Khi đó với mọi $i \neq j$, ta có $\langle f(\mathbf{e}_i), f(\mathbf{e}_j) \rangle = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$ và $\|f(\mathbf{e}_i)\| = \sqrt{\langle f(\mathbf{e}_i), f(\mathbf{e}_i) \rangle} = \sqrt{\langle \mathbf{e}_i, \mathbf{e}_j \rangle} = \|\mathbf{e}_i\| = 1$. Suy ra $(f(\mathbf{e}_1), f(\mathbf{e}_2), \ldots, f(\mathbf{e}_n))$ là một hệ trực chuẩn, cho nên nó là một cơ sở trực chuẩn của $U$.

(2)⇒(3) Giả sử $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là một cơ sở trực chuẩn và $A$ là ma trận của $f$ đối với cơ sở $\mathscr{B}$. Vì $(f(\mathbf{e}_1), f(\mathbf{e}_2), \ldots, f(\mathbf{e}_n))$ là cơ sở trực chuẩn nên $A$ là ma trận chuyển từ cơ sở trực chuẩn $\mathscr{B}$ sang cơ sở trực chuẩn $(f(\mathbf{e}_1), f(\mathbf{e}_2), \ldots, f(\mathbf{e}_n))$, cho nên $A^{-1} = \overline{A}^t$, hay $A$ là ma trận unita.

(3)⇒(4) Giả sử $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là một cơ sở trực chuẩn và $A$ là ma trận của $f$ đối với cơ sở $\mathscr{B}$. Khi đó ta có $A^{-1} = \overline{A}^c$. Mặt khác $A^{-1}$ và $\overline{A}^t$ lần lượt là ma trận của $f^{-1}$ và $f^*$ đối với cơ sở $\mathscr{B}$; do đó $f^{-1} = f^*$.

(4)⇒(5) Với mọi $\mathbf{x} \in U$, ta có

$$\begin{aligned}\|f(\mathbf{x})\| = \sqrt{\langle f(\mathbf{x}), f(\mathbf{x})\rangle} &= \sqrt{\langle \mathbf{x}, f^*(f(\mathbf{x}))\rangle} = \\ &= \sqrt{\langle \mathbf{x}, f^{-1}(f(\mathbf{x}))\rangle} = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle} = \|\mathbf{x}\|.\end{aligned}$$

Vậy $f$ bảo toàn độ dài.

(5)⇒(1) Với mọi $\mathbf{x}, \mathbf{y} \in U$, ta có

$$\begin{aligned}&\langle f(\mathbf{x}+\mathbf{y}), f(\mathbf{x}+\mathbf{y})\rangle \\ &\quad = \langle f(\mathbf{x}) + f(\mathbf{y})\rangle + \langle f(\mathbf{x}) + f(\mathbf{y})\rangle \\ &\quad = \langle f(\mathbf{x}), \mathbf{f}(\mathbf{x})\rangle + \langle f(\mathbf{x}), f(\mathbf{y})\rangle + \langle f(\mathbf{y}), f(\mathbf{x})\rangle + \langle f(\mathbf{y}), f(\mathbf{y})\rangle \\ &\quad = \|f(\mathbf{x})\|^2 + \langle f(\mathbf{x}), f(\mathbf{y})\rangle + \overline{\langle f(\mathbf{x}), f(\mathbf{y})\rangle} + \|f(\mathbf{y})\|^2\end{aligned}$$

$$\begin{aligned}\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle &= \langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle \\ &= \|\mathbf{x}\|^2 + \langle \mathbf{x}, \mathbf{y}\rangle + \overline{\langle \mathbf{x}, \mathbf{y}\rangle} + \|\mathbf{y}\|^2.\end{aligned}$$

Do $f$ bảo toàn độ dài vector cho nên

(4.1.3) $$\langle f(\mathbf{x}), f(\mathbf{y})\rangle + \overline{\langle f(\mathbf{x}), f(\mathbf{y})\rangle} = \langle \mathbf{x}, \mathbf{y}\rangle + \overline{\langle \mathbf{x}, \mathbf{y}\rangle}.$$

Từ đó ta cũng có $\langle f(i\mathbf{x}), f(\mathbf{y})\rangle + \overline{\langle f(i\mathbf{x}), f(\mathbf{y})\rangle} = \langle i\mathbf{x}, \mathbf{y}\rangle + \overline{\langle i\mathbf{x}, \mathbf{y}\rangle}$ suy ra $i\langle f(\mathbf{x}), f(\mathbf{y})\rangle + \overline{i}\overline{\langle f(\mathbf{x}), f(\mathbf{y})\rangle} = i\langle \mathbf{x}, \mathbf{y}\rangle + \overline{i}\overline{\langle \mathbf{x}, \mathbf{y}\rangle}$. Do đó ta có

(4.1.4) $$\langle f(\mathbf{x}), f(\mathbf{y})\rangle - \overline{\langle f(\mathbf{x}), f(\mathbf{y})\rangle} = \langle \mathbf{x}, \mathbf{y}\rangle - \overline{\langle \mathbf{x}, \mathbf{y}\rangle}.$$

Từ (4.1.3) và (4.1.4) ta được $\langle f(\mathbf{x}), f(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ với mọi $\mathbf{x}, \mathbf{y} \in U$, hay $f$ là phép biến đổi unita. □

## 4.2 Tính chất giá trị riêng vector riêng của phép biến đổi unita

**4.2.1 Định lý.** *Mọi giá trị riêng của phép biến đổi unita đều có modulus bằng* 1.

*Chứng minh.* Giả sử $\lambda$ là giá trị riêng của phép biến đổi unita $f$ và $\mathbf{x}$ là vector riêng tương ứng. Khi đó, ta có

$$\langle \mathbf{x}, \mathbf{x}\rangle = \langle f(\mathbf{x}), f(\mathbf{x})\rangle = \langle \lambda\mathbf{x}, \lambda\mathbf{x}\rangle = \lambda\overline{\lambda}\langle \mathbf{x}, \mathbf{x}\rangle.$$

Do đó $\lambda\overline{\lambda} = 1$ hay $|\lambda| = 1$. □

**4.2.2 Định lý.** *Cho $f$ là phép biến đổi unita của không gian unita $U$. Trong $U$ luôn tồn tại cơ sở trực chuẩn gồm toàn vector riêng của $f$.*

*Chứng minh.* Ta chứng minh qui nạp theo số chiều $n$ của $U$. Khi $n = 1$, gọi $\mathbf{e}_1$ là cơ sở trực chuẩn của $U$. Vì $f(\mathbf{e}_1) \in U$ nên $f(\mathbf{e}_1) = \lambda \mathbf{e}_1$. Suy ra $\mathbf{e}_1$ là vector riêng của $f$. Vậy định lý đúng với $n = 1$.

Giả sử định lý đúng với mọi không gian unita $n-1$ chiều. Xét không gian unita $U$ có số chiều $n$. Gọi $\lambda_1$ là một giá trị riêng của $f$ và vector riêng tương ứng là $\mathbf{e} \neq \mathbf{0}$. Khi đó $\mathbf{e}_1 = \frac{1}{\|\mathbf{e}\|}\mathbf{e}$ cũng là vector riêng của $f$ ứng với giá trị $\lambda_1$ nhưng có độ dài bằng 1, nghĩa là $f(\mathbf{e}_1) = \lambda_1 \mathbf{e}_1$. Gọi $E$ là không gian con của $U$ sinh bởi $\mathbf{e}_1$, và đặt $F = E^{\perp}$. Khi đó với mọi $\mathbf{x} \in F$, ta có

$$\langle \lambda_1 \mathbf{e}_1, f(\mathbf{x}) \rangle = \langle f(\mathbf{e}_1), f(\mathbf{x}) \rangle = \langle \mathbf{e}_1, \mathbf{x} \rangle = 0,$$

suy ra $f(\mathbf{x}) \in F$. Vậy $F$ là bất biến qua $f$. Do đó $f|_F$ là một phép biến đổi unita trên $F$. Theo giả thiết qui nạp, trong $F$ có một cơ sở trực chuẩn $(\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n)$ gồm toàn vector riêng của $f$, và khi đó $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ là cơ sở cần tìm. □

**4.2.3 Hệ quả.** *$A$ là một ma trận unita khi và chỉ khi tồn tại một ma trận $T$ sao cho ma trận $T^{-1}AT$ là ma trận chéo mà phần tử trên đường chéo là các số phức đơn vị.*

*Chứng minh.* Nếu $A$ là ma trận unita thì $A$ là ma trận của một phép biến đổi unita $f$ đối với cơ sở trực chuẩn $\mathscr{A}$ nào đó trong không gian unita $U$. Theo định lý trên trong $U$ có cơ sở trực chuẩn $\mathscr{B}$ gồm toàn vector riêng của $f$, do đó ma trận của $f$ đối với cơ sở $\mathscr{B}$ là ma trận chéo. Mặt khác, nếu $T$ là ma trận chuyển từ cơ sở $\mathscr{A}$ sang cơ sở $\mathscr{B}$ thì $T$ là ma trận unita và ta có ma trận $f$ đối với cơ sở $\mathscr{B}$ là $T^{-1}AT$.

Ngược lại, vì ma trận đảo của ma trận unita là ma trận unita và tích của hai ma trận unita là ma trận unita nên ta dễ dàng suy ra điều ngược lại do $A = T(T^{-1}AT)T^{-1}$. □

### Bài tập

**1**) Cho $f$ là phép biến đổi unita của không gian unita $n$ chiều $U$. Chứng minh rằng nếu $L$ là không gian con bất biến của $f$ thì phần bù trực giao $L^\perp$ cũng bất biến qua $f$.

**2**) Chứng minh rằng nếu hai phép biến đổi unita của một không gian unita giao hoán được với nhau thì trong $U$ tồn tại một cơ sở trực chuẩn gồm toàn vector riêng chung cho cả hai phép biến đổi unita trên.

**3**) Chứng minh rằng hai vector riêng ứng với hai giá trị riêng khác nhau của một phép biến đổi unita là trực giao.

**4**) Cho ma trận $A = \dfrac{1}{9}\begin{bmatrix} 4+3i & 4i & -6-2i \\ -4i & 4-3i & -2-6i \\ 6+2i & -2-6i & 1 \end{bmatrix}$. Hãy tìm ma trận chéo $B$ và ma trận unita $Q$ sao cho $B = Q^{-1}AQ$.

# § 5 Phép biến đổi đối xứng và đối xứng lệch

## 5.1 Phép biến đổi đối xứng

**5.1.1 Định nghĩa.** Phép biến đổi tuyến tính $f$ của không gian unita $U$ gọi là **phép biến đổi tuyến tính đối xứng** nếu $f = f^*$. Điều này tương đương với

$$\langle f(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, f(\mathbf{y})\rangle \qquad \text{với mọi } \mathbf{x}, \mathbf{y} \in U.$$

**5.1.2 Định lý.** *Phép biến đổi tuyến tính $f$ của không gian unita $U$ là phép biến đổi tuyến tính đối xứng khi và chỉ khi ma trận của nó đối với cơ sở trực chuẩn của $U$ là* ***ma trận Hermite****, nghĩa là ma trận $A$ có tính chất $A = \overline{A}^t = A^H$.*

*Chứng minh.* Ta biết rằng nếu ma trận của $f$ đối với cơ sở trực chuẩn nào đó là $A$ thì đối với cơ sở này ma trận của $f^*$ là $\overline{A}^t$ (Định lý 3.2.3). Do đó, $f$ là phép biến đổi tuyến tính đối xứng khi và chỉ khi $A = \overline{A}^t$. □

**5.1.3 Định lý.** *Mọi giá trị riêng của phép biến đổi tuyến tính đối xứng là số thực.*

*Chứng minh.* Giả sử $\lambda$ là một giá trị riêng của phép biến đổi tuyến tính đối xứng $f$. Gọi vector riêng tương ứng là $\mathbf{e}$. Ta có

$$\lambda\langle\mathbf{e},\mathbf{e}\rangle = \langle\lambda\mathbf{e},\mathbf{e}\rangle = \langle f(\mathbf{e}),\mathbf{e}\rangle = \langle\mathbf{e}, f(\mathbf{e})\rangle = \langle\mathbf{e},\lambda\mathbf{e}\rangle = \overline{\lambda}\langle\mathbf{e},\mathbf{e}\rangle.$$

Suy ra $\lambda = \overline{\lambda}$ hay $\lambda$ là một số thực. □

**5.1.4 Định lý.** *Trong không gian unita luôn tồn tại cơ sở trực chuẩn gồm toàn vector riêng của một phép biến đổi tuyến tính đối xứng cho trước trên không gian đó.*

*Chứng minh.* Hoàn toàn tương tự như chứng minh Định lý 4.2.2, xin dành cho bạn đọc xem như bài tập. □

**5.1.5 Định lý.** *Phép biến đổi tuyến tính $f$ của một không gian unita là phép biến đổi tuyến tính đối xứng khi và chỉ khi ma trận của $f$ đối với một cơ sở trực chuẩn thích hợp là ma trận chéo với các phần tử thực.*

*Chứng minh.* Giả sử $f$ là phép biến đổi tuyến tính đối xứng trong không gian unita. Khi đó theo Định lý 5.1.4 trong không gian unita này tồn tại cơ sở trực chuẩn $\mathscr{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ sao cho mọi vector trong cơ sở đều là vector riêng. Theo Định lý 5.1.3 mọi giá trị riêng của $f$ đều là thực, nghĩa là $f(\mathbf{a}_i) = \lambda_i\mathbf{a}_i$ với $\lambda_i \in \mathbb{R}$ và $i = 1, 2, \ldots, n$. Do đó ma trận của $f$ đối với cơ sở $\mathscr{A}$ là ma trận chéo với các phần tử nằm trên đường chéo là $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Ngược lại, giả sử ma trận của $f$ đối với cơ sở trực chuẩn $\mathscr{B}$ là ma trận chéo $D$ với các phần tử thực. Khi đó rõ ràng $D^H = D$. Do đó theo Định lý 5.1.2 ta có $f$ là phép biến đổi tuyến tính đối xứng. □

**5.1.6 Hệ quả.** *Ma trận vuông $A$ trên $\mathbb{C}$ là ma trận Hermite đối xứng khi và chỉ khi tồn tại ma trận unita $T$ sao cho $T^{-1}AT$ là ma trận chéo thực.*

*Chứng minh.* Tương tự như chứng minh Hệ quả 4.2.3, xin dành cho bạn đọc xem như bài tập. □

## 5.2 Phép biến đổi đối xứng lệch

**5.2.1 Định nghĩa.** Phép biến đổi tuyến tính $f$ trong không gian unita $U$ gọi là **phép biến đổi tuyến tính đối xứng lệch** nếu $f = -f^*$. Điều này tương đương với

$$\langle f(\mathbf{x}), \mathbf{y}\rangle = -\langle \mathbf{x}, f(\mathbf{y})\rangle \qquad \text{với mọi } \mathbf{x}, \mathbf{y} \in U.$$

Ma trận vuông $A$ trên $\mathbb{C}$ thỏa $\overline{A}^t = -A$ được gọi là **ma trận Hermite đối xứng lệch**.

**5.2.2 Định lý.** *Phép biến đổi tuyến tính $f$ của không gian unita $U$ là phép biến đổi tuyến tính đối xứng lệch khi và chỉ khi ma trận của nó đối với cơ sở trực chuẩn của $U$ là ma trận Hermite đối xứng lệch.*

*Chứng minh.* Ta đã biết nếu ma trận của $f$ đối với một cơ sở trực chuẩn nào đó là $A$ thì đối với cơ sở này ma trận của $f^*$ là $\overline{A}^t$ (Định lý 3.2.3). Do đó ta có kết quả: *$f$ là phép biến đổi tuyến tính đối xứng lệch khi và chỉ khi $A = -\overline{A}^c$.* □

**5.2.3 Định lý.** *Phép biến đổi tuyến tính $f$ trong không gian unita $U$ là đối xứng khi và chỉ khi $if$ là phép biến đổi tuyến tính đối xứng lệch.*

*Chứng minh.* Ta đã biết nếu ma trận của $f$ đối với một cơ sở trực chuẩn nào đó là $A$ thì đối với cơ sở này ma trận của $f^*$ là $\overline{A}^t$ và của $if$ là $iA$. Mặt khác, ta cũng đã biết $f$ là phép biến đổi tuyến tính đối xứng khi và chỉ khi $A = \overline{A}^t$. Ta biến đổi tương đương đẳng thức ma trận này được

$$A = \overline{A}^t \iff -iA = \overline{iA}^t \iff -(iA) = \overline{iA}^t$$

Đẳng thức cuối tương đương với $if$ là phép biến đổi tuyến tính đối xứng lệch. □

**5.2.4 Định lý.** *Mọi giá trị riêng của một phép biến đổi tuyến tính đối xứng lệch đều có phần thực bằng* 0.

*Chứng minh.* Giả sử $\lambda$ là giá trị riêng của phép biến đổi tuyến tính đối xứng lệch $f$ và **e** là vector riêng tương ứng. Ta có

$$\lambda\langle \mathbf{e}, \mathbf{e}\rangle = \langle \lambda\mathbf{e}, \mathbf{e}\rangle = \langle f(\mathbf{e}), \mathbf{e}\rangle = -\langle \mathbf{e}, f(\mathbf{e})\rangle = -\langle \mathbf{e}, \lambda\mathbf{e}\rangle = -\overline{\lambda}\langle \mathbf{e}, \mathbf{e}\rangle.$$

Suy ra $\lambda = -\overline{\lambda}$ hay $\lambda + \overline{\lambda} = 0$. Do đó $\mathrm{Re}(\lambda) = 0$. □

**5.2.5 Định lý.** *Trong không gian unita luôn tồn tại cơ sở trực chuẩn gồm toàn vector riêng của một phép biến đổi tuyến tính đối xứng lệch cho trước trên không gian đó.*

**5.2.6 Định lý.** *Phép biến đổi tuyến tính $f$ của một không gian unita là phép biến đổi tuyến tính đối xứng lệch khi và chỉ khi ma trận của $f$ đối với một cơ sở trực chuẩn thích hợp là ma trận chéo với các phần tử trên đường chéo chính là các số phức có phần thực bằng* $0$.

**5.2.7 Hệ quả.** *Ma trận vuông $A$ là ma trận Hermite đối xứng lệch khi và chỉ khi tồn tại ma trận unita $T$ sao cho $T^{-1}AT$ là một ma trận chéo với các phần tử trên đường chéo chính là các số phức có phần thực bằng* $0$.

## 5.3 Phép biến đổi tuyến tính chuẩn tắc

Ta nhận thấy rằng một phép biến đổi tuyến tính $f$ của một không gian unita $n$ chiều $U$ dù là phép biến đổi unita, phép biến đổi tuyến tính đối xứng hay là phép biến đổi tuyến tính đối xứng lệch đều có thể tìm được một cơ sở trực chuẩn thích hợp của $U$ sao cho ma trận $A$ của $f$ có dạng chéo

$$A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \qquad \text{với } \lambda_i \in \mathbb{C} \text{ với mọi } i = 1, 2, \dots, n$$

Khi đó đối với cùng cơ sở trực chuẩn này, ma trận của phép biến đổi tuyến tính liên hợp $f^*$ là $\overline{A}^t$, cũng có dạng chéo.

$$\overline{A}^t = \begin{bmatrix} \overline{\lambda}_1 & 0 & \dots & 0 \\ 0 & \overline{\lambda}_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \overline{\lambda}_n \end{bmatrix}$$

Dễ dàng thấy rằng $A\overline{A}^t = \overline{A}^t A$, do đó $f \circ f^* = f^* \circ f$. Đặc trưng này dẫn đến định nghĩa sau.

**5.3.1 Định nghĩa.** Phép biến đổi tuyến tính $f$ của không gian unita $U$ được gọi là **phép biến đổi tuyến tính chuẩn tắc** nếu $f \circ f^* = f^* \circ f$.

Như vậy tất cả các phép biến đổi unita, phép biến đổi tuyến tính đối xứng, phép biến đổi tuyến tính đối xứng lệch đều là phép biến đổi tuyến tính chuẩn tắc.

**5.3.2 Bổ đề.** *Giả sử $f$ và $g$ là hai phép biến đổi tuyến tính của không gian unita $U$ và $\lambda$ là giá trị riêng của $f$. Nếu $f \circ g = g \circ f$ thì không gian con riêng $E(\lambda) = \{\mathbf{x} \in U : f(\mathbf{x}) = \lambda\mathbf{x}\}$ là không gian con bất biến của $f$.*

*Chứng minh.* Ta đã biết $E(\lambda)$ là không gian con riêng bất biến của $f$. Với mọi $\mathbf{x} \in E(\lambda)$ ta có

$$f(g(\mathbf{x})) = f \circ g(\mathbf{x}) = g \circ f(\mathbf{x}) = g(f(\mathbf{x})) = g(\lambda\mathbf{x}) = \lambda g(\mathbf{x}).$$

Suy ra $g(\mathbf{x}) \in E(\lambda)$, nghĩa là $E(\lambda)$ là không gian con bất biến của $g$. $\square$

**5.3.3 Bổ đề.** *Giả sử $f$ và $g$ là hai phép biến đổi tuyến tính của không gian unita $U$ thỏa $f \circ g = g \circ f$. Khi đó trong $U$ tồn tại ít nhất một vector riêng chung cho cả $f$ và $g$.*

*Chứng minh.* Giả sử $\lambda$ là một giá trị riêng nào đó của $f$. Theo Bổ đề 5.3.2, $E(\lambda)$ là không gian con bất biến của $g$, do đó $g|_{E(\lambda)}$ là phép biến đổi tuyến tính trên $E(\lambda)$. Do đó trong không gian này tồn tại vector riêng $\mathbf{x}_0$ của $g|_{E(\lambda)}$ (cũng là của $g$) ứng với giá trị riêng $\mu$ nào đó. Vì $\mathbf{x}_0 \neq \mathbf{0}$ nên đó cũng là vector riêng của $f$. $\square$

**5.3.4 Định lý.** *Phép biến đổi tuyến tính $f$ của không gian unita $U$ là phép biến đổi tuyến tính chuẩn tắc khi và chỉ khi tồn tại một cơ sở trực chuẩn của $U$ mà đối với cơ sở trực chuẩn này ma trận của $f$ và của $f^*$ đồng thời là ma trận chéo.*

*Chứng minh.* Do tích hai ma trận chéo là giao hoán nên nếu phép biến đổi tuyến tính $f$ và $f^*$ đối với một cơ sở trực chuẩn nào đó là các ma trận chéo thì $f \circ f^* = f^* \circ f$. Khi đó $f$ là phép biến đổi tuyến tính chuẩn tắc.

Ngược lại, ta sẽ chứng minh bằng qui nạp theo số chiều $n$ của $U$. Khi $n = 1$, vì $f$ giao hoán được với $f^*$ nên theo Bổ đề 5.3.3 trong $U$ tồn tại vector riêng $\mathbf{e}_1$ của các hai phép biến đổi tuyến tính $f$ và $f^*$. Chuẩn hóa $\mathbf{e}_1$ ta được cơ sở trực chuẩn của $U$ gồm toàn vector riêng của $f$ và $f^*$. Giả sử điều cần chứng minh đúng với mọi không gian unita $n-1$ chiều. Giả sử $U$ là một không gian unit $n$ chiều. Khi đó trong $U$ tồn tại vector riêng đơn vị $\mathbf{e}_1$ chung cho cả $f$ và $f^*$. Đặt $U_1 = \langle \mathbf{e}_1 \rangle^\perp$, khi đó $\dim(U_1) = n - 1$. Ta thấy với mọi $\mathbf{x} \in U_1$

$$\langle f^*(\mathbf{x}), \mathbf{e}_1 \rangle = \langle \mathbf{x}, f(\mathbf{e}_1) \rangle = \langle \mathbf{x}, \lambda \mathbf{e}_1 \rangle = 0$$

suy ra $f^*(\mathbf{x}) \in U_1$. Tương tự với mọi $\mathbf{x} \in U_1$ ta có

$$\langle f(\mathbf{x}), \mathbf{e}_1 \rangle = \langle \mathbf{x}, f^*(\mathbf{e}_1) \rangle = \langle \mathbf{x}, \mu \mathbf{e}_1 \rangle = 0$$

suy ra $f(\mathbf{x}) \in U_1$. Vậy $U_1$ bất biến qua cả $f$ và $f^*$. Do đó $f$ và $f^*$ *là* hai phép biến đổi tuyến tính giao hoán trên $U_1$. Theo giả thiết qui nạp trong $U_1$ có cơ sở trực chuẩn $(\mathbf{e}_2, \dots, \mathbf{e}_n)$ gồm toàn vector riêng của $f$ và $f^*$. Khi đó $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ là cơ sở trực chuẩn cần tìm. □

## Bài tập

**1)** Cho phép biến đổi tuyến tính $f$ của không gian unita $\mathbb{C}$ mà ma trận đối với cơ sở tự nhiên là $A = \begin{bmatrix} 3 & -i & 0 \\ i & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. Hãy tìm một cơ sở trực chuẩn của $\mathbb{C}^3$ gồm toàn vector riêng của $f$.

**2)** Cho $f$ là phép biến đổi tuyến tính đối xứng lệch của không gian unita $U$. Chứng minh rằng nếu $L$ là không gian con bất biến đối với $f$ thì phần bù trực giao $L^\perp$ cũng bất biến đối với $f$.

**3)** Chứng minh Định lý 5.1.4.

**4**) Chứng minh Hệ quả 5.1.6.

**5**) Chứng minh rằng nếu **x** là vector riêng của phép biến đổi tuyến tính chuẩn tắc $f$ ứng với giá trị riêng $\lambda$ thì **x** cũng là vector riêng của phép biến đổi tuyến tính liên hợp $f^*$ ứng với giá trị riêng $\overline{\lambda}$.

**6**) Chứng minh rằng hai vector riêng của một phép biến đổi tuyến tính chuẩn tắc ứng với hai giá trị riêng khác nhau là trực giao.

**7**) Giả sử **e** là một vector riêng của phép biến đổi tuyến tính chuẩn tắc $f$. Chứng minh rằng không gian con $L$ gồm tất cả các vector trực giao với **e** là bất biến qua $f$.

**8**) Chứng minh rằng điều kiện cần và đủ để phép biến đổi tuyến tính $f$ là chuẩn tắc là mọi vector riêng của $f$ cũng là vector riêng của $f^*$.

# Tài liệu tham khảo


[1] Nguyễn Viết Đông, Lê Anh Vũ. *Toán Cao Cấp. Tập 2: Đại số tuyến tính. Tái bản lần thứ tư*. NXB Giáo Dục, 2005.

[2] Terry Lawson. *Linear Algebra*. John Willey, New York, 1996.

[3] Nguyễn Văn Mậu, Lê Ngọc Lăng, Phạm Thế Long, Nguyễn Minh Tuấn. *Các Đề Thi Olympic Toán Sinh Viên Toàn Quốc*. NXB Giáo Dục, 2006.

[4] Hans Schneider and George Phillip Barker. *Matrices and Linear Algebra. Second edition*. Dover Publications, New York, 1989.

[5] Gilbert Strang. *Linear Algebra and its Applications. Third edition*. Harcourt Brace Jovanovich, Fort Worth, 1988.

[6] Nguyễn Duy Thuận, Phí Mạnh Ban, Nông Quốc Chinh. *Đại Số Tuyến Tính*. NXB Đại học Sư phạm, Hà Nội, 2003.

[7] Bùi Tường Trí. *Đại Số Tuyến Tính, dành cho sinh viên khoa Toán giai đoạn 2*. Trường ĐHSP TP. Hồ Chí Minh, 2000.

# NỘI DUNG ĐỀ THI

**Câu 1.** **(a)** (1 điểm) Cho $A$ và $B$ là các ma trận đối xứng cấp $n$. Chứng minh rằng $AB$ đối xứng khi và chỉ khi $AB = BA$.

**(b)** (0.75 điểm) Tìm ma trận nghịch đảo của $\begin{bmatrix} 1 & 4 & 3 \\ -1 & -2 & 0 \\ 2 & 2 & 3 \end{bmatrix}$, nếu có.

**(c)** (1 điểm) Dùng phương pháp Gauss-Jordan giải hệ phương trình $\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 2 \\ x_1 + x_2 + x_3 + 2x_4 + 2x_5 = 3 \\ x_1 + x_2 + x_3 + 2x_4 + 3x_5 = 2 \end{cases}$

**Câu 2.** **(a)** (0.5 điểm) Cho $U$ và $V$ là hai không gian con hữu hạn chiều của một không gian vector nào đó. Chứng minh rằng $\dim(U + V) = \dim U + \dim V$ khi và chỉ khi $U \cap V = \{\mathbf{0}\}$.

**(b)** (1 điểm) Tìm một cơ sở và số chiều của không gian con sinh bởi hệ vector

$$((1, 0, -1, 1), (1, 1, -1, 0), (0, 1, 0, -1), (1, -1, 1, 0)).$$

**Câu 3.** **(a)** (0.75 điểm) Cho $\mathbf{a} = (-1, 2, 3)$, $\mathbf{b} = (3, 4, 2)$, $\mathbf{x} = (2, 6, 6)$ trong $\mathbb{R}^3$. Vector $\mathbf{x}$ có thuộc không gian con sinh bởi hai vector $\mathbf{a}, \mathbf{b}$ không? Vì sao?

**(b)** (1 điểm) Chứng minh hệ $\mathscr{A} = (\mathbf{a}_1 = (1, 1, 1), \mathbf{a}_2 = (1, 2, 2), \mathbf{a}_3 = (2, 3, 4))$ là một cơ sở của $\mathbb{R}^3$. Hãy tìm tọa độ của vector $\mathbf{x} = (3, 2, 5)$ đối với cơ sở $\mathscr{A}$.

**(c)** (0.75 điểm) Tìm ma trận của toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ xác định bởi

$$f(x_1, x_2, x_3) = (2x_3, x_2 + 3x_1, 2x_1 - x_3),$$

trong cơ sở chính tắc và tìm hạng của $f$.

**(d)** (1 điểm) Cho ánh xạ tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi $f(x_1, x_2, x_3) = (x_2 - 2x_1, 3x_3 - 4x_2)$. Tìm một cơ sở và số chiều của $\ker f$, và $f$ có phải là đơn cấu hay không? Vì sao?

**Câu 4.** **(a)** (1.25 điểm) Ma trận $A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 3 \\ 1 & 1 & -1 \end{bmatrix}$ có chéo hóa được không (trên $\mathbb{R}$)? Vì sao? Nếu nó chéo hóa được, hãy tìm ma trận làm chéo hóa $A$ và ma trận chéo tương ứng.

**(b)** (1 điểm) Cho một ánh xạ tuyến tính $f : \mathbb{R}^2 \to \mathbb{R}^3$ xác định bởi

$$f(x_1, x_2) = (x_1, 2x_1 - x_2, 3x_1 + 2x_2).$$

Cho $\mathscr{A} = ((1,0),(2,1))$ là cơ sở trong $\mathbb{R}^2$, và $\mathscr{B} = ((1,0,0), (2,1,0),(3,2,1))$ là cơ sở trong $\mathbb{R}^3$. Tìm ma trận của $f$ trong cặp cơ sở $(\mathscr{A}, \mathscr{B})$.

# NỘI DUNG ĐỀ THI

**Câu 1.** **(a)** (1 điểm) Tìm hạng của dạng song tuyến tính $\xi$ trên $\mathbb{R}^3$ xác định bởi: với $\mathbf{x} = (x_1, x_2, x_3)$ và $\mathbf{y} = (y_1, y_2, y_3)$ ta có $\xi(\mathbf{x}, \mathbf{y}) = 2x_1y_1 + 2x_1y_2 - x_1y_3 + x_2y_1 - x_2y_2 + 2x_3y_1 + x_3y_3$. Tìm ma trận của $\xi$ đối với cơ sở $\mathscr{A} = ((-1,1,-1),(1,1,1), (0,-1,1))$.

**(b)** (1 điểm) Đưa dạng toàn phương $p(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + 4x_1x_3 + 5x_2^2 - 4x_2x_3 - x_3^2$ trên $\mathbb{R}^3$ về dạng chuẩn tắc, và tìm cơ sở $p$-chuẩn tắc tương ứng.

**(c)** (1 điểm) Phân loại dạng toàn phương $p(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + 2x_1x_3 + (a-1)x_2^2 + 2x_2x_3 + a^2x_3^2$ trên $\mathbb{R}^3$ theo tham số $a$.

**Câu 2.** **(a)** (1 điểm) Xác định $a$ để ánh xạ $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ xác định bởi: với $\mathbf{x} = (x_1, x_2, x_3)$ và $\mathbf{y} = (y_1, y_2, y_3)$ ta có $f(\mathbf{x}, \mathbf{y}) = 3x_1y_1 - x_1y_2 + 2x_1y_3 - x_2y_1 + x_2y_2 + ax_2y_3 + 2x_3y_1 + ax_3y_2 + (1-a)x_3y_3$, là một tích vô hướng trên $\mathbb{R}^3$.

**(b)** (1,5 điểm) Cho $U$ là không gian con của không gian Euclid $\mathbb{R}^4$ sinh bởi hai vector $\mathbf{a} = (1,-2,2,1)$ và $\mathbf{b} = (-1,2,-1,0)$.

Tìm một cơ sở trực chuẩn của phần bù trực giao $U^{\perp}$ của $U$. Tìm thành phần trực giao của vector **x** $= (4, 3, 2, 1)$ đối với $U$.

**(c)** (1,5 điểm) Cho dạng toàn phương $p(x, y, z) = x^2 - 2xy + y^2 + 2z^2$ trên $\mathbb{R}^3$. Hãy tìm một cơ sở trực chuẩn $p$-chính tắc của dạng toàn phương $p$.

**Câu 3.** **(a)** (1,5 điểm) Tìm dạng chính tắc $E(\lambda)$ của $\lambda$-ma trận

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

Tìm ma trận Jordan $J$ (nếu có) sao cho dạng chính tắc của $J - \lambda I_3$ cũng là $E(\lambda)$.

**(b)** (1,5 điểm) Tìm dạng chuẩn tắc Jordan $J$ (nếu có) của ma trận $\begin{bmatrix} 0 & 1 & -1 \\ 2 & 1 & -2 \\ 3 & 3 & -4 \end{bmatrix}$ và tìm ma trận $P$ sao cho $P^{-1}AP = J$.

# Tra cứu