HỌC VIỆN KỸ THUẬT QUÂN SỰ
KHOA CÔNG NGHỆ THÔNG TIN

HY ĐỨC MẠNH

Bài giảng: Đại số tuyến tính và Hình học giải tích

Tài liệu học tập cho sinh viên tại Học viện KTQS

Lưu hành nội bộ

Hà Nội — 2014

Mục lục

Lời nói đầu

Bài giảng "Đại số tuyến tính và hình học giải tích" được viết theo đề cương chương trình của Bộ môn Toán - Khoa Công nghệ Thông tin - Học viện Kỹ thuật Quân sự. Tài liệu biên soạn dựa trên các giáo trình của Học viện kỹ thuật quân sự và một số giáo trình dành cho sinh viên các trường đại học kỹ thuật trong và ngoài nước. Đây là *tài liệu cá nhân* biên soạn giảng dạy cho các lớp của chương trình tiên tiến Việt-Nga (75 tiết), cũng như các lớp học viên quân sự và dân sự (60 tiết) tại Học viện. Vì thời lượng học môn này đối với các đối tượng học viên (trừ các lớp chương trình TTVN) đã giảm so với những năm trước đây (chỉ còn 60 tiết) nên hầu hết các kết quả cơ bản chỉ được đưa ra mà không có chứng minh, để hiểu sâu sắc vấn đề sinh viên cần tự đọc chứng minh trong các sách giáo khoa cho môn học này.

Đặc biệt nhấn mạnh rằng khi học viên đọc bài giảng này cần kèm theo hai tài liệu bắt buộc là "đề cương chi tiết môn học" và "đề cương chi tiết bài giảng" đã được các cấp phê duyệt và công bố trên trang web của Khoa Công nghệ Thông tin (http://fit.mta.edu.vn/subjectstat_DH.htm). Đối với những mục (phần) không có trong hai đề cương trên xem là phần đọc thêm của học viên.

Phần bài tập sau mỗi bài sinh viên làm theo yêu cầu và hướng dẫn của "đề cương chi tiết" bài giảng (bài tập trong [4]).

Vì bài giảng biên soạn bằng Latex theo cấu trúc định sẵn gần giống với sách giáo khoa nhưng *không phải là sách giáo khoa*. Để học tập đạt kết quả tốt sinh viên cần có các tài liệu bắt buộc là [3], [4].

Trong tài liệu những tính chất sẽ thường được viết dưới dạng các *mệnh đề*, các kết quả quan trọng được phát biểu trong các *định lý*. Bên cạnh các vấn đề cơ bản của môn học, trong bài giảng chúng tôi có đưa thêm các kiến thức bổ trợ khác (ví dụ như thực hành tính toán số trên phần mềm *Maple*).

Cuối cùng, trong quá trình biên soạn khó tránh khỏi có sai sót chúng tôi hoan nghênh sự phát hiện của học viên để kịp thời sửa chữa.

Tháng 2 năm 2014

Những kí hiệu

$\mathbb{K}$	trường nào đó
$\mathbb{N}$	tập hợp số tự nhiên
$\mathbb{Z}$	tập hợp số nguyên
$\mathbb{Q}$	tập hợp số hữu tỉ
$\mathbb{R}$	tập hợp số thực
$\mathbb{C}$	tập hợp số phức
$\emptyset$	tập hợp rỗng
deg	bậc của đa thức
$M_{m\times n}(\mathbb{K})$	tập các ma trận cỡ $m \times n$ trên $\mathbb{K}$
$M_n(\mathbb{K})$	tập các ma vuông cấp n trên $\mathbb{K}$
$GL_n(\mathbb{K})$	tập các ma vuông cấp n khả nghịch
A^T	ma trận chuyển vị của ma trận A
$\det(A)$	định thức ma trận A
$Trace(A)$	vết của ma trận A
$rank(A)$	hạng của ma trận A
$\sim$	tương đương hoặc đồng dạng giữa hai ma trận
$span$	bao tuyến tính
$dim(V)$	chiều của không gian V
$Im(f)$	không gian ảnh của ánh xạ f
$Ker(f)$	không gian nhân (hạch) của ánh xạ f
$\langle .,. \rangle$	tích vô hướng
$\mathbb{E}^n$	không gian Euclide thực n chiều
$\perp$	trực giao
$\|\|.\|\|$	chuẩn (độ dài)

Chương 1

Ma trận, định thức, hệ phương trình tuyến tính

1.1 Logic, tập hợp, ánh xạ và cấu trúc đại số

1.1.1 Logic mệnh đề và vị từ

Định nghĩa

Định nghĩa 1. *Mệnh đề là các khẳng định mà ta có thể biết nó đúng hoặc sai. Ta thường ký hiệu mệnh đề bởi các chữ cái in hoa A, B, C,...*

Ví dụ 1.

– "Hà nội là thủ đô Việt nam" - mệnh đề đúng.

– Trên tập $\mathbb{R}$ xét quan hệ "nhỏ hơn", khi đó mệnh đề "1<0" là mệnh đề sai.

Khi mệnh đề A đúng là nói mệnh đề nhận giá trị đúng và viết là "A-1", "A-true" (A-t) hoặc "A-đúng" (A-đ), ngược lại ta nói A nhận giá trị sai và viết là "A-0", "A-false" (A-f) hay "A-sai" (A-s). Mệnh đề chỉ nhận hai giá trị đúng hoặc sai và không có khả năng thứ ba.

Các phép toán

a) *Phép tuyển $\vee$ (hoặc, hoặc là):* Giả sử A, B - 2 mệnh đề. A$\vee$B (đọc là A hoặc B, A tuyển B) cũng là một mệnh đề, nó chỉ nhận giá trị sai khi cả A và B đều sai còn đúng trong các trường hợp còn lại.

b) *Phép hội $\wedge$ (và):* Giả sử A, B - 2 mệnh đề. A$\wedge$B (đọc là A và B, A hội B) cũng là một mệnh đề, nó chỉ nhận giá trị đúng khi cả A và B đều đúng còn sai trong các trường hợp còn lại.

A	B	A∨B
1	1	1
1	0	1
0	1	1
0	0	0

Bảng 1.1: Bảng giá trị logic phép tuyển

A	B	A∧B
1	1	1
1	0	0
0	1	0
0	0	0

Bảng 1.2: Bảng giá trị logic phép hội

c) *Phép kéo theo →:* Giả sử A, B - 2 mệnh đề. A→B (đọc là A kéo theo B, A suy ra B) cũng là một mệnh đề, nó chỉ nhận giá trị sai khi A đúng kéo theo B sai. A còn gọi là giả thiết, B gọi là kết luận.

A	B	A→B
1	1	1
1	0	0
0	1	1
0	0	1

Bảng 1.3: Bảng giá trị logic phép kéo theo

d) *Phép tương đương ⇔:* Giả sử A, B - 2 mệnh đề. A⇔B (đọc là A tương đương B) cũng là một mệnh đề, nó chỉ nhận giá trị đúng khi A và B cùng đúng hoặc cùng sai.

A	B	A⇔B
1	1	1
1	0	0
0	1	0
0	0	1

Bảng 1.4: Bảng giá trị logic phép kéo theo

d) *Phép phủ định ⌉:* Giả sử A là mệnh đề. ⌉A (đọc là phủ định A) cũng là một mệnh đề và nó nhận giá trị ngược lại với giá trị A.

A	$\rceil$A
1	0
0	1

Bảng 1.5: Bảng giá trị logic phép phủ định

Công thức và định lý

Từ các mệnh đề ban đầu người ta xây dựng các mệnh đề mới thông qua sử dụng 5 phép toán trên. Các mệnh đề ban đầu gọi là sơ cấp, còn các mệnh đề nhận được gọi là công thức. Các công thức luôn nhận giá trị đúng gọi là công thức hằng đúng, chúng ta chỉ quan tâm các công thức này, các công thức hằng đúng còn gọi là "Định lý" hay "Định luật".

Ví dụ 2. $(A \to B) \to C$ - là một công thức

$(A \to B) \Leftrightarrow (\rceil B \to \rceil A)$ - là một công thức hằng đúng (định lý).

Dưới đây là một số công thức hằng đúng quan trọng:

a) *Giao hoán:*

$$A \vee B \Leftrightarrow B \vee A$$

$$A \wedge B \Leftrightarrow B \wedge A$$

b) *Kết hợp:*

$$(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$$

$$(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$$

c) *Phân phối*

$$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$$

$$A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$$

d) *Lũy đẳng:*

$$A \vee A \Leftrightarrow A$$

$$A \wedge A \Leftrightarrow A$$

e) *Hấp thụ:*

$$(A \vee B) \wedge A \Leftrightarrow A$$

$$(A \wedge B) \vee A \Leftrightarrow A$$

f) *Công thức De Morgan:*

$$\rceil(A \wedge B) \Leftrightarrow (\rceil A) \vee (\rceil B)$$

$$\rceil(A \vee B) \Leftrightarrow (\rceil A) \wedge (\rceil B)$$

g) Công thức chứng minh phản chứng:

$$\rceil(A \to B) \Leftrightarrow A \wedge (\rceil B)$$

Để chứng minh các công thức là hằng đúng ta thay tất cả các giá trị có thể của các mệnh đề sơ cấp, lập bảng giá trị logic từ đó đưa đến kết luận.

Mệnh đề lượng từ

a) Tập hợp , phần tử: Tập hợp là một khái niệm không định nghĩa được mà chỉ có thể mô tả nó.

Ví dụ 3. Tập hợp các sinh viên lớp K48-A.

Ký hiệu tập hợp bởi các chữ cái in hoa A, B, C,... các phần tử của tập hợp ký hiệu bởi chữ cái thường a,b, c,... Ta viết $a \in A$ để chỉ a là phần tử của tập hợp A.

b) Hàm mệnh đề: Ta nói $f(x_1, .., x_n)$ là một mệnh đề n-ngôi xác định trên tập A nếu với mọi ($\forall$) $a_1, ..., a_n \in A$ thì $f(a_1, ..., a_n)$ là một mệnh đề.

Ví dụ 4. $A = \mathbb{R}$, khi đó với $x, y \in \mathbb{R}$ thì "$x \geq y$" là hàm mệnh đề hai ngôi xác định trên A.

c) Lượng từ

i) Lượng từ tồn tại (riêng): Giả sử $f(x)$ là một hàm mệnh đề xác định trên A, mệnh đề "$\exists x f(x)$" - đọc là "tồn tại x để $f(x)$" - nó nhận giá trị đúng khi có $a \in A$ để $f(a)$ là đúng.

ii) Lượng từ phổ biến (chung): Giả sử $f(x)$ là một hàm mệnh đề xác định trên A, mệnh đề "$\forall x f(x)$" - đọc là "với mọi x để $f(x)$" - nó nhận giá trị đúng khi với mỗi $a \in A$ bất kỳ thì $f(a)$ là đúng.

Người ta có thể xây dựng mệnh đề có chứa nhiều lượng từ.

Định lý 1.1.1. *Phủ định của mệnh đề có chứa lượng từ là mệnh đề nhận được bằng cách thay các lượng từ chung thành các lượng từ riêng và hàm mệnh đề thay bằng phủ định của nó.*

Ví dụ 5. $\rceil(\forall x f(x)) \Leftrightarrow \exists x \rceil f(x)$

1.1.2 Tập hợp

Tập hợp và phép toán trên tập hợp

a) Khái niệm: Như trên đã nói, tập hợp là khái niệm không được định nghĩa, người ta ta chỉ mô tả tập hợp. Ký hiệu $a \in A$ để chỉ phần tử a thuộc tập A. Tập không có phần tử nào gọi là *tập rỗng* và ký hiệu là $\emptyset$. Hai tập A và B gọi là *bằng nhau* nếu chúng có chứa các phần tử giống nhau. Tập A gọi là *con* của tập B (hay là B chứa A) nếu mọi phần tử của A đều thuộc tập B, ký hiệu là $A \subseteq B$, vậy:

$$A = B \Leftrightarrow (A \subseteq B) \wedge (B \subseteq A)$$

Quy ước tập $\emptyset$ là con của mọi tập hợp.

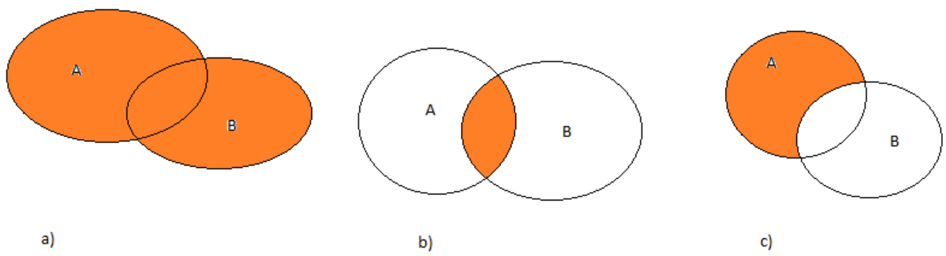
b) Các phép toán trên tập hợp: Giả sử A và B là các tập hợp

i) *Phép hợp:* $A \cup B$ đọc là A *hợp* B là tập các phần tử thuộc ít nhất một trong hai tập hợp đó (Hình 1.1a))

$$A \cup B = \{x | (x \in A) \vee (x \in B)\}$$

ii) *Phép giao:* $A \cap B$ đọc là A *giao* B là tập các phần tử thuộc cả A và B(Hình 1.1b))

$$A \cap B = \{x | (x \in A) \wedge (x \in B)\}$$

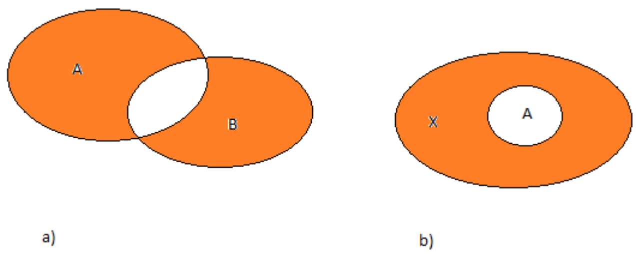


Hình 1.1: Biểu đồ Venn: Hợp, giao và hiệu của hai tập hợp

iii) *Hiệu của hai tập hợp:* $A \setminus B$ đọc là A *trừ* B là tập các phần tử thuộc cả A và không thuộc B (Hình 1.1c))

$$A \setminus B = \{x : (x \in A) \wedge (x \notin B)\}$$

iv) *Hiệu đối xứng của hai tập hợp:* $A \triangle B$ là tập hợp được xác định như sau: $A \triangle B = (A \setminus B) \cup (B \setminus A)$ (Hình 3.2a)).



Hình 1.2: Biểu đồ Venn: Hiệu đối xứng và phần bù

v) *Phần bù:* Giả sử X là một tập hợp và A là tập con của X. Phần bù của A trong X ký hiệu là $\overline{A}$ (hay $C_X A$) và xác định bởi $\overline{A} = X \setminus A$ (Hình 3.2b)).

c) Các tính chất:

i) Giao hoán:

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

ii) Kết hợp:

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

iii) Phân phối:

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

iv) Lũy đẳng:

$$A \cup A = A \cap A = A$$

v) Hấp thụ:

$$(A \cup B) \cap A = A$$
$$(A \cap B) \cup A = A$$

vi) Công thức De Morgan

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$
$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

Tổng quát: $\overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \overline{A_i}$ và $\overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \overline{A_i}$.

Chứng minh các tính chất trên có thể dựa vào các luật tương ứng của logic mệnh đề.

a) Tích Descartes và quan hệ thứ tự bộ phận

Định nghĩa 2. *Giả sử A, B là các tập hợp. Tích Descartes của A và B ký hiệu là $A \times B$ là tập hợp gồm các phần tử có dạng (a, b) ở đó $a \in A$ và $b \in B$.*

Ta có thể mở rộng định nghĩa cho trường hợp tổng quát tích Descartes của n tập hợp : $A_1 \times A_2 \times ... \times A_n$. Khi $A_1 = A_2 = ... = A_n = A$ ta viết là A^n.

Định nghĩa 3. *Giả sử X là một tập hợp. Một quan hệ hai ngôi (hay quan hệ) trên X là một tập con $\mathcal{R}$ của X^2. Nếu $(x, y) \in \mathcal{R}$ ta nói x quan hệ $\mathcal{R}$ với y và viết là $x\mathcal{R}y$, vậy $x\mathcal{R}y \Leftrightarrow (x, y) \in \mathcal{R}$.*

Ví dụ 6. Quan hệ "x chia hết cho y" là quan hệ hai ngôi trên $\mathbb{N}$

$$\mathcal{R} = \{(x, y) \in \mathbb{N}^2 : x \vdots y\}$$

Các tính chất thường gặp:

i) Phản xạ: $\mathcal{R}$ gọi là phản xạ nếu $\forall x \in X \quad x\mathcal{R}x$ tức là $(x, x) \in \mathcal{R}$.

ii) Đối xứng: $\mathcal{R}$ - đối xứng nếu $(\forall x)(\forall y)(x\mathcal{R}y \to y\mathcal{R}x)$

iii) Bắc cầu: $\mathcal{R}$ - bắc cầu nếu $(\forall x)(\forall y)(\forall z)[(x\mathcal{R}y) \wedge (y\mathcal{R}z) \to x\mathcal{R}z]$

iv) Phản đối xứng: $\mathcal{R}$ - phản đối xứng nếu $(\forall x)(\forall y)[(x\mathcal{R}y) \wedge (y\mathcal{R}x) \to x = y]$.

Một quan hệ $\mathcal{R}$ gọi là *tương đương* nếu quan hệ đó có các tính chất i), ii) và iii).

Nếu $\mathcal{R}$ là quan hệ tương đương trên X và phần tử $x \in X$, tập con của X gồm tất cả các phần tử có quan hệ $\mathcal{R}$ với x gọi là lớp tương đương của phần tử x, ký hiệu là $[x]_{\mathcal{R}}$ hay đơn giản là $[x]$. Rõ ràng, phần tử x luôn thuộc lớp tương đương của chính nó và hai lớp tương đương hoặc trùng nhau hoặc không giao nhau. Ta nói rằng tập các lớp tương đương tạo thành một *phân hoạch* của X.

Ví dụ 7. Giả sử có một số nguyên dương n, ta định nghĩa một quan hệ hai ngôi $\mathcal{R}$ trên $\mathbb{Z}$ như sau:

$$\forall x, y \in \mathbb{Z}, x\mathcal{R}y \Leftrightarrow x - y \vdots n$$

Quan hệ này được gọi là quan hệ đồng dư modulo n. Nếu $x\mathcal{R}y$ ta viết là $x \equiv y \ (mod \ n)$. Dễ dàng kiểm tra quan hệ đồng dư modulo n có các tính chất i), ii) và iii), do đó nó là quan hệ tương đương. Có thể chứng minh được tập $\mathbb{Z}$ với quan hệ đồng dư modulo n phân hoạch thành n lớp tương đương $\{[0], [1], ..., [n-1]\}$ (bài tập), ta ký hiệu

$$\mathbb{Z}_n = \{[0], [1], ..., [n-1]\}$$

Quan hệ $\mathcal{R}$ gọi là *thứ tự bộ phận* (hay *từng phần*) nếu có các tính chất i), iii) và iv). Khi đó ta nói tập X với quan hệ thứ tự này là *được sắp một phần*.

Quan hệ thứ tự bộ phận trong X gọi là thứ tự *hoàn toàn* (hay *toàn phần*) nếu với mọi a, b trong X ta có $a\mathcal{R}b$ hoặc $b\mathcal{R}a$, khi đó tập X là *được sắp hoàn toàn*.

Ví dụ 8. Tập số thực $\mathbb{R}$ với quan hệ $\leq$ là tập được sắp hoàn toàn.

Giả sử $\mathcal{R}$ là quan hệ thứ tự bộ phận trong X, tập hợp $A \subseteq X$. Phần tử a_0 gọi là *bé nhất* trong A nếu $\forall a \in A$ ta có $a_0\mathcal{R}a$ tức là $(a_0, a) \in \mathcal{R}$.

Một tập được sắp hoàn toàn sẽ được gọi là *được sắp tốt* khi và chỉ khi mọi tập con khác rỗng của nó đều có phần tử bé nhất.

Ví dụ 9. Tập hợp $\mathbb{N}$ với quan hệ $\leq$ là tập được sắp tốt, còn tập $\mathbb{Z}$ cũng với quan hệ này không phải được sắp tốt do $\mathbb{Z}$ không có phần tử bé nhất.

b) Nguyên lý quy nạp toán học trên tập số tự nhiên $\mathbb{N}$

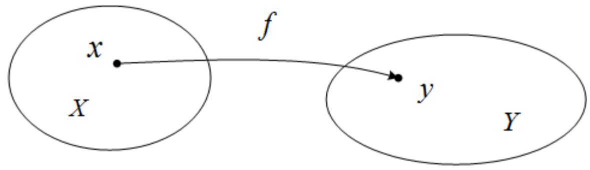
Định lý 1.1.2. *Mệnh đề $f(n)$ phụ thuộc $n \in \mathbb{N}$ sẽ đúng cho mọi n nếu thỏa mãn hai điều kiện:*
i) $f(1)$ - đúng
ii) Từ $f(k)$ - đúng kéo theo $f(k+1)$ - đúng với mọi $k \in \mathbb{N}$.

Chứng minh: Ta chứng minh bằng phản chứng. Giả sử tồn tại m để $f(m)$ sai. Do tập $\mathbb{N}$ với quan hệ $\leqslant$ là được sắp tốt nên nếu gọi $M \subseteq \mathbb{N}$ là tập hợp $\{m : f(m) - \text{sai}\}$ thì M có phần tử bé nhất, ký hiệu m_0, hiển nhiên do giả thiết i) nên $m_0 \neq 1, m_0 \geq 2$. Vì $f(m_0)$ - sai và m_0 bé nhất nên $f(m_0 - 1)$ - đúng. Tuy nhiên theo ii) từ đây lại có $f(m_0 - 1 + 1) = f(m_0)$ - đúng. Điều đó dẫn đến mâu thuẫn. Mâu thuẫn này chứng tỏ giả thiết phản chứng là sai, tức là ta có điều phải chứng minh. ▶

1.1.3 Ánh xạ. Lực lượng của tập hợp.

a) Các định nghĩa: Giả sử X, Y là các tập hợp.

Định nghĩa 4. *Ánh xạ f từ X vào Y ký hiệu là $f : X \to Y$ là một quy tắc cho tương ứng mỗi phần tử $x \in X$ với một và chỉ một phần tử $y \in Y$. Khi đó ta nói y là ảnh của x và viết là $y = f(x)$. X - gọi là tập xác định của ánh xạ f hay tập nguồn, Y - gọi là tập đích.*



Hình 1.3: Ánh xạ

Xét hai ánh xạ $f : X \to Y$ và $g : X \to Y$. f và g gọi là *bằng nhau* và viết là $f = g$ nếu $f(x) = g(x), \forall x \in X$.

Định nghĩa 5. *Giả sử $f : X \to Y$, $A \subset X$, $B \subset Y$. Tập hợp $f(A) := \{y \in Y : \exists x \in A, y = f(x)\}$ gọi là tập ảnh của A. Tập hợp $f^{-1}(B) = \{x \in X : f(x) \in B$ gọi là nghịch ảnh của B bởi f.*

Quy ước $f(\varnothing) = \varnothing$.

Ví dụ 10. Các hàm số đã học ở phổ thông là những ánh xạ, ví dụ $\sin : \mathbb{R} \to [-1, 1]$ với $x \mapsto \sin(x)$.

Tính chất:

i) $A \subset f^{-1}(f(A))$, $f(f^{-1}(B)) \subset B$

ii) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$, $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$

iii) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$, $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$

Định nghĩa 6. *Ta nói*

i) f - toàn ánh (lên, tràn ánh) nếu $f(X) = Y$, tức là $\forall y \in Y \exists x \in X : y = f(x)$ hay nói cách khác $f^{-1}(y)$ có không ít hơn một phần tử.

ii) f - đơn ánh nếu $\forall x, x' \in X$ từ $x \neq x' \to f(x) \neq f(x')$. Vậy f - đơn ánh khi và chỉ khi với $y \in Y$ tập $f^{-1}(y)$ có không quá một phần tử.

iii) f - song ánh nếu nó vừa đơn ánh vừa toàn ánh, tức là $\forall x \in X \exists! y \in Y : y = f(x)$. Một song ánh $X \to X$ còn gọi là một phép thế trên X.

Ví dụ 11. $f(x) = \sin x$ là toàn ánh vì với mọi $\alpha \in [-1, 1]$ tồn tại $x = \arcsin \alpha + 2k\pi, k \in \mathbb{Z}$ để $f(x) = \alpha$.

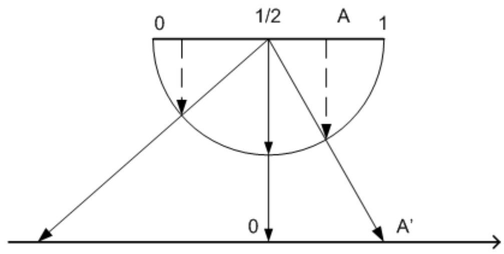
Ví dụ 12. $f : A \to f(A) \subset Y$ mà đơn ánh sẽ là song ánh.

Định nghĩa 7. *Giả sử A, B là hai tập hợp, các phần tử của chúng thuộc một loại nào đó. Nếu có một song ánh (tương ứng 1-1) giữa các phần tử của A và B thì ta nói rằng A tương đương B ký hiệu là $A \sim B$.*

Dễ thấy rằng quan hệ tương đương này thực sự là quan hệ tương đương theo định nghĩa ở trên.

Ví dụ 13. Xét tập $\mathbb{N} = \{1, 2, ..., n, ...\}$ và tập $M = \{2, 4, ..., 2n, ...\}$. Phép tương ứng $n \Leftrightarrow 2n$ là tương ứng 1-1.

Ví dụ 14. Khoảng $(0; 1)$ tương đương với trục số thực $\mathbb{R}$. Có nhiều cách chứng minh điều này, có thể xét phép tương ứng $1-1$ như trong Hình 1.4.



Hình 1.4: Tương ứng $1-1$ giữa $(0; 1)$ và $\mathbb{R}$

Định nghĩa 8. *Lực lượng của tập hợp A bất kỳ là "cái chung" có trong tất cả các tập hợp tương đương với A. Nếu A hữu hạn thì lực lượng A chính là số phần tử (không trùng nhau) trong A. Lực lượng của A ký hiệu là $|A|$. Theo định nghĩa $A \sim B$ nếu $|A| = |B|$.*

Tập hợp tương đương với tập $\mathbb{N}$ gọi là tập đếm được. Ví dụ $\mathbb{Z}, \mathbb{Q}$ là đếm được (bài tập).

b) Ánh xạ ngược:

Định nghĩa 9. *Cho hai ánh xạ $f : X \to Y$ và $g : Y \to Z$, ánh xạ từ X vào Z xác định bởi $x \mapsto g(f(x))$ gọi là hợp thành (tích) của g và f ký hiệu là $g \circ f$ (hay gf).*

Chú ý rằng $g \circ f$ chỉ xác định khi tập đích của f trùng với tập nguồn của g.

Tính chất:

i) $h \circ (g \circ f) = (h \circ g) \circ f$

ii) $(g \circ f)(A) = g(f(A))$

iii) $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B)), \forall B \subset Z$

iv) $g \circ f$ - đơn ánh thì f - đơn ánh

v) $g \circ f$ - toàn ánh thì g - toàn ánh.

Định nghĩa 10. *i) Ánh xạ $Id_X : X \to X$ sao cho $Id_X(x) = x, \forall x \in X$ gọi là ánh xạ đồng nhất trên X.*
ii) Giả sử $f : X \to Y$ là một ánh xạ. f gọi là khả nghịch nếu tồn tại ánh xạ $g : Y \to X$ sao cho $g \circ f = Id_X$ và $f \circ g = Id_Y$. Khi đó g gọi là ánh xạ ngược hay nghịch đảo của f và ký hiệu là f^{-1}.

Ví dụ 15. Các hàm ngược đã biết ở phổ thông cho ta các ví dụ về ánh xạ ngược.

Định lý 1.1.3. *(Tồn tại ánh xạ ngược) Ánh xạ $f : X \to Y$ có f^{-1} khi và chỉ khi f là song ánh.*

1.1.4 Sơ lược về cấu trúc đại số

Nhóm, vành, trường:

a) Phép toán hai ngôi (phép toán trong):

Định nghĩa 11. *Giả sử X tập hợp khác rỗng. Một phép toán hai ngôi trên X là một ánh xạ $\circ : X \times X \to X : (x, y) \mapsto x \circ y$. Phép toán $\circ$ gọi là hợp thành của x và y.*

Ví dụ 16. Phép cộng $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (x, y) \mapsto x + y$

Các tính chất thường gặp:

i) Kết hợp: $\forall x, y, z \in X$ ta có $x \circ (y \circ z) = (x \circ y) \circ z$

ii) Giao hoán: $\forall x, y \in X : x \circ y = y \circ x$

iii) Phân phối: Giả sử có hai phép toán hai ngôi $*$ và $\circ$ trên X. Phép toán $*$ gọi là có phân phối bên trái với phép toán $\circ$ nếu $\forall x, y, z \in X : x * (y \circ z) = (x * y) \circ (x * z)$. Tương tự ta có tính chất phân phối bên phải. Nếu không nói gì ta hiểu phép toán có tính chất phân phối cả hai phía.

Định nghĩa 12. *Phần tử $e \in X$ gọi là phần tử trung hòa của phép toán $\circ$ nếu $\forall x \in X : x \circ e = e \circ x = x$. Nếu ký hiệu $\circ$ theo lối cộng $(+)$ thì phần tử* *trung hòa ký hiệu là* 0 *(không), còn theo lối nhân* $(\cdot)$ *thì phần tử trung hòa ký hiệu là* 1 *(một).*

Định nghĩa 13. *Giả sử* X *có phần tử trung hòa* e*. Ta nói* x *là phần tử khả đối xứng nếu tồn tại* $x' \in X : x \circ x' = x' \circ x = e$*. Phần tử* x' *gọi là phần tử đối xứng của* x*. Khi* $\circ$ *viết theo lối cộng ta ký hiệu* x' *là* $-x$ *và gọi là phần tử đối, còn lối nhân ta ký hiệu là* x^{-1} *và gọi là phần tử nghịch đảo.*

Dễ thấy phần tử trung hòa là duy nhất. Nếu X với phép toán $\circ$ có tính chất kết hợp và phần tử trung hòa e thì phần tử đối xứng cũng là duy nhất.

b) Nhóm:

Định nghĩa 14. *Giả sử tập hợp* $X \neq \varnothing$ *với phép toán hai ngôi* $\circ$ *đã cho. Ký hiệu* $(X, \circ)$ *gọi là một nhóm nếu phép toán là kết hợp, có phần tử trung hòa e và mọi phần tử đều khả đối xứng (khả nghịch). Ngoài ra nếu* $\circ$ *có tính chất giao hoán thì nhóm được gọi là nhóm giao hoán (hay Abel).*

Ta thường ký hiệu $(X, \circ, e)$ để chỉ một nhóm với phần tử trung hòa e.

Ví dụ 17. $(\mathbb{R}, +, 0)$ và $(\mathbb{Q} \setminus \{0\}, \cdot, 1)$ là các nhóm Abel.

c) Vành: Xét hai phép toán trên tập $X \neq \varnothing$ là

$$+ : X \times X \to X : (x, y) \mapsto x + y$$

và

$$\cdot : X \times X \to X : (x, y) \mapsto x \cdot y$$

Định nghĩa 15. *Ta nói tập* X *với hai phép toán trên lập thành một vành* $(X, +, \cdot)$ *nếu:*

i) $(X, +, 0)$ *là nhóm Abel*

ii) Phép nhân $(\cdot)$ *có tính chất kết hợp và phân phối với phép cộng* $(+)$*.*

Ngoài ra nếu phép nhân có đơn vị 1 thì ta nói *vành có đơn vị*, phép nhân có tính chất giao hoán thì gọi là *vành giao hoán*.

Ví dụ 18. $(\mathbb{Z}, +, \cdot)$ và $(\mathbb{R}, +, \cdot)$ là các vành giao hoán, có đơn vị.

d) Trường:

Định nghĩa 16. *Trường là một vành giao hoán có đơn vị* $1 \neq 0$ *và mọi phần tử khác* 0 *đều khả nghịch (đối với phép nhân).*

Ví dụ 19. $\mathbb{Q}, \mathbb{R}, \mathbb{Z}_p$ (với p nguyên tố) là các trường (bài tập).

1.1.5 Số phức:

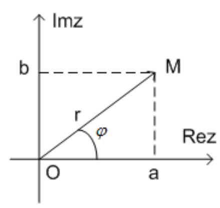
a) Định nghĩa: Trên $\mathbb{R}^2$ trang bị hai phép toán như sau:

- Phép cộng $(+)$: $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$

- Phép nhân $\cdot$: $(a_1, b_1) \cdot (a_2, b_2) = (a_1 \cdot a_2 - b_1 \cdot b_2, a_1 \cdot b_2 + a_2 \cdot b_1)$ ở đó $(a_1, b_1)^{-1} = \left(\frac{a_1}{a_1^2+b_1^2}, \frac{-b_1}{a_1^2+b_1^2}\right)$ với $(a_1, b_1) \neq (0, 0)$.

Dễ thấy phần tử $(0, 0)$ là phần tử trung hòa của phép cộng, $(1, 0)$ là phần tử đơn vị của phép nhân. Khi đó có thể kiểm tra được $(\mathbb{R}^2, +, \cdot)$ là một trường và gọi là *trường số phức*, ký hiệu là $\mathbb{C}$. Vậy mỗi số phức $z \in \mathbb{C}$ là $z = (a, b) \in \mathbb{R}^2$. Vì mỗi số thực $x \in \mathbb{R}$ ta có thể đồng nhất với $(x, 0) \in \mathbb{C}$, khi đó có thể coi $\mathbb{R} \subset \mathbb{C}$.

b) Đơn vị ảo, dạng đại số của số phức: Ta gọi số phức $(0, 1)$ là i và gọi là đơn vị ảo. Như vậy $(0, 1) \equiv i$ nên $(0, b) = ib$. Khi đó $i^2 = (0, 1) \cdot (0, 1) = (-1, 0)$, ta đồng nhất nó với -1 theo lập luận trên, từ đó $i^2 = -1$. Vậy với $z \in \mathbb{C}$ thì $z = (a, b) = (a, 0) + (0, b) = a + ib$ và được gọi là *dạng đại số* của số phức. Ký hiệu $Rez = a, Imz = b$ gọi là phần thực và phần ảo tương ứng. Với việc cho tương ứng số phức $z = (a, b) \in \mathbb{R}^2$, đặt điểm $M = (a, b)$ trên mặt phẳng, ta có thể biểu diễn số phức trên mặt phẳng tọa độ Descartes như là $z = \overrightarrow{OM}$. Trục hoành Ox được gọi là trục thực, trục tung Oy gọi là trục ảo, mặt phẳng biểu diễn số phức gọi là mặt phẳng phức (Hình 1.5).



Hình 1.5: Biểu diễn số phức trên mặt phẳng

Định nghĩa 17. *Cho số phức $z = a + ib$, khi đó $a - ib$ gọi là liên hợp của z ký hiệu là $\overline{z}$.*

Dễ thấy rằng: $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$ và $z \cdot \overline{z} = |z|^2$ ở đây $|z|^2 = a^2 + b^2$.

c) Dạng lượng giác của số phức: Cho số phức $z = a + ib \neq 0$, có thể viết lại z như sau:

$$z = \sqrt{a^2 + b^2} \left(\frac{a}{\sqrt{a^2 + b^2}} + \frac{b}{\sqrt{a^2 + b^2}} i \right)$$

Đặt $r := |z| = \sqrt{a^2 + b^2}$ gọi là modul của z, góc φ giữa $\overrightarrow{OM}$ với trục thực gọi là argument của z ký hiệu $\varphi := arg(z)$. Dễ thấy:

$$\begin{cases} a = r\cos\varphi \\ b = r\sin\varphi \end{cases}$$

khi đó $z = r(\cos\varphi + i\sin\varphi)$ gọi là *dạng lượng giác* của số phức. Với mỗi số thực φ ta đặt

$$e^{i\varphi} := \cos\varphi + i\sin\varphi$$

thì $z = re^{i\varphi}$ gọi là *dạng mũ* của số phức.

Dễ thấy module và argumen của số phức có các tính chất đơn giản sau:

i) $|\overline{z}| = |z|\,;\ |z_1 z_2| = |z_1|\,|z_2|\,;\ |z^n| = |z|^n$

ii) $\arg(z_1 + z_2) = \arg(z_1) + \arg(z_2)\,;\ \arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2)\,;\ \arg(z^n) = n\arg(z)$

Công thức Moivre: Giả sử $z = r(\cos\varphi + i\sin\varphi)$, khi đó

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

Dễ dàng chứng minh công thức này bằng quy nạp theo n.

d) Căn bậc n của số phức: Giả sử ta có số phức dạng lượng giác $z = r(\cos\varphi + i\sin\varphi)$. Ta gọi *căn bậc* n của số phức z là tập hợp

$$\mathbb{C}_{\sqrt[n]{z}} = \{\mathrm{w} \in \mathbb{C} : \mathrm{w}^n = z\}$$

Để hiểu rõ ta viết tập hợp này một cách tường minh hơn. Đặt:

$$\mathrm{w} = \rho(\cos\theta + \mathrm{i}\sin\theta)$$

Theo công thức Moivre ta có:

$$\mathrm{w}^n = \rho^n(\cos n\theta + \mathrm{i}\sin n\theta)$$

do vậy

$$\begin{cases} \rho^n \mathrm{cosn}\theta = r\cos\varphi \\ \rho^n \mathrm{sinn}\theta = r\mathrm{sin}\varphi \end{cases}$$

từ đó ta có

$$\begin{cases} \rho = \sqrt[n]{r} \\ \theta = \frac{\varphi + k2\pi}{n}, k \in \mathbb{Z} \end{cases}$$

Khi đó căn bậc n của z viết lại là

$$\mathbb{C}_{\sqrt[n]{z}} = \left\{ \sqrt[n]{r}\left(\cos\frac{\varphi + k2\pi}{n} + \mathrm{i}\sin\frac{\varphi + k2\pi}{n}\right), k = 0; 1; ...; n-1 \right\}$$

Vậy căn bậc n của $z \neq 0$ có đúng n giá trị khác nhau.

Vành đa thức:

Giả sử $\mathbb{K}$ là trường, một *đa thức* (một biến) trên $\mathbb{K}$ là biểu thức dạng

$$p(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n$$

ở đó x là biến, $a_i \in \mathbb{K}$ với $i = \overline{1, n}$ là các hệ số, $n \in \mathbb{Z}^+$.

Nếu tất cả các hệ số a_i bằng 0 ta có *đa thức hằng không* và viết là $p(x) = 0$.

Nếu $a_0 \neq 0 = a_1 = ... = a_n$ thì $p(x)$ gọi là *đa thức hằng.*

Nếu $a_n \neq 0$ thì $p(x)$ gọi là *đa thức bậc n*, ký hiệu bậc của đa thức là $deg(p) = n$. Với cách hiểu như vậy thì đa thức hằng là các đa thức bậc 0, đa thức hằng 0 không có bậc (đôi khi người ta coi nó có bậc $-\infty$). Nếu $p(x)$ có dạng $ax^n, a \in \mathbb{K}$ thì gọi là *đơn thức.*

Hai đa thức $p(x)$ và $q(x)$ gọi là *bằng nhau*, viết là $p(x) = q(x)$, nếu các hệ số tương ứng của chúng bằng nhau.

Giả sử ta có hai đa thức

$$p(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n$$

và

$$q(x) = b_0 + b_1x + b_2x^2 + ... + b_mx^m$$

khi đó *tổng* và *tích* của hai đa thức định nghĩa như sau:

$$p(x)+q(x) = \begin{cases} (a_0 + b_0) + ... + (a_n + b_n)x^n + b_{n+1}x^{n+1} + ... + b_mx^m, m \geq n \\ (a_0 + b_0) + ... + (a_m + b_m)x^m + b_{m+1}x^{m+1} + ... + b_nx^n, m < n \end{cases}$$

và

$$p(x)q(x) = c_0 + c_1x + ... + c_{m+n}x^{m+n}$$

ở đó $c_k = \sum\limits_{i=0}^{k} a_ib_{k-i}, k = \overline{0, m+n}$. Dễ thấy

i) $\deg(p(x) + q(x)) \leq max\{\deg(p(x)), \deg(q(x))\}$

ii) $\deg(p(x)q(x)) = \deg(p(x)) + \deg(q(x))$

Ký hiệu $\mathbb{K}[x]$ là tập tất cả các đa thức (một biến) trên trường $\mathbb{K}$, khi đó có thể chứng minh $\mathbb{K}[x]$ với hai phép toán trên lập thành một vành giao hoán có đơn vị (đơn vị ở đây là đa thức hằng bằng 1) (bài tập).

Định lý 1.1.4. *(Phép chia Euclide) Giả sử $\mathbb{K}$ là một trường, và $p(x), q(x) \in \mathbb{K}[x]$, $q(x) \neq 0$. Khi đó tồn tại duy nhất các đa thức $h(x)$ và $r(x)$ sao cho $p(x) = h(x)q(x) + r(x)$ với $deg(r) < deg(q)$.*

Các đa thức $h(x)$ và $r(x)$ lần lượt gọi là *thương* và *phần dư* trong phép chia $p(x)$ cho $q(x)$. Trong tính toán ta có thể thực hiện phép chia đa thức như sắp đặt chia số nguyên.

Ví dụ 20. Để chia đa thức $2x^3 - x^2 + x - 1$ cho đa thức $x^2 + 1$, ta thực hiện như trong Hình 1.6.

$$\begin{array}{c|c} 2x^3 - x^2 + x - 1 & x^2 + 1 \\ \cline{2-2} 2x^3 + 2x & 2x - 1 \\ \hline -x^2 - x - 1 & \\ -x^2 - 1 & \\ \hline -x & \end{array}$$

Hình 1.6: Chia đa thức

Cho hai đa thức $p(x), q(x) \in \mathbb{K}[x]$, ở đó $\mathbb{K}$ là trường và $q(x) \neq 0$. Nếu tồn tại đa thức $h(x) \in \mathbb{K}[x]$ sao cho $p(x) = h(x)q(x)$ thì ta nói $p(x)$ *chia hết* cho $q(x)$ (hay $q(x)$ là *ước* của $p(x)$ trong $\mathbb{K}[x]$).

Một đa thức $d(x)$ là ước của cả $p(x)$ và $q(x)$ gọi là *ước chung* của $p(x)$ và $q(x)$. Nếu $d(x)$ là ước chung của $p(x)$ và $q(x)$ đồng thời nó chia hết cho mọi ước chung khác của $p(x)$ và $q(x)$ thì gọi là ước chung lớn nhất của chúng, viết tắt là UCLN, ký hiệu là $d(x) = (p(x), q(x))$.

Để đảm bảo tính duy nhất của UCLN người ta quy ước hệ số bậc cao nhất của UCLN là 1. Để tìm UCLN ta dùng thuật chia Euclide bằng cách thực hiện một số hữu hạn các phép chia liên tiếp sau đây:

$$\begin{gathered} p(x) = q(x)h(x) + r(x), \deg(r) < \deg(q) \\ q(x) = r(x)h_1(x) + r_1(x), \deg(r_1) < \deg(r) \\ \cdots \\ r_{k-2}(x) = r_{k-1}(x)h_k(x) + r_k(x), \deg(r_k) < \deg(r_{k-1}) \\ r_{k-1}(x) = r_k(x)h_{k+1}(x) \end{gathered}$$

đa thức cuối cùng khác 0 trong dãy phép chia trên là $r_k(x)$ và sau khi lấy $r_k(x)$ chia cho hệ số bậc cao nhất của nó ta được UCLN của $p(x)$ và $q(x)$.

Từ thuật toán Euclide ta thấy rằng, nếu $d(x) = (p(x), q(x))$ thì có thể tìm được các đa thức $u(x), v(x) \in \mathbb{K}[x]$ sao cho

$$p(x)u(x) + q(x)v(x) = d(x)$$

Một *nghiệm* trên $\mathbb{K}$ của $p(x)$ là phần tử $\alpha \in \mathbb{K}$ sao cho khi thay vào $p(x)$ ta được biểu thức đồng nhất bằng 0.

Định lý 1.1.5. *(Bezout) Phần tử $\alpha \in \mathbb{K}$ là nghiệm của $p(x) \in \mathbb{K}[x]$ khi và chỉ khi $p(x)$ chia hết cho $x - \alpha$.*

Như vậy nếu α là nghiệm $p(x)$ thì tồn tại đa thức $h(x)$ sao cho

$$p(x) = (x - \alpha)h(x)$$

Nếu có một đa thức $r(x)$ sao cho $p(x) = (x - \alpha)^k r(x) \quad (k \in \mathbb{N})$ nhưng không thể biểu diễn $p(x)$ dưới dạng $p(x) = (x - \alpha)^{k+1} s(x)$ với $s(x) \in \mathbb{K}[x]$ thì k gọi là *bội* của nghiệm α, và α gọi là *nghiệm bội* k của $p(x)$. Khi $\mathbb{K} = \mathbb{C}$ ta có kết quả sau đây và thường gọi là *Định lý cơ bản của đại số học*.

Định lý 1.1.6. *(Định lý cơ bản của đại số học) Mọi đa thức hệ số phức bậc $n \geq 1$ có ít nhất một nghiệm phức. Nói cách khác, một đa thức cấp n sẽ có đủ n nghiệm phức kể cả bội.*

Cho $p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n \in \mathbb{K}[x]$ và $\alpha \in \mathbb{K}$. Ta có thể dùng sơ đồ Horner (xem Hình (1.7)) để tìm $h(x) = b_0 + b_1 x + b_2 x^2 + ... + b_{n-1} x^{n-1}$ và $r = p(\alpha)$ trong thuật chia Euclide $p(x) = (x - \alpha)h(x) + r$

	a_n	a_{n-1}	$\cdots$	a_1	a_0
α	$b_{n-1} = a_n$	$b_{n-2} = a_{n-1} + \alpha b_{n-1}$	$\cdots$	$b_0 = a_1 + \alpha b_1$	$r = a_0 + \alpha b_0$

Hình 1.7: Sơ đồ Horner

Ví dụ 21. Cho đa thức $p(x) = x^4 + x^3 - 2x^2 + 1$ trên $\mathbb{Q}[x]$ và $\alpha = 2 \in \mathbb{Q}$, sử dụng sơ đồ trên ta có:

	1	1	−2	0	1
2	1	3	4	8	17

Hình 1.8: Áp dụng sơ đồ Horner

Vậy $p(x) = (x - 2)h(x) + r$, trong đó $h(x) = x^3 + 3x^2 + 4x + 8$ còn $r = p(2) = 17$

Khai triển Taylor của đa thức: Cho $p(x) \in \mathbb{K}[x]$ và $deg(p) = n$. Với mỗi $\alpha \in \mathbb{K}$ đa thức trên có thể khai triển duy nhất dưới dạng:

$$p(x) = \sum_{k=0}^{n} c_k (x - \alpha)^k.$$

Không khó để chứng minh điều này dựa trên định lý về phép chia Euclide. Nhờ lược đồ Horner có thể thu được các hệ số c_k từ bảng sau

	a_n	a_{n-1}	$\cdots$	a_1	a_0
α	a_n	$*$	$\cdots$	$*$	$\boxed{c_0}$
α	a_n	$*$	$\cdots$	$\boxed{c_1}$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$		
α	a_n	$\boxed{c_{n-1}}$			
α	$\boxed{c_n = a_n}$				

Hình 1.9: Sơ đồ Horner cho khai triển Taylor

Ví dụ 22. Phân tích đa thức $p(x) = x^4 + x^3 - 2x^2 + 1$ theo các lũy thừa của $x - 2$. Ta lập sơ đồ Horner (xem Hình 1.10).

	1	1	−2	0	1
2	1	3	4	8	$\boxed{17}$
2	1	5	14	$\boxed{36}$	
2	1	7	$\boxed{28}$		
2	1	$\boxed{9}$			
2	$\boxed{1}$				

Hình 1.10: Sử dụng sơ đồ Horner khai triển Taylor đa thức

Từ đó ta có $p(x) = (x-2)^4 + 9(x-2)^3 + 28(x-2)^2 + 36(x-2) + 17$.

Chú ý 1. *Với $\mathbb{K}$ là các trường thông thường $(\mathbb{Q}, \mathbb{R}, \mathbb{C})$, từ lý thuyết giải tích có thể thấy các hệ số c_k chính là đạo hàm cấp k của $p(x)$ tại α, tức là*

$$c_k = \frac{p^{(k)}(\alpha)}{k!}$$

Công thức Viet. Cho đa thức $p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n \in \mathbb{K}[x], a_n \neq 0$. Giả sử $p(x)$ có n nghiệm (kể cả bội) là $\alpha_1, .., \alpha_n \in \mathbb{K}$. Khi đó ta có

$$p(x) = a_n(x - \alpha_1)...(x - \alpha_n)$$

Khai triển vế phải và so sánh các lũy thừa cùng bậc ta được công thức Viet biểu thị các hệ số theo các nghiệm của nó:

$$\begin{cases} \frac{a_{n-1}}{a_n} = -\sum\limits_{i=1}^{n} \alpha_i \\ \frac{a_{n-2}}{a_n} = -\sum\limits_{1\le i<j\le n}^{n} \alpha_i\alpha_j \\ \vdots \\ \frac{a_{n-k}}{a_n} = (-1)^k \sum\limits_{1\le i_1<\ldots<i_k\le n}^{n} \alpha_{i_1}\alpha_{i_2}\ldots\alpha_{i_k} \\ \vdots \\ \frac{a_0}{a_n} = (-1)^n \alpha_1\alpha_2\ldots\alpha_n \end{cases}$$

Rõ ràng việc hoán vị các nghiệm đa thức không làm công thức Viet thay đổi.

Quy tắc đổi dấu Descartes. Quy tắc dấu Descartes dưới đây là một công cụ hữu ích để tìm số không điểm (nghiệm) của một đa thức một biến thực khi mà ta không nhìn vào đồ thị của nó. Ở đây ta hiểu số lần đổi đổi dấu của một đa thức là số lần đổi dấu của các hệ số (khác không) trong đa thức đó (đa thức được viết theo chiều tăng hoặc giảm của lũy thừa).

Ví dụ 23. Đa thức

$$p\left(x\right) = x^4 + x^2 - 2$$

có 1 lần đổi dấu.

Quy tắc Descartes được phát biểu như sau:

Định lý 1.1.7. *Giả sử N là số không điểm (nghiệm) dương của đa thức một biến thực*

$$p\left(x\right) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

còn S là số lần đổi dấu của đa thức. Khi đó $S \ge N$ và $S - N$ là một số chẵn.

Dễ dàng thu được hệ quả tương tự đối với số không điểm âm khi ta xét đa thức $p(-x)$.

Ví dụ 24. Vì đa thức

$$p\left(x\right) = x^4 + x^2 - 2$$

có 1 lần đổi dấu nên theo quy tắc dấu Descartes đa thức có đúng một không điểm dương. Xét

$$p\left(-x\right) = x^4 + x^2 - 2$$

cũng có một lần đổi dấu nên đa thức có đúng một không điểm âm. Như vậy đa thức có hai nghiệm thực: một âm, một dương, cặp nghiệm còn lại chính là cặp nghiệm phức liên hợp. Không khó để thấy các nghiệm của đa thức này là $\pm 1, \pm 2i$.

1.2 Ma trận

1.2.1 Ma trận

Giả sử $\mathbb{K}$ là một trường (thực hoặc phức), m, n là các số tự nhiên.

Định nghĩa 18. *Ma trận cấp $m \times n$ trên trường $\mathbb{K}$ là một mảng chữ nhật gồm m hàng, n cột với $m \times n$ phần tử $a_{ij}, i = \overline{1;m}, j = \overline{1;n}$ dạng*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

và viết đơn giản ở dạng $A = (a_{ij})_{m \times n}$.

$m \times n$ gọi là kích thước (cấp, cỡ) ma trận (theo thói quen ta người ta thường dùng từ "cấp" cho ma trận vuông), tên các ma trận thường được đặt bởi các chữ cái in hoa $A, B, ...$ Tập các ma trận cỡ $m \times n$ ký hiệu bởi $M_{m \times n}(\mathbb{K})$. Khi $m = n$ ta nói ma trận vuông cấp n, tập các ma trận vuông cấp n ký hiệu $M_n(\mathbb{K})$.

Hai ma trận gọi là "*bằng nhau*" nếu các phần tử tương ứng của chúng bằng nhau.

Ma trận O là ma trận gồm các phần tử bằng 0, tức là $a_{ij} = 0 \forall i, j$.

Ma trận đơn vị cấp n là ma trận vuông trên $\mathbb{K}$ với các phần tử trên đường chéo chính bằng 1, các phần tử còn lại bằng 0, ký hiệu là $E_n = diag(1, 1, ..., 1)$ hoặc đơn giản là E khi đã biết cấp của nó (một số tài liệu ký hiệu I_n) dạng

$$E = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$

Nếu dùng ký hiệu Kronecker

$$\delta_{\text{ij}} = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$$

thì $E = (\delta_{\mathrm{ij}})_n$.

Khi $m = 1$ hoặc $n = 1$ ta được các *ma trận một hàng hoặc một cột* (hay gọi là các *vector hàng (cột)*).

1.2.2 Các phép toán trên ma trận

Định nghĩa 19. *Giả sử có hai ma trận cỡ $m \times n$*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

khi đó ta viết

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

và gọi là tổng của hai ma trận A và B.

Mệnh đề 1.2.1. *Giả sử A, B, C là các ma trận cỡ $m \times n$, O là ma trận các phần tử đều là 0, khi đó:*
i) $A + B = B + A$
ii) $A + (B + C) = (A + B) + C$
iii) $A + O = A$ và
iv) Tồn tại ma trận A' sao cho: $A + A' = O$, khi đó ký hiệu A' là $-A$.

Như vậy tập các ma trận với phép toán cộng lập thành một nhóm Abel.

Định nghĩa 20. *Giả sử có ma trận cấp $m \times n$*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

và hằng số $c \in \mathbb{K}$. Chúng ta viết

$$cA = \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & \vdots & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{pmatrix}$$

và gọi là tích của ma trận A với hằng số c.

Mệnh đề 1.2.2. *Giả sử A, B là các ma trận cỡ $m \times n$, và $c, d \in \mathbb{K}$, khi đó:*
i) $c(A + B) = cA + cB$
ii) $(c + d)A = (cA + dA)$
iii) $0A = O$ *và*
iv) $c(dA) = (cd)A$.

Đặc biệt với $c = -1 \in \mathbb{K}$, thay vì viết $(-1)A$ ta viết là $-A$. Dễ thấy các tính chất trên suy ra từ tính chất phép toán trên trường $\mathbb{K}$

Định nghĩa 21. *Giả sử có hai ma trận cỡ $m \times n$ và $n \times p$ tương ứng là*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, B = \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \vdots & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix}$$

khi đó tích của hai ma trận A và B là ma trận cỡ $m \times p$ được cho dưới dạng

$$AB = \begin{pmatrix} q_{11} & \cdots & q_{1p} \\ \vdots & \vdots & \vdots \\ q_{m1} & \cdots & q_{mp} \end{pmatrix}$$

trong đó phần tử q_{ij} của ma trận tích được xác định bởi

$$q_{\mathrm{ij}} = \sum_{k=1}^{n} a_{ik} b_{kj}, i = \overline{1; m}, j = \overline{1; p}$$

Chú ý 2. *Để thực hiện được phép nhân hai ma trận thì số cột ma trận đầu tiên phải bằng số hàng ma trận thứ hai.*

Khi $A \in M_n(\mathbb{K})$ thì $A.A$ viết là A^2, quy ước $A^n = A.A....A$ (n ma trận A nhân với nhau, ở đó $A^0 = E$). Nếu xét đa thức

$$p_m(x) = \sum_{k=0}^{m} a_k x^k$$

thì khi thay biến x bởi ma trận $A \in M_n(\mathbb{K})$ ta được

$$p_m(A) = \sum_{k=0}^{m} a_k A^k.$$

Rõ ràng $p_m(A) \in M_n(\mathbb{K})$.

Các tính chất của phép nhân hai ma trận cho dưới trong mệnh đề sau đây:

Mệnh đề 1.2.3. *i) (Luật kết hợp) Giả sử A, B, C là các ma trận cỡ $m \times n$, $n \times p$ và $p \times r$ tương ứng, khi đó: $A(BC) = A(BC)$*
ii) (Luật phân phối phải) Giả sử A là ma trận cỡ $m \times n$, B, C là các ma trận cỡ $n \times p$, khi đó ta có $A(B + C) = AB + AC$, tương tự ta cũng có luật phân phối bên trái.
iii) Giả sử A là ma trận cỡ $m \times n$, B là ma trận cỡ $n \times p$ và $c \in \mathbb{K}$, khi đó $c(AB) = (cA)B = A(cB)$.

Dễ thấy phép nhân ma trận nói chung không giao hoán. Từ các tính chất trên có thể thấy tập các ma trận vuông cấp n với phép cộng và nhân ma trận lập thành một *vành*.

Ví dụ 25. Một công ty Z có ba cửa hàng I, II, III cùng bán 4 loại mặt hàng: Tivi, điều hòa, tủ lạnh, lò vi sóng với giá bán (triệu đồng/chiếc) lần lượt cho bởi ma trận cột A, lượng hàng bán được trong ngày của các cửa hàng lần lượt được cho bởi các hàng của ma trận B

$$A = \begin{pmatrix} 5 \\ 9 \\ 8 \\ 3 \end{pmatrix}, B = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 1 & 2 & 1 & 2 \\ 3 & 2 & 0 & 1 \end{pmatrix}$$

khi đó tích

$$BA = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 1 & 2 & 1 & 2 \\ 3 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 9 \\ 8 \\ 3 \end{pmatrix} = \begin{pmatrix} 28 \\ 37 \\ 36 \end{pmatrix}$$

cho số tiền các cửa hàng I, II, III bán được trong ngày.

Giả sử $A = (a_{ij})_{m \times n} \in M_{m \times n}(\mathbb{K})$. *Chuyển vị của ma trận A* ký hiệu là A^T thu được bằng cách viết lại các hàng thành các cột:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, A^T = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \vdots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$$

Mệnh đề 1.2.4. *i)* $(A + B)^T = A^T + B^T, \forall A, B \in M_{m \times n}(\mathbb{K})$
ii) $(cA)^T = cA^T, \forall A \in M_{m \times n}(\mathbb{K}), c \in \mathbb{K}$
iii) $(AB)^T = B^T A^T, \forall A \in M_{m \times n}(\mathbb{K}), B \in M_{n \times p}(\mathbb{K})$
iv) $(A^T)^T = A, \forall A \in M_{m \times n}(\mathbb{K})$.

Dễ dàng chứng minh các mệnh đề này từ định nghĩa của phép chuyển vị.

Trong thực tế ta thường gặp một số ma trận vuông có tính chất khá đẹp, có thể liệt kê vài loại sau đây.

Giả sử $A \in M_n(\mathbb{K})$, ma trận A gọi là *đối xứng* nếu $A^T = A$ và *phản đối xứng* nếu $A^T = -A$.

Ma trận A gọi là *tam giác trên (dưới)* nếu các phần tử phía dưới (phía trên) đường chéo chính bằng 0. Viết theo ký hiệu toán học, ma trận tam giác trên là ma trận

$$A = (a_{\mathrm{ij}})_{n\times n} : a_{\mathrm{ij}} = 0, \forall i > j$$

Đặc biệt nếu các phần tử ngoài đường chéo chính đều bằng 0 thì ta có *ma trận đường chéo*, ma trận E là trường hợp đặc biệt của loại này.

Ma trận A được gọi là *khả nghịch* nếu tồn tại ma trận B sao cho $AB = BA = E$, khi đó B được ký hiệu là A^{-1} (về ma trận khả nghịch ta sẽ nghiên cứu ở phần sau). Ma trận A gọi là *trực giao* nếu $A^{-1} = A^T$.

Chú ý 3. *Bằng cách chia ma trận thành các khối thích hợp, ta có thể lập một ma trận gồm các khối, trên đó ta cũng có thể trang bị các phép toán cộng và nhân như các ma trận bình thường. Ma trận kiểu như vậy gọi là ma trận khối.*

1.3 Định thức

1.3.1 Định thức và tính chất

Giả sử có ma trận vuông $A \in M_n(\mathbb{K})$. *Định thức cấp n* của ma trận A, ký hiệu là $\det(A)$ (hay $|A|$) đó là một số và được định nghĩa truy hồi như sau:

Định thức ma trận cấp $n = 1$: $A = (a_{11})$ thì $\det(A) = a_{11}$.

Định thức ma trận cấp $n = 2$:

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

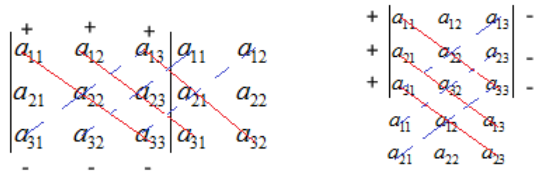
Tương tự ta có thể tính định thức cấp $n = 3$

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} +$$
$$+a_{13}\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Khai triển ra ta được

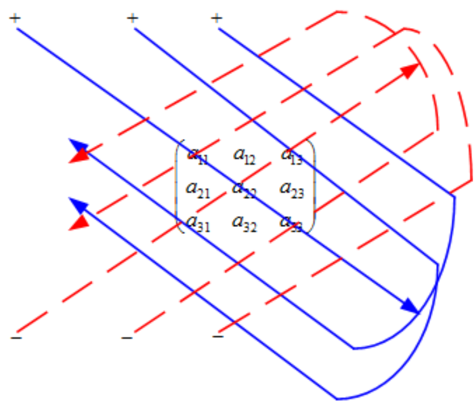
$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Có thể khái quát cách tính định thức cấp 3 bởi sơ đồ như Hình 1.11.



Hình 1.11: Sơ đồ tính định thức cấp 3

hoặc bởi sơ đồ như hình Hình 1.12.



Hình 1.12: Sơ đồ tính định thức cấp 3

Giả sử ta tính được định thức cấp $n-1$, ta đưa ra định nghĩa định thức cấp $n \geq 2$. Trước hết ta ký hiệu $\overline{A_i^j}$ là ma trận vuông cấp $n-1$ nhận được từ ma trận $A = (a_{ij})_{n\times n}$ cấp n bằng cách bỏ đi hàng i cột j, đặt $A_{ij} = (-1)^{i+j}\det\left(\overline{A_i^j}\right)$ gọi là *phần phụ đại số* (còn gọi là *cofactor*) của a_{ij}, định thức (cấp $n-1$) $\det\left(\overline{A_i^j}\right)$ gọi là *minor* ứng với a_{ij}. Ta có định nghĩa định thức cấp n sau.

Định nghĩa 22. $\det(A) = \sum\limits_{k=1}^{n} a_{1k}A_{1k}$

Công thức định nghĩa định thức còn gọi là công thức khai triển định thức theo hàng 1, tổng quát có thể chứng minh được định lý sau:

Định lý 1.3.1. *(Xem [5], [16])*
i) $\det(A) = \sum\limits_{k=1}^{n} a_{ik}A_{ik}, \forall i = \overline{1;n}$
ii) $\sum\limits_{k=1}^{n} a_{ik}A_{jk} = 0, \forall i \neq j$

Công thức i) được gọi là công thức khai triển định thức theo hàng i bất kỳ.

Một ma trận vuông gọi là *không suy biến* nếu định thức của nó khác 0, ngược lại gọi là *suy biến.*

Hệ quả sau đây thu được trực tiếp từ định nghĩa và định lý trên:

Hệ quả 1.3.1. *i) Nếu một hàng (cột) của định thức gồm toàn các phần tử bằng* 0 *thì định thức bằng* 0.
ii) Với ma trận A bất kỳ và hằng số $\alpha \in \mathbb{K}$ *thì*

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ \alpha a_{k1} & \cdots & \alpha a_{kn} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

Từ đây ta dễ dàng có được $\det(\alpha A) = \alpha^n \det(A)$.
iii) Định thức ma trận tam giác (trên hoặc dưới) bằng tích các phần tử trên đường chéo chính.

Định lý 1.3.2. *(Xem [16]) Với ma trận vuông cấp n bất kỳ ta có*

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum \pm a_{1i_1}a_{2i_2}...a_{ni_n}$$

ở đó tổng chạy trên tất cả các hoán vị $(i_1, i_2, ..., i_n)$ *của các số* $1, 2, ..., n$. *Dấu cộng hay trừ của các số hạng phụ thuộc vào tính chẵn lẻ của hoán vị* $(i_1, i_2, ..., i_n)$, *rõ ràng số các số hạng trong dấu tổng là* $n!$.

Chú ý 4. *Có những quan điểm khác nhau khi tiếp cận khái niệm định thức, cách tiếp cận trên dễ dàng hơn đối với sinh viên kỹ thuật tuy nhiên có hạn chế là rất khó chứng minh chặt chẽ về mặt lý thuyết. Để hiểu sâu phần này sinh viên cần đọc kỹ bài định thức trong sách giáo khoa ([3]).*

Tính chất của định thức:

Mệnh đề 1.3.1. *Đổi chỗ 2 hàng của định thức làm định thức đổi dấu.*

Hệ quả 1.3.2. *Định thức có hai hàng bằng nhau (hoặc tỉ lệ nhau) thì định thức bằng* 0.

Mệnh đề 1.3.2. *Đối với một ma trận vuông A bất kỳ thì*

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ \alpha a_{k1} + \beta b_{k1} & \cdots & \alpha a_{kn} + \beta b_{kn} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ \alpha a_{k1} & \cdots & \alpha a_{kn} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ \beta b_{k1} & \cdots & \beta b_{kn} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

Dễ dàng chứng minh tính chất này từ công thức khai triển định thức theo hàng k.

Hệ quả 1.3.3. *Lấy một hàng của định thức nhân với một số $\alpha \neq 0$ rồi cộng vào một hàng (cột) khác thì định thức không đổi.*

Mệnh đề 1.3.3. *Đối với một ma trận vuông A bất kỳ thì* $\det(A^T) = \det(A)$

Từ mệnh đề trên ta thấy rằng tất cả các tính chất của định thức đúng với hàng thì cũng đúng với cột. Hơn nữa, trong công thức khai triển định thức theo hàng bất kỳ, cố định chỉ số cột, cho chỉ số hàng chạy ta vẫn được định thức (khai triển theo cột).

1.3.2 Cách tính định thức:

Các phép biến đổi dạng: 1) nhân một hàng (cột) của định thức với hằng số khác không; 2) nhân một hàng (cột) của định thức với hằng số khác không rồi cộng vào hàng (cột khác); 3) đổi chỗ hai hàng (cột) gọi là các *phép biến đổi sơ cấp*. Có thể sử dụng các phép biến đổi sơ cấp để tính định thức bằng định nghĩa của ma trận dạng tam giác.

Ví dụ 26. Biến đổi sơ cấp hàng tính định thức

$$\begin{vmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{vmatrix} = -1$$

Ngoài cách tính định thức bằng biến đổi sơ cấp hoặc cách khai triển theo hàng i bất kỳ theo định nghĩa, ta có thể tính định thức dựa vào định lý Laplace dưới đây. Trước hết, ký hiệu $\overline{A^{j_1 j_2 \ldots j_k}_{i_1 i_2 \ldots i_k}}$ là ma trận nhận được bằng cách gạch bỏ các hàng có thứ tự $i_1, i_2, \ldots, i_k$ và các cột $j_1, j_2, \ldots, j_k$. Còn ma trận $A^{j_1 j_2 \ldots j_k}_{i_1 i_2 \ldots i_k}$ là ma trận mà các phần tử của nó nằm trên giao của các hàng $i_1, i_2, \ldots, i_k$ và các cột $j_1, j_2, \ldots, j_k$. Khi đó $(-1)^{i_1+\ldots+i_k+j_1+\ldots+j_k} \det\left(\overline{A^{j_1 j_2 \ldots j_k}_{i_1 i_2 \ldots i_k}}\right)$ gọi là phần phụ đại số của $A^{j_1 j_2 \ldots j_k}_{i_1 i_2 \ldots i_k}$.

Định lý 1.3.3. *(Laplace)*

$$\det(A) = \sum_{1 \le j_1 < \ldots < j_k \le n} (-1)^{i_1+\ldots+i_k+j_1+\ldots+j_k} \det\left(A^{j_1 j_2 \ldots j_k}_{i_1 i_2 \ldots i_k}\right) \det\left(\overline{A^{j_1 j_2 \ldots j_k}_{i_1 i_2 \ldots i_k}}\right)$$

ở đó tổng được lấy theo tất cả các $(j_1, \ldots, j_k)$ *sao cho* $1 \le j_1 < \ldots. < j_k \le n$.

Định lý trên được chứng minh trong [3].

Ví dụ 27. Tính định thức của

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 1 & 3 \\ 4 & 2 & 1 \end{pmatrix}$$

Khai triển theo hàng 1 và 2, ta có

$$|A^{12}_{12}| = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix}, |A^{13}_{12}| = \begin{vmatrix} 1 & 3 \\ -2 & 3 \end{vmatrix}, |A^{23}_{12}| = \begin{vmatrix} 1 & 3 \\ -2 & 3 \end{vmatrix}$$

như vậy

$$\det(A) = (-1)^{1+2+1+2}.1.|A^{12}_{12}| + (-1)^{1+2+1+3}.2.|A^{13}_{12}| + (-1)^{1+2+2+3}.4.|A^{23}_{12}| = -1$$

Hệ quả 1.3.4.

$$\det(AB) = \det(A)\det(B)$$

1.4 Hạng ma trận, ma trận nghịch đảo

1.4.1 Hạng của ma trận

Giả sử ma trận $A \in M_{m\times n}(\mathbb{K})$.

Định nghĩa 23. *Nếu tồn tại số r sao cho $0 \le r \le min\{m,n\}$ mà*
i) Tồn tại định thức con cấp $r \neq 0$ (định thức của ma trận lập từ các phần tử nằm trên giao của r hàng r cột nào đó của A)
ii) Mọi định thức con cấp $r+1$ (nếu có) bằng 0, số r (duy nhất!) như thế gọi là hạng của ma trận A và ký hiệu là $rank(A)$.

Chú ý 5. *Rõ ràng $r = 0$ khi và chỉ khi $a_{ij} = 0$, và tất nhiên r lớn nhất là bằng $min\{m,n\}$. Các phép biến đổi sơ cấp không làm thay đổi hạng của ma trận vì chúng không làm thay đổi tính chất bằng 0 hay khác 0 của định thức.*

Ví dụ 28. Trong lý thuyết điều khiển toán học, hệ phương trình vi phân

$$\dot{x}(t) = \mathrm{Ax}(t) + Bu(t)$$

mô tả một hệ thống điều khiển nào đó, ở đây $x(t) = (x_1(t), ..., x_n(t)) \in \mathbb{R}^n$ là vector trạng thái ($\dot{x}(t)$ là chỉ đạo hàm của vector $x(t)$), $u(t) = (u_1(t), ..., u_m(t)) \in \mathbb{R}^m$ gọi là vector điều khiển, A, B là các ma trận hằng cỡ $n \times n$ và $n \times m$ tương ứng (hệ tuyến tính dừng). Cặp trạng thái (x^0, x^1) gọi là *điều khiển được* sau thời gian $t_1 > 0$ nếu tồn tại điều khiển $u(t)$ sao cho nghiệm $x(t, x_0, u)$ của hệ thỏa mãn

$$x(0, x^0, u) = x^0, x(t_1, x^0, u) = x^1$$

và gọi là *điều khiển hoàn toàn* nếu bất kỳ hai trạng thái x^0, x^1 sẽ tìm được thời gian $t_1 > 0$ sao cho cặp (x^0, x^1) điều khiển được sau thời gian t_1. Kalman chứng minh được rằng hệ điều khiển trên là điều khiển hoàn toàn khi và chỉ khi

$$rank\left(B, AB, ..., A^{n-1}B\right) = n$$

điều kiện này gọi là tiêu chuẩn hạng Kalman cho tính điều khiển được của hệ tuyến tính dừng. Ví dụ, xét tính điều khiển được của hệ

$$\begin{cases} \dot{x}_1 = x_2 + u \\ \dot{x}_2 = x_1 + 2x_2 + 2u \end{cases}$$

Ta có

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Dễ thấy rằng

$$rank\,(B, AB) = rank \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} = 2 = n$$

nên hệ đã cho là điều khiển được hoàn toàn.

Định nghĩa 24. *Ma trận $A \in M_{m\times n}(\mathbb{K})$ gọi là hình thang nếu*
i) Các hàng khác 0 (tức là hàng ít nhất một phần tử khác 0) nằm trên các hàng bằng 0
ii) Với hai hàng khác 0 phần tử khác 0 đầu tiên của hàng dưới luôn nằm bên phải cột chứa phần tử khác 0 đầu tiên ở hàng trên.

Ví dụ 29.

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 2 & 4 \end{pmatrix}, B = \begin{pmatrix} 1 & -1 & 3 & 2 & 5 \\ 0 & 0 & 2 & 3 & 5 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Đối với một ma trận hình thang việc tìm hạng của ma trận rõ ràng đơn giản hơn (hạng chính là số hàng khác 0), ví dụ ở trên có thể thấy $rank(A) = 3, rank(B) = 3$. Từ nhận xét phía trên ta có thể đưa ra cách tìm hạng ma trận bằng phép biến đổi sơ cấp. Trước hết ta nhắc lại các phép biến đổi sơ cấp (cho ma trận):

a) Đổi chỗ hai hàng (cột) $h_i \leftrightarrow h_j (c_i \leftrightarrow c_j)$.

b) Nhân số $\alpha \neq 0$ với hàng (cột)i: $\alpha h_i \to h_i (\alpha c_i \to c_i)$.

c) Nhân số $\alpha \neq 0$ với hàng (cột) i rồi cộng vào hàng (cột) j: $\alpha h_i + h_j \to h_j (\alpha c_i + c_j \to c_j)$.

Như vậy ta có thể sử dụng các phép biến đổi sơ cấp đưa ma trận về dạng hình thang rồi kết luận hạng của ma trận. Phương pháp này thường được gọi là phép khử Gauss.

Ví dụ 30. Tìm hạng ma trận bằng biến đổi sơ cấp

$$\begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix} \xrightarrow[\substack{h_1 - h_2 \to h_2 \\ h_1 - h_3 \to h_3}]{} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \xrightarrow[c_2 \leftrightarrow c_3]{} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

khi đó hạng ma trận bằng 3 (bước biến đổi sơ cấp cuối cùng thực ra không cần thiết).

Không khó để chứng minh các tính chất sau của hạng ma trận.

Mệnh đề 1.4.1. *i)Giả sử tích AB xác định, khi đó $rank(AB) \leq min\{rank(A), rank(B)\}$*
ii) Giả sử B là ma trận vuông và tích AB xác định. Nếu $\det(B) \neq 0$ thì $rank(AB) = rank(A)$.

1.4.2 Điều kiện tồn tại ma trận nghịch đảo

Giả sử $A = (a_{ij})_{n\times n} \in M_n(\mathbb{K})$, E là ma trận đơn vị cấp n. Nếu tồn tại ma trận $B \in M_n(\mathbb{K})$ sao cho $AB = BA = E$ thì ma trận A gọi là khả nghịch. Ma trận B xác định như trên là duy nhất vì nếu tồn tại $B_i \in M_n(\mathbb{K})$ sao cho $AB_1 = B_1A = E$ thì

$$B = BE = B(AB_1) = (BA)B_1 = EB_1 = B_1$$

Ma trận B được gọi là ma trận nghịch đảo của A và ký hiệu là A^{-1}. Ký hiệu $GL_n(\mathbb{K})$ là tập các ma trận vuông cấp n khả nghịch, dễ dàng kiểm tra tập này với phép nhân ma trận lập thành một nhóm và gọi là *nhóm tuyến tính tổng quát cấp n.*

Định lý 1.4.1. *Giả sử $A \in M_n(\mathbb{K})$. A khả nghịch khi và chỉ khi* $\det(A) \neq 0$.

Chứng minh: *Điều kiện cần:* Giả sử A khả nghịch, khi đó tồn tại ma trận $B \in M_n(\mathbb{K})$ sao cho $AB = BA = E$. Từ đây $\det(AB) = 1$ hay $\det(A)\det(B) = 1$, do đó $\det(A) \neq 0$.

Điều kiện đủ: Giả sử $\det(A) = d \neq 0$. Xét

$$B = \begin{pmatrix} \frac{A_{11}}{d} & \dots & \frac{A_{n1}}{d} \\ \vdots & \vdots & \vdots \\ \frac{A_{1n}}{d} & \dots & \frac{A_{nn}}{d} \end{pmatrix}$$

ở đó A_{ik} là các phần phụ đại số của a_{ik} trong ma trận A. Khi đó tích $AB = (c_{ij})_{n\times n}$, ở đó

$$c_{\text{ij}} = \frac{1}{d}\sum_{k=1}^{n} a_{ik}A_{jk} = \frac{1}{d}\begin{cases} \det(A), i = j \\ 0, i \neq j \end{cases} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$$

tức là $AB = E$ hay A khả nghịch. ▶

Chú ý 6. *Từ chứng minh trên suy ra rằng* $A^{-1} = \frac{1}{d}\overline{A}^T$, *trong đó* $\overline{A}^T$ *là ma trận các phần phụ đại số của* A. *Và cũng dễ dàng nhận được* $\det(A^{-1}) = \frac{1}{\det(A)}$.

Không khó để chứng minh các tính chất sau đây của ma trận khả nghịch:

Mệnh đề 1.4.2. *i) Nếu* A, B *là các ma trận vuông cấp* n *khả nghịch thì* $(AB)^{-1} = B^{-1}A^{-1}$.
ii) Nếu A *là ma trận khả nghịch thì* $(A^{-1})^{-1} = A$.

Mệnh đề 1.4.3. *(Đọc thêm)*
i) Nếu $A \in M_n(\mathbb{K})$ *là ma trận không suy biến và* c, d *là các ma trận cỡ* $n \times 1$ *(ma trận cột) sao cho* $1 + d^TA^{-1}c \neq 0$ *thì tổng* $A + cd^T$ *không suy biến và*

$$\left(A + cd^T\right)^{-1} = A^{-1} - \frac{A^{-1}cd^TA^{-1}}{1 + d^TA^{-1}c}$$

ii) Tổng quát ta có Công thức Sherman-Morrison, nếu C, D *là các ma trận cỡ* $n \times k$ *sao cho* $(1 + D^TA^{-1}C)^{-1}$ *tồn tại thì*

$$\left(A + CD^T\right)^{-1} = A^{-1} - A^{-1}C\left(E + D^TA^{-1}C\right)^{-1}D^TA^{-1}$$

Giả sử $A = (a_{ij})_{n\times n} \in M_n(\mathbb{K})$, ma trận $(A - \lambda E)$ gọi là ma trận đặc trưng của A còn $\det(A - \lambda E)$ là một đa thức bậc n của λ trên $\mathbb{K}$ gọi là đa thức đặc trưng:

$$P_n(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}\alpha_{n-1}\lambda^{n-1} + ... - \alpha_1\lambda + \alpha_0$$

Không khó để thấy $\alpha_0 = \det(A)$ còn $\alpha_{n-1} = \sum\limits_{i=1}^{n} a_{ii}$ gọi là *vết* của ma trận A ký hiệu là $Trace(A)$.

Định lý 1.4.2. *(Định lý Hamilton-Cayley) Mọi ma trận vuông* A *đều là nghiệm của đa thức đặc trưng của nó, tức là* $P_n(A) = O$.

Chứng minh của định lý có thể tìm thấy trong [3], [14] (*Đọc thêm*).

1.4.3 Tìm ma trận nghịch đảo bằng biến đổi sơ cấp

Đối với ma trận vuông A bất kỳ, bằng phép biến đổi sơ cấp có thể đưa nó về ma trận dạng đường chéo. Bản chất của các phép biến đổi sơ cấp này là nhân ma trận A với một ma trận không suy biến. Ví dụ khi đổi chỗ hai hàng i và j của ma trận A thực chất là nhân bên trái ma trận A với $C_{ij} = (c_{lh})_{n\times n}$ ở đó

$$\begin{cases} c_{\text{ij}} = c_{ji} = 1 \\ c_{ii} = c_{jj} = 0, c_{kk} = 1, k \neq i, k \neq j \\ c_{lh} = 0, l \neq h, (l,h) \neq (i,j), (l,h) \neq (j,i) \end{cases}$$

Dễ thấy $\det(C_{ij}) = -1$. Nếu đổi chỗ hai cột thì nhân ma trận này ở bên phải của A.

Phép biến đổi sơ cấp nhân hàng thứ i với $\lambda \neq 0$ là nhân bên trái ma trận A với ma trận $C_{ii} = (c_{lh})_{n\times n}$ ở đó

$$\begin{cases} c_{ii} = \lambda \\ c_{kk} = 1, k \neq i \\ c_{lh} = 0, l \neq h \end{cases}$$

không khó để thấy $\det(C_{ii}) = \lambda$. Nếu nhân bên phải với ma trận này là nhân cột thứ i với λ.

Tương tự phép biến đổi sơ cấp nhân hàng thứ i với λ rồi cộng vào hàng thứ j thực chất là nhân bên trái ma trận A với ma trận $C_{ij} = (c_{lh})_{n\times n}$ ở đó

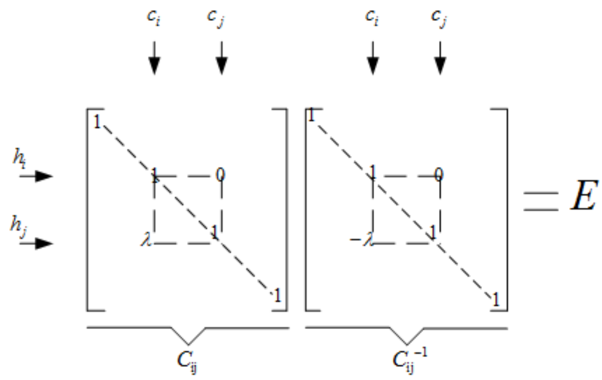
$$\begin{cases} c_{kk} = 1, \forall k \\ c_{ji} = \lambda \\ c_{lh} = 0, \forall l \neq h, (l,h) \neq (i,j) \end{cases}$$

dễ thấy $\det(C_{ij}) = 1$. Nếu nhân bên phải với ma trận C_{ij} này chính là nhân cột thứ j với λ rồi cộng vào cột thứ i. Ta nhận thấy rằng nếu ma trận C_{ij} có dạng trên thì $C_{ij}^{-1} = (c_{lh})_{n\times n}$ ở đó

$$\begin{cases} c_{kk} = 1, \forall k \\ c_{ji} = -\lambda \\ c_{lh} = 0, \forall l \neq h, (l,h) \neq (i,j) \end{cases}$$

Điều này dễ thấy qua Hình 1.13.

Nếu A là ma trận khả nghịch thì chỉ bằng phép biến đổi sơ cấp hàng có thể đưa ma trận A về ma trận đơn vị E. Tức là tồn tại ma trận khả nghịch B sao cho $BA = E$, ở đó B là tích các ma trận có các dạng trên. Nhưng từ đó ta có ngay $BE = A^{-1}$. Tức là các phép biến đổi sơ cấp hàng đưa A về



Hình 1.13: Ma trận C_{ij}^{-1}

E thì ngược lại đưa E thành A^{-1}. Như vậy ta có thể sử dụng biến đổi sơ cấp hàng tìm ma trận nghịch đảo theo sơ đồ (thường được gọi là phép khử Gauss-Jordan)

$$(A|E) \to \left(E|A^{-1}\right)$$

hoặc

$$(E|A) \to \left(A^{-1}|E\right).$$

Tương tự chúng ta có thể sử dụng biến đổi sơ cấp cột tìm ma trận nghịch đảo theo sơ đồ sau

$$\left(\frac{A}{E}\right) \to \left(\frac{E}{A^{-1}}\right)$$

Ví dụ 31.

$$(A|E) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\substack{-h_1+h_2\to h_2 \\ -h_1+h_3\to h_3}]{} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[\substack{h_3+h_2\to h_2 \\ h_3+h_1\to h_1}]{} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & -2 & 1 & 1 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[\substack{2h_2+h_1\to h_1 \\ -h_2\to h_2 \\ -h_3\to h_3}]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & 2 & 3 \\ 0 & 1 & 0 & 2 & -1 & -1 \\ 0 & 0 & 1 & 1 & 0 & -1 \end{array}\right)$$

từ đó ta có

$$A^{-1} = \begin{pmatrix} -4 & 2 & 3 \\ 2 & -1 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

1.4.4 Phân tích LU và LUP

Ma trận càng đơn giản thì làm việc với nó càng dễ dàng. Ma trận tam giác dưới và trên là những ma trận đơn giản như vậy. Tiếp theo đây ta phân tích một ma trận khả nghịch $A \in GL_n(\mathbb{K})$ thành tích của hai ma trận tam giác dưới L và trên U, cả L, U đều khả nghịch. Người ta gọi phân tích đó là *phân tích LU* của A.

Phân tích về các ma trận tam giác kiểu như vậy có ứng dụng lớn trong giải quyết các bài toán giải hệ phương trình cũng như tính định thức. Người ta chứng minh được mọi ma trận cấp n khả nghịch thỏa mãn phân tích LU khi và chỉ khi các định thức con chính khác 0, phân tích sẽ duy nhất nếu thêm điều kiện các phần tử trên đường chéo của ma trận L (hoặc U) bằng 1.

Để tìm phân tích này ta làm như sau:

Bước 1: Biến đổi sơ cấp hàng ma trận A thành ma trận tam giác trên U (không có đổi chỗ các hàng). Như đã biết, bản chất của quá trình này là nhân A với dãy ma trận không suy biến dạng tam giác dưới, giả sử dãy đó là $C = C_k ... C_2 C_1$, ta có

$$U = C_k ... C_2 C_1 A = CA.$$

Bước 2: Do $LU = A$ nên tìm được L bằng công thức

$$L = {C_1}^{-1} ... {C_{k-1}}^{-1} {C_k}^{-1} = C^{-1}.$$

Ví dụ 32. Phân tích LU ma trận

$$A = \begin{pmatrix} 6 & 18 & 3 \\ 2 & 12 & 1 \\ 4 & 15 & 3 \end{pmatrix}.$$

Ta biến đổi sơ cấp A về U như sau

$$A = \begin{pmatrix} 6 & 18 & 3 \\ 2 & 12 & 1 \\ 4 & 15 & 3 \end{pmatrix} \xrightarrow{-h_1/3 + h_2 \to h_2} \begin{pmatrix} 6 & 18 & 3 \\ 0 & 6 & 0 \\ 4 & 15 & 3 \end{pmatrix} \xrightarrow{-2h_1/3 + h_3 \to h_3}$$

$$\begin{pmatrix} 6 & 18 & 3 \\ 0 & 6 & 0 \\ 0 & 3 & 1 \end{pmatrix} \xrightarrow{-h_2/2 + h_3 \to h_3} \begin{pmatrix} 6 & 18 & 3 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{pmatrix} = U.$$

Ma trận C là

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{2}{3} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{3} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Từ đó

$$L = C^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{2}{3} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & \frac{1}{2} & 1 \end{pmatrix}.$$

Vậy

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 6 & 18 & 3 \\ 0 & 6 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Tổng quát hơn với một ma trận vuông khả nghịch A ta có phân tích LUP, đó là phân tích dạng

$$PA = LU$$

ở đó L, U vẫn là các ma trận tam giác như trên, P là ma trận nhận được trong biến đổi sơ cấp hàng (ở đây là đổi chỗ các hàng, một số tài liệu gọi là *ma trận hoán vị*) của A.

Ví dụ 33. Xét ma trận

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 3 & 4 \\ -1 & -1 & 1 \end{pmatrix}.$$

Rõ ràng, ngay từ đầu phân tích LU ma trận này là không thể. Trước hết ta đổi chỗ hàng đầu và hàng cuối của ma trận A cho nhau để được phần tử hàng đầu tiên, cột đầu tiên khác 0, tức là ta chọn ma trận P với

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

từ đó

$$PA = \begin{pmatrix} -1 & -1 & 2 \\ 2 & 3 & 4 \\ 0 & 1 & 2 \end{pmatrix}.$$

Lập lại quy trình như trong ví dụ trên đối với ma trận PA ta có

$$C_2 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, C_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Và ta có phân tích $PA = LU$, ở đó

$$U = \begin{pmatrix} -1 & -1 & 2 \\ 0 & 1 & 8 \\ 0 & 0 & -6 \end{pmatrix}, L = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Một ma trận vuông A cấp n suy biến có $rank(A) < n$ cũng có thể có phân tích LUP, tuy nhiên ta không xét ở đây.

Lý luận tương tự khi tìm ma trận nghịch đảo, ta có thuật toán sau đây rất hữu ích trong việc tìm phân tích LU (LUP) của ma trận.

Bước 1: Viết ma trận khối dạng $(E|A)$ (hoặc $(E|PA)$).

Bước 2: Biến đổi sơ cấp hàng (không có phép đổi chỗ) đưa $(E|A)$ (hoặc $(E|PA)$) về dạng $(L^{-1}|U)$.

Chú ý ta có thể cải tiến thuật toán trên như sau: khi biến đổi sơ cấp hàng nếu phần tử khử là a_{jj} với $j > 1$ (phần tử này luôn khác không) thì ma trận bên trái giữ nguyên các cột từ $1, 2, ..., j-1$. Ma trận cuối cùng không còn là L^{-1}. Tuy nhiên khi đó ma trận L nhận được từ ma trận này bằng cách giữ nguyên các a_{ij} với $j \geq i$, còn với $j < i$ thì lấy $-a_{ij}$ (điều này thu được do tính chất phép biến đổi sơ cấp).

Ví dụ 34. Trở lại ví dụ phân tích LU ma trận

$$A = \begin{pmatrix} 6 & 18 & 3 \\ 2 & 12 & 1 \\ 4 & 15 & 3 \end{pmatrix}.$$

Có thể xem các thao tác thực hiện phân tích LU như Hình 1.14.

$$(E|A) = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6 & 18 & 3 \\ 0 & 1 & 0 & 2 & 12 & 1 \\ 0 & 0 & 1 & 4 & 15 & 3 \end{array}\right) \xrightarrow[\frac{-2h_1}{3}+h_3\to h_3]{\frac{-h_1}{3}+h_2\to h_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6 & 18 & 3 \\ \boxed{-1/3} & 1 & 0 & 0 & 6 & 0 \\ \boxed{-2/3} & 0 & 1 & 0 & 3 & 1 \end{array}\right)$$

$$\xrightarrow{\frac{-h_2}{2}+h_3\to h_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6 & 18 & 3 \\ -1/3 & 1 & 0 & 0 & 6 & 0 \\ -2/3 & \boxed{-1/2} & 1 & 0 & 0 & 1 \end{array}\right) \Rightarrow L = \begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 2/3 & 1/2 & 1 \end{pmatrix}$$

Hình 1.14: Phân tích LU

1.5 Hệ phương trình tuyến tính

1.5.1 Các định nghĩa và ví dụ

Định nghĩa 25. *Ta gọi hệ phương trình tuyến tính tổng quát m phương trình, n ẩn trên trường $\mathbb{K}$ $(m, n \in \mathbb{N})$ là hệ phương trình:*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \quad\quad\quad\quad \ldots. \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{1.1}$$

trong đó $a_{ij}, b_i (i = \overline{1; m}, j = \overline{1; n})$ cho trước trên $\mathbb{K}$, x_j là các ẩn cần tìm.

Định nghĩa 26. *Nghiệm của hệ là bộ $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$ mà khi thay $x_j = \lambda_j (j = \overline{1; n})$ ta được các đẳng thức trên $\mathbb{K}$. Nếu hệ có nghiệm ta nói hệ tương thích, ngược lại ta nói hệ không tương thích.*

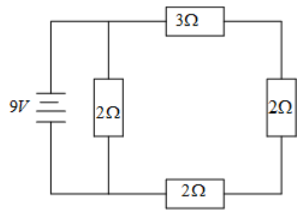
Giải hệ phương trình là đi tìm tập hợp nghiệm của nó. Khi các số hạng ở vế phải bằng 0 ta gọi hệ là *hệ phương trình thuần nhất*. Đôi khi ta viết (1.1) gọn lại dưới dạng

$$\sum_{j=1}^{n} a_{\text{ij}}x_j = b_i, i = \overline{1; m} \tag{1.2}$$

Nếu đặt $A = (a_{ij})_{m\times n} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ thì (1.1) có thể viết lại ở dạng

$$Ax = b \tag{1.3}$$

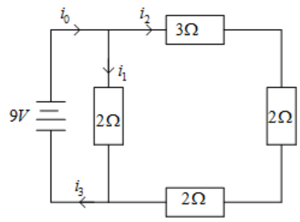
và gọi là dạng ma trận của hệ phương trình tuyến tính.



Hình 1.15: Mạch điện song song

Ví dụ 35. Xét mạch điện mắc song song được cho trong Hình 1.15

Ta sẽ đi tính cường độ dòng điện qua mỗi điện trở. Theo định luật Kirchoff tổng các dòng chạy vào một nút bằng tổng các dòng chạy ra khỏi nút đó, mặt khác chú ý rằng cường độ dòng trên đoạn mạch mắc nối tiếp là như nhau. Giả sử các cường độ dòng điện được phân bố như trong Hình 1.16. theo định luật Ohm cho đoạn mạch mắc song song, hiệu điện thế đoạn



Hình 1.16: Phân bố cường độ dòng điện

mạch có điện trở bên trái bằng hiệu điện thế đoạn mạch mắc song song có ba điện trở còn lại và bằng $9V$. Áp dụng luật Kirchoff ta có hệ sau đây để xác định các cường độ dòng điện

$$\begin{cases} i_0 - i_1 - i_2 = 0 \\ i_1 + i_2 - i_3 = 0 \\ 2i_1 = 9 \\ 7i_2 = 9 \end{cases}$$

giải hệ này ta có ngay $i_0 = i_3 = \frac{81}{14}(A), i_1 = \frac{9}{2}(A), i_2 = \frac{9}{7}(A)$.

1.5.2 Hệ Cramer

Khi $m = n$ (1.1) được gọi là *hệ Cramer* nếu $\det(A) \neq 0$. Ký hiệu A_{x_j} là ma trận A thay cột thứ j bởi cột hệ số tự do b.

Định lý 1.5.1. *(công thức Cramer) Hệ Cramer có nghiệm duy nhất được cho bởi công thức*

$$x_j = \frac{\det\left(A_{x_j}\right)}{\det(A)}, j = \overline{1;n}.$$

Chứng minh: Giả sử $\det(A) \neq 0$, hệ phương trình viết dưới dạng (1.5) là $Ax = b$. Vì $\det(A) \neq 0$ nên tồn tại ma trận A^{-1}, nhân hai vế của hệ với A^{-1} ta được $x = A^{-1}b$. Mặt khác ta có

$$A^{-1} = \frac{1}{\det(A)}\overline{A}^T = \frac{1}{\det(A)}\begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & \vdots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix}$$

trong đó $\overline{A}$ là ma trận các phần phụ đại số của ma trận A. Từ đây rút ra

$$x_j = \frac{1}{\det(A)}\left(A_{1j}b_1 + A_{2j}b_2 + ... + A_{nj}b_n\right) = \frac{1}{\det(A)}\sum_{k=1}^{n} b_k A_{kj}.$$

Xét định thức

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & b_1 & a_{1j+1} & \cdots & a_{1_n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & b_n & a_{nj+1} & \cdots & a_{nn} \end{vmatrix}$$

phân tích định thức này theo cột j ta có ngay

$$D = \sum_{k=1}^{n} b_k A_{kj} = \det\left(A_{x_j}\right)$$

từ đó

$$x_j = \frac{\det\left(A_{x_j}\right)}{\det\left(A\right)}, j = \overline{1;n}$$

ta có điều phải chứng minh. ▶

Ví dụ 36. Xét hệ phương trình $Ax = b$, ở đó

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 3 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Dễ thấy $\det(A) = -1$, theo công thức Cramer

$$x_1 = \frac{\det\begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & 2 \\ 3 & 0 & 3 \end{pmatrix}}{\det\left(A\right)} = -3, x_2 = \frac{\det\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 2 & 3 & 3 \end{pmatrix}}{\det\left(A\right)} = -4$$

$$x_3 = \frac{\det\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 2 & 0 & 3 \end{pmatrix}}{\det(A)} = 3.$$

Một phương pháp khác để giải hệ Cramer là sử dụng phân tích LU. Xét hệ phương trình $Ax = b$. Nếu ma trận A đưa được về dạng $A = LU$, thì đầu tiên đặt $y = Ux$, ta giải hệ $Ly = b$ trước sau đó giải tiếp hệ $Ux = y$. Vì các ma trận L, U là tam giác nên việc giải này rất dễ dàng.

Trong trường hợp phải đổi chỗ các hàng của ma trận A thì ta có thể sử dụng phân tích LUP để giải hệ. Để ý rằng hệ $Ax = b$ tương đương với $PAx = Pb$ do ma trận P không suy biến. Việc giải hệ tiếp theo sẽ làm như với cách giải trong phân tích LU ở trên.

1.5.3 Điều kiện cần và đủ để hệ tổng quát có nghiệm

Xét hệ (1.1), ma trận A gắn thêm cột hệ số vào cột thứ $n+1$ gọi là ma trận mở rộng của hệ

$$\widetilde{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

Định lý 1.5.2. *(Cronecker-Capelli) i) Hệ (1.1) có nghiệm khi và chỉ khi $rank(A) = rank(\widetilde{A})$.*
ii) Nếu $rank(A) = rank(\widetilde{A}) = r$ thì hệ (1.1) tương đương với hệ r phương trình nằm trên r hàng có định thức con cấp r khác 0.

Hệ quả 1.5.1. *Hệ n phương trình tuyến tính thuần nhất n ẩn có nghiệm không tầm thường khi và chỉ khi định thức ma trận hệ bằng 0.*

Chứng minh định lý trên có thể xem trong [3]. Từ định lý này ta thấy nếu hệ (1.1) tương thích thì sẽ có r ẩn giải theo $n - r$ ẩn còn lại, r ẩn đó gọi là các ẩn *phụ thuộc*, $n - r$ ẩn còn lại gọi là *tự do*. Nhờ phép khử Gauss ta đưa ra các bước giải hệ phương trình tuyến tính tổng quát như sau:

Bước 1: Viết ma trận mở rộng của hệ $\widetilde{A} = (A|b)$, sử dụng các biến đổi sơ cấp hàng đưa ma trận này về dạng hình thang (chú ý rằng có thể sử dụng đổi vị trí các cột nhưng phải đánh số lại các ẩn).

Bước 2: Kiểm tra điều kiện $rank(A) = rank(\widetilde{A}) = r$, nếu không thỏa mãn thì kết luận hệ vô nghiệm, bài toán dừng lại. Nếu điều kiện này thỏa mãn hệ có r ẩn phụ thuộc vào $n - r$ ẩn tự do.

Bước 3: Kết luận nghiệm (giải r ẩn theo $n-r$ ẩn tự do hoặc vô nghiệm).

Ví dụ 37. Xét hệ dạng (1.3) trong đó

$$A = \begin{pmatrix} 1 & 2 & -1 & 3 & 0 \\ -1 & -2 & 2 & -2 & -1 \\ 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 2 & 2 & -1 \end{pmatrix}; b = \begin{pmatrix} 1 \\ 1 \\ 6 \\ 7 \end{pmatrix}.$$

Khi đó ma trận mở rộng của hệ là

$$\widetilde{A} = \left(\begin{array}{ccccc|c} 1 & 2 & -1 & 3 & 0 & 1 \\ -1 & -2 & 2 & -2 & -1 & 1 \\ 1 & 2 & 0 & 4 & 0 & 6 \\ 0 & 0 & 2 & 2 & -1 & 7 \end{array}\right).$$

Sử dụng biến đổi sơ cấp hàng đưa ma trận về dạng

$$A = \left(\begin{array}{ccccc|c} \underline{1} & 2 & -1 & 3 & 0 & 1 \\ 0 & 0 & \underline{1} & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & \underline{1} & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Dễ thấy điều kiện hạng thỏa mãn ($r = 3$) số ẩn tự do là 2, ở đây là 2 ẩn $x_2,\ x_4$, ta chọn $x_2 = s, x_4 = t$ thì nghiệm hệ có thể viết dưới dạng

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 5 \\ 0 \\ 3 \end{pmatrix} + s \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -4 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Nếu cho các ẩn tự do nhận các giá trị cụ thể (ví dụ bằng 0) ta thu được một *nghiệm riêng* của hệ. Về cấu trúc nghiệm của hệ tổng quát ta sẽ nghiên cứu kỹ hơn trong chương sau. Tuy nhiên từ ví dụ trên có thể thấy nếu cho cột hệ số tự do toàn bộ là 0 thì quá trình giải ta thu được *nghiệm tổng quát* hệ thuần nhất. Và không khó để rút ra nhận xét rằng, *nghiệm tổng quát của hệ không thuần nhất bằng nghiệm tổng quát hệ thuần nhất cộng với một nghiệm riêng của hệ không thuần nhất.*

1.6 Thực hành tính toán trên Maple

Maple là phần mềm toán học phổ biến với giao diện và cấu trúc lệnh dễ sử dụng cho mọi đối tượng, nó cung cấp đầy đủ các công cụ để tính toán số cho nhiều chuyên ngành trong đó có đại số tuyến tính. Ở đây chúng tôi cung cấp một số lệnh cơ bản trong tính toán thực hành với các ma trận, định thức và hệ phương trình tuyến tính (người học có thể tìm hiểu thêm qua các tài liệu hướng dẫn sử dụng và mục HELP trên trình duyệt). Hiện tại phiên bản mới nhất là Maple 17, tuy nhiên có thể cài các phiên bản trước đó để sử dụng vì không có nhiều thay đổi về cấu trúc lệnh.

1.6.1 Các phép toán và ký hiệu đặc biệt

i) Các phép toán cộng trừ nhân chia: $+, -, *, /$, lũy thừa dùng dấu mũ.
ii) Các hàm sơ cấp: $sin(x), cos(x), tan(x), cotan(x), exp(x), ln(x), log[a](x),$ $abs(x), max(x1, x2, ..), min(x1, x2, \ldots), sqrt(x), GAMMA(x), Beta(x, y)$
iii) Các hằng số: $Pi, I, infinity, true$
iv) Lệnh gán: $A :=$ biểu thức.

1.6.2 Tính toán với các biểu thức đại số

Thông thường ta khởi động công việc bằng lệnh *restart;*

- Khai triển biểu thức đại số: Dùng lệnh *expand (biểu thức);*
- Phân tích đa thức thành nhân tử: Dùng lệnh *factor (biểu thức);*
- Đơn giản biểu thức đại số: Dùng lệnh *simplify (biểu thức);*
- Tối giản phân thức: Lệnh *normal (phân thức);*
- Giải (bất) phương trình và hệ (bất) phương trình: Các lệnh

solve(phương trình, {danh sách biến});

solve({phương trình 1, phương trình 2,...} , {danh sách biến});

Ví dụ 38.

$> factor(x^4 - 10 * x^3 + 35 * x^2 - 50 * x + 24);$

$$(x - 1) * (x - 2) * (x - 3) * (x - 4)$$

$> bt := cos(x)^5 + sin(x)^4 + 2 * cos(x)^2 - 2 * sin(x)^2 - cos(2 * x);$

$$cos(x)^5 + sin(x)^4 + 2 * cos(x)^2 - 2 * sin(x)^2 - cos(2 * x)$$

$> simplify(bt);$

$$cos(x)^4 * (cos(x) + 1)$$

$> sys := \{2*x+5*y-4*z = 9, 3*x+5*y+2*z = 12, 4*x-y+5*z = -3\};$

$$\{2*x+5*y-4*z = 9, 3*x+5*y+2*z = 12, 4*x-y+5*z = -3\}$$

$> solve(sys, \{x, y, z\});$

$$\left\{x = -\frac{33}{37}, y = \frac{99}{37}, z = \frac{24}{37}\right\}$$

1.6.3 Tính toán trên ma trận

Để thực hiện tính toán trên ma trận ta sử dụng gói lệnh *linalg*.

- Khai báo ma trận: có hai cách khai báo

Cách 1: A:=matrix(m, n, [dãy (hàng) phần tử]);

Cách 2: A:= array([[hàng 1],[hàng 2],...,[hàng n]]);

- Khai báo vector (dù máy tính cho dạng hàng nhưng ta ngầm hiểu là dạng cột): v:=vector([dãy (hàng phần tử)]);

- Phép cộng và nhân ma trận với lệnh *evalm*, ngoài ra có thể dùng lệnh *multiply* để nhân ma trận.

- Tính định thức với lệnh det

- Ma trận chuyển vị lệnh *transpose*

- Ma trận khả nghịch dùng lệnh *inverse*

- So sánh hai ma trận A, B có bằng nhau không (giá trị logic là *true* hoặc *false*) lệnh *equal(A, B)*

- Tính vết ma trận bằng lệnh *trace*

- Đưa ma trận A về dạng hình thang bằng biến đổi dòng (dạng Gauss-Jordan) bởi lệnh *gaussjord(A);* muốn biết thêm thông tin về hạng của ma trận dùng lệnh *gaussjord(A,'r');* và để gọi hạng ma trận ra ta đánh lệnh r;

- Đưa ma trận vuông A về dạng tam giác trên bằng biến đổi dòng bởi lệnh *gausselim(A);* muốn biết thêm thông tin về hạng của ma trận và định thức của A dùng lệnh *gausselim(A,'r','d');* và để gọi hạng ma trận cùng định thức ra ta đánh thêm các lệnh r; d;

- Giải phương trình ma trận $Ax = b$ dùng lệnh *linsolve* (A, b); nếu máy không trả lời tức là hệ vô nghiệm.

Ví dụ 39.

$> with(linalg);$

BlockDiagonal, GramSchmidt, JordanBlock, LUdecomp, QRdecomp, Wronskian, addcol, addrow, adj, adjoint, angle, augment, backsub, band, basis, bezout, blockmatrix, charmat, charpoly, cholesky, col, coldim, colspace, colspan, companion, concat, cond, copyinto, crossprod, curl, definite, delcols, delrows, det, diag, diverge, dotprod, eigenvals, eigenvalues, eigenvectors, eigenvects, entermatrix, equal, exponential, extend, ffgausselim, fibonacci, forwardsub, frobenius, gausselim, gaussjord, geneqns, genmatrix, grad, hadamard, hermite, hessian, hilbert, htranspose, ihermite, indexfunc, innerprod, intbasis, inverse, ismith, issimilar, iszero, jacobian, jordan, kernel, laplacian, leastsqrs, linsolve, matadd, matrix, minor, minpoly, mulcol, mulrow, multiply, norm, normalize, nullspace, orthog, permanent, pivot, potential, randmatrix, randvector, rank, ratform, row, rowdim, rowspace, rowspan, rref, scalarmul, singularvals, smith, stackmatrix, submatrix, subvector, sumbasis, swapcol, swaprow, sylvester, toeplitz, trace, transpose, vandermonde, vecpotent, vectdim, vector, wronskian

$> A := array([[1, 2, 3], [1, 1, 2], [1, 1, 1]]);$

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$

$> B := array([[3, 5, 4], [2, 3, 3], [2, 2, 2]]);$

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

$> evalm(A + B);$

$$\begin{bmatrix} 4 & 7 & 7 \\ 3 & 4 & 5 \\ 3 & 3 & 3 \end{bmatrix}$$

$> evalm(A\& * B);$

$$\begin{bmatrix} 13 & 17 & 16 \\ 9 & 12 & 11 \\ 7 & 10 & 9 \end{bmatrix}$$

$> multiply(A, B);$

$$\begin{bmatrix} 13 & 17 & 16 \\ 9 & 12 & 11 \\ 7 & 10 & 9 \end{bmatrix}$$

$> \det(A);$

$$1$$

$> inverse(A);$

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

$> M := array([[1, 2, 3, 4], [2, 3, 5, 6], [-1, 3, 1, 4]]);$

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

$> gaussjord(M, 'r')$

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

$> r;$

$$3$$

$> gausselim(B, 'r', 'd')$

$$\begin{bmatrix} 3 & 5 & 4 \\ 0 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 0 & -2 \end{bmatrix}$$

$> r;$

$$3$$

$> d;$

$$2$$

$> b := array([1, 2, 3]);$

$$b := [1\ 2\ 3]$$

$> linsolve(A, b);$

$$[4\ 0\ -1]$$

Chú ý rằng nếu đánh $*$ mà máy báo lỗi thì đánh lại là $\&*$.

Ngoài ra chúng ta có thể dùng gói lệnh *LinearAlgebra* với các cú pháp lệnh khá gần với cách viết thông thường được sử dụng trong các sách giáo khoa về đại số tuyến tính hiện nay.

Chương 2

Không gian vector và ánh xạ tuyến tính

2.1 Không gian vector và không gian vector con

2.1.1 Định nghĩa

Giả sử $\mathbb{K}$ là một trường, V là tập khác rỗng, các phần tử của V gọi là *vector* và ký hiệu bởi $a, b, ..$, các phần tử của $\mathbb{K}$ gọi là *vô hướng (scalar)* ký hiệu bởi $\lambda, \theta, ...$ Trên V trang bị hai phép toán

- Phép cộng:

$$\begin{gathered} + : V \times V \to V \\ (a, b) \mapsto a + b \end{gathered}$$

- Phép nhân ngoài:

$$\begin{gathered} \cdot : \mathbb{K} \times V \to V \\ (\lambda, a) \mapsto \lambda a \end{gathered}$$

Ta nói V với hai phép toán trên lập thành một *không gian vector* (hay *không gian tuyến tính*) trên $\mathbb{K}$ (viết tắt là $\mathbb{K}$ - không gian vector) nếu tám tiên đề sau thỏa mãn:

i) Với mọi vector $a, b, c \in V$ ta có $(a + b) + c = a + (b + c)$

ii) Tồn tại vector $0 \in V$ sao cho $a + 0 = 0 + a = a, \forall a \in V$

iii) Với mọi vector $a \in V$ tồn tại vector $-a$ để $a + (-a) = (-a) + a = 0$

iv) Với mọi vector $a, b \in V$ thì $a + b = b + a$

v) Với mọi vô hướng $\lambda, \eta \in \mathbb{K}$, vector $a \in V$ đều có: $(\lambda + \eta)a = \lambda a + \eta a$

vi) Với mọi vô hướng $\lambda \in \mathbb{K}$, vector $a, b \in V$ đều có: $\lambda(a + b) = \lambda a + \lambda b$

vii) Với mọi vô hướng $\lambda, \eta \in \mathbb{K}$, vector $a \in V$ thì $(\lambda\eta)a = \lambda(\eta a)$

viii) Với mọi vector $a \in V$ thì $1a = a$, ở đó 1 là đơn vị trường $\mathbb{K}$.

Không khó để thấy bốn tiên đề đầu nói rằng V là một nhóm Abel với phép toán cộng vector, hai tiên đề tiếp theo nói rằng phép nhân vô hướng có tính chất phân phối với phép cộng vô hướng, phân phối với phép cộng vector, hai tiên đề cuối là tính kết hợp của phép nhân vô hướng và nhân vô hướng với đơn vị trường $\mathbb{K}$. Nếu không có gì hiểu lầm ta hiểu ký hiệu 0 vừa là vector, vừa là ký hiệu vô hướng $0 \in \mathbb{K}$.

Ví dụ 40. Trường $\mathbb{K}$ trên chính nó cũng là một không gian vector, khi đó mỗi phần tử vừa coi là vector vừa coi là vô hướng.

Ví dụ 41. Không gian $\mathbb{K}^n = \{x = (x_1, ..., x_n) : x_i \in \mathbb{K}\}$ với các phép toán định nghĩa như sau. Với mọi $x, y \in \mathbb{K}^n, x = (x_1, ..., x_n), y = (y_1, ..., y_n)$ và $\lambda \in \mathbb{K}$ thì:

$$x + y = (x_1 + y_1, ..., x_n + y_n)$$

$$\lambda x = (\lambda x_1, ..., \lambda x_n)$$

Dễ dàng kiểm tra $\mathbb{K}^n$ là $\mathbb{K}$ - không gian vector. Với $\mathbb{K} = \mathbb{R}$, và $n = 2$ (hoặc $n = 3$), chúng ta có liên hệ hình học trên mặt phẳng (hoặc không gian) đã được làm quen ở bậc học phổ thông.

Ví dụ 42. Giả sử $D \subset \mathbb{K}$ là một tập nào đó. Xét tập hợp tất cả các hàm số nhận giá trị trong $\mathbb{K}$ là $F(D) = \{f : D \to \mathbb{K}\}$ với phép toán cộng ánh xạ và nhân vô hướng xác định như sau:

$$(f + g)(x) = f(x) + g(x)$$

$$(\lambda f)(x) = \lambda f(x)$$

Khi đó có thể thấy $F(D)$ là $\mathbb{K}$ - không gian vector.

Ví dụ 43. Xét hệ phương trình thuần nhất

$$Ax = 0$$

ở đó $A = (a_{ij})_{m\times n} \in M_{m\times n}(\mathbb{K})$ và $x = (x_1, ..., x_n)^T \in \mathbb{K}^n$. Tập tất cả các nghiệm của hệ này lập thành một không gian vector, gọi là *không gian nghiệm* hệ thuần nhất.

Ví dụ 44. Xét phương trình vi phân sau

$$y'' + ay' + by = 0$$

trong đó a, b là các hằng số thực, phương trình này được gọi là phương trinh thuần nhất cấp hai. Lý thuyết phương trình vi phân chỉ ra rằng tập hai nghiệm bất kỳ (hai hàm khả vi thỏa mãn phương trình) tạo thành một không gian vector.

Khẳng định sau đây dễ dàng suy ra từ định nghĩa

Mệnh đề 2.1.1. *Trong không gian vector V, với mọi $a, b, c \in V$, $\lambda \in \mathbb{K}$ thì*
1) Vector 0 và vector đối $-a$ là duy nhất
2) Phương trình ẩn vector $x + a = b$ có duy nhất nghiệm $x = b - a$
3) Nếu $a + c = b + c$ thì có thể giản ước được, tức là khi đó có $a = b$
4) $0a = 0 = \lambda 0$, ở đây 0 là vô hướng trong phép nhân đầu tiên, là vector trong phép nhân thứ hai.
5) $\lambda a = 0$ khi và chỉ khi hoặc $\lambda = 0$ hoặc $a = 0$
6) $-(\lambda a) = (-\lambda)a$

Định nghĩa 27. *Giả sử V là $\mathbb{K}$ - không gian vector. Tập hợp $W \subseteq V$ gọi là một không gian vector con của V nếu đối với phép toán cộng vector và nhân vô hướng cảm sinh trên W thì W cũng là không gian vector.*

Nói cách khác W đóng với phép cộng vector và nhân vô hướng, hay là

$$a + b \in W, \forall a, b \in W$$

$$\lambda a \in W, \forall \lambda \in \mathbb{K}, \forall a \in \mathrm{W}$$

cũng có thể viết ngắn gọn lại là

$$\lambda a + \eta b \in \mathrm{W}, \forall \lambda, \eta \in \mathbb{K}, \forall a, b \in \mathrm{W}$$

Trong thực hành để xem một tập hợp con của V có là không gian con ta thường kiểm tra điều kiện cuối cùng này.

Giả sử $a_1, ..., a_m$ là các vector của $\mathbb{K}$ - không gian vector V nào đó. Một *tổ hợp tuyến tính* của m vector này là vector $\sum\limits_{i=1}^{m} \lambda_i a_i \in V$, ở đó λ_i là các phần tử nào đó của trường $\mathbb{K}$. Tập các tổ hợp tuyến tính của m vector này gọi là *bao tuyến tính* của chúng ký hiệu là $span\{a_1, ..., a_m\}$, như vậy

$$span\,\{a_1, ..., a_n\} = \left\{ a : a = \sum_{i=1}^{m} \lambda_i a_i, \lambda_i \in \mathbb{K} \right\} \subseteq V$$

Ví dụ 45. Dễ dàng kiểm tra được các không gian con sau đây

i) V và $\{0\}$ (không gian chỉ có vector 0) là các không gian con tầm thường của V.

ii)W $= \{x = (x_1, 0, ..., 0) : x_1 \in \mathbb{K}\}$ là không gian con của $\mathbb{K}^n$.

iii) Giả sử $a_1, ..., a_m$ là các vector của $\mathbb{K}$ - không gian vector V, khi đó $span\,\{a_1, ..., a_m\}$ là không gian vector con của V và gọi là không gian con sinh bởi $a_1, ..., a_m$, đôi khi còn viết là $\langle a_1, ..., a_m\rangle$.

iv) Ký hiệu $\mathbf{P}_n[x]$ là tập các đa thức một biến x bậc không quá n trên trường $\mathbb{K}$. Khi đó $\mathbf{P}_n[x]$ là không gian vector con của $\mathbb{K}[x]$ với phép toán cộng đa thức và phép nhân đa thức với phần tử trường $\mathbb{K}$.

v) Giả sử $D = (a, b) \subset \mathbb{K}$, ký hiệu $C(a, b)$ là tập tất cả các hàm liên tục trên (a, b), khi đó $C(a, b)$ là không gian vector con của không gian $F(a, b)$. Ký hiệu $C^1(a, b)$ là tập các hàm khả vi liên lục trên (a, b) (tập các hàm số khả vi và đạo hàm của chúng liên tục trên (a, b)) thì $C^1(a, b)$ lại là không gian con của $C(a, b)$, tổng quát ta có dãy các không gian con "lồng nhau": $C^k(a, b)$ là không gian con $C^{k-1}(a, b)$.

2.1.2 Hạng hệ hữu hạn vector. Cơ sở và chiều

a) Hệ độc lập tuyến tính và phụ thuộc tuyến tính

Định nghĩa 28. *Hệ vector $\{a_1, ..., a_n\}$ trong $\mathbb{K}$ - không gian vector V gọi là phụ thuộc tuyến tính nếu tồn tại $\lambda_1, ..., \lambda_n \in \mathbb{K}$ không đồng thời bằng* 0 *sao cho tổ hợp tuyến tính* $\sum\limits_{i=1}^{m} \lambda_i a_i = 0$

Hệ vector không phụ thuộc tuyến tính gọi là *độc lập tuyến tính*, tức là từ tổ hợp tuyến tính $\sum\limits_{i=1}^{m} \lambda_i a_i = 0$ suy ra $\lambda_1 = ... = \lambda_n = 0$.

Quy ước rằng, hệ rỗng $\emptyset$ là độc lập tuyến tính. Vector 0 là tổ hợp tuyến tính tầm thường của hệ $\emptyset$ và 0 là vector duy nhất biểu thị qua hệ $\emptyset$.

Ví dụ 46. Hệ n vector $a_1, ..., a_n \in \mathbb{K}^n$ là độc lập tuyến tính khi và chỉ khi định thức ma trận tạo bởi các cột tọa độ của chúng bằng 0. Thật vậy, giả sử ta có $\sum\limits_{i=1}^{n} \lambda_i a_i = 0$ trong đó $a_j = (a_{1j}, ..., a_{nj})^T$, tức là

$$\begin{cases} \lambda_1 a_{11} + \lambda_2 a_{12} + ... + \lambda_n a_{1n} = 0 \\ \cdots \\ \lambda_1 a_{n1} + \lambda_2 a_{n2} + ... + \lambda_n a_{nn} = 0 \end{cases}$$

Hệ độc lập tuyến tính tương đương với $\lambda_i \equiv 0, \forall i = \overline{1; n}$, nhưng theo lý thuyết hệ phương trình tuyến tính thuần nhất, điều này có nghĩa là $\det(A) \neq 0$, ở đó $A = (a_{ij})$ là ma trận của hệ.

Hệ độc lập và phụ thuộc tuyến tính có tính chất sau:

Mệnh đề 2.1.2. *i) Hệ một vector a là độc lập tuyến tính khi và chỉ khi $a \neq 0$*
ii) Hệ $\{a_1, .., a_n\}$ độc lập tuyến tính thì $a_i \neq 0, \forall i = \overline{1; n}$
iii) Hệ n vector $\{a_1, .., a_n\}$ phụ thuộc tuyến tính khi và chỉ khi có một vector biểu diễn tuyến tính qua $n - 1$ vector còn lại
iv) Giả sử ta có hệ $\{a_i\}_{i \in I}$, ở đó I là tập chỉ số nào đó và $J \subset I$. Nếu hệ $\{a_j\}_{j \in J}$ phụ thuộc tuyến tính thì hệ $\{a_i\}_{i \in I}$ phụ thuộc tuyến tính và nếu $\{a_i\}_{i \in I}$ độc lập tuyến tính thì $\{a_j\}_{j \in J}$ độc lập tuyến tính.

Không khó để chứng minh một kết quả sau đây

Định lý 2.1.1. *Giả sử $\{a_1, ..., a_m\}$ và $\{b_1, ..., b_n\}$ là hai hệ vector trong $\mathbb{K}$ - không gian vector V. Nếu hệ $\{a_1, ..., a_m\}$ độc lập tuyến tính và mỗi vector biểu thị tuyến tính qua hệ $\{b_1, ..., b_n\}$ thì $m \leq n$.*

Hai hệ vector tùy ý trong $\mathbb{K}$ - không gian vetcor V gọi là *tương đương* nếu chúng có thể biểu thị tuyến tính qua nhau.

Dễ dàng suy ra hệ quả sau đây từ định lý trên

Hệ quả 2.1.1. *Hai hệ vector $\{a_1, ..., a_m\}$ và $\{b_1, ..., b_n\}$ tương đương nhau và cả hai cùng độc lập tuyến tính thì $m = n$.*

Định nghĩa 29. *Hệ $\{a_i\}_{i \in I}$ trong $\mathbb{K}$ - không gian vetcor V, với I là tập chỉ số nào đó. Hệ $\{a_j\}_{j \in J \subset I}$ gọi là hệ con độc lập tuyến tính tối đại của hệ ban đầu nếu thêm bất kỳ vector $a_i, i \in I \backslash J$ nào thì hệ đều phụ thuộc tuyến tính.*

Định lý 2.1.2. *Nếu hệ con $\{a_1, ..., a_n\}$ là độc lập tuyến tính tối đại của $\{a_i\}_{i \in I}$ thì mọi vector $a_i, i \in I$ đều biểu thị tuyến tính duy nhất qua $\{a_1, ..., a_n\}$.*

Chứng minh: Nếu $i \in \{1, 2, ..., n\}$ thì dễ thấy có thể biểu thị tuyến tính duy nhất dưới dạng

$$a_i = \sum_{j=1}^{n} \lambda_j a_j, \lambda_j = \begin{cases} 1, j = i \\ 0, j \neq i \end{cases}$$

Nếu $i \notin \{1, 2, ..., n\}$, do $\{a_1, ..., a_n\}$ độc lập tuyến tính tối đại nên $\{a_1, ..., a_n, a_i\}$ là phụ thuộc tuyến tính và

$$a_i = \sum_{j=1}^{n} \lambda_j a_j$$

cần chứng minh biểu diễn này là duy nhất. Giả sử rằng

$$a_i = \sum_{j=1}^{n} \alpha_j a_j$$

khi đó ta có

$$\sum_{j=1}^{n} \lambda_j a_j = \sum_{j=1}^{n} \alpha_j a_j$$

hay

$$\sum_{j=1}^{n} (\lambda_j - \alpha_j)\, a_j = 0$$

do $\{a_1, ..., a_n\}$ độc lập tuyến tính nên từ đây ta có $\lambda_j = \alpha_j, \forall j = \overline{1; n}$. ▶

Chú ý 7. *i) Một hệ hữu hạn vector luôn có thể xây dựng một hệ con độc lập tuyến tính tối đại.*
ii) Hai hệ con độc lập tuyến tính tối đại của một hệ vector thì tương đương nhau.

Những nhận xét trên rất hữu ích và coi như phần bài tập cho người đọc.
b) Hạng của hệ hữu hạn vector.

Định nghĩa 30. *Xét hệ hữu hạn vector $\{a_i\}_{i \in I}$ (I hữu hạn), mọi hệ con độc lập tuyến tính tối đại của hệ này đều tương đương nhau và có cùng số vector, số đó gọi là hạng của hệ vector, ký hiệu là $rank\{a_i\}_{i \in I}$.*

Dễ thấy rằng

i) Nếu $rank\{a_i\}_{i \in I} = r$ thì $0 \le r \le |I|$, hơn thế $r = |I|$ khi và chỉ khi hệ $\{a_i\}_{i \in I}$ độc lập tuyến tính, $r = 0$ khi và chỉ khi $a_i = 0, \forall i$.

ii) Hai hệ hữu hạn vector tương đương nhau thì chúng có cùng hạng.

iii) Nếu có hệ vector $\{b_j\}_{j \in J}$ biểu thị tuyến tính được qua hệ $\{a_i\}_{i \in I}$ thì $rank\{a_i\}_{i \in I} = rank\left(\{a_i\}_{i \in I} \cup \{b_j\}_{j \in J}\right)$.

c) Hệ sinh và cơ sở

Định nghĩa 31. *Giả sử V là $\mathbb{K}$ không gian vector. Một hệ vector trong V gọi là hệ sinh của V nếu mọi vector của V đều biểu thị tuyến tính qua hệ này.*

Định nghĩa 32. *Một hệ sinh của V độc lập tuyến tính gọi là một cơ sở của V.*

Như vậy, cơ sở là hệ sinh độc lập tuyến tính, hiển nhiên khi đó hệ độc lập tuyến tính tối đại. Vì vậy mọi vector trong V đều biểu thị tuyến tính duy nhất qua nó. Nếu V có một hệ sinh hữu hạn thì ta nói V là $\mathbb{K}$ - *không gian vector hữu hạn sinh (chiều).*

Định nghĩa 33. *Số vector cơ sở của $\mathbb{K}$ - không gian vector V được gọi là chiều của của không gian này, ký hiệu là $dim_{\mathbb{K}}V$, hay đơn giản là $dimV$ nếu như đã xác định V trên trường $\mathbb{K}$.*

Mỗi không gian vector hữu hạn chiều đều có nhiều cơ sở (là các hệ độc lập tuyến tính tối đại) nhưng mỗi cơ sở đều có cùng số vector. Tổng quát ta có nhận xét sau

Chú ý 8. *Trong không gian vector V n chiều ta có*
i) Mọi hệ có số vector lớn hơn n đều phụ thuộc tuyến tính.
ii) Mọi cơ sở có đúng n vector, mọi hệ độc lập tuyến tính gồm n vector đều là cơ sở.
iii) Mọi hệ độc lập tuyến tính đều có thể bổ sung để được cơ sở của V.
iv) Mọi vector của V đều biểu thị tuyến tính duy nhất qua cơ sở cố định.

Ví dụ 47. $\mathbb{K}^n$ là $\mathbb{K}$ - không gian vector hữu hạn chiều. Thật vậy, không khó để thấy hệ n vector $\{e_i\}_{i=\overline{1;n}}$ ở đó

$$e_1 = (1, 0, 0, ..., 0), e_2 = (0, 1, 0, ..., 0), e_n = (0, ..., 0, 1)$$

là một hệ sinh của $\mathbb{K}^n$ vì với mọi $x = (x_1, x_2, ..., x_n) \in \mathbb{K}^n$ đều có biểu diễn

$$x = \sum_{i=1}^{n} x_i e_i$$

Rõ ràng $\{e_i\}_{i=\overline{1;n}}$ là độc lập tuyến tính, do đó nó là cơ sở của $\mathbb{K}^n$ (gọi là cơ sở chính tắc) và $dim\mathbb{K}^n = n$.

Ví dụ 48. Ký hiệu $\mathbb{K}[x]$ là tập hợp tất cả các đa thức trên trường $\mathbb{K}$, khi đó $\mathbb{K}[x]$ là không gian vector với phép cộng đa thức và nhân đa thức với phần tử trường $\mathbb{K}$. Có thể chứng minh $\mathbb{K}[x]$ không hữu hạn chiều (bài tập).

Ví dụ 49. Giả sử $\mathbb{K} = \mathbb{C}$ và không gian vector của chúng ta là $\mathbb{C}^n$. Chúng ta nhớ lại rằng $\mathbb{C} = \mathbb{R}^2$, và như vậy $\mathbb{C}^n$ giống như là $\mathbb{R}^{2n}$. Thế thì chiều của $\mathbb{C}^n$ là n hay $2n$? Ở đây chúng ta cần hiểu rằng $\mathbb{C}^n$ vừa là một không gian vector trên $\mathbb{C}$ nhưng cũng là không gian vector trên $\mathbb{R}$, và rằng $dim_{\mathbb{C}}\mathbb{C}^n = n$

còn $dim_{\mathbb{R}}\mathbb{C}^n = 2n$. Vì vậy khi chúng ta nói chiều của $\mathbb{C}^n$, chúng ta cần phân biệt giữa chúng, chúng ta nói *chiều phức* (bằng n) hoặc *chiều thực* (bằng $2n$).

2.1.3 Tọa độ của vector trong cơ sở. Đổi cơ sở

Giả sử V là $\mathbb{K}$ - không gian vector n chiều. Trong V cho một cơ sở cố định $\{e_1, ..., e_n\}$ (*ký hiệu hình thức,* $(e) = (e_1, ..., e_n)$ *biểu diễn dưới dạng hàng, mỗi phần tử lại là một vector*).Vì mọi vector $x \in V$ biểu thị tuyến tính duy nhất qua $\{e_1, ..., e_n\}$ nên tồn tại duy nhất bộ $(x_1, ..., x_n)$, $x_i \in \mathbb{K}$ sao cho

$$x = \sum_{i=1}^{n} x_i e_i$$

như vậy ta có một song ánh từ V vào $\mathbb{K}^n$.

Bộ $(x_1, x_2, ..., x_n)$ gọi là *tọa độ* của x trong cơ sở $\{e_1, ..., e_n\}$ và viết là $x_{(e)} = (x_1, ..., x_n)^T$ (tức là một ma trận cột). Từ nay để tránh nhầm lẫn, ta viết $x_{(e)} = (x_1, ..., x_n)^T = [x_i]_{(e)}$.

Ví dụ 50. Trong $\mathbb{K}^n$ xét *cơ sở chính tắc* $\{e_1, ..., e_n\}$, nếu xem (e) là một ma trận (cột i là tọa độ vector e_i) thì đây chính là ma trận đơn vị E.

Vì V có nhiều cơ sở nên mỗi vector trong cơ sở khác nhau sẽ biểu diễn khác nhau. Giả sử V có cơ sở $\{e_1, ..., e_n\}$, bên cạnh đó $\{e'_1, ..., e'_n\}$ cũng là cơ sở của V và

$$x_{(e)} = [x_i]_{(e)}; x_{(e')} = \left[x'_j\right]_{(e')}$$

Giả sử rằng mỗi vector e'_j trong cơ sở $\{e_1, ..., e_n\}$ có tọa độ là $[c_{ij}]_{(e)}$ tức là

$$e'_j = \sum_{i=1}^{n} c_{\mathrm{ij}} e_i \tag{2.1}$$

Ma trận $C = (c_{ij})_{n\times n} \in M_n(\mathbb{K})$ gồm các cột tọa độ của e'_j trong (e) được gọi là *ma trận chuyển cơ sở từ* (e) *sang* (e'), và viết là $C : (e) \to (e')$. Không khó để thấy rằng

$$(e') = (e)C \tag{2.2}$$

Ta có

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{j=1}^{n} x'_j e'_j = \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} c_{\mathrm{ij}} e_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} c_{\mathrm{ij}} x'_j \right) e_i$$

Do x biểu thị tuyến tính duy nhất qua $\{e_1, ..., e_n\}$ nên

$$x_i = \sum_{j=1}^{n} c_{\text{ij}} x'_j \tag{2.3}$$

hay viết dưới dạng ma trận là

$$[x_i]_{(e)} = C[x'_j]_{(e')} \tag{2.4}$$

Các công thức 2.3 và 2.4 gọi là các công thức đổi tọa độ.

Ví dụ 51. Trong $\mathbb{R}^3$ xét cơ sở chính tắc $\{e_1, e_2, e_3\}$, cho cơ sở thứ hai $\{e'_1, e'_2, e'_3\}$ ở đó ma trận chuyển cơ sở (các cột là tọa độ của e'_j trong (e)) là

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Trong cơ sở (e) vector x có tọa độ là $x_{(e)} = (1, 1, 1)^T$. Ta đi tìm tọa độ của x trong cơ sở $\{e'_1, e'_2, e'_3\}$. Giả sử $x_{(e')} = (x'_1, x'_2, x'_3)^T$, theo công thức 2.4 ta có

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$$

từ đây giải hệ phương trình ta được $x_{(e')} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)^T$.

Chú ý 9. *Nếu C là ma trận chuyển cơ sở từ $\{e_1, ..., e_n\}$ sang $\{e'_1, ..., e'_n\}$ thì C^{-1} là ma trận chuyển cơ sở từ $\{e'_1, ..., e'_n\}$ sang $\{e_1, ..., e_n\}$.*

2.1.4 Định lý về hạng ma trận

Giả sử $A = (a_{ij})_{m\times n} \in M_{m\times n}(\mathbb{K})$, hàng thứ i của ma trận là $r_i = (a_{i1}, a_{i2}, ..., a_{in}) \in \mathbb{K}^n, i = \overline{1; m}$, còn cột thứ j của ma trận là $c_j = (a_{1j}, a_{2j}, ..., a_{mj})^T \in \mathbb{K}^m, j = \overline{1; n}$. Như vậy có thể xem ma trận A cấu thành từ m vector hàng r_i $(i = \overline{1; m})$ hoặc n cột c_j $(j = \overline{1; n})$. Định lý dưới đây cho ta mối liên hệ giữa hạng của ma trận và hạng của hệ các vector hàng (cột) của ma trận.

Định lý 2.1.3. *Hạng của ma trận A bằng hạng của hệ vector hàng (cột) của nó, tức là*

$$rank\,(A) = rank\,(r_1, ..., r_m) = rank\,(c_1, ..., c_n)$$

Phần chứng minh có thể tìm thấy trong [3], [14]. Nhờ kết quả trên chúng ta đưa bài toán tìm hạng của hệ vector về bài toán tìm hạng ma trận đã biết ở chương trước.

Ví dụ 52. Giả sử trong không gian vector $\mathbb{K}^m$ xét n vector $v_1, ..., v_n$. Ma trận tạo bởi n vector cột là $A = (v_1, ..., v_n) \in M_{m\times n}(\mathbb{K})$. Hệ vector $\{v_1, ..., v_n\}$ độc lập tuyến tính khi hệ phương trình $Ax = 0$, ở đó $x = (x_1, ..., x_n)^T$, chỉ có nghiệm tầm thường, điều này có nghĩa là $rank(A) = n \le m$.

Bài toán tìm cơ sở, chiều của không gian con sinh bởi một hệ vector $\{a_1, a_2, ..., a_n\}$, tức là không gian con $span\{a_1, a_2, ..., a_n\}$. Để giải quyết bài toán trên chúng ta làm như sau:

Bước 1: Tìm hạng của hệ vector bằng cách viết ma trận tạo thành từ các hàng. Dùng biến đổi sơ cấp hàng đưa ma trận về dạng hình thang.

Bước 2: Số hàng khác 0 trong ma trận hình thang chính là hạng của hệ vector hay chính là chiều của không gian con sinh bởi hệ này còn sở của không gian con sinh bởi hệ chính là các vector hàng khác 0 của ma trận hình thang.

Ví dụ 53. Tìm cơ sở chiều của không gian con sinh bởi các vector hàng sau đây

$$a_1 = (1, 2, 0, 1)\,; a_2 = (1, 1, 2, 0)\,; a_3 = (1, 3, -2, 2)\,; a_4 = (-1, 1, -6, 2)$$

Để giải quyết bài toán, ta viết ma trận tạo thành từ các hàng vector và biến đổi sơ cấp đưa về ma trận hình thang

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 3 & -2 & 2 \\ -1 & 1 & -6 & 2 \end{pmatrix} \xrightarrow[h_1+h_4\to h_4]{\substack{-h_1+h_2\to h_2 \\ -h_1+h_3\to h_3}} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 2 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & 3 & -6 & 3 \end{pmatrix}$$

ma trận cuối tương đương với ma trận

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

nên ta có ngay chiều của không gian con sinh bởi các vector ban đầu là 2, còn cơ sở của không gian con là $\{b_1, b_2\}$, trong đó

$$b_1 = (1, 2, 0, 1)\,; b_2 = (0, 1, -2, 1)\,.$$

2.1.5 Không gian tổng và không gian giao. Tổng trực tiếp

Giả sử V là $\mathbb{K}$ - không gian vector, W_1, W_2 là các không gian con của V. Hiển nhiên rằng $W_1 \cap W_2$ đóng đối với phép cộng vector và nhân vector với vô hướng, do đó $W_1 \cap W_2$ cũng là không gian con của V và gọi là *không gian giao* của W_1 và W_2.

Ký hiệu

$$\mathrm{W}_1 + \mathrm{W}_2 = \{x = a_1 + a_2 : a_1 \in \mathrm{W}_1, a_2 \in \mathrm{W}_2\}$$

khi đó $W_1 + W_2$ cũng là không gian vector con của V và gọi là *không gian tổng* của W_1 và W_2.

Thật vậy, lấy hai vector bất kỳ $a, b \in W_1 + W_2$, theo định nghĩa tức là $a = a_1 + a_2, b = b_1 + b_2$ trong đó $a_1, b_1 \in W_1$ còn $a_2, b_2 \in W_2$. Với hai vô hướng bất kỳ $\alpha, \beta \in \mathbb{K}$ ta có

$$\alpha a + \beta b = \alpha\,(a_1 + a_2) + \beta\,(b_1 + b_2) = (\alpha a_1 + \beta b_1) + (\alpha a_2 + \beta a_2) \in \mathrm{W}_1 + \mathrm{W}_2$$

tức là $W_1 + W_2$ là không gian vector con của V.

Định lý 2.1.4. *Giả sử V là $\mathbb{K}$ - không gian vector hữu hạn chiều, W_1, W_2 là các không gian con của V, khi đó*

$$\dim(\mathrm{W}_1 + \mathrm{W}_2) = \dim(\mathrm{W}_1) + \dim(\mathrm{W}_2) - \dim(\mathrm{W}_1 \cap \mathrm{W}_2)$$

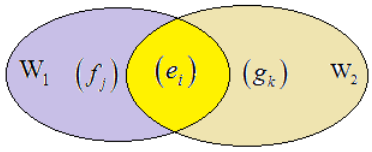
Chứng minh: Giả sử $dimW_1 = m, dimW_2 = n$, còn $dim(W_1 \cap W_2) = p$ và $\{e_1, ..., e_p\}$ là cơ sở của $W_1 \cap W_2$. Bổ sung thêm các vector để $\{e_1, ..., e_p, f_{p+1}, ..., f_m\}$ trở thành cơ sở của W_1 và $\{e_1, ..., e_p, g_{p+1}, ..., g_n\}$ thành cơ sở của W_2. Ta sẽ chứng minh rằng hệ các vector $\{e_1, ..., e_p, f_{p+1}, ..., f_m, g_{p+1}, ..., g_n\}$ độc lập tuyến tính và là cơ sở của $W_1 + W_2$. Thật vậy, xét tổ hợp tuyến tính

$$\alpha_1 e_1 + ... + \alpha_p e_p + \beta_{p+1} f_{p+1} + ... + \beta_m f_m + \gamma_{p+1} g_{p+1} + ... + \gamma_n g_n = 0$$

trong đó các vô hướng thuộc trường $\mathbb{K}$.

Ta có

$$\alpha_1 e_1 + ... + \alpha_p e_p + \beta_{p+1} f_{p+1} + ... + \beta_m f_m = -\left(\gamma_{p+1} g_{p+1} + ... + \gamma_n g_n\right)$$



.

Hình 2.1: Không gian tổng và giao

vì vế phải đẳng thức nằm trong W_2, còn vế trái nằm trong W_1 nên

$$-\left(\gamma_{p+1}g_{p+1}+...+\gamma_n g_n\right)\in \mathrm{W}_1\cap \mathrm{W}_2$$

Do đó tồn tại $\eta_1,...,\eta_p\in\mathbb{K}$ sao cho

$$-\left(\gamma_{p+1}g_{p+1}+...+\gamma_n g_n\right)=\eta_1 e_1+...+\eta_p e_p$$

hay là

$$\eta_1 e_1+...+\eta_p e_p+\gamma_{p+1}g_{p+1}+...+\gamma_n g_n=0$$

nhưng $\{e_1,...,e_p,g_{p+1},...,g_n\}$ là cơ sở W_2 nên chúng độc lập tuyến tính, do đó

$$\eta_1=..=\eta_p=\gamma_{p+1}=...=\gamma_n=0$$

Từ đây suy ra rằng

$$\alpha_1 e_1+...+\alpha_p e_p+\beta_{p+1}f_{p+1}+...+\beta_m f_m=0$$

và cũng với lý do tương tự như trên ta có

$$\alpha_1=...=\alpha_p=\beta_{p+1}=...=\beta_m=0$$

Tức là $\{e_1,...,e_p,f_{p+1},...,f_m,g_{p+1},...,g_n\}$ độc lập tuyến tính. Lấy vector bất kỳ $x\in W_1+W_2$, khi đó $x=x_1+x_2$ với $x_1\in W_1, x_2\in W_2$. Rõ ràng x_1 biểu thị tuyến tính qua $\{e_1,...,e_p,f_{p+1},...,f_m\}$, còn x_2 biểu thị tuyến tính qua $\{e_1,...,e_p,g_{p+1},...,g_n\}$, do đó x biểu thị tuyến tính qua $\{e_1,...,e_p,f_{p+1},...,f_m,g_{p+1},...,g_n\}$. Nói cách khác $\{e_1,...,e_p,f_{p+1},...,f_m,g_{p+1},...,g_n\}$ là hệ sinh của W_1+W_2. Một hệ sinh độc lập tuyến tính chính là một cơ sở, tức là chúng ta chứng minh được $\{e_1,...,e_p,f_{p+1},...,f_m,g_{p+1},...,g_n\}$ là cơ sở của W_1+W_2. Hơn thế

$$\dim\left(\mathrm{W}_1{+}\mathrm{W}_2\right)=m+n-p=\dim W_1+\dim W_2-\dim\left(\mathrm{W}_1\cap \mathrm{W}_2\right)$$

Ta có điều phải chứng minh. ▶

Nếu $W_1 \cap W_2 = \{0\}$ thì $W_1 + W_2$ gọi là một *tổng trực tiếp* và ký hiệu là $\mathrm{W}_1 \oplus \mathrm{W}_2$. Khi đó

$$\dim (\mathrm{W}_1 \oplus \mathrm{W}_2) = \dim W_1 + \dim W_2$$

Bài toán tìm cơ sở chiều của các không gian con W_1, W_2, không gian tổng $W_1 + W_2$ và giao $W_1 \cap W_2$. Để giải quyết ta tiến hành làm ba bài toán nhỏ như sau.

Bài toán 1: Tìm cơ sở chiều các không gian con W_1 và W_2, giả sử cơ sở của chúng lần lượt là $\{a_1, a_2, ..., a_l\}$ và $\{b_1, b_2, ..., b_m\}$.

Bài toán 2: Để tìm cơ sở chiều của không gian tổng $W_1 + W_2$ ta tìm cơ sở chiều của không gian con sinh bởi hệ vector $\{a_1, a_2, ..., a_l, b_1, b_2, ..., b_m\}$.

Bài toán 3: Để tìm cơ sở chiều của không gian giao $W_1 \cap W_2$ ta tìm cơ sở chiều không gian nghiệm hệ thuần nhất với $l + m$ ẩn là $x_1, ..., x_l, y_1, ..., y_m$ sau

$$x_1 a_1 + x_2 a_2 + ... + x_l a_l = y_1 b_1 + y_2 b_2 + ... + y_m b_m$$

Bài toán 1 và 2 được giải quyết dựa trên bài toán tìm cơ sở chiều không gian con ở phía trên. Đối với bài toán 3, chúng ta sẽ giải quyết sau khi nghiên cứu kỹ cấu trúc không gian nghiệm hệ thuần nhất trong phần sau.

2.2 Ánh xạ tuyến tính

2.2.1 Khái niệm ánh xạ tuyến tính và toán tử tuyến tính

Định nghĩa 34. *Giả sử V, W là các $\mathbb{K}$ - không giac vector. Ánh xạ f từ V vào W gọi là một ánh xạ tuyến tính (đồng cấu tuyến tính) nếu nó bảo toàn các phép toán trên V, tức là*

i) $f(x + y) = f(x) + f(y), \forall x, y \in V$

ii) $f(\lambda x) = \lambda f(x), \forall \lambda \in \mathbb{K}, x \in V$.

Một ánh xạ tuyến tính từ V vào chính nó được gọi là một *tự đồng cấu tuyến tính* hay *toán tử tuyến tính* trên V.

Có thể viết lại hai điều kiện trong định nghĩa thành một điều kiện là

$$f\left(\lambda x + \beta y\right) = \lambda f\left(x\right) + \beta f\left(y\right), \forall \lambda, \beta \in \mathbb{K}, x, y \in V.$$

Trong trường hợp tổng quát, ta có

$$f\left(\sum_{i=1}^{n}\lambda_i x_i\right)=\sum_{i=1}^{n}\lambda_i f\left(x_i\right),\forall\lambda_i\in\mathbb{K},x_i\in V.$$

Không khó để thấy rằng nếu f là ánh xạ tuyến tính thì $f(0)=0$ và $f(-x)=-f(x),\forall x\in V.$

Giả sử $\mathbb{K}$ là một trường, V,W là các $\mathbb{K}$ - không gian vector, các ví dụ dưới đây là ánh xạ tuyến tính.

Ví dụ 54. Ánh xạ hằng 0 được cho như sau $0:V\to W$ sao cho $0(x)=0,\forall x\in V.$

Ví dụ 55. Ánh xạ đồng nhất $Id:V\to V$ sao cho $Id(x)=x,\forall x\in V.$

Ví dụ 56. Ánh xạ $f:\mathbb{K}^n\to\mathbb{K}$ sao cho

$$f(x)=\sum_{i=1}^{n}\alpha_i x_i,\forall x=(x_1,...,x_n)\in\mathbb{K}^n,\alpha_i\in\mathbb{K}.$$

Ví dụ 57. Giả sử $A\in M_{m\times n}(\mathbb{K})$, ánh xạ $f:\mathbb{K}^n\to\mathbb{K}^m$ sao cho $f(x)=Ax,\forall x=(x_1,...,x_n)^T\in\mathbb{K}^n.$

Ví dụ 58. Ánh xạ đạo hàm các đa thức $D:\mathbb{R}[x]\to\mathbb{R}[x]$ sao cho $D\left(p(x)\right)=p'\left(x\right),\forall p(x)\in\mathbb{R}[x].$

Ví dụ 59. Ánh xạ tích phân các hàm liên tục $I:C\left[a,b\right]\to\mathbb{R}$ sao cho

$$I\left(x\right)=\int\limits_a^b x\left(t\right)dt,\forall x(.)\in C\left[a,b\right]$$

Ví dụ 60. Ánh xạ $f:\mathbb{R}^2\to\mathbb{R}$ xác định bởi

$$f\left(x,y\right)=\begin{cases}\dfrac{x^2+y^2}{x+y},x+y\neq 0\\ 0,x+y=0\end{cases}$$

thỏa mãn điều kiện thứ hai của định nghĩa ánh xạ tuyến tính nhưng không thỏa mãn điều kiện đầu tiên.

Ví dụ 61. Ánh xạ $f:\mathbb{C}\to\mathbb{C}$ xác định bởi

$$f\left(z\right)=\overline{z}$$

thỏa mãn điều kiện cộng tính (điều kiện thứ nhất) nhưng không phải ánh xạ tuyến tính.

Ký hiệu $Hom(V, W)$ (*Homomorphism*) là tập tất cả ánh xạ tuyến tính từ V vào W. Trên $Hom(V, W)$ trang bị phép cộng xác định bởi

$$(f+g)(x) = f(x) + g(x), \forall f, g \in Hom(V, W)$$

và phép nhân với phần tử của $\mathbb{K}$ xác định bởi

$$(\lambda f)(x) = \lambda f(x), \forall \lambda \in \mathbb{K}$$

Ánh xạ 0 chính là phần tử trung hòa của phép cộng, phần tử đối của f là ánh xạ $-f$. Không khó để thấy tập $Hom(V, W)$ với hai phép toán trên lập thành một $\mathbb{K}$ - không gian vector. Khi $W = \mathbb{K}$, ta nói $Hom(V, \mathbb{K})$ là không gian liên hợp của V ký hiệu là V^*. Cũng với các phép toán trên, tập các tự đồng cấu tuyến tính, ký hiệu là $End(V)$ (*Endomorphism*), cũng là không gian vector.

Dưới đây chúng ta đưa ra định lý xác định một ánh xạ tuyến tính.

Định lý 2.2.1. *Giả sử V là $\mathbb{K}$ - không gian vector n chiều với cơ sở $\{e_1, ..., e_n\}$ và một hệ n vector tùy ý cho trước $\{w_1, w_2, ..., w_n\}$ trong $\mathbb{K}$ - không gian vector W. Khi đó tồn tại duy nhất một ánh xạ tuyến tính $f : V \to W$ sao cho $f(e_i) = w_i, \forall i = \overline{1; n}$.*

Chứng minh: Giả sử $x \in V$, vì $\{e_1, ..., e_n\}$ là cơ sở nên tồn tại duy nhất bộ $(x_1, .., x_n)$ sao cho

$$x = \sum_{i=1}^{n} x_i e_i, x_i \in \mathbb{K}$$

Ta xây dựng ánh xạ f từ V vào W thỏa mãn $f(e_i) = w_i, i = \overline{1; n}$.

Đặt $f(x) = \sum\limits_{i=1}^{n} x_i \mathrm{w}_i$, khi đó f thực sự là ánh xạ vì bộ $(x_1, .., x_n)$ là duy nhất nên $f(x)$ là duy nhất, hơn thế f là ánh xạ tuyến tính. Thật vậy $\forall x, y \in V : x = \sum\limits_{i=1}^{n} x_i e_i, y = \sum\limits_{i=1}^{n} y_i e_i$ và $\forall \alpha, \beta \in \mathbb{K}$ ta có

$$f(\alpha x + \beta y) = \sum_{i=1}^{n} (\alpha x_i + \beta y_i) \mathrm{w}_i = \alpha f(x) + \beta f(y).$$

Để ý rằng $e_i = 0e_1 + ... + 0e_{i-1} + e_i + 0e_{i+1} + ... + 0e_n$ nên $f(e_i) = 0\mathrm{w}_1 + ... + 0\mathrm{w}_{i-1} + \mathrm{w}_i + 0\mathrm{w}_{i+1} + ... + 0\mathrm{w}_n = \mathrm{w}_i$.

Chúng ta cần chứng minh ánh xạ như vậy xác định duy nhất. Giả sử g là một ánh xạ tuyến tính từ V vào W cũng thỏa mãn $g(e_i) = w_i, \forall i = \overline{1; n}$.

Khi đó $\forall x = \sum\limits_{i=1}^{n} x_i e_i \in V$ thì

$$g(x) = \sum_{i=1}^{n} x_i g(e_i) = \sum_{i=1}^{n} x_i \mathrm{w}_i = \sum_{i=1}^{n} x_i f(e_i) = f(x)$$

hay $f \equiv g$. ▶

2.2.2 Ảnh và nhân của ánh xạ tuyến tính

Giả sử V, W là các $\mathbb{K}$ - không gian vector, $f \in Hom(V, W)$.

Định nghĩa 35. *Hạch (hay nhân) của ánh xạ tuyến tính f là tập hợp ký hiệu bởi $Ker(f)$ và xác định bởi*

$$Kerf = \{x \in V : f(x) = 0\}$$

Ảnh của ánh xạ tuyến tính f, ký hiệu là Imf, và xác định bởi

$$Imf = \{y \in W : y = f(x), \forall x \in V\}$$

Mệnh đề 2.2.1. *$Kerf$ và Imf là các không gian con của V và W tương ứng.*

Định lý 2.2.2. *Giả sử Ω, U lần lượt là các không gian vector con của V, W tương ứng. Khi đó*
i) $f(\Omega)$ là $\mathbb{K}$ - không gian vector con của W.
ii) $f^{-1}(U)$ là $\mathbb{K}$ - không gian vector con của V.
iii) Ω là không gian vector con hữu hạn chiều của V thì $f(\Omega)$ cũng hữu hạn chiều và $dimf(\Omega) \leq dim\Omega$.

Chứng minh: Việc chứng minh i) và ii) không khó, chú ý là $f(\Omega)$ và $f^{-1}(U)$ không rỗng. Ta chứng minh iii). Giả sử $dim\Omega = m$ và $\{e_1, ..., e_m\}$ là cơ sở Ω, khi đó với mỗi $y \in f(\Omega)$ tồn tại $x = \sum\limits_{i=1}^{m} x_i e_i$ sao cho

$$y = \sum_{i=1}^{m} x_i f(e_i)$$

như vậy y biểu thị tuyến tính qua $\{f(e_1), ..., f(e_m)\}$ hay $\{f(e_1), ..., f(e_m)\}$ là một hệ sinh của $f(\Omega)$. Vì vậy $f(\Omega)$ cũng hữu hạn chiều và $dimf(\Omega) \leq m = dim\Omega$. ▶

Chú ý 10. *i) Nếu $f : V \to W$ và V hữu hạn chiều thì $dim Imf \leq n = dimV$.*

ii) Giả sử $f : V \to W$ là ánh xạ tuyến tính và $\{v_i\}_{i=\overline{1;m}}$ là hệ bất kỳ, khi đó nếu $\{v_i\}_{i=\overline{1;m}}$ phụ thuộc tuyến tính thì $\{f(v_i)\}_{i=\overline{1;m}}$ là phụ thuộc tuyến tính. Ngược lại, nếu $\{f(v_i)\}_{i=\overline{1;m}}$ độc lập tuyến tính thì $\{v_i\}_{i=\overline{1;m}}$ độc lập tuyến tính.

Hệ quả 2.2.1. *Ánh xạ tuyến tính không làm tăng hạng của một hệ vector.*

Định nghĩa 36. *Giả sử $f : V \to W$ là ánh xạ tuyến tính. Chiều của Imf gọi là hạng của ánh xạ f, ký hiệu $rank(f)$, còn chiều của $Kerf$ gọi là số khuyết của f, ký hiệu là $def(f)$.*

Từ định nghĩa ta thấy ngay rằng

$$0 \leqslant rank\,(f) \leqslant \dim V$$
$$0 \leqslant \operatorname{def}(f) \leqslant \dim V$$

Định nghĩa 37. *Giả sử $f : V \to W$ là ánh xạ tuyến tính. Ánh xạ f gọi là đơn cấu (toàn cấu, đẳng cấu) nếu f là đơn ánh (toàn ánh, song ánh).*

Định lý 2.2.3. *Giả sử $f : V \to W$ là ánh xạ tuyến tính.*
i) f đơn cấu khi và chỉ khi $Kerf = \{0\}$
ii) f toàn cấu khi và chỉ khi $Imf = W$

Chứng minh: Rõ ràng ii) là hiển nhiên theo định nghĩa toàn ánh. Ta chứng minh i). Giả sử f đơn ánh, nếu $f(x) = 0 = f(0)$ thì $x = 0$, do đó $Kerf = \{0\}$. Ngược lại, giả sử $Kerf = \{0\}$. Nếu $f(x_1) = f(x_2)$ với $x_1, x_2 \in V$ thì ta có $f(x_1 - x_2) = 0$, hay $x_1 = x_2$, tức là f đơn ánh. ▶

Từ định lý trên ta thấy rằng, f đơn cấu khi và chỉ khi $def(f) = 0$, còn f toàn cấu khi và chỉ khi $rank(f) = dimW$.

Ví dụ 62. Ánh xạ $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi

$$f(x_1, x_2, x_3) = (x_1, 0).$$

Ta có

$$\text{Imf} = \{(x_1, 0) : x_1 \in \mathbb{R}\}$$
$$\text{Kerf} = \{(0, x_2, x_3) : x_2, x_3 \in \mathbb{R}\}.$$

Ví dụ 63. Ánh xạ $f : \mathbb{R}^3 \to \mathbb{R}^2$ xác định bởi

$$f(x_1, x_2, x_3) = (x_1, x_2).$$

Ta có

$$\text{Imf} = \{(x_1, x_2) : x_1, x_2 \in \mathbb{R}\} = \mathbb{R}^2$$

$$\text{Kerf} = \{(0, 0, x_3) : x_3 \in \mathbb{R}\}.$$

Ví dụ 64. Ánh xạ $f : \mathbb{R}^2 \to \mathbb{R}^3$ xác định bởi

$$f(x_1, x_2) = (x_1, x_2, x_1 - 2x_2).$$

Ta có

$$\text{Imf} = \{(x_1, x_2, x_1 - 2x_2) : x_1, x_2 \in \mathbb{R}\}$$

$$\text{Kerf} = \{(0, 0, 0)\} = \{0\}.$$

Không khó để chứng minh kết quả sau đây

Định lý 2.2.4. *Giả sử $f : V \to W$ là ánh xạ tuyến tính, V, W là các $\mathbb{K}$ - không gian vector hữu hạn chiều*
i) f đơn cấu khi và chỉ khi f biến một hệ độc lập tuyến tính bất kỳ thành một hệ độc lập tuyến tính
ii) f toàn ánh khi và chỉ khi f biến một hệ sinh của V thành hệ sinh trong W
iii) f đẳng cấu khi và chỉ khi f biến một cơ sở V thành cơ sở của W.

Định lý 2.2.5. *Giả sử $f : V \to W$ là ánh xạ tuyến tính, khi đó*

$$\dim \operatorname{Im} f + \dim Kerf = n$$

Chứng minh: Giả sử $dimKerf = s$ và $\{e_1, ..., e_s\}$ là cơ sở của $Kerf$. Ta bổ sung $\{e_1, ..., e_s, v_{s+1}, ..., v_n\}$ để được cơ sở của V. Ta sẽ chứng minh rằng $\{f(v_{s+1}), ..., f(v_n)\}$ là cơ sở của Imf. Thật vậy, với mọi $x \in V$ ta có

$$x = \sum_{i=1}^{s} \alpha_i e_i + \sum_{j=s+1}^{n} \beta_j v_j$$

do đó

$$f(x) = f\left(\sum_{i=1}^{s} \alpha_i e_i + \sum_{j=s+1}^{n} \beta_j v_j\right) = \sum_{j=s+1}^{n} \beta_j f(v_j)$$

tức là $\{f(v_j)\}_{j=\overline{s+1;n}}$ là hệ sinh của Imf. Ta chứng minh hệ sinh này độc lập tuyến tính. Xét tổ hợp tuyến tính bất kỳ

$$\sum_{j=s+1}^{n} \beta_j f(v_j) = 0 \Rightarrow f\left(\sum_{j=s+1}^{n} \beta_j v_j\right) = 0$$

do vậy $\sum\limits_{j=s+1}^{n} \beta_j v_j \in Kerf$, nhưng khi đó ta có

$$\sum_{j=s+1}^{n} \beta_j v_j = \sum_{i=1}^{s} \lambda_i e_i$$

hay là

$$\sum_{j=s+1}^{n} \beta_j v_j - \sum_{i=1}^{s} \lambda_i e_i = 0$$

do $\{e_1, ..., e_s, v_{s+1}, ..., v_n\}$ là cơ sở của V nên

$$\beta_j = \lambda_i = 0, \forall i = \overline{1; s}, j = \overline{s+1; n}$$

tức là $\{f(v_j)\}_{j=\overline{s+1;n}}$ độc lập tuyến tính, do đó là cơ sở của Imf. Từ đó

$$\dim \operatorname{Im} f + \dim Kerf = n - s + s = n$$

Ta có ngay điều phải chứng minh. ▶

2.2.3 Ánh xạ tuyến tính ngược

Định lý 2.2.6. *Giả sử $f : V \to W$ là đẳng cấu tuyến tính, khi đó ánh xạ ngược của một ánh xạ tuyến tính cũng là ánh xạ tuyến tính.*

Chứng minh: Giả sử $f : V \to W$ là song ánh, ta chứng minh rằng f^{-1} là ánh xạ tuyến tính. Thật vậy $\forall y_1, y_2 \in W$ và $\forall \alpha, \beta \in \mathbb{K}$ ta có

$$f\left(f^{-1}(\alpha y_1 + \beta y_2)\right) = \left(f \circ f^{-1}\right)(\alpha y_1 + \beta y_2) = Id_{\mathrm{W}}(\alpha y_1 + \beta y_2)$$

ở đó Id_W là ánh xạ đồng nhất trên W. Nhưng ánh xạ Id_W cũng là tuyến tính nên

$$Id_{\mathrm{W}}(\alpha y_1 + \beta y_2) = \alpha Id_{\mathrm{W}}(y_1) + \beta Id_{\mathrm{W}}(y_2) = \alpha f\left(f^{-1}(y_1)\right) + \beta f\left(f^{-1}(y_2)\right)$$

Do f đẳng cấu nên từ

$$f\left(f^{-1}(\alpha y_1 + \beta y_2)\right) = f\left(\alpha f^{-1}(y_1) + \beta f^{-1}(y_2)\right)$$

ta có ngay

$$f^{-1}(\alpha y_1 + \beta y_2) = \alpha f^{-1}(y_1) + \beta f^{-1}(y_2)$$

▶

Hệ quả 2.2.2. *Giả sử $f : V \to V$ là toán tử tuyến tính, khi đó các mệnh đề sau tương đương*
i) f đẳng cấu
ii) Tồn tại toán tử tuyến tính ngược của f
iii) f đơn cấu
iv) f toàn cấu

2.2.4 Ma trận và biểu thức tọa độ ánh xạ tuyến tính

Giả sử $f : V \to W$ là ánh xạ tuyến tính, $dimV = n, dimW = m$, $\{e_1, ..., e_n\}$ và $\{w_1, ..., w_m\}$ lần lượt là cơ sở của V và W. Vì $f(e_j) \in W^m, j = \overline{1; n}$ nên biểu diễn được qua (w)

$$f(e_j) = \sum_{i=1}^{m} a_{\text{ij}} \text{w}_i$$

Ánh xạ f hoàn toàn xác định khi biết ảnh của cơ sở tức là nếu ta biết các phần tử $a_{ij}, i = \overline{1; m}, j = \overline{1; n}$ thì sẽ hoàn toàn xác định được f.

Ma trận $A = (a_{ij})_{m \times n}$ được gọi là *ma trận của ánh xạ tuyến tính* f trong cặp cơ sở $((e), (w))$. Chúng ta có thể biểu diễn điều này như sau

$$(f(e)) = (w)A \tag{2.5}$$

Như vậy, cột thứ j của ma trận A chính là tọa độ của vector $f(e_j)$ trong cơ sở $\{w_1, ..., w_m\}$. Khi không nói gì khác ta hiểu ma trận của ánh xạ tuyến tính là trong một cơ sở cố định nào đó.

Nếu f là tự đồng cấu tuyến tính thì A sẽ là một ma trận vuông. Có thể chứng minh được có các song ánh giữa $Hom(V, W)$ với $M_{m \times n}(\mathbb{K})$ và $End(V)$ với $M_n(\mathbb{K})$, các song ánh này bảo toàn phép toán tổng các ánh xạ tuyến tính và phép nhân phần tử trường $\mathbb{K}$ với ánh xạ tuyến tính.

Giả sử $A = (a_{ij})_{m \times n}$ là ma trận của ánh xạ tuyến tính $f : V \to W$ trong cặp cơ sở $((e), (w))$, với mỗi $x \in V$ chúng ta muốn thiết lập liên hệ giữa các tọa độ $x_{(e)}$ và $f(x)_{(w)}$.

Giả sử

$$x = \sum_{j=1}^{n} x_j e_j$$

hay là $x_{(e)} = [x_j]_{(e)}$ và

$$f(x) = \sum_{i=1}^{m} y_i \text{w}_i$$

hay là $f(x)_{(\mathrm{w})} = [y_i]_{(\mathrm{w})}$. Khi đó

$$\sum_{i=1}^{n} y_i \mathrm{w}_i = f(x) = \sum_{j=1}^{n} x_j f(e_j) = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{n} a_{\mathrm{ij}} \mathrm{w}_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{\mathrm{ij}} x_j \right) \mathrm{w}_i$$

từ đây ta có

$$y_i = \sum_{j=1}^{m} a_{\mathrm{ij}} x_j, i = \overline{1; m} \tag{2.6}$$

hay viết dưới dạng tọa độ cột là

$$[y_i]_{(w)} = A[x_j]_{(e)} \tag{2.7}$$

Nếu như ta đã ngầm hiểu ánh xạ tuyến tính trong cặp cơ sở cố định nào đó thì ta viết là

$$f(x) = Ax \tag{2.8}$$

ở đó **tọa độ các vecto luôn hiểu là cho dưới dạng cột**. Các công thức từ 2.6 đến 2.8 gọi là biểu thức tọa độ của ánh xạ tuyến tính.

Nếu chỉ cho ánh xạ tuyến tính mà không nói gì về cặp cơ sở $((e), (w))$ ta ngầm hiểu rằng ánh xạ được cho trong cặp cơ sở chính tắc của V và W. Ma trận trong cặp cơ sở chính tắc gọi là *ma trận chính tắc.*

Ví dụ 65. Tìm ma trận chính tắc của ánh xạ tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^3$ cho bởi

$$f(x_1, x_2, x_3, x_4) = (x_1 - x_2 + x_3 - x_4, x_2 + x_3, x_1)$$

Trong cơ sở chính tắc của $\mathbb{R}^4$ và $\mathbb{R}^3$, ta có

$$f(1,0,0,0) = (1,0,1)^T, f(0,1,0,0) = (-1,1,0)^T,$$
$$f(0,0,1,0) = (1,1,0)^T, f(0,0,0,1) = (-1,0,0)^T$$

nên ma trận của f sẽ là

$$A = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Định lý 2.2.7. *Giả sử M là ma trận ánh xạ tuyến tính $f : V \to W$ trong cặp cơ sở $((e), (w))$, N là ma trận ánh xạ tuyến tính $g : W^m \to \Omega^p$ trong cặp cơ sở $((w), (v))$. Khi đó ma trận của ánh xạ $g \circ f$ là NM với cặp cơ sở là $((e), (v))$.*

Chứng minh: Giả sử

$$\{e_1, ..., e_n\}, \{\mathrm{w}_1, ..., \mathrm{w}_m\}, \{v_1, ..., v_p\}$$

lần lượt là cơ sở của V, W, Ω, còn $M = (a_{ij})_{m\times n}, N = (b_{ki})_{p\times m}$ là các ma trận tương ứng của f và g trong các cặp cơ sở đã cho. Ta có

$$f(e_j) = \sum_{i=1}^{m} a_{\mathrm{ij}}\mathrm{w}_i, g(\mathrm{w}_i) = \sum_{k=1}^{p} b_{ki}v_k$$

từ đó

$$(g \circ f)(e_j) = g(f(e_j)) = g\left(\sum_{i=1}^{m} a_{\mathrm{ij}}\mathrm{w}_i\right) = \sum_{i=1}^{m} a_{\mathrm{ij}} g(\mathrm{w}_i)$$

hay

$$(g \circ f)(e_j) = \sum_{i=1}^{m} a_{\mathrm{ij}}\left(\sum_{k=1}^{p} b_{ki}v_k\right) = \sum_{k=1}^{p}\left(\sum_{i=1}^{m} b_{ki}a_{\mathrm{ij}}\right)v_k$$

Đặt $c_{kj} = \sum\limits_{i=1}^{m} b_{ki}a_{\mathrm{ij}}$, thì trong cặp cơ sở $((e), (v))$ ma trận của $g \circ f$ là $NM = (c_{kj})_{p\times n}$. ▶

Tương tự có thể chỉ ra nếu A, B là ma trận của các ánh xạ tuyến tính $f, g \in Hom(V, W)$ trong cặp cơ sở nào đó thì $A + B$ chính là ma trận của ánh xạ $f + g$. Đặc biệt nếu $f, g \in End(V)$ là các toán tử tuyến tính với các ma trận A, B tương ứng trong cơ sở nào đó thì có thể chứng minh được $f + g, \lambda f, f \circ g$ sẽ có ma trận tương ứng là $A + B, \lambda A, BA$.

Định lý 2.2.8. *Giả sử $f \in Hom(V, W)$ có ma trận trong cặp cơ sở nào đó là A, khi đó $rank(f) = rank(A)$.*

Chứng minh: Giả sử $\dim V = n, \dim W = m$, A là ma trận của f trong cặp cơ sở $((e), (w))$. Theo định nghĩa

$$rank(f) = \dim \mathrm{Im} f = rank(f(e_1), ..., f(e_n))$$

tức là hạng của các vector cột $f(e_j) = (a_{1j}, ..., a_{mj})^T, j = \overline{1; n}$. Nhưng đây chính là các cột của ma trận A, ta có ngay điều phải chứng minh. ▶

2.2.5 Không gian nghiệm hệ phương trình thuần nhất

Xét hệ phương trình tuyến tính thuần nhất m phương trình n ẩn:

$$\sum_{j=1}^{n} a_{\mathrm{ij}}x_j = 0, x_i \in \mathbb{K}, i = \overline{1; m} \tag{2.9}$$

Hệ này có thể viết dưới dạng ma trận là

$$\text{Ax} = 0, A = (a_{\text{ij}})_{m\times n}, x = (x_1, ..., x_n)^T \tag{2.10}$$

Ký hiệu $\aleph_0$ là tập tất cả các nghiệm của hệ này.

Định lý 2.2.9. *Xét hệ phương trình 2.9 (hoặc 2.10)*
i) Tập $\aleph_0$ tạo thành một không gian vector con của $\mathbb{K}^n$.
ii) Nếu $rank(A) = r$ thì $dim\aleph_0 = n - r$.

Chứng minh: Ta xem A là ma trận của $f : \mathbb{K}^n \to \mathbb{K}^m$ trong cặp cơ sở $((e), (w))$ nào đó, tức là trong cặp cơ sở này ta có

$$f(x) = \text{Ax} = 0$$

Như vậy, tập các nghiệm của hệ 2.10 chính là những phần tử của không gian $Kerf$, do đó $\aleph_0 = Kerf$ và là một không gian con của $\mathbb{K}^n$. Nếu $rank(A) = r$ thì từ Định lý 2.2.8 ta có $rank(A) = dimImf = r$, do đó

$$dim\aleph_0 = dimKerf = dim\mathbb{K}^n - dimImf = n - r$$

Ta có điều phải chứng minh. ►

Một tổ hợp tuyến tính tất cả các vector cơ sở của $\aleph_0$ gọi là *nghiệm tổng quát* của hệ 2.10.

Xét hệ tuyến tính tổng quát

$$\text{Ax} = b \tag{2.11}$$

ở đó $A = (a_{\text{ij}})_{m\times n}, x = (x_1, ..., x_n)^T, b = (b_1, ..., b_m)^T$. Theo ký hiệu chương trước ma trận mở rộng của hệ sẽ là $\widetilde{A} = (A|b)$. Định lý Cronecker-Capelli khẳng định rằng: hệ (2.11) có nghiệm khi và chỉ khi $rank(A) = rank(\widetilde{A})$. Ta có thể chứng minh định lý một cách đơn giản như sau.

Coi mỗi cột của A là một vector, ký hiệu chúng lần lượt là $c_1, c_2, ..., c_n$. Như vậy hệ phương trình có thể viết lại như là

$$x_1c_1 + x_2c_2 + ... + x_nc_n = b$$

Hệ tương thích tương đương với vector b biểu diễn tuyến tính qua các vector cột $c_1, c_2, ..., c_n$. Tức là hệ các vector $\{c_1, c_2, ..., c_n, b\}$ phụ thuộc tuyến tính. Điều này tương đương với

$$rank(A) = rank(c_1, c_2, ..., c_n) = rank(c_1, c_2, ..., c_n, b) = rank(\widetilde{A})$$

Nếu gọi $\overline{x}$ là nghiệm tổng quát của hệ thuần nhất 2.10, còn a là nghiệm riêng của 2.11 thì nghiệm tổng quát của hệ 2.11 bằng tổng của nghiệm tổng quát hệ thuần nhất 2.10 và nghiệm riêng 2.11, tức là

$$x = \overline{x} + a$$

Tiếp theo chúng ta xét bài toán: Tìm cơ sở, chiều của $Kerf$ và Imf khi biết ma trận của ánh xạ tuyến tính f. Để giải quyết bài toán này ta tiến hành như sau:
Bước 1: Sử dụng biến đổi sơ cấp hàng đưa ma trận A về ma trận hình thang A', giả sử $rank(A) = rank(A') = r$, khi đó $dimImf = r$. Để tìm cơ sở của Imf ta chọn r vector cột của ma trận A độc lập tuyến tính.
Bước 2: Giải hệ phương trình thuần nhất $A'x = 0$, cơ sở không gian nghiệm $\aleph_0$ hệ này chính là cơ sở của $Kerf$.

Ví dụ 66. Tìm cơ sở chiều của $Kerf$ và Imf của ánh xạ tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^3$ cho bởi

$$f\left(x_1, x_2, x_3, x_4\right) = \left(x_1 - x_2 + x_3 - x_4, x_2 + x_3, x_1\right)$$

Ta có ma trận của ánh xạ tuyến tính f là

$$A = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Dễ thấy

$$\det \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} = -2 \neq 0$$

nên $rank(A) = 3$, tức là $dimImf = 3$, chọn ba vector cột đầu tiên trong ma trận A độc lập tuyến tính làm cơ sở của Imf. Tuy nhiên có thể thấy rằng $Imf = \mathbb{R}^3$, vì vậy hoàn toàn có thể lấy cơ sở Imf chính là cơ sở chính tắc của $\mathbb{R}^3$, tức là

$$Imf = \left\langle (1, 0, 0)^T;\ (0, 1, 0)^T;\ (0, 0, 1)^T \right\rangle$$

Vì $dimImf = 3$ nên $dimKerf = 4 - 3 = 1$. Giải hệ phương trình $Ax = 0$

ta có ngay

$$\begin{cases} x_1 = 0 \\ x_2 = -t \\ x_3 = t \\ x_4 = 2t \end{cases}$$

do đó chọn được cơ sở của $Kerf$ là $a = (0, -1, 1, 2)^T$.

2.2.6 Ma trận của ánh xạ tuyến tính khi đổi cơ sở

Giả sử $f \in Hom(V, W)$, $\dim V = n, \dim W = m$. Trong cặp cơ sở $((e), (w))$, ánh xạ tuyến tính f có ma trận A. Câu hỏi đặt ra là: trong một cặp cơ sở mới $((e'), (w'))$, ánh xạ tuyến tính f có ma trận biểu diễn như thế nào?

Giả sử B là ma trận chuyển cơ sở (e) sang (e'), C là ma trận chuyển cơ sở (w) sang (w'). Từ công thức 2.2, ta có

$$(e') = (e)\,B; \quad (\mathrm{w}') = (\mathrm{w})\,C$$

Nếu gọi $\widetilde{A}$ là ma trận của ánh xạ tuyến tính f trong cặp cơ sở mới $((e'), (w'))$, thì theo công thức 2.5 chúng ta có

$$(f\,(e)) = (\mathrm{w})\,A$$

và

$$(f\,(e')) = (\mathrm{w'})\,\widetilde{A}$$

ở đó $(f(e)) = (f(e_1), ..., f(e_n))$, $(f(e')) = (f(e'_1), ..., f(e'_n))$. Từ đây chúng ta có

$$(f\,(e')) = (\mathrm{f}\,(e))\,B = (\mathrm{w})\,AB$$

do vậy

$$(\mathrm{w'})\,\widetilde{A} = (\mathrm{w})\,AB \Rightarrow (\mathrm{w})\,C\widetilde{A} = (\mathrm{w})\,AB \Rightarrow C\widetilde{A} = AB$$

từ đó ta có

$$\widetilde{A} = C^{-1}AB \tag{2.12}$$

Đặc biệt, giả sử $f \in End(V)$ là toán tử tuyến tính và A là ma trận của f trong cơ sở $\{e_1, ..., e_n\}$, còn $\widetilde{A}$ là ma trận của f trongcơ sở $\{e'_1, ..., e'_n\}$. Nếu ma trận chuyên cơ sở $\{e_1, ..., e_n\}$ sang $\{e'_1, ..., e'_n\}$ là C thì ta có

$$\widetilde{A} = C^{-1}AC \tag{2.13}$$

Khi xét hai ma trận cùng cỡ, đôi lúc chúng ta quan tâm tới các tính chất "tương tự" giữa chúng.

i) Ma trận tương đương: Hai ma trận $A, B \in M_{m\times n}(\mathbb{K})$ gọi là *tương đương*, viết là $A \sim B$ nếu tồn tại các ma trận vuông $P \in M_m(\mathbb{K})$ và $Q \in M_n(\mathbb{K})$ khả nghịch để $B = PAQ$. Dễ thấy quan hệ này thỏa mãn định nghĩa về quan hệ tương đương nói ở chương đầu tiên.

ii) Ma trận đồng dạng: Hai ma trận vuông $A, B \in M_n(\mathbb{K})$ gọi là *đồng dạng*, viết là $A \sim B$ nếu tồn tại ma trận vuông $P \in M_n(\mathbb{K})$ khả nghịch để $B = P^{-1}AP$. Dễ thấy quan hệ đồng dạng cũng là quan hệ tương đương nhưng mạnh hơn quan hệ tương đương.

Như vậy hai ma trận của ánh xạ tuyến tính (toán tử tuyến tính) trong hai cơ sở khác nhau là tương đương (đồng dạng) với nhau.

Ví dụ 67. Giả sử toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ trong cơ sở chính tắc $(e) = (e_1, e_2, e_3)$ cho bởi

$$f(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, x_1 + x_2 + 2x_3, x_1 + x_2 + x_3)$$

Ta hãy tìm ma trận của f trong cơ sở $(e_1 + e_2, e_2 + e_3, e_1 + e_3)$.

Dễ thấy,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

Đặt $(e') = (e'_1, e'_2, e'_3) = (e_1 + e_2, e_2 + e_3, e_1 + e_3)$, dễ thấy rằng $(e') = (e)C$, ở đó

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Gọi $\widetilde{A}$ là ma trận của ánh xạ tuyến tính f trong cơ sở (e'), áp dụng công thức 2.13 ta có ngay

$$\widetilde{A} = C^{-1}AC = \begin{pmatrix} \frac{3}{2} & 3 & \frac{5}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{3}{2} & 2 & \frac{3}{2} \end{pmatrix}$$

Cũng có thể làm ngắn gọn như sau. Áp dụng công thức 2.5 thì

$$(f(e')) = (e')\widetilde{A}$$

Với $(f(e')) = (f(e'_1), f(e'_2), f(e'_3))$ (vì f tuyến tính dễ dàng tính được) và $(e') = (e'_1, e'_2, e'_3)$, ta xem là các ma trận mà mỗi cột là tọa độ của $f(e'_j)$ hoặc e_j $(j = \overline{1;3})$ trong cơ sở (e), tức là

$$F = (f(e')) = \begin{pmatrix} 3 & 5 & 4 \\ 2 & 3 & 3 \\ 2 & 2 & 2 \end{pmatrix}, T = (e') = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Do đó

$$\widetilde{A} = T^{-1}F = \begin{pmatrix} \frac{3}{2} & 3 & \frac{5}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{3}{2} & 2 & \frac{3}{2} \end{pmatrix}$$

2.3 Trị riêng và vector riêng

2.3.1 Trị riêng và vector riêng của toán tử tuyến tính

Giả sử V là $\mathbb{K}$ - không gian vector, $f \in End(V)$ là toán tử tuyến tính, W là một không gian vector con của V. W được gọi là *bất biến* đối với f hay là f - *bất biến* nếu như $f(W) \subset W$, tức là

$$f(\mathrm{w}) \in \mathrm{W}, \forall \mathrm{w} \in \mathrm{W}$$

Ví dụ 68. V và $\{0\}$ là các không gian con f - bất biến tầm thường. Ngoài ra, dễ dàng kiểm tra $Kerf$ và Imf cũng các không gian con f - bất biến.

Vector $v \in V$, $v \neq 0$ được gọi là *vector riêng* của toán tử tuyến tính f nếu tồn tại $\lambda \in \mathbb{K}$ sao cho $f(v) = \lambda v$. Phần tử λ gọi là *trị riêng* của toán tử tuyến tính f, khi đó ta nói v là vector riêng ứng với trị riêng λ.

Mệnh đề 2.3.1. *Không gian vector sinh bởi $v \neq 0$ là f - bất biến khi và chỉ khi v là vector riêng của f.*

Chứng minh: Đặt $L = \langle v \rangle$ là không gian con sinh bởi $v \neq 0$, giử sử L là f - bất biến. Khi đó theo định nghĩa $f(L) \subset L$. Do L là không gian một chiều nên tồn tại $\lambda \in \mathbb{K}$ để $f(v) = \lambda v \in L$. Điều này có nghĩa v là một vector riêng.

Ngược lại, nếu $v \neq 0$ là vector riêng ứng trị riêng $\lambda \in \mathbb{K}$ thì với vector bất kỳ $a \in Ł, a \neq 0$ ta có $a = \alpha v$, ở đó $\alpha \in \mathbb{K}$. Ta có

$$f(a) = f(\alpha v) = \alpha f(v) = \alpha \lambda v = \lambda a \in L$$

nói cách khác L là f - bất biến ▶

Rõ ràng từ chứng minh trên ta thấy, nếu v là vector riêng riêng ứng với trị riêng λ thì $\alpha v (\alpha \neq 0)$ cũng là vector riêng ứng với trị riêng λ.

Đặt

$$V(\lambda) = \{x \in V : f(x) = \lambda x\}.$$

Dễ thấy $V(\lambda)$ là không gian con bất biến của f, hơn thế $V(\lambda) \neq \{0\}$ khi và chỉ khi λ là trị riêng của f. Thật vậy, do $f - \lambda Id_V$ cũng là một ánh xạ tuyến tính và $V(\lambda) \equiv \mathrm{Ker}\,(f - \lambda Id_V)$ nên khẳng định đầu là hiển nhiên. Khẳng định còn lại thu được từ định nghĩa của vector riêng.

Khi $\lambda \in \mathbb{K}$ là một trị riêng của toán tử tuyến tính f thì $V(\lambda)$ gọi là *không gian con riêng* ứng với trị riêng λ.

Giả sử trong cơ sở $\{e_1, ..., e_n\}$ của V toán tử tuyến tính f có ma trận A. Từ lý thuyết ánh xạ tuyến tính ở trên dẫn ra rằng, với mọi $\lambda \in \mathbb{K}$ toán tử tuyến tính $f - \lambda Id_V$ trong cơ sở $\{e_1, ..., e_n\}$ có ma trận là $A - \lambda E$, ở đó E là ma trận đơn vị (ánh xạ đồng nhất trong mọi cơ sở đều có ma trận là E). Nếu λ là một trị riêng của f thì vector $x \neq 0$ là vector riêng ứng với nó khi và chỉ khi $x \in V(\lambda) = Ker\,(f - \lambda Id_V)$, cũng tức là x thỏa mãn hệ phương trình

$$(A - \lambda E)\, x = 0 \tag{2.14}$$

Như vậy $x = (x_1, ..., x_n)^T \neq 0$ là vector riêng của f khi và chỉ khi nó là nghiệm không tầm thường của hệ 2.14. Điều này xảy ra khi và chỉ khi

$$\det\,(A - \lambda E) = 0 \tag{2.15}$$

Có thể thấy $\det\,(A - \lambda E)$ là một đa thức bậc n, đa thức này không thay đổi dù chúng ta đổi cơ sở của V.

Thật vậy, giả sử trong có sở mới $\{e'_1, ..., e'_n\}$ của V toán tử tuyến tính có ma trận là $\widetilde{A}$, ma trận chuyển cơ sở từ $\{e_1, ..., e_n\}$ sang $\{e'_1, ..., e'_n\}$ là C khi đó $\widetilde{A} = C^{-1}AC$. Ta có

$$\begin{aligned} \det\left(\widetilde{A} - \lambda E\right) &= \det\left(C^{-1}AC - \lambda C^{-1}EC\right) \\ &= \det\left(C^{-1}\,(A - \lambda E)\,C\right) = \det\,(A - \lambda E) \end{aligned}$$

Đa thức $\det\,(A - \lambda E)$ được gọi là *đa thức đặc trưng* của toán tử tuyến tính f (hay của ma trận A). Phương trình 2.15 gọi là *phương trình đặc trưng.*

Giả sử $A = (a_{ij})_{n\times n}$ thì

$$P_n(\lambda) = \det(A - \lambda E) = (-1)^n\lambda^n + (-1)^{n-1}Trace(A)\lambda^{n-1} + ... + \det(A)$$

Định lý 2.3.1. *Phần tử $\lambda \in \mathbb{K}$ là trị riêng của toán tử tuyến tính f khi và chỉ khi λ là nghiệm phương trình đặc trưng 2.15.*

Chứng minh: λ là trị riêng của f khi và chỉ khi $V(\lambda) = Ker(f - \lambda Id_V) \neq \{0\}$, điều này tức là $dimKer(f - \lambda Id_V) > 0$. Nhưng như vậy thì

$$rank(f - \lambda Id_V) = n - \dim Ker(f - \lambda Id_V) < n$$

nó tương đương với $rank(A - \lambda E) < n$. Mà điều nảy xảy ra khi và chỉ khi $P_n(\lambda) = \det(A - \lambda E) = 0$, tức là λ là nghiệm phương trình 2.15. ▶

Định nghĩa 38. *Tập tất cả các trị riêng của toán tử tuyến tính f gọi là phổ của nó, ký hiệu là $\sigma(f)$.*

Bài toán tìm trị riêng và vector riêng ta giải quyết như sau:
Bước 1: Lập đa thức đặc trưng $P_n(\lambda)$, giải phương trình đặc trưng $P_n(\lambda) = 0$.
Bước 2: Với mỗi λ tìm được trong bước 1 ta giải hệ phương trình $(A-\lambda E)x = 0$ để tìm vector riêng x ứng với trị riêng λ.

Ví dụ 69. Cho toán tử tuyến tính $f \in End(\mathbb{R}^3)$ có ma trận chính tắc là

$$A = \begin{pmatrix} 3 & -2 & -2 \\ 3 & -1 & -3 \\ 1 & -2 & 0 \end{pmatrix}$$

Hãy tìm trị riêng và vector riêng của f.

Ta có phương trình đặc trưng của f là

$$\det(A - \lambda E) = -\lambda^3 + 2\lambda^2 + \lambda - 2 = 0$$

Giải phương trình này được các trị riêng là ± 1, 2. Với $\lambda_1 = 1$ tìm được vector riêng $v_1 = (1, 0, 1)^T$, với $\lambda_2 = -1$ tìm được vector riêng $v_2 = (1, 1, 1)^T$, với $\lambda_3 = 2$ tìm được vector riêng $v_3 = (0, 1, -1)^T$.

2.3.2 Chéo hóa ma trận

Để nghiên cứu toán tử tuyến tính $f \in End(V)$ ta nghiên cứu ma trận của nó, ma trận của f càng đơn giản thì hình dung về f sẽ càng dễ hiểu. Nhiệm vụ của chúng ta là tìm một cơ sở của V để trong cơ sở này ma trận của f có dạng đơn giản nhất có thể. Cụ thể ở đây chúng ta tìm cơ sở của V để f có ma trận dạng đường chéo.

Toán tử tuyến tính f gọi là *chéo hóa được* nếu trong V tìm được (ít nhất) một cơ sở để trong đó ma trận của f có dạng chéo. Việc tìm một cơ sở như vậy gọi là chéo hóa f.

Định lý 2.3.2. *Toán tử tuyến tính $f \in End(V)$ có ma trận chéo khi và chỉ khi tồn tại một cơ sở V gồm toàn vector riêng.*

Dễ dàng được chứng minh định lý này dựa trên định nghĩa ma trận của toán tử tuyến tính.

Mệnh đề 2.3.2. *Các vector riêng ứng với trị riêng khác nhau của toán tử tuyến tính f là độc lập tuyến tính.*

Chứng minh: Ta chứng minh bằng quy nạp.

Với $k = 1$ ta có một vector riêng $v_1 \neq 0$, do đó độc lập tuyến tính, khẳng định của mệnh đề đúng.

Giả sử mệnh đề đúng với $k = m$ tức là ta có m vector riêng $v_1, ..., v_m$ ứng với các trị riêng đôi một khác nhau $\lambda_1, ..., \lambda_m$ độc lập tuyến tính. Ta phải chứng minh đúng với $k = m + 1$, tức là hệ các vector riêng $\{v_1, ..., v_{m+1}\}$ ứng với các trị riêng đôi một khác nhau $\lambda_1, ..., \lambda_{m+1}$ độc lập tuyến tính. Xét tổ hợp tuyến tính

$$\alpha_1 e_1 + ... + \alpha_m e_m + \alpha_{m+1} e_{m+1} = 0 \tag{2.16}$$

tác động toán tử f vào hai vế của 2.16 và sử dụng giả thiết các vector riêng ta có

$$\alpha_1 \lambda_1 e_1 + ... + \alpha_m \lambda_m e_m + \alpha_{m+1} \lambda_{m+1} e_{m+1} = 0 \tag{2.17}$$

Nhân hai vế của 2.16 với λ_{m+1} rồi trừ cho 2.17 ta có

$$\alpha_1 \left(\lambda_{m+1} - \lambda_1\right) e_1 + ... + \alpha_m \left(\lambda_{m+1} - \lambda_m\right) e_m = 0$$

Nhưng $\{e_1, ..., e_m\}$ độc lập tuyến tính và các trị riêng đôi một khác nhau nên

$$\alpha_1 = ... = \alpha_m = 0$$

Từ đó thay vào 2.16 ta có $\alpha_{m+1} = 0$, tức là hệ $m+1$ vector riêng $\{v_1, ..., v_{m+1}\}$ độc lập tuyến tính. Theo nguyên lý quy nạp ta có điều phải chứng minh. ▶
Ta có một hệ quả hiển nhiên như sau.

Hệ quả 2.3.1. *Nếu toán tử tuyến tính f có n trị riêng phân biệt $\lambda_1, ..., \lambda_n$ thì f chéo hóa được và ma trận của f trong cơ sở gồm toàn vector riêng tương ứng với các trị riêng này là $A = diag(\lambda_1, ..., \lambda_n)$.*

Như vậy hệ quả trên cho ta một điều kiện đủ (không phải điều kiện cần) để chéo hóa một toán tử tuyến tính.

Mệnh đề 2.3.3. *Nếu λ_0 là nghiệm bội k của phương trình đặc trưng 2.15 thì ứng với nó có không quá k vector riêng độc lập tuyến tính.*

Chứng minh: Giả sử λ_0 là nghiệm bội $k \geq 1$ của phương trình đặc trưng 2.15 của toán tử tuyến tính f, tức là

$$\det\left(A - \lambda E\right) = \left(\lambda - \lambda_0\right)^k p\left(\lambda\right)$$

Ngoài ra, giả sử rằng ứng với λ_0 có m vector riêng độc lập tuyến tính là $v_1, ..., v_m$, hiển nhiên là $m \leq n = dimV$. Ta bổ sung thêm $n - m$ vector nữa để $v_1, ..., v_m, v_{m+1}, ..., v_n$ trở thành cơ sở của V. Trong cơ sở này, giả sử toán tử tuyến tính f có ma trận biểu diễn là B, rõ ràng

$$\det\left(B - \lambda E\right) = \left(\lambda - \lambda_0\right)^m q\left(\lambda\right)$$

Nhưng phương trình đặc trưng bất biến qua phép đổi cơ sở nên

$$\left(\lambda - \lambda_0\right)^m q\left(\lambda\right) = \det\left(A - \lambda E\right) = \left(\lambda - \lambda_0\right)^k p\left(\lambda\right)$$

Từ đó ta có $\left(\lambda - \lambda_0\right)^m$ phải là ước của $\det\left(A - \lambda E\right)$, tức là $m \leq k$. Ta có điều phải chứng minh. ►

Nếu toán tử tuyến tính f chéo hóa được thì nó phải có đủ n trị riêng kể cả bội. Đó là điều kiện cần để f chéo hóa được. Như thế giả sử f có đủ n trị riêng kể cả bội, để chéo hóa được f thì ứng với trị riêng nào đó có bội là m ta phải tìm được đủ m vector riêng độc lập tuyến tính.

Từ các kết quả trên ta đưa ra các bước giải bài toán chéo hóa toán tử tuyến tính f (hay ma trận của nó).

Bước 1: Giải phương trình đặc trưng tìm tất cả các trị riêng. Giả sử giải được các trị riêng $\lambda_1, ..., \lambda_k$ với bội tương ứng $n_1, ..., n_k$.

Bước 2: Kiểm tra điều kiện cần:

- Nếu $k = n$ toán tử f chéo hóa được.
- Nếu $n_1 + n_2 + ... + n_k < n$ toán tử f không chéo hóa được.

- Nếu $n_1 + n_2 + ... + n_k = n$ tìm các vector riêng độc lập tuyến tính ứng với mỗi trị riêng. Nếu tồn tại λ_i mà số vector riêng ứng với nó nhỏ hơn n_i thì f không chéo hóa được.

Bước 3: Nếu f chéo hóa được và $\{v_{11}, ..., v_{1n_1}, ..., v_{k1}, ..., v_{kn_k}\}$ là cơ sở gồm toàn vector riêng, đặt ma trận P bằng các cột tọa độ của những vector riêng theo thứ tự này

$$P = (v_{11}, ..., v_{1n_1}, ..., v_{k1}, ..., v_{kn_k})$$

Khi đó ma trận của toán tử f trong cơ sở này là

$$\overline{A} = P^{-1}AP = \begin{pmatrix} \lambda_1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \lambda_1 & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \lambda_k & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \lambda_k \end{pmatrix}$$

Chú ý rằng toán tử tuyến tính f có thể chéo hóa được trên $\mathbb{K} = \mathbb{C}$ nhưng lại không chéo hóa được trên $\mathbb{K} = \mathbb{R}$.

Ví dụ 70. Cho ma trận chính tắc của toán tử tuyến tính $f \in End(\mathbb{R}^3)$ là

$$A = \begin{pmatrix} -14 & -28 & -44 \\ -7 & -14 & -23 \\ 9 & 18 & 29 \end{pmatrix}$$

Hãy chéo hóa f nếu được.

Phương trình đặc trưng là

$$\det(A - \lambda E) = -\lambda^3 + \lambda^2 + 2\lambda = 0$$

Giải phương trình này được các trị riêng khác nhau $\lambda_1 = -1,\ \lambda_2 = 0,\ \lambda_3 = 2$, do đó f chéo hóa được. Các vector riêng tương ứng là

$$v_1 = \begin{pmatrix} -8 \\ -1 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

từ đó ta có

$$P = (v_1, v_2, v_3) = \begin{pmatrix} -8 & -2 & -1 \\ -1 & 1 & -1 \\ 3 & 0 & 1 \end{pmatrix}, \overline{A} = P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Việc đưa toán tử tuyến tính hay ma trận về dạng đường chéo có rất nhiều ích lợi trong việc tính toán. Ví dụ, một ma trận dạng đường chéo khi lấy lũy thừa thì chỉ cần lũy thừa các phần tử đường chéo, tức là

$$(diag\,(\lambda_1, ..., \lambda_n))^m = diag\,(\lambda_1^m, ..., \lambda_n^m)$$

Ví dụ 71. Dãy Fibonacci là một dãy số tự nhiên $\{a_n\}_{n\geq 0}$ bắt đầu bằng $a_0 = 0$ và $a_1 = 1$, mỗi số sau đó bằng tổng hai số đứng ngay trước nó. Có thể mô tả dãy này dưới dạng một *hệ động lực* như sau

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}$$

với $n \geq 1$.

Đặt

$$F = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

khi đó ta thấy rằng

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = F \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = F^2 \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} = ... = F^n \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}$$

Đặt

$$v_0 = \begin{pmatrix} a_1 \\ a_0 \end{pmatrix}, v_n = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$$

thì dãy Fibonacci có thể viết lại là $v_n = F^n v_0$.

Tổng quát hơn, xét một hệ động lực

$$v_n = Av_{n-1}$$

trong đó $A \in M_m(\mathbb{K}), v \in \mathbb{K}^m$. Khi đó tương tự như trên, nếu cố định $v_0 = (\alpha_1, ..., \alpha_m)^T$ thì ta cũng thu được

$$v_n = A^n v_0$$

Nếu ma trận A có dạng chéo, $A = diag(\lambda_1, ..., \lambda_m)$ thì

$$A^n = diag(\lambda_1^n, ..., \lambda_m^n)$$

Khi đó

$$v_n = (\lambda_1^n \alpha_1, ..., \lambda_m^n \alpha_m)$$

Trở lại với dãy $Fibonacci$, ta sẽ chéo hóa ma trận F. Phương trình đặc trưng của F là

$$\lambda^2 - \lambda + 1 = 0$$

Giải phương trình này ta được các trị riêng

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2}$$

Dễ thấy ứng với trị riêng λ_1 ta có vector riêng là $(\lambda_1, 1)^T$, ứng với trị riêng λ_2 ta có vector riêng là $(\lambda_2, 1)^T$. Đặt

$$P = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}$$

thì ta có

$$F = P \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} P^{-1}$$

và

$$F^n = P \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix} P^{-1}$$

Dễ dàng tính được

$$a^n = \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} = \frac{1}{\sqrt{5}.2^n}\left(\left(1+\sqrt{5}\right)^n - \left(1-\sqrt{5}\right)^n\right)$$

Từ đây rút ra

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{\lambda_1^{n+1} - \lambda_2^{n+1}}{\lambda_1^n - \lambda_2^n} = \lambda_1 \approx 1.6$$

Điều này có nghĩa là với n lớn, hai số liên tiếp trong dãy sẽ tỉ lệ nhau với hằng số tỉ lệ λ_1. Giới hạn trên liên quan tới *tỉ lệ vàng* xuất hiện rất nhiều trong các hiện tượng tự nhiên và kỹ thuật.

2.4 Thực hành tính toán trên Maple

Trước hết khởi động Maple và dùng gói lệnh *linalg*, khai báo ma trận như trong chương trước

- Để tìm ma trận đặc trưng $\lambda E - A$ ta dùng lệnh *charmat(A,lambda);*

- Tính đa thức đặc trưng dùng lệnh *charpoly(A,lambda);*

- Tìm trị riêng bởi lệnh *eigenvals;*

- Tìm vector riêng của ma trận bằng lệnh *eigenvects(A);* Sau khi dùng lệnh này máy sẽ cho kêt quả trong các móc vuông ([]), mỗi móc vuông sẽ có các thông tin về trị riêng, bội đại số của trị riêng, các vector cơ sở của không gian con riêng ứng với trị riêng.

- Tìm cơ sở của không gian con bằng lệnh *basis;*

- Tìm cơ sở cho không gian hạch của toán tử tuyến tính (ma trận) A bằng lệnh *kernel(A);*

Ví dụ 72.

$> restart;$

$> with(linalg);$

BlockDiagonal, GramSchmidt, JordanBlock, LUdecomp, QRdecomp, Wronskian, addcol, addrow, adj, adjoint, angle, augment, backsub, band, basis, bezout, blockmatrix, charmat, charpoly, cholesky, col, coldim, colspace, colspan, companion, concat, cond, copyinto, crossprod, curl, definite, delcols, delrows, det, diag, diverge, dotprod, eigenvals, eigenvalues, eigenvectors, eigenvects, entermatrix, equal, exponential, extend, ffgausselim, fibonacci, forwardsub, frobenius, gausselim, gaussjord, geneqns, genmatrix, grad, hadamard, hermite, hessian, hilbert, htranspose, ihermite, indexfunc, innerprod, intbasis, inverse, ismith, issimilar, iszero, jacobian, jordan, kernel, laplacian, leastsqrs, linsolve, matadd, matrix, minor, minpoly, mulcol, mulrow, multiply, norm, normalize, nullspace, orthog, permanent, pivot, potential, randmatrix, randvector, rank, ratform, row, rowdim, rowspace, rowspan, rref, scalarmul, singularvals, smith, stackmatrix, submatrix, subvector, sumbasis, swapcol, swaprow, sylvester, toeplitz, trace, transpose, vandermonde, vecpotent, vectdim, vector, wronskian

$> A := array([[1,2,3],[1,1,2],[1,1,1]]);$

$$A := \begin{bmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{bmatrix}$$

$> charmat(A, lambda);$

$$A := \begin{bmatrix} \lambda - 4 & 5 & -2 \\ -5 & \lambda - 7 & -3 \\ -6 & 9 & \lambda - 4 \end{bmatrix}$$

$> charpoly(A, lambda);$

$$\lambda^3 - \lambda^2$$

$> eigenvals(A);$

$$0, 0, 1$$

$> eigenvects(A);$

$$[1, 1, \{[1\ 1\ 1]\}], [0, 2, \{[1\ 2\ 3]\}]$$

$> kernel(A);$

$$\{[1\ 2\ 3]\}$$

$> v1 := vector([1,2,3,4]);$

$$v1 := [1\ 2\ 3\ 4]$$

$> v2 := vector([-1,0,2,-2]);$

$$v2 := [-1\ 0\ 2\ \ -2]$$

$> v := vector([2,1,-3,-2]);$

$$v3 := [2\ 1\ \ -3\ \ -2]$$

$> v4 := vector([2,3,2,0]);$

$$v4 := [2\ 3\ 2\ 0]$$

$> basis(v1, v2, v3, v4);$

$$\{v1, v2, v3\}$$

Chương 3

Hình học trong không gian Euclide

3.1 Dạng toàn phương trong không gian vector

3.1.1 Dạng song tuyến tính đối xứng và dạng toàn phương

Giả sử V là $\mathbb{R}$ - không gian vector.

Định nghĩa 39. *Một ánh xạ $\psi : V \times V \to \mathbb{R}$ gọi là dạng song tuyến trên V nếu ψ là hàm tuyến tính với từng biến khi biến kia cố định, tức là*

$$\psi(\lambda x + \beta x', y) = \lambda\psi(x, y) + \beta\psi(x', y)$$
$$\psi(x, \lambda y + \beta y') = \lambda\psi(x, y) + \beta\psi(x, y')$$

với mọi $x, x', y, y' \in V; \lambda, \beta \in \mathbb{R}$. Ngoài ra, ψ sẽ gọi là dạng song tuyến tính đối xứng nếu thoả mãn thêm điều kiện

$$\psi(x, y) = \psi(y, x)$$

với mọi $x, y \in V$.

Định nghĩa 40. *Ánh xạ $q : V \to \mathbb{R}$ gọi là một dạng toàn phương trên V nếu có một dạng song tuyến tính ψ trên V sao cho*

$$q(x) = \psi(x, x), \forall x \in V$$

Nếu ψ là một dạng song tuyến tính đối xứng thì nó được gọi là *dạng cực* của dạng toàn phương q.

Từ định nghĩa dạng toàn phương ta thấy rằng

$$q(\lambda x) = \lambda^2 q(x), \forall x \in V, \lambda \in \mathbb{R}$$

Ví dụ 73. Giả xử $f_i : V \to \mathbb{R}$ với $i = 1; 2$ là các hàm tuyến tính. Dễ thấy $\psi(x, y) = f_1(x)f_2(y)$ là một dạng song tuyến tính, $q(x) = f_1(x)f_2(x)$ là một dạng toàn phương.

Ví dụ 74. Hai dạng song tuyến tính ψ_1, $\psi_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ cho bởi

$$\psi_1 (x, y) = x_1y_1 - x_1y_2 + 3x_2y_2$$
$$\psi_2 (x, y) = x_1y_1 + 2x_1y_2 - 3x_2y_1 + 3x_2y_2$$

xác định cùng một dạng toàn phương

$$q (x) = x_1^2 - x_1x_2 + 3x_2^2$$

Ví dụ 75. Giả sử $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ sao cho $\psi (x, y) = \sum\limits_{i=1}^{n} x_iy_i$. Dễ thấy ψ là một dạng xong tuyến tính đối xứng, $q(x) = \psi(x, x)$ là dạng toàn phương.

Ví dụ 76. Giả sử $V = C[a, b]$ là không gian các hàm liên tục trên đoạn $[a, b]$, xét ánh xạ $\psi : V \times V \to \mathbb{R}$ cho bởi

$$\psi (x, y) = \int\limits_a^b x (t)y (t)\, dt$$

Khi đó ψ là một dạng song tuyến tính đối xứng, còn

$$q (x) = \int\limits_a^b x^2 (t)dt$$

là một dạng toàn phương.

Giả sử V là $\mathbb{R}$ - không gian vector n chiều, ψ là một dạng song tuyến tính đối xứng trên V, q là dạng toàn phương xác định bởi ψ. Giả sử $\{e_1, ..., e_n\}$ là một cơ sở của V. Với mọi $x, y \in V$ thì

$$x = \sum_{i=1}^{n} x_ie_i,\ y = \sum_{j=1}^{n} y_je_j$$

Ta có

$$\psi (x, y) = \psi \left(\sum_{i=1}^{n} x_ie_i, \sum_{j=1}^{n} y_je_j \right) = \sum_{i,j=1}^{n} x_iy_j\psi (e_i, e_j) \tag{3.1}$$

và

$$q (x) = \sum_{i,j=1}^{n} x_ix_j\psi (e_i, e_j) \tag{3.2}$$

Đặt $\psi (e_i, e_j) = a_{\mathrm{ij}}, i, j = \overline{1; n}$, ma trận $A = (a_{ij})_{n\times n} \in M_n(\mathbb{R})$ gọi là ma trận của ψ (hay q) trong cơ sở đã cho. Vì ψ là dạng song tuyến tính đối xứng nên $a_{ij} = a_{ji}$, tức là A là ma trận đối xứng. Do tọa độ của vector luôn được hiểu là cho dưới dạng cột, ta có thể viết lại ψ và q như là

$$\psi\left(x, y\right) = x^T A y \tag{3.3}$$

và

$$q\left(x\right) = x^T A x \tag{3.4}$$

Định lý sau đây cho ta mối liên hệ giữa một dạng toàn phương và dạng cực của nó.

Định lý 3.1.1. *Với mỗi dạng toàn phương $q(x)$ trong V, tồn tại duy nhất dạng song tuyến tính đối xứng $\psi(x, y)$ là dạng cực của $q(x)$.*

Như vậy, dạng toàn phương và dạng cực của nó thiết lập một song ánh $1:1$.

Chứng minh: Rõ ràng một dạng song tuyến tính đối xứng $\psi(x, y)$ thiết lập một dạng toàn phương tương ứng $q(x)$.

Ngược lại, giả sử có một dạng toàn phương $q(x)$, ta xác định ψ bởi công thức duy nhất

$$\psi\left(x, y\right) = \frac{q\left(x + y\right) - q\left(x\right) - q\left(y\right)}{2}, \forall x, y \in \mathbb{V}$$

Dễ dàng kiểm tra được đây là một dạng song tuyến tính đối xứng. ▶

Ở trên ta đã xét dạng song tuyến tính và dạng toàn phương trong cơ sở $\{e_1, ..., e_n\}$ của V được biểu diễn bởi các công thức 3.3 và 3.4. Giả sử trong V có một cơ sở khác là $\{e'_1, ..., e'_n\}$, khi đó ma trận dạng song tuyến tính và dạng toàn phương mới $\overline{A}$ biểu diễn như thế nào?

Gọi C là ma trận chuyển cơ sở từ (e) sang (e'), tức là $(e') = (e)C$, như đã biết

$$x_{(e)} = C x_{(e')};\ y_{(e)} = C y_{(e')}$$

do đó trong cơ sở mới ta có

$$\begin{aligned}\psi\left(x, y\right) &= x_{(e)}^T A y_{(e)} = x_{(e')}^T C^T A C y_{(e')}\\ q\left(x\right) &= x_{(e)}^T A x_{(e)} = x_{(e')}^T C^T A C x_{(e')}\end{aligned}$$

tức là $\overline{A} = C^T A C$.

Vì $\det(C) \neq 0$ nên $rank(A) = rank(\overline{A})$.

Giả sử ψ là dạng song tuyến tính đối xứng xác định dạng toàn phương q trên V. Ta nói vector $x \in V$ là *trực giao* với $y \in V$ nếu như $\psi(x, y) = 0$, ký hiệu là $x \perp y$, và đọc x là ψ - trực giao với y (hay q - trực giao) hoặc đơn giản là trực giao nếu dạng song tuyến tính và dạng toàn phương đã xác định.

Nếu x trực giao với chính nó thì ta gọi là ψ - *đẳng hướng* (hay q - *đẳng hướng*).

Giả sử $S \subset V$ là tập con khác rỗng. Vector x gọi là *trực giao với* S nếu nó trực giao với mọi vector trong S, viết là $x \perp S$. Ký hiệu

$$S^{\perp} = \{x \in V : x \perp S\}$$

dễ dàng chứng minh được $S^{\perp}$ là không gian con của V.

Không gian con $V^{\perp}$ của V được gọi là *nhân* của dạng song tuyến tính đối xứng ψ (hoặc q), ký hiệu là $Ker\psi$ (hoặc $Kerq$). Như vậy

$$\mathrm{Ker}\psi \equiv \mathrm{Ker}q = \{x \in V : \psi\left(x, y\right) = 0, \forall y \in V\}$$

Hạng của dạng song tuyến tính đối xứng ψ (hoặc dạng toàn phương q) ký hiệu là $rank(\psi)$ (hoặc $rank(q)$) xác định bởi

$$rank(\psi) \equiv \mathrm{rank}(q) = n - dimKer\psi = n - dimKerq$$

Nếu $rank(\psi) = rank(q) = n$ ta nói ψ và q là *không suy biến*, ngược lại ta nói chúng *suy biến*.

Vậy ψ và q không suy biến khi và chỉ khi $dimKer\psi = dimKerq = 0$ hay là $\mathrm{Ker}\psi \equiv \mathrm{Ker}q = \{0\}$.

Giả sử $\{e_1, ..., e_n\}$ là cơ sở của V, trong cơ sở này ψ và q có ma trận là $A = (a_{ij})_{n\times n}$, ở đó $a_{ij} = \psi(e_i, e_j), \forall i, j = \overline{1; n}$. Ta có

$$x \in K\mathrm{er}\psi = V^{\perp} \Leftrightarrow x \perp y, \forall y \in V$$

nhưng trong cơ sở (e) ta có $y = \sum\limits_{i=1}^{n} y_i e_i$, do đó

$$x \perp e_i, \forall i = \overline{1; n}$$

hay

$$\psi\left(x, e_i\right) = 0, \forall i = \overline{1; n}$$

Mà trong cơ sở (e) thì $x = \sum\limits_{j=1}^{n} x_j e_j$, do vậy

$$\psi(x, e_i) = \sum_{i=1}^{n} x_j \psi(e_j, e_i) = \sum_{j=1}^{n} a_{\text{ij}} x_j = 0, \forall i = \overline{1; n}$$

Đây là hệ phương trình thuần nhất xác định $Ker\psi = Kerq = V^{\perp}$. Hiển nhiên là $dimKer\psi = n - rank(A)$ bởi vì $Ker\psi$ là không gian nghiệm hệ thuần nhất này.

Vậy theo định nghĩa hạng của dạng song tuyến tính đối xứng ở trên ta có

$$rank(\psi) \equiv rank(q) = rank(A) \tag{3.5}$$

3.1.2 Đưa dạng toàn phương về dạng chính tắc

Từ định lý 3.1.1 ở trên, giữa các dạng toàn phương và dạng song tuyến tính đối xứng cực của nó có một song ánh nên từ nay trở đi chúng ta chỉ nghiên cứu các dạng toàn phương.

Giả sử V là $\mathbb{R}$ - không gian vector n chiều, $\{e_1, ..., e_n\}$ là một cơ sở của V.

Định nghĩa 41. *Dạng toàn phương $q(x)$ trong cơ sở đã cho gọi là dạng chính tắc nếu như trong cơ sở này mà trận của $q(x)$ có dạng đường chéo, tức là*

$$q(x) = \sum_{i=1}^{r} \alpha_i x_i^2 \tag{3.6}$$

ở đó $r = rank(q)$. Khi đó cơ sở được gọi là cơ sở chính tắc của dạng toàn phương $q(x)$.

Định lý 3.1.2. *Mọi dạng toàn phương $q(x)$ trong V đều tồn tại cơ sở chính tắc để $q(x)$ có dạng chính tắc.*

Chứng minh: Giả sử $q(x)$ là một dạng toàn phương trên V, trong cơ sở $\{e_1, ..., e_n\}$ nào đó, $q(x)$ có ma trận biểu biễn là $A = (a_{ij})_{n \times n}, A = A^T$, tức là

$$q(x) = \sum_{i,j=1}^{n} a_{\text{ij}} x_i x_j$$

Ta làm dần theo từng biến. Trước hết ta đưa $q(x) = q(x_1, x_2, ..., x_n)$ về dạng

$$q\left(x_1, ..., x_n\right) = \alpha_1 x_1^2 + q_1\left(x_2, ..., x_n\right)$$

trong đó $q_1(x_1, ..., x_n)$ là dạng toàn phương của $n-1$ biến, sau đó tiếp tục đưa $q_1(x_1, ..., x_n)$ về dạng

$$q_1\left(x_2, ..., x_n\right) = \alpha_2 x_2^2 + q_2\left(x_3, ..., x_n\right)$$

Quá trình trên cứ như vậy cho đến khi đưa được về dạng chính tắc. Ta chỉ cần chứng minh cho bước đầu tiên. Có hai trường hợp phủ định nhau.

Trường hợp 1. $a_{11}, ..., a_{nn}$ không đồng thời bằng 0, nếu cần có thể đánh số lại cơ sở, ta giả thiết rằng $a_{11} \neq 0$. Khi đó

$$q\left(x\right) = a_{11}x_1^2 + 2\mathrm{a}_{12}x_1x_2 + ... + 2\mathrm{a}_{1n}x_1x_n + \sum_{i,j=2}^{n} a_{\mathrm{ij}}x_ix_j$$

Ta viết lại như sau

$$q\left(x\right) = a_{11}\left[x_1^2 + 2x_1\left(\frac{\mathrm{a}_{12}}{a_{11}}x_2 + ... + \frac{\mathrm{a}_{1n}}{a_{11}}x_n\right) + \left(\frac{\mathrm{a}_{12}}{a_{11}}x_2 + ... + \frac{\mathrm{a}_{1n}}{a_{11}}x_n\right)^2\right]$$
$$-a_{11}\left(\frac{\mathrm{a}_{12}}{a_{11}}x_2 + ... + \frac{\mathrm{a}_{1n}}{a_{11}}x_n\right)^2 + \sum_{i,j=2}^{n} a_{\mathrm{ij}}x_ix_j$$

Đặt

$$\begin{cases} x_1' = x_1 + \dfrac{\mathrm{a}_{12}}{a_{11}}x_2 + ... + \dfrac{\mathrm{a}_{1n}}{a_{11}}x_n \\ x_2' = x_2 \\ \vdots \\ x_n' = x_n \end{cases}$$

thì

$$q\left(x\right) = a_{11}{x_1'}^2 + q_1\left(x_2', ..., x_n'\right)$$

Không khó để thấy

$$\begin{cases} x_1 = x_1' - \dfrac{\mathrm{a}_{12}}{a_{11}}x_2' - ... - \dfrac{\mathrm{a}_{1n}}{a_{11}}x_n' \\ x_2 = x_2' \\ \vdots \\ x_n = x_n' \end{cases}$$

và ma trận đổi cơ sở này sẽ là

$$C = \begin{pmatrix} 1 & -\frac{\mathrm{a}_{12}}{a_{11}} & -\frac{\mathrm{a}_{13}}{a_{11}} & \cdots & -\frac{\mathrm{a}_{1n}}{a_{11}} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \det(C) = 1$$

Trường hợp 2. Nếu $a_{11} = ... = a_{nn} = 0$, vì $rank(q) = rank(A) = r > 0$ nên tồn tại a_{ij} nào đó khác 0, nếu cần đánh số lại các vector cơ sở, giả sử $a_{12} \neq 0$. Xét phép đổi biến

$$\begin{cases} x_1 = x'_1 + x'_2 \\ x_2 = x'_1 - x'_2 \\ x_3 = x'_3 \\ \vdots \\ x_n = x'_n \end{cases}$$

Khi đó

$$q(x) = 2\mathrm{a}_{12}x_1x_2 + ... + 2\mathrm{a}_{1n}x_1x_n + \sum_{i,j=2}^{n} a_{\mathrm{ij}}x_ix_j = 2\mathrm{a}_{12}\left({x'_1}^2 - {x'_2}^2\right) + ...$$

tức là ta sẽ đưa về dạng toàn phương trong trường hợp 1. Ở đây ma trận chuyển cơ sở là

$$C = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \det(C) = -2$$

Quá trình trên sẽ lặp (hữu hạn) cho đến khi đưa được $q(x)$ về dạng chính tắc. ▶

Phương pháp trong chứng minh trên gọi là *phương pháp Lagrange.*

Ví dụ 77. Dạng toàn phương trong một cơ sở nào đó của $\mathbb{R}^3$ được cho bởi

$$q(x) = x_1^2 + 2x_1x_2 + 2\mathrm{x}_2^2 + 4\mathrm{x}_2x_3 + 5\mathrm{x}_3^2$$

với ma trận biểu diễn là

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 5 \end{pmatrix}$$

Ta viết lại $q(x)$ như sau

$$q\left(x\right) = \left(x_1 + x_2\right)^2 + \left(x_2 + 2\mathrm{x}_3\right)^2 + \mathrm{x}_3^2$$

Xét phép đổi biến

$$\begin{cases} x_1' = x_1 + x_2 \\ x_2' = x_2 + 2\mathrm{x}_3 \\ x_3' = x_3 \end{cases} \Leftrightarrow \begin{cases} x_1 = x_1' - x_2' + 2x_3' \\ x_2 = x_2' - 2x_3' \\ x_3 = x_3' \end{cases}$$

Từ đó ta có ma trận chuyển cơ sở C

$$C = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, C^T AC = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

và $q(x) = {x_1'}^2 + {x_2'}^2 + {x_3'}^2$.

Ngoài phương pháp Lagrange đưa dạng toàn phương về dạng chính tắc còn có nhiều phương pháp khác nữa, ví dụ phương pháp Jacobi (xem [3],...). Dưới đây ta đưa ra một phương pháp biến đổi sơ cấp ma trận đối xứng để đưa dạng toàn phương về dạng chính tắc. Để làm được điều này ta tiến hành như sau

Bước 1: Viết ma trận khối dạng $(A|E)$, ở đó A là ma trận dạng toàn phương.

Bước 2: Biến đổi sơ cấp song song các hàng và cột ma trận khối $(A|E)$ (để không làm mất tính đối xứng của A) đưa về ma trận khối $(\overline{A}|C^T)$, ở đó $\overline{A}$ là ma trận chéo.

Bước 3: Với phép đổi biến $x = Cy$, ta đưa dạng toán phương về dạng chính tắc với ma trận là $\overline{A} = C^T AC$.

Ví dụ 78. Đưa dạng toàn phương trong ví dụ trên về dạng chính tắc bằng phương pháp này

$$q\left(x\right) = x_1^2 + 2x_1x_2 + 2\mathrm{x}_2^2 + 4\mathrm{x}_2x_3 + 5\mathrm{x}_3^2$$

Ma trận của dạng toàn phương là

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 5 \end{pmatrix}$$

Khi đó ta có

$$(A|E) = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 5 & 0 & 0 & 1 \end{array}\right) \xrightarrow{-h_1+h_2\to h_2} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 5 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{-c_1+c_2\to c_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 5 & 0 & 0 & 1 \end{array}\right) \xrightarrow{-h_2+h_3\to h_3}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 5 & -1 & 1 & 1 \end{array}\right) \xrightarrow{-c_2+c_3\to c_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 5 & -1 & 1 & 1 \end{array}\right) = \left(\overline{A}|C^T\right)$$

Xét phép đổi biến $x = Cy$, thì dạng chính tắc của $q(x)$ là

$$q(x) = y_1^2 + y_2^2 + 4y_3^2$$

Cả hai phương pháp đưa dạng toàn phương này về dạng chính tắc đều có một điểm "giống nhau". Đó là tất cả các hệ số của dạng chuẩn tắc đều dương. Để hiểu rõ hơn lý do tại sao lại vậy chúng ta nghiên cứu về chỉ số quán tính của dạng toán phương trong phần sau.

3.1.3 Luật quán tính.

Giả sử trong một cơ sở nào đó của V, dạng toàn phương $q(x)$ có dạng chính tắc

$$q(x) = \sum_{i=1}^{r} \alpha_i x_i^2, r = rank(q)$$

giả sử rằng

$$q(x) = \sum_{i=1}^{p} \alpha_i x_i^2 + \sum_{j=p+1}^{r} \alpha_j x_j^2$$

ở đó

$$\begin{cases} \alpha_1, ..., \alpha_p > 0 \\ \alpha_{p+1}, ..., \alpha_r < 0 \end{cases}$$

Xét phép đổi biến

$$\begin{cases} x_i = \dfrac{1}{\sqrt{\alpha_i}} x'_i, i = \overline{1;p} \\ x_i = \dfrac{1}{\sqrt{-\alpha_i}} x'_i, i = \overline{p+1;r} \\ x_i = x'_i, i > r \end{cases}$$

thì trong cơ sở mới $\{e'_1, ..., e'_n\}$ này $q(x)$ có dạng

$$q(x) = \sum_{i=1}^{p} {x'_i}^2 - \sum_{j=p+1}^{r} {x'_j}^2$$

Dạng này gọi là *dạng chuẩn tắc* của $q(x)$. Số p gọi là *chỉ số quán tính dương* còn $t = r - p$ gọi là *chỉ số quán tính âm*, chỉ số dương và âm gọi chung là *chỉ số quán tính*. Hiệu số $\sigma = p - t$ gọi là *ký số* dạng toàn phương.

Định lý 3.1.3. *(Luật quán tính của Sylvester) Chỉ số quán tính không phụ thuộc cơ sở của V mà trong đó $q(x)$ có dạng chuẩn tắc.*

Chứng minh: Giả sử trong cơ sở $\{e_1, ..., e_n\}$, $q(x)$ có dạng

$$q(x) = \sum_{i=1}^{p} x_i^2 - \sum_{j=p+1}^{r} x_j^2$$

và trong cơ sở $\{e'_1, ..., e'_n\}$ $q(x)$ có dạng

$$q(x) = \sum_{i=1}^{m} {x'_i}^2 - \sum_{j=m+1}^{r} {x'_j}^2$$

Ta cần chứng minh $p = m$. Đặt

$$\mathrm{W} = span\{e_1, ..., e_p\}$$
$$Z = span\{e'_{m+1}, ..., e'_n\}$$

là các không gian con p chiều và $n - m$ chiều tương ứng. Khi đó, với mỗi vector $x \in W$ ta có

$$x = x_1e_1 + ... + x_pe_p = x_1e_1 + ... + x_pe_p + 0.e_{p+1} + ... + 0.e_n$$

tức là đối với cơ sở (e) thì $x = (x_1, ..., x_p, 0, ..., 0)^T$. Theo giả thiết thì

$$q(x) = \sum_{i=1}^{p} x_i^2 \geq 0, \forall x \in \mathrm{W}$$

và

$$\mathrm{q}\left(x\right)=0\Leftrightarrow x_1=...=x_p=0\Leftrightarrow x=0\in W\Leftrightarrow x=0\in V$$

Tương tự, với mỗi $x\in Z$ thì

$$x=x'_{m+1}e'_{m+1}+...+x'_ne'_n$$

và

$$q\left(x\right)=-\sum_{j=m+1}^{r}{x'_j}^2\leqslant 0,\forall x\in Z$$

Ta cũng có

$$q\left(x\right)=0\Leftrightarrow x'_{m+1}=...=x'_r=0$$

Nếu $x\in W\cap Z$ thì

$$\begin{cases} q\left(x\right)\geqslant 0\\ q\left(x\right)\leqslant 0\end{cases}$$

Do đó $q(x)=0$, và điều này có nghĩa là $x=0\in V$, tức là $W\cap Z=\{0\}$ hay $dim(W\cap Z)=0$, từ đó

$$\dim W+\dim Z=\dim\left(\mathrm{W}+Z\right)+\dim\left(\mathrm{W}\cap Z\right)=\dim\left(\mathrm{W}+Z\right)\leqslant n$$

hay là $p+n-m\leq n$, cũng tức là $p\leq m$. Chứng minh tương tự như trên ta cũng suy ra $m\leq p$. Do đó $p=m$. ▶

Như vậy theo luật quán tính, dạng chuẩn tắc của dạng toàn phương thực $q(x)$ trên V là duy nhất.

3.1.4 Dạng toàn phương xác định dấu

Giả sử $q(x)$ là dạng toàn phương trên V và $\psi(x,y)$ là dạng cực của nó.

Định nghĩa 42. *Dạng toàn phương $q(x)$ gọi là không âm (không dương) nếu $q(x)=\psi(x,x)\geq 0$ $(\leq 0),\forall x\in V$. Ngược lại nếu tồn tại $x,y\in V$ để*

$$q(x)=\psi(x,x)>0>q(y)=\psi(y,y)$$

thì dạng toàn phương gọi là không xác định dấu.

Dạng toàn phương là xác định dương (âm) nếu nó là không âm (không dương) và dấu đẳng thức xảy ra khi và chỉ khi $x=0$.

Dạng song tuyến tính đối xứng gọi là xác định dương nếu nó cho một dạng toàn phương xác định dương.

Ví dụ 79. Cho dạng toàn phương

$$q(x) = \sum_{i=1}^{n} \alpha_i x_i^2$$

- $q(x)$ không âm khi và chỉ khi $\alpha_i \geq 0, \forall i = \overline{1;n}$
- $q(x)$ xác định dương khi và chỉ khi $\alpha_i > 0, \forall i = \overline{1;n}$

Định lý 3.1.4. *Dạng toàn phương $q(x)$ xác định dương khi và chỉ khi*

$$\begin{cases} rank(q) = \dim V \\ \sigma = p - t = n \end{cases}$$

Chứng minh: Giả sử $q(x)$ xác định dương và $\{e_1, ..., e_n\}$ là cơ sở chính tắc, khi đó trong cơ sở này

$$q(x) = \sum_{i=1}^{r} \alpha_i x_i^2, r = rank(q)$$

- Nếu chỉ số quán tính âm $t > 0$ thì tồn tại $\alpha_i < 0$, ta chọn $x = x_i e_i$ thì $q(x) = \alpha_i x_i^2 < 0$, trái giả thiết. Do đó $t = 0$, tức là $\sigma = p = r$.

- Nếu $r = rank(q) < n$, ta chọn $x = (0, .., 0, x_{r+1}, ..., x_n)^T$ với $x_{r+1}, ..., x_n$ không đồng thời bằng 0. Khi đó

$$q(x) = \alpha_1.0 + ... + \alpha_r.0 + 0.x_{r+1}^2 + ... + 0.x_n^2 = 0$$

trái giả thiết $q(x)$ xác định dương, do đó $rank(q) = n$.

Ngược lại, nếu $rank(q) = dimV$ và $\sigma = p - t = n$ thì $p = n$, do đó tồn tại cơ sở chuẩn tắc để

$$q(x) = \sum_{i=1}^{n} x_i^2 > 0, \forall x \neq 0$$

tức là $q(x)$ xác định dương. ▶

Chúng ta cũng có kết quả tương tự cho trường hợp xác định âm. Định lý dưới đây cho ta một tiêu chuẩn cần và đủ để một dạng toàn phương xác định dương.

Định lý 3.1.5. *(Sylvester) Giả sử $q(x)$ là dạng toàn phương với ma trận biểu diễn $A = (a_{ij})_{n\times n}$. Điều kiện cần và đủ để $q(x)$ xác định dương là các định thức con chính của ma trận A đều dương, tức là*

$$\Delta_1 = a_{11} > 0, \Delta_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} > 0, ..., \Delta_n = \det(A) > 0$$

Chứng minh định lý trên có thể xem trong [3], [6], [14].

Hệ quả 3.1.1. *Dạng toàn phương $q(x)$ xác định âm khi và chỉ khi*

$$(-1)^i \Delta_i > 0, \forall i = \overline{1; n}.$$

Ví dụ 80. Giả sử $f(x) = f(x_1, ..., x_n)$ là một hàm thực trơn của n biến có điểm tới hạn tại $a \in \mathbb{R}^n$, tức là tất cả các đạo hàm riêng tại điểm này bằng 0

$$\frac{\partial f}{\partial x_i}(a) = 0$$

Ma trận *Hessian* tại a được xác định bởi ma trận đối xứng cấp n

$$H_f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{n \times n}$$

Một kết quả quan trọng sau đây của giải tích nói rằng nếu $x^T H_f x > 0$ với mọi $x \neq 0$ thì hàm f có cực tiểu địa phương tại a. Việc chứng minh kết quả này dựa trên lý thuyết đại số tuyến tính mà cụ thể chính là sử dụng định lý Sylvester ở trên.

3.2 Không gian Euclide

3.2.1 Tích vô hướng

Định nghĩa 43. *Giả sử V là $\mathbb{R}$ - không gian vector. Một dạng song tuyến tính đối xứng, xác định dương $\psi(x, y)$ trên V được gọi là một tích vô hướng, ký hiệu là $\langle x, y \rangle$.*

Như vậy, tích vô hướng là ánh xạ $< ., . >: V \times V \to \mathbb{R}$ thỏa mãn

E_1) $\langle x, y \rangle = \langle y, x \rangle, \forall x, y \in V$

E_2) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \forall x, y, z \in V$

E_3) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle, \forall x \in V, \lambda \in \mathbb{R}$

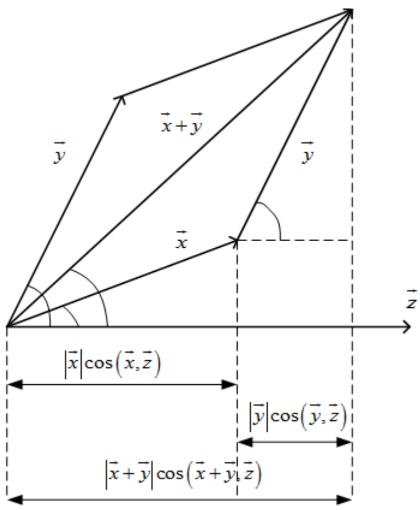
E_4) $\langle x, x \rangle \geq 0, \forall x \in V, \langle x, x \rangle = 0 \Leftrightarrow x = 0$

Một hệ quả hiển nhiên từ định nghĩa đó là $\langle x, 0 \rangle = 0$.

Ví dụ 81. Trong hình học phổ thông ta đã biết tích vô hướng của hai vector $\overrightarrow{x}, \overrightarrow{y}$ là

$$\langle \overrightarrow{x}, \overrightarrow{y} \rangle = |\overrightarrow{x}| |\overrightarrow{y}| \cos(\overrightarrow{x}, \overrightarrow{y})$$

không khó để kiểm tra các điều kiện E_1, E_3, E_4 của tích vô hướng, điều kiện E_2 chứng minh dễ dàng bằng hình học (xem Hình 3.1).



Hình 3.1: Tính chất E_2

Ví dụ 82. Trong $\mathbb{R}^n$ thì

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

là một tích vô hướng.

Ví dụ 83. Trên tập các ma trận $M_n(\mathbb{R})$ xây dựng tích vô hướng

$$\langle A, B\rangle = Trace\left(A^T B\right)$$

Có thể kiểm tra được đây thực sự là một tích vô hướng. Thật vậy, ta xem mỗi ma trận A được biểu diễn như là n vector cột dạng $A = (a_1, a_2, ..., a_n)$, như vậy nếu $B = (b_1, b_2, ..., b_n)$ thì

$$\langle A, B\rangle = Trace\left(A^T B\right) = \sum_{i=1}^{n} a_i^T b_i$$

dễ dàng kiểm tra các điều kiện của tích vô hướng dựa trên tính chất vết ma trận vuông như $Trace(A + B) = Trace(A) + Trace(B)$, $Trace(\lambda A) = \lambda Trace(A)$, $Trace(A^T B) = Trace(B^T A)$. Riêng tính chất xác định dương ta có ngay

$$\langle A, A\rangle = Trace\left(A^T A\right) = \sum_{i=1}^{n} a_i^T a_i = \sum_{i=1}^{n} \|a_i\|^2 \geqslant 0$$

Ví dụ 84. Trong $C[a, b]$ thì

$$\langle x, y\rangle = \int_a^b x(t) y(t)\, dt$$

cũng cho ta một tích vô hướng.

Có muôn vàn cách xây dựng tích vô hướng theo định nghĩa, chỉ cần ta có một dạng song tuyến tính đối xứng xác định dương (theo tiêu chuẩn Sylvester thì tất cả các định thức con chính ma trận biểu diễn đều dương) cho ta một tích vô hướng.

Không gian vector thực V trên đó trang bị một tích vô hướng gọi là *không gian Euclide*.

Dưới đây chúng ta đưa ra một vài bất đẳng thức tích vô hướng quan trọng.

3.2.2 Bất đẳng thức tích vô hướng

Bất đẳng thức Cauchy - Bunhia - Schwartz (CBS)

Định lý 3.2.1. *(Bất đẳng thức Cauchy - Bunhia - Schwartz)*

$$\langle x, y\rangle^2 \leqslant \langle x, x\rangle \langle y, y\rangle \tag{3.7}$$

Chứng minh: Nếu $x = 0$ hoặc $y = 0$ bất đẳng thức hiển nhiên đúng. Giả sử $x, y \neq 0$, ta có

$$\langle x + \lambda y, x + \lambda y\rangle \geqslant 0, \forall x, y \in V, \lambda \in \mathbb{R}$$

khai triển ra ta được

$$\lambda^2 \langle y, y\rangle + 2\lambda \langle x, y\rangle + \langle x, x\rangle \geqslant 0, \forall x, y \in V, \lambda \in \mathbb{R}$$

Đây là tam thức bậc hai xảy ra với mọi $\lambda \in \mathbb{R}$ khi và chỉ khi

$$\langle x, y\rangle^2 - \langle x, x\rangle \langle y, y\rangle \leqslant 0$$

Đẳng thức xảy ra nếu $x + \lambda y = 0$, tức là x, y phụ thuộc tuyến tính. ▶

Bất đẳng thức Minkowski

Với mọi $x \in V$ thì $\langle x, y\rangle \geq 0$, ta xác định số $\|x\| = \sqrt{\langle x, x\rangle}$ và gọi là *chuẩn* của vector x (đôi khi còn gọi là *modul* hay *độ dài* của x). Không khó để thấy $\|x\|$ có các tính chất sau:

i) $\|x\| \geq 0, \forall x \in V$, đẳng thức xảy ra khi và chỉ khi $x = 0$.

ii) $\|\lambda x\| = |\lambda| \, \|x\|, \forall x \in V, \lambda \in \mathbb{R}$.

iii) $\|x + y\| \leqslant \|x\| + \|y\|$, đẳng thức xảy ra khi và chỉ khi x, y phụ thuộc tuyến tính.

Tính chất cuối cùng chính là bất đẳng thức Minkowski. Ta chứng minh bất đẳng thức này. Ta có

$$\|x+y\|^2 = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$$

Sử dụng định bất đẳng thức CBS ta có

$$\|x+y\|^2 \leqslant \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

Và dễ dàng thu được bất đẳng thức Minkowski. Đẳng thức xảy ra khi và chỉ khi x, y phụ thuộc tuyến tính.

Không gian V được gọi là *định chuẩn* nếu với mọi $x \in V$ xác định một chuẩn thỏa mãn các điều kiện i)-iii) ở trên.

Chuẩn $\|x - y\|$ được gọi là *khoảng cách* giữa x và y, ký hiệu $d(x, y)$. Khoảng cách này có tính chất

$$d(x, y) \leq d(x, z) + d(z, y)$$

và gọi là bất đẳng thức tam giác, bất đẳng thức này suy ra trực tiếp từ bất đẳng thức Minkowski.

3.2.3 Cơ sở trực chuẩn, quá trình trực chuẩn hóa Gram-Schmidt

Giả sử trên V trang bị một tích vô hướng. Hai vector được gọi là *trực giao*, ký hiệu $x \perp y$ nếu $\langle x, y\rangle = 0$.

Mệnh đề 3.2.1. *Giả sử $v_1, .., v_m \in V$ là các vector đôi một trực giao, khi đó*

$$\left\|\sum_{i=1}^{m} v_i\right\|^2 = \sum_{i=1}^{m} \|v_i\|^2$$

Mệnh đề trên chứng minh một cách dễ dàng. Khi $m = 2$ ta có *công thức Pythagore.*

Hệ vector $\{e_1, ..., e_m\}$ trong V gọi là *hệ trực giao* nếu $e_i \neq 0$ và $e_i \perp e_j$ với mọi $i \neq j$, ở đó $i, j = \overline{1; m}$, tức là

$$\langle e_i, e_j\rangle = \begin{cases} 0, i \neq j \\ \|e_i\|^2, i = j \end{cases}$$

Hệ trực giao mà $\|e_i\| = 1$ gọi là *hệ trực chuẩn.*

Mệnh đề 3.2.2. *Hệ trực giao là độc lập tuyến tính.*

Cơ sở $\{e_1, ..., e_n\}$ gọi là một *cơ sở trực giao (trực chuẩn)* nếu nó là một hệ trực giao (trực chuẩn).

Nếu $\{e_1, ..., e_n\}$ là một cơ sở trực giao thì $\{e'_1, ..., e'_n\}$, ở đó $e'_i = \frac{e_i}{||e_i||}$, là cơ sở trực chuẩn. Quá trình như vậy gọi là *chuẩn hóa* cơ sở trực giao.

Định lý 3.2.2. *(Gram-Schmidt) Nếu hệ các vector $\{v_1, ..., v_m\}$ là độc lập tuyến tính bất kỳ trong V thì tồn tại hệ trực chuẩn $\{e_1, ..., e_m\}$ sao cho $e_i \in Span\{v_1, ..., v_i\}, \ \forall i = \overline{1; m}$.*

Chứng minh: Ta chứng minh bằng quy nạp.

Với $j = 1$, đặt $e_1 = \frac{v_1}{||v_1||}$. Giả sử xây dựng được hệ $\{e_1, ..., e_j\}$ trực chuẩn và $e_k \in span\{v_1, ..., v_k\}, \ \forall k = \overline{1; j}$. Ta chỉ ra cách xây dựng e_{j+1}.

Đặt $e'_{j+1} = v_{j+1} + \alpha_1 e_1 + ... + \alpha_j e_j$, ở đó các α_i xác định sau và đòi hỏi $e'_{j+1} \perp e_k, \forall k = \overline{1; j}$. Như vậy

$$\left\langle e'_{j+1}, e_k \right\rangle = 0 \Leftrightarrow \langle v_{j+1}, e_k \rangle + \alpha_k \|e_k\| = 0$$

Do $\|e_k\| = 1$ nên $\alpha_k = -\langle v_{j+1}, e_k \rangle, \ \forall k = \overline{1; j}$ Tức là

$$e'_{j+1} = v_{j+1} - \sum_{k=1}^{j} \langle v_{j+1}, e_k \rangle e_k$$

Vì $\{v_1, ..., v_j, v_{j+1}\}$ độc lập tuyến tính nên v_{j+1} không biểu diễn tuyến tính qua $\{e_1, ..., e_j\}$ (vì $span\{e_1, ..., e_j\} = span\{v_1, ..., v_j\}$), do đó $e'_{j+1} \neq 0$.

Đặt

$$e_{j+1} = \frac{e'_{j+1}}{\left\|e'_{j+1}\right\|}$$

Ta được hệ trực chuẩn $\{e_1, ..., e_{j+1}\}$ thỏa mãn định lý. ▶

Hệ quả 3.2.1. *Nếu $\{v_1, ..., v_n\}$ là một cơ sở bất kỳ của V thì tồn tại cơ sở trực chuẩn $\{e_1, ..., e_n\}$ sao cho $e_j \in span\{v_1, ..., v_j\}$.*

Định lý 3.2.3. *Nếu $\{e_1, ..., e_n\}$ là một cơ sở trực chuẩn của V thì với mọi $x \in V$ ta có*

$$x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i.$$

Có thể chứng minh định lý một cách dễ dàng.

Chú ý 11. *Như vậy mọi không gian vector Euclide V n chiều đều có cơ sở trực chuẩn. Mọi hệ trực chuẩn trong không gian V đều có thể bổ sung để tạo thành cơ sở trực chuẩn. Giả sử $\{e_1, ..., e_n\}$ là cơ sở trực chuẩn thì với mọi $x, y \in V$ ở đó $x = \sum\limits_{i=1}^{n} x_i e_i,\ y = \sum\limits_{i=1}^{n} y_i e_i$ ta có*

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Mệnh đề 3.2.3. *Giả sử có hai cơ sở trực chuẩn $\{e_1, ..., e_n\}$ và $\{e'_1, ..., e'_n\}$ trong không gian Euclide V. Khi đó ma trận chuyển cơ sở $C : (e) \to (e')$ là ma trận trực giao.*

Chứng minh: Ta có $(e') = (e)C$, do đó $(e')^T = C^T(e)^T$. Xét ma trận các tích vô hướng

$$(e')^T (e') = \begin{pmatrix} \langle e'_1, e'_1\rangle & \cdots & \langle e'_1, e'_n\rangle \\ \vdots & \vdots & \vdots \\ \langle e'_n, e'_1\rangle & \vdots & \langle e'_n, e'_n\rangle \end{pmatrix} = E$$

tương tự thì $(e)^T (e) = E$, điều này tức là

$$C^T (e)^T (e) C = C^T E C = C^T C = E$$

hay $C^T = C^{-1}$. ▶

Ví dụ 85. Trực chuẩn hóa Gram-schmidt hệ vector

$$v_1 = (1, 1, 1), v_2 = (0, 1, 1), v_3 = (0, 0, 1)$$

Rõ ràng hệ này độc lập tuyến tính.

Đặt $w_1 = v_1,\ e_1 = \frac{w_1}{\|w_1\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$. Tiếp theo, ta có

$$\mathrm{w}_2 = v_2 - \langle v_2, e_1\rangle e_1 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right) \Rightarrow e_2 = \frac{\mathrm{w}_2}{\|\mathrm{w}_2\|} = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

và

$$\mathrm{w}_3 = v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2 = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

từ đó

$$e_3 = \frac{\mathrm{w}_3}{\|\mathrm{w}_3\|} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Ma trận chuyển cơ sở trực chuẩn thành trực chuẩn là ma trận trực giao, vì vậy ma trận trực giao đóng vai trò quan trọng. Ví dụ dưới đây cho ta tất cả các dạng của ma trận trực giao cấp hai.

Ví dụ 86. Tìm biểu diễn của ma trận trực giao cấp hai. Giả sử ma trận trực giao cấp hai có dạng

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Theo giả thiết $AA^T = E$, do đó $\det(A) = \pm 1$. Từ đó ta có các hệ phương trình để tìm các phần tử của A là

$$\begin{cases} a^2 + c^2 = 1 \\ ab + c\mathrm{d} = 0 \\ b^2 + d^2 = 1 \\ a\mathrm{d} - bc = 1 \end{cases}$$

hoặc

$$\begin{cases} a^2 + c^2 = 1 \\ ab + c\mathrm{d} = 0 \\ b^2 + d^2 = 1 \\ a\mathrm{d} - bc = -1 \end{cases}$$

Giải hệ tìm được ma trận A chỉ có hai biểu diễn là

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, a^2 + b^2 = 1$$

hoặc

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, a^2 + b^2 = 1$$

Đặt $a = \cos\varphi, b = \sin\varphi$ thì chúng ta viết lại A ở dạng

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

hoặc

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

3.2.4 Phân tích QR

Giả sử rằng $A \in M_{n\times m}(\mathbb{R})$ là ma trận gồm m cột độc lập tuyến tính, A viết dưới dạng các vector cột là $A = (v_1, v_2, ..., v_m)$. Trực chuẩn hoá Gram-Schmidt các vector $v_1, v_2, ..., v_m$ ta được các vector $e_1, e_2, ..., e_m$, mặt khác từ chứng minh định lý 3.2.2 ta thấy rằng

$$v_k = \sum_{i=1}^{k} \langle v_k, e_i \rangle\, e_i$$

vì vậy ta có thể viết

$$A = (v_1, v_2, ..., v_m) = (e_1, e_2, ..., e_m)\begin{pmatrix} \langle v_1, e_1\rangle & \langle v_2, e_1\rangle & \cdots & \langle v_m, e_1\rangle \\ 0 & \langle v_2, e_2\rangle & \cdots & \langle v_m, e_2\rangle \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \langle v_m, e_m\rangle \end{pmatrix} = QR$$

Như vậy Q là ma trận các cột trực giao, còn R là ma trận vuông cấp m các hệ số khi khai triển các vector v_k theo cơ sở trực chuẩn thu được từ quá trình trực chuẩn hóa Gram-Schmidt các vector này (hệ số Fourier). Rõ ràng $\langle v_i, e_i\rangle \neq 0$ vì vậy R khả nghịch. Ta có thể phát biểu kết quả này dưới dạng định lý sau.

Định lý 3.2.4. *Giả sử $A \in M_{n\times m}(\mathbb{R})$ với $rank(A) = m$, khi đó có thể phân tích*

$$A = QR$$

trong đó Q là ma trận có các cột trực giao, còn R là ma trận tam giác trên cấp m khả nghịch.

Từ định lý trên ta thấy, nếu A là ma trận vuông cấp n khả nghịch thì Q là ma trận trực giao cấp n. Một điều chú ý thêm là phân tích QR nói chung không duy nhất.

Ví dụ 87. Cho ma trận

$$A = (v_1, v_2, v_3) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Tìm một phân tích QR của nó. Để giải quyết bài toán, đầu tiên ta trực chuẩn hóa Gram-Schmidt các vector v_1, v_2, v_3, ví dụ ta tìm được

$$e_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)^T, e_2 = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)^T, e_3 = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$$

Khi đó

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Còn

$$R = \begin{pmatrix} \langle v_1, e_1 \rangle & \langle v_2, e_1 \rangle & \langle v_3, e_1 \rangle \\ 0 & \langle v_2, e_2 \rangle & \langle v_3, e_2 \rangle \\ 0 & 0 & \langle v_3, e_3 \rangle \end{pmatrix} = \begin{pmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

3.3 Không gian con trực giao và hình chiếu

Định nghĩa 44. *Giả sử V là không gian Euclide thực, W, Z là hai không gian con của V. W và Z gọi là trực giao nhau, viết là $W \perp Z$, nếu $w \perp z, \forall w \in W, z \in Z$. Nếu $V = W \oplus Z$ và $W \perp Z$ thì W, Z là các không gian con bù trực giao nhau.*

Mệnh đề 3.3.1. *Giả sử W, Z là các không gian con của không gian Eulcid V.*

a) $\{0\} \perp W$ với mọi W - không gian con của V.

b) Nếu $W \perp Z$ thì $W \cap Z = \{0\}$.

c) $W^{\perp} = \{v \in V : v \perp w, \forall w \in W\}$ là phần bù trực giao của W.

Giả sử x, y là các vector khác 0 trong không gian Euclide V, *góc* giữa x và y ký hiệu $\widehat{x, y} \in [0, \pi]$ xác định từ

$$\cos(\widehat{x, y}) = \frac{\langle x, y \rangle}{\|x\| \, \|y\|}$$

Định lý 3.3.1. *Giả sử V là không gian Euclide n chiều, $\{v_1, ..., v_m\}$ là hệ trực chuẩn các vector trong V, $W = span\{v_1, ..., v_m\}$, với mọi $x \in V$, đặt*

$$w_1 = \sum_{i=1}^{m} \langle x, v_i \rangle v_i$$

$$w_2 = x - w_1$$

Khi đó

a) $w_1 \in W$

b) $w_2 \perp W$.

Chứng minh:

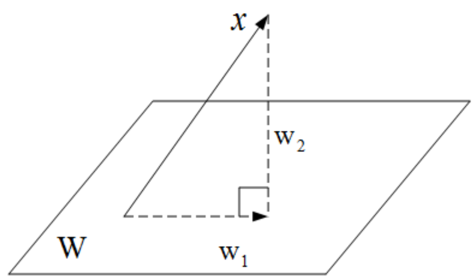
a) Hiển nhiên.

b) Ta có

$$\langle w_2, v_i\rangle = \langle x - w_1, v_i\rangle = \langle x, v_i\rangle - \langle x, v_i\rangle \|v_i\|^2 = 0$$

Do đó $w_2 \perp W$. ▶

Vector w_1 gọi là *hình chiếu* của x lên W, ký hiệu $w_1 = ch_W x$, còn w_2 gọi là *thành phần của x trực giao với W*.



Hình 3.2: Hình chiếu trực giao

Mệnh đề 3.3.2. *Giả sử vector $v \in V$ và W là không gian con của V, khi đó*

$$\min_{w\in W} \|v - w\| = \|v - ch_W v\|$$

Chứng minh: Đặt $v_0 = ch_W v$, ta có

$$\|v - w\|^2 = \langle v - w, v - w\rangle = \langle v - v_0 + v_0 - w, v - v_0 + v_0 - w\rangle$$
$$= \|v - v_0\|^2 + \|v_0 - w\|^2 + 2\langle v - v_0, v_0 - w\rangle = \|v - v_0\|^2 + \|v_0 - w\|^2$$

vì $v_0 - w \in W,\ v - v_0 \in W^{\perp}$. Do đó

$$\|v - w\| \geqslant \|v - v_0\|$$

dấu đẳng thức xảy ra khi và chỉ khi $w = v_0$. ▶

3.4 Chéo hóa trực giao ma trận đối xứng

3.4.1 Toán tử tự liên hợp

Định nghĩa 45. *Giả sử V là không gian Euclide thực, một toán tử tuyến tính $f : V \to V$ gọi là tự liên hợp (hay đối xứng) nếu với mọi $x, y \in V$ ta có*

$$\langle f(x), y\rangle = \langle x, f(y)\rangle$$

Nếu V là không gian Euclide phức thì toán tử f còn gọi là *tự liên hợp Hermite.*

Ví dụ 88. Toán tử đồng nhất $Id : V \to V$ là tự liên hợp.

Ví dụ 89. Xét toán tử $T : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ sao cho $T(A) = A^T$. Tập $M_n(\mathbb{R})$ trang bị tích vô hướng $\langle A, B\rangle = Trace(A^T B)$. Khi đó T là tự liên hợp.

Ví dụ 90. Xét toán tử tuyến tính $f : \mathbb{R}^n \to \mathbb{R}^n$ sao cho $f(x) = Ax$, ở đó $A = A^T$. Khi đó f là tự liên hợp.

Định lý 3.4.1. *Giả sử V là không gian Euclide thực, điều kiện cần và đủ để toán tử tuyến tính $f : V \to V$ tự liên hợp trên V là trong một cơ sở trực chuẩn nào đó ma trận của f có dạng đối xứng.*

Chứng minh: Chú ý rằng, vì tọa độ mỗi vector luôn biểu diễn dạng cột nên tích vô hướng của hai vector có thể biểu diễn như tích hai ma trận, tức là

$$\langle x, y\rangle = x^T y$$

Bây giờ giả sử toán tử tuyến tính f tự liên hợp và $\{e_1, .., e_n\}$ là một cơ sở trực chuẩn nào đó. Nếu A là ma trận của f trong cơ sở này, theo lý thuyết chương trước ta có $f(e_i) = Ae_i$, như vậy

$$\langle f(e_i), e_j\rangle = \langle Ae_i, e_j\rangle = \langle e_i, f(e_j)\rangle = \langle e_i, Ae_j\rangle$$

tức là

$$e_i^T A^T e_j = {e_i}^T A e_j$$

hay $A = A^T$. Điều ngược lại không khó (xem ví dụ phía trên). ▶

Định lý 3.4.2. *Trị riêng của toán tử tự liên hợp là số thực.*

Chứng minh: Giả sử f là toán tử tự liên hợp trên không gian Euclide V thực n chiều, $A = (a_{ij})_{n\times n}$ là ma trận của f trong một cơ sở trực chuẩn nào đó, $A = A^T$. Giả sử $\lambda = a + ib$ là trị riêng của f, khi đó vector riêng ứng với λ là $u + iv$, ở đó u, v là các vector thực. Như vậy

$$A(u + iv) = (a + ib)(u + iv)$$

Lấy liên hợp hai vế ta có

$$A(u - iv) = (a - ib)(u - iv)$$

Xét tích vô hướng

$$(u - iv)^T A (u + iv) = (a + ib) \left(\|u\|^2 + \|v\|^2 \right)$$

mặt khác, vế trái của biểu thức này có thể viết lại là

$$\left((u - iv)^T A \right) (u + iv) = (A (u - iv))^T (u + iv) = (a - ib) \left(\|u\|^2 + \|v\|^2 \right)$$

Do $u + iv$ là vector riêng nên $\|u\|^2 + \|v\|^2 \neq 0$, từ đó suy ra $a + ib = a - ib$ hay $b = 0$. Điều này có nghĩa là λ thực. ▶

Mệnh đề 3.4.1. *Giả sử f là tự liên hợp, v và w là hai vector riêng ứng với λ, μ là hai trị riêng khác nhau của f, khi đó $v \perp w$.*

Chứng minh: Ta có

$$\langle f(v), w \rangle = \langle \lambda v, w \rangle = \langle v, f(w) \rangle = \langle v, \mu w \rangle$$

Do $\lambda \neq \mu$ nên $v \perp w$. ▶

Định lý 3.4.3. *Luôn tồn tại cơ sở trực chuẩn nào đó để toán tử tự liên hợp có dạng đường chéo.*

Chứng minh: Giả sử λ_1 là một trị riêng của toán tử tự liên hợp f, $\lambda_1 \in \mathbb{R}$. Giả sử e_1 là trị riêng ứng với λ_1, ta có $f(e_1) = \lambda_1 e_1$, $L_1 = span\{e_1\}$ là không gian con f - bất biến. Gọi L_2 là không gian con trực giao với e_1, tức là $L_2 = L(e_1)^{\perp}$, thì $dim L_2 = n - 1$ (do $V = L_1 \oplus L_2$). Có thể thấy L_2 cũng là f - bất biến. Thật vậy, với mọi $x \in L_2$ ta có

$$\langle e_1, \text{f(x)} \rangle = \langle f(e_1), x \rangle = \langle \lambda_1 e_1, x \rangle = \lambda_1 \langle e_1, x \rangle = 0$$

Do đó $f(x) \in L_2$, hay L_2 là f - bất biến. Khi đó tồn tại $e_2 \in L_2$ sao cho $f(e_2) = \lambda_2 e_2, \lambda_2 \in \mathbb{R}$. Gọi $L_3 = span\{e_1, e_2\}^{\perp}$ thì $dim L_3 = n - 2$, có thể kiểm tra được L_3 là f - bất biến. Cứ như vậy ta có thể xây dựng được n vector riêng độc lập tuyến tính $e_1, e_2, ..., e_n$ ứng với các trị riêng $\lambda_1, \lambda_2, ..., \lambda_n$. Rõ ràng hệ các vector $\{e_1, e_2, ..., e_n\}$ là trực giao, chuẩn hóa hệ này ta có

$$e'_k = \frac{e_k}{\|e_k\|}, f(e'_k) = \frac{f(e_k)}{\|e_k\|} = \frac{\lambda_k e_k}{\|e_k\|} = \lambda_k e'_k$$

Trong cơ sở đã chuẩn hóa này ta có ma trận biểu diễn là $\overline{A} = diag(\lambda_1, ..., \lambda_n)$. ▶

Hệ quả 3.4.1. *Mọi ma trận đối xứng thực đều tồn tại ma trận trực giao C sao cho C^TAC có dạng chéo.*

Chứng minh: Vì A đối xứng nên nó là ma trận của toán tử tự liên hợp trong một cơ sở trực chuẩn $(e) = (e_1, ..., e_n)$ nào đó. Theo định lý trên tồn tại cơ sở trực chuẩn (e') để trong cơ sở này toán tử tự liên hợp có ma trận $\overline{A}$ dạng chéo. Vì ma trận chuyển cơ sở C từ (e) sang (e') là ma trận trực giao nên ta có $\overline{A} = C^TAC$. ▶

Chéo hóa toán tử tự liên hợp thực chất là chéo hóa ma trận đối xứng $A = A^T$ của nó. Từ đây chúng ta có các bước để chéo hóa ma trận đối xứng như sau.

Bước 1: Giải phương trình đặc trưng $\det(A - \lambda E) = 0$.

Bước 2: Gọi $\lambda_1, ..., \lambda_k$ là các trị riêng với bội $n_1, ..., n_k$ tương ứng. Hiển nhiên $n_1 + ... + n_k = n$. Với mỗi λ_i tìm được n_i vector riêng độc lập tuyến tính, trực chuẩn hóa Gram-Schmidt hệ này ta được n_i vector trực chuẩn là $e_{i1}, ..., e_{in_i}$.

Bước 3: Cơ sở mới thu được là $(e) = (e_{11}, ..., e_{1n_1}, ..., e_{k1}, ..., e_{kn_k})$. Gọi ma trận C là ma trận tạo thành từ các cột vector cơ sở này thì

$$\overline{A} = C^TAC = \begin{pmatrix} \lambda_1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \lambda_1 & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \lambda_k & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \lambda_k \end{pmatrix}$$

Ví dụ 91. Cho toán tử tự liên hợp trong $\mathbb{R}^3$ với ma trận biểu diễn

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Ta đi tìm ma trận trực giao C sao cho C^TAC có dạng chéo.

Ta có

$$\det(A - \lambda E) = (\lambda - 1)^2 (4 - \lambda) = 0$$

Với $\lambda = 1$ giải được các vector riêng $v_1 = (-1, 1, 0)^T, v_2 = (-1, 0, 1)$, trực chuẩn hóa Gram-Schmidt ta được

$$e_1 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)^T, e_2 = \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)^T$$

Với $\lambda = 4$ tìm được vector riêng $e_3 = (1, 1, 1)$, chuẩn hóa ta được $e_3 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)^T$. Như vậy

$$C = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}, C^T AC = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Giả sử $q(x)$ là một dạng toàn phương trên không gian Euclide V n chiều. Tìm một cơ sở trực chuẩn để $q(x)$ có dạng chính tắc được gọi là *chính tắc hóa trực giao dạng toàn phương*. Như vậy, nếu cho ma trận $A = (a_{ij})_{n\times n}$ của dạng toàn phương $q(x)$ trong một cơ sở trực chuẩn (ma trận đối xứng) thì chính tắc hóa trực giao dạng toàn phương này chính là tìm ma trận trực giao C để $C^T MC$ có dạng chéo.

Từ đó ta có phương pháp giải *bài toán đưa dạng toàn phương về dạng chính tắc bằng phương pháp chéo hóa trực giao* giống như bài toán chéo hóa trực giao ma trận đối xứng, nhưng bước cuối cùng thì ta kết luận dạng chính tắc của dạng toàn phương. Tức là ta xét phép đổi biến $x = Cy$, ở đó C là ma trận trực giao trong lời giải bài toán chéo hóa trực giao, ta thu được dạng chính tắc là

$$q(x) = y^T C^T ACy$$

Bài toán tìm max, min của dạng toàn phương trong không gian Euclide

$$\max_{\|x\|=1} q(x), \min_{\|x\|=1} q(x)$$

Giống như lập luận ở trên tồn tại ma trận trực giao C sao cho $x = Cy$ và

$$q(x) = \sum_{i=1}^{n} \lambda_i y_i^2$$

trong đó $\lambda_1 \leq ... \leq \lambda_n$ là các trị riêng của ma trận dạng toàn phương $q(x)$. Khi đó

$$\lambda_1 \|y\| \leqslant q(x) \leqslant \lambda_n \|y\|$$

ở đó $\|x\| = \|Cy\| = \|y\| = 1$. Do đó $\min q(x) = \lambda_1$ tại các y ứng với λ_1 và $\max q(x) = \lambda_n$ tại các y ứng với λ_n. Để xác định giá trị lớn nhất và nhỏ nhất đạt tại các điểm x trong hệ tọa độ đầu tiên chỉ cần viết lại $x = Cy$.

Toán tử tự liên hợp f gọi là *xác định không âm* nếu như $\langle f(x), x\rangle \geqslant 0$ với mọi $x \in V$, viết là $f \geq 0$. Tương tự có thể định nghĩa toán tử tự liên hợp *xác định âm, xác định dương...*

Mệnh đề 3.4.2. *Toán tử tự liên hợp f xác định không âm khi và chỉ khi tất cả các trị riêng của nó đều không âm.*

Chứng minh: Giả sử $\lambda_1, ..., \lambda_n$ là các trị riêng (kể cả bội) của toán tử tự liên hợp $f \geq 0$. Khi đó tồn tại một cơ sở trực chuẩn gồm toàn vector riêng $e_1, ..., e_n$. Ta có

$$0 \leqslant \langle f(e_i), e_i\rangle = \lambda_i \langle e_i, e_i\rangle = \lambda_i, \forall i = \overline{1; n}$$

Ngược lại, giả sử tất cả các trị riêng $\lambda_i \geq 0$ và $e_1, ..., e_n$ là cơ sở trực chuẩn ứng với nó. Khi đó $\forall x \in V : x = \sum\limits_{i=1}^{n} x_i e_i$ ta có

$$\langle f(x), x\rangle = \sum_{i=1}^{n} \lambda_i {x_i}^2 \langle e_i, e_i\rangle = \sum_{i=1}^{n} \lambda_i {x_i}^2 \geqslant 0$$

►

Không khó để chứng minh một kết quả tương tự sau

Định lý 3.4.4. *Dạng toàn phương $q(x)$ trên không gian Euclide với ma trận biểu diễn A là:*
i) Xác định âm khi và chỉ khi tất cả các trị riêng của A đều âm
ii) Xác định dương khi và chỉ khi tất cả các trị riêng của A đều dương
iii) Không xác định dấu nếu tồn tại các trị riêng của A trái dấu nhau.

3.4.2 Phổ của toán tử tự liên hợp

Giả sử V là $\mathbb{K}$ không gian vector n chiều có một phân tích thành tổng trực tiếp của k không gian con $V_1, ..., V_k$, tức là

$$V = V_1 \oplus V_2 \oplus ... \oplus V_k$$

Điều này có nghĩa là với mỗi $x \in V$, ta có biểu diễn duy nhất

$$x = v_1 + v_2 + ... + v_k, v_i \in V_i$$

Xét ánh xạ $P_i : V \to V$ là phép chiếu V lên V_i xây dựng như sau

$$P_i(x) = v_i$$

Dễ dàng kiểm tra được P_i là toán tử tuyến tính trong V. Thật vậy, với $y \in V$ và y có biểu diễn duy nhất $y = w_1 + w_2 + ... + w_k, w_i \in V_i$, ta có

$$P_i(\alpha x + \beta y) = \alpha v_i + \beta w_i = \alpha P\left(x\right) + \beta P\left(y\right), \forall \alpha, \beta \in \mathbb{K}$$

Không khó để chứng minh các tính chất sau đây của toán tử chiếu P_i.

Mệnh đề 3.4.3. *Giả sử $P_i, i = \overline{1;k}$ là toán tử chiếu trên V, khi đó*
i) $P_i^2 = P_i$
ii) $P_i \circ P_j = P_j \circ P_i = 0$ *nếu* $i \neq j$
iii) $\sum\limits_{i=1}^{k} P_i = Id_V$

Bây giờ giả sử V là không gian Euclide n chiều và $f : V \to V$ là một toán tử tự liên hợp trên đó. Theo trên, khi đó tồn tại cơ sở gồm toàn vector riêng của V là $\{e_1, ..., e_n\}$ ứng với các trị riêng $\lambda_1, ..., \lambda_n$. Đặt $V_i = V(\lambda_i)$ là không gian con sinh bởi $e_i, i = \overline{1;n}$ thì hiển nhiên

$$V = V_1 \oplus V_2 \oplus .... \oplus V_n$$

Giả sử P_i là toán tử chiếu V lên V_i, P_i gọi là các *toán tử chiếu trực giao* (do các V_i đôi một trực giao nhau). Có thể thấy P_i là tự liên hợp trong V. Thật vậy,

$$\left\langle P_i\left(x\right), y\right\rangle = \left\langle x_i e_i, \sum_{j=1}^{n} y_j e_j \right\rangle = x_i y_i = \left\langle \sum_{j=1}^{n} x_j e_j, y_i e_i \right\rangle = \left\langle x, P_i\left(y\right)\right\rangle$$

Với vector $x \in V$ bất kỳ $x = \sum\limits_{i=1}^{n} x_i e_i$ thì

$$f\left(x\right) = \sum_{i=1}^{n} x_i f\left(e_i\right) = \sum_{i=1}^{n} \lambda_i x_i e_i = \sum_{i=1}^{n} \lambda_i P_i\left(x\right)$$

Biểu thức hình thức

$$f = \sum_{i=1}^{n} \lambda_i P_i$$

gọi là biểu diễn phổ của toán tử tự liên hợp f. Nếu gọi A là ma trận của toán tử tự liên hợp f trong cơ sở trực chuẩn gồm toàn vector riêng $\{e_1, ..., e_n\}$ và $\mathbb{P}_i$ là ma trận của toán tử chiếu trực giao P_i thì ta có biểu diễn

$$A = \sum_{i=1}^{n} \lambda_i \mathbb{P}_i$$

Dễ thấy

$$P_i x = x_i e_i = \langle x, e_i \rangle e_i = e_i {e_i}^T x$$

Do đó, ma trận $\mathbb{P}_i$ của toán tử P_i chính là $\mathbb{P}_i = e_i e_i^T$.

Trở lại vấn đề hình chiếu trực giao ở trên. Nếu $W = span\{v_1, ..., v_m\}$ là không gian con của không gian Euclide n chiều V, ở đó $\{v_1, ..., v_m\}$ là hệ trực chuẩn, khi đó hình chiếu của $x \in V$ trên W là

$$w_1 = \sum_{i=1}^{m} \langle x, e_i \rangle e_i = \sum_{i=1}^{m} e_i e_i^T x = \sum_{i=1}^{m} P_i x$$

Toán tử chiếu P lên W chính là $P = P_1 + ... + P_m$, ma trận của toán tử này là $\mathbb{P} = \sum\limits_{i=1}^{m} e_i e_i^T$.

Ví dụ 92. Trong $\mathbb{R}^3$ tìm

$$\min_{w \in W} \|x - w\|$$

ở đó $x = (1, 1, 1)^T$ còn W là không gian con sinh bởi $v_1 = (1, 0, 2)^T$ và $v_2 = (0, 1, -1)^T$.

Đầu tiên trực chuẩn hóa cơ sở của W ta được

$$W = span\{e_1, e_2\} : e_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^T; e_2 = \left(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right)^T$$

Ta biết rằng min đạt được tại $w_1 = Px$, ở đó ma trận của P là

$$\mathbb{P} = e_1 e_1^T + e_2 e_2^T = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

Từ đây ta có

$$\min_{\mathrm{w} \in W} \|x - w\| = \frac{1}{\sqrt{3}}$$

tại $w_1 = \left(\frac{4}{3}, \frac{2}{3}, \frac{2}{3}\right)^T$.

3.5 Phân loại các đường cong và mặt cong bậc hai

3.5.1 Phương trình siêu mặt bậc hai

Dưới đây ta sẽ phân loại các siêu mặt bậc hai tổng quát trong không gian Euclide thực n chiều nào đó. Để tránh nhầm lẫn, ta ký hiệu không gian đó là $\mathbb{E}^n$, và để dễ hình dung ta xem mỗi vector $x = (x_1, ..., x_n)$ trong cơ sở trực chuẩn $\{e_1, ..., e_n\}$ của $\mathbb{E}^n$ là một *điểm*. Điểm 0 ký hiệu là O, và gọi là điểm gốc của không gian $\mathbb{E}^n$.

Định nghĩa 46. *Một siêu mặt bậc hai trong $\mathbb{E}^n$ là tập các điểm $x \in \mathbb{E}^n$ thỏa mãn phương trình*

$$q(x) + 2L(x) + c = 0 \tag{3.8}$$

ở đó, $q(x)$ là một dạng toàn phương, $L(x)$ là một dạng tuyến tính trong $\mathbb{E}^n$, $c \in \mathbb{R}$ là hằng số nào đó.

Giả sử trong cơ sở $\{e_1, ..., e_n\}$ của $\mathbb{E}^n$, phương trình siêu mặt (3.8) có dạng

$$\sum_{i,j=1}^{n} a_{\text{ij}} x_i x_j + 2\sum_{j=1}^{n} b_j x_j + c = 0 \tag{3.9}$$

Định lý 3.5.1. *Trong không gian Euclide $\mathbb{E}^n$ luôn tồn tại cơ sở trực chuẩn nào đó để phương trình siêu mặt bậc hai hai có một trong ba dạng chính tắc bậc hai sau:*

i) Dạng I: $\sum\limits_{i=1}^{r} \alpha_i y_i^2 = 1,\ \alpha_i \neq 0, i = \overline{1;r}, 1 \leq r \leq n$

ii) Dạng II: $\sum\limits_{i=1}^{r} \alpha_i y_i^2 = 0,\ \alpha_i \neq 0, i = \overline{1;r}, 1 \leq r \leq n$

iii) Dạng III: $\sum\limits_{i=1}^{r} \alpha_i y_i^2 = 2py_{r+1},\ \alpha_i \neq 0, i = \overline{1;r}, p > 0, 1 \leq r \leq n-1$

Chứng minh: Như ta đã biết, luôn tồn tại cơ sở trực chuẩn $\{e'_1, ..., e'_n\}$ gồm toàn các vector riêng ứng với các trị riêng $\lambda_1, ..., \lambda_n$ của $A = (a_{ij})_{n\times n}$ để dạng toàn phương $q(x)$ có dạng chính tắc. Giả sử C là ma trận chuyển cơ sở trực chuẩn $\{e_1, ..., e_n\}$ thành $\{e'_1, ..., e'_n\}$, khi đó với phép đổi biến $x = Ct$ (giữ nguyên gốc O), siêu mặt sẽ có dạng

$$\sum_{i=1}^{n} \lambda_i t_i^2 + 2\sum_{i=1}^{n} d_i t_i + e = 0 \tag{3.10}$$

Trong các trị riêng $\lambda_1, ..., \lambda_n$, giả sử có r giá trị khác 0, và không sợ nhầm lẫn, ta vẫn ký hiệu các giá trị khác 0 đó là $\lambda_1, ..., \lambda_r$, khi đó (3.10) có thể viết lại dưới dạng

$$\sum_{i=1}^{r} \lambda_i t_i^2 + 2\sum_{i=1}^{r} d_i t_i + e = 0 \tag{3.11}$$

Nếu ta tịnh tiến điểm gốc O sang I (vẫn trong cơ sở $\{e'_1, ..., e'_n\}$) như sau

$$\begin{cases} t_i = z_i - \dfrac{d_i}{\lambda_i}, i = \overline{1;r} \\ t_i = z_j, j = \overline{r+1;n} \end{cases}$$

thì trong cơ sở (e') gốc I mới, phương trình siêu mặt có dạng

$$\sum_{i=1}^{r} \lambda_i z_i^2 + \sum_{i=r+1}^{n} c' z_i + d' = 0, 1 \leq r \leq n, \lambda_i \neq 0 \tag{3.12}$$

i) Nếu $c'_i = 0, i = \overline{r+1;n}$ và $d' \neq 0$ khi đó (3.12) có dạng I.

ii) Nếu $c'_i = 0, i = \overline{r+1;n}$ và $d' = 0$ khi đó (3.12) có dạng II.

iii) Nếu $r < n$ và tồn tại $c'_i \neq 0, i = \overline{r+1;n}$, ví dụ $c'_{r+1} \neq 0$. Đặt

$$p = \sqrt{\sum_{j=r+1}^{n} {c'_i}^2}, g_i = \frac{{c'}_i}{p}, i = \overline{r+1;n}$$

ta đưa siêu mặt về dạng

$$\sum_{i=1}^{r} \lambda_i z_i^2 + 2p\left(\sum_{i=r+1}^{n} g_i z_i + \frac{d'}{2p}\right) = 0 \tag{3.13}$$

ở đó các g_i thỏa mãn $\sum\limits_{i=r+1}^{n} g_i^2 = 1$. Tiếp tục xét phép đổi gốc mới từ I sang K như sau

$$\begin{cases} y_i = z_i, i = i = \overline{1;r} \\ y_{r+1} = -\displaystyle\sum_{i=r+1}^{n} g_i z_i - \frac{d'}{2p} \\ y_j = \displaystyle\sum_{k=r+1}^{n} g_{jk} z_k, j = \overline{r+1;n} \end{cases}$$

ở đó g_{jk} được chọn sao cho ma trận phép đổi biến này vẫn là ma trận trực giao (để đảm bảo cơ sở mới vẫn là trực chuẩn). Khi đó siêu mặt sẽ có dạng

$$\sum_{i=1}^{r} \alpha_i y_i^2 = 2py_{r+1}, \alpha_i = -\lambda_i, 1 \leqslant r \leqslant n \tag{3.14}$$

Đây chính là dạng III. Ta có điều phải chứng minh. ▶

Sau đây ta đưa ra tên gọi của một số siêu mặt bậc hai:

i) Siêu mặt có dạng I với $r = n$ với các $\alpha_i > 0, i = \overline{1;n}$, gọi là *siêu mặt Ellipsoid* $n-1$-chiều. Phương trình này có thể viết lại dưới dạng

$$\sum_{i=1}^{n} \frac{x_i^2}{a_i^2} = 1$$

ii) Siêu mặt có dạng I với $r = n$ với các $\alpha_i, i = \overline{1;n}$, khác dấu nhau gọi là *siêu mặt Hyperboloid* và có thể viết lại dưới dạng

$$\sum_{i=1}^{k} \frac{x_i^2}{a_i^2} - \sum_{j=k+1}^{n} \frac{x_j^2}{a_j^2} = 1$$

iii) Siêu mặt bậc hai có phương trình dạng II với $r = n$ và các hệ số $\alpha_i, i = \overline{1;n}$, mang dấu khác nhau gọi là *siêu mặt nón* (thực).

iv) Siêu mặt bậc hai có phương trình dạng III với $r = n-1$ và các hệ số $\alpha_i, i = \overline{1;n-1}$, cùng dấu gọi là *siêu mặt Paraboloid Elliptic*, còn các hệ số $\alpha_i, i = \overline{1;n-1}$, có dấu khác nhau gọi là *siêu mặt Paraboloid Hyperbolic.*

v) Siêu mặt bậc hai có phương trình dạng I, II với $r < n$ và dạng III với $r < n-1$ gọi là các *siêu mặt trụ* (Elliptic, Hyeperbolic, Parabolic...)

Cũng phải chú ý rằng đôi khi các siêu mặt suy biến thành một điểm, hoặc tập rỗng (còn gọi là *siêu mặt ảo*).

3.5.2 Phân loại các đường cong và mặt cong bậc hai

Đường cong bậc hai trên mặt phẳng

Trong trường hợp không gian Euclide $n = 2$ chiều ta có các đường cong bậc hai trên mặt phẳng Euclide, và $n = 3$ ta có các mặt cong bậc hai trong không gian Euclide. Chúng ta hiểu mặt phẳng hoặc không gian Euclide như là các không gian tọa độ thực trên đó có trang bị tích vô hướng, mỗi vector đồng nhất với một điểm, vector 0 đồng nhất với gốc O. Chúng ta thường sử dụng hệ truc tọa độ trực chuẩn Descartes (đã quen thuộc ở bậc học dưới) để mô tả trong trường hợp hai hoặc ba chiều. Trên không gian đó có thể xây dựng các khái niệm khoảng cách, góc... như đã nói trong các bài trước.

Đường bậc hai tổng quát trên mặt phẳng với hệ tọa độ trực chuẩn Descartes Oxy có dạng

$$a_{11}x^2 + 2\mathrm{a}_{12}xy + a_{22}y^2 + a_1x + a_2y + c = 0 \tag{3.15}$$

ở đó các hệ số a_{11}, a_{12}, a_{22} không đồng thời bằng 0. Ma trận của dạng toàn phương là $A = A^T$ có thể đưa về dạng chéo bằng biển bổi trực giao, nghĩa là tồn tại ma trận C sao cho

$$C^T AC = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

ở đó C là ma trận trực giao. Vì C là cấp hai nên nó chỉ có một trong hai dạng trong ví dụ đã đưa ra ở bài chéo hóa trực giao, ta chọn

$$C = \begin{pmatrix} cos\varphi & -\sin\varphi \\ \sin\varphi & cos\varphi \end{pmatrix}$$

và xét phép đổi biến

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

Ta có thể đưa dạng toàn phương về dạng chính tắc

$$\lambda_1 x'^2 + \lambda_2 y'^2$$

Phép đổi biến ở trên thực chất là quay hệ tọa độ ban đầu đi một góc φ và khi đó đường bậc hai có dạng

$$\lambda_1 x'^2 + 2\lambda_2 y'^2 + a_1' x' + a_2' y' + c = 0 \tag{3.16}$$

Tiếp tục sử dụng phép tịnh tiến gốc ta có thể đưa về một trong các loại sau đây

1) Ellipse (hoặc đường tròn)

$$\frac{x^2}{\mathrm{a}^2} + \frac{y^2}{b^2} = 1$$

2) Hyperbola

$$\frac{x^2}{\mathrm{a}^2} - \frac{y^2}{b^2} = 1$$

3) Ellipse ảo

$$\frac{x^2}{\mathrm{a}^2} + \frac{y^2}{b^2} = -1$$

4) Cặp đường thẳng ảo cắt nhau (tại một điểm thực)

$$\frac{x^2}{\mathrm{a}^2} + \frac{y^2}{b^2} = 0$$

5) Cặp đường thẳng cắt nhau

$$\frac{x^2}{\mathrm{a}^2} - \frac{y^2}{b^2} = 0$$

6) Parabola

$$x^2 = 2py$$

7) Cặp đường thẳng song song

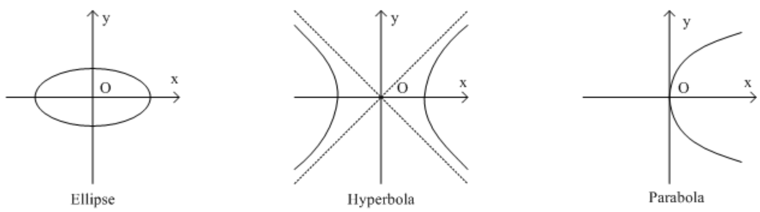
$$\frac{x^2}{\mathrm{a}^2} = 1$$

8) Cặp đường thẳng ảo song song

$$\frac{x^2}{\mathrm{a}^2} = -1$$

9) Cặp đường thẳng trùng nhau

$$\frac{x^2}{\mathrm{a}^2} = 0$$



Hình 3.3: Các đường Conic

Các đường bậc hai Ellipse, Hyperbola, Parabola là những đường Conic đã được học trong bậc học phổ thông.

Mặt cong bậc hai trong không gian

Trong không gian hệ tọa độ trực chuẩn Decartes $Oxyz$ mặt bậc hai tổng quát là tập các điểm thỏa mãn phương trình đại số

$$a_{11}x^2 + 2\mathrm{a}_{12}xy + 2\mathrm{a}_{13}x\mathrm{z} + a_{22}y^2 + 2\mathrm{a}_{23}yz + a_{33}z^2 + 2\mathrm{a}_1x + 2\mathrm{a}_2y + 2\mathrm{a}_3z + c = 0. \tag{3.17}$$

Xét dạng toàn phương với ma trận là $A = (a_{ij})_{3\times 3}$, ở đó $A = A^T$. Tồn tại ma trận chuyển cơ sở C để

$$C^TAC = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

Ma trận C là ma trận trực giao, có thể chọn C sao cho $\det(C) = 1$, ví dụ

$$C = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Phép đổi biến

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = C \begin{pmatrix} \mathrm{x}' \\ y' \\ z' \end{pmatrix}$$

là phép quay hệ trục tọa độ đi một góc φ. Mặt cong khi đó có dạng

$$\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + 2\mathrm{a'}_1 x' + 2\mathrm{a}'_2 y' + 2\mathrm{a}'_3 z' + c = 0 \tag{3.18}$$

Tịnh tiến gốc tọa độ nếu cần ta sẽ đưa mặt bậc hai về một trong các dạng sau

1) Ellipsoid (cầu)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

2) Ellipsoid ảo

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1$$

3) Nón ảo

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$$

4) Hyperboloid 1 tầng

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

5) Hyperboloid 2 tầng

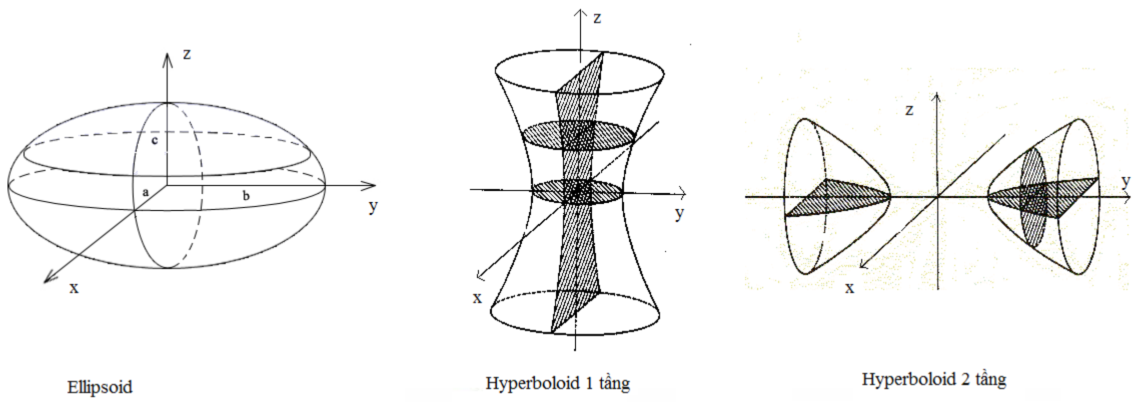
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

6) Nón Elliptic

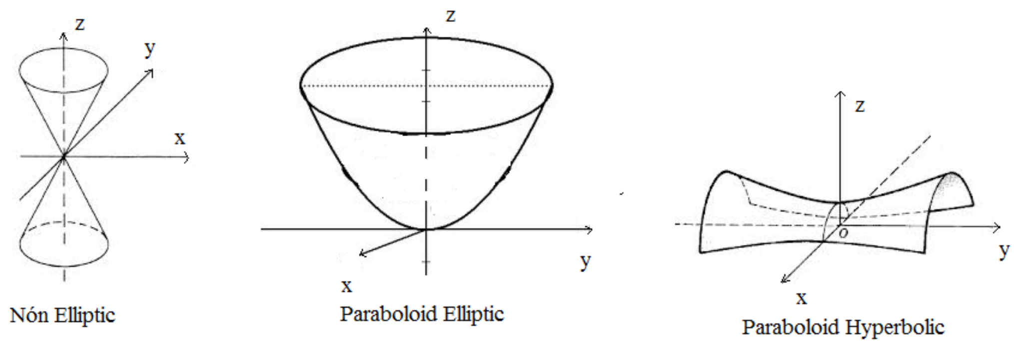
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$

7) Paraboloid Elliptic

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 2z$$



Hình 3.4: Mặt Elipsoid và Hyperboloid 1, 2 tầng



Hình 3.5: Nón Elliptic, Paraboloid Elliptic, Paraboloid Hyperbolic

8) Paraboloid Hyperbolic (yên ngựa)

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2z$$

9) Trụ Elliptic

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

10) Trụ Elliptic ảo

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1$$

11)Trụ Parabolic

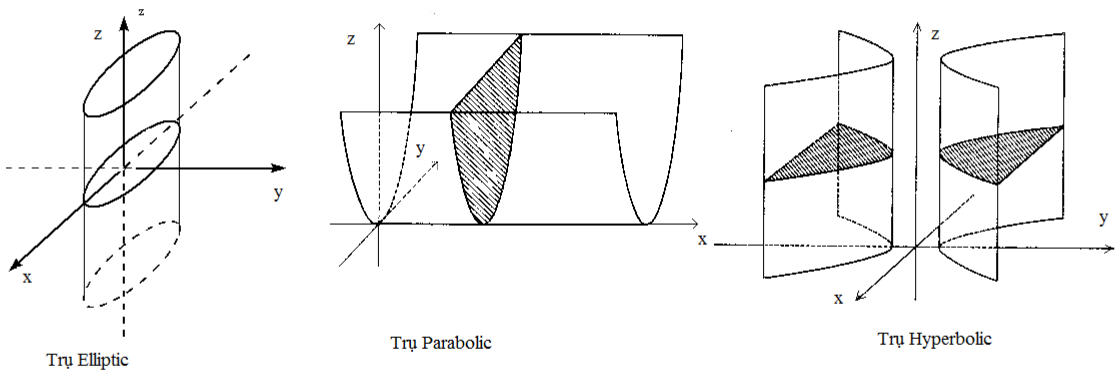
$$y^2 = 2px$$

12) Trụ Hyperbolic

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

13) Cặp mặt phẳng ảo liên hợp

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$$



Hình 3.6: Các mặt trụ

14) Cặp mặt phẳng cắt nhau

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$$

15) Cặp mặt phẳng thực song song

$$\frac{x^2}{a^2} = 1$$

16) Cặp mặt phẳng ảo song song

$$\frac{x^2}{a^2} = -1$$

17) Cặp mặt phẳng trùng nhau

$$\frac{x^2}{a^2} = 0$$

3.6 Thực hành tính toán trên Maple

Như các chương trước chúng ta vẫn làm việc trong môi trường *linalg*.

- Để tính tích vô hướng của hai vector u, v ta dùng lệnh $dotprod(u, v);$ hoặc $dotprod(u, v, orthogonal);$

- Tìm cơ sở trực giao của không gian vector sinh bởi một họ các vector bằng lệnh $GramSchmidt(\{v1, v2, ...\});$

Ví dụ 93.

$> v1 := vector([1, 2, 3]);$

$$v1 := [1\ 2\ 3]$$

$> v2 := vector([-1, 0, -2]);$

$$v1 := [-1\ 0\ \ -2]$$

$> v3 := vector([2, 1, -3]);$

$$v3 := [2\ 1\ \ -3]$$

$> dotprod(v1, v2);$

$$-7$$

$> GramSchmidt(v1, v2, v3);$

$$\left\{[1, 2, 3], \left[-\frac{1}{2}, 1, -\frac{1}{2}\right], \left[\frac{20}{7}, \frac{5}{7} - \frac{10}{7}\right]\right\}$$

Để vẽ đồ thị đường cong và mặt cong trong Maple ta cần dùng gói lệnh *plots*.

- Vẽ đồ thì hàm số $y = f(x)$ ta dùng lệnh có cú pháp

$> plot(f(x), x = a..b, y = c..d,$title='tiêu đề');

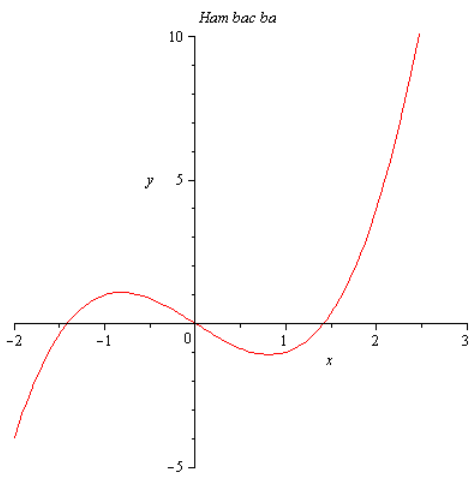
Ví dụ 94. Vẽ đồ thị hàm số $y = x^3 - 2x$

$> restart;$

$> with(plots);$

$> plot(x^3 - 2 * x, x = -2..3, y = -5..10, title =' Hambacba');$

Ta có kết quả trên Hình 3.7.



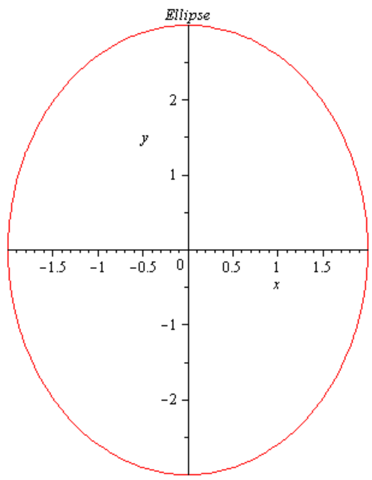
Hình 3.7: Đồ thị hàm bậc ba

- Vẽ đồ thị hàm ẩn $f(x, y) = 0$ dùng cú pháp lệnh

$> implicitplot(f(x, y) = 0, x = a..b, y = c..d);$

Ví dụ 95. Vẽ đồ thị ellipse $\frac{x^2}{4} + \frac{y^2}{9} = 1$ (Hình 3.8)

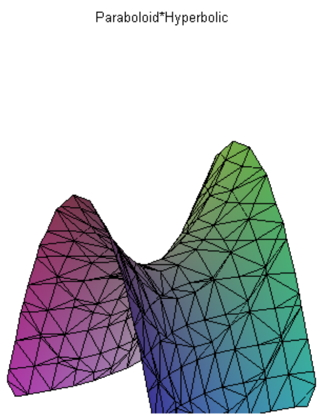
$> implicitplot((1/4) * x^2 + (1/9) * y^2 = 1, x = -2..2, y = -3..3, title =' Ellipse');$



Hình 3.8: Ellipse

- Để vẽ mặt cong bậc hai ta dùng lệnh *implicitplot3d*.

Ví dụ 96. Vẽ mặt yên ngựa $\frac{x^2}{4} - \frac{y^2}{9} = 2z$ (Hình 3.9)



Hình 3.9: Paraboloid Hyperbolic

Bảng 3.1: Bảng trị riêng và vector riêng của ma trận đặc biệt

Ma trận	Trị riêng λ	Vector riêng e
Ma trận đối xứng $A = A^T$	$\forall\lambda \in \mathbb{R}$	$e_i^T e_j = 0$
Ma trận trực giao $A^T = A^{-1}$	$\forall\lvert\lambda\rvert = 1$	$\overline{e_i}^T e_j = 0$
Ma trận phản đối xứng thực $A^T = -A$	$\forall\lambda$ thuần ảo	$\overline{e_i}^T e_j = 0$
Ma trận Hermite (phức) $\overline{A}^T = A$	$\forall\lambda \in \mathbb{R}$	$\overline{e_i}^T e_j = 0$
Ma trận xác định dương $x^T A x > 0$	$\forall\lambda > 0$	$e_i{}^T e_j = 0$
Ma trận đồng dạng $B = M^{-1}AM$	$\lambda_A = \lambda_B$	$e_A = Me_B$
Ma trận chiếu $P = P^2 = P^T$	$\lambda = 1;\ 0$	K.gian sinh bởi các cột; $KerP$
Ma trận hạng 1 tức $A = uv^T$	$\lambda = v^T u;\ 0, ..., 0$	u; Toàn bộ k.gian $u^\perp$
Ma trận nghịch đảo A^{-1}	$1/\lambda_A$	Giữ nguyên vector riêng của A
Ma trận dịch chuyển $A + cE$	$\lambda_A + c$	Giữ nguyên vector riêng của A

Tài liệu tham khảo

[1] Lê Tuấn Hoa. Đại số tuyến tính qua các ví dụ và bài tập // NXB ĐHQGHN. 2005. 406T.

[2] Nguyễn Đình Trí (Chủ biên). Toán cao cấp. Tập 1. Đại số tuyến tính và hình học giải tích // NXB GD. 1997.

[3] Nguyễn Xuân Viên. Đại số tuyến tính // NXB Học viện KTQS. 2014. 252T.

[4] Nguyễn Xuân Viên (Chủ biên). Bài tập đại số tuyến tính và hình học giải tích // NXB QĐND. 2010. 251T.

[5] A. R. G. Heesterman. Matrices and Their Roots: A Textbook of Matrix Algebra // WS. 1990. 444p.

[6] G. Hadley. Linear Algebra // Addison-Wesley Publishing Company, INC. 1961. 290p.

[7] S. Lang. Linear Algebra. Third Edition // Springer. 2004. 296p.

[8] D. Serre. Matrices: Theory and Applications // Springer. 2002. 219p.

[9] G. Strang. Introduction to Linear Algebra, 4th, Wellesley-Cambridge Press. 2009. 574p.

[10] Maple Getting Started Guide. Copyright © Maplesoft, a division of Waterloo Maple Inc. 2005.

[11] П.С. Александров. Курс аналитической геометрии и линейной алгебры //Изд. Наука. 1979. 511с.

[12] Ильин В.А., Позняк Э.Г. Аналитическая геометрия // Изд. Наука. 1999. 223с.

[13] Ильин В.А., Позняк Э.Г. Линейная алгебра // Изд. Наука. 1999. 294с.

[14] Ф.Р. Гантмахер. Теория Матриц // Изд. Наука. 1966. 576с.

[15] И. М. Гельфанд. Лекции по линейной алгебре // Изд. Наука. 1974. 271с.

[16] Мальцев А.И. Основы линейной алгебры, М., Наука 1970. 400с.

Index