BÙI XUÂN HẢI (chủ biên)

ĐẠI SỐ TUYẾN TÍNH VÀ ỨNG DỤNG (Tập 2)

09/2019

BÙI XUÂN HẢI (chủ biên), TRẦN NGỌC HỘI,
TRỊNH THANH ĐÈO, LÊ VĂN LUYỆN

Khoa Toán - Tin học
Trường Đại học Khoa học Tự nhiên
ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH

Đại số tuyến tính và ứng dụng

(Tập 2)

(LƯU HÀNH NỘI BỘ)

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN, ĐHQG-HCM, 09/2019

Mục lục

Chương 6

DẠNG CHÍNH TẮC JORDAN

Trong toàn bộ chương này, ta ký hiệu K là một trường, V là không gian vec tơ n chiều trên K. Hơn nữa, nếu không giải thích gì thêm thì ta ký hiệu f là một toán tử tuyến tính trên V ($f \in L_K(V)$) và A là ma trận vuông cấp n trên K ($A \in M_n(K)$).

§1. Đa thức triệt tiêu

Định nghĩa 1.1. Cho đa thức $Q(t) = a_m t^m + \cdots + a_1 t + a_0 \in K[t]$.

i) Nếu $f \in L_K(V)$ thì toán tử $Q(f) := a_m f^m + \cdots + a_1 f + a_0 \mathrm{Id}_V$ được gọi là *đa thức theo toán tử f*;

ii) Nếu $A \in M_n(K)$ thì ma trận $Q(A) := a_m A^m + \cdots + a_1 A + a_0 I_n$ được gọi là *đa thức theo ma trận A*.

Ví dụ 1. Nếu $f \in L(\mathbb{R}^2)$ xác định bởi $f(x,y) = (x+2y, -x+y)$ và $Q(t) := 2t^2 - 3t + 4$ thì $Q(f) = 2f^2 - 3f + 4\mathrm{Id}_{\mathbb{R}^2}$. Mà

$$\begin{aligned}
\mathrm{Id}_{\mathbb{R}^2}(x;y) &= (x,y);\\
f(x,y) &= (x+2y, -x+y);\\
f^2(x,y) &= f(x+2y, -x+y) = (-x+4y, -2x-y).
\end{aligned}$$

Do đó

$$\begin{aligned}
Q(f)(x,y) &= 2(-x+4y, -2x-y) - 3(x+2y, -x+y) + 4(x,y)\\
&= (-x+2y, -x-y).
\end{aligned}$$

Ví dụ 2. Nếu $A = \begin{pmatrix} 1 & 2 \\ 3 & -2 \end{pmatrix}$ và $Q(t) := 2t^2 - 3t + 4$ thì

$$\begin{aligned} Q(A) &= 2A^2 - 3A + 4I_2 \\ &= 2\begin{pmatrix} 7 & -2 \\ -3 & 10 \end{pmatrix} - 3\begin{pmatrix} 1 & 2 \\ 3 & -2 \end{pmatrix} + 4\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 15 & -10 \\ -15 & 30 \end{pmatrix}. \end{aligned}$$

Mệnh đề 1.2. *Nếu $P, Q \in K[t]$ thì*

- $P(f)\circ Q(f) = Q(f)\circ P(f), \forall f \in L_K(V)$;
- $P(A).Q(A) = Q(A).P(A), \forall A \in M_n(K)$.

Định nghĩa 1.3. Cho $f \in L_K(V), A \in M_n(K), Q(t) \in K[t]$. Khi đó:

i) Nếu $Q(f) = 0$ thì ta nói $Q(t)$ là *đa thức triệt tiêu* toán tử f;

ii) Nếu $Q(A) = 0$ thì ta nói $Q(t)$ là *đa thức triệt tiêu* ma trận A.

Mệnh đề 1.4. i) *Nếu $Q(t) \in K[t]$ là đa thức triệt tiêu toán tử $f \in L_K(V)$ và λ là trị riêng của f thì λ là nghiệm của $Q(t)$.*

ii) *Nếu $Q(t) \in K[t]$ là đa thức triệt tiêu ma trận $A \in M_n(K)$ và λ là trị riêng của A thì λ là nghiệm của $Q(t)$.*

Chứng minh. i) Giả sử v là một vectơ riêng của f ứng với trị riêng λ và $Q(t) = a_m t^m + \cdots + a_1 t + a_0$ là đa thức triệt tiêu f. Khi đó

$$a_m f^m + \cdots + a_1 f + a_0 \mathrm{Id}_V = 0,$$

nên

$$(a_m f^m + \cdots + a_1 f + a_0 \mathrm{Id}_V)(v) = 0,$$

nghĩa là

$$a_m f^m(v) + \cdots + a_1 f(v) + a_0 \mathrm{Id}_V(v) = 0. \tag{1}$$

Mặt khác, do v là trị riêng của f nên $f(v) = \lambda v$. Suy ra

$$f^k(v) = \lambda^k v, \forall k \in \mathbb{N}.$$

Do đó (1) kéo theo $(a_m\lambda^m + \cdots + a_1\lambda + a_0)v = 0$. Mà $v \neq 0$ nên $a_m\lambda^m + \cdots + a_1\lambda + a_0 = 0$, nghĩa là $Q(\lambda) = 0$.

ii) Chứng minh tương tự i). $\square$

Từ mệnh đề trên ta thấy rằng, nếu toán tử f thỏa mãn $f^2 = f$ thì các trị riêng của f chỉ có thể là 0 hoặc 1; nếu $f^3 = f$ thì các trị riêng của f chỉ có thể là $0, 1$ hoặc -1. Tuy nhiên, không phải tất cả các nghiệm của $Q(t)$ đều là trị riêng của f. Chẳng hạn, nếu $f = \mathrm{Id}_V$ thì đa thức $Q(t) = t^2 - t$ triệt tiêu f nhưng 0 không phải là trị riêng của f.

Câu hỏi đầu tiên mà ta có thể đặt ra là: Phải chăng mọi toán tử tuyến tính $f \in L_K(V)$ đều tồn tại một đa thức $0 \neq Q(t) \in K[t]$ triệt tiêu f? Câu trả lời là khẳng định. Thật vậy, nếu $\dim_K(V) = n$ thì $L_K(V) \cong M_n(K)$, suy ra $\dim_K(L_K(V)) = n^2$. Do đó các phần tử $\mathrm{Id}_V, f, f^2, \ldots, f^{n^2}$ phụ thuộc tuyến tính trong $L_K(V)$, suy ra tồn tại các phần tử $a_0, a_1, a_2, \ldots, a_{n^2} \in K$, không đồng thời bằng 0 sao cho

$$a_0\mathrm{Id}_V + a_1 f + a_2 f^2 + \cdots + a_{n^2} f^{n^2} = 0.$$

Vậy $Q(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_{n^2} t^{n^2}$ là đa thức triệt tiêu f.

Định lý 1.5 (Cayley-Hamilton). i) *Nếu $f \in L_K(V)$ thì đa thức đặc trưng $P_f(t)$ của f là đa thức triệt tiêu f, nghĩa là $P_f(f) = 0$;*

ii) *Nếu $A \in M_n(K)$ thì đa thức đặc trưng $P_A(t)$ của A là đa thức triệt tiêu A, nghĩa là $P_A(A) = 0$.*

Chứng minh. Ta chỉ cần chứng minh i). Gọi $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ là một cơ sở của V và $A = [f]_{\mathcal{B}}$. Như vậy,

$$f(u_i) = \sum_{j=1}^{n} a_{ji} u_j,\ 1 \leq i \leq n.$$

Hơn nữa $u_i = \sum\limits_{j=1}^{n} \delta_{ij} u_j$ với $\delta_{ij} = \begin{cases} 1, & \text{nếu } i = j \\ 0, & \text{nếu } i \neq j, \end{cases}$ nên với mọi i, $1 \leq i \leq n$, ta có

$$\sum_{j=1}^{n} (a_{ji}\mathrm{Id}_V - \delta_{ij} f)(u_j) = 0.$$

Gọi $B = (b_{ij})$ là ma trận vuông cấp n với các hệ số là các toán tử tuyến tính được xác định bởi $b_{ij} = a_{ji}\mathrm{Id}_V - \delta_{ij} f$. Do

$$P_f(t) = \det(A - tI_n) = \det(A - tI_n)^{\top},$$

mà ma trận này có các hệ số là những đa thức

$$[(A - tI_n)^{\top}]_{ij} = a_{ij}\mathrm{Id}_V - \delta_{ij} t,$$

nên $P_f(f) = \det B$. Như vậy để chứng minh $P_f(f) = 0$, ta cần chứng minh

$$(\det B)(u_k) = 0 \text{ với mọi } k = 1, \ldots, n.$$

Thật vậy, do định nghĩa của B, các vectơ $u_1, u_2, \ldots u_n$ thỏa mãn:

$$\sum_{j=1}^{n} b_{ij}(u_j) = 0, \quad 1 \leq i \leq n.$$

Gọi $\overline{B} = \mathrm{adj}(B)$ là ma trận phó của B. Khi đó

$$(\det B) I_n = \overline{B} B.$$

Suy ra, với mọi $1 \leq k \leq n$ ta có

$$\delta_{kj} \det B = \sum_{i=1}^{n} \overline{B}_{ki} b_{ij}.$$

Lấy ảnh của u_j theo hai vế, rồi lấy tổng theo j,

$$\sum_{j=1}^{n} \delta_{kj} \det B(u_j) = \sum_{j=1}^{n} \Big(\sum_{i=1}^{n} \overline{B}_{ki} b_{ij}(u_j) \Big).$$

Suy ra $\det B(u_k) = \sum\limits_{i=1}^{n} \overline{B}_{ki} \Big(\sum\limits_{j=1}^{n} b_{ij}(u_j) \Big) = 0.$ $\square$

Dưới đây là một kết quả rất quan trọng về các đa thức triệt tiêu.

Bổ đề 1.6 (Bổ đề căn bản). *Cho $f \in L_K(V)$ và $Q(t)$ là đa thức triệt tiêu f. Khi đó, nếu $Q_1, \dots, Q_p$ là các đa thức nguyên tố cùng nhau sao cho $Q(t) = Q_1(t) \dots Q_p(t)$ thì*

$$V = \mathrm{Ker} Q_1(f) \oplus \cdots \oplus \mathrm{Ker} Q_p(f).$$

Phát biểu tương tự cho đa thức triệt tiêu ma trận.

Chứng minh. Ta chứng minh bằng qui nạp theo p.

+) *Trường hợp* $p = 1$:

Với $p = 1$ ta có $Q(t) = Q_1(t)$, nên $Q_1(f) = Q(f) = 0$ và $V = \mathrm{Ker} Q_1(f)$.

+) *Trường hợp* $p = 2$:

Giả sử $Q(t) = Q_1(t)Q_2(t)$, trong đó Q_1 và Q_2 nguyên tố cùng nhau. Khi đó, tồn tại các đa thức U_1 và U_2 sao cho $U_1Q_1 + U_2Q_2 = 1$. Suy ra

$$U_1(f)Q_1(f) + U_2(f)Q_2(f) = \mathrm{Id}_V.$$

Do đó, với mọi $u \in V$ ta có

$$u = U_1(f)Q_1(f)(u) + U_2(f)Q_2(f)(u), \tag{1}$$

kéo theo

$$V = \mathrm{Im}(U_1(f)Q_1(f)) + \mathrm{Im}(U_2(f)Q_2(f)). \tag{2}$$

Vì $Q_1(f)Q_2(f) = Q(f) = 0$ nên $Q_2(f)U_1(f)Q_1(f) = 0$, suy ra

$$\mathrm{Im}(U_1(f)Q_1(f)) \subseteq \mathrm{Ker} Q_2(f).$$

Hoàn toàn tương tự, ta có $\mathrm{Im}(U_2(f)Q_2(f)) \subseteq \mathrm{Ker} Q_1(f)$. Do đó, từ (2) ta được

$$V = \mathrm{Ker} Q_1(f) + \mathrm{Ker} Q_2(f).$$

Giả sử $u \in \mathrm{Ker} Q_1(f) \cap \mathrm{Ker} Q_2(f)$. Từ (1) suy ra ngay $u = 0$. Vậy

$$V = \mathrm{Ker} Q_1(f) \oplus \mathrm{Ker} Q_2(f).$$

+) *Trường hợp* $p > 2$:

Ta có $Q(t) = (Q_1(t)\dots Q_{p-1}(t))Q_p(t)$. Đặt

$$\widetilde{Q}(t) = Q_1(t)\dots Q_{p-1}(t),$$

ta có $\widetilde{Q}(t)$ và $Q_p(t)$ nguyên tố cùng nhau. Theo trường hợp $p = 2$ ta có

$$V = \mathrm{Ker}\widetilde{Q}(f) \oplus \mathrm{Ker}Q_p(f).$$

Đặt $W = \mathrm{Ker}\widetilde{Q}(f)$ và $\overline{f} = f|_W$. Ta chứng minh $\overline{f} \in L_K(W)$. Thật vậy, $\forall u \in W$, ta có $\overline{f}(u) = f(u)$. Do $u \in W$ nên $\widetilde{Q}(f)(u) = 0$, suy ra

$$\widetilde{Q}(f)f(u) = f\widetilde{Q}(f)(u) = 0.$$

Hơn nữa $\widetilde{Q}(f)f(u) = \widetilde{Q}(f)(f(u))$, do đó $f(u) \in \mathrm{Ker}\widetilde{Q}(f)$.

Ngoài ra, do $\overline{f} = f|_W$ nên $\widetilde{Q}(\overline{f}) = 0$. Áp dụng giả thiết qui nạp, ta được

$$W = \mathrm{Ker}Q_1(\overline{f}) \oplus \dots \oplus \mathrm{Ker}Q_{p-1}(\overline{f}).$$

Với mọi $i \in \overline{1, p-1}$, ta có

$$\begin{aligned}\mathrm{Ker}Q_i(\overline{f}) &= \{u \in W \mid Q_i(\overline{f})(u) = 0\}\\ &= \{u \in W \mid Q_i(f)(u) = 0\} \subseteq \mathrm{Ker}Q_i(f).\end{aligned}$$

Ngược lại, giả sử $u \in \mathrm{Ker}Q_i(f)$. Khi đó,

$$\widetilde{Q}(f)(u) = Q_1(f)\dots Q_{i-1}(f)Q_{i+1}(f)\dots Q_{p-1}(f)Q_i(f)(u) = 0,$$

suy ra $u \in W$. Điều này chứng tỏ $u \in \mathrm{Ker}Q_i(\overline{f})$. Do đó

$$\mathrm{Ker}Q_i(f) = \mathrm{Ker}Q_i(\overline{f}), \forall i \in \overline{1, p-1}.$$

Vậy $V = \mathrm{Ker}Q_1(f) \oplus \dots \oplus \mathrm{Ker}Q_{p-1}(f) \oplus \mathrm{Ker}Q_p(f)$. □

Hệ quả 1.7. *Cho $f \in L_K(V)$ và $P_f(t)$ là đa thức đặc trưng của f. Giả sử $P_f(t)$ phân rã trên K thành*

$$P_f(t) = (-1)^n (t - \lambda_1)^{\alpha_1} \dots (t - \lambda_p)^{\alpha_p}.$$

Khi đó $V = \mathrm{Ker}(f - \lambda_1 \mathrm{Id})^{\alpha_1} \oplus \cdots \oplus \mathrm{Ker}(f - \lambda_p \mathrm{Id})^{\alpha_p}$.

Với các ký hiệu như trên, ta gọi

$$N(\lambda_i) := \mathrm{Ker}(f - \lambda_i \mathrm{Id})^{\alpha_i}$$

là *không gian đặc trưng* của f ứng với trị riêng λ_i.

Mệnh đề 1.8. *Cho $f \in L_K(V)$ và λ là trị riêng của f. Khi đó:*

i) $E(\lambda) \subseteq N(\lambda)$.

ii) $f(N(\lambda)) \subseteq N(\lambda)$.

Chứng minh. i) Nếu $u \in E(\lambda)$ thì $(f - \lambda \mathrm{Id})(u) = 0$, nên $(f - \lambda \mathrm{Id})^{\alpha}(u) = 0$ (α là số bội của λ), hay $u \in N(\lambda)$. Do đó $E(\lambda) \subseteq N(\lambda)$.

ii) Giả sử $u \in N(\lambda)$. Khi đó $(f - \lambda \mathrm{Id})^{\alpha}(u) = 0$ nên $f(f - \lambda \mathrm{Id})^{\alpha}(u) = 0$, nghĩa là $(f - \lambda \mathrm{Id})^{\alpha} f(u) = 0$. Do đó $f(u) \in N(\lambda)$. □

Ví dụ 3. Giả sử $f \in L_K(V)$ sao cho $f^2 = f$. Khi đó $Q(t) = t(t-1)$ là đa thức triệt tiêu f. Vậy, theo Bổ đề 1.6, $V = \mathrm{Ker} f \oplus \mathrm{Ker}(f - \mathrm{Id})$. Ta thấy V là tổng trực tiếp của các không gian riêng của f, suy ra f chéo hóa được.

Dưới đây là một kết quả tổng quát hơn ví dụ vừa xét.

Định lý 1.9. *Toán tử tuyến tính f chéo hóa được khi và chỉ khi tồn tại một đa thức phân rã trên K, có toàn nghiệm đơn và triệt tiêu f.*

Chứng minh. Giả sử f chéo hóa được. Khi đó tồn tại cơ sở $\mathcal{B} = (u_1, \dots, u_n)$ của V gồm toàn các vectơ riêng của f. Gọi $\lambda_1, \dots, \lambda_p$ là các trị riêng đôi một khác nhau của f. Khi đó, với mọi $u \in \mathcal{B}$, tồn tại một $\lambda_i, 1 \leq i \leq p$ sao cho $(f - \lambda_i \mathrm{Id})(u) = 0$. Suy ra

$$(f - \lambda_1 \mathrm{Id}) \dots (f - \lambda_p \mathrm{Id})(u) = 0, \forall u \in \mathcal{B}.$$

Vậy đa thức $Q(t) = (t - \lambda_1) \dots (t - \lambda_p)$ phân rã trên K, chỉ có toàn nghiệm đơn và triệt tiêu f.

Ngược lại, giả sử $Q(t) = (t - \lambda_1) \dots (t - \lambda_p)$ là đa thức tiệt tiêu f (với $\lambda_i \neq \lambda_j, \forall i \neq j$). Khi đó theo Bổ đề 1.6, ta có

$$V = \mathrm{Ker}(f - \lambda_1 \mathrm{Id}) \oplus \dots \oplus \mathrm{Ker}(f - \lambda_p \mathrm{Id}) = E(\lambda_1) \oplus \dots \oplus E(\lambda_p).$$

Suy ra f chéo hóa được. □

§2. Đa thức tối tiểu

Định nghĩa 2.1. Đa thức có hệ số bậc cao nhất bằng 1 được gọi là *đa thức đơn khởi*. Đa thức đơn khởi có bậc nhỏ nhất triệt tiêu toán tử tuyến tính f được gọi là *đa thức tối tiểu* của f, ký hiệu bởi m_f.

Mệnh đề 2.2. *Đa thức $Q(t) \in K[t]$ triệt tiêu f khi và chỉ khi $Q(t)$ chia hết cho $m_f(t)$ trong $K[t]$.*

Chứng minh. Giả sử $Q(f) = 0$, ta chia $Q(t)$ cho $m_f(t)$:

$$Q(t) = P(t).m_f(t) + R(t), \quad \deg(R) < \deg(m_f).$$

Do $Q(f) = 0$ nên $R(f) = 0$, kéo theo $R(t) = 0$ (do m_f là đa thức tối tiểu của f). Do đó

$$Q(t) = P(t) m_f(t).$$

Ngược lại, nếu $Q(t) = P(t) m_f(t)$ thì

$$Q(f) = P(f) m_f(f) = 0,$$

nên $Q(t)$ triệt tiêu f. □

Hệ quả 2.3. *Cho $f \in L_K(V)$. Khi đó $m_f(t)$ là ước của $P_f(t)$.*

Chứng minh. Áp dụng Định lý Cayley-Hamilton và Mệnh đề 2.2 □

Hệ quả 2.4. *Đa thức tối tiểu của một toán tử tuyến tính là duy nhất.*

Chứng minh. Giả sử m_1 và m_2 là hai đa thức tối tiểu của toán tử f. Khi đó, theo Mệnh đề 2.2, m_1 là ước của m_2 và m_2 là ước của m_1. Do m_1 và m_2 đều là các đa thức đơn khởi nên $m_1 = m_2$. □

Mệnh đề 2.5. *Cho $f \in L_K(V)$. Khi đó tập nghiệm của $m_f(t)$ trùng với tập nghiệm của $P_f(t)$.*

Chứng minh. Hệ quả 2.3 suy ra mỗi nghiệm của $m_f(t)$ đều là nghiệm của $P_f(t)$; Mệnh đề 1.4 suy ra mỗi nghiệm của $P_f(t)$ đều là nghiệm của $m_f(t)$. □

Ví dụ 4. Cho $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$. Ta có

$$P_A(t) = -(t+1)(t+2)(t-3).$$

Áp dụng Mệnh đề 2.5 suy ra $m_A(t) = (t+1)(t+2)(t-3)$.

Ví dụ 5. Cho $A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$. Ta có

$$P_A(t) = -(t-1)(t+2)^2.$$

Áp dụng Mệnh đề 2.5 suy ra $m_A(t) = (t-1)(t+2)$ hoặc $m_A(t) = (t-1)(t+2)^2$. Mà $(A - I_3)(A + 2I_3) = 0$, nên $m_A(t) = (t-1)(t+2)$.

Định lý 2.6. *Toán tử tuyến tính $f \in L_K(V)$ chéo hóa được khi và chỉ khi đa thức tối tiểu $m_f(t)$ của f phân rã trên K và có toàn nghiệm đơn.*

Chứng minh. Chiều đảo được suy ra từ Định lý 1.9. Ngược lại, giả sử f chéo hóa được. Khi đó

$$P_f(t) = (-1)^n (t-\lambda_1)^{\alpha_1} \dots (t-\lambda_p)^{\alpha_p}, \ \lambda_i \neq \lambda_j, \forall i \neq j.$$

Theo chứng minh Định lý 1.9 đa thức $Q(t) = (t-\lambda_1)\dots(t-\lambda_p)$ triệt tiêu f. Áp dụng Mệnh đề 2.5 suy ra $m_f(t) = Q(t)$. □

Ví dụ 6. Xét ma trận $A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$.

Đa thức tối tiểu của A là

$$m_A(t) = (t-1)(t-2).$$

Theo Định lý 2.6, A chéo hóa được.

Ví dụ 7. Xét ma trận $A = \begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

Đa thức đặc trưng của A là $P_A(t) = -(t-1)^3$, nên

$$m_A(t) \in \{t-1;\ (t-1)^2;\ (t-1)^3\}.$$

Theo Định lý 2.6, A chéo hóa được khi và chỉ khi $m_A(t) = t-1$, nghĩa là $A - I_3 = 0$. Nhưng do $A \neq I_3$, nên A không chéo hóa được.

Ví dụ 8. Xét ma trận $A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$.

Đa thức đặc trưng của A là

$$P_A(t) = -(t-1)(t-2)^2,$$

do đó

$$m_A(t) \in \{(t-1)(t-2);\ (t-1)(t-2)^2\}.$$

Theo Định lý 2.6, ta có A chéo hóa được khi và chỉ khi $m_A(t) = (t-1)(t-2)$, nghĩa là $(A - I_3)(A - 2I_3) = 0$. Nhưng bằng cách tính toán trực tiếp ta thấy rằng

$$(A - I_3)(A - 2I_3) \neq 0,$$

do đó A không chéo hóa được.

§3. Sự tam giác hóa

Mệnh đề 3.1. *Mọi ma trận tam giác trên đều đồng dạng với một ma trận tam giác dưới nào đó.*

Chứng minh. Giả sử $A \in M_n(K)$ là ma trận tam giác trên. Chọn $f \in L_K(K^n)$ sao cho ma trận biểu diễn f theo cơ sở chính tắc $\mathcal{B}_0 = (e_1, \ldots, e_n)$ là A. Xét cơ sở $\mathcal{B} = (e_n, \ldots, e_1)$, ta có

$$[f]_{\mathcal{B}} = (\mathcal{B}_0 \to \mathcal{B})^{-1}[f]_{\mathcal{B}_0}(\mathcal{B}_0 \to \mathcal{B}).$$

Đặt $A' := [f]_{\mathcal{B}}$ và $P := (\mathcal{B}_0 \to \mathcal{B})$, ta thấy A' là ma trận tam giác dưới và $A' = P^{-1}AP$. Suy ra A' đồng dạng với A. □

Định nghĩa 3.2. Một ma trận $A \in M_n(K)$ được gọi là *tam giác hóa được* nếu A đồng dạng với một ma trận tam giác.

Bài toán đặt ra là khi nào thì ma trận $A \in M_n(K)$ tam giác hóa được? Do Mệnh đề 3.1, ta chỉ cần xét khi nào ma trận A đồng dạng với một ma trận tam giác trên. Theo ngôn ngữ của các toán tử tuyến tính thì vấn đề đặt ra là khi nào một toán tử tuyến tính được biểu diễn bằng một ma trận tam giác trên theo một cơ sở nào đó.

Định lý 3.3. *Toán tử $f \in L_K(V)$ tam giác hóa được khi và chỉ khi đa thức đặc trưng của f phân rã trên K.*

Chứng minh. Giả sử f tam giác hóa được và $\mathcal{B} = (u_1, \ldots, u_n)$ là một cơ sở của V sao cho

$$[f]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Suy ra

$$P_f(t) = \det \begin{pmatrix} \lambda_1 - t & & * \\ & \ddots & \\ 0 & & \lambda_n - t \end{pmatrix} = (\lambda_1 - t) \ldots (\lambda_n - t),$$

nghĩa là $P_f(t)$ phân rã trên K.

Ngược lại, giả sử $P_f(t)$ phân rã trên K. Ta chứng minh bằng qui nạp rằng toán tử f tam giác hóa được.

Nếu $n = 1$ thì không có gì để chứng minh. Giả sử $n > 1$ và khẳng định đúng với $n - 1$. Gọi $\lambda_1 \in K$ là một nghiệm nào đó của $P_f(t)$ và u_1 là một vectơ riêng ứng với trị riêng λ_1. Ta bổ túc thêm $n-1$ vectơ vào (u_1) để có một cơ sở $\mathcal{C} = (u_1, u_2, \ldots, u_n)$ của V. Khi đó

$$A = [f]_{\mathbb{C}} = \left(\begin{array}{c|ccc} \lambda_1 & b_2 & \ldots & b_n \\ \hline & & & \\ 0 & & B & \\ & & & \end{array}\right),$$

với B là ma trận vuông cấp $n - 1$. Xét không gian con $W = \langle u_2, \ldots, u_n \rangle$ và toán tử tuyến tính $g \in L_K(W)$ sao cho ma trận biểu diễn g theo cơ sở $(u_2, \ldots, u_n)$ là B. Ta có

$$P_f(t) = \det(A - tI_n) = (\lambda_1 - t)\det(B - tI_{n-1}) = (\lambda_1 - t)P_g(t).$$

Do $P_f(t)$ phân rã trên K nên $P_g(t)$ cũng phân rã trên K, do đó theo giả thiết qui nạp ma trận B tam giác hóa được. Vậy tồn tại một cơ sở $(v_2, \dots, v_n)$ của W sao cho ma trận biểu diễn g theo cơ sở này là ma trận tam giác trên. Khi đó ma trận biểu diễn f theo cơ sở $(u_1, v_2, \dots, v_n)$ cũng có dạng tam giác trên. □

Hệ quả 3.4. *Mọi ma trận $A \in M_n(\mathbb{C})$ đều tam giác hóa được.*

Nhận xét 3.5. Nếu ma trận A đồng dạng với ma trận tam giác A' thì trên đường chéo chính của A' chỉ toàn là các trị riêng của A.

Hệ quả 3.6. *Cho $A \in M_n(\mathbb{R})$ và $Sp_{\mathbb{C}}(A) = \{\lambda_1, \dots, \lambda_n\}$ là phổ của A trên $\mathbb{C}$. Khi đó:*

i) $tr(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$;

ii) $\det A = \lambda_1 \lambda_2 \dots \lambda_n$.

Chứng minh. Do các ma trận đồng dạng đều có cùng vết và cùng định thức nên những điều cần chứng minh là hiển nhiên. □

Ví dụ 9. Ma trận $A = \begin{pmatrix} -4 & 0 & -2 \\ 0 & 1 & 0 \\ 5 & 1 & 3 \end{pmatrix}$ có đa thức đặc trưng là

$$P_A(t) = -(t+2)(t-1)^2,$$

nên theo Định lý 3.3, A tam giác hóa được trên $\mathbb{R}$. Ta xem A như ma trận biểu diễn toán tử tuyến tính f theo cơ sở chính tắc. Khi đó tồn tại cơ sở $\mathcal{B} = (u_1, u_2, u_3)$ sao cho ma trận biểu diễn f theo $\mathcal{B}$ có dạng tam giác trên

$$[f]_{\mathcal{B}} = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & -2 \end{pmatrix}.$$

Tìm u_1: Nhận xét rằng u_1 chính là vectơ riêng ứng với trị riêng $\lambda_1 = 1$. Xét hệ phương trình $(A - I_3)X = 0$, ta có

$$A - I_3 = \begin{pmatrix} -5 & 0 & -2 \\ 0 & 0 & 0 \\ 5 & 1 & 2 \end{pmatrix} \xrightarrow[d_2 \leftrightarrow d_3]{d_3 + d_1} \begin{pmatrix} -5 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Cho $x_3 = 5$ suy ra $x_1 = -2$. Vậy ta có thể lấy $u_1 = (-2, 0, 5)$.

Tìm u_2: Ta có $f(u_2) = au_1 + u_2$, nên $(f - \mathrm{Id})(u_2) = au_1$. Do đó

$$\begin{pmatrix} -5 & 0 & -2 \\ 0 & 0 & 0 \\ 5 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = a \begin{pmatrix} -2 \\ 0 \\ 5 \end{pmatrix}.$$

Giải hệ phương trình trên

$$\left(\begin{array}{ccc|c} -5 & 0 & -2 & -2a \\ 0 & 0 & 0 & 0 \\ 5 & 1 & 2 & 5a \end{array}\right) \xrightarrow[d_2 \leftrightarrow d_3]{d_3 + d_1} \left(\begin{array}{ccc|c} -5 & 0 & -2 & -2a \\ 0 & 1 & 0 & 3a \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Chọn $a = -1$ và $x_3 = 4$, suy ra $x_1 = -2, x_2 = -3$. Do đó ta có thể lấy $u_2 = (-2, -3, 4)$.

Tìm u_3: Ta biết rằng tồn tại vectơ riêng v ứng với trị riêng $\lambda_2 = -2$, nghĩa là $f(v) = -2v$. Ta có thể chọn $u_3 = v, b = c = 0$. Ta có

$$A + 2I_3 = \begin{pmatrix} -2 & 0 & -2 \\ 0 & 3 & 0 \\ 5 & 1 & 5 \end{pmatrix} \xrightarrow[d_3 - 5d_1]{-\frac{1}{2}d_1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 0 \\ 0 & 1 & 0 \end{pmatrix} \xrightarrow[d_3 - d_2]{\frac{1}{3}d_2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Do đó có thể lấy $u_3 = (-1, 0, 1)$.

Kiểm tra dễ dàng u_1, u_2, u_3 độc lập tuyến tính, do đó chúng tạo thành một cơ sở của $\mathbb{R}^3$. Ma trận biểu diễn f theo cơ sở $\mathcal{B}$ là

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

Ma trận chuyển từ cơ sở chính tắc sang cơ sở $\mathcal{B}$ là

$$P = \begin{pmatrix} -2 & -2 & -1 \\ 0 & -3 & 0 \\ 5 & 4 & 1 \end{pmatrix}.$$

Cuối cùng ta có $A' = P^{-1}AP$.

§4. Dạng tam giác khối

Ta biết rằng, nếu đa thức đặc trưng $P_f(t)$ của toán tử $f \in L_K(V)$ phân rã trên K thì f tam giác hóa được, nghĩa là tồn tại cơ sở $\mathcal{B}$ của V sao cho ma trận biểu diễn f theo $\mathcal{B}$ có dạng tam giác (trên hoặc dưới). Trong mục này ta tiếp tục việc rút gọn toán tử f sao cho có thể "tốt" hơn nữa. Cụ thể, nếu $P_f(t)$ phân rã trên K thì ta có thể đưa f về dạng tam giác khối.

Bổ đề 4.1. *Giả sử $V = V_1 \oplus \ldots \oplus V_p$, trong đó V_i là các không gian con bất biến đối với f. Khi đó, nếu $\mathcal{B}_1, \ldots, \mathcal{B}_p$ tương ứng là các cơ sở của $V_1, \ldots, V_p$ thì ma trận của f theo cơ sở $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_p)$ là*

$$[f]_{\mathcal{B}} = \begin{pmatrix} \boxed{M_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{M_p} \end{pmatrix} =: \mathrm{diag}(M_1, \ldots, M_n),$$

trong đó M_i là ma trận biểu diễn của hạn chế của f lên V_i theo cơ sở $\mathcal{B}_i$.

Chứng minh. Giả sử $\mathcal{B}_1 = (u_{11}, \ldots, u_{1,n_1}), \ldots, \mathcal{B}_p = (u_{p1}, \ldots, u_{p,n_p})$. Do

$f(V_i) \subseteq V_i$ với mọi $i = \overline{1,p}$ nên

$$\begin{cases} f(u_{11}) = a_{11}u_{11} + \ldots + a_{n_1 1}u_{1,n_1}; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(u_{1,n_1}) = a_{1n_1}u_{11} + \ldots + a_{n_1 n_1}u_{1,n_1} \end{cases}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\begin{cases} f(u_{p1}) = b_{11}u_{p1} + \ldots + b_{n_p 1}u_{p,n_p}; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(u_{p,n_p}) = b_{1n_p}u_{p1} + \ldots + b_{n_p n_p}u_{p,n_p}. \end{cases}$$

Do đó $[f]_{\mathcal{B}} = \operatorname{diag}(M_1, \ldots, M_p)$. □

Định lý 4.2 (Rút gọn theo dạng tam giác khối). *Cho $f \in L_K(V)$. Giả sử đa thức đặc trưng của f phân rã trên K thành:*

$$P_f(t) = (-1)^n(t-\lambda_1)^{\alpha_1}\ldots(t-\lambda_p)^{\alpha_p},\ \lambda_i \neq \lambda_j, \forall i \neq j.$$

Khi đó tồn tại cơ sở $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_p)$ của V, trong đó $\mathcal{B}_i$ là một cơ sở của $N(\lambda_i)$ sao cho

$$[f]_{\mathcal{B}} = \operatorname{diag}(M_1, \ldots, M_p),$$

với M_i là ma trận biểu diễn của hạn chế của f lên không gian đặc trưng $N(\lambda_i)$ theo cơ sở $\mathcal{B}_i$ và M_i có dạng tam giác trên.

Chứng minh. Vì $N(\lambda_i)$ bất biến đối với f nên theo Hệ quả 1.7 và Định lý 4.2, tồn tại cơ sở $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_p)$ của V sao cho

$$[f]_{\mathcal{B}} = \operatorname{diag}(M_1, \ldots, M_p),$$

trong đó $\mathcal{B}_i$ là cơ sở của $N(\lambda_i), M_i = [f_i]_{\mathcal{B}_i}$, với f_i là hạn chế của f lên $N(\lambda_i)$. Do đó ta chỉ cần chứng minh rằng M_i tam giác hóa được và $Sp_K(M_i) = \{\lambda_i, \ldots, \lambda_i\}$ là đủ.

Theo định nghĩa, $N(\lambda_i) = \operatorname{Ker}(f - \lambda_i \operatorname{Id})^{\alpha_i}$ nên

$$(f - \lambda_i \operatorname{Id})^{\alpha_i}(u) = 0, \forall u \in N(\lambda_i).$$

Vậy $(\lambda-\lambda_i)^{\alpha_i}$ là đa thức triệt tiêu f_i, do đó đa thức tối tiểu của f_i có dạng

$$m_{f_i}(t) = (t-\lambda_i)^{\gamma_i}, \text{ với } 1 \leq \gamma_i \leq \alpha_i.$$

Áp dụng Mệnh đề 2.2, suy ra đa thức đặc trưng của f_i có dạng

$$P_{f_i}(t) = (-1)^{\delta_i}(t-\lambda_i)^{\delta_i}, \text{ với } \gamma_i \leq \delta_i.$$

Vậy $P_{f_i}(t)$ phân rã trên K và $Sp_K(f_i) = \{\lambda_i, \ldots, \lambda_i\}$, kéo theo M_i tam giác hóa được. Ta còn cần phải chứng minh $\delta_i = \alpha_i, \forall i \in \overline{1,p}$. Thực vậy,

$$\begin{aligned}
P_f(t) &= |M_1 - tI_{\delta_1}| \ldots |M_p - tI_{\delta_p}| \\
&= P_{f_1}(t) \ldots P_{f_p}(t) = (-1)^{\delta_1+\ldots+\delta_p}(t-\lambda_1)^{\delta_1} \ldots (t-\lambda_p)^{\delta_p} \\
&= (-1)^n (t-\lambda_1)^{\alpha_1} \ldots (t-\lambda_p)^{\alpha_p}.
\end{aligned}$$

Do đó $\delta_i = \alpha_i, \forall i \in \overline{1,p}$.

Để thuận tiện, ta giữ nguyên các ký hiệu về cơ sở và ma trận biểu diễn các toán tử khi ta đưa về dạng tam giác trên trong quá trình như trên. Như vậy ta có điều cần chứng minh. $\square$

Ví dụ 10. Cho $A = \begin{pmatrix} 1 & -1 & 2 & 2 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \end{pmatrix}$. Xem A như ma trận của toán tử tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^4$ theo cơ sở chính tắc $\mathcal{B}_0 = (e_1, e_2, e_3, e_4)$.

Ta có $P_f(t) = t^2(t+1)(t-3)$ nên tồn tại cơ sở $\mathcal{B} = (u_1, u_2, u_3, u_4)$ của $\mathbb{R}^4$ sao cho $[f]_{\mathcal{B}}$ có dạng tam giác khối

$$[f]_{\mathcal{B}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

nghĩa là: $f(u_1) = -u_1,\ f(u_2) = 3u_2,\ f(u_3) = 0,\ f(u_4) = au_3.$

Ta có u_1, u_2, u_3 lần lượt là vectơ riêng của f ứng với các trị riêng $\lambda_1 = -1, \lambda_2 = 3, \lambda_3 = 0$. Có thể lấy $u_1 = (-2, 0, 1, 1)$, $u_2 = (2, 0, 1, 1)$, $u_3 = (1, 1, 0, 0)$.

Cuối cùng cho $a = 1$ ta tính được $u_4 = (0, 1, 1, 0)$.

Dễ dàng kiểm tra $\mathcal{B} = (u_1, u_2, u_3, u_4)$ là cơ sở của $\mathbb{R}^4$. Ma trận chuyển cơ sở từ $\mathcal{B}_0$ sang $\mathcal{B}$ là

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

Nhận xét rằng, nếu một ma trận tam giác hóa được thì ta có thể đưa nó về dạng tam giác khối để áp dụng vào giải một số bài toán ma trận. Chẳng hạn, việc nâng lên lũy thừa của một ma trận tam giác nói chung là không thể được. Tuy nhiên, nếu ta đưa nó về dạng tam giác khối thì việc nâng lên lũy thừa sẽ đưa về tính lũy thừa của các khối trên đường chéo chính. Hơn nữa, mỗi khối như vậy đều là ma trận tam giác mà tất cả các phần tử nằm trên đường chéo đều bằng nhau. Vậy bài toán thực chất đưa về việc nâng lên lũy thừa của ma trận tam giác dạng $A = \begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}$.

Phân tích $A = \lambda I + N$, trong đó $N = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$. Khi đó, đa thức tối tiểu của N có dạng $m_N(t) = t^p$, với $p \leq n$, nên $N^p = 0$. Hơn nữa, λI và N giao hoán với nhau nên ta dễ dàng tính được $A^k = (\lambda I + N)^k$ bằng cách áp dụng công thức khai triển nhị thức Newton.

§5. Dạng chính tắc Jordan

Định nghĩa 5.1. Giả sử $\lambda \in K$ và $n \geq 2$. Ta gọi một *khối Jordan* ứng với giá trị λ là một ma trận vuông cấp n có dạng:

$$J(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

Cụ thể hơn, $J(\lambda)$ là ma trận tam giác trên mà mọi phần tử thuộc đường chéo đều bằng λ; mọi phần tử bên trên đường chéo một mức đều bằng 1; tất cả các phần tử còn lại đều bằng 0.

Nếu $n = 1$ thì ta quy ước $J(\lambda) = (\ \lambda\ )$.

Ví dụ 11. Ta có các khối Jordan cấp $1, 2, 3$ ứng với giá trị λ là:

$$(\ \lambda\ ), \quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Bổ đề 5.2. *Giả sử $J = J(\lambda)$ là một khối Jordan cấp n. Khi đó:*

i) *Đa thức đặc trưng của J là: $P_J(t) = (-1)^n(t-\lambda)^n$.*

ii) *Đa thức tối tiểu của J là $m_J(t) = (t-\lambda)^n$.*

iii) *Chiều của không gian riêng $E(\lambda)$ là $\dim E(\lambda) = 1$.*

Chứng minh. Dễ dàng chứng minh bằng tính toán trực tiếp. □

Định nghĩa 5.3. Cho $T \in L_K(V)$. Ta nói T là *toán tử lũy linh* nếu tồn tại m nguyên dương sao cho $T^m = 0$. Số nguyên dương k nhỏ nhất thỏa mãn $T^k = 0$ được gọi là *bậc lũy linh* của T.

Mệnh đề 5.4. *Cho $T \in L_K(V)$. Toán tử T lũy linh nếu và chỉ nếu tồn tại cơ sở $\mathcal{B}$ của V sao cho $[T]_{\mathcal{B}}$ là ma trận tam giác trên ngặt (ma trận tam giác trên mà mọi phần tử thuộc đường chéo đều bằng 0).*

Chứng minh. Giả sử T là toán tử lũy linh bậc k. Khi đó đa thức tối tiểu của T có dạng $m_T(t) = t^k$, nên đa thức đặc trưng của T là $P_T(t) = (-1)^n t^n$. Do đó T tam giác hóa được, nghĩa là tồn tại một cơ sở $\mathcal{B}$ của V sao cho $[T]_{\mathcal{B}}$ là ma trận tam giác trên ngặt. Điều ngược lại là hiển nhiên. □

Định nghĩa 5.5. Cho $T \in L_K(V)$. Nếu tồn tại vectơ $u \in V$ sao cho tập hợp $\mathcal{B} = \{u, T(u), \ldots, T^{n-1}(u)\}$ tạo thành một cơ sở của V thì ta nói T là *toán tử xyclic* và $\mathcal{B}$ là *cơ sở xyclic* của V.

Bổ đề 5.6. *Cho T là toán tử lũy linh bậc $k > 1$. Khi đó, các điều sau tương đương:*

i) *$k = n$, nghĩa là $P_T(t) = (-1)^n t^n$ và $m_T(t) = t^n$.*

ii) *T là toán tử xyclic.*

iii) *Tồn tại cơ sở $\mathcal{B}$ của V sao cho $[T]_{\mathcal{B}} = J(0)$.*

Chứng minh. (i) $\Rightarrow$ (ii). Nếu $k = n$ thì $T^{n-1} \neq 0$, nên tồn tại $0 \neq u \in V$ sao cho $T^{n-1}(u) \neq 0$. Xét đẳng thức

$$a_1 u + a_2 T(u) + \cdots + a_n T^{n-1}(u) = 0, \text{ với } a_1, a_2, \ldots, a_n \in K.$$

Tác động T^{n-1} lên hai vế, nhận được $a_1 T^{n-1}(u) = 0$, nên $a_1 = 0$. Tiếp tục tác động T^{n-2} lên hai vế, nhận được $a_2 = 0$. Cứ tiếp tục như thế, cuối cùng ta được $a_1 = a_2 = \cdots = a_n = 0$, nên $\mathcal{B} = \{u, T(u), \ldots, T^{n-1}(u)\}$ độc lập tuyến tính. Do đó $\mathcal{B}$ là cơ sở xyclic của V, nghĩa là T là toán tử xyclic.

(ii) $\Rightarrow$ (iii). Giả sử $\mathcal{B} = \{u, T(u), \ldots, T^{n-1}(u)\}$ là cơ sở xyclic của V. Khi đó $[u]_{\mathcal{B}} = J(0)$.

(iii) $\Rightarrow$ (i). Áp dụng Bổ đề 5.2 ta được $k = n$. $\square$

Bổ đề 5.7. *Giả sử $T \in L_K(V)$ là toán tử lũy linh bậc $k > 1$. Khi đó:*

$$0 = \operatorname{Ker}(T^0) \subsetneq \operatorname{Ker}(T^1) \subsetneq \operatorname{Ker}(T^2) \subsetneq \ldots \subsetneq \operatorname{Ker}(T^k) = V.$$

Chứng minh. Đặt $K_p := \operatorname{Ker}(T^p)$. Nếu $u \in K_p$ thì $T^p(u) = 0$, nên $T^{p+1}(u) = 0$, nghĩa là $u \in K_{p+1}$. Do đó $K_p \subseteq K_{p+1}$. Mặt khác, nếu tồn tại $p \leq k-1$ sao cho $K_p = K_{p+1}$ thì $K_p = K_{p+1} = \ldots = K_k = V$. Do đó T có bậc lũy linh $\leq p$, mâu thuẫn. $\square$

Bổ đề 5.8. *Giả sử $T \in L_K(V)$ là toán tử lũy linh bậc $k > 1$. Khi đó, tồn tại các không gian con không tầm thường $W_1, \ldots, W_k$ của V sao cho*

i) $K_p = K_{p-1} \oplus W_p$, *với* $1 \leq p \leq k, K_p = \operatorname{Ker}(T^p)$.

ii) $T(W_p) \subseteq W_{p-1}$, *với* $2 \leq p \leq k$.

Chứng minh. Với $p = k$, chọn W_k là một phần bù nào đó của K_{k-1} trong $K_k = V$. Giả sử ta xây dựng được các không gian con $W_k, W_{k-1}, \ldots, W_p$ thỏa mãn (i) và (ii). Ta sẽ xây dựng không gian con W_{p-1} sao cho nó cũng thỏa mãn (i) và (ii). Trước hết, giả sử $u \in W_p$. Do $W_p \subseteq K_p$ nên $T^p(u) = 0$, nghĩa là $T^{p-1}(T(u)) = 0$. Do đó $T(u) \in K_{p-1}$. Vậy $T(W_p) \subseteq K_{p-1}$. Tiếp theo, giả sử $v \in T(W_p) \cap K_{p-2}$. Khi đó tồn tại $u \in W_p$ sao cho $v = T(u)$. Mà $T^{p-2}(v) = 0$, nên $T^{p-1}(u) = 0$, nghĩa là $u \in K_{p-1}$. Vậy, $u \in K_{p-1} \cap W_p = 0$, kéo theo $v = T(u) = 0$. Do đó $T(W_p) \cap K_{p-2} = 0$. Như vậy $K_{p-2} \oplus T(W_p) \subseteq K_{p-1}$. Do đó tồn tại không gian con G_{p-1} sao cho

$$K_{p-1} = K_{p-2} \oplus T(W_p) \oplus G_{p-1}.$$

Đặt $W_{p-1} := T(W_p) \oplus G_{p-1}$. Dễ thấy W_{p-1} thỏa mãn (i) và (ii). $\square$

Hệ quả 5.9. *Với các ký hiệu $W_1, W_2, \ldots, W_k$ như trên, ta có*

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

Chứng minh. Ta có

$$V = K_k = K_{k-1} \oplus W_k = K_{k-2} \oplus W_{k-1} \oplus W_k = \cdots = W_1 \oplus \cdots \oplus W_k.$$

$\square$

Tiếp theo, ta sẽ xây dựng một cơ sở của V bằng cách xây dựng cơ sở cho mỗi không gian con W_p, sau đó sắp xếp lại các vectơ cơ sở bằng cách phân hoạch nó thành các cơ sở xyclic của các không gian con của V. Theo cơ sở cuối cùng này ma trận biểu diễn f sẽ có dạng khối đường chéo mà trên đường chéo chính là các khối Jordan.

Bổ đề 5.10. *Với các ký hiệu như trên, nếu $\mathcal{B} = (u_1, \ldots, u_r)$ là một cơ sở của W_p thì $T(\mathcal{B}) := (T(u_1), \ldots, T(u_r))$ là một họ độc lập tuyến tính trong W_{p-1}.*

Chứng minh. Giả sử $\lambda_1, \ldots, \lambda_r \in K$ sao cho

$$\lambda_1 T(u_1) + \cdots + \lambda_r T(u_r) = 0.$$

Khi đó, $T(\lambda_1 u_1 + \cdots + \lambda_r u_r) = 0$, nên

$$\lambda_1 u_1 + \cdots + \lambda_r u_r \in \mathrm{Ker} T = K_1 \subseteq K_{p-1}.$$

Do đó

$$\lambda_1 u_1 + \cdots + \lambda_r u_r \in W_p \cap K_{p-1} = 0,$$

kéo theo $\lambda_1 = \cdots = \lambda_r = 0$. $\square$

Bây giờ ta sẽ lần lượt xây dựng cơ sở cho W_i, bắt đầu từ $i = k$. Trong W_k (mà ta còn ký hiệu là G_k), lấy một cơ sở nào đó. Tiếp theo, bằng qui nạp lùi, ta xây dựng cơ sở trong W_{p-1} như sau: Ta có $W_{p-1} = T(W_p) \oplus G_{p-1}$. Lấy ảnh qua T của cơ sở đã xây dựng trong W_p hợp với một cơ sở của G_{p-1} ta được một cơ sở của W_{p-1}. Bằng cách này ta xây dựng được các cơ sở $\mathcal{B}_k$, $\mathcal{B}_{k-1}$, ..., $\mathcal{B}_1$ tương ứng của $W_k, W_{k-1} \dots, W_1$ như sau:

$\mathcal{B}_k = (u_1, \dots, u_{n_k})$;

$\mathcal{B}_{k-1} = (T(u_1), \dots, T(u_{n_k}), \omega_1, \dots, \omega_{n_{k-1}})$;

$\mathcal{B}_{k-2} = (T^2(u_1), \dots, T^2(u_{n_k}), T(\omega_1), \dots, T(\omega_{n_{k-1}}), z_1, \dots, z_{n_{k-2}})$

....

Trong đó $(u_1, \dots, u_{n_k})$ là cơ sở của G_k; $(\omega_1, \dots, \omega_{n_{k-1}})$ là cơ sở của G_{k-1}; $(z_1, \dots, z_{n_{k-2}})$ là cơ sở của G_{k-2}; ...

Bổ đề 5.11. *Cho* $0 \neq u \in G_p$. *Đặt* $H_p(u) := \langle u, T(u), \dots, T^{p-1}(u)\rangle$. *Khi đó*

i) $\dim H_p(u) = p$.

ii) $H_p(u)$ *là không gian con bất biến đối với* T.

iii) $T|_{H_p(u)}$ *là toán tử xyclic.*

Chứng minh. Dễ dàng chứng minh. □

Bây giờ ta đã có đủ điều kiện để nhận được định lý chính dưới đây.

Định lý 5.12. *Cho* $u \in L_K(V)$ *là một toán tử lũy linh. Khi đó* V *phân tích được thành tổng trực tiếp các không gian con bất biến đối với* T *sao cho hạn chế của* T *lên mỗi không gian con này là một toán tử xyclic.*

Chứng minh. Qua cách xây dựng các không gian con $H_p(u)$ ta nhận thấy rằng V là tổng trực tiếp của các không gian con

$$H_k(u_1), \ldots, H_k(u_{n_k}), H_{k-1}(\omega_1), \ldots, H_{k-1}(\omega_{n_{k-1}}), \ldots$$

Hơn nữa, mọi không gian con nói trên đều bất biến đối với T và hạn chế của T lên mỗi không gian con đó là một toán tử xyclic. $\square$

Dạng tam giác khối nói chung đã có thể dùng khá tốt cho những ứng dụng. Tuy nhiên, ta có thể tiến hành rút gọn trong từng khối cho đến khi nhận được một dạng mà theo một nghĩa nào đó là *đơn giản nhất*. Đó chính là *dạng chính tắc Jordan* mà ta sẽ đề cập tới trong phần ngay sau đây.

Định nghĩa 5.13. Ma trận khối đường chéo dạng $J := \operatorname{diag}(J_1, \ldots, J_k)$ mà mỗi khối J_i đều là khối Jordan được gọi là ma trận *dạng Jordan*.

Ví dụ 12. Các ma trận sau là ma trận dạng Jordan

$$\left(\begin{array}{cc|cc|c|c} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 2 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right); \left(\begin{array}{ccc|c} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right); \left(\begin{array}{cc|cc} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right).$$

Nhận xét 5.14. Mọi ma trận đường chéo cấp n đều là ma trận dạng Jordan, với n khối Jordan là n ma trận vuông cấp 1.

Định lý sau chứng tỏ rằng mọi ma trận vuông đều đồng dạng với ma trận dạng Jordan.

Định lý 5.15 (Jordan). *Cho toán tử tuyến tính $f \in L_K(V)$ có p trị riêng phân biệt λ_1, ..., λ_p. Giả sử đa thức đặc trưng của f có dạng $P_f(t) = (-1)^n(t-\lambda_1)^{n_1} \ldots (t-\lambda_p)^{n_p}$ và đa thức tối tiểu của f có dạng*

$m_f(t) = (t-\lambda_1)^{k_1}\dots(t-\lambda_p)^{k_p}$. Khi đó tồn tại cơ sở $\mathcal{B}$ của V sao cho

$$[f]_{\mathcal{B}} = \begin{pmatrix} \boxed{\widetilde{J}(\lambda_1)} & & & 0 \\ & \boxed{\widetilde{J}(\lambda_2)} & & \\ & & \ddots & \\ 0 & & & \boxed{\widetilde{J}(\lambda_p)} \end{pmatrix}.$$

Trong đó, với mỗi i, $\widetilde{J}(\lambda_i)$ là khối đường chéo cấp n_i, gồm $\gamma_i = \dim E(\lambda_i)$ khối Jordan dạng $J(\lambda_i)$, khối lớn nhất có cấp bằng k_i.

Chứng minh. 1) Nếu $p=1$, nghĩa là $P_f(t) = (-1)^n(t-\lambda)^n$. Do $P_f(t)$ phân rã trên K nên f tam giác hóa được. Do đó tồn tại cơ sở $\mathcal{B}'$ của V sao cho

$$[f]_{\mathcal{B}'} = \begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}.$$

Đặt $T := f - \lambda \mathrm{Id}_V$, suy ra

$$[T]_{\mathcal{B}'} = [f]_{\mathcal{B}'} - \lambda I_n = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

Do đó, theo Mệnh đề 5.4, T là toán tử lũy linh. Gọi k là bậc lũy linh của T. Do ma trận biểu diễn toán tử $\lambda\mathrm{Id}_V$ theo mọi cơ sở của V đều là λI_n nên từ Định lý 5.12, tồn tại một cơ sở của V sao cho trong đó ma trận biểu diễn f có dạng khối đường chéo mà trên đường chéo chính là các khối Jordan dạng $J(\lambda)$. Theo cách xây dựng các không gian con $H_p(u)$ trong chứng minh Định lý 5.12 thì không gian con có số chiều lớn nhất chính là $H_k(u)$. Do đó cấp của khối Jordan lớn nhất là k. Vì mỗi khối Jordan chỉ chứa đúng một vectơ riêng nên số các khối Jordan đúng bằng số chiều của không gian riêng $E(\lambda)$.

2) Trong trường hợp tổng quát, ta phân tích V thành tổng trực tiếp của các không gian đặc trưng, sau đó áp dụng trường hợp 1) đối với mỗi không gian đặc trưng. Như vậy định lý đã được chứng minh hoàn toàn. $\square$

Đối với toán tử tuyến tính $f \in L_K(V)$, việc tìm cơ sở và ma trận biểu diễn f theo cơ sở ấy như đã làm trong Định lý Jordan được gọi là việc đưa f về *dạng chính tắc Jordan*. Rõ ràng bằng ngôn ngữ ma trận ta cũng có thể nói tới việc đưa ma trận về dạng chính tắc Jordan. Dựa vào Định lý trên, ta có phương pháp đưa một ma trận A (thực hiện tương tự cho một toán tử f) về dạng chính tắc Jordan như sau:

1. Tìm và phân tích đa thức đặc trưng của A dưới dạng

$$P_A(t) = (-1)^n (t - \lambda_1)^{n_1} \ldots (t - \lambda_p)^{n_p},\ \lambda_i \neq \lambda_j \text{ với mọi } i \neq j.$$

2. Tìm và phân tích đa thức tối tiểu của A dưới dạng

$$m_A(t) = (t - \lambda_1)^{k_1} \ldots (t - \lambda_p)^{k_p}.$$

3. Với mỗi i, đặt $N_i := A - \lambda_i I_n$ và ký hiệu $E(\lambda_i, r)$ là không gian nghiệm của hệ phương trình $N_i^r X = 0$.

- Tính $d_{i1} := \dim E(\lambda_i, 1)$; $d_{i2} := \dim E(\lambda_i, 2) - d_{i1}; \ldots$
- Xây dựng tập hợp $\mathcal{B}_i$ là hợp của d_{i1} tập con như sau:

+ Tìm d_{i1} vectơ độc lập tuyến tính thuộc $E(\lambda_i, 1)$ và phân bổ vào mỗi tập con của $\mathcal{B}_i$ một vectơ;
+ Tìm d_{i2} vectơ độc lập tuyến tính thuộc $E(\lambda_i, 2)$ nhưng không thuộc $E(\lambda_i, 1)$ và phân bổ vào d_{i2} tập con đầu tiên của $\mathcal{B}_i$, mỗi tập con của $\mathcal{B}_i$ một vectơ.
+ Cứ tiếp tục quá trình trên đến khi $\mathcal{B}_i$ có đủ n_i vectơ.

- Đặt $\widetilde{J}(\lambda_i)$ là ma trận khối đường chéo cấp n_i gồm d_{i1} khối Jordan dạng $J(\lambda_i)$, mỗi khối có cấp tương ứng bằng số vectơ được phân bổ cho khối từ $\mathcal{B}_i$.

4. Đặt $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_p$; $P = (\mathcal{B}_0 \to \mathcal{B})$ ta được dạng chính tắc Jordan của A là $P^{-1}AP = \operatorname{diag}(\widetilde{J}(\lambda_1), \ldots, \widetilde{J}(\lambda_p))$.

Ví dụ 13. Tìm dạng chính tắc Jordan của $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 1 \\ -2 & -4 & -4 & -1 \end{pmatrix}$.

Giải. Xem A như ma trận biểu diễn toán tử tuyến tính $f \in L_{\mathbb{R}}(\mathbb{R}^4)$.

- Đa thức đặc trưng của A là $P_A(t) = (t-1)^4$.

- Đặt $N := A - I_4$. Do $N \neq 0$ và $N^2 = 0$ nên đa thức tối tiểu của A là $m_A(t) = (t-1)^2$.

- Ta có $N = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 2 & 1 \\ -2 & -4 & -4 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$,

  nên hệ $NX = 0$ có vô số nghiệm, xác định bởi

  $$(x_1, x_2, x_3, x_4) = (-2a - 2b - c; a; b; c), a, b, c, d \in \mathbb{R}.$$

  Nghiệm căn bản của hệ là $u_1 = (-2; 1; 0; 0)$, $u_2 = (-2; 0; 1; 0)$, $u_3 = (-1; 0; 0; 1)$. Do đó $d_{11} = \dim E(1) = 3$ và cơ sở của $E(1)$ là $\{u_1, u_2, u_3\}$.

- Phân bổ u_1, u_2, u_3 vào ba tập hợp $\{u_1\}, \{u_2\}, \{u_3\}$.

- Do $N^2 = 0$ nên $d_{12} = \dim E(1,2) - d_{11} = 4 - 3 = 1$. Hơn nữa, do $N^2 = 0$ nên $E(1,2) = \mathbb{R}^4$, nên ta chỉ cần chọn vectơ v sao cho

$\{u_1, u_2, u_3, v\}$ độc lập tuyến tính, chẳng hạn, chọn $v = (1, 0, 0, 0)$ và phân bổ v vào tập hợp thứ nhất, thành $\{u_1, v\}$.

- Từ các tính toán ở trên ta được:

- Dạng chính tắc Jordan của A gồm 3 khối Jordan, trong đó khối lớn nhất có cấp 2.

$$J = \left(\begin{array}{cc|c|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right).$$

- Ma trận P thỏa mãn $P^{-1}AP = J$ là

$$P = (u_1^\top \ v^\top \ u_2^\top \ u_3^\top) = \begin{pmatrix} -2 & 1 & -2 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$\square$

Ví dụ 14. Tìm dạng chính tắc Jordan của $A = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

Giải. Xem A là ma trận biểu diễn toán tử tuyến tính $f \in L_{\mathbb{R}}(\mathbb{R}^4)$.

- Đa thức đặc trưng của A là $P_A(t) = (t-1)(t-2)^3$.
- Đặt $N_1 = A - I_4$, $N_2 = (A - 2I_4)$. Ta có $N_1 N_2^2 \neq 0$ nên đa thức tối tiểu của A là $m_A(t) = (t-1)(t-2)^2$.
- Ta có $\dim E(1) = 1$ nên $\widetilde{J}(1) = 1$.

- Ta có $\dim E(2) = 2$, nên $\widetilde{J}(2)$ có hai khối Jordan, khối thứ nhất có cấp $k_2 = 2$, khối thứ hai có cấp $n_2 - k_2 = 1$, nên

$$\widetilde{J}(2) = \begin{pmatrix} \boxed{\begin{matrix} 2 & 1 \\ 0 & 2 \end{matrix}} & \\ & \boxed{2} \end{pmatrix}.$$

- Vậy dạng chính tắc Jordan của A là

$$J = \left(\begin{array}{c|cc|c} 1 & 0 & 0 & 0 \\ \hline 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right).$$

- Cơ sở cần tìm là $\mathcal{B} = (u_1, u_2, v, u_3)$, trong đó $\{u_1\}$ là cơ sở của $E(1)$, $\{u_2, u_3\}$ là cơ sở của $E(2)$, v là nghiệm của hệ $(A - 2I_4)X = u_2^\top$.

□

Ví dụ 15. Tìm dạng chính tắc Jordan của $A = \begin{pmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{pmatrix}$.

Giải. Đa thức đặc trưng của A là $P_A(t) = (t-2)^4$. Đặt $N = (A - 2I_4)$. Ta có $N^2 \neq 0$ và $N^3 = 0$ nên đa thức tối tiểu của A là $m_A(t) = (t-2)^2$. Mặt khác, do A là ma trận cấp 4 nên trong dạng chính tắc Jordan của gồm hai khối dạng $J_1 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ hoặc một khối dạng J_1 và hai khối dạng $J_2 = (\,2\,)$. Nhưng $r(N) = 2$ nên $\dim E(2) = 4 - 2 = 2$ nên trường hợp thứ hai không xảy ra. Vậy dạng chính tắc Jordan của A là

$$J = \left(\begin{array}{cc|cc} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right).$$

$\square$

Bài tập

Bài 1. Tìm đa thức tối tiểu của ma trận:

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

Bài 2. Giả sử toán tử tuyến tính $f : \mathbb{R}^3 \to \mathbb{R}^3$ có ma trận biểu diễn theo cơ sở chính tắc là

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

Hãy tìm đa thức tối tiểu của f và phân tích $\mathbb{R}^3$ thành tổng trực tiếp của các không gian đặc trưng.

Bài 3. Tìm đa thức tối tiểu của $A = \begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}$. Ma trận A có chéo hóa được trên trường số thực $\mathbb{R}$ hay không?

Bài 4. Tìm một ma trận $A \in M_3(\mathbb{R})$ sao cho đa thức tối tiểu của A là $m_A(t) = t^2$.

Bài 5. Cho V là không gian vectơ n chiều trên $\mathbb{K}$ và f là toán tử tuyến tính trên V. Giả sử tồn tại $k \in \mathbb{N}$ sao cho $f^k = 0$. Chứng minh$f^n = 0$.

Bài 6. Tìm ma trận B sao cho $B^2 = A$.

a) $A = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 3 & 0 \\ -8 & 2 & 4 \end{pmatrix}$.

c) $A = \begin{pmatrix} 6 & -6 & 5 \\ -4 & -1 & 10 \\ 7 & -6 & 4 \end{pmatrix}$.

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

Bài 7. Cho ma trận $A = \begin{pmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix}$.

a) Chứng minh A không chéo hóa được.

b) Hãy tam giác hóa ma trận A.

Bài 8. Tìm dạng tam giác của A và chỉ rõ ma trận khả nghịch P sao cho $P^{-1}AP$ có dạng tam giác:

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

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

Bài 9. Tìm dạng chính tắc Jordan của các ma trận

a) $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 1 \\ -2 & -5 & -4 & -1 \end{pmatrix}$;

c) $A = \begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}$;

b) $A = \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$;

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

Bài 10. Giả sử toán tử $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ có ma trận biểu diễn theo cơ sở chính tắc là

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

a) Tính đa thức tối tiểu của f. Từ đó rút ra kết luận gì về tính chéo hóa của toán tử f?

b) Tìm cơ sở $\mathcal{B}$ của $\mathbb{R}^3$ sao cho $[f]_{\mathcal{B}}$ có dạng chính tắc Jordan. Từ đó hãy chỉ ra một cơ sở cho mỗi không gian đặc trưng của f.

Bài 11. Cho $A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$. Hãy tính A^n với n nguyên dương.

Bài 12. Cho $A = \begin{pmatrix} -1 & 1 & 1 & 0 \\ -1 & 2 & 1 & -1 \\ 5 & -3 & -2 & 5 \\ 4 & -2 & -2 & 3 \end{pmatrix}$.

a) Tìm một ma trận Jordan A' đồng dạng với A và chỉ rõ ma trận khả nghịch P thỏa mãn $A' = P^{-1}AP$.

b) Tính A^n, với n nguyên dương.

Chương 7

KHÔNG GIAN EUCLID

Trong chương này ngoại trừ những trường hợp riêng sẽ được nói rõ, ta chỉ xét các không gian vectơ trên trường số thực $\mathbb{R}$. Trong các chương trước ta đã khảo sát các không gian vectơ tổng quát. Tuy nhiên, khái niệm không gian vectơ chưa mở rộng một cách đầy đủ các không gian 2 hoặc 3 chiều của hình học giải tích. Chẳng hạn, cho đến nay ta vẫn chưa đề cập đến tích vô hướng, độ dài vectơ hay góc giữa hai vectơ,... và vì vậy chúng ta chưa phát triển được lý thuyết hình học metric phong phú đã biết trong trường hợp 2 hoặc 3 chiều. Trong chương này chúng ta sẽ bổ sung cho những khiếm khuyết đó.

§1. Tích vô hướng và không gian Euclid

Định nghĩa 1.1. Cho V là không gian vectơ trên $\mathbb{R}$. Ánh xạ

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R};\ (u, v) \mapsto \langle u, v\rangle$$

được gọi là một *tích vô hướng* trong V nếu với mọi $u, v, w \in V$ và với mọi $\alpha, \beta \in \mathbb{R}$, ta có

i) $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$;

ii) $\langle u, \alpha v + \beta w\rangle = \alpha\langle u, v\rangle + \beta\langle u, w\rangle$;

iii) $\langle u, v\rangle = \langle v, u\rangle$;

iv) $\langle u, u\rangle \geq 0$; $\langle u, u\rangle = 0$ khi và chỉ khi $u = 0$.

Một không gian vectơ hữu hạn chiều với tích vô hướng được gọi là một *không gian Euclid*.

Ví dụ 16. Tập hợp tất cả các vectơ trong không gian $\mathbb{R}^2$ và trong không gian $\mathbb{R}^3$ với tích vô hướng quen thuộc đã được định nghĩa trong các sách giáo khoa về toán sơ cấp là không gian Euclid.

Ví dụ 17. Không gian vectơ $\mathbb{R}^n$ là không gian Euclid với tích vô hướng của vectơ $u = (x_1, \ldots, x_n)$ và vectơ $v = (y_1, \ldots, y_n)$ được xác định bởi:

$$\langle u, v\rangle := x_1y_1 + \cdots + x_ny_n.$$

Tích vô hướng này được gọi là *tích vô hướng chính tắc* trong $\mathbb{R}^n$.

Ví dụ 18. Với $u = (x_1, x_2, x_3), v = (y_1, y_2, y_3) \in \mathbb{R}^3$ ta định nghĩa

$$\langle u, v\rangle := x_1y_1 + 2x_2y_2 + 3x_3y_3 + x_1y_2 + x_2y_1.$$

Dễ thấy các tính chất (i)-(iii) trong Định nghĩa 1.1 được thỏa mãn. Hơn nữa

$$\begin{aligned}\langle u, u\rangle &= x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2\\ &= x_1^2 + 2x_1x_2 + x_2^2 + x_2^2 + 3x_3^2\\ &= (x_1 + x_2)^2 + x_2^2 + 3x_3^2 \geq 0;\\ \langle u, u\rangle = 0 &\Leftrightarrow x_1 + x_2 = x_2 = x_3 = 0\\ &\Leftrightarrow x_1 = x_2 = x_3 = 0.\end{aligned}$$

Do đó (iv) cũng thỏa mãn. Do đó phép toán trên xác định một tích vô hướng trên $\mathbb{R}^3$ (khác tích vô hướng chính tắc).

Ví dụ 19. Ánh xạ $\langle A, B\rangle := tr(A^\top B)$ xác định một tích vô hướng trong không gian $M_2(\mathbb{R})$ các ma trận vuông cấp 2 trên $\mathbb{R}$.

Ví dụ 20. Nếu V là một không gian vectơ với tích vô hướng $\langle\ ,\ \rangle_V$ và W là không gian con của V thì ánh xạ $\langle\ ,\ \rangle_W$ xác định bởi

$$\langle u, v\rangle_W := \langle u, v\rangle_V, \forall u, v \in W$$

là một tích vô hướng trong W.

Định nghĩa 1.2. Cho V là không gian Euclid V với tích vô hướng $\langle\ ,\ \rangle$ và $u \in V$. Khi đó:

- Ký hiệu $||u|| := \sqrt{\langle u, u\rangle}$ được gọi là *chuẩn* (hay *độ dài*) của u.
- Nếu $||u|| = 1$ thì ta nói u là một *vectơ đơn vị*.

Nhận xét 1.3. Cho $u \in V$ và $\lambda \in \mathbb{R}$. Khi đó:

i) $||u|| \geq 0$; $||u|| = 0 \Leftrightarrow u = 0$.

ii) $||\lambda u|| = |\lambda|.||u||$.

Ví dụ 21. a) Trong không gian Euclid ở Ví dụ 16, độ dài của một vectơ chính là độ dài quen thuộc mà ta đã biết trong Hình học sơ cấp.

b) Độ dài của vectơ $u = (x_1, \ldots, x_n)$ trong không gian ở Ví dụ 17 được xác định như sau:

$$||u|| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

c) Độ dài của vectơ $u = (x_1, x_2, x_3)$ trong không gian ở Ví dụ 18 được xác định như sau:

$$||u|| = \sqrt{(x_1 + x_2)^2 + x_2^2 + 3x_3^2}.$$

Mệnh đề 1.4. *Cho V là không gian Euclid V với tích vô hướng $\langle\ ,\ \rangle$. Khi đó, với mọi $u, v \in V$ ta có*

i) $\langle u, v\rangle^2 \leq ||u||^2\, ||v||^2$ (Bất đẳng thức Cauchy-Schwarz).

Dấu '=' xảy ra khi và chỉ khi u và v phụ thuộc tuyến tính.

ii) $||u + v|| \leq ||u|| + ||v||$ (Bất đẳng thức tam giác).

Dấu '=' xảy ra khi và chỉ khi tồn tại $\lambda \geq 0$ sao cho $v = \lambda u$ hoặc $u = \lambda v$.

Chứng minh. i) Nếu $||u|| = 0$ thì bất đẳng thức hiển nhiên được thỏa mãn. Giả sử $||u|| \neq 0$. Khi đó, với mọi $\lambda \in \mathbb{R}$ ta có $||\lambda u + v||^2 \geq 0$, nghĩa là

$$\lambda^2 ||u||^2 + 2\lambda\langle u, v\rangle + ||v||^2 \geq 0.$$

Vế trái của bất đẳng thức là một tam thức bậc hai theo λ, nên tam thức này nhận giá trị không âm với mọi λ khi và chỉ khi biệt số $\Delta \leq 0$, nghĩa là

$$\langle u, v\rangle^2 \leq ||u||^2\,||v||^2.$$

Dấu '=' xảy ra khi và chỉ khi $||u + \lambda v||^2 = 0$, hay $u + \lambda v = 0$, nghĩa là u và v phụ thuộc tuyến tính.

ii) Ta có $||u+v||^2 = ||u||^2 + ||v||^2 + 2\langle u, v\rangle \leq ||u||^2 + ||v||^2 + 2||u||\,||v|| = (||u|| + ||v||)^2$. Do đó $||u + v|| \leq ||u|| + ||v||$. Nếu $v = \lambda u$ hoặc $u = \lambda v$ thì hiển hiên dấu '=' xảy ra. Ngược lại, nếu dấu '=' xảy ra thì

$$\langle u, v\rangle = |\langle u, v\rangle| = ||u||\,||v||.$$

Đẳng thức thứ hai suy ra u và v phụ thuộc tuyến tính, nghĩa là tồn tại $\lambda \geq 0$ sao cho $v = \lambda u$ hoặc $u = \lambda v$. Đẳng thức thứ nhất suy ra $\lambda \geq 0$. □

Nhận xét 1.5. a) Nếu $0 \neq u, v \in V$ thì $\dfrac{|\langle u, v\rangle|}{||u||\,||v||} \leq 1$, nên tồn tại $\theta \in [0, \pi]$ sao cho $\cos\theta = \dfrac{\langle u, v\rangle}{||u||\,||v||}$. Ta gọi θ là *góc* (không định hướng) giữa u và v.

b) Tích vô hướng có thể được biểu diễn thông qua chuẩn như sau:

$$\langle u, v\rangle = \frac{||u+v||^2 - ||u||^2 - ||v||^2}{2}.$$

§2. Cơ sở trực giao và cơ sở trực chuẩn

Định nghĩa 2.1. Cho V là không gian Euclid với tích vô hướng $\langle \, , \, \rangle$ và $u, v \in V$, $S = \{u_1, \dots, u_m\} \subseteq V$. Khi đó:

i) Nếu $\langle u, v\rangle = 0$ thì ta nói u *trực giao* với v, ký hiệu $u \perp v$

ii) Nếu $\langle u_i, u_j\rangle = 0$ với mọi $i \neq j$ thì ta nói S là một *tập trực giao*

iii) Tập trực giao chỉ gồm các vectơ đơn vị được gọi là *tập trực chuẩn*.

iv) Nếu cơ sở $\mathcal{B}$ của V là một tập trực giao (tập trực chuẩn) thì $\mathcal{B}$ được gọi là *cơ sở trực giao* (*cơ sở trực chuẩn*) của V.

Mệnh đề 2.2. *i) Mọi hệ trực giao không chứa vectơ* 0 *đều độc lập tuyến tính.*

ii) Nếu $(u_1, \dots, u_m)$ *là một hệ trực giao không chứa vectơ* 0 *thì*

$$\left(\frac{u_1}{||u_1||}, \dots, \frac{u_m}{||u_m||}\right)$$

là một hệ trực chuẩn.

Chứng minh. i) Giả sử $S = (u_1, \dots, u_m)$ là một hệ trực giao không chứa vectơ 0. Xét đẳng thức

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0.$$

Với mỗi $i, 1 \leq i \leq m$, nhân hai vế cho u_i và sử dụng tính chất $\langle u_i, u_i\rangle \neq 0$ và $\langle u_i, u_j\rangle = 0$ với mọi $j \neq i$ ta được $\alpha_i = 0$, nghĩa là S độc lập tuyến tính.

ii) Dễ dàng chứng minh. $\square$

Giả sử V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và $\mathcal{B} = (e_1, \ldots, e_n)$ là một cơ sở của V. Khi đó, ma trận $A = (a_{ij}) \in M_n(\mathbb{R})$ xác định bởi $a_{ij} = \langle e_i, e_j \rangle$, được gọi là ma trận biểu diễn tích vô hướng $\langle\ ,\ \rangle$ theo cơ sở $\mathcal{B}$, ký hiệu $A = \langle\ ,\ \rangle_{\mathcal{B}}$.

Ví dụ 22. Với $u = (x_1, x_2, x_3), v = (y_1, y_2, y_3) \in \mathbb{R}^3$, ta đã định nghĩa tích vô hướng (xem Ví dụ 18).

$$\langle u, v \rangle := x_1y_1 + 2x_2y_2 + 3x_3y_3 + x_1y_2 + x_2y_1.$$

Ma trận biểu diễn tích vô hướng này theo cơ sở chính tắc $\mathcal{B}_0$ của $\mathbb{R}^3$ là

$$\langle\ ,\ \rangle_{\mathcal{B}_0} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Mệnh đề 2.3. *Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và $\mathcal{B}$ là một cơ sở của V. Khi đó, với mọi $u, v \in V$ ta có*

$$\langle u, v \rangle = [u]_{\mathcal{B}}^{\top} \langle\ ,\ \rangle_{\mathcal{B}} [v]_{\mathcal{B}}.$$

Chứng minh. Dễ dàng kiểm tra trực tiếp. □

Hệ quả 2.4. *Trong không gian Euclid V với tích vô hướng $\langle\ ,\ \rangle$, cơ sở $\mathcal{B} = (e_1, \ldots, e_n)$ là cơ sở trực chuẩn nếu và chỉ nếu $\langle\ ,\ \rangle_{\mathcal{B}} = I_n$.*

Trong Mệnh đề 2.3, biểu thức của tích vô hướng được viết thông qua việc chọn cơ sở $\mathcal{B}$. Khi đó công thức tính $\langle u, v \rangle$ là

$$\langle u, v \rangle = [u]_{\mathcal{B}}^{\top} \langle\ ,\ \rangle_{\mathcal{B}} [v]_{\mathcal{B}}.$$

Nếu ta chọn một cơ sở $\mathcal{B}'$ khác thì $\langle\ ,\ \rangle_{\mathcal{B}'}$ tất nhiên sẽ thay đổi. Tuy nhiên các ma trận $\langle\ ,\ \rangle_{\mathcal{B}}$ và $\langle\ ,\ \rangle_{\mathcal{B}'}$ có mối liên hệ mật thiết với nhau thông qua Mệnh đề sau:

Mệnh đề 2.5. *Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và $\mathcal{B}, \mathcal{B}'$ là hai cơ sở của V. Khi đó*

$$\langle\ ,\ \rangle_{\mathcal{B}'} = (\mathcal{B} \to \mathcal{B}')^\top \langle\ ,\ \rangle_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}'),$$

trong đó $(\mathcal{B} \to \mathcal{B}')$ là ma trận chuyển cơ sở từ $\mathcal{B}$ sang $\mathcal{B}'$.

Chứng minh. Với mọi $u, v \in V$, ta có

$$[u]_{\mathcal{B}} = (\mathcal{B} \to \mathcal{B}')[u]_{\mathcal{B}'} \text{ và } [v]_{\mathcal{B}} = (\mathcal{B} \to \mathcal{B}')[v]_{\mathcal{B}'}.$$

Hơn nữa theo Mệnh đề 2.3 ta có

$$\langle u, v\rangle = [u]_{\mathcal{B}}^\top \langle\ ,\ \rangle_{\mathcal{B}} [v]_{\mathcal{B}}.$$

Do đó

$$\begin{aligned}\langle u, v\rangle &= ((\mathcal{B} \to \mathcal{B}')[u]_{\mathcal{B}'})^\top \langle\ ,\ \rangle_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}')[v]_{\mathcal{B}'}\\ &= [u]_{\mathcal{B}'}^\top (\mathcal{B} \to \mathcal{B}')^\top \langle\ ,\ \rangle_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}') [v]_{\mathcal{B}'}.\end{aligned}$$

Mặt khác

$$\langle u, v\rangle = [u]_{\mathcal{B}'}^\top \langle\ ,\ \rangle_{\mathcal{B}'} [v]_{\mathcal{B}'}$$

Suy ra

$$[u]_{\mathcal{B}'}^\top \langle\ ,\ \rangle_{\mathcal{B}'} [v]_{\mathcal{B}'} = [u]_{\mathcal{B}'}^\top (\mathcal{B} \to \mathcal{B}')^\top \langle\ ,\ \rangle_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}') [v]_{\mathcal{B}'}.$$

Bằng cách chọn u và v lần lượt là các vectơ của cơ sở $\mathcal{B}'$, ta suy ra được

$$\langle\ ,\ \rangle_{\mathcal{B}'} = (\mathcal{B} \to \mathcal{B}')^\top \langle\ ,\ \rangle_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}').$$

□

Định lý 2.6 (Gram-Schmidt)**.** *Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$. Khi đó, mọi không gian con của V đều có cơ sở trực chuẩn.*

Chứng minh. Giả sử W là một không gian con của V và $S = (v_1, \ldots, v_p)$ là một cơ sở của W. Do Mệnh đề 2.2, ta chỉ cần xây dựng một cơ sở trực giao $(u_1, \ldots, u_p)$ cho W là đủ. Việc chọn các vectơ $u_1, \ldots, u_p$ được thực hiện thông qua thuật toán sau, gọi là ***quá trình trực giao hóa Gram-Schmidt***.

Bước 1. Đặt $u_1 := v_1$;

Bước 2: Tìm λ_1 sao cho $u_2 := v_2 + \lambda_1 u_1$ trực giao với u_1, nghĩa là

$$0 = \langle u_2, u_1 \rangle = \langle v_2 + u_1, u_1 \rangle = \langle v_2, u_1 \rangle + \lambda_1 \langle u_1, u_1 \rangle.$$

Do đó ta có thể chọn $\lambda_1 = -\dfrac{\langle v_2, u_1 \rangle}{||u_1||^2}$.

Bước 3: Tìm λ_1, λ_2 sao cho $u_3 := v_3 + \lambda_1 u_1 + \lambda_2 u_2$ trực giao với u_1 và u_2, nghĩa là

$$\begin{cases} 0 = \langle u_3, u_1 \rangle = \langle v_3 + \lambda_1 u_1 + \lambda_2 u_2, u_1 \rangle = \langle v_3, u_1 \rangle + \lambda_1 ||u_1||^2 \\ 0 = \langle u_3, u_2 \rangle = \langle v_3 + \lambda_1 u_1 + \lambda_2 u_2, u_2 \rangle = \langle v_3, u_2 \rangle + \lambda_1 ||u_2||^2 \end{cases}$$

Do đó ta có thể chọn

$$\lambda_1 = -\frac{\langle v_3, u_1 \rangle}{||u_1||^2}, \quad \lambda_2 = -\frac{\langle v_3, u_2 \rangle}{||u_2||^2}.$$

Cứ tiếp tục quá trình trên, ta tìm được vectơ

$$u_p := v_p + \lambda_1 u_1 + \cdots + \lambda_{p-1} u_{p-1}, \quad \text{với } \lambda_i = -\frac{\langle v_p, u_i \rangle}{||u_i||^2}.$$

Như vậy ta đã xây dựng được một họ các vectơ trực giao $(u_1, \ldots, u_p)$ Bây giờ ta chỉ cần chứng minh

$$\langle u_1, \ldots, u_p \rangle = \langle v_1, \ldots, v_p \rangle.$$

Ta có $\langle u_1 \rangle = \langle v_1 \rangle$. Giả sử $1 < i \leq p - 1$ và $\langle u_1, \ldots, u_i \rangle = \langle v_1, \ldots, v_i \rangle$. Khi đó mỗi một vectơ $u_k (1 \leq k \leq i)$ đều là tổ hợp tuyến tính của các vectơ $v_1, \dots, v_i$. Theo cách xây dựng thì u_{i+1} là tổ hợp tuyến tính của các vectơ $v_{i+1}, u_1, \dots, u_i$, do đó u_{i+1} cũng là tổ hợp tuyến tính của các vectơ $v_{i+1}, v_1, \dots, v_i$, nghĩa là $\langle u_1, \dots, u_{i+1}\rangle \subseteq \langle v_1, \dots, v_{i+1}\rangle$. Hoàn toàn tương tự ta cũng có $\langle v_1, \dots, v_{i+1}\rangle \subseteq \langle u_1, \dots, u_{i+1}\rangle$. □

Ví dụ 23. Trong không gian Euclid $\mathbb{R}^4$ với tích vô hướng chính tắc, cho W là không gian con sinh bởi các vectơ $v_1 = (1, 1, 0, 0)$, $v_2 = (1, 0, -1, 1)$, $v_3 = (0, 1, 1, 1)$. Hãy tìm một cơ sở trực chuẩn cho W.

Lời giải. Nhận xét rằng các vectơ v_1, v_2, v_3 độc lập tuyến tính nên chúng tạo thành một cơ sở của W.

Đặt $u_1 := v_1$ và $u_2 := v_2 + \lambda_1 u_1$, với

$$\lambda_1 = -\frac{\langle v_2, u_1\rangle}{||u_1||^2} = -\frac{1}{2},$$

nghĩa là

$$u_2 = (1, 0, -1, 1) + \Big(-\frac{1}{2}\Big)(1, 1, 0, 0) = \frac{1}{2}\big(1, -1, -2, 2\big).$$

Nhận xét rằng, nếu thay u_2 bởi $u_2' = \alpha u_2, \alpha \in \mathbb{R}$ thì u_1 và u_2' vẫn trực giao với nhau. Do đó ta có thể lấy $u_2 = (1, -1, -2, 2)$. Tiếp theo, đặt

$$u_3 := v_3 + \lambda_1 u_1 + \lambda_2 u_2,$$

với

$$\lambda_1 = -\frac{\langle v_2, u_1\rangle}{||u_1||^2} = -\frac{1}{2}, \lambda_2 = -\frac{\langle v_3, u_2\rangle}{||u_2||^2} = -\frac{1}{10},$$

nghĩa là $u_3 = \frac{2}{5}(-1, 1, 2, 3)$. Tuy nhiên, ta có thể lấy $u_3 = (-1, 1, 2, 3)$. Khi đó $\mathcal{B} = (u_1, u_2, u_3)$ là một cơ sở trực giao của W. Trực chuẩn hóa cơ sở $\mathcal{B}$ ta nhận được cơ sở trực chuẩn $\mathcal{B}_c = (e_1, e_2, e_3)$ của W, với e_1, e_2, e_3 như sau: $e_1 = \dfrac{1}{\sqrt{2}}\big(1, 1, 0, 0\big), e_2 = \dfrac{1}{\sqrt{10}}\big(1, -1, -2, 2\big), e_3 = \dfrac{1}{\sqrt{15}}\big(-1, 1, 2, 3\big)$. □

Mệnh đề 2.7. *Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và $\mathcal{B} = (u_1, \dots, u_n)$ là một cơ sở trực chuẩn của V. Khi đó, với mọi $u \in V$, ta có $u = \langle u, u_1\rangle u_1 + \cdots + \langle u, u_n\rangle u_n$.*

Chứng minh. Giả sử $u = x_1u_1 + x_2u_2 + \cdots + x_nu_n$. Khi đó

$$\langle u, u_i\rangle = x_1\langle u_1, u_i\rangle + x_2\langle u_2, u_i\rangle + \cdots + x_n\langle u_n, u_i\rangle.$$

Vì $\mathcal{B}$ là cơ sở trực chuẩn nên $\langle u_i, u_j\rangle = \delta_{ij}$ Suy ra $\langle u, u_i\rangle = x_i$. □

Định nghĩa 2.8. Cho V là một không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và S là tập con khác rỗng của V. Tập tất cả các vectơ u thuộc V sao cho u trực giao với mọi vectơ trong S được gọi là *phần bù trực giao* của S, ký hiệu bởi $S^\perp$, nghĩa là

$$S^\perp := \{u \in V \mid \langle u, v\rangle = 0, \forall v \in S\}.$$

Dễ thấy $0^\perp = V$ và $V^\perp = 0$.

Mệnh đề 2.9. *Cho V là một không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và S là tập con khác rỗng của V. Khi đó $S^\perp$ là không gian con của V.*

Không gian $S^\perp$ còn được gọi là *không gian con trực giao* với S.

Chứng minh. Dễ dàng chứng minh. □

Mệnh đề 2.10. *Cho V là một không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và W là không gian con của V. Khi đó*

$$V = W \oplus W^\perp.$$

Chứng minh. Đặt $n = \dim V$ và $m = \dim W$. Theo Định lý Gram-Schmidt (Định lý 2.6), tồn tại cơ sở trực chuẩn $\mathcal{B} = (u_1, \ldots, u_m)$ của W. Bổ sung $n - m$ vectơ vào $\mathcal{B}$ để được một cơ sở của V và trực giao hóa cơ sở này bằng quá trình trực giao hóa Gram-Schmidt, ta được một cơ sở trực chuẩn của V có dạng

$$\mathcal{B}' = (u_1, \ldots, u_m, \ldots, u_n).$$

Đặt $W' = \langle u_{m+1}, \ldots, u_n \rangle$, suy ra $V = W + W'$. Hơn nữa, do $\mathcal{B}'$ là cơ sở trực chuẩn của V nên $W' \subseteq W^\perp$. Ngược lại, với mọi $u \in W^\perp$, từ Mệnh đề 2.7 ta được

$$\begin{aligned} u &= [\langle u, u_1 \rangle u_1 + \cdots + \langle u, u_m \rangle u_m] + \langle u, u_{m+1} \rangle u_{m+1} + \cdots + \langle u, u_n \rangle u_n \\ &= \langle u, u_{m+1} \rangle u_{m+1} + \cdots + \langle u, u_n \rangle u_n, \end{aligned}$$

nên $u \in W'$. Vậy $W^\perp = W'$. Cuối cùng, nếu $u \in W \cap W^\perp$ thì $\langle u, u \rangle = 0$ nên $u = 0$. Do đó $W \cap W^\perp = \{0\}$. Vậy $V = W \oplus W^\perp$. □

Từ mệnh đề trên, ta suy ra ngay Hệ quả sau:

Hệ quả 2.11. *Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và W là không gian con của V. Khi đó*

i) $\dim(W^\perp) = \dim(V) - \dim(W)$.

ii) $(W^\perp)^\perp = W$.

Từ công thức $V = W \oplus W^\perp$ trong Mệnh đề 2.10 suy ra mỗi vectơ $u \in V$ đều viết được một cách duy nhất dưới dạng

$$u = x_0 + v, \qquad \text{với } x_0 \in W, v \in W^\perp.$$

Ta gọi x_0 là *hình chiếu trực giao* của u lên W, ký hiệu $x_0 =: pr_W(u)$.

Mệnh đề sau đây cho ta một cách tính hình chiếu trực giao của một vectơ u lên không gian con W của V.

Mệnh đề 2.12. *Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và W là không gian con của V. Giả sử $(u_1, \ldots, u_m)$ là một cơ sở trực chuẩn của W. Khi đó, với mọi $u \in V$ ta có:*

$$pr_W(u) = \langle u, u_1\rangle u_1 + \cdots + \langle u, u_m\rangle u_m.$$

Chứng minh. Gọi $(u_{m+1}, \ldots, u_n)$ là một cơ sở trực chuẩn của $W^\perp$. Khi đó, theo Mệnh đề 2.10 ta có $(u_1, \ldots, u_m, u_{m+1}, \ldots, u_n)$ là một cơ sở trực chuẩn của V. Đặt

$$v_1 = \langle u, u_1\rangle u_1 + \cdots + \langle u, u_m\rangle u_m; \quad v_2 = \langle u, u_{m+1}\rangle u_{m+1} + \cdots + \langle u, u_n\rangle u_n.$$

Áp dụng Hệ quả 2.7, nhận được $u = v_1 + v_2$. Lưu ý rằng $v_1 \in W$ và $v_2 \in W^\perp$. Áp dụng Mệnh đề 2.7, suy ra $pr_W(u) = v_1$. □

Ví dụ 24. Trong không gian Euclid $\mathbb{R}^4$ với tích vô hướng chính tắc, cho W là không gian con sinh bởi các vectơ $v_1 = (1, 1, 0, 0)$, $v_2 = (1, 0, -1, 1)$, $v_3 = (0, 1, 1, 1)$. Tìm hình chiếu trực giao của vectơ $u = (1, 2, 0, 3)$ lên W.

Lời giải. Từ kết quả của Ví dụ 23 ta được một cơ sở trực chuẩn của W là: $\mathcal{B}_c = (e_1, e_2, e_3)$, với

$$e_1 = \frac{1}{\sqrt{2}}\big(1, 1, 0, 0\big), e_2 = \frac{1}{\sqrt{10}}\big(1, -1, -2, 2\big), e_3 = \frac{1}{\sqrt{15}}\big(-1, 1, 2, 3\big).$$

Ta có

- $\langle u, e_1\rangle e_1 = \dfrac{1}{2}(1, 2, 0, 0)$,
- $\langle u, e_2\rangle e_2 = \dfrac{1}{10}(1, 2, 0, 12)$,
- $\langle u, e_3\rangle e_3 = \dfrac{1}{15}(1, 2, 0, 27)$.

Do đó hình chiếu trực giao của u lên W là

$$pr_W(u) = \langle u, e_1\rangle e_1 + \langle u, e_2\rangle e_2 + \langle u, e_3\rangle e_3 = \frac{1}{3}(2, 4, 0, 9).$$

□

Định nghĩa 2.13. Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$; W là không gian con của V và $u, v \in V$. Đặt

$$d(u, v) := ||u - v|| \quad \text{và} \quad d(u, W) = \min\{d(u, w), w \in W\}.$$

Ta gọi $d(u, v)$ là *khoảng cách* giữa u và v; $d(u, W)$ là *khoảng cách* từ u đến W.

Bổ đề 2.14. *Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và $u, v, w \in V$. Khi đó:*

i) $d(u, v) \geq 0$; $d(u, v) = 0$ *khi và chỉ khi* $u = v$.

ii) $d(u, v) = d(v, u)$.

iii) $d(u, w) \leq d(u, v) + d(v, w)$.

Chứng minh. (i) và (ii) được suy ra từ định nghĩa khoảng cách và chuẩn.

(iii) Ta có $d(u, w) = ||u - w|| = ||(u - v) + (v - w)|| \leq ||u - v|| + ||v - w|| = d(u, v) + d(v, w)$. □

Mệnh đề 2.15. *Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$; W là không gian con của V và $u \in V$. Khi đó*

$$d(u, W) = ||u - pr_W(u)||.$$

Chứng minh. Đặt $v = pr_W(u)$ và $u_0 = u - v$. Suy ra $v \in W$ và $u_0 \in W^\perp$.

Do $v \in W$ nên $d(u, W) \leq d(u, v)$. Ngược lại, với mọi $w \in W$ ta có $\langle u, w\rangle = \langle u_0 + v, w\rangle = \langle u_0, v\rangle + \langle v, w\rangle = \langle v, w\rangle$ và $\langle u, v\rangle = \langle u_0 + v, v\rangle =$

$\langle u_0, v\rangle + \langle v, v\rangle = ||v||^2$. Do đó $d(u,w)^2 - d(u,v)^2 = ||u-w||^2 - ||u-v||^2 = ||w||^2 - ||v||^2 - 2\langle u, w\rangle + 2\langle u, v\rangle = ||w||^2 - 2\langle v, w\rangle + ||v||^2 = (||w|| - ||v||)^2 \geq 0$. Suy ra $d(u,w) \geq d(u,v)$. Vậy $d(u,v) = \min\{d(u,w), w \in W\}$. $\square$

Ví dụ 25. Với u và W như trong Ví dụ 24, ta có $pr_W(u) = \frac{1}{3}(2,4,0,9)$, nên $u - pr_W(u) = \frac{1}{3}(1,2,0,0)$. Do đó $d(u,W) = ||u - pr_W(u)|| = \frac{\sqrt{5}}{3}$.

§3. Toán tử đối xứng

Trong các chương 4, 5 và 6 ta đã nghiên cứu tính chất của các toán tử tuyến tính trong không gian vectơ trên trường K, với K là trường số thực $\mathbb{R}$ hoặc trường số phức $\mathbb{C}$. Do không gian Euclid trước hết cũng là không gian vectơ nên nó thỏa mãn tất cả những tính chất đã đề cập đến trong các chương kể trên. Ngoài ra, không gian Euclid còn được trang bị tích vô hướng nên nó còn có thêm những tính chất khác nữa mà những không gian vectơ thông thường không có. Trong mục này và mục tiếp theo ta sẽ đề cập đến những tính chất như vậy.

Định nghĩa 3.1. Cho V là không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ và f là toán tử tuyến tính trên V. Ta nói f là một *toán tử đối xứng* (hay *toán tử tự liên hợp*) nếu

$$\langle f(u), v\rangle = \langle u, f(v)\rangle, \forall u, v \in V.$$

Mệnh đề 3.2. *Ma trận biểu diễn toán tử đối xứng trên V theo một cơ sở trực chuẩn của V là ma trận đối xứng.*

Chứng minh. Giả sử $\mathcal{B} = (e_1, \ldots, e_n)$ là một cơ sở trực chuẩn của V. Khi đó, từ Mệnh đề 2.3 và Hệ quả 2.4 ta được

$$[f(u)]_{\mathcal{B}}^{\top}[v]_{\mathcal{B}} = [u]_{\mathcal{B}}^{\top}[f(v)]_{\mathcal{B}},$$

nghĩa là

$$[u]_{\mathcal{B}}^{\top}\,[f]_{\mathcal{B}}^{\top}\,[v]_{\mathcal{B}} = [u]_{\mathcal{B}}^{\top}\,[f]_{\mathcal{B}}\,[v]_{\mathcal{B}}.$$

Lần lượt thay u và v bởi các vectơ của cơ sở $\mathcal{B}$, ta có $[f]_{\mathcal{B}}^{\top} = [f]_{\mathcal{B}}$. Vậy $[f]_{\mathcal{B}}$ là ma trận đối xứng. □

Định lý 3.3. *Cho V là một không gian Euclid với tích vô hướng $\langle\ ,\ \rangle$ trên $\mathbb{R}$ và f là một toán tử tuyến tính trên V. Khi đó*

i) *Mọi trị riêng của f đều là số thực.*

ii) *f chéo hóa được.*

iii) *Các không gian con riêng của f đôi một trực giao.*

Chứng minh. Gọi $\mathcal{B}$ là một cơ sở trực chuẩn của V và đặt $A = [f]_{\mathcal{B}}$ và $p(t)$ là đa thức đặc trưng của A.

i) Gọi λ là một nghiệm (trên $\mathbb{C}$) của $p(t)$. Khi đó, hệ $(A - \lambda I_n)X = 0$ có nghiệm không tầm thường $u = (\alpha_1, \ldots, \alpha_n)$, với $\alpha_i \in \mathbb{C}$. Đặt $\overline{u} = (\overline{\alpha_1}, \ldots, \overline{\alpha_n})$, trong đó $\overline{\alpha_i}$ là số phức liên hợp của α_i. Khi đó

$$u.\overline{u}^{\top} = \sum_{i=1}^{n} \alpha_i \overline{\alpha_i} = \sum_{i=1}^{n} |\alpha_i|^2 > 0.$$

Do A là ma trận đối xứng nên $A^{\top} = A$. Do đó, từ đẳng thức $A.u^{\top} = \lambda u^{\top}$ ta suy ra $\lambda u = (A.u^{\top})^{\top} = uA^{\top} = uA$. Hơn nữa, A là ma trận thực nên $\overline{A} = A$. Suy ra $\overline{\lambda}(u.\overline{u}^{\top}) = u(\overline{\lambda u^{\top}}) = u.\overline{A.u^{\top}} = u.\overline{A}.\overline{u}^{\top} = u.A.\overline{u}^{\top} = \lambda u.u^{\top}$. Do đó $\lambda \in \mathbb{R}$.

ii) Ta chứng minh ii) bằng qui nạp theo $n = \dim V$. Hiển nhiên ii) đúng với $n = 1$. Với $n > 1$, giả sử ii) đúng với mọi không gian có số chiều bằng $n - 1$. Xét một trị riêng λ của f và u là một vectơ riêng ứng với trị riêng λ. Đặt $H := \langle u \rangle^{\perp}$. Khi đó $\dim(H) = n - 1$. Trước hết ta chứng minh H là không gian con bất biến đối với f. Thật vậy, với mọi $v \in H$, ta có $\langle u, f(v) \rangle = \langle f(u), v \rangle = \langle \lambda u, v \rangle = \lambda \langle u, v \rangle = 0$, nghĩa là $f(v) \in H$. Do H là không gian con bất biến đối với f nên $\tilde{f} := f|_H$ là toán tử tuyến tính trên H. Hơn nữa, hiển nhiên $\tilde{f}$ cũng là toán tử đối xứng, nên theo giả thiết qui nạp, $\tilde{f}$ chéo hoá được. Suy ra tồn tại một cơ sở $\mathcal{B}$ của H gồm toàn các vectơ riêng của $\tilde{f}$ (cũng là của f). Khi đó $\{u\} \cup \mathcal{B}$ là một cơ sở của V gồm toàn các vectơ riêng của f. Vậy f chéo hóa được.

iii) Giả sử $\mu \neq \lambda$ là hai trị riêng khác nhau của f; u là vectơ riêng của f ứng với trị riêng λ; v là vectơ riêng của f ứng với μ. Khi đó $\langle f(u), v\rangle = \langle u, f(v)\rangle$ nên $\langle \lambda u, v\rangle = \langle u, \mu v\rangle$, nghĩa là $(\lambda - \mu)\langle u, v\rangle = 0$. Mà $\lambda \neq u$, nên $\langle u, v\rangle = 0$. $\square$

Lưu ý rằng một ma trận phức đối xứng không nhất thiết chéo hóa được trên $\mathbb{R}$ hoặc thậm chí trên $\mathbb{C}$.

Ví dụ sau đây minh họa cho nhận xét trên.

Ví dụ 26. Xét $A = \begin{pmatrix} 0 & 1 \\ 1 & 2i \end{pmatrix}$. Ta có đa thức đặc trưng của A là

$$p(t) = |A - tI_2| = \begin{vmatrix} -t & 1 \\ 1 & 2i - t \end{vmatrix} = (t - i)^2.$$

Do đó A chéo hóa được khi và chỉ khi đa thức tối tiểu của A có dạng $m(t) = t - i$. Nhưng $m(A) = A - i.I_2 = \begin{pmatrix} -i & 1 \\ 1 & i \end{pmatrix} \neq 0$ nên ta có điều mâu thuẫn. Vậy A không chéo hóa được.

§4. Toán tử trực giao

Trong mục này ta sẽ nghiên cứu những toán tử tuyến tính của một không gian Euclid bảo toàn chuẩn của các vectơ, nghĩa là nghiên cứu những toán tử f thỏa mãn tính chất $||f(u)|| = ||u||$.

Định nghĩa 4.1. Cho V là một không gian Euclid và f là một toán tử tuyến tính trên V. Ta nói f là một *toán tử trực giao* nếu

$$||f(u)|| = ||u||, \ \forall u \in V.$$

Mệnh đề 4.2. *Cho V là không gian Euclid và f là toán tử tuyến tính trên V. Khi đó f là toán tử trực giao khi và chỉ khi*

$$\langle f(u), f(v)\rangle = \langle u, v\rangle, \forall u, v \in V.$$

Chứng minh. Chiều đảo là hiển nhiên. Chiều thuận được suy ra từ Nhận xét 1.5b. □

Định nghĩa 4.3. Một ma trận $A \in M_n(\mathbb{R})$ được gọi là *ma trận trực giao* nếu A khả nghịch và $A^{-1} = A^\top$. Tập hợp tất cả các ma trận trực giao cấp n trên $\mathbb{R}$ được ký hiệu bởi $O(n, \mathbb{R})$.

Nhận xét 4.4. Tập hợp $O_n(\mathbb{R})$ là một nhóm con (với phép nhân ma trận) của $\mathrm{GL}_n(\mathbb{R})$, gọi là *nhóm trực giao*. Hơn nữa, nếu $A \in O(n, \mathbb{R})$ thì $\det A = \pm 1$.

Mệnh đề 4.5. *Cho V là không gian Euclid; $\mathcal{B}$ là một cơ sở trực chuẩn của V và f là toán tử tuyến tính trên V. Khi đó f là toán tử trực giao khi và chỉ khi $[f]_\mathcal{B}$ là ma trận trực giao.*

Chứng minh. Đặt $A = [f]_\mathcal{B}$. Khi đó, với mọi $u, v \in V$ ta có $\langle f(u), f(v)\rangle = \langle u, v\rangle \Leftrightarrow [f(u)]_\mathcal{B}^\top [f(v)]_\mathcal{B} = [u]_\mathcal{B}^\top [v]_\mathcal{B} \Leftrightarrow ([f]_\mathcal{B} [u]_\mathcal{B})^\top ([f]_\mathcal{B} [v]_\mathcal{B}) = [u]_\mathcal{B}^\top [v]_\mathcal{B}$ $\Leftrightarrow [u]_\mathcal{B}^\top (A^\top A) [v]_\mathcal{B} = [u]_\mathcal{B}^\top [v]_\mathcal{B}$.

Do đó, nếu $A^{-1} = A^\top$ thì $\langle f(u), f(v)\rangle = \langle u, v\rangle, \forall u, v \in V$, nghĩa là f trực giao. Ngược lại, nếu f trực giao thì bằng cách chọn u, v lần lượt là các vectơ cơ sở của $\mathcal{B}$ ta được $A^\top A = I_n$, nghĩa là A trực giao. □

Hệ quả 4.6. *Nếu f là một toán tử trực giao thì* $\det f = \pm 1$. *Nói riêng, f là một tự đẳng cấu.*

Mệnh đề 4.7. *Cho f là toán tử tuyến tính trên V và $\mathcal{B} = (e_1, \ldots, e_n)$ là cơ sở trực chuẩn của V. Khi đó f là một toán tử trực giao khi và chỉ khi $f(\mathcal{B}) := (f(e_1), \ldots, f(e_n))$ cũng là cơ sở trực chuẩn của V.*

Chứng minh. *Chiều thuận.* Giả sử f là một toán tử trực giao. Theo Hệ quả 4.6, f là một tự đẳng cấu, do đó f biến một cơ sở của V thành một cơ sở của V, nghĩa là $f(\mathcal{B})$ là cơ sở của V. Hơn nữa, do f trực giao nên với mọi i, j ta có

$$\langle f(e_i), f(e_j)\rangle = \langle e_i, e_j)\rangle = \delta_{ij}.$$

Vậy $f(\mathcal{B})$ cũng là cơ sở trực chuẩn của V.

Chiều đảo. Giả sử $(f(\mathcal{B})$ là cơ sở trực chuẩn của V. Suy ra

$$\langle f(e_i), f(e_j)\rangle = \delta_{ij} = \langle e_i, e_j\rangle \text{ với mọi } i, j.$$

Do đó, với mọi $u, v \in V$, bằng cách biểu diễn

$$u = \sum_{i=1}^{n} x_i e_i \text{ và } v = \sum_{j=1}^{n} y_j e_j$$

ta được

$$\langle f(u), f(v)\rangle = \sum_{i,j=1}^{n} x_i y_j \langle f(e_i), f(e_j)\rangle = \sum_{i,j=1}^{n} x_i y_j \langle e_i, e_j\rangle = \langle u, v\rangle.$$

Vậy f là toán tử trực giao. $\square$

Ví dụ 27. Xét $A = \dfrac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$. Bằng cách tính toán trực tiếp, ta được $A^\top A = I_3$, nên A là ma trận trực giao. Nhưng ta cũng có thể kiểm tra

A trực giao bằng cách khác như sau: Xét không gian Euclid $\mathbb{R}^3$ với tích vô hướng chính tắc. Đặt $u_1 = \frac{1}{3}(2,2,-1)$, $u_2 = \frac{1}{3}(-1,2,2)$, $u_3 = \frac{1}{3}(2,-1,2)$. Gọi f là toán tử tuyến tính trên $\mathbb{R}^3$ thỏa mãn $f(e_i) = u_i, i \in \{1,2,3\}$, với $\mathcal{B}_0 = (e_1,e_2,e_3)$ là cơ sở chính tắc của $\mathbb{R}^3$. Do f biến cơ sở trực chuẩn thành cơ sở trực chuẩn nên theo Mệnh đề 4.7, f là toán tử trực giao. Mà A là ma trận biểu diễn f theo cơ sở trực chuẩn $\mathcal{B}_0$ nên A là ma trận trực giao.

Mệnh đề 4.8. *Nếu $\mathcal{B}$ và $\mathcal{B}'$ là hai cơ sở trực chuẩn của V thì ma trận chuyển cơ sở từ $\mathcal{B}$ sang $\mathcal{B}'$ là ma trận trực giao.*

Chứng minh. Giả sử $\mathcal{B} = (e_1,\dots,e_n)$ và $\mathcal{B}' = (e'_1,\dots,e'_n)$. Gọi f là toán tử tuyến tính xác định bởi $f(e_i) = e'_i, \forall i \in \overline{1,n}$. Theo Mệnh đề 4.7, ta có f là toán tử trực giao. Suy ra $[f]_{\mathcal{B}}$ là ma trận trực giao. Hơn nữa $[f]_{\mathcal{B}} = (\mathcal{B} \to \mathcal{B}')$. Như vậy $(\mathcal{B} \to \mathcal{B}')$ ma trận trực giao. $\square$

Bài tập

Bài 1. Với giá trị nào của $\lambda \in \mathbb{R}$ các ánh xạ dưới đây xác định tích vô hướng trong không gian $\mathbb{R}^3$:

a) $\langle u,v\rangle = x_1y_1 + 10x_2y_2 + 6x_1y_2 + \lambda x_3y_3 - x_2y_3 - x_3y_2$.

b) $\langle u,v\rangle = 2x_1y_1 + 7x_1y_2 + 7x_2y_1 + 8x_2y_2 - 3x_3y_3 + \lambda x_2y_3 + \lambda x_3y_2$.

Bài 2. Xét không gian Euclid $\mathbb{R}^n$ với tích vô hướng chính tắc. Chứng minh rằng với mọi $u = (x_1,\dots,x_n) \in \mathbb{R}^n$, ta có

$$\Big(\sum_{i=1}^n x_i\Big)^2 \le n\Big(\sum_{i=1}^n x_i^2\Big).$$

Bài 3. Cho không gian vectơ $M_n(\mathbb{R})$ gồm các ma trận vuông cấp n trên trường số thực $\mathbb{R}$.

a) Với $A = (a_{ij}) \in M_n(\mathbb{R})$, hãy tính vết $tr(AA^\top)$ theo a_{ij}. Suy ra $|tr(A)| \leq \sqrt{n.tr(AA^\top)}$.

b) Chứng minh rằng ánh xạ $(A, B) \mapsto tr(AB^\top)$ xác định một tích vô hướng trong không gian $M_n(\mathbb{R})$.

Bài 4. Trong không gian Euclid $\mathbb{R}^3$ với tích vô hướng chính tắc, cho P là mặt phẳng được xác định bởi phương trình $x_1 - 2x_2 + x_3 = 0$ và π là phép chiếu trực giao từ $\mathbb{R}^3$ xuống P. Hãy viết ma trận biểu diễn π theo cơ sở chính tắc của $\mathbb{R}^3$.

Bài 5. Trong không gian Euclid $\mathbb{R}^3$ với tích vô hướng chính tắc, cho $u_1 = (1, 0, 1), u_2 = (2, 1, 0)$ và $u_3 = (1, 1, 1)$.

a) Chứng minh rằng $\mathcal{B} = (u_1, u_2, u_3)$ là một cơ sở của $\mathbb{R}^3$.

b) Áp dụng quá trình trực giao hóa Gram-Schmidt để xây dựng một cơ sở trực chuẩn $\mathcal{B}' = (e_1, e_2, e_3)$ từ cơ sở $\mathcal{B}$ của $\mathbb{R}^3$.

Bài 6. Trong không gian Euclid $\mathbb{R}^3$ với tích vô hướng chính tắc, cho W là không gian sinh bởi các vectơ $u_1 = (1, 0, 1), u_2 = (2, 1, 2)$ và $u_3 = (1, 1, 1)$. Hãy tìm một cơ sở trực chuẩn cho không gian W.

Bài 7. Cho $P, Q \in \mathbb{R}[x]$, đặt $\langle P, Q\rangle := \int\limits_0^1 P(x)Q(x)dx$. Chứng minh rằng $\langle P, Q\rangle$ là một tích vô hướng trên không gian các đa thức $\mathbb{R}[x]$.

Bài 8. Với $n \geq 0$, xét tích phân suy rộng $I_n = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} dx$.

a) Chứng minh rằng tích phân I_n luôn hội tụ và $I_{2k+1} = 0, \forall k \geq 0$.

b) Chứng minh $I_n = (n-1)I_{n-2}, \forall n \geq 2$. Suy ra giá trị của I_{2k}.

c) Với $P(x), Q(x) \in \mathbb{R}[x]$, ta đặt $\langle P(x), Q(x)\rangle = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}} P(x)Q(x)dx$. Chứng minh rằng $\langle P(x), Q(x)\rangle$ là một tích vô hướng trên $\mathbb{R}[x]$.

d) Xét không gian con $\mathbb{R}_2[x]$ của không gian $\mathbb{R}[x]$ với tích vô hướng như trên. Hãy tính khoảng cách từ x^3 đến $\mathbb{R}_2[x]$.

Bài 9. Xét ánh xạ $\langle\ ,\ \rangle : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ xác định bởi:

$$\langle u, v\rangle = x_1y_1+2x_2y_2+4x_3y_3+18x_4y_4+x_1y_3+x_3y_1+2x_2y_4+2x_4y_2+6x_3y_4+6x_4y_3,$$

với $u = (x_1, x_2, x_3, x_4)$, $v = (y_1, y_2, y_3, y_4)$.

a) Chứng minh rằng $\langle\ ,\ \rangle$ là một tích vô hướng trong $\mathbb{R}^4$.

b) Viết ma trận biểu diễn tích vô hướng này theo cơ sở chính tắc.

c) Cho W là không gian nghiệm của hệ phương trình:

$$\begin{cases} x_1 - x_2 + 7x_3 - x_4 = 0; \\ x_2 - 2x_4 = 0. \end{cases}$$

Hãy tìm một cơ sở của $W^\perp$.

Bài 10. Trong không gian Euclide với tích vô hướng chính tắc, cho W là không gian sinh bởi các vectơ $u_1 = (2, 1, -2, 4), u_2 = (-2, 1, -1, -6), u_3 = (-2, 3, -4, -8)$.

a) Tìm một cơ sở cho mỗi không gian con W và $W^\perp$.

b) Cho $u = (5, 5, -3, 1) \in \mathbb{R}^4$. Tìm hình chiếu trực giao $pr_W(u)$ của u xuống W và tính khoảng cách $d(u, W)$ từ u đến W.

Bài 11. Cho $A = (a_{ij}) \in O_n(\mathbb{R})$. Chứng minh rằng, nếu $\det A = 1$ thì mỗi phần tử a_{ij} của A đều bằng phần bù đại số của nó.

Bài 12. Cho V là một không gian Euclid và f là một phép biến đổi trực giao trong V.

a) Chứng minh rằng $\operatorname{Ker}(f - \operatorname{Id}_V) = \operatorname{Im}(f - \operatorname{Id}_V)^\perp$.

b) Chứng minh rằng, nếu $(f - \operatorname{Id}_V)^2 = 0$ thì $f = \operatorname{Id}_V$.

Bài 13. Trong không gian Euclid $\mathbb{R}^3$ với tích vô hướng chính tắc, cho toán tử tuyến tính f có ma trận biểu diễn theo cơ sở chính tắc là:

$$A = \begin{pmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}.$$

Hãy xây dựng một cơ sở trực chuẩn của $\mathbb{R}^3$ từ các vectơ riêng f.

Bài 14. Trong không gian Euclid $\mathbb{R}^3$ với tích vô hướng chính tắc, cho toán tử tuyến tính f có ma trận biểu diễn theo cơ sở chính tắc là

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

Chứng minh rằng f là toán tử trực giao.

Chương 8

KHÔNG GIAN UNITA

§1. Không gian unita

Định nghĩa 1.1. *Không gian unita* V là một không gian vectơ phức hữu hạn chiều có trang bị một tích vô hướng, tức là một ánh xạ $V \times V \to V$; $(u, v) \mapsto \langle u, v\rangle$ và thỏa mãn các tính chất sau với mọi $u, v, u_1, u_2 \in V$ và với mọi $\alpha \in \mathbb{C}$:

i) $\langle u, v\rangle = \overline{\langle v, u\rangle}$

ii) $\langle u, u\rangle \in \mathbb{R}^+ \cup \{0\}$ và $\langle u, u\rangle = 0 \Leftrightarrow u = 0$

iii) $\langle u_1 + u_2, v\rangle = \langle u_2, v\rangle + \langle u_2, v\rangle$

iv) $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$

Nhận xét 1.2.
1) Dễ dàng kiểm tra các tính chất sau của không gian unita:

i) $\langle u, v_1 + v_2\rangle = \langle u, v_1\rangle + \langle u, v_2\rangle$;

ii) $\langle u, \alpha v\rangle = \overline{\alpha}\,\langle u, v\rangle$.

2) Không gian vectơ con của không gian unita cũng là không gian unita.

Ví dụ 28. Không gian vectơ $\mathbb{C}^n$ là không gian unita với tích vô hướng

$$\langle u, v\rangle = x_1\overline{y_1} + \cdots + x_n\overline{y_n}, \text{với} u = (x_1, \ldots, x_n),\ v = (y_1, \ldots, y_n) \in \mathbb{C}^n.$$

Trong chương này, khi nói đến không gian $\mathbb{C}^n$ mà không có chú thích gì thêm thì ta hiểu đó là không gian unita với tích vô hướng như trên.

Ví dụ 29. Không gian các đa thức hệ số phức $\mathbb{C}[x]$ là không gian unita với tích vô hướng

$$\langle P, Q\rangle = \int\limits_0^1 P(x)\overline{Q(x)}\,dx \text{ với mọi } P, Q \in \mathbb{C}[x].$$

Trong không gian unita, chuẩn hay độ dài của các vectơ được định nghĩa tương tự như trong không gian Euclid. Khi đó, các tính chất cơ bản của chuẩn, bất đẳng thức Cauchy-Schwarz, bất đẳng thức tam giác trong không gian Euclid vẫn còn đúng trong không gian unita. Cũng vậy, trong không gian unita, sự trực giao giữa các vectơ được định nghĩa hoàn toàn tương tự như trong không gian Euclid, từ đó ta cũng có các khái niệm như không gian con trực giao, ma trận biểu diễn của tích vô hướng, cơ sở trực giao, cơ sở trực chuẩn, quá trình trực giao hóa Gram- Schmidt, khoảng cách giữa các vectơ, hình chiếu trực giao của một vectơ lên một không gian con,... Nói chung, các tính chất liên quan đến sự trực giao trong không gian Euclid vẫn còn đúng trong không gian Uniata. Ở đây ta chỉ nêu lại một số kết quả có một ít thay đổi so với không gian Euclid

Mệnh đề 1.3. *Cho V là một không gian unita, $\mathcal{B}$ là một cơ sở của V và $\langle\ ,\ \rangle_{\mathcal{B}}$ là ma trận biểu diễn của tích vô hướng tương ứng trong cơ sở $\mathcal{B}$. Khi đó, với mọi $u, v \in V$ ta có $\langle u, v\rangle = [u]_{\mathcal{B}}^{\top}\langle\ ,\ \rangle_{\mathcal{B}}\overline{[v]_{\mathcal{B}}}$.*

Mệnh đề 1.4. *Cơ sở $\mathcal{B}$ của không gian unita V là cơ sở trực chuẩn khi và chỉ khi với mọi $u, v \in V$ ta có $\langle u, v\rangle = [u]_{\mathcal{B}}^{\top}\overline{[v]_{\mathcal{B}}}$.*

Mệnh đề 1.5. *Cho V là một không gian unita, $\mathcal{B}$ và $\mathcal{B}'$ là hai cơ sở của V. Khi đó $\langle\ ,\ \rangle_{\mathcal{B}'} = (\mathcal{B} \to \mathcal{B}')^{\top}\langle\ ,\ \rangle_{\mathcal{B}}\overline{(\mathcal{B} \to \mathcal{B}')}$.*

§2. Toán tử tuyến tính liên hợp

Định lý 2.1. *Cho V là không gian unita và f là một phiếm hàm tuyến tính trên V. Khi đó tồn tại duy nhất vectơ $v \in V$ sao cho $f(u) = \langle u, v\rangle$ với mọi $u \in V$.*

Chứng minh. Cho $\mathcal{B} = (e_1, \ldots, e_n)$ là một cơ sở trực chuẩn của V. Đặt $v = \sum\limits_{j=1}^{n} \overline{f(e_j)}.e_j$ và gọi f_v là phiếm hàm tuyến tính trên V xác định bởi $f_v(u) = \langle u, v\rangle$. Khi đó, $f_v(e_i) = \left\langle e_i, \sum\limits_{j=1}^{n} \overline{f(e_j)}\, e_j \right\rangle = f(e_i)$ với mọi $1 \le i \le n$, nên $f = f_v$, nghĩa là $f(u) = \langle u, v\rangle$ với mọi $u \in V$. Nếu tồn tại $w \in V$ sao cho $f(u) = \langle u, w\rangle$ với mọi $u \in V$ thì $\langle u, v - w\rangle = \langle u, v\rangle - \langle u, w\rangle = f(u) - f(u) = 0$ với mọi $u \in V$, nên $v - w \in V^{\perp}$. Suy ra $w = v$. $\square$

Định lý 2.2. *Cho V là không gian unita và f là một toán tử tuyến tính trên V. Khi đó tồn tại duy nhất một toán tử tuyến tính f^* trên V sao cho $\langle f(u), v\rangle = \langle u, f^*(v)\rangle$ với mọi $u, v \in V$. Ta gọi f^* là toán tử tuyến tính liên hợp của f trên V.*

Chứng minh. Cho vectơ v bất kỳ trong V. Khi đó ánh xạ $u \mapsto \langle f(u), v\rangle$ là toán tử tuyến tính trên V nên theo Định lý 2.1 tồn tại duy nhất một vectơ $v' \in V$ sao cho $\langle f(u), v\rangle = \langle u, v'\rangle$ với mọi $u \in V$. Gọi f^* là ánh xạ định bởi $f^*(v) = v'$. Khi đó $\langle f(u), v\rangle = \langle u, f^*(v)\rangle$ với mọi $u \in V$. Mặt khác, với mọi $u, v, w \in V$ và với mọi $\alpha \in \mathbb{C}$, ta có $\langle u, f^*(\alpha v + w)\rangle = \langle f(u), \alpha v + w\rangle$ $= \overline{\alpha}\langle f(u), v\rangle + \langle f(u), w\rangle = \overline{\alpha}\langle u, f^*(v)\rangle + \langle u, f^*(w)\rangle = \langle u, \alpha f^*(v) + f^*(w)\rangle$. Suy ra $f^*(\alpha v + w) = \alpha f^*(v) + f^*(w)$, nên f^* là toán tử tuyến tính. Nếu g cũng là một toán tử tuyến tính trên V thỏa mãn $\langle f(u), v\rangle = \langle u, g(v)\rangle$ với mọi $u, v \in V$ thì $\langle u, f^*(v) g(v)\rangle = \langle u, f^*(v)\rangle - \langle u, g(v)\rangle = \langle f(u), v\rangle - \langle f(u), v\rangle = 0$ với mọi $u, v \in V$, nên $f^*(v) = g(v)$, nghĩa là $g = f^*$. $\square$

Định nghĩa 2.3. Cho $A = (a_{ij}) \in Mm \times n(\mathbb{C})$. *Ma trận liên hợp chuyển vị* của A, ký hiệu là A^*, là ma trận $(a'_{ji}) \in M_{n\times m}(\mathbb{C})$ xác định bởi $a'_{ji} = \overline{a_{ij}}$ với mọi $1 \le i \le m, 1 \le j \le n$.

Định lý 2.4. *Cho V là không gian unita và f là một toán tử tuyến tính trên V. Khi đó, nếu A ma trận biểu diễn của f và B là ma trận biểu diễn của toán tử tuyến tính liên hợp f^* theo cùng một cơ sở trực chuẩn của V thì $B = A^*$.*

Chứng minh. Giả sử $\mathcal{B} = (e_1, \dots, e_n)$ là một cơ sở trực chuẩn của V. Khi đó mỗi vectơ $u \in V$ được viết dưới dạng $u = \sum_{j=1}^{n} \langle u, e_j\rangle e_j$ nên với mọi $1 \le i \le n$, ta có $f(e_i) = \sum_{j=1}^{n} \langle f(e_i), e_j\rangle e_j$. Suy ra $[f]_{\mathcal{B}} = (a_{ij})$, với $a_{ij} = \langle f(e_i), e_j\rangle$. Áp dụng kết quả trên cho f^* ta được $[f^*]_{\mathcal{B}} = (a'_{ij})$, với $a'_{ij} = \langle f^*(e_i), e_j\rangle = \overline{\langle e_j, f^*(e_i)\rangle} = \overline{\langle f(e_j), e_i\rangle} = \overline{a_{ij}}$, nghĩa là $[f^*]_{\mathcal{B}} = [f]^*_{\mathcal{B}}$. □

Định lý 2.5. *Cho V là không gian unita hữu hạn chiều trên $\mathbb{C}$; f, g là hai toán tử tuyến tính trên V và $\alpha \in \mathbb{C}$. Khi đó:*

i) $(f^*)^* = f$;

ii) $(\alpha f)^* = \overline{\alpha}\, f^*$;

iii) $(f + g)^* = f^* + g^*$;

iv) $(fg)^* = g^* f^*$;

v) *Nếu f là đẳng cấu thì* $(f^{-1})^* = (f^*)^{-1}$.

Chứng minh. Với mọi $u, v \in V$ ta có:

i) $\langle u, (f^*)^*(v)\rangle = \langle f^*(u), v\rangle = \overline{\langle v, f^*(u)\rangle} = \overline{\langle f(v), u\rangle} = \langle u, f(v)\rangle$, nên $(f^*)^* = f$.

ii) $\langle u, (\alpha f)^*(v)\rangle = \langle \alpha f(u), v\rangle = \alpha\langle f(u), v\rangle = \alpha\langle u, f^*(v)\rangle = \langle u, \overline{\alpha} f^*(v)\rangle$, nên $(\alpha f)^* = \overline{\alpha} f^*$.

iii) $\langle u, (f+g)^*(v)\rangle = \langle (f+g)(u), v\rangle = \langle f(u), v\rangle + \langle g(u), v\rangle = \langle u, f^*(v)\rangle + \langle u, g^*(v)\rangle = \langle u, f^*(v) + g^*(v)\rangle$, nên $(f+g)^* = f^* + g^*$.

iv) $\langle u, (fg)^*(v)\rangle = \langle (fg)(u), v\rangle = \langle f(g(u)), v\rangle = \langle g(u), f^*(v)\rangle = \langle u, g^*(f^*(v))\rangle$, nên $(fg)^* = g^* f^*$.

v) $(f^{-1})^* f^* = (f f^{-1})^* = \mathrm{Id}_V^* = \mathrm{Id}_V$. Tương tự $f^*(f^{-1})^* = \mathrm{Id}_V$, nên $(f^{-1})^* = (f^*)^{-1}$. □

Từ định lý trên ta được kết quả sau:

Định lý 2.6. *Với $A, B \in M_n(\mathbb{C})$ và $\alpha \in \mathbb{C}$, ta có*

i) $(A^*)^* = A$;

ii) $(\alpha A)^* = \overline{\alpha}\, A^*$;

iii) $(A+B)^* = A^* + B^*$;

iv) $(AB)^* = B^* A^*$;

v) *Nếu A khả nghịch thì* $(A^{-1})^* = (A^*)^{-1}$.

§3. Toán tử chuẩn tắc

Định nghĩa 3.1. Toán tử tuyến tính f trên không gian unita hữu hạn chiều được gọi là *toán tử chuẩn tắc* nếu f và f^* giao hoán nhau, nghĩa là $f\,f^* = f^* f$.

Bổ đề 3.2. *Cho V là không gian vectơ hữu hạn chiều trên trường K và f, g là hai toán tử tuyến tính giao hoán nhau trên V. Khi đó mọi không* *gian riêng $E(\lambda)$ của f đều bất biến bởi g, nghĩa là $g(E(\lambda)) \subseteq E(\lambda)$.*

Chứng minh. Cho $\lambda \in K$ là một trị riêng của f. Với mọi $u \in E(\lambda)$ ta có $f(u) = \lambda u$ nên $f(g(u)) = (fg)(u) = (gf)(u) = g(f(u)) = g(\lambda u) = \lambda g(u)$. Do đó $g(u) \in E(\lambda)$. $\square$

Bổ đề 3.3. *Cho V là không gian unita hữu hạn chiều, f là toán tử tuyến tính trên V và W là một không gian con của V bất biến bởi f. Khi đó không gian con trực giao $W^\perp$ bất biến bởi toán tử liên hợp f^* của f.*

Chứng minh. Với mọi $u \in W$, $v \in W^\perp$ ta có $f(u) = \lambda u$ nên $f(u) \in W$ và $\langle u, f^*(v)\rangle = \langle f(u), v\rangle = 0$. Do đó $f^*(v) \in W^\perp$. $\square$

Định lý sau đây cho thấy cấu trúc của các toán tử chuẩn tắc trên các không gian unita hữu hạn chiều.

Định lý 3.4. *Toán tử tuyến tính f trên không gian unita V hữu hạn chiều là toán tử chuẩn tắc khi và chỉ khi tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_\mathcal{B}$ là một ma trận chéo.*

Chứng minh. $(\Leftarrow)$ Giả sử tồn tại cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_\mathcal{B}$ là ma trận chéo. Khi đó $[f^*]_\mathcal{B} = {[f]_\mathcal{B}}^*$ cũng là một ma trận chéo nên $[f\,f^*]_\mathcal{B} = [f^*]_\mathcal{B}[f]_\mathcal{B} = [f]_\mathcal{B}[f^*]_\mathcal{B} = [f^*f\,]_\mathcal{B}$. Suy ra $f\,f^* = f^*f$.

$(\Rightarrow)$ Ta chứng minh bằng quy nạp theo $n = \dim V$. Với $n = 1$, kết quả trên là hiển nhiên. Giả sử kết quả trên đúng cho mọi không không unita U có $\dim U \leq n-1$. Do tính đóng đại số của $\mathbb{C}$, f có ít nhất một trị riêng trong $\mathbb{C}$. Gọi $\lambda \in \mathbb{C}$ là một trị riêng của f và $W = E(\lambda)$, ta có $V = W \oplus W^\perp$. Hiển nhiên W bất biến bởi f, nên theo Bổ đề 3.3, $W^\perp$ bất biến bởi f^*. Do tính chuẩn tắc của f, theo Bổ đề 3.2, W cũng bất biến bởi f^*. Từ đây, lại theo Bổ đề 3.3, $W^\perp$ bất biến bởi $(f^*)^* = f$. Như vậy, W và $W^\perp$ đều bất biến bởi f và f^*, do đó $f|_W$ và $f|_{W^\perp}$ lần lượt là các toán tử chuẩn tắc trên các không gian unita W và $W^\perp$. Do $\dim W$ và $\dim W^\perp$ đều nhỏ hơn n nên theo quy nạp, tồn tại các cơ sở trực chuẩn $\mathcal{B}_1$ của W và $\mathcal{B}_2$ của $W^\perp$ sao cho $[f|_W]_{\mathcal{B}_1}$, $[f|_{W^\perp}]_{\mathcal{B}_2}$ là các ma trận chéo. Đặt $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$, ta được $\mathcal{B}$ là cơ sở trực chuẩn của V và $[f]_{\mathcal{B}}$ là ma trận chéo. □

§4. Toán tử unita

Định nghĩa 4.1. Cho V là không gian unita và f là toán tử tuyến tính trên V. Ta nói f là *toán tử unita* nếu f bảo toàn tích vô hướng, nghĩa là $\langle f(u), f(v)\rangle = \langle u, v\rangle$ với mọi $u, v \in V$.

Định lý 4.2. *Cho V là không gian unita hữu hạn chiều và f là toán tử tuyến tính trên. Khi đó f là toán tử unita khi và chỉ khi f bảo toàn độ dài của các vectơ, nghĩa là $||f(u)|| = ||u||$ với mọi $u \in V$. Nói riêng, mọi toán tử unita đều là đẳng cấu tuyến tính.*

Chứng minh. Chiều thuận là hiển nhiên nên ta chỉ cần chứng minh chiều đảo. Giả sử f bảo toàn độ dài của các vectơ. Khi đó với mọi $u, v \in V$, ta có $||f(u+v)||^2 = ||u+v||^2$, nên $||f(u)||^2 + 2\mathrm{Re}\langle f(u), f(v)\rangle + ||f(v)||^2 = ||u||^2 + 2\mathrm{Re}\langle u, v\rangle + ||v||^2$, nghĩa là $\mathrm{Re}\langle f(u), f(v)\rangle = \mathrm{Re}\langle u, v\rangle$.

Áp dụng kết quả trên cho các vectơ iu và v ta được $\mathrm{Re}\langle f(iu), f(v)\rangle = \mathrm{Re}\langle iu, v\rangle$, nên $\mathrm{Im}\langle f(u), f(v)\rangle = \mathrm{Im}\langle u, v\rangle$. Do đó $\langle f(u), f(v)\rangle = \langle u, v\rangle$, nghĩa là f là toán tử unita. Mặt khác, nếu $u \in \mathrm{Ker} f$ thì $||u|| = ||f(u)|| = 0$, nên $u = 0$, nghĩa là $\mathrm{Ker} f = 0$, do đó f là đơn cấu. Mà V có chiều hữu hạn nên f cũng là đẳng cấu. □

Định lý 4.3. *Toán tử tuyến tính f là toán tử unita khi và chỉ khi $f^* = f^{-1}$. Nói riêng, mọi toán tử unita đều chuẩn tắc.*

Chứng minh. Ta có f là toán tử unita $\Leftrightarrow \forall u, v \in V, \langle f(u), f(v)\rangle = \langle u, v\rangle$ $\Leftrightarrow \forall u, v \in V, \langle u, (f^* f)(v)\rangle = \langle u, v\rangle \Leftrightarrow \forall v \in V, (f^* f)(v) = v \Leftrightarrow f^* f = \mathrm{Id}_V$ $\Leftrightarrow f^* = f^{-1}$. Khẳng định còn lại là hiển nhiên. □

Mệnh đề 4.4. *Cho f và g là các toán tử unita. Khi đó f^{-1} và fg cũng là toán tử unita.*

Chứng minh. Theo Định lý 4.2 và 2.5, f^{-1} tồn tại và $(f^{-1})* = (f^*)^{-1} = (f^{-1})^{-1} = f$, nên f^{-1} cũng là toán tử unita. Cũng theo Định lý 2.5 ta có $(fg)^* = g^* f^* = g^{-1} f^{-1} = (fg)^{-1}$, nên fg cũng là toán tử unita. □

Định nghĩa 4.5. Ma trận $A \in M_n(\mathbb{C})$ được gọi là *ma trận unita* nếu $AA^* = I_n$, nghĩa là $A^* = A^{-1}$.

Từ định nghĩa trên và Định lý 2.6 ta có nhận xét sau:

Nhận xét 4.6. Tích hai ma trận unita cùng cấp cũng là ma trận unita; mọi ma trận unita đều khả nghịch và nghịch đảo của chúng cũng là ma trận unita.

Từ các Định lý 2.4 và 4.3 ta suy ra ngay kết quả sau:

Định lý 4.7. *Toán tử tuyến tính f trên không gian unita V là toán tử unita khi và chỉ khi ma trận biểu diễn của f theo một cơ sở trực chuẩn của V là ma trận unita.*

Định lý 4.8. *Toán tử tuyến tính f trên không gian unita V là toán tử unita khi và chỉ khi f biến mọi cơ sở trực chuẩn của V thành cơ sở trực chuẩn của V.*

Chứng minh. ($\Rightarrow$) Giả sử f là toán tử unita trên không gian unita V và $\mathcal{B} = (u_1, \ldots, u_n)$ là một cơ sở trực chuẩn của V. Khi đó, theo Định lý 4.2, f là một đẳng cấu nên $f(\mathcal{B}) = (f(u_1), \ldots, f(u_n))$ cũng là một cơ sở của V. Vì

f bảo toàn tích vô hướng của các vectơ nên $\langle f(u_i), f(u_j)\rangle = \langle u_i, u_j\rangle = \delta_{ij}$. Do đó $f(\mathcal{B})$ là cơ sở trực chuẩn của V.

($\Leftarrow$) Giả sử f là toán tử tuyến tính biến cơ sở trực chuẩn $\mathcal{B} = (u_1, \ldots, u_n)$ của V thành cơ sở trực chuẩn $f(\mathcal{B}) = (f(u_1), \ldots, f(u_n))$. Với mọi $u, v \in V$ ta có $[f(u)]_{f(\mathcal{B})} = [u]_{\mathcal{B}}$, $[f(v)]_{f(\mathcal{B})} = [v]_{\mathcal{B}}$ nên theo Mệnh đề 1.4 ta có $\langle f(u), f(v)\rangle = [f(u)]_{f(\mathcal{B})}^{\top}\ \overline{[f(v)]_{f(\mathcal{B})}} = [u]_{\mathcal{B}}^{\top}\ \overline{[v]_{\mathcal{B}}} = \langle u, v\rangle$. $\square$

Bổ đề 4.9. *Mọi trị riêng của toán tử unita đều có môđun bằng 1.*

Chứng minh. Giả sử f là toán tử unita và $\lambda \in \mathbb{C}$ là một trị riêng của f. Khi đó tồn tại vectơ $0 \neq u \in V$ sao cho $f(u) = \lambda u$. Suy ra $\langle u, u\rangle = \langle f(u), f(u)\rangle = \langle \lambda u, \lambda u\rangle = \lambda\overline{\lambda}\langle u, u\rangle = |\lambda|^2\langle u, u\rangle$. Mà $u \neq 0$ nên $\langle u, u\rangle > 0$, do đó $|\lambda| = 1$. $\square$

Định lý 4.10. *Cho V là không gian unita hữu hạn chiều và f là toán tử tuyến tính trên V. Khi đó f là toán tử unita khi và chỉ khi tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_{\mathcal{B}}$ là ma trận chéo, với các hệ số trên đường chéo có môđun bằng 1.*

Chứng minh. ($\Rightarrow$) Giả sử f là toán tử unita. Khi đó, theo Định lý 4.3, f là toán tử chuẩn tắc. Do đó từ Định lý 3.4 và 4.7 ta được kết luận của chiều thuận.

($\Leftarrow$) Giả sử tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_{\mathcal{B}}$ là ma trận chéo, với các hệ số trên đường chéo đều có môđun bằng 1. Chú ý rằng mọi số phức λ có môđun bằng 1 đều thỏa mãn $\overline{\lambda} = \lambda^{-1}$, nên $[f]_{\mathcal{B}}$ là ma trận unita, do đó f là toán tử unita. $\square$

Định lý 4.11. *Ma trận $A \in M_n(\mathbb{C})$ là ma trận unita khi và chỉ khi A là ma trận chuyển cơ sở từ một cơ sở trực chuẩn sang một cơ sở trực chuẩn của một không gian unita n chiều.*

Chứng minh. ($\Rightarrow$) Cho $A \in M_n(\mathbb{C})$ là ma trận unita. Do $A^*A = I_n$ nên A khả nghịch và các vectơ cột của A tạo thành một cơ sở trực chuẩn $\mathcal{B}$ của không gian unita $\mathbb{C}^n$. Khi đó A là ma trận chuyển cơ sở từ cơ sở chính tắc (trực chuẩn) của $\mathbb{C}^n$ sang cơ sở $\mathcal{B}$.

($\Leftarrow$) Giả sử $A = (a_{ij})$ là ma trận chuyển cơ sở từ cơ sở trực chuẩn $\mathcal{B}_1 = (u_1, \ldots, u_n)$ sang cơ sở trực chuẩn $\mathcal{B}_2 = (v_1, \ldots, v_n)$ của không gian unita n chiều V. Khi đó

$$[AA^*]_{ij} = \sum_{k=1}^{n} a_{ik}.\overline{a_{jk}} = \left\langle \sum_{k=1}^{n} a_{ik}u_k, \sum_{k=1}^{n} a_{jk}v_k \right\rangle = \langle u_i, v_j \rangle = \delta_{ij},$$

nghĩa là $AA^* = I_n$. □

Bây giờ, từ các Định lý 4.8 và Định lý 4.11 ta suy ra ngay kết quả sau.

Hệ quả 4.12. *Toán tử tuyến tính f trên không gian unita V là toán tử unita khi và chỉ khi ma trận biểu diễn của f theo các cặp cơ sở trực chuẩn của V là ma trận unita.*

Định lý 4.13. *Ma trận $A \in M_n(\mathbb{C})$ là ma trận unita khi và chỉ khi tồn tại ma trận unita P sao cho $P^{-1}AP$ là ma trận chéo với các hệ số trên đường chéo có môđun bằng 1.*

Chứng minh. Do mọi ma trận chéo với các hệ số trên đường chéo có môđun bằng 1 đều là ma trận unita nên chiều đảo là hiển nhiên. Ta chứng minh phần thuận. Xét toán tử tuyến tính f_A trên $\mathbb{C}^n$ xác định bởi $f_A(u) = Au$. Khi đó $[f_A]_{\mathcal{B}_0} = A$ với $\mathcal{B}_0$ là cơ sở (trực chuẩn) chính tắc của $\mathbb{C}^n$. Vì A là ma trận unita nên f_A là toán tử unita. Theo Định lý 4.10 tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của $\mathbb{C}^n$ sao cho $[f]_{\mathcal{B}}$ là ma trận chéo với các hệ số trên đường chéo có môđun bằng 1. Đặt $P = (\mathcal{B}_0 \to \mathcal{B})$. Theo Định lý 4.11, P là ma trận unita và $P^{-1}AP = [f]_{\mathcal{B}}$. □

§5. Toán tử Hermite

Định nghĩa 5.1. Toán tử tuyến tính f trên không gian unita V được gọi là *toán tử Hermite* nếu $\langle f(u), v\rangle = \langle u, f(v)\rangle$ với mọi $u, v \in V$.

Định lý 5.2. *Toán tử tuyến tính f là toán tử Hermite khi và chỉ khi $f^* = f$. Nói riêng, mọi toán tử Hermite đều chuẩn tắc.*

Chứng minh. Ta có f là toán tử Hermite khi và chỉ khi $\forall u, v \in V, \langle f(u), v\rangle = \langle u, f(v)\rangle \Leftrightarrow \forall u, v \in V, \langle u, f^*(v)\rangle = \langle u, f(v)\rangle \Leftrightarrow \forall v \in V, f^*(v) = f(v) \Leftrightarrow f^* = f$. Khẳng định còn lại là hiển nhiên. □

Mệnh đề 5.3. *Nếu f là toán tử Hermite và là đẳng cấu thì ánh xạ ngược f^{-1} cũng là toán tử Hermite.*

Chứng minh. Từ Định lý 4.2 ta được $(f^{-1})^* = (f^*)^{-1} = f^{-1}$, nên f^{-1} cũng là toán tử Hermite. □

Định nghĩa 5.4. Ma trận $A \in M_n(\mathbb{C})$ được gọi là *ma trận Hermite* nếu $A^* = A$.

Kết quả sau được suy trực tiếp từ Định lý 2.6.

Nhận xét 5.5. Nếu A là ma trận Hermite khả nghịch thì A^{-1} cũng là ma trận Hermite.

Từ các Định lý 2.4 và 5.2 ta suy ra ngay kết quả sau.

Định lý 5.6. *Toán tử tuyến tính f trên không gian unita V là toán tử Hermite khi và chỉ khi ma trận biểu diễn của f theo một cơ sở trực chuẩn của V là ma trận Hermite.*

Bổ đề 5.7. *Mọi trị riêng của toán tử Hermite đều là số thực.*

Chứng minh. Giả sử f là toán tử Hermite trên không gian unita V và $\lambda \in \mathbb{C}$ là một trị riêng của f. Khi đó tồn tại $0 \neq u \in V$ sao cho $f(u) = \lambda u$, nên $\langle f(u), u\rangle = \langle \lambda u, u\rangle = \lambda\langle u, u\rangle$ và $\langle u, f(u)\rangle = \langle u, \lambda u\rangle = \overline{\lambda}\langle u, u\rangle$. Mà $\langle f(u), u\rangle = \langle u, f(u)\rangle$ và $\langle u, u\rangle > 0$, nên $\lambda = \overline{\lambda}$, nghĩa là $\lambda \in \mathbb{R}$. □

Định lý 5.8. *Toán tử tuyến tính f trên không gian unita V là toán tử Hermite khi và chỉ khi tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_{\mathcal{B}}$ là ma trận chéo với hệ số thực.*

Chứng minh. ($\Rightarrow$) Nếu f là toán tử Hermite thì, theo Mệnh đề 5.3, f là toán tử chuẩn tắc. Do đó chiều thuận của định lý được suy ra trực tiếp từ Định lý 3.4 và Bổ đề 5.7.

($\Leftarrow$) Nếu tồn tại cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_{\mathcal{B}}$ là ma trận chéo với hệ số thực thì $[f]_{\mathcal{B}}$ là ma trận Hermite nên f là toán tử Hermite. □

Định lý 5.9. *Ma trận $A \in M_n(\mathbb{C})$ là ma trận Hermite khi và chỉ khi tồn tại một ma trận unita P sao cho $P^{-1}AP$ là ma trận chéo với hệ số thực.*

Chứng minh. Do mọi ma trận chéo với hệ số thực đều là ma trận Hermite nên chiều đảo là hiển nhiên. Ta chứng minh chiều thuận. Xét toán tử tuyến tính f_A trên $\mathbb{C}^n$ xác định bởi $f_A(u) = Au$. Khi đó $[f_A]_{\mathcal{B}_0} = A$ với $\mathcal{B}_0$ là cơ sở (trực chuẩn) chính tắc của $\mathbb{C}^n$. Vì A là ma trận Hermite nên f_A là toán tử Hermite. Theo Định lý 5.8 tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của $\mathbb{C}^n$ sao cho ma trận biểu diễn $[f]_{\mathcal{B}}$ là một ma trận chéo với hệ số thực. Bây giờ, chọn P là ma trận chuyển cơ sở từ $\mathcal{B}_0$ sang $\mathcal{B}$ thì P là ma trận unita theo Định lý 4.11 và $P^{-1}AP = [f]_{\mathcal{B}}$. Định lý được chứng minh. □

§6. Toán tử phản Hermite

Định nghĩa 6.1. Toán tử tuyến tính f trên không gian unita V được gọi là *toán tử phản Hermite* nếu $\langle f(u), v\rangle = -\langle u, f(v)\rangle$ với mọi $u, v \in V$.

Định lý 6.2. *Toán tử tuyến tính f là toán tử Hermite khi và chỉ khi $f^* = -f$. Nói riêng, mọi toán tử phản Hermite đều chuẩn tắc.*

Chứng minh. Ta có f là toán tử phản Hermite khi và chỉ khi $\forall u, v \in V$, $\langle f(u), v\rangle = -\langle u, f(v)\rangle \Leftrightarrow \forall u, v \in V$, $\langle u, f^*(v)\rangle = \langle u, -f(v)\rangle \Leftrightarrow \forall v \in V$, $f^*(v) = -f(v) \Leftrightarrow f^* = -f$. Khẳng định còn lại là hiển nhiên. □

Mệnh đề 6.3. *Nếu f là toán tử phản Hermite và là đẳng cấu thì f^{-1} cũng là toán tử phản Hermite.*

Chứng minh. Từ Định lý 2.5, ta được $(f^{-1})^* = (f^*)^{-1} = (-f)^{-1} = -f^{-1}$, nên f^{-1} cũng là toán tử phản Hermite. □

Định nghĩa 6.4. Ma trận $A \in M_n(\mathbb{C})$ được gọi là *ma trận phản Hermite* nếu $A^* = -A$.

Nhận xét sau được suy ra trực tiếp từ Định lý 2.6.

Nhận xét 6.5. Nếu A là ma trận phản Hermite khả nghịch thì A^{-1} cũng là ma trận phản Hermite.

Từ các Định lý 2.4 và 6.2 ta suy ra ngay kết quả sau.

Định lý 6.6. *Toán tử tuyến tính f trên không gian unita V là toán tử phản Hermite khi và chỉ khi ma trận biểu diễn của f theo một cơ sở trực chuẩn của V là một ma trận phản Hermite.*

Bổ đề 6.7. *Mọi trị riêng khác 0 của một toán tử phản Hermite đều thuần ảo.*

Chứng minh. Giả sử f là toán tử phản Hermite trên không gian unita V và $0 \neq \lambda \in \mathbb{C}$ là trị riêng của f. Khi đó tồn tại $0 \neq u \in V$ sao cho $f(u) = \lambda u$, nên $\langle f(u), u\rangle = \langle \lambda u, u\rangle = \lambda\langle u, u\rangle$ và $\langle u, f(u)\rangle = \langle u, \lambda u\rangle = \overline{\lambda}\langle u, u\rangle$. Mà $\langle f(u), u\rangle = -\langle u, f(u)\rangle$ và $\langle u, u\rangle > 0$ nên $\overline{\lambda} = -\lambda$, hay λ thuần ảo. □

Định lý 6.8. *Toán tử tuyến tính f trên không gian unita V là toán tử phản Hermite khi và chỉ khi tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_{\mathcal{B}}$ là ma trận chéo với các hệ số trên đường chéo hoặc bằng 0 hoặc thuần ảo.*

Chứng minh. ($\Rightarrow$) Nếu f là toán tử phản Hermite thì, theo Định lý 6.2, f là toán tử chuẩn tắc, nên chiều thuận của định lý được suy trực tiếp từ Định lý 3.4 và Bổ đề 6.7.

($\Leftarrow$) Nếu tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_{\mathcal{B}}$ là ma trận chéo với các hệ số trên đường chéo hoặc bằng 0 hoặc thuần ảo thì $[f]_{\mathcal{B}}$ là ma trận phản Hermite, nên f là toán tử phản Hermite. □

Định lý 6.9. *Ma trận $A \in M_n(\mathbb{C})$ là ma trận phản Hermite khi và chỉ khi tồn tại một ma trận unita P sao cho $P^{-1}AP$ là ma trận chéo với các hệ số trên đường chéo hoặc bằng 0 hoặc thuần ảo.*

Chứng minh. Chiều đảo là hiển nhiên, nên ta chỉ cần chứng minh chiều thuận. Xét toán tử tuyến tính f trên $\mathbb{C}^n$ xác định bởi $f(u) = Au$. Khi đó $[f]_{\mathcal{B}_0} = A$ với $\mathcal{B}_0$ là cơ sở (trực chuẩn) chính tắc của $\mathbb{C}^n$. Do A là ma trận phản Hermite nên f là toán tử phản Hermite. Theo Định lý 6.8, tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của $\mathbb{C}^n$ sao cho $[f]_{\mathcal{B}}$ là ma trận với các hệ số trên đường chéo hoặc bằng 0 hoặc thuần ảo. Đặt $P = (\mathcal{B}_0 \to \mathcal{B})$, suy ra P là ma trận unita (do Định lý 4.11) và $P^{-1}AP = [f]_{\mathcal{B}}$. □

Bài tập

Bài 1. Xác định ma trận chuyển vị liên hợp A^* của ma trận A:

a) $A = \begin{pmatrix} 3+7i & 0 \\ 2i & 4-i \end{pmatrix}$.   b) $A = \begin{pmatrix} 2-3i & 3+i \\ -1+2i & 2+i \end{pmatrix}$.

Bài 2. Xác định ma trận chuyển vị liên hợp A^* của ma trận A:

a) $A = \begin{pmatrix} i & 1 & -i \\ 1 & -i & i \\ -i & i & 1 \end{pmatrix}$.   b) $A = \begin{pmatrix} 1+2i & 2-i & -1+3i \\ 3-2i & 2+i & 4+2i \\ 3+i & 2-2i & -1+i \end{pmatrix}$.

Bài 3. Xác định A có là ma trận unita hay không bằng cách thực hiện phép nhân AA^*?

a) $A = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$.   c) $A = \frac{1}{2}\begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}$.

b) $A = \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}$.   d) $A = \frac{\sqrt{2}}{2}\begin{pmatrix} i & i \\ i & -i \end{pmatrix}$.

Bài 4. Chứng tỏ A là ma trận unita bằng cách kiểm tra các vectơ dòng của A tạo thành cơ sở trực chuẩn của $\mathbb{C}^2$ (và C^3). Xác định A^{-1}.

a) $A = \frac{1}{5}\begin{pmatrix} -4 & 3i \\ 3 & 4i \end{pmatrix}$.   b) $A = \frac{1}{2}\begin{pmatrix} 1+i & -1-i \\ \sqrt{2} & \sqrt{2} \end{pmatrix}$.

Bài 5. Ma trận nào sau đây là ma trận Hermite?

a) $A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$.   b) $A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

c) $A = \begin{pmatrix} 1 & 3-i \\ 3+i & i \end{pmatrix}$. d) $A = \begin{pmatrix} 0 & 3-2i \\ 3-2i & 4 \end{pmatrix}$.

Bài 6. Ma trận nào sau đây là ma trận Hermite?

a) $A = \begin{pmatrix} 0 & 2+i & 1 \\ 2-i & i & 0 \\ 1 & 0 & 1 \end{pmatrix}$. c) $A = \begin{pmatrix} -1 & 2 & 3 \\ 2 & 0 & -1 \\ 3 & -1 & 4 \end{pmatrix}$.

b) $A = \begin{pmatrix} 0 & i & 1 \\ 2+i & i & 0 \\ 0 & 1 & 0 \end{pmatrix}$. d) $A = \begin{pmatrix} 3 & 2-i & -3i \\ 2+i & 0 & 1-i \\ 3i & 1+i & 0 \end{pmatrix}$.

Bài 7. Tìm ma trận unita P sao cho P^*AP là ma trận chéo.

a) $A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$.

b) $A = \begin{pmatrix} 0 & 2-i \\ 2+i & 4 \end{pmatrix}$.

c) $A = \begin{pmatrix} 4 & 2+2i \\ 2-2i & 6 \end{pmatrix}$.

d) $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & -1+i \\ 0 & -1-i & 0 \end{pmatrix}$.

e) $A = \begin{pmatrix} 3 & 2-i & -3i \\ 2+i & 0 & 1-i \\ 3i & 1+i & 0 \end{pmatrix}$.

Bài 8. Chứng minh rằng mọi ma trận Hermite $A \in M_n(C)$ đều biểu diễn được dưới dạng $A = B + iC$, trong đó B là ma trận đối xứng với hệ số thực, C là ma trận phản xứng với hệ số thực. Áp dụng chứng minh trên để biểu diễn ma trận $A = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}$.

Bài 9. Chứng minh rằng mọi ma trận $A \in M_n(\mathbb{C})$ đều có thể biểu diễn được dưới dạng $A = B + iC$, trong đó B, C là các ma Hermite. Áp dụng chứng minh trên để biểu diễn ma trận $A = \begin{pmatrix} i & 2 \\ 2+i & 1-2i \end{pmatrix}$.

Bài 10. Trong không gian unita $V = \mathbb{C}^2$, xét cơ sở $\mathcal{B}$ và toán tử f xác định bởi:

$$\mathcal{B} = \{(2, 1-i), (3+4i, -2)\}, \;\; [f]_{\mathcal{B}} = \begin{pmatrix} 1+i & -3 \\ 5 & 3+2i \end{pmatrix}.$$

a) Tìm biểu thức của f^*.

b) Xét xem f có là toán tử chuẩn tắc hay không?

Bài 11. Trong không gian unita $V = \mathbb{C}^2$, cho toán tử f thỏa mãn

$$f(1-2i, 3) = (4, -2+7i), \; f(4, 2+3i) = (-1-6i, -2).$$

a) Tìm biểu thức của f^*.

b) Xét xem f có là toán tử chuẩn tắc hay không?

Bài 12. Cho V là không gian unita hữu hạn chiều và f là một toán tử Hermite trên V. Chứng minh rằng:

a) $||(\mathrm{Id}_V + if)(u)|| = ||(\mathrm{Id}_V - if)(u)||$ với mọi $u \in V$.

b) $(\mathrm{Id}_V + if)$ và $(\mathrm{Id}_V - if)$ là các song ánh.

c) $(\mathrm{Id}_V + if)(\mathrm{Id}_V - if)^{-1}$ là toán tử unita.

Bài 13. Cho V là không gian unita hữu hạn chiều và f là một toán tử tuyến tính trên V. Chứng minh rằng:

a) f là toán tử Hermite $\Leftrightarrow \langle f(u), u\rangle$ là số thực với mọi $u \in V$.

b) f xác định dương $\Leftrightarrow \langle f(u), u\rangle$ là số thực dương với mọi $u \in V$.

c) Nếu $g, h \in L(V)$ thỏa mãn f^*g là toán tử Hermite và $gg^* + hh^* = \mathrm{Id}_V$ thì $g + ih$ là toán tử unita.

d) Tồn tại duy nhất cặp toán tử g, h thỏa mãn $f = g + ih$. Hơn nữa, f chuẩn tắc khi và chỉ khi $gh = hg$.

Chương 9

DẠNG SONG TUYẾN TÍNH VÀ DẠNG TOÀN PHƯƠNG

Trong chương này ký hiệu K để chỉ trường số thực $\mathbb{R}$ hay trường số phức $\mathbb{C}$ và V để chỉ không gian vectơ hữu hạn chiều trên trường K.

§1. Dạng song tuyến tính

Định nghĩa 1.1. Cho V là một không gian vectơ trên trường K và $f : V \times V \to K$ là một ánh xạ. Ta nói f là một *dạng song tuyến tính* trên V nếu $f(u, v)$ tuyến tính theo từng biến $u, v \in V$, nghĩa là với mọi $u, u_1, u_2, v, v_1, v_2 \in V$ và với mọi $\alpha, \beta \in \mathbb{K}$ ta có:

(i) $f(\alpha u_1 + u_2, v) = \alpha f(u_1, v) + f(u_2, v)$;

(ii) $f(u, \beta v_1 + v_2) = \beta f(u, v_1) + f(u, v_2)$.

Nếu $f(u, v) = f(v, u)$ với mọi $u, v \in V$ thì ta nói f là *dạng song tuyến tính đối xứng*.

Ví dụ 30. Một tích vô hướng trên không gian Euclide V là một dạng song tuyến tính trên V.

Định nghĩa 1.2. Cho f là một dạng song tuyến tính trên V và $\mathcal{B} = (u_1, \ldots, u_n)$. Đặt $a_{ij} = f(u_i, u_j)$ và $A = (a_{ij})$. Khi đó A được gọi là *ma trận của dạng song tuyến tính* f theo cơ sở $\mathcal{B}$, ký hiệu $A = [f]_{\mathcal{B}}$.

Với mọi $u, v \in V$, đặt $[u]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ và $[v]_{\mathcal{B}} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$. Khi đó

$$u = \alpha_1 u_1 + \cdots + \alpha_n u_n \quad \text{và} \quad v = \beta_1 u_1 + \cdots + \beta_n u_n,$$

nên

$$f(u, v) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \alpha_i \beta_j,$$

nghĩa là

$$f(u, v) = [u]_{\mathcal{B}}^{\top} [f]_{\mathcal{B}} [v]_{\mathcal{B}}.$$

Đẳng thức trên được gọi là *biểu thức tọa độ của dạng song tuyến tính f theo cơ sở $\mathcal{B}$*.

Ma trận của f theo cơ sở chính tắc của K^n được gọi tắt là *ma trận của f* và biểu thức tọa độ của f theo cơ sở chính tắc được gọi là *biểu thức của f*.

Ví dụ 31. Xét dạng song tuyến tính f trên $\mathbb{R}^3$ định bởi:

$$f(u, v) = x_1 y_1 + 2x_1 y_2 - 4x_1 y_3 - x_2 y_2 + 3x_2 y_3 + x_3 y_1 + 9x_3 y_2,$$

với $u = (x_1, x_2, x_3), v = (y_1, y_2, y_3)$. Khi đó ma trận của f là

$$A = \begin{pmatrix} 1 & 2 & -4 \\ 0 & -1 & 3 \\ 1 & 9 & 0 \end{pmatrix}.$$

Nhận xét 1.3. Dạng song tuyến tính f trên không gian vectơ V là dạng song tuyến tính đối xứng khi và chỉ khi ma trận của f theo một cơ sở của V là ma trận đối xứng.

§2. Dạng toàn phương

Định nghĩa 2.1. Cho V là một không gian vectơ trên K và f là một dạng song tuyến tính đối xứng trên V. Khi đó ánh xạ $Q : V \to K$ xác định bởi $Q(u) = f(u, u)$ được gọi là *dạng toàn phương* trên V ứng với dạng song tuyến tính đối xứng f. Ta cũng nói f là *dạng cực* của dạng toàn phương Q.

Để đơn giản, ta gọi một dạng toàn phương trên không gian vectơ thực (t.ư. phức) là một *dạng toàn phương thực* (t.ư. *dạng toàn phương phức*). Một dạng toàn phương trên K^n còn được gọi là *dạng toàn phương n biến*.

Nhận xét 2.2. Nếu Q là một dạng toàn phương thì dạng cực f của Q hoàn toàn xác định thông qua công thức:

$$f(u, v) = \frac{Q(u + v) - Q(u) - Q(v)}{2}.$$

Định nghĩa 2.3. Nếu Q là một dạng toàn phương trên V ứng với dạng song tuyến tính đối xứng f và $\mathcal{B}$ là một cơ sở bất kỳ của V thì ma trận $[f]_{\mathcal{B}}$ còn được gọi là *ma trận của dạng toàn phương* Q theo cơ sở $\mathcal{B}$, ký hiệu bởi $[Q]_{\mathcal{B}}$.

Do f là dạng song tuyến tính đối xứng nên ma trận của dạng toàn phương Q theo một cơ sở bất kỳ luôn luôn là một ma trận đối xứng. Hơn nữa, với $\mathcal{B} = (u_1, \ldots, u_n)$ là một cơ sở của V và $u = \alpha_1 u_1 + \cdots + \alpha_n u_n$ thuộc V ta có

$$Q(u) = [u]_{\mathcal{B}}^{\top} [Q]_{\mathcal{B}} [u]_{\mathcal{B}} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \text{ với } a_{ij} = a_{ji}. \tag{1}$$

Đảo lại, (1) xác định dạng toàn phương Q trên V có ma trận theo cơ sở $\mathcal{B}$

là $A = (a_{ij})_{n\times n}$. Ta cũng thường viết (1) dưới dạng

$$Q(u) = [u]_{\mathcal{B}}^{\top}\,[Q]_{\mathcal{B}}\,[u]_{\mathcal{B}} = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{1\leq i<j\leq n} 2a_{ij}x_ix_j \tag{$1'$}$$

Ta gọi (1) và ($1'$) là *biểu thức tọa độ của dạng toàn phương* Q theo cơ sở $\mathcal{B}$. Đặc biệt, ma trận của dạng toàn phương Q theo cơ sở chính tắc của K^n cũng được gọi là ma trận của Q và biểu thức tọa độ của Q theo cơ sở chính tắc cũng được gọi là biểu thức của Q.

Ví dụ 32. Xét dạng toàn phương Q trên $\mathbb{R}^3$ có biểu thức:

$$Q(x_1, x_2, x_3) = x_1^2 - 3x_2^2 + 2x_1x_2 - 4x_1x_3 + 8x_2x_3.$$

Ta có ma trận của Q là

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

và $Q(u) = u.A.u^{\top}$, với $u = (x_1, x_2, x_3)$.

Ví dụ 33. Giả sử Q là dạng toàn phương trên $\mathbb{R}^3$ có ma trận là

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

Khi đó biểu thức của Q là

$$Q(u) = u.A.u^{\top} = 2x_1^2 - 4x_3^2 + 6x_1x_2 - 2x_1x_3 + 10x_2x_3,$$

trong đó $u = (x_1, x_2, x_3)$.

Định lý 2.4. *Nếu $\mathcal{B}, \mathcal{B}'$ là hai cơ sở của V và f là một dạng song tuyến tính trên V thì*

$$[f]_{\mathcal{B}'} = (\mathcal{B} \to \mathcal{B}')^\top [f]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}').$$

Chứng minh. Giả sử $\mathcal{B} = (u_1, \ldots, u_n), \mathcal{B}' = (v_1, \ldots, v_n)$. Khi đó với mọi u, v thuộc V ta có

$$\begin{aligned} f(u,v) = [u]_{\mathcal{B}}^\top [f]_{\mathcal{B}} [v]_{\mathcal{B}} &= ((\mathcal{B} \to \mathcal{B}')[u]_{\mathcal{B}'})^\top [f]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}')[v]_{\mathcal{B}'} \\ &= [u]_{\mathcal{B}'}^\top (\mathcal{B} \to \mathcal{B}')^\top [f]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}')[v]_{\mathcal{B}'}. \end{aligned}$$

Do đó $[f]_{\mathcal{B}'} = (\mathcal{B} \to \mathcal{B}')^\top [f]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}')$. □

Hệ quả 2.5. *Cho Q là một dạng toàn phương trên V và $\mathcal{B}, \mathcal{B}'$ là hai cơ sở của V. Khi đó*

$$[Q]_{\mathcal{B}'} = (\mathcal{B} \to \mathcal{B}')^\top [Q]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}').$$

Nhận xét 2.6. Cho Q là một dạng toàn phương trên V; $\mathcal{B}, \mathcal{B}'$ là hai cơ sở của V và $P = (B \to \mathcal{B}')$ là ma trận chuyển cơ sở từ $\mathcal{B}$ sang B'. Khi đó, nếu biểu thức tọa độ của dạng toàn phương Q theo cơ sở $\mathcal{B}$ là

$$Q(u) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \le i < j \le n} 2a_{ij} x_i x_j, \text{với } u = x_1 u_1 + \cdots + x_n u_n \qquad (*)$$

và biểu thức tọa độ của dạng toàn phương Q theo cơ sở $\mathcal{B}'$ là

$$Q(u) = \sum_{i=1}^{n} b_{ii} y_i^2 + \sum_{1 \le i < j \le n} 2b_{ij} y_i y_j \text{với } u = y_1 v_1 + \cdots + y_n v_n \qquad (**)$$

thì phép biến đổi tọa độ không suy biến

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

đưa biểu thức của dạng toàn phương Q từ $(*)$ về $(**)$.

Ngược lại, với mỗi phép biến đổi tọa độ không suy biến

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \text{ với } P \text{ khả nghịch,}$$

thì biểu thức tọa độ của Q theo cơ sở $\mathcal{B}'$ thỏa mãn $(\mathcal{B} \to \mathcal{B}') = P$ có dạng $(**)$, trong đó các b_{ij} được xác định bởi công thức

$$B = P^\top AP, \quad B = (b_{ij}), A = (a_{ij}) \in M_n(K).$$

Định nghĩa 2.7. Cho Q là một dạng toàn phương trên không gian vectơ n chiều V và $\mathcal{B}$ là một cơ sở của V. Khi đó, hạng của ma trận $[Q]_\mathcal{B}$ được gọi là *hạng* của Q, ký hiệu $\operatorname{rank}(Q)$ (hay $r(Q)$).

Từ Hệ quả 2.5 cho thấy hạng của Q không phụ thuộc vào cách chọn cơ sở $\mathcal{B}$. Hiển nhiên $\operatorname{rank}(Q) \le n = \dim V$. Nếu $\operatorname{rank}(Q) = n$ thì ta nói Q *không suy biến*. Ngược lại, nếu $\operatorname{rank}(Q) < n$ thì Q *suy biến*.

Ví dụ 34. Xét Q là dạng toàn phương 3 biến thực xác định bởi:

$$Q(x_1, x_2, x_3) = x_1^2 - 3x_2^2 + 2x_1x_2 - 4x_1x_3 + 8x_2x_3.$$

a) Tìm hạng và khảo sát tính không suy biến của Q.

b) Tìm biểu thức toạ độ của Q theo cơ sở $\mathcal{B} = (u_1, u_2, u_3)$ của $\mathbb{R}^3$, trong đó $u_1 = (1, -1, 0); u_2 = (-1, 2, 1); u_3 = (2, 0, 3)$ và chỉ ra phép biến đổi toạ độ không suy biến tương ứng.

Giải. Ma trận của Q là

$$A := [Q]_{\mathcal{B}_0} = \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 4 \\ -2 & 4 & 0 \end{pmatrix}.$$

a) Ta có det$A = -20$ nên $r(Q) = r(A) = 3$. Do đó Q không suy biến.

b) Ma trận chuyển từ cơ sở chính tắc $\mathcal{B}_0$ sang cơ sở $\mathcal{B}$ là

$$(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}.$$

Do đó ma trận của Q theo cơ sở $\mathcal{B}$ là

$$\begin{aligned}
[Q]_{\mathcal{B}} &= (\mathcal{B}_0 \to \mathcal{B})^\top [Q]_{\mathcal{B}_0} (\mathcal{B}_0 \to \mathcal{B}) \\
&= \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 4 \\ -2 & 4 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} \\
&= \begin{pmatrix} -4 & 2 & -18 \\ 2 & 5 & 28 \\ -18 & 28 & -20 \end{pmatrix}.
\end{aligned}$$

Vậy biểu thức tọa độ của Q theo cơ sở $\mathcal{B}$ xác định bởi:

$$\begin{aligned}
Q(u) &= [u]_{\mathcal{B}}^\top \, [Q]_{\mathcal{B}} \, [u]_{\mathcal{B}} \\
&= (y_1,\ y_2,\ y_3) \begin{pmatrix} -4 & 2 & -18 \\ 2 & 5 & 28 \\ -18 & 28 & -20 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\
&= -4y_1^2 + 5y_2^2 - 20y_3^2 + 4y_1y_2 - 36y_1y_3 + 56y_2y_3,
\end{aligned}$$

với $u = y_1u_1 + y_2u_2 + y_3u_3 \in \mathbb{R}^3$.

Phép biến đổi tọa độ không suy biến tương ứng là

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (\mathcal{B}_0 \to \mathcal{B}) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

hay $\begin{cases} x_1 &= y_1 - y_2 + 2y_3 \\ x_2 &= -y_1 + 2y_2 \\ x_3 &= y_2 + 3y_3. \end{cases}$ $\square$

Ví dụ 35. Trong không gian $\mathbb{R}^3$, xét cơ sở $\mathcal{B} = (u_1, u_2, u_3)$, với $u_1 = (1,-1,0), u_2 = (-1,2,1), u_3 = (2,0,3)$ và cho Q là dạng toàn phương có biểu thức tọa độ theo cơ sở $\mathcal{B}$ như sau:

$$Q(u) = x_1^2 - 3x_2^2 + 2x_1x_2 - 4x_1x_3 + 8x_2x_3, \quad \text{với } [u]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

a) Tìm hạng của Q và khảo sát tính không suy biến của Q.

b) Tìm biểu thức của Q và chỉ ra phép biến đổi tọa độ không suy biến tương ứng.

Giải. Ma trận của Q theo cơ sở $\mathcal{B}$ là $A := [Q]_{\mathcal{B}} = \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 4 \\ -2 & 4 & 0 \end{pmatrix}$.

a) Ta có $\det A = -20$ nên $r(Q) = r(A) = 3$. Do đó Q không suy biến.

b) Ta có $(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ nên

$$(\mathcal{B} \to \mathcal{B}_0) = (\mathcal{B}_0 \to \mathcal{B})^{-1} = \begin{pmatrix} 6 & 5 & -4 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}.$$

Do đó ma trận của Q (theo cơ sở chính tắc $\mathcal{B}_0$) là

$$\begin{aligned}
[Q]_{\mathcal{B}_0} &= (\mathcal{B} \to \mathcal{B}_0)^\top [Q]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}_0) \\
&= \begin{pmatrix} 6 & 3 & -1 \\ 5 & 3 & -1 \\ -4 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 4 \\ -2 & 4 & 0 \end{pmatrix} \begin{pmatrix} 6 & 5 & -4 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 45 & 34 & -30 \\ 34 & 24 & -22 \\ -30 & -22 & 20 \end{pmatrix}
\end{aligned}$$

Suy ra biểu thức của Q theo (cơ sở $\mathcal{B}_0$) là

$$Q(y_1, y_2, y_3) = 45y_1^2 + 24y_2^2 + 20y_3^2 + 68y_1y_2 - 60y_1y_3 - 44y_2y_3.$$

Phép biến đổi tọa độ không suy biến tương ứng là

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (\mathcal{B} \to \mathcal{B}_0) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 6 & 5 & -4 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3, \end{pmatrix}$$

nghĩa là $\begin{cases} x_1 = 6y_1 + 5y_2 - 4y_3 \\ x_2 = 3y_1 + 3y_2 - 2y_3 \\ x_3 = -y_1 - y_2 + y_3. \end{cases}$ □

§3. Dạng chính tắc của dạng toàn phương

Định nghĩa 3.1. Cho Q là một dạng toàn phương trên V và $\mathcal{B} = (u_1, \ldots, u_n)$ là một cơ sở của V. Khi đó, nếu $[Q]_{\mathcal{B}}$ là ma trận đường chéo, nghĩa là biểu thức tọa độ của Q theo cơ sở $\mathcal{B}$ có dạng

$$Q(u) = \sum_{i=1}^{n} a_{ii}x_i^2, \text{ với } u = x_1u_1 + \cdots + x_nu_n \quad (1)$$

thì biểu thức này được gọi là *dạng chính tắc* của Q và $\mathcal{B}$ được gọi là *cơ sở Q-chính tắc*.

Định lý 3.2. *Cho Q là một dạng toàn phương trên V. Khi đó trong V tồn tại một cơ sở Q-chính tắc.*

Chứng minh. Nếu $Q = 0$ thì Q có dạng chính tắc trong mọi cơ sở của V. Nếu $Q \neq 0$ thì việc xây dựng một cơ sở Q-chính tắc được thực hiện thông qua thuật toán sau, gọi là ***Thuật toán Lagrange***:

Giả sử biểu thức của dạng toàn phương Q theo cơ sở $\mathcal{B} = (u_1, \ldots, u_n)$ định bởi

$$Q(u) = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{1\le i<j\le n} 2a_{ij}x_ix_j \qquad (*)$$

Để đưa Q về dạng chính tắc, ta áp dụng quy nạp để chuyển về trường hợp đưa dạng toàn phương $n-1$ biến về dạng chính tắc. Bài toán được chia thành 2 trường hợp như sau:

Trường hợp 1. Tồn tại i sao cho $a_{ii} \neq 0$

Sau khi đánh số lại các phần tử của cơ sở $\mathcal{B}$ nếu cần, ta có thể giả sử $a_{11} \neq 0$. Khi đó

$$\begin{aligned}
Q(u) &= a_{11}\left(x_1^2 + 2x_1\sum_{i=2}^{n}\frac{a_{1i}}{a_{11}}x_i\right) + (\text{những số hạng không chứa } x_1) \\
&= a_{11}\left(x_1 + \sum_{i=2}^{n}\frac{a_{1i}}{a_{11}}x_i\right)^2 + (\text{một dạng toàn phương của } x_2, \ldots, x_n) \\
&= a_{11}y_1^2 + \sum_{i,j=2}^{n} b_{ij}y_iy_j,
\end{aligned}$$

với phép biến đổi tọa độ không suy biến tương ứng là:

$$\begin{cases} y_1 &= x_1 + \displaystyle\sum_{i=2}^{n}\frac{a_{1i}}{a_{11}}x_i \\ y_j &= x_j, \ \forall j \geq 2. \end{cases}$$

Khi đó Q được đưa về dạng toàn phương $n-1$ biến $Q_1 = \sum\limits_{i,j=2}^{n} b_{ij}y_iy_j$.

Trường hợp 2. $a_{ii} = 0$ với mọi i và tồn tại $j \neq i$ sao cho $a_{ij} \neq 0$

Sau khi đánh số lại các phần tử của cơ sở $\mathcal{B}$ nếu cần, ta có thể giả sử $a_{12} \neq 0$. Thực hiện phép biến đổi tọa độ không suy biến

$$\begin{cases} x_1 &= y_1 + y_2 \\ x_2 &= y_1 - y_2 \\ x_j &= y_j,\ \forall j \geq 3. \end{cases}$$

Khi đó $2a_{12}x_1x_2 = 2a_{12}(y_1^2 - y_2^2)$, nên

$$Q(u) = \sum_{i,j=1}^{n} a_{ij}x_ix_j = \sum_{i,j=1}^{n} b_{ij}y_iy_j.$$

Ta có hệ số của y_1^2 là $2a_{12} \neq 0$ nên bài toán được đưa về Trường hợp 1 đã xét.

$\square$

Ví dụ 36. Đưa dạng toàn phương sau về dạng chính tắc và chỉ ra cơ sở Q-chính tắc cùng phép biến đổi tọa độ không suy biến tương ứng:

$$Q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 - 12x_2x_3.$$

Giải. Đặt $u = (x_1, x_2, x_3)$, ta có

$$\begin{aligned} Q(u) &= x_1^2 + 2x_1(2x_2 - x_3) + 2x_2^2 - 3x_3^2 - 12x_2x_3 \\ &= (x_1 + 2x_2 - x_3)^2 - (2x_2 - x_3)^2 + 2x_2^2 - 3x_3^2 - 12x_2x_3 \\ &= (x_1 + 2x_2 - x_3)^2 - 2x_2^2 - 8x_2x_3 - 4x_3^2 \\ &= (x_1 + 2x_2 - x_3)^2 - 2(x_2^2 + 4x_2x_3 + 4x_3^2) + 4x_3^2 \\ &= (x_1 + 2x_2 - x_3)^2 - 2(x_2 + 2x_3)^2 + 4x_3^2 \end{aligned}$$

Thực hiện phép biến đổi tọa độ không suy biến

$$\begin{cases} y_1 = x_1 + 2x_2 - x_3 \\ y_2 = x_2 + 2x_3 \\ y_3 = x_3, \end{cases}$$

ta được

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 5 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Như vậy ta đã đưa được Q về dạng chính tắc

$$Q(u) = y_1^2 - 2y_2^2 + 4y_3^2, \quad \text{với } [u]_{\mathcal{B}} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

trong đó cơ sở Q-chính tắc $\mathcal{B} = (u_1, u_2, u_3)$ được xác định bởi:

$$(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} 1 & -2 & 5 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix},$$

nghĩa là $u_1 = (1,0,0); u_2 = (-2,1,0); u_3 = (5,-2,1)$. □

Ví dụ 37. Đưa dạng toàn phương thực sau đây về dạng chính tắc; Chỉ ra cơ sở Q-chính tắc và phép biến đổi tọa độ không suy biến tương ứng:

$$Q(x_1, x_2, x_3) = x_1x_2 + 2x_1x_3 - 2x_2x_3.$$

Giải. Đặt $x_1 = y_1 + y_2; x_2 = y_1 - y_2; x_3 = y_3$, ta được

$$\begin{aligned} Q(u) &= y_1^2 - y_2^2 + 2(y_1 + y_2)y_3 - 2(y_1 - y_2)y_3 \\ &= y_1^2 - y_2^2 + 4y_2y_3 \\ &= y_1^2 - (y_2 - 2y_3)^2 + 4y_3^2. \end{aligned}$$

Đặt

$$\begin{cases} z_1 = y_1 \\ z_2 = y_2 - 2y_3 \\ z_3 = y_3 \end{cases} \Leftrightarrow \begin{cases} y_1 &= z_1 \\ y_2 &= z_2 + 2z_3 \\ y_3 &= z_3 \end{cases}$$

ta đưa được Q về dạng chính tắc

$$Q(u) = z_1^2 - z_2^2 + 4z_3^2.$$

Phép biến đổi tọa độ không suy biến tương ứng là

$$\begin{cases} x_1 &= z_1 + z_2 + 2z_3 \\ x_2 &= z_1 - z_2 - 2z_3 \\ x_3 &= z_3 \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

Cơ sở Q-chính tắc tương ứng là $\mathcal{B} = (u_1, u_2, u_3)$ của $\mathbb{R}^3$ thỏa mãn:

$$(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

nghĩa là $u_1 = (1, 1, 0); u_2 = (1, -1, 0); u_3 = (2, -2, 1)$. □

§4. Dạng chính tắc trực giao của dạng toàn phương

Định nghĩa 4.1. Cho V là không gian Euclid và Q là một dạng toàn phương trên V. Cơ sở $\mathcal{B}$ của V được gọi là một *cơ sở Q-chính tắc trực giao* nếu $\mathcal{B}$ là cơ sở trực chuẩn và là cơ sở Q-chính tắc của V. Khi đó biểu thức tọa độ của Q theo cơ sở $\mathcal{B}$ được gọi là *dạng chính tắc trực giao* của Q.

Định lý 4.2. *Cho V là không gian Euclid và Q là một dạng toàn phương trên V. Khi đó cơ sở Q-chính tắc trực giao của V luôn tồn tại.*

Chứng minh. Xét $\mathcal{B}_0$ là một cơ sở trực chuẩn của V. Khi đó $[Q]_{\mathcal{B}_0}$ là ma trận đối xứng thực nên chéo hóa trực giao được, nghĩa là tồn tại ma trận trực giao P sao cho $P^{-1}[Q]_{\mathcal{B}_0}P$ là ma trận đường chéo. Gọi $\mathcal{B}$ là cơ sở của V sao cho $(\mathcal{B}_0 \to \mathcal{B}) = P$. Khi đó

$$[Q]_{\mathcal{B}} = (\mathcal{B}_0 \to \mathcal{B})^{\top} [Q]_{\mathcal{B}_0} (\mathcal{B}_0 \to \mathcal{B}) = P^{\top} [Q]_{\mathcal{B}_0} P = P^{-1} [Q]_{\mathcal{B}_0} P$$

nên $[Q]_{\mathcal{B}}$ là ma trận đường chéo, do đó $\mathcal{B}$ là cơ sở Q-chính tắc. Mặt khác, do P là ma trận trực giao nên $\mathcal{B}$ là cơ sở trực chuẩn của V. Vậy $\mathcal{B}$ là cơ sở Q-chính tắc trực giao của V. □

Từ chứng minh trên ta thấy, để đưa Q về dạng chính tắc trực giao ta dùng phép biến đổi tọa độ $[u]_{\mathcal{B}_0} = (\mathcal{B}_0 \to \mathcal{B})[u]_{\mathcal{B}}$. Do $(\mathcal{B}_0 \to \mathcal{B})$ là ma trận trực giao nên ta gọi phép biến đổi trên là *phép biến đổi tọa độ trực giao*.

Nhận xét 4.3. Giả sử Q có dạng chính tắc trực giao là

$$Q(u) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2, \tag{1}$$

với $\mathcal{B} = (u_1, \dots, u_n)$ là cơ sở Q-chính tắc trực giao tương ứng và $u = x_1u_1 + \cdots + x_nu_n$. Khi đó $\lambda_1, \dots, \lambda_n$ gồm tất cả các trị riêng của $[Q]_{\mathcal{B}}$ (kể cả số bội) và không phụ thuộc vào việc chọn cơ sở $\mathcal{B}$.

Thật vậy, từ (1) ta suy ra $[Q]_{\mathcal{B}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, nên hiển nhiên $\lambda_1, \dots, \lambda_n$ là tất cả các trị riêng của $[Q]_{\mathcal{B}}$. Bây giờ cho $\mathcal{B}' = (u'_1, \dots, u'_n)$ là một cơ sở Q-chính tắc trực giao khác của V. Khi đó, với $u = y_1u'_1 + \cdots + y_nu'_n$ ta có $Q(u)$ có dạng $Q(u) = \lambda'_1 y_1^2 + \cdots + \lambda'_n y_n^2$. Theo chứng minh trên, $\lambda'_1, \dots, \lambda'_n$ là các trị riêng của $[Q]_{\mathcal{B}'}$. Theo Hệ quả 2.5, $[Q]_{\mathcal{B}'} = (\mathcal{B} \to \mathcal{B}')^{\top} [Q]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}')$. Do $\mathcal{B}, \mathcal{B}'$ là các cơ sở trực chuẩn của V nên $(\mathcal{B} \to \mathcal{B}')$ là ma trận trực giao, nghĩa là $(\mathcal{B} \to \mathcal{B}')^{\top} = (\mathcal{B} \to \mathcal{B}')^{-1}$. Do đó $[Q]_{\mathcal{B}'} = (\mathcal{B} \to \mathcal{B}')^{-1} [Q]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}')$. Vậy $[Q]_{\mathcal{B}}$ đồng dạng với $[Q]'_{\mathcal{B}}$, nghĩa là hai dãy $\lambda_1, \dots, \lambda_n$ và $\lambda'_1, \dots, \lambda'_n$ trùng nhau.

Cho V là không gian Euclid n chiều và Q là một dạng toàn phương trên V. Khi đó, để đưa Q về dạng chính tắc trực giao và chỉ ra cơ sở Q-chính tắc trực giao cũng như phép biến đổi tọa độ trực giao tương ứng, ta thực hiện thông qua thuật toán sau, gọi là ***Thuật toán chính tắc trực giao***.

Bước 1. Tìm một cơ sở trực chuẩn $\mathcal{B}_0$ của V và xác định $A := [Q]_{\mathcal{B}_0}$.

Bước 2. Chéo hóa trực giao ma trận A và tìm ma trận trực giao P sao cho

$$P^{-1}AP = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Bước 3. Tìm cơ sở $\mathcal{B} = (u_1, \ldots, u_n)$ thỏa mãn $P = (\mathcal{B}_0 \to \mathcal{B})$. Khi đó $\mathcal{B}$ là cơ sở Q-chính tắc trực giao và dạng chính tắc trực giao của Q là

$$Q(u) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2, \quad \text{với } u = x_1u_1 + \cdots + x_nu_n.$$

Phép biến đổi tọa độ trực giao tương ứng xác định bởi

$$[u]_{\mathcal{B}_0} = (\mathcal{B}_0 \to \mathcal{B})[u]_{\mathcal{B}}.$$

Ví dụ 38. Đưa dạng toàn phương sau về dạng chính tắc trực giao; chỉ ra cơ sở Q-chính tắc trực giao và phép biến đổi tọa độ trực giao tương ứng:

$$Q(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

Giải. Gọi $\mathcal{B}_0$ là cơ sở chính tắc của $\mathbb{R}^3$. Suy ra

$$A := [Q]_{\mathcal{B}_0} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Đa thức đặc trưng của A là $p(t) = (1+t)^2(2-t)$. Do đó A có 2 trị riêng là $\lambda_1 = -1$ (kép) và $\lambda_2 = 2$ (đơn).

- Ta có $A - \lambda_1 I_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$

nên hệ $(A - \lambda_1 I_3)X = 0$ có nghiệm xác định bởi

$$(x_1, x_2, x_3) = (-\alpha - \beta, \alpha, \beta), \quad \alpha, \beta \in \mathbb{R}.$$

Do đó $\mathcal{B}_1 = (u_1, u_2)$, với $u_1 = (-1, 1, 0); u_2 = (-1, 0, 1)$, là cơ sở của $E(\lambda_1)$ và $\dim E(\lambda_1) = 2$.

• Ta có $A - \lambda_2 I_3 = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$,
nên hệ $(A - \lambda_2 I_3)X = 0$ có nghiệm xác định bởi

$$(x_1, x_2, x_3) = (\alpha, \alpha, \alpha), \quad \alpha \in \mathbb{R}.$$

Do đó $\mathcal{B}_2 = (u_3)$, với $u_3 = (1, 1, 1)$, là cơ sở của $E(\lambda_2)$, và $\dim E(\lambda_2) = 1$.

• Bằng quá trình trực giao hóa Gram-Schmidt, ta được cơ sở trực giao của $E(\lambda_1)$ là (v_1, v_2), với

$$v_1 = u_1; \;\; v_2 = u_2 - \frac{\langle u_1, v_1 \rangle}{\langle v_1, v_1 \rangle}.v_1 = \Big(\frac{1}{2}, -\frac{1}{2}, 1\Big).$$

Chuẩn hóa (v_1, v_2) ta được cơ sở trực chuẩn $\mathcal{B}'_1 = (w_1, w_2)$ của $E(\lambda_1)$, với

$$w_1 := \frac{v_1}{||v_1||} = \Big(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\Big); \quad w_2 := \frac{v_2}{||v_2||} = \Big(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\Big).$$

• Chuẩn hóa (u_3) ta được cơ sở trực chuẩn $\mathcal{B}'_2 = (w_3)$ của $E(\lambda_2)$, với

$$w_3 = \frac{u_3}{||u_3||} = \Big(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big).$$

Đặt $\mathcal{B} = \mathcal{B}'_1 \cup \mathcal{B}'_2 = (w_1, w_2, w_3)$. Ta có $\mathcal{B}$ là cơ sở trực chuẩn của $\mathbb{R}^3$ và

$$(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Như vậy dạng chính tắc trực giao của Q là

$$Q(u) = \lambda_1 y_1^2 + \lambda_1 y_2^2 + \lambda_2 y_3^2 = -y_1^2 - y_2^2 + 2y_3^2, \quad \text{với } [u]_{\mathcal{B}} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Cơ sở chính tắc trực giao tương ứng là $\mathcal{B} = (w_1, w_2, w_3)$; phép biến đổi tọa độ trực giao tương ứng là

$$[u]_{\mathcal{B}_0} = (\mathcal{B}_0 \to \mathcal{B})[u]_{\mathcal{B}} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

nghĩa là
$$\begin{cases} x_1 = -\frac{1}{\sqrt{2}} y_1 - \frac{1}{\sqrt{6}} y_2 + \frac{1}{\sqrt{3}} y_3 \\ x_2 = \frac{1}{\sqrt{2}} y_1 - \frac{1}{\sqrt{6}} y_2 + \frac{1}{\sqrt{3}} y_3 \\ x_3 = \frac{2}{\sqrt{6}} y_2 + \frac{1}{\sqrt{3}} y_3 \end{cases}$$

§5. Chỉ số quán tính, dạng toàn phương xác định dấu

Định nghĩa 5.1. Cho V là không gian vectơ n chiều trên $\mathbb{R}$ và $\mathcal{B} = (u_1, \dots, u_n)$ là một cơ sở của V. Giả sử Q là một dạng toàn phương trên V có biểu thức tọa độ theo cơ sở $\mathcal{B}$ có dạng

$$Q(u) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2, \text{ với } u = x_1 u_1 + \cdots + x_n u_n, \quad (1)$$

trong đó r, s là các số nguyên thỏa mãn $0 \le s \le r \le n$. Khi đó ta nói $\mathcal{B}$ là *cơ sở Q-chuẩn tắc* và (1) là dạng chuẩn tắc của Q.

Định lý 5.2. *Cho V là không gian vectơ n chiều trên $\mathbb{R}$ và Q là một dạng toàn phương trên V. Khi đó luôn tồn tại cơ sở Q-chuẩn tắc của V.*

Chứng minh. Theo Định lý 3.2 tồn tại $\mathcal{B}$ là cơ sở Q-chính tắc của V sao cho biểu thức tọa độ của Q theo cơ sở $\mathcal{B}$ có dạng

$$Q(u) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2,$$

Bằng cách đánh số lại nếu cần ta có thể giả sử $\lambda_1, \ldots, \lambda_s > 0; \lambda_{s+1}, \ldots, \lambda_r < 0; \lambda_{r+1} = \ldots = \lambda_n = 0$. Dùng phép biến đổi toạ độ không suy biến

$$y_j := \begin{cases} \sqrt{\lambda_j} x_j, & \text{nếu } \leq j \leq s \\ \sqrt{-\lambda_j} x_j, & \text{nếu } s+1 \leq j \leq r \\ x_j, & \text{nếu } r+1 \leq j \leq n \end{cases}$$

ta được dạng chuẩn tắc của Q là:

$$Q(u) = y_1^2 + \cdots + y_s^2 - y_{s+1}^2 - \cdots - y_r^2.$$

Cơ sở Q-chuẩn tắc cần tìm là cơ sở tương ứng với phép biến đổi tọa độ trên. □

Định lý 5.3. *Cho Q là một dạng toàn phương trên V và $\mathcal{B}$ là cơ sở Q-chuẩn tắc của V. Khi đó biểu thức tọa độ của Q theo cơ sở $\mathcal{B}$ có dạng*

$$Q(u) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2.$$

trong đó s và r không phụ thuộc vào cách chọn cơ sở $\mathcal{B}$.

Chứng minh. Ta có $r = \operatorname{rank}(Q)$ nên r không phụ thuộc vào $\mathcal{B}$. Giả sử $\dim V = n$ và $\mathcal{B}_1 = (u_1, \ldots, u_n), \mathcal{B}_2 = (v_1, \ldots, v_n)$ là hai cơ sở Q-chuẩn tắc của V sao cho biểu thức tọa độ của Q trong $\mathcal{B}_1, \mathcal{B}_2$ lần lượt là

$$Q(u) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 \tag{1}$$

$$Q(u) = y_1^2 + \cdots + y_t^2 - y_{t+1}^2 - \cdots - y_r^2 \tag{2}$$

Đặt $V_1 = \langle u_1, \dots, u_s \rangle$ và $V_2 = \langle v_{t+1}, \dots, v_n \rangle$. Khi đó, nếu $u \in V_1 \cap V_2$ thì tồn tại các α_i, β_j sao cho $u = \alpha_1 u_1 + \cdots + \alpha_s u_s$ và $u = \beta_{t+1} v_{t+1} + \cdots + \beta_n v_n$. Suy ra

$$Q(u) = \alpha_1^2 + \cdots + \alpha_s^2 \geq 0 \;\text{ và }\; Q(u) = -\beta_{t+1}^2 - \cdots - \beta_r^2 \leq 0.$$

Do đó $\alpha_1 = \cdots = \alpha_s = 0$, hay $u = 0$, nghĩa là $V_1 \cap V_2 = \{0\}$. Do đó

$$n \geq \dim(V_1 + V_2) = \dim V_1 + \dim V_2 = s + (n - t).$$

Suy ra $s \leq t$. Tương tự ta cũng có $t \leq s$. Vậy $s = t$. □

Dựa vào định lý trên ta thấy hệ số s và r không phụ thuộc vào cách chọn cơ sở của V, nên ta có định nghĩa sau:

Định nghĩa 5.4. Ta gọi

- s là *chỉ số dương quán tính* của Q;
- $r - s$ là *chỉ số âm quán tính* của Q;
- $(s, r - s)$ là *cặp chỉ số quán tính* của Q;
- $s - (r - s) = 2s - r$ là *ký số* của Q.

Ví dụ 39. Xét lại Ví dụ trong 36, ta thấy dạng toàn phương Q có dạng chính tắc là

$$Q(u) = y_1^2 - 2y_2^2 + 4y_3^2.$$

Do đó Q có chỉ số dương quán tính là 2; chỉ số âm quán tính là 1; cặp chỉ số quán tính là $(2, 1)$ và ký số là 1.

Định nghĩa 5.5. Cho V là không gian vectơ n chiều trên $\mathbb{R}$ và Q là một dạng toàn phương trên V. Ta nói:

i) *Q xác định dương* nếu $Q(u) > 0$ với mọi $0 \neq u \in V$.

ii) *Q xác định âm* nếu $Q(u) < 0$ với mọi $0 \neq u \in V$.

Mệnh đề 5.6. *Cho V là không gian vectơ n chiều trên $\mathbb{R}$ và Q là một dạng toàn phương trên V. Khi đó:*

(i) Q xác định dương $\Leftrightarrow$ dạng cực của Q là một tích vô hướng trên V.

(ii) Q xác định dương $\Leftrightarrow$ Q có chỉ số dương quán tính bằng n.

(iii) Q xác định âm $\Leftrightarrow$ Q có chỉ số âm quán tính bằng n.

Chứng minh. i) Dễ dàng chứng minh.

ii) *Chiều đảo.* Nếu Q có chỉ số dương quán tính bằng n. Khi đó tồn tại cơ sở Q-chuẩn tắc $\mathcal{B}$ của V sao cho biểu thức tọa độ của Q theo cơ sở $\mathcal{B}$ là:

$$Q(u) = x_1^2 + \cdots + x_n^2, \quad \text{với } u = x_1u_1 + \cdots + x_nu_n.$$

Nếu $u \neq 0$ thì tồn tại i sao cho $x_i \neq 0$, nên $Q(u) > 0$. Do đó Q xác định dương.

Chiều thuận. Giả sử Q xác định dương nhưng chỉ số dương quán tính của Q khác n. Gọi $\mathcal{B} = (u_1, \ldots, u_n)$ là một cơ sở Q-chính tắc của V. Khi đó biểu thức tọa độ của Q trong $\mathcal{B}$ có dạng

$$Q(u) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2,$$

trong đó tồn tại i sao cho $\lambda_i \leq 0$. Lấy $u = u_i$, ta được $u \neq 0$ và $Q(u) = a_i \leq 0$, mâu thuẫn với tính xác định dương của Q.

iii) Q xác định âm $\Leftrightarrow -Q$ xác định dương nên iii) được suy ra từ ii). □

Hệ quả 5.7. *Mọi dạng toàn phương xác định dương hay xác định âm đều không suy biến.*

§6. Tiêu chuẩn Sylvester và Thuật toán Jacobi

Cho $A = (a_{ij}) \in M_n(\mathbb{R})$. Định thức của ma trận sinh bởi các dòng $1, \ldots, k$ và các cột $1, \ldots, k$ của A được gọi là *định thức con chính cấp* k của A, ký hiệu bởi

$$\Delta_k = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \ldots & \ldots & \ldots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}.$$

Định lý 6.1 (Tiêu chuẩn Sylvester)**.** *Cho V là không gian vectơ n chiều trên $\mathbb{R}$, $\mathcal{B}$ là một cơ sở của V, Q là một dạng toàn phương trên V và $A = [Q]_{\mathcal{B}}$. Khi đó:*

i) *Q xác định dương $\Leftrightarrow$ mọi định thức con chính của A đều dương.*

ii) *Q xác định âm $\Leftrightarrow$ mọi định thức con chính cấp chẵn của A đều dương và mọi định thức con chính cấp lẻ của A đều âm.*

Chứng minh. i) Chiều thuận là hiển nhiên. Ngược lại, giả sử mọi định thức con chính của A đều dương. Gọi f là dạng cực của Q và giả sử $\mathcal{B} = (u_1, \ldots, u_n)$. Tương tự quá trình trực chuẩn hóa Gram–Schmidt, ta xây dựng được cơ sở f-trực giao $\mathcal{B}' = (v_1, \ldots, v_n)$ từ cơ sở $\mathcal{B}$ bằng cách chọn các α_{ij} sao cho

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & 0 & \ldots & 0 \\ \alpha_{21} & \alpha_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \tag{$*$}$$

và với mỗi $k \le n$, $f(u_j, v_k) = \begin{cases} 0, & \text{nếu } j < k, \\ 1, & \text{nếu } j = k. \end{cases}$

• Chọn α_{11} sao cho $v_1 = \alpha_{11}u_1$ và $f(u_1, v_1) = 1$, nghĩa là

$$1 = f(u_1, \alpha_{11}u_1) = \alpha_{11}f(u_1, u_1) = \alpha_{11}a_{11} = \alpha_{11}\Delta_1.$$

Do đó ta có thể chọn $\alpha_{11} = \dfrac{1}{\Delta_1}$.

• Với $k > 1$, giả sử đã xác định được các $\alpha_{ij}, 1 \leq i < k, 1 \leq j \leq i$. Khi đó

$$(*) \Leftrightarrow \begin{cases} a_{11}\alpha_{k1} + a_{12}\alpha_{k2} + \dots + a_{1k}\alpha_{kk} = 0 \\ a_{21}\alpha_{k1} + a_{22}\alpha_{k2} + \dots + a_{2k}\alpha_{kk} = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{k-1,1}\alpha_{k1} + a_{k-1,2}\alpha_{k2} + \dots + a_{k-1,k}\alpha_{kk} = 0 \\ a_{k1}\alpha_{k1} + a_{k2}\alpha_{k2} + \dots + a_{kk}\alpha_{kk} = 1. \end{cases} \qquad (**)$$

Ma trận hệ số của hệ $(**)$ có định thức bằng $\Delta_k \neq 0$ nên các hệ số α_{kj} được xác định duy nhất, trong đó

$$\alpha_{kk} = \frac{1}{\Delta_k}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,k-1} & 0 \\ a_{21} & a_{22} & \cdots & a_{2,k-1} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{k-1,1} & a_{k-1,2} & \cdots & a_{k-1,k-1} & 0 \\ a_{k1} & a_{k2} & \cdots & a_{k,k-1} & 1. \end{vmatrix} = \frac{\Delta_{k-1}}{\Delta_k} \neq 0.$$

Ta có

$$\begin{vmatrix} \alpha_{11} & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \alpha_{11}\alpha_{22}\dots\alpha_{nn} = \frac{1}{\Delta_1}.\frac{\Delta_1}{\Delta_2}\dots\frac{\Delta_{n-1}}{\Delta_n} = \frac{1}{\Delta_n} \neq 0,$$

nên $\mathcal{B}'$ là cơ sở của V. Mặt khác, từ $(*)$ ta suy ra

$$f(v_i, v_j) = \delta_{ij}\alpha_{ij} \text{ với mọi } 1 \leq i, j \leq n,$$

nên $\mathcal{B}'$ là cơ sở f-trực giao của V.

Ma trận của Q theo cơ sở $\mathcal{B}'$ là

$$B := [Q]_{\mathcal{B}'} = \begin{pmatrix} Q(v_1) & 0 & \dots & 0 \\ 0 & Q(v_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & Q(v_n) \end{pmatrix}.$$

Với mỗi k, gọi A_k, B_k lần lượt là các ma trận có từ A, B bằng cách xoá đi $n-k$ dòng cuối và $n-k$ cột cuối. Khi đó A_k, B_k lần lượt là các ma trận của dạng toàn phương Q thu hẹp lên không gian sinh bởi $\{u_1, \dots, u_k\}$ trong các cơ sở $\mathcal{B}_k = (u_1, \dots, u_k)$ và $\mathcal{B}'_k = (v_1, \dots, v_k)$. Gọi $P_k = (\mathcal{B}_k \to \mathcal{B}'_k)$, ta được $B_k = (P_k)^\top A_k P_k$. Tương tự quá trình trực chuẩn hoá Gram-Schmidt, ta được P_k là ma trận tam giác trên có tất cả các hệ số trên đường chéo đều khác 0. Do đó $\det(P_k) = \det(P_k^\top) \neq 0$. Suy ra

$$\det(A_k) = \alpha \det(\mathcal{B}_k) = \alpha\, Q(v_1) \dots Q(v_k) \ (\alpha > 0).$$

Ta có Q xác định dương $\Leftrightarrow Q(v_k) > 0$ với mọi $k \Leftrightarrow \Delta_k > 0$ với mọi k.

ii) Do Q xác định âm $\Leftrightarrow -Q$ xác định dương nên ii) được suy ra từ i). $\square$

Giả sử dạng toàn phương Q có biểu thức của theo cơ sở $\mathcal{B} = (u_1, \dots, u_n)$ là $Q(u) = \sum\limits_{i,j=1}^{n} a_{ij} x_i x_j$ và giả sử ma trận $A = (a_{ij})$ có các định thức con chính $\Delta_k \neq 0$ với mọi k. Khi đó, từ chứng minh trên, ta có thuật toán để tìm cơ sở Q-chính tắc $\mathcal{B}' = (v_1, \dots, v_n)$ và xác định biểu thức của Q theo cơ sở $\mathcal{B}'$ như sau, gọi là ***Thuật toán Jacobi***:

Bước 1: Tính các định thức con chính Δ_k và kiểm tra $\Delta_k \neq 0$ với mọi k.

Bước 2: Đặt $\Delta_0 = 1$ và với mỗi k, đặt $\lambda_k := \dfrac{\Delta_{k-1}}{\Delta_k}$. Khi đó biểu thức của Q theo cơ sở $\mathcal{B}' = (v_1, \dots, v_n)$ là:

$$Q(u) = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2, \qquad \text{với } u = y_1 v_1 + \dots + y_n v_n.$$

Bước 3: Xác định cơ sở $\mathcal{B}' = (v_1, \ldots, v_n)$ như sau:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & 0 & \ldots & 0 \\ \alpha_{21} & \alpha_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix},$$

trong đó $\alpha_{kk} = \lambda_k$ và $\begin{pmatrix} \alpha_{k1} \\ \vdots \\ \alpha_{kk} \end{pmatrix} = A_k^{-1} \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$,

Ví dụ 40. Đưa dạng toàn phương thực sau về dạng chính tắc bằng thuật toán Jacobi:

$$Q(x, y, z) = 2x^2 + 9y^2 + 9z^2 + 8xy + 4xz + 12yz.$$

Giải. Ma trận của Q theo cơ sở chính tắc $\mathcal{B}_0 = (u_1, u_2, u_3)$ của $\mathbb{R}^3$ là

$$A = \begin{pmatrix} 2 & 4 & 2 \\ 4 & 9 & 6 \\ 2 & 6 & 9 \end{pmatrix}.$$

Các định thức con chính của A là

$$\Delta_1 = 2; \quad \Delta_2 = \begin{vmatrix} 2 & 4 \\ 4 & 9 \end{vmatrix} = 2; \quad \Delta_3 = \begin{vmatrix} 2 & 4 & 2 \\ 4 & 9 & 6 \\ 2 & 6 & 9 \end{vmatrix} = 6.$$

Theo tiêu chuẩn Sylvester, dạng chính tắc của Q theo cơ sở $\mathcal{B}' = (v_1, v_2, v_3)$ là

$$Q(u) = \frac{1}{\Delta_1}.y_1^2 + \frac{\Delta_1}{\Delta_2}.y_2^2 + \frac{\Delta_2}{\Delta_3}.y_3^2 = \frac{1}{2}y_1^2 + y_2^2 + \frac{1}{3}y_3^2.$$

Trong đó v_1, v_2, v_3 được xác định bởi:

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} \alpha_{11} & 0 & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} \alpha_{11} & 0 & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix},$$

trong đó

$$\begin{aligned} \alpha_{11} &= a_{11}^{-1} = \frac{1}{2}; \\ \begin{pmatrix} \alpha_{21} \\ \alpha_{22} \end{pmatrix} &= \begin{pmatrix} 2 & 4 \\ 4 & 9 \end{pmatrix}^{-1} . \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \\ \begin{pmatrix} \alpha_{31} \\ \alpha_{32} \\ \alpha_{33} \end{pmatrix} &= \begin{pmatrix} 2 & 4 & 2 \\ 4 & 9 & 6 \\ 2 & 6 & 9 \end{pmatrix}^{-1} . \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}. \end{aligned}$$

Vậy $v_1 = \left(\frac{1}{2}; 0; 0\right)$, $v_2 = (-2; 1; 0)$, $v_3 = \left(1; -\frac{2}{3}; \frac{1}{3}\right)$. □

Ví dụ 41. Tìm $m \in \mathbb{R}$ để dạng toàn phương sau xác định dương:

$$Q(x, y, z) = x^2 + my^2 + (m+3)z^2 - 2xy + 4xz - 6yz.$$

Giải.. Ma trận của Q là

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & m & -3 \\ 2 & -3 & m+3 \end{pmatrix}.$$

Các định thức con chính của A là

$$\Delta_1 = 1; \quad \Delta_2 = \begin{vmatrix} 1 & -1 \\ -1 & m \end{vmatrix} = m-1; \quad \Delta_3 = \begin{vmatrix} 1 & -1 & 2 \\ -1 & m & -3 \\ 2 & -3 & m+3 \end{vmatrix} = m(m-2).$$

Theo tiêu chuẩn Sylvester ta có

$$Q \text{ xác định dương} \Leftrightarrow \Delta_1 > 0, \Delta_2 > 0, \Delta_3 > 0 \Leftrightarrow m > 2.$$

□

Bài tập

Bài 1. Đưa các dạng toàn phương Q về dạng chính tắc, chỉ ra cơ sở Q-chính tắc và phép biến đổi tọa độ không suy biến tương ứng, trong đó $u = (x_1, x_2, x_3) \in \mathbb{R}^3$.

a) $Q(u) = x_1^2 + 5x_2^2 - 4x_3^2 + 2x_1x_2 - 4x_1x_3$.

b) $Q(u) = 4x_1^2 + x_2^2 + x_3^2 - 4x_1x_2 + 4x_1x_3 - 3x_2x_3$.

c) $Q(u) = 2x_1^2 + 18x_2^2 + 8x_3^2 - 12x_1x_2 + 8x_1x_3 - 27x_2x_3$.

d) $Q(u) = 2x_1x_2 - 3x_1x_3 - 6x_2x_3$.

e) $Q(u) = 3x_1x_2 + 4x_1x_3 + 12x_2x_3$.

Bài 2. Đưa các dạng toàn phương Q về dạng chính tắc trực giao, chỉ ra cơ sở Q-chính tắc trực giao và phép biến đổi tọa độ trực giao tương ứng, trong đó $u = (x_1, x_2, x_3) \in \mathbb{R}^3$.

a) $Q(u) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$.

b) $Q(u) = 3x_1^2 + 3x_3^2 + 4x_1x_2 + 4x_1x_3 - 2x_2x_3$.

c) $Q(u) = x_1^2 + x_2^2 + 5x_3^2 - 6x_1x_2 - 2x_1x_3 + 2x_2x_3$.

d) $Q(u) = x_1^2 + x_2^2 + x_3^2 + 4x_1x_2 + 4x_1x_3 + 4x_2x_3$.

e) $Q(u) = 6x_1^2 + 5x_2^2 + 7x_3^2 + 4x_1x_2 + 4x_1x_3$.

Bài 3. Tìm tham số m để dạng toàn phương Q xác định dương, trong đó $u = (x_1, x_2, x_3) \in \mathbb{R}^3$.

a) $Q(u) = 2x_1^2 + x_2^2 + 3x_3^2 + 2mx_1x_2 + 2x_1x_3$.

b) $Q(u) = x_1^2 + x_2^2 + 5x_3^2 + 2mx_1x_2 - 2x_1x_3 + 4x_2x_3$.

c) $Q(u) = 5x_1^2 + x_2^2 + mx_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3$.

d) $Q(u) = 2x_1^2 + 2x_2^2 + x_3^2 + 2mx_1x_2 + 6x_1x_3 + 2x_2x_3$.

e) $Q(u) = x_1^2 + 4x_2^2 + x_3^2 + 2mx_1x_2 + 10x_1x_3 + 6x_2x_3$.

Bài 4. Cho V là không gian vectơ n chiều trên $\mathbb{R}$ và Q là một dạng toàn phương trên V. Chứng minh rằng:

a) Nếu $Q(u) \neq 0, \forall\, 0 \neq u \in V$ thì Q xác định dương hoặc xác định âm.

b) Nếu Q không suy biến và tồn tại $0 \neq u \in V$ sao cho $Q(u) = 0$ thì $Q(V) = \mathbb{R}$.

c) Nếu Q suy biến thì tồn tại $0 \neq u \in V$ sao cho $Q(u) = 0$. Cho ví dụ chứng tỏ trong trường hợp này $Q(V)$ có thể bằng $\mathbb{R}$, có thể khác $\mathbb{R}$.

Bài 5. Cho dạng toàn phương:

$$Q(x, y, z) = x^2 + 3y^2 - 4xy + 2xz + 2yz.$$

Tìm tất cả các không gian con W của $\mathbb{R}^3$ sao cho $\dim W = 2$ và $Q(u) > 0$ với mọi $u \in W \setminus \{0\}$.

Bài 6. Cho V là không gian hữu hạn chiều trên $\mathbb{R}$, Q là một dạng toàn phương xác định dương trên V và f là dạng cực của Q. Chứng minh rằng với mọi $u, v \in V$ ta có

a) $|f(u, v)| \leq \sqrt{Q(u)Q(v)}$.

b) $\sqrt{Q(u+v)} \leq \sqrt{Q(u)} + \sqrt{Q(v)}$.

Bài 7. Chứng minh rằng một dạng toàn phương thực là xác định dương khi và chỉ khi ma trận A của nó (trong một cơ sở nào đó) được biểu diễn dưới dạng $A = C^\top C$, với C là một ma trận thực khả nghịch.

Bài 8. Cho Q_1 và Q_2 là hai dạng toàn phương thực, trong đó một trong hai dạng là xác định dương. Chứng minh rằng có thể đưa đồng thời Q_1 và Q_2 về dạng chính tắc (nghĩa là tồn tại một cơ sở $\mathcal{B}$ sao cho các ma trận của Q_1 và Q_2 theo cơ sở này đều là ma trận chéo). Chỉ ra rằng giả thiết xác định dương không thể bỏ được.

Bài 9. Cho Q là một dạng toàn phương thực có biểu thức tọa độ theo cơ sở $\mathcal{B} = (u_1, \ldots, u_n)$ có dạng

$$Q(u) = f_1^2 + \cdots + f_s^2 - f_{s+1}^2 - \cdots - f_{s+t}^2, \text{ với } u = x_1u_1 + \cdots + x_nu_n,$$

trong đó $f_1, \ldots, f_r$ là các dạng tuyến tính thực của các biến $x_1, \ldots, x_n$. Chứng minh rằng Q có chỉ số dương quán tính không vượt quá s và chỉ số âm quán tính không vượt quá t.

Bài 10. Cho Q_1 và Q_2 là hai dạng toàn phương trên V. Ta nói $\varphi \in \mathrm{End}(V)$ đưa Q_1 về Q_2 nếu $Q_2(u) = Q_1(\varphi(u))$ với mọi $u \in V$. Nếu tồn tại một phép biến đổi tuyến tính không suy biến trên V đưa Q_1 về Q_2 thì ta nói Q_1 và Q_2 tương đương. Chứng minh rằng

a) Hai dạng toàn phương là tương đương khi và chỉ khi chúng có cùng dạng chính tắc.

b) Hai dạng toàn phương thực là tương đương khi và chỉ khi chúng có cùng các chỉ số quán tính.

Bài 11. Cho Q_1 và Q_2 là hai dạng toàn phương thực. Chứng minh rằng nếu từ mỗi dạng này có thể đưa về dạng kia bằng một phép biến đổi tuyến tính (có thể suy biến) thì Q_1 và Q_2 tương đương.

Chỉ mục