ĐẠI SỐ A2

Bản nháp 2

Tháng 7/2016

Mục lục

Chương 6

DẠNG CHÍNH TẮC JORDAN

6.1 Sự tam giác hóa

Nếu một ma trận vuông A chéo hóa được thì, như chúng ta đã thấy ở Chương 5, nó có thể có khá nhiều ứng dụng thú vị. Tuy nhiên, không phải ma trận vuông nào cũng chéo hóa được. Vậy, chúng ta phải làm gì nếu A không chéo hóa được? Tồn tại những dạng rút gọn khác của ma trận vuông (hay của toán tử tuyến tính) mà những ứng dụng của chúng cũng rất đáng được quan tâm. Trong mục này chúng ta sẽ xem xét một sự rút gọn như vậy, được gọi là sự *tam giác hóa*.

Mệnh đề 6.1.1. *Mọi ma trận tam giác trên đều đồng dạng với một ma trận tam giác dưới nào đó.*

Chứng minh. Giả sử A là ma trận tam giác trên cấp n với các hệ sô trên trường $\mathbb{K}$. Ta chọn $f \in \mathrm{End}_{\mathbb{K}}(\mathbb{K}^n)$ 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 ma trận A' là ma trận tam giác dưới và $A' = P^{-1}AP$. Suy ra A' đồng dạng với A. ■

Bài toán đặt ra là khi nào thì ma trận $A \in M_n(\mathbb{K})$ đồng dạng với một ma trận tam giác? Do Mệnh đề 6.1.1 nên ta chỉ cần xét khi nào ma trận A đồng dạng với một ma trận tam giác trên (hay A được gọi là *tam giác hóa được*). 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 6.1.2. *Toán tử $f \in \mathrm{End}_{\mathbb{K}}(V)$ tam giác hóa được khi và chỉ khi đa thức đặc trưng của f phân rã trên $\mathbb{K}$.*

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

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

Từ đó ta có

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

nghĩa là $P_f(\lambda)$ phân rã trên $\mathbb{K}$.

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

Nếu $n = 1$ thì không có gì để chứng minh. Vậy, giả sử $n > 1$ và khẳng định đúng với $n - 1$. Gọi $\lambda_1 \in \mathbb{K}$ là một nghiệm nào đó của $P_f(\lambda)$ 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

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

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 $g \in \mathrm{End}_{\kappa}(W)$ sao cho ma trận biểu diễn g theo cơ sở $(u_2, \ldots, u_n)$ là B. Ta có

$$P_f(\lambda) = \det(A - \lambda I_n) = (\lambda_1 - \lambda)\det(B - \lambda I_{n-1}) = (\lambda_1 - \lambda) P_g(\lambda).$$

Vì $P_f(\lambda)$ phân rã trên $\mathbb{K}$ nên $P_g(\lambda)$ cũng phân rã trên $\mathbb{K}$, do đó theo giả thiết qui nạp ma trận B tam giác hóa được. Vậy tồn tại một cơ sở $(v_2, \ldots, v_n)$ 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 $(u_1, v_2, \ldots, v_n)$ cũng có dạng tam giác trên. ■

Hệ quả 6.1.3. *Mọi ma trận $A \in M_n(\mathbb{C})$ đều tam giác hóa được.*

Nhận xét 6.1.4. 1) 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.

2) Mọi ma trận $A \in M_n(\mathbb{R})$ đều tam giác hóa được trên $\mathbb{C}$.

Hệ quả 6.1.5. *Cho ma trận $A \in M_n(\mathbb{R})$ và phổ của A là $Sp_{\mathbb{C}}(A) = \{\lambda_1, \dots, \lambda_n\}$. Khi đó ta có*

$$tr(A) = \lambda_1 + \cdots + \lambda_n \textit{ và } \det A = \lambda_1 \dots \lambda_n.$$

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 6.1.6. Ma trận $A = \begin{pmatrix} -4 & 0 & -2 \\ 0 & 1 & 0 \\ 5 & 1 & 3 \end{pmatrix}$ có đa thức đặc trưng $P_A(\lambda) = -(\lambda+2)(\lambda-1)^2$ nên theo Định lý 6.1.2, 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 một 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

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

Ta sẽ tính các vectơ u_1, u_2 và u_3. 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 $u_1 = (-2, 0, 5)$.

Tính u_2: Ta có $f(u_2) = au_1 + u_2 \Rightarrow (f - Id)(u_2) = au_1$. 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

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

Cho $a = -1, x_3 = 4$. Ta có $x_1 = -2, x_2 = -3$. Vậy ta có thể lấy $u_2 = (-2, -3, 4)$.

Tính 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 $A' = P^{-1}AP$.

6.2 Đa thức triệt tiêu. Định lý Hamilton - Calley

Cho V là một không gian vectơ trên trường $\mathbb{K}$ và $Q \in \mathbb{K}[t]$:

$$Q(t) = a_m t^m + a_{m-1}t^{m-1} + \cdots + a_1 t + a_0.$$

Nếu $f \in \mathrm{End}_{\mathbb{K}}(V)$ thì ta ký hiệu $Q(f)$ là một toán tử tuyến tính trên V xác định bởi

$$Q(f) = a_m f^m + a_{m-1} f^{m-1} + \cdots + a_1 f + a_0 Id_V.$$

Nhận xét 6.2.1. Nếu $P, Q \in \mathbb{K}[t]$ thì

$$P(f)Q(f) = Q(f)P(f), \quad \forall f \in \mathrm{End}_{\mathbb{K}}(V).$$

Định nghĩa 6.2.2. Cho $f \in \mathrm{End}_{\mathbb{K}}(V)$ và $Q(t) \in \mathbb{K}[t]$. Ta nói $Q(t)$ là *đa thức triệt tiêu* toán tử f nếu $Q(f) = 0$.

Mệnh đề 6.2.3. *Giả sử $Q(t)$ là đa thức triệt tiêu toán tử f và λ là một trị riêng của f. Khi đó λ là nghiệm của $Q(t)$.*

Chứng minh. Gọi v là một vectơ riêng của f ứng với trị riêng λ. Khi đó $f^k(v) = \lambda^k v, \forall k \in \mathbb{N}$. Giả sử

$$Q(t) = a_m t^m + a_{m-1} t^{m-1} + \cdots + a_1 t + a_0$$

là đa thức triệt tiêu f. Khi đó ta có

$$\begin{aligned}
& a_m f^m + a_{m-1} f^{m-1} + \cdots + a_1 f + a_0 Id_V = 0 \\
\Rightarrow\ & (a_m f^m + a_{m-1} f^{m-1} + \cdots + a_1 f + a_0 Id_V)(v) = 0 \\
\Rightarrow\ & a_m f^m(v) + a_{m-1} f^{m-1}(v) + \cdots + a_1 f(v) + a_0 Id_V(v) = 0 \\
\Rightarrow\ & a_m \lambda^m v + a_{m-1} \lambda^{m-1} v + \cdots + a_1 \lambda v + a_0 v = 0 \\
\Rightarrow\ & (a_m \lambda^m + a_{m-1} \lambda^{m-1} + \cdots + a_1 \lambda + a_0) v = 0.
\end{aligned}$$

Do $v \neq 0$ nên từ đó suy ra

$$a_m \lambda^m + a_{m-1} \lambda^{m-1} + \cdots + a_1 \lambda + a_0 = 0 \text{ hay } Q(\lambda) = 0. \qquad \blacksquare$$

Áp dụng mệnh đề vừa chứng minh ta thấy rằng nếu toán tử f thỏa $f^2 = f$ thì các giá 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, cũng cần thiết lưu ý rằng không phải tất cả các nghiệm của $Q(t)$ đều là trị riêng của f. Ví dụ, nếu $f = Id_V$ 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 đối với mọi toán tử tuyến tính $f \in \mathrm{End}_{\mathbb{K}}(V)$ đều tồn tại một đa thức $0 \neq Q(t) \in \mathbb{K}[t]$ triệt tiêu f? Câu trả lời là khẳng định. Thật vậy, nếu $\dim_{\mathbb{K}}(V) = n$ thì $\mathrm{End}_{\mathbb{K}}(V) \cong M_n(\mathbb{K})$, suy ra $\dim_{\mathbb{K}}(\mathrm{End}_{\mathbb{K}}(V)) = n^2$. Do đó các phần tử $Id_V, f, f^2, \ldots, f^{n^2}$ phụ thuộc tuyến tính trong $\mathrm{End}_{\mathbb{K}}(V)$, suy ra tồn tại các phần tử $a_0, a_1, a_2, \ldots, a_{n^2} \in \mathbb{K}$, không phải tất cả đều bằng 0 sao cho

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

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 Hamilton - Calley mà ta sẽ chứng minh dưới đây cho thấy đa thức đặc trưng của f là đa thức triệt tiêu đa thức triệt tiêu f.

Định lý 6.2.4 (Hamilton - Calley). *Cho f là một toán tử tuyến tính trên không gian hữu hạn chiều V. Khi đó đa thức đặc trưng $P_f(\lambda)$ triệt tiêu f,* *nghĩa là* $P_f(f) = 0$.

Gọi $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ là một cơ sở của V và A là ma trận biểu diễn f theo cơ sở $\mathcal{B}$. Như vậy,

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

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 ta có

$$\sum_{j=1}^{n} (a_{ji} Id_V - \delta_{ij} f)(u_j) = 0, \quad 1 \leq i \leq n$$

Gọi $B = (b_{ij})$ là ma trận vuông với các hệ số là các toán tử tuyến tính được xác định bởi

$$b_{ij} = a_{ji} Id_V - \delta_{ij} f.$$

Do $P_f(x) = \det(A - xI_n) = \det(A - xI_n)^\top$, mà ma trận này có các hệ số là những đa thức

$$[(A - xI_n)^\top]_{ij} = a_{ij} Id_V - \delta_{ij} x,$$

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

$$(\det B)(u_k) = 0, \quad \forall k = 1, \ldots, n.$$

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

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

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

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

Suy ra, với mỗi $1 \leq k \leq n$ 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,

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

Suy ra

$$\det B(u_k) = \sum_{i=1}^{n} \overline{B}_{ki} \left(\sum_{j=1}^{n} b_{ij}(u_j) \right) = 0.$$

Dưới đây ta sẽ chứng minh một kết quả rất quan trọng về các đa thức triệt tiêu.

Bổ đề 6.2.5 (Bổ đề căn bản). *Cho $f \in \mathrm{End}_{\mathbb{K}}(V)$ và $Q(t) = Q_1(t) \dots Q_p(t)$, trong đó $Q_1, \dots, Q_p$ là những đa thức nguyên tố cùng nhau. Khi đó, nếu $Q(t)$ triệt tiêu f thì*

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

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

Nếu $p = 1$ thì $Q(t) = Q_1(t)$, do đó nếu $Q(f) = 0$ thì $Q_1(f) = 0$ 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 là những đa thức nguyên tố cùng nhau. Khi đó, tồn tại những đa thức U_1 và U_2 sao cho

$$U_1 Q_1 + U_2 Q_2 = 1.$$

Từ đó suy ra

$$U_1(f)Q_1(f) + U_2(f)Q_2(f) = Id_V.$$

Do đó $\forall u \in V$, 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 $Q_1(f)Q_0(f) = 0$ 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

$$\operatorname{Im}(U_2(f)Q_2(f)) \subseteq \operatorname{Ker} Q_1(f).$$

Do đó, từ (2) ta có $V = \operatorname{Ker} Q_1(f) + \operatorname{Ker} Q_2(f)$. Giả sử $u \in \operatorname{Ker} Q_1(f) \cap \operatorname{Ker} Q_2(f)$. Từ (1) suy ra ngay $u = 0$. Vậy

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

Trường hợp $p > 2$: Ta có

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

Đặt $\tilde{Q}(t) = Q_1(t) \dots Q_{p-1}(t)$, ta có $\tilde{Q}(t)$ và $Q_p(t)$ là những đa thức nguyên tố cùng nhau. Theo trường hợp $p = 2$ ta có

$$V = \operatorname{Ker} \tilde{Q}(f) \oplus \operatorname{Ker} Q_p(f).$$

Đặt $W = \operatorname{Ker} \tilde{Q}(f)$ và $\overline{f} = f|_W$. Ta chứng minh $\overline{f} \in \operatorname{End}_K(W)$. Thật vậy, $\forall u \in W$, ta có $\overline{f}(u) = f(u)$. Do $u \in W$ nên $\tilde{Q}(f)(u) = 0$, suy ra $\tilde{Q}(f)f(u) = f\tilde{Q}(f)(u) = 0$. Hơn nữa $\tilde{Q}(f)f(u) = \tilde{Q}(f)(f(u))$, do đó $f(u) \in \operatorname{Ker} \tilde{Q}(f)$.

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

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

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

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

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

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

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

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

Vậy

$$V = \operatorname{Ker} Q_1(f) \oplus \cdots \oplus \operatorname{Ker} Q_{p-1}(f) \oplus \operatorname{Ker} Q_p(f). \qquad \blacksquare$$

Hệ quả 6.2.6. *Cho f là một toán tử tuyến tính trong không gian vectơ n chiều V trên $\mathbb{K}$ và giả sử đa thức đặc trưng $P_f(\lambda)$ phân rã trên $\mathbb{K}$:*

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

Khi đó

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

Định nghĩa 6.2.7. Cho f là một toán tử tuyến tính trong không gian vectơ n chiều V trên $\mathbb{K}$ và giả sử đa thức đặc trưng $P_f(\lambda)$ phân rã trên $\mathbb{K}$:

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

Ta gọi

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

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

Theo Hệ quả 6.2.6, không gian V được phân tích thành tổng trực tiếp của các không gian đặc trưng.

Nhận xét 6.2.8.

1) Không gian riêng luôn nằm trong không gian đặc trưng (ứng với cùng một trị riêng)

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

Thật vậy, nếu $u \in E(\lambda)$ thì $(f - \lambda Id)(u) = 0$, suy ra $(f - \lambda Id)^\alpha(u) = 0$ (α là số bội của λ) hay $u \in N(\lambda)$.

2) Không gian đặc trưng là bất biến đối với f, nghĩa là

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

Thật vậy, giả sử $u \in N(\lambda)$. Khi đó

$$\begin{aligned} & (f - \lambda Id)^\alpha(u) = 0 \Rightarrow f(f - \lambda Id)^\alpha(u) = 0 \\ \Rightarrow\ & (f - \lambda Id)^\alpha f(u) = 0 \Rightarrow f(u) \in N(\lambda). \end{aligned}$$

Để kết thúc mục này, dưới đây ta sẽ xét một ứng dụng của Bổ đề căn bản.

Giả sử $f \in \operatorname{End}_{\mathbb{K}}(V)$ sao cho $f^2 = f$. Khi đó $Q(t) = t(t-1)$ là đa thức triệt tiêu f. Vậy, theo Bổ đề 6.2.5, $V = \operatorname{Ker} f \oplus \operatorname{Ker}(f - Id)$. 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 chúng ta vừa xét.

Định lý 6.2.9. *Toán tử tuyến tính f chéo hóa được khi và chỉ khi tồn tại một đa thức phân rã trên $\mathbb{K}$, có toàn nghiệm đơn và triệt tiêu f.*

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

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

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

Ngược lại, giả sử

$$Q(t) = (t - \lambda_1) \ldots (t - \lambda_p), \lambda_i \neq \lambda_j, \forall i \neq j$$

và $Q(t)$ triệt tiêu f. Khi đó theo Bổ đề 6.2.5, ta có

$$\begin{aligned} V &= \operatorname{Ker}(f - \lambda_1 Id) \oplus \cdots \oplus \operatorname{Ker}(f - \lambda_p Id) \\ &= E(\lambda_1) \oplus \cdots \oplus E(\lambda_p). \end{aligned}$$

Suy ra f chéo hóa được. ■

6.3 Đa thức tối tiểu

Định nghĩa 6.3.1. Đa thức đơn khởi bậc nhỏ nhất triệt tiêu toán tử tuyến tính f được gọi là *đa thức tối tiểu* của f và ký hiệu là m_f.

Mệnh đề 6.3.2. *Đa thức $Q(t) \in K[t]$ triệt tiêu f khi và chỉ khi $Q(t)$ chia hết cho $m_f(t)$ trong $\mathbb{K}[t]$.*

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

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

(theo qui ước, đa thức 0 có bậc $-\infty$). Vì $Q(f) = 0$ nên suy ra $R(f) = 0$, kéo theo $R(t) = 0$ do định nghĩa đa thức tối tiểu. Do đó $Q(t) = P(t)m_f(t)$.

Ngược lại, nếu $Q(t) = P(t)m_f(t)$ thì $Q(f) = P(f)m_f(f) = 0$, nghĩa là $Q(t)$ triệt tiêu f. ■

Hệ quả 6.3.3. *m_f là ước của P_f.*

Chứng minh. Áp dụng Định lý Hamilton-Calley và Mệnh đề 6.3.2 ■

Hệ quả 6.3.4. *Đa thức tối tiểu 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 6.3.2, 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 từ đó suy ra $m_1 = m_2$. ■

Mệnh đề 6.3.5. *Tập nghiệm của m_f trùng với tập nghiệm của P_f.*

Chứng minh. Từ Hệ quả 6.3.3 suy ra mỗi nghiệm của m_f đều là nghiệm của P_f. Từ Mệnh đề 6.2.3 suy ra mỗi nghiệm của P_f đều là nghiệm của m_f.

Ví dụ 6.3.6. a) 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 6.3.5 suy ra $m_A(t) = (t+1)(t+2)(t-3)$.

b) 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 6.3.5 suy ra

$$m_A(t) = \left[\begin{array}{l} (t-1)(t+2), \\ (t-1)(t+2)^2. \end{array} \right.$$

Ta có $(A - I_3)(A + 2I_3) = 0$. Vậy $m_A(t) = (t-1)(t+2)$.

Định lý 6.3.7. *Toán tử tuyến tính f chéo hóa được khi và chỉ khi đa thức tối tiểu của nó phân rã trên $\mathbb{K}$ và có toàn nghiệm đơn.*

Chứng minh. Điều kiện đủ là đúng do Định lý 6.2.9. 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 6.2.9 đa thức $Q(t) = (t-\lambda_1)\dots(t-\lambda_p)$ triệt tiêu f. Áp dụng Mệnh đề 6.3.5 suy ra $m_f(t) = Q(t)$. ■

Ví dụ 6.3.8. a) 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 6.3.7, A chéo hóa được.

b) Ma trận $A = \begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ có đa thức đặc trưng $P_A(t) = -(t-1)^3$ nên

$$m_A(t) = \left[\begin{array}{l} t-1 \\ (t-1)^2, \\ (t-1)^3. \end{array} \right.$$

Theo Định lý 6.3.7, A chéo hóa được $\Leftrightarrow m_A(t) = t - 1 \Leftrightarrow A - I_3 = 0$. Do $A \neq I_3$ nên từ đó suy ra A không chéo hóa được.

c) Ma trận $A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$ có đa thức đặc trưng là $P_A(t) = -(t-1)(t-2)^2$, do đó

$$m_A(t) = \left[\begin{array}{l} (t-1)(t-2), \\ (t-1)(t-2)^2. \end{array} \right.$$

Theo Định lý 6.3.7, ta có

$$A \text{ chéo hóa được} \Leftrightarrow m_A(t) = (t-1)(t-2) \Leftrightarrow (A - I_3)(A - 2I_3) = 0.$$

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.

6.4 Dạng tam giác khối

Trong mục 6.1 ta đã biết nếu đa thức đặc trưng của toán tử f trên không gian vectơ hữu hạn chiều phân rã trên trường cơ sở $\mathbb{K}$ thì f tam giác hóa được, nghĩa là tồn tại một cơ sở $\mathcal{B}$ của V sao cho trong đó ma trận biểu diễn f theo $\mathcal{B}$ có dạng tam giác (trên hoặc dưới). Nói chung, điều này cũng đã cho chúng ta những ứng dụng khá tốt. Trong mục này chúng ta tiếp tục việc rút gọn toán tử f sao cho có thể "tốt" hơn nữa. Chính xác hơn, nếu đa thức đặc trưng của f phân rã trên $\mathbb{K}$ thì ta có thể đưa f về dạng tam giác khối.

Bổ đề 6.4.1. *Cho $V = V_1 \oplus \ldots \oplus V_p$, 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

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

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 $\mathcal{B}_1 = (u_1, \ldots, u_{n_1}), \ldots, \mathcal{B}_p = (u_1, \ldots, u_{n_p})$. Vì $f(V_i) \subseteq V_i, \forall i = \overline{1,p}$ nên ta có

$$\begin{cases} f(u_1) = a_{11}u_1 + \ldots + a_{n_1 1}u_{n_1}; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(u_{n_1}) = a_{1n_1}u_1 + \ldots + a_{n_1 n_1}u_{n_1} \end{cases}$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\begin{cases} f(u_1) = b_{11}u_1 + \ldots + b_{n_p 1}u_{n_p}; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(u_{n_p}) = b_{1n_p}u_1 + \ldots + b_{n_p n_p}u_{n_p}. \end{cases}$$

Do đó

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

Định lý 6.4.2 (Rút gọn theo dạng tam giác khối). *Cho f là một toán tử tuyến tính trên không gian vectơ n chiều trên trường $\mathbb{K}$. Giả sử đa thức đặc trưng của f phân rã trên $\mathbb{K}$:*

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

Khi đó tồn tại một 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 6.2.6 và Định lý 6.4.2, tồn tại một cơ sở $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_p)$ của V sao cho

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

trong đó $\mathcal{B}_i$ là một 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 còn cần chứng minh rằng M_i tam giác hóa được và $Sp_{\mathbb{K}}(M_i) = \{\lambda_i, \ldots, \lambda_i\}$ là đủ.

Theo định nghĩa, $N(\lambda_i) = \operatorname{Ker}(f - \lambda_i Id)^{\alpha_i}$ nên $(f - \lambda_i Id)^{\alpha_i}(u) = 0, \forall u \in N(\lambda_i)$. Vậy $(\lambda - \lambda_i)^{\alpha_i}$ là đa thức triệt tiêu f_i, do đó đa thức tối tiểu của f_i có dạng

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

Áp dụng Mệnh đề 6.3.2, suy ra đa thức đặc trưng của f_i có dạng

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

Vậy $P_{f_i}(\lambda)$ phân rã trên $\mathbb{K}$ và $Sp_{\mathbb{K}}(f_i) = \{\lambda_i, \dots, \lambda_i\}$, 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, ta có

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

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

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. ■

Ví dụ 6.4.3. Cho ma trận $A = \begin{pmatrix} 1 & -1 & 2 & 2 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \end{pmatrix}$. Xem A như ma trận của toán tử tuyến tính $f : \mathbb{R}^4 \to \mathbb{R}^4$ theo cơ sở chính tắc $\mathcal{B}_0 = (e_1, e_2, e_3, e_4)$. Vì $P_f(\lambda) = \lambda^2(\lambda + 1)(\lambda - 3)$ nên tồn tại cơ sở $\mathcal{B} = (u_1, u_2, u_3, u_4)$ sao cho theo cơ sở này ma trận của f có dạng tam giác khối

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

Điều này có nghĩa

$$\begin{cases} f(u_1) = -u_1, \\ f(u_2) = 3u_2, \\ f(u_3) = 0, \\ f(u_4) = au_3. \end{cases}$$

Ta có vectơ u_1 là vectơ riêng của f ứng với trị riêng $\lambda = -1$. Có thể lấy

$u_1 = (-2, 0, 1, 1)$. Vectơ u_2 là vectơ riêng ứng với trị riêng $\lambda = 3$. Có thể lấy $u_2 = (2, 0, 1, 1)$. Vectơ u_3 là vectơ riêng ứng với trị riêng $\lambda = 0$. Có thể lấy $u_3 = (1, 1, 0, 0)$. Cuối cùng cho $a = 1$ ta tính được $u_4 = (0, 1, 1, 0)$. Dễ dàng kiểm tra rằng các vectơ u_1, u_2, u_3, u_4 độc lập tuyến tính, do đó chúng tạo thành một cơ sở của $\mathbb{R}^4$ mà ta cần tìm. Ma trận chuyển cơ sở từ $\mathcal{B}_0$ sang $\mathcal{B}$ là

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

Như vậy chúng ta đã thấy rằng nếu một ma trận tam giác hóa được thì ta có thể làm được nhiều hơn thế nữa bằng cách đưa ma trận biểu diễn nó về dạng tam giác khối. Điều này có lợi hơn trong các ứng dụng. Thực vậy, nếu đối với một ma trận tam giác việc nâng lên lũy thừa nói chung là không thể được. Tuy nhiên, nếu chúng 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. Về phần mình, mỗi khối như vậy đều là một ma trận tam giác mà tất cả các phần tử nằm trên đường chéo chính 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 $N = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$. Ta chứng minh tồn tại $p \leq n$ sao cho $N^p = 0$. Thực vậy, ta có $P_N(t) = (-1)^n t^n$, nên $m_N(t) = t^p$, với $p \leq n$. Do $m_N(N) = 0$ nên $N^p = 0$. Bây giờ nhận xét rằng λI và N giao hoán với nhau nên ta dễ dàng tính được $A^k = (\lambda I + N)^k$ bằng cách khai triển nhị thức Newton.

6.5 Toán tử lũy linh

Trong mục này V luôn luôn là không gian vectơ n chiều trên trường $\mathbb{K}$.

Định nghĩa 6.5.1. Giả sử $\lambda \in \mathbb{K}$ và $n \geq 2$, ta gọi ma trận vuông cấp n dạng sau đây là một *khối Jordan*:

$$J(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}.$$

Nếu cấp của khối bằng 1 thì ta qui ước $J(\lambda) = (\lambda)$.

Bổ đề 6.5.2. *Giả sử $J = J(\lambda)$ là một khối Jordan cấp n. Khi đó ta có những điều khẳng định sau đây:*

(i) $P_J(t) = (-1)^n(t-\lambda)^n$.

(ii) $m_J(t) = (t-\lambda)^n$.

(iii) $\dim E(\lambda) = 1$.

Chứng minh. Bằng cách tính toán trực tiếp ta thấy ngay những điều cần chứng minh. ■

Định nghĩa 6.5.3. Ta nói toán tử tuyến tính u trong không gian V là một *toán tử lũy linh* nếu tồn tại một số nguyên $m \geq 1$ sao cho $u^m = 0$. Số nguyên dương β nhỏ nhất thỏa $u^\beta = 0$ được gọi là *bậc lũy linh* của u.

Mệnh đề 6.5.4. *Toán tử tuyến tính u là lũy linh nếu và chỉ nếu tồn tại một cơ sở của V sao cho trong đó ma trận biểu diễn u có dạng tam giác trên mà trên đường chéo chính chỉ toàn là 0.*

Chứng minh. Giả sử u là toán tử lũy linh bậc β. Khi đó đa thức tối tiểu $m_u(t) = t^\beta$, kéo theo đa thức đặc trưng $P_u(t) = (-1)^n t^n$. Do đó u tam giác hóa được, nghĩa là tồn tại một cơ sở của V mà trong đó ma trận biểu diễn u có dạng tam giác trên với các phần tử 0 nằm trên đường chéo chính. Điều ngược lại là hiển nhiên. ■

Định nghĩa 6.5.5. Cho u là toán tử tuyến tính trong không gian vectơ V. Nếu tồn tại một vectơ x trong V sao cho các vectơ $x, u(x), \ldots, u^{n-1}(x)$ tạo thành một cơ sở của V thì ta nói u là *toán tử xyclic* và $\mathcal{B} = (x, u(x), \ldots, u^{n-1}(x))$ là *cơ sở xyclic*.

Bổ đề 6.5.6. *Cho u là toán tử lũy linh với bậc lũy linh $\beta > 1$. Khi đó, các điều kiện dưới đây tương đương:*

(i) $\beta = n$, nghĩa là $P_u(t) = (-1)^n t^n$ và $m_u(t) = t^n$.

(ii) u là toán tử xyclic.

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

Chứng minh. (i) $\Rightarrow$ (ii). Giả sử điều kiện (i) được thỏa mãn. Khi đó $u^{n-1} \neq 0$. Do đó tồn tại $0 \neq x \in V$ sao cho $u^{n-1}(x) \neq 0$. Giả sử

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

Tác động u^{n-1} lên hai vế, nhận được $a_1 u^{n-1}(x) = 0$. Từ đó suy ra $a_1 = 0$. Bây giờ tác động u^{n-2} lên hai vế, nhận được $a_2 = 0$. Cứ tiếp tục như thế, cuối cùng nhận được $a_1 = a_2 = \cdots = a_n = 0$. Vậy, $(x, u(x), \ldots, u^{n-1}(x))$ là cơ sở xyclic của V.

(ii) $\Rightarrow$ (iii). Giả sử $(x, u(x), \ldots, u^{n-1}(x))$ là cơ sở xyclic của V. Đặt $\mathcal{B} = (u^{n-1}(x), \ldots, u^2(x), u(x), x)$. Khi đó ta có

$$[u]_{\mathcal{B}} = J(0).$$

(iii) $\Rightarrow$ (i). Nếu có (iii) thì theo Bổ đề 6.5.2 ta có $\beta = n$. ■

Trong các bổ đề tiếp theo ta luôn giả thiết u là toán tử lũy linh với bậc lũy linh $\beta > 1$.

Bổ đề 6.5.7. *Với mọi $p \in \overline{0, \beta}$, ký hiệu $K_p := \operatorname{Ker}(u^p)$. Khi đó ta có dãy tăng ngặt sau đây:*

$$0 = K_0 \subset K_1 \subset K_2 \subset \ldots \subset K_{\beta-1} \subset K_\beta = V.$$

Chứng minh. Vì $u^p(x) = 0 \Rightarrow u^{p+1}(x) = 0$, nên $K_p \subseteq K_{p+1}$. Mặt khác, nếu tồn tại $p \in \overline{1, \beta - 1}$ sao cho $K_p = K_{p+1}$ thì ta có

$$K_p = K_{p+1} = \ldots = K_\beta = V.$$

Điều này dẫn đến p là bậc lũy linh của u, là điều mâu thuẫn vì $p < \beta$. ■

Bổ đề 6.5.8. *Tồn tại các không gian con $M_1, \ldots, M_\beta$ không tầm thường thỏa mãn các điều kiện:*

(i) $K_p = K_{p-1} \oplus M_p, p \in \overline{1, \beta}$;

(ii) $u(M_p) \subseteq M_{p-1}, p \in \overline{2, \beta}$.

Chứng minh. Với $p = \beta$, chọn M_β là một phần bù nào đó của $K_{\beta-1}$ trong $K_\beta = V$. Giả sử đã xây dựng được các không gian con $M_\beta, M_{\beta-1}, \ldots, M_p$ thỏa mãn các điều kiện (i) và (ii). Ta sẽ xây dựng không gian con M_{p-1} sao cho nó cũng thỏa mãn (i) và (ii). Trước hết ta kiểm tra các điều kiện sau đây:

a) $u(M_p) \subseteq K_{p-1}$;

b) $u(M_p) \cap K_{p-2} = 0$.

Giả sử $x \in M_p$. Vì $M_p \subseteq K_p$ nên $u^p(x) = 0$, hay $u^{p-1}(u(x)) = 0$, nghĩa là $u(x) \in K_{p-1}$. Ta đã chứng minh (a).

Bây giờ, giả sử $y \in u(M_p) \cap K_{p-2}$. Khi đó, ta có $y = u(x), x \in M_p$ và $u^{p-2}(y) = 0$, kéo theo $u^{p-1}(x) = 0$, nghĩa là $x \in K_{p-1}$. Vậy, $x \in K_{p-1} \cap M_p = 0$, suy ra $x = 0$, kéo theo $y = u(x) = 0$. Ta đã chứng minh (b). Kết hợp (a) và (b) ta có

$$K_{p-2} \oplus u(M_p) \subseteq K_{p-1}.$$

Do đó tồn tại G_{p-1} là phần bù của $K_{p-2} \oplus u(M_p)$ trong K_{p-1}, nghĩa là

$$K_{p-1} = K_{p-2} \oplus u(M_p) \oplus G_{p-1}.$$

Bây giờ nếu đặt

$$M_{p-1} = u(M_p) \oplus G_{p-1}$$

thì ta thấy M_{p-1} thỏa các điều kiện (i) và (ii). ■

Bổ đề 6.5.9. *Với các ký hiệu $M_1, M_2, \ldots, M_\beta$ như trên, ta có*

$$V = M_1 \oplus M_2 \oplus \ldots \oplus M_\beta.$$

Chứng minh. Ta có

$$\begin{aligned} V = K_\beta &= K_{\beta-1} \oplus M_\beta \\ &= K_{\beta-2} \oplus M_{\beta-1} \oplus M_\beta \\ &= \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &= M_1 + \oplus M_2 \oplus \ldots \oplus M_\beta. \end{aligned}$$ ■

Tiếp theo, ta sẽ xây dựng một cơ sở của V bằng cách xây dựng cơ sở cho mỗi không gian con M_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 6.5.10. *Ảnh của một cơ sở của M_p là một họ độc lập tuyến tính trong M_{p-1}.*

Chứng minh. Giả sử $(u_1, \ldots, u_r)$ là một cơ sở của M_p và $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ sao cho

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

Khi đó, $u(\lambda_1 u_1 + \cdots + \lambda_r u_r) = 0$, hay $\lambda_1 u_1 + \cdots + \lambda_r u_r \in \operatorname{Ker} u = K_1 \subseteq K_{p-1}$. Từ đó suy ra $\lambda_1 u_1 + \cdots + \lambda_r u_r \in M_p \cap K_{p-1} = 0$, kéo theo $\lambda_1 = \cdots = \lambda_r = 0$. ■

Bây giờ ta sẽ lần lượt xây dựng cơ sở cho M_i bắt đầu từ $i = \beta$. Trong M_β (mà ta còn ký hiệu là G_β), lấy một cơ sở bất kỳ nào đó. Tiếp theo, bằng qui nạp lùi ta sẽ xây dựng cơ sở trong M_{p-1} như sau:

Ta có $M_{p-1} = u(M_p) \oplus G_{p-1}$. Lấy ảnh qua u của cơ sở đã xây dựng trong M_p hợp với một cơ sở của G_{p-1} ta sẽ có một cơ sở của M_{p-1}. Bằng cách này ta xây dựng được các cơ sở $\mathcal{B}_\beta, \mathcal{B}_{\beta-1} \ldots, \mathcal{B}_1$ tương ứng của $M_\beta, M_{\beta-1} \ldots, M_1$:

$$\mathcal{B}_\beta = \underbrace{(u_1, \ldots, u_{n_\beta})}_{G_\beta},$$

$$\mathcal{B}_{\beta-1} = (u(u_1), \ldots, u(u_{n_\beta}), \underbrace{\omega_1, \ldots, \omega_{n_{\beta-1}}}_{G_{\beta-1}}),$$

$$\mathcal{B}_{\beta-2} = (u^2(u_1), \ldots, u^2(u_{n_\beta}), u(\omega_1), \ldots, u(\omega_{n_{\beta-1}}), \underbrace{z_1, \ldots, z_{n_{\beta-2}}}_{G_{\beta-2}}),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\mathcal{B}_1 = (u^{\beta-1}(u_1), \ldots, u^{\beta-1}(u_{n_\beta}), u^{\beta-2}(\omega_1), \ldots, u^{\beta-2}(\omega_{n_{\beta-1}}),$$
$$u^{\beta-3}(z_1), \ldots, u^{\beta-3}(z_{n_{\beta-2}}), \ldots, \underbrace{x_1, \ldots, x_{n_1}}_{G_1}).$$

Với $p \in \overline{1, \beta}$ và $0 \neq x \in G_p$, đặt

$$I_p(x) := \langle x, u(x), \ldots, u^{p-1}(x) \rangle.$$

Khi đó những khẳng định trong bổ đề dưới đây là hiển nhiên.

Bổ đề 6.5.11. *Với mọi $p \in \overline{1, \beta}$, ta có*

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

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

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

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

Định lý 6.5.12. *Cho u là một toán tử lũy linh trong không gian vectơ V. Khi đó V phân tích thành tổng trực tiếp các không gian con bất biến đối với u, sao cho hạn chế của u lên mỗi không gian con này là một toán tử xyclic.*

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

$$H_\beta(u_1), \ldots, H_\beta(u_{n_\beta}), H_{\beta-1}(\omega_1), \ldots, H_{\beta-1}(\omega_{n_{\beta-1}}),$$
$$\ldots, H_1(x_1), \ldots, H_1(x_{n_1}).$$

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

6.6 Dạng chính tắc Jordan

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 định lý dưới đây.

Định lý 6.6.1 (Jordan). *Cho $f \in \mathrm{End}_{\mathbb{K}}(V)$ sao cho $P_f(t)$ phân rã trên $\mathbb{K}$.*

1) Giả sử f chỉ có một trị riêng và

$$P_f(t) = (-1)^n(t-\lambda)^n, m_f(t) = (t-\lambda)^\beta, dimE(\lambda) = \gamma.$$

Khi đó, tồn tại một cơ sở $\mathcal{B}$ của V sao cho

$$[f]_{\mathcal{B}} = \begin{pmatrix} \boxed{J_1(\lambda)} & & & 0 \\ & \boxed{J_2(\lambda)} & & \\ & & \ddots & \\ 0 & & & \boxed{J_\gamma(\lambda)} \end{pmatrix} =: \widetilde{J}(\lambda),$$

trong đó:

a) $J_k(\lambda)$ là khối Jordan;

b) cấp của khối lớn nhất là β;

c) số các khối Jordan là γ.

2) Nếu f có các trị riêng khác nhau $\alpha_1, \dots, \alpha_p$ và

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

thì tồn tại một cơ sở $\mathcal{B}$ của V sao cho

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

Chứng minh. 1) Giả sử $P_f(t) = (-1)^n(t-\lambda)^n$. Vì $P_f(t)$ phân rã trên $\mathbb{K}$ nên f tam giác hóa được. Do đó tồn tại một cơ sở $\mathcal{B}'$ sao cho

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

Phân tích $A = \lambda I_n + N$, với $N = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} = [u]_{\mathcal{B}'}$. Lưu ý rằng $f = \lambda Id_V + u$. Theo Mệnh đề 6.5.4, u là toán tử lũy linh. Gọi β là bậc lũy linh của u. Vì ma trận biểu diễn toán tử λId_V trong mọi cơ sở đều là λI_n nên áp dụng Định lý 6.5.12, ta tìm được một cơ sở 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 $H_p(x)$ trong chứng minh Định lý 6.5.12 thì không gian con có số chiều lớn nhất chính là $H_\beta(x)$. Do đó cấp của khối Jordan lớn nhất là β. 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) Trước hết phân tích V thành tổng trực tiếp của các không gian đặc trưng, sau đó áp dụng Phần 1) đối với mỗi không gian đặc trưng. Như vậy định lý đã được chứng minh hoàn toàn. ■

Đối với toán tử tuyến tính f, 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.

Ví dụ 6.6.2. Cho ma trận

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

Xem A như ma trận biểu diễn toán tử tuyến tính $f \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^4)$. Ta sẽ tìm cơ sở của $\mathbb{R}^4$ trong đó f có dạng chính tắc Jordan.

Ta có $P_A(\lambda) = (\lambda - 1)^4$. Do $A - I_4 \neq 0$ và $(A - I_4)^2 = 0$ nên $m_A(\lambda) = (\lambda - 1)^2$. Ta tính được $\dim E(1) = 3$. Theo 1) của Định lý 6.6.1, dạng chính tắc Jordan A' của A gồm 3 khối Jordan, trong đó khối lớn nhất có cấp 2. Vậy

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

Gọi (u_1, u_2, u_3, u_4) là cơ sở trong đó f có dạng chính tắc nói trên. Ta thấy các vectơ u_1, u_2 và u_3 cần phải chọn sao cho chúng tạo thành cơ sở của không gian riêng $E(1)$. Hơn nữa u_3 cần phải được chọn sao cho ta có thể tìm được u_4 thỏa $f(u_3) = u_3 + u_4$ và các vectơ u_1, u_2, u_3, u_4 độc lập tuyến tính với nhau. Bằng cách tính toán trực tiếp ta có thể lấy các vectơ sau:

$$u_1 = (0, 1, 0, -2), u_2 = (1, 0, 0, -1), u_3 = (0, 0, 1, -2), u_4 = (1, 0, 0, 0).$$

Ví dụ 6.6.3. $A = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$

Xem A là ma trận biểu diễn toán tử tuyến tính $f \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^4)$. Ta có $P_f(\lambda) = (\lambda-1)(\lambda-2)^3$. Theo Hệ quả 6.2.6, $\mathbb{R}^4$ phân tích thành tổng trực tiếp của các không gian đặc trưng $N(1) = \mathrm{Ker}\,(f - Id)$ và $N(2) = \mathrm{Ker}\,(f - 2Id)^3$, nghĩa là

$$\mathbb{R}^4 = N(1) \oplus N(2).$$

Theo Định lý 6.4.2, nếu f_1 là hạn chế của f lên không gian đặc trưng $N(1)$ thì $P_{f_1}(\lambda) = \lambda - 1$. Tương tự, nếu f_2 là hạn chế của f lên không gian đặc trưng $N(2)$ thì $P_{f_2}(\lambda) = (\lambda - 2)^3$. Theo Định lý 6.6.1, Phần 2), tồn tại cơ sở trong đó dạng chính tắc Jordan của f là ma trận khối đường chéo gồm hai khối: khối thứ nhất là dạng chính tắc Jordan của f_1 và khối thứ hai là dạng chính tắc Jordan của f_2, mà các dạng chính tắc này được xác định bởi Định lý 6.6.1, Phần 1). Do $\dim E(2) = 2$, nên ma trận biểu diễn f_2 có hai khối Jordan. Do $\dim N(2) = 3$ nên ma trận biểu diễn f_2 có một khối Jordan cấp 1 và một khối Jordan cấp 2. Vậy dạng chính tắc Jordan của A là

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

Cơ sở cần tìm sẽ bao gồm một vectơ riêng u_1 ứng với trị riêng $\lambda_1 = 1$, hai vectơ riêng u_2, u_3 là các vectơ cơ sở của không gian riêng $E(2)$ ứng với trị riêng $\lambda_2 = 2$ và cuối cùng là vectơ u_4 được tính từ công thức $f(u_4) = u_3 + 2u_4$.

BÀI TẬP

Bài 6.1. Đối với mỗi ma trận dưới đây hãy đưa về dạng tam giác và chỉ rõ ma trận khả nghịch P làm tam giác hóa nó:

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

Bài 6.2. Giải hệ phương trình vi phân $\dfrac{dX}{dt} = AX$, với A là ma trận trong Bài 6.1.

Bài 6.3. Tìm đa thức tối tiểu của các ma trận trong Bài 6.1.

Bài 6.4. Giả sử toán tử tuyến tính $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ 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 6.5. Tìm dạng chính tắc Jordan của các ma trận

$$a) A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 1 \\ -2 & -5 & -4 & -1 \end{pmatrix};\ b) \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix};$$

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

Bài 6.6. Cho V là không gian vectơ n chiều trên $\mathbb{K}$ và f là toán tử tuyến tính trên V. Giả sử có $k \in \mathbb{N}$ sao cho $f^k = 0$. Chứng minh$f^n = 0$.

Bài 6.7. Tìm một ma trận vuông A cấp 3 có đa thức tối tiểu $P_A(t) = t^2$.

Bài 6.8. 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 của $\mathbb{R}^3$ sao cho trong đó ma trận biểu diễn của f có dạng chính tắc Jordan. Từ đó hãy chỉ ra một cơ sở cho mỗi không gian đặc trưng của f.

Bài 6.9. Tìm đa thức tối tiểu của ma trận $A = \begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}$. Ma trận A có chéo hóa được trên trường số thực $\mathbb{R}$ hay không?

Bài 6.10. Cho $A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$. Hãy tính A^n đối với mọi số nguyên dương n.

Bài 6.11. Tìm ma trận B sao cho B chéo hóa được trên $\mathbb{R}$ và $B^2 =$

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

Bài 6.12. Cùng một câu hỏi như trong Bài 6.9 cho các trường hợp sau:

a) $A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ 0 & 2 & 2 \end{pmatrix}$. b) $A = \begin{pmatrix} 6 & -6 & 5 \\ -4 & -1 & 10 \\ 7 & -6 & 4 \end{pmatrix}$.

Bài 6.13. Cho ma trận

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

a) Tìm đa thức đặc trưng của A.

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

c) Áp dụng Câu b) để tính A^n, với n là một số nguyên dương bất kỳ.

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 $\mathbb{R}$.

7.1 Tích vô hướng và không gian Euclid

Trong các chương trước chúng 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 chúng 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.

Định nghĩa 7.1.1. Cho V là không gian vectơ . Ánh xạ

$$\begin{array}{rccc} \langle , \rangle : & V \times V & \longrightarrow & \mathbb{R} \\ & (u, v) & \longmapsto & \langle u, v \rangle \end{array}$$

được gọi là một *tích vô hướng* trong V nếu $\forall\, u, v, w \in V, \forall\, \alpha, \beta \in \mathbb{R}$, 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) $\langle u, u \rangle \geq 0$, trong đó $\langle u, u \rangle = 0$ nếu và chỉ nếu $u = 0$.

Định nghĩa 7.1.2. Ta gọi một không gian vectơ hữu hạn chiều với tích vô hướng là một *không gian Euclid.*

Sau đây là một số ví dụ về các không gian Euclid.

Ví dụ 7.1.3. Tập hợp tất cả các vectơ trong không gian thực 3 chiều với tích vô hướng quen thuộc đã được định nghĩa trong các sách giáo khoa về toán sơ cấp là một không gian Euclid.

Ví dụ 7.1.4. Cho không gian vectơ $V = \mathbb{R}^n$, với $u = (x_1, \ldots, x_n)$ và $v = (y_1, \ldots, y_n)$ ta định nghĩa

$$\langle u, v \rangle := x_1 y_1 + \cdots + x_n y_n.$$

Khi đó V là không gian Euclid. Tích vô hướng vừa định nghĩa được gọi là *tích vô hướng chính tắc* trong $\mathbb{R}^n$.

Ví dụ 7.1.5. Với $u = (x_1, x_2, x_3), v = (y_1, y_2, y_3) \in \mathbb{R}^3$ ta định nghĩa

$$\langle u, v \rangle := x_1 y_1 + 2x_2 y_2 + 3x_3 y_3 + x_1 y_2 + x_2 y_1.$$

Dễ dàng thấy rằng các tính chất (i)-(iii) trong Định nghĩa 7.1.1 được thỏa mãn. Hơn nữa, những tính toán dưới đây cho thấy tính chất (iv) cũng được thỏa mãn.

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

Từ đó suy ra

$$\begin{aligned}\langle u, u \rangle = 0 &\Leftrightarrow x_1 + x_2 = x_2 = x_3 = 0 \\ &\Leftrightarrow x_1 = x_2 = x_3 = 0.\end{aligned}$$

Ví dụ 7.1.6. Xét không gian vectơ $M_2(\mathbb{R})$ gồm các ma trận vuông cấp 2 trên trường số thực $\mathbb{R}$. Khi đó $\langle A, B \rangle := Tr(A^\top B)$ là một tích vô hướng trong $M_2(\mathbb{R})$.

Ví dụ 7.1.7. Với các đa thức $P, Q \in \mathbb{R}[x]$, ta định nghĩa

$$\langle P, Q \rangle := \int_0^1 P(x) Q(x) dx.$$

Hiển nhiên các tính chất (i)-(iii) trong Định nghĩa 7.1.1 được thỏa mãn. Ta sẽ chứng tỏ tính chất (iv) cũng được thỏa mãn. Thật vậy, ta có $\langle P, P \rangle =$

$\int_0^1 P(x)^2 dx \geq 0$. Giả sử $\langle P, P \rangle = 0$. Vì $P(x)$ là một hàm liên tục và $P(x)^2 \geq 0$ nên từ điều kiện $\int_0^1 P(x)^2 dx = 0$ suy ra $P(x) = 0$ với $x \in [0, 1]$. Do đa thức $P(x)$ chỉ có thể có một số hữu hạn nghiệm nên từ đó suy ra $P(x) = 0$.

Ví dụ 7.1.8. Cho W là một không gian con của không gian vectơ V. Giả sử trong V có tích vô hướng $\langle , \rangle_V$. Với mọi $u, v \in W$, định nghĩa

$$\langle u, v \rangle_W := \langle u, v \rangle_V.$$

Dễ thấy đây là một tích vô hướng trong W.

Định nghĩa 7.1.9. Xét không gian Euclid V, ta nói *chuẩn* hay *độ dài* của vectơ u, ký hiệu $||u||$, là số thực $\sqrt{\langle u, u \rangle}$, nghĩa là $||u|| = \sqrt{\langle u, u \rangle}$. Nếu một vectơ có độ dài bằng 1 thì ta sẽ nói nó là một *vectơ đơn vị*.

Nhận xét 7.1.10. Cho $u \in V$ và $\lambda \in \mathbb{R}$, ta có

i) $||u|| \geq 0$. Hơn nữa $||u|| = 0 \Leftrightarrow u = 0$.

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

Ví dụ 7.1.11. a) Trong không gian Euclid ở Ví dụ 7.1.3 độ dài của các vectơ xác định như trong Định nghĩa 7.1.9 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 7.1.4 được xác định như sau:

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

c) Độ dài của vectơ $P(t)$ trong không gian ở Ví dụ 7.1.7 là

$$||P(x)|| = \sqrt{\int_a^b P(x)^2 dx}.$$

Bổ đề 7.1.12 (Bất đẳng thức Cauchy-Schwarz). *Với mọi $u, v \in V$ ta có*

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

Hơn nữa, dấu "$=$" xảy ra khi và chỉ khi u và v phụ thuộc tuyến tính.

Chứng minh. Nếu $||u|| = ||v|| = 0$ thì $u = v = 0$ và bất đẳng thức hiển nhiên được thỏa mãn.

Giả sử $||v|| \neq 0$ và $\lambda \in \mathbb{R}$ là một số thực bất kỳ. Ta có

$$\begin{aligned}
&||u + \lambda v||^2 \geq 0 \\
&\Rightarrow ||u||^2 + ||\lambda v||^2 + 2\langle u, \lambda v\rangle \geq 0 \\
&\Rightarrow \lambda^2 ||v||^2 + 2\lambda\langle u, v\rangle + ||u||^2 \geq 0.
\end{aligned}$$

Vế trái của bất đẳng thức sau cùng là một tam thức bậc hai theo λ. Để tam thức này luôn nhận giá trị không âm đối với mọi λ thì điều kiện cần và đủ là biệt số $\Delta' \leq 0$, nghĩa là

$$\langle u, v\rangle^2 - ||u||^2\,||v||^2 \leq 0$$

hay

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

Bây giờ, giả sử dấu " = " xảy ra, nghĩa là $\langle u, v\rangle^2 = ||u||^2\,||v||^2$. Khi đó tam thức bậc hai nói trên có nghiệm kép, nghĩa là tồn tại $\lambda \in \mathbb{R}$ sao cho

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

hay $||u + \lambda v||^2 = 0$. Từ đó suy ra $u + \lambda v = 0$ hay u và v là các vectơ phụ thuộc tuyến tính. ■

Mệnh đề 7.1.13 (Bất đẳng thức tam giác). *Với mọi $u, v \in V$ ta có*

$$||u + v|| \leq ||u|| + ||v||.$$

Hơn nữa, khi $u \neq 0$, dấu $=$ xảy ra khi và chỉ khi tồn tại $\lambda \geq 0$ sao cho $v = \lambda u$ hoặc $u = \lambda v$.

Chứng minh. Ta có

$$\begin{aligned}
||u + v||^2 &= ||u||^2 + ||v||^2 + 2\langle u, v\rangle \\
&\leq ||u||^2 + ||v||^2 + 2|\langle u, v\rangle| \\
&\leq ||u||^2 + ||v||^2 + 2||u||\,||v|| \quad \text{(bất đẳng thức C-S)} \\
&= (||u|| + ||v||)^2.
\end{aligned}$$

Suy ra $||u + v|| \leq ||u|| + ||v||$.

Nếu $v = \lambda u$, với $\lambda \geq 0$ thì ta có

$$\begin{aligned}
||u + v|| &= ||u + \lambda u|| = ||(1 + \lambda u)u|| \\
&= (1 + \lambda)||u|| = ||u|| + \lambda||u|| \\
&= ||u|| + ||\lambda u|| = ||u|| + ||v||.
\end{aligned}$$

Ngược lại, giả sử
$$||u+v|| = ||u|| + ||v||.$$
Khi đó
$$\begin{aligned}||u+v||^2 &= ||u||^2 + ||v||^2 + 2\langle u, v\rangle \\ &= ||u||^2 + ||v||^2 + 2||u||\,||v||\end{aligned}$$

Từ đó suy ra $\langle u, v\rangle = ||u||\,||v||$, kéo theo $\langle u, v\rangle^2 = ||u||^2\,||v||^2$. Theo Bổ đề 7.1.12, u và v phụ thuộc tuyến tính. Giả sử, chẳng hạn $u \neq 0$ và $v = \lambda u$. Khi đó từ bất đẳng thức C-S ta còn có $|\langle u, v\rangle| = ||u||\,||v||$, suy ra $\langle u, v\rangle = |\langle u, v\rangle|$. Thay $v = \lambda u$ vào đẳng thức cuối cùng, nhận được $\lambda\,||u|| = |\lambda|\,||u||$. Từ đó suy ra $\lambda \geq 0$. ■

Giả sử u và v là hai vectơ khác không của V. Áp dụng bất đẳng thức C-S, ta có
$$\frac{|\langle u, v\rangle|}{||u||\,||v||} \leq 1.$$
Từ đó suy ra tồn tại duy nhất một góc $\theta \in [0, \pi]$ sao cho
$$\cos\theta = \frac{\langle u, v\rangle}{||u||\,||v||} \leq 1.$$
Ta gọi θ là *góc* (không định hướng) giữa các vectơ u và v. Góc giữa vectơ 0 và một vectơ u bất kỳ được xem là tùy ý.

Cuối cùng, để kết thúc tiết này, lưu ý rằng tích vô hướng có thể được biểu diễn qua chuẩn bởi công thức dưới đây:
$$\langle u, v\rangle = \frac{1}{2}(||u+v||^2 - ||u||^2 - ||v||^2).$$

7.2 Sự trực giao

Định nghĩa 7.2.1. Cho V là một không gian Euclid với tích vô hướng $\langle , \rangle$.

(i) Ta nói các vectơ $u, v \in V$ *trực giao* với nhau, và viết $u \perp v$, nếu $\langle u, v\rangle = 0$.

(ii) Nếu $A \subseteq V$ là một tập con khác $\emptyset$ của V thì ta đặt
$$A^{\perp} := \{u \in V \mid \langle u, a\rangle = 0, \forall a \in A\}.$$

Khi đó $A^\perp$ là một không gian con của V và ta gọi $A^\perp$ là *không gian con trực giao* với A.

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

Bây giờ giả sử V là không gian vectơ trên trường K và V^* là không gian đối ngẫu của nó. Nếu W là không gian con của V thì đặt

$$W^0 := \{f \in V^* \mid f(v) = 0, \forall v \in W\}.$$

Dễ thấy W^0 là không gian con của V^* và ta gọi nó là *linh hóa tử* của W. Hiển nhiên, nếu $\{v_1, \ldots, v_p\}$ là một cơ sở của W thì

$$W^0 = \{f \in V^* \mid f(v_1) = \cdots = f(v_p) = 0\}.$$

Mệnh đề 7.2.2. *Nếu V là không gian vectơ hữu hạn chiều trên K và W là không gian con của V thì*

$$\dim V = \dim W + \dim W^0.$$

Chứng minh. Giả sử $\dim V = n$ và $\{v_1, \ldots, v_p\}$ là một cơ sở của W. Bổ túc thêm các vectơ của V vào tập hợp nói trên để nhận được một cơ sở của V:

$$\mathcal{B} = \{v_1, \ldots, v_p, v_{p+1}, \ldots, v_n\}.$$

Gọi $\mathcal{B}^* = \{\rho_1, \ldots, \rho_p, \rho_{p+1}, \ldots, \rho_n\}$ là cơ sở đối ngẫu của $\mathcal{B}$. Ta sẽ chứng minh $\{\rho_{p+1}, \ldots, \rho_n\}$ là cơ sở của W^0.

Với mọi $k \in \overline{p+1, n}$ ta có $\rho_k(v_1) = \cdots = \rho_k(v_p) = 0$, suy ra $\rho_k \in W^0$. Do $\rho_{p+1}, \ldots, \rho_n\}$ là các vectơ độc lập tuyến tính nên ta chỉ cần chứng minh chúng sinh ra W^0 là đủ. Vậy, xét $\forall f \in W^0$ và $\forall u \in V$. Ta có

$$u = x_1v_1 + \cdots + x_pv_p + x_{p+1}v_{p+1} + \cdots + x_nv_n.$$

Khi đó $f(u) = x_{p+1}f(v_{p+1}) + \cdots + x_nf(v_n)$. Đặt $\lambda_k = f(v_k)$, với mọi $k \in \overline{p+1, n}$, ta có

$$f(u) = \lambda_{p+1}\rho_{p+1}(u) + \cdots + \lambda_n\rho_n(u).$$

Từ đó suy ra $f = \lambda_{p+1}\rho_1 + \cdots + \lambda_n\rho_n$. ∎

Trở lại với không gian Euclid n chiều V. Như trên đã nhận xét, $V^* \simeq V$. Dưới đây ta sẽ xây dựng một đẳng cấu tự nhiên giữa V và V^*.

Mệnh đề 7.2.3. *Cho V là không gian Euclid với tích vô hướng $\langle , \rangle$. Ánh xạ*

$$\begin{array}{rccc} \sigma : & V & \longrightarrow & V^* \\ & v & \longmapsto & \sigma(v), \end{array}$$

trong đó

$$\begin{array}{rccc} \sigma(v) : & V & \longrightarrow & \mathbb{R}^* \\ & u & \longmapsto & \langle u, v\rangle \end{array}$$

là một đẳng cấu giữa V và V^*. Hơn nữa, nếu W là một không gian con của V thì $\sigma(W^{\perp}) = W^0$.

Chứng minh. Dễ dàng kiểm tra σ là một ánh xạ tuyến tính. Do $\dim(V) = \dim(V^*)$ nên để chứng minh σ là đẳng cấu ta chỉ cần chứng minh σ là đơn cấu là đủ. Vậy, giả sử $v \in V$ sao cho $\sigma(v) = 0$. Điều này có nghĩa là $\langle u, v\rangle = 0, \forall u \in V$. Nói riêng, lấy $u = v$ ta có $\langle v, v\rangle = 0$, kéo theo $v = 0$. Vậy σ là đơn cấu, kéo theo σ là đẳng cấu.

Tiếp theo ta có

$$\begin{aligned} \sigma^{-1}(W^0) &= \{v \in V \mid \sigma(v) \in W^0\} \\ &= \{v \in V \mid \sigma(v)(u) = 0, \forall u \in W\} \\ &= \{v \in V \mid \langle u, v\rangle = 0, \forall u \in W\} = W^{\perp}. \end{aligned}$$

Do σ là đẳng cấu nên từ đó suy ra $\sigma(W^{\perp}) = W^0$. ■

Hệ quả 7.2.4. *Nếu W là không gian con của không gian Euclid V thì*

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

Mệnh đề 7.2.5. *Nếu W là không gian con của không gian Euclid V thì*

(i) $V = W \oplus W^{\perp}$.

(ii) $W^{\perp\perp} := (W^{\perp})^{\perp} = W$.

Chứng minh. (i) Từ nhận xét rằng $W \cap W^{\perp} = 0$ và từ Hệ quả 7.2.4 suy ra ngay $V = W \oplus W^{\perp}$.

(ii) $\forall u \in W, \forall v \in W^{\perp}$ ta có $\langle u, v\rangle = 0$, suy ra $u \in W^{\perp\perp}$. Vậy $W \subseteq W^{\perp\perp}$. Áp dụng Hệ quả 7.2.4, ta có

$$\begin{aligned} \dim(W^{\perp\perp}) &= \dim(V) - \dim(W^{\perp}) \\ &= \dim(V) - (\dim(V) - \dim)(W) = \dim(W). \end{aligned}$$

Từ đó suy ra $\dim(W^{\perp\perp}) = \dim(W)$, kéo theo $W^{\perp\perp} = W$. ■

7.3 Cơ sở trực giao và cơ sở trực chuẩn. Quá trình trực giao hóa Gram-Schmidt

Định nghĩa 7.3.1. Cho V là không gian Euclid n chiều và $\mathcal{B} = (e_1, \dots, e_n)$ là một cơ sở của V.

(i) Ta nói $\mathcal{B}$ là *cơ sở trực giao* nếu $\langle e_i, e_j \rangle = 0, \forall i \neq j$.

(ii) Ta nói $\mathcal{B}$ là *cơ sở trực chuẩn* nếu $\langle e_i, e_j \rangle = \delta_{ij}, \forall i \neq j$, trong đó δ_{ij} là số Kronecker.

Hiển nhiên nếu $(e_1, \dots, e_n)$ là cơ sở trực giao thì $(\frac{e_1}{||e_1||}, \dots, \frac{e_n}{||e_n||})$ là cơ sở trực chuẩn.

Mệnh đề 7.3.2. *Cho $\mathcal{B} = (e_1, \dots, e_n)$ là một cơ sở trực chuẩn và u là một vectơ bất kỳ của không gian Euclid V. Khi đó ta có*

$$u = \langle u, e_1 \rangle e_1 + \cdots + \langle u, e_n \rangle e_n.$$

Chứng minh. Giả sử $u = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$. Khi đó

$$\langle u, e_i \rangle = x_1 \langle e_1, e_i \rangle + x_2 \langle e_2, e_i \rangle + \cdots + x_n \langle e_n, e_i \rangle.$$

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

Định lý 7.3.3. *Trong một không gian Euclid bất kỳ luôn tồn tại các cơ sở trực chuẩn.*

Chứng minh. Do nhận xét phía trên nên ta chỉ cần chứng minh sự tồn tại cơ sở trực giao là đủ. Điều này sẽ được chứng minh bằng qui nạp theo n. Nếu $n = 1$ thì không có điều gì để chứng minh. Giả sử điều khẳng định là đúng cho những không gian số chiều bé hơn n. Xét một vectơ $0 \neq v \in V$ và đặt $W = \langle v \rangle^{\perp}$. Khi đó $V = \langle v \rangle \oplus W$ và $\dim(W) = n - 1$. Theo giả thiết qui nạp trong W ta tìm được cơ sở trực giao, chẳng hạn $(u_1, \dots, u_{n-1})$. Đặt $u_n = v$, hiển nhiên ta có một cơ sở trực giao của V là $(u_1, \dots, u_{n-1}, u_n)$. ■

Mệnh đề dưới đây cho ta một điều kiện cần và đủ để một cơ sở là trực chuẩn.

Mệnh đề 7.3.4. *Cho $\mathcal{B} = (e_1, \dots, e_n)$ là một cơ sở của không gian Euclid V. Khi đó, $\mathcal{B}$ là cơ sở trực chuẩn nếu và chỉ nếu đối với mọi vectơ u, v của*

V ta có

$$\langle u, v\rangle = \sum_{i=1}^{n} x_i y_i,$$

trong đó $[u]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, [v]_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ *là tọa độ của* u *và* v *theo cơ sở* $\mathcal{B}$.

Chứng minh. Giả sử $\mathcal{B}$ là cơ sở trực chuẩn, ta có

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

Để đơn giản hóa, trong trường hợp này, ta có thể viết lại

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

Điều ngược lại là hiển nhiên. ■

Từ mệnh đề vừa chứng minh ta suy ra ngay hệ quả sau:

Hệ quả 7.3.5. *Cho* $\mathcal{B} = (e_1, \ldots, e_n)$ *là cơ sở trực chuẩn trong không gian Euclid* V. *Khi đó ta có phép đẳng cấu sau đây giữa* V *và không gian Euclid* $\mathbb{R}^n$ *với tích vô hướng chính tắc:*

$$\begin{array}{rccc} \varphi_{\mathcal{B}}: & V & \longrightarrow & \mathbb{R}^n \\ & u = \sum_{i=1}^{n} x_i e_i & \longmapsto & (x_1, \ldots, x_n). \end{array}$$

Từ công thức $V = W \oplus W^{\perp}$ trong Mệnh đề 7.2.5 suy ra mỗi vectơ $u \in V$ đều viết được một cách duy nhất dưới dạng $u = x_0 + v$, trong đó $x_0 \in W$ và $v \in W^{\perp}$. Ta gọi x_0 là *hình chiếu trực giao* của u lên W và ký hiệu là $x_0 = pr_W(u)$. 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 7.3.6. *Cho* V *là không gian Euclid và* W *là một không gian con của* V. *Giả sử* $(e_1, \ldots, e_m)$ *là một cơ sở trực chuẩn của* W *và* u *là một vectơ bất kỳ của* V. *Khi đó ta có công thức*

$$pr_W(u) = \langle u, e_1\rangle e_1 + \cdots + \langle u, e_m\rangle e_m.$$

Chứng minh. Gọi $(e_{m+1}, \ldots, e_n)$ là một cơ sở trực chuẩn của phần bù trực giao $W^{\perp}$. Khi đó, theo Mệnh đề 7.2.5 ta có $(e_1, \ldots, e_m, e_{m+1}, \ldots, e_n)$ là một cơ sở trực chuẩn của V. Áp dụng Hệ quả 7.3.2, nhận được

$$u = (\langle u, e_1\rangle e_1 + \cdots + \langle u, e_m\rangle e_m) + (\langle u, e_{m+1}\rangle e_{m+1} + \cdots + \langle u, e_n\rangle e_n).$$

Lưu ý rằng $\langle u, e_1\rangle e_1 + \cdots + \langle u, e_m\rangle e_m \in W$ và

$$\langle u, e_{m+1}\rangle e_{m+1} + \cdots + \langle u, e_n\rangle e_n \in W^{\perp}.$$

Do đó áp dụng Mệnh đề 7.3.2, suy ra

$$pr_W(u) = \langle u, e_1\rangle e_1 + \cdots + \langle u, e_m\rangle e_m.$$ ■

Quá trình trực giao hóa Gram-Schmidt

Qua Hệ quả 7.3.5 ta thấy rằng có thể đồng nhất một không gian Euclid n chiều V với không gian $\mathbb{R}^n$ cùng tích vô hướng chính tắc. Tuy nhiên khi đó cần phải xây dựng được trong V một cơ sở trực chuẩn. Dưới đây ta sẽ mô tả một thuật toán cho phép nhận được một cơ sở trực giao từ một cơ sở bất kỳ của V (như đã nói phía trên, từ một cơ sở trực giao ta dễ dàng nhận được cơ sở trực chuẩn). Một thuật toán như vậy thường được gọi là *quá trình trực giao hóa Gram-Schmidt*.

Định lý 7.3.7. *Cho $(v_1, \ldots, v_p)$ là một họ các vectơ độc lập tuyến tính của không gian Euclid V và $W = \langle v_1, \ldots, v_p\rangle$ là không gian con của V sinh bởi các vectơ nói trên. Khi đó, từ các vectơ $v_1, \ldots, v_p$ ta có thể xây dựng một cơ sở trực chuẩn cho W.*

Nói riêng, từ một cơ sở bất kỳ của V ta có thể xây dựng được một cơ sở trực chuẩn của V.

Chứng minh. Như đã nhận xét ở trên, ta chỉ cần xây dựng một cơ sở trực giao $(u_1, \ldots, u_p)$ cho W là đủ.

Đặt $u_1 := v_1$

$u_2 := v_2 + \lambda_1 u_1$, với $\lambda_1 \in \mathbb{R}$ sao cho $u_2 \perp u_1$.

Với điều kiện này ta có

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

Do $u_1 \neq 0$ nên từ đó suy ra

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

Tiếp theo, tìm u_3 dưới dạng

$u_3 := v_3 + \lambda_1 u_1 + \lambda_2 u_2$, với $\lambda_1, \lambda_2 \in \mathbb{R}$ sao cho $u_3 \perp u_1$ và $u_3 \perp u_2$.

Tìm λ_1 như sau:

$$\begin{aligned} 0 = \langle u_3, u_1 \rangle &= \langle v_3 + \lambda_1 u_1 + \lambda_2 u_2, u_1 \rangle \\ &= \langle v_3, u_1 \rangle + \lambda_1 ||u_1||^2 \quad (\text{do } \langle u_2, u_1 \rangle = 0). \end{aligned}$$

Từ đó suy ra $\lambda_1 = -\frac{\langle v_3, u_1 \rangle}{||u_1||^2}$. Hoàn toàn tương tự, nhận được $\lambda_2 = -\frac{\langle v_3, u_2 \rangle}{||u_2||^2}$.

Giả sử đã tìm được các vectơ trực giao $u_1, \ldots, u_{p-1}$. Ta sẽ tìm vectơ u_p dưới dạng sau

$$u_p = v_p + \lambda_1 u_1 + \cdots + \lambda_{p-1} u_{p-1}.$$

Từ điều kiện $u_p \perp u_i$ ta tìm được $\lambda_i = -\frac{\langle v_p, u_i \rangle}{||u_i||^2}$. 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 $v_1, \ldots, v_i$. Theo cách xây dựng thì u_{i+1} là tổ hợp tuyến tính của các vectơ $v_{i+1}, u_1, \ldots, u_i$, do đó u_{i+1} cũng là tổ hợp tuyến tính của các vectơ $v_{i+1}, v_1, \ldots, v_i$. Ta đã chứng minh

$$\langle u_1, \ldots, u_{i+1} \rangle \subseteq \langle v_1, \ldots, v_{i+1} \rangle.$$

Hoàn toàn tương tự ta cũng có

$$\langle v_1, \ldots, v_{i+1} \rangle \subseteq \langle u_1, \ldots, u_{i+1} \rangle. \qquad \blacksquare$$

Ví dụ 7.3.8. Trong không gian Euclid $\mathbb{R}^4$ với tích vô hướng chính tắc cho vectơ $u = (1, 2, 0, 3)$ và cho không gian con W được sinh ra bởi các vectơ

$$v_1 = (1, 1, 0, 0), v_2 = (1, 0, -1, 1), v_3 = (0, 1, 1, 1).$$

Hãy tìm hình chiếu trực giao của u lên W.

Giải. Trước hết ta sẽ dùng quá trình trực giao hóa Gram-Schmidt để xây dựng một cơ sở trực chuẩn cho W, sau đó áp dụng công thức trong Mệnh đề 7.3.6 để tính hình chiếu trực giao của u lên W. 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 $u_1 := v_1$

$u_2 := v_2 + \lambda_1 u_1$, với $\lambda_1 = -\frac{\langle v_2, u_1\rangle}{||u_1||^2} = -\frac{1}{2}$. Từ đó

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

Nhận xét rằng nếu ta thay u_2 bởi $u_2' = \alpha u_2, \alpha \in \mathbb{R}$ thì các vectơ u_1 và u_2' vẫn trực giao với nhau. Do đó ta có thể lấy $u_2 = (1, -1, -2, 2)$. Bây giờ tìm u_3 dưới dạng

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

với $\lambda_1 = -\frac{\langle v_2, u_1\rangle}{||u_1||^2} = -\frac{1}{2}$ và $\lambda_2 = -\frac{\langle v_3, u_2\rangle}{||u_2||^2} = -\frac{1}{10}$. Do đó $u_3 = \frac{2}{5}(-1, 1, 2, 3)$. Tuy nhiên ta có thể lấy $u_3 = (-1, 1, 2, 3)$. Trực chuẩn hóa cơ sở (u_1, u_2, u_3) ta nhận được cơ sở trực chuẩn của W như sau

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

Ta có

- $\langle u, e_1\rangle e_1 = \frac{1}{\sqrt{2}}(1, 2, 0, 0)\frac{1}{\sqrt{2}}(1, 1, 0, 0) = \frac{1}{2}(1, 2, 0, 0)$,
- $\langle u, e_2\rangle e_2 = \frac{1}{\sqrt{10}}(1, -2, 0, 6)\frac{1}{\sqrt{10}}(1, -1, -2, 2) = \frac{1}{10}(1, 2, 0, 12)$,
- $\langle u, e_2\rangle e_2 = \frac{1}{\sqrt{15}}(-1, 2, 0, 9)\frac{1}{\sqrt{15}}(-1, 1, 2, 3) = \frac{1}{15}(1, 2, 0, 27)$.

Vậy hình chiếu trực giao của u lên W là

$$pr_W(u) = \langle u, e_1\rangle e_1 + \langle u, e_2\rangle e_2 + \langle u, e_3\rangle e_3 = \frac{1}{30}(20, 40, 0, 90).$$

7.4 Khoảng cách trong không gian Euclid

Định nghĩa 7.4.1. Cho u và v là hai vectơ trong không gian Euclid V. Số thực không âm $||u - v||$ được gọi là *khoảng cách* giữa các vectơ u và v và được ký hiệu là $d(u, v)$. Vậy

$$d(u, v) = ||u - v||.$$

Bổ đề 7.4.2. *Đối với mọi vectơ u, v trong không gian Euclid V ta có những khẳng định sau đây:*

(i) $d(u, v) = 0$ *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 ngay từ định nghĩa khoảng cách và chuẩn.

(iii) Áp dụng bất đẳng thức tam giác ta có

$$||u - w|| = ||(u - v) + (v - w)|| \leq ||u - v|| + ||v - w||.$$

Từ đó suy ra điều cần chứng minh. ■

Định nghĩa 7.4.3. Cho W là một không gian con của không gian Euclid V và u là một vectơ của V. Ta gọi khoảng cách giữa u và hình chiếu trực giao của nó lên W là *khoảng cách từ u đến W* và ký hiệu là $d(u, W)$. Vậy

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

Mệnh đề 7.4.4. *Khoảng cách từ một vectơ đến một không gian con là khoảng cách ngắn nhất (nhỏ nhất) từ vectơ ấy đến các vectơ của không gian con đã cho.*

Chứng minh. Giả sử u là một vectơ và W là một không gian con của không gian Euclid V. Đặt $w = pr_W(u)$, ta cần chứng minh $||u - v|| \geq ||u - w||, \forall v \in W$. Ta có

$$\begin{aligned}
& ||u - v|| \geq ||u - w|| \\
\Leftrightarrow\ & ||u - v||^2 \geq ||u - w||^2 \\
\Leftrightarrow\ & ||u||^2 + ||v||^2 - 2\langle u, v\rangle \geq ||u||^2 + ||w||^2 - 2\langle u, w\rangle \\
\Leftrightarrow\ & ||v||^2 - 2\langle w, v\rangle \geq ||w||^2 - 2\langle w, w\rangle \\
\Leftrightarrow\ & ||v||^2 - (||w + v||^2 - ||w||^2 - ||v||^2) \geq -||w||^2 \\
\Leftrightarrow\ & 2||v||^2 + 2||w||^2 - ||w + v||^2 \geq 0.
\end{aligned}$$

Theo bất đẳng thức tam giác ta có

$$||w + v|| \leq ||w|| + ||v||$$

hay

$$||w + v||^2 \leq ||w||^2 + ||v||^2 + 2||w||\,||v||.$$

Từ đó suy ra

$$\begin{aligned}
2||v||^2 + 2||w||^2 - ||v + w||^2 &\geq 2||v||^2 + 2||w||^2 \\
&\quad - (||v||^2 + ||w||^2 + 2||v||\,||w||) \\
&= (||v|| - ||w||)^2 \geq 0.
\end{aligned}$$

■

7.5 Ma trận biểu diễn của tích vô hướng

Giả sử V là không gian Euclid với tích vô hướng $\langle , \rangle$ và $\mathcal{B} = (e_1, \dots, e_n)$ là một cơ sở của V. Với các vectơ $u = \sum_{i=1}^n x_i e_i$ và $v = \sum_{i=1}^n y_i e_i$ bất kỳ trong V, ta có

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

Đặt $a_{ij} = \langle e_i, e_j\rangle$ và gọi ma trận vuông $A = (a_{ij})$ cấp n là ma trận biểu diễn tích vô hướng $\langle , \rangle$ theo cơ sở $\mathcal{B}$. Ta dùng ký hiệu $\langle , \rangle_{\mathcal{B}}$ để chỉ ma trận nói trên.

Ví dụ 7.5.1. Xét tích vô hướng trong Ví dụ 7.1.5. 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

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

Khi đó ma trận biểu diễn tích vô hướng này theo cơ sở chính tắc là

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

Mệnh đề 7.5.2. *Cho V là một 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. ■

Mệnh đề 7.5.3. *Trong không gian Euclid V, cơ sở $\mathcal{B} = (e_1, \dots, e_n)$ là trực chuẩn nếu và chỉ nếu $\langle , \rangle_{\mathcal{B}} = I_n$.*

Chứng minh. $\mathcal{B}$ là cơ sở trực chuẩn $\Leftrightarrow \langle u_i, u_j\rangle = \delta_{ij} \Leftrightarrow \langle , \rangle_{\mathcal{B}} = I_n$. ■

Trong Mệnh đề 7.5.2 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 $\mathcal{B}'$ nào khác thì ma trận $\langle , \rangle_{\mathcal{B}'}$ biểu diễn tích vô hướng $\langle , \rangle$ theo cơ sở $\mathcal{B}'$ 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.

Mệnh đề 7.5.4. *Cho V là một 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 7.5.2 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

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

7.6 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 $\mathbb{K}$, với $\mathbb{K}$ là trường số thực $\mathbb{R}$ hoặc trường số phức $\mathbb{C}$. Vì không gian Euclid trước hết cũng là một 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 bình thường không có. Trong mục này và mục tiếp theo ta sẽ đề cập đến những tính chất như vậy.

Định nghĩa 7.6.1. Ta nói toán tử f trong không gian Euclid V là một *toán* *tử đối xứng* (hay *toán tử tự liên hợp*) nếu

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

Để thấy ý nghĩa của khái niệm toán tử đối xứng, ta hãy xét một cơ sở trực chuẩn $\mathcal{B} = (e_1, \ldots, e_n)$ của V. Khi đó, theo Mệnh đề 7.3.4, công thức trong Định nghĩa 7.6.1 được viết lại dưới dạng ma trận sau:

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

hay

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

Lần lượt cho u và v là các vectơ của cơ sở $\mathcal{B}$, ta có $[f]_{\mathcal{B}}^\top = [f]_{\mathcal{B}}$. Vậy f là toán tử đối xứng khi và chỉ khi ma trận biểu diễn f theo một cơ sở trực chuẩn (do đó trong mọi cơ sở trực chuẩn) là ma trận đối xứng.

Định lý 7.6.2. *Cho f là toán tử đối xứng trong không gian Euclid. Khi đó ta có:*

(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 với nhau.

Chứng minh. (i) Gọi A là ma trận biểu diễn toán tử f theo một cơ sở trực chuẩn. Khi đó, theo nhận xét phía trên thì A là ma trận đối xứng thực. Xét đa thức đặc trưng $P_A(t)$ và gọi λ là một nghiệm bất kỳ trong $\mathbb{C}$ của nó. Khi đó hệ phương trình tuyến tính thuần nhất

$$(A - \lambda I_n)X = 0 \tag{1}$$

với $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in M_{n\times 1}(\mathbb{C})$ có nghiệm không tầm thường. Đặt $\overline{X} = \begin{pmatrix} \overline{x_1} \\ \overline{x_2} \\ \vdots \\ \overline{x_n} \end{pmatrix}$, trong đó $\overline{x_i}$ là số phức liên hợp của x_i. Khi đó

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

Ta có

$$AX = \lambda X \Rightarrow \overline{AX} = \overline{\lambda X} \Rightarrow A\overline{X} = \overline{\lambda X}. \tag{2}$$

Vì $A^\top = A$ nên ta có

$$(AX)^\top \overline{X} = X^\top A^\top \overline{X} = X^\top (A\overline{X}). \tag{3}$$

Kết hợp (1), (2) và (3), nhận được

$$(\lambda X)^\top \overline{X} = X^\top (\overline{\lambda X}).$$

Từ đó suy ra

$$\lambda(X^\top \overline{X}) = \overline{\lambda}(X^\top \overline{X}).$$

Nhưng như ta đã thấy $X^\top \overline{X}$ là một số thực dương nên từ đẳng thức cuối cùng suy ra $\lambda = \overline{\lambda}$, nghĩa là $\lambda \in \mathbb{R}$.

(ii) Ta sẽ chứng minh điều khẳng định bằng qui nạp theo số chiều n của V rằng tồn tại trong V một cơ sở gồm toàn các vectơ riêng. Nếu $n = 1$ thì không có gì để chứng minh. Vậy, giả sử $n > 1$ và điều khẳng định đúng đối với những 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 f là không gian con bất biến đối với f nên $\tilde{f} := f|_H$ là toán tử tuyến tính trong không gian Euclid H. Hơn nữa, hiển nhiên $\tilde{f}$ cũng là toán tử đối xứng. Vậy, theo giả thiết qui nạp, toán tử $\tilde{f}$ 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 ứng với λ, v là vectơ riêng ứng với μ. Khi đó

$$\begin{aligned} & \langle f(u), v \rangle = \langle u, f(v) \rangle \\ \Rightarrow\ & \langle \lambda u, v \rangle = \langle u, \mu v \rangle \\ \Rightarrow\ & \lambda \langle u, v \rangle = \mu \langle u, v \rangle \\ \Rightarrow\ & (\lambda - \mu)\langle u, v \rangle = 0 \\ \Rightarrow\ & \langle u, v \rangle = 0. \end{aligned}$$

■

Cuối cùng lưu ý rằng một ma trận phức đối xứng không nhất thiết chéo hóa được trên $\mathbb{R}$ hoặc thậm chí trên $\mathbb{C}$. Sau đây là một ví dụ minh họa.

Ví dụ 7.6.3. Cho $A = \begin{pmatrix} 0 & \alpha \\ \alpha & \beta \end{pmatrix}$ với $\alpha, \beta \in \mathbb{C}$. Ta có

$$P_A(\lambda) = |A - \lambda I_2| = \begin{vmatrix} -\lambda & \lambda \\ \lambda & \beta - \lambda \end{vmatrix} = \lambda^2 - \beta\lambda - \alpha^2.$$

Nếu $\Delta = \beta^2 + 4\alpha^2 = 0$ (là điều có thể xảy ra đối với các số phức α và β) thì $P_A(\lambda)$ có một nghiệm kép $\lambda = \frac{\beta}{2}$. Khi đó $P_A(\lambda) = \left(\lambda - \frac{\beta}{2}\right)^2$. Vậy A chéo hóa được khi và chỉ khi $m_A(\lambda) = \lambda - \frac{\beta}{2}$. Nghĩa là $m_A(A) = A - \frac{\beta}{2}I_2 = 0$ hay $A = \frac{\beta}{2}I_2$, nhưng điều này mâu thuẫn. Vậy A không chéo hóa được.

7.7 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 7.7.1. 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

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

Mệnh đề 7.7.2. *Cho f là toán tử tuyến tính trên không gian Euclid V. Khi đó những điều sau tương đương:*

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

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

(iii) Nếu $\mathcal{B}$ là một cơ sở trực chuẩn của V. Đặt $A = [f]_{\mathcal{B}}$ thì $A^\top A = I_n = AA^\top$. Nói riêng, A là ma trận khả nghịch và $\det A = \pm 1$.

Chứng minh. (i) $\Rightarrow$ (ii). Chỉ việc cho $u = v$.

(ii) $\Rightarrow$ (i). Ta có

$$\begin{aligned}\langle f(u), f(v)\rangle &= \frac{1}{2}(||f(u) + f(v)||^2 - ||f(u)||^2 - ||f(v)||^2)\\ &= \frac{1}{2}(||f(u+v)||^2 - ||f(u)||^2 - ||f(v)||^2)\\ &= \frac{1}{2}(||u+v||^2 - ||u||^2 - ||v||^2) = \langle u, v\rangle.\end{aligned}$$

(i) $\Leftrightarrow$ (iii). Vì $\mathcal{B}$ là cơ sở trực chuẩn nên với mọi $u, v \in V$ ta có

$$\begin{aligned}
&\langle f(u), f(v)\rangle = \langle u, v\rangle \\
\Leftrightarrow\ & [f(u)]_{\mathcal{B}}^{\top}[f(v)]_{\mathcal{B}} = [u]_{\mathcal{B}}^{\top}[v]_{\mathcal{B}} \\
\Leftrightarrow\ & ([f]_{\mathcal{B}}[u]_{\mathcal{B}})^{\top}([f]_{\mathcal{B}}[v]_{\mathcal{B}}) = [u]_{\mathcal{B}}^{\top}[v]_{\mathcal{B}} \\
\Leftrightarrow\ & [u]_{\mathcal{B}}^{\top}([f]_{\mathcal{B}}^{\top}[f]_{\mathcal{B}})[v]_{\mathcal{B}} = [u]_{\mathcal{B}}^{\top}[v]_{\mathcal{B}} \\
\Leftrightarrow\ & [f]_{\mathcal{B}}^{\top}[f]_{\mathcal{B}} = I_n \text{ hay} A^{\top}A = I_n.
\end{aligned}$$

Phần khẳng định còn lại của (iii) là hiển nhiên. ■

Hệ quả 7.7.3. *Nếu f là một toán tử trực giao thì* $\det f = \pm 1$. *Nói riêng, f là một tự đẳng cấu.*

Mệnh đề 7.7.4. *Cho f là toán tử tuyến tính trên V. Khi đó f là một toán tử trực giao khi và chỉ khi nó biến một cơ sở trực chuẩn thành một cơ sở trực chuẩn. Để f thỏa tính chất nói trên thì điều kiện đủ là tồn tại một cơ sở trực chuẩn sao cho f biến nó thành một cơ sở trực chuẩn.*

Chứng minh. Giả sử f là một toán tử trực giao. Theo Hệ quả 7.7.3, f là một tự đẳng cấu, do đó f biến cơ sở thành cơ sở. Nếu $\mathcal{B} = (e_1, \ldots, e_n)$ là cơ sở trực chuẩn thì

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

Vậy $(f(e_1), \ldots, f(e_n))$ cũng là cơ sở trực chuẩn.

Ngược lại, giả sử tồn tại cơ sở trực chuẩn $\mathcal{B} = (e_1, \ldots, e_n)$ sao cho $(f(e_1), \ldots, f(e_n))$ cũng là cơ sở trực chuẩn. Xét các vectơ $u, v \in V$:

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

Do $(e_1, \ldots, e_n)$ và $(f(e_1), \ldots, f(e_n))$ là các cơ sở trực chuẩn nên ta có

$$\begin{aligned}
\langle f(u), f(v)\rangle &= \langle \sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\rangle \\
&= \sum_{i,j=1}^{n} x_i y_j \langle f(e_i), f(e_j)\rangle \\
&= \sum_{i,j=1}^{n} x_i y_j \delta_{ij} = \sum_{i=1}^{n} x_i y_i = \langle u, v\rangle.
\end{aligned}$$

Vậy f là toán tử trực giao. ■

Định nghĩa 7.7.5. Tập hợp

$$O(n, \mathbb{R}) := \{A \in M_n(\mathbb{R}) \| A^\top A = I_n\}$$

là một nhóm đối với phép nhân và được gọi là *nhóm trực giao*.

Mỗi ma trận $A \in O(n, \mathbb{R})$ được gọi là một *ma trận trực giao*. Mỗi ma trận trực giao đều biểu diễn một toán tử trực giao theo một cơ sở trực chuẩn của một không gian Euclid.

Ví dụ 7.7.6. Ma trận

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

là ma trận trực giao. Ta có thể kiểm tra điều này bằng cách thực hiện phép nhân ma trận

$$A^\top A = I_3.$$

Nhưng ta cũng có thể kiểm tra 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 $f(e_i) = u_i, i \in \{1, 2, 3\}$. Vì f biến cơ sở trực chuẩn thành cơ sở trực chuẩn nên theo Mệnh đề 7.7.4, f là *toán tử trực giao*. Mà A là ma trận biểu diễn f theo cơ sở trực chuẩn nên A là ma trận trực giao.

Mệnh đề 7.7.7. *Ma trận chuyển cơ sở từ một cơ sở trực chuẩn sang một cơ sở trực chuẩn là một ma trận trực giao.*

Chứng minh. Giả sử $\mathcal{B} = (e_1, \ldots, e_n)$ và $\mathcal{B}' = (e'_1, \ldots, e'_n)$ là hai cơ sở trực chuẩn. Gọi f là toán tử thỏa $f(e_i) = e'_i, \forall i \in \overline{1, n}$. Theo Mệnh đề 7.7.4, 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. ■

BÀI TẬP

Bài 7.1. 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 7.2. Xét không gian Euclid $\mathbb{R}^n$ với tích vô hướng chính tắc. Chứng minh rằng với mọi $u = (x_1, \ldots, x_n) \in \mathbb{R}^n$, ta có

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

Bài 7.3. 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}. Qua đó hãy chứng minh rằng $|Tr(A)| \leq \sqrt{nTr(AA^\top)}$.

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 7.4. Xét không gian Euclid $\mathbb{R}^3$ với tích vô hướng chính tắc.

a) Cho P là mặt phẳng trong $\mathbb{R}^3$ 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.

b) Cho các vectơ $u_1 = (1, 0, 1), u_2 = (2, 1, 0)$ và $u_3 = (1, 1, 1)$. Chứng minh rằng $\mathcal{B} = (u_1, u_2, u_3)$ là một cơ sở của $\mathbb{R}^3$. Xét xem $\mathcal{B}$ có phải là cơ sở trực chuẩn không. Nếu $\mathcal{B}$ không phải là cơ sở trực chuẩn thì hãy sử dụng quá trình trực chuẩn hóa Gram-Schmidt để xây dựng từ $\mathcal{B}$ một cơ sở trực chuẩn $\mathcal{B}' = (e_1, e_2, e_3)$.

c) Cho các vectơ $v_1 = (1, 0, 1), v_2 = (2, 1, 2)$ và $v_3 = (1, 1, 1)$. Hãy tìm số chiều và một cơ sở trực chuẩn cho không gian con sinh bởi các vectơ v_1, v_2, v_3.

Bài 7.5. Với $n \geq 0$, xét tích phân suy rộng

$$I_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^n e^{-\dfrac{u^2}{2}}\, dx.$$

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

b) Chứng minh công thức truy hồi

$$I_n = (n-1)I_{n-2}, \forall n \geq 2.$$

Áp dụng để tính I_{2k}.

c) Định nghĩa ánh xạ

$$\langle , \rangle : \mathbb{R}[u] \times \mathbb{R}[u] \longrightarrow \mathbb{R}$$

như sau:

$$\forall P, Q \in \mathbb{R}[u], \langle P, Q\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\dfrac{u^2}{2}} P(u)Q(u)dx.$$

Chứng minh rằng ánh xạ nói trên là một tích vô hướng.

d) Xét không gian con $\mathbb{R}_2[u]$ của $\mathbb{R}[u]$. Hãy tính khoảng cách từ u^3 đến $\mathbb{R}_2[u]$.

Bài 7.6. Xét ánh xạ $\langle , \rangle : \mathbb{R}^4 \times \mathbb{R}^4 \longrightarrow \mathbb{R}$ cho bởi:

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

a) Chứng minh rằng ánh xạ này 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 con của $\mathbb{R}^4$ xác định bởi hệ phương trình tuyến tính thuần nhất

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

Hãy tìm một cơ sở của $W^{\perp}$.

Bài 7.7. Trong không gian Euclide với tích vô hướng chính tắc cho các vectơ $u_1 = (2, 1, -2, 4), u_2 = (-2, 1, -1, -6), u_3 = (-2, 3, -4, -8)$. Gọi $W = \langle u_1, u_2, u_3\rangle$ là không gian con của $\mathbb{R}^4$ sinh ra bởi các vectơ u_1, u_2, u_3 và $W^{\perp}$ là không gian con của $\mathbb{R}^4$ trực giao với W.

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 khoảng cách $d(u, W)$ từ u đến W.

Bài 7.8. Cho $A \in O(n, \mathbb{R})$. Chứng minh rằng, nếu $\det A = 1$ thì mỗi phần tử a_{ij} của A đều bằng phần bù đại số của nó.

Bài 7.9. Cho f là một phép biến đổi trực giao trong không gian Euclid V.

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

b) Chứng minh rằng, nếu $(f - Id_V)^2 = 0$ thì $f = Id_V$.

Bài 7.10. Xây dựng một cơ sở trực chuẩn của $\mathbb{R}^3$ từ các vectơ riêng của toán tử $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ có ma trận biểu diễn theo cơ sở chính tắc là

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

Bài 7.11. Toán tử $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ có ma trận theo cơ sở chính tắc là

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

Hãy chứng minh rằng f là toán tử trực giao trong không gian Euclid $\mathbb{R}^3$ với tích vô hướng chính tắc.

Chương 8

KHÔNG GIAN UNITA

8.1 Khái niệm về không gian Unita

Định nghĩa 8.1.1. *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

$$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (u, v) & \longmapsto & \langle u, v \rangle \end{array}$$

thỏa các tính chất sau: Với mọi $u, v, u_1, u_2 \in V, \alpha \in \mathbb{C}$ ta có

(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 8.1.2. 1) Với ký hiệu như trong Định nghĩa 8.1.1, ta suy ra ngay 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) $\langle u, \alpha v \rangle \in \overline{\alpha} \, \langle u, v \rangle$

2) Một không gian vectơ con của một không gian Unita cũng là một không gian Unita.

Ví dụ 8.1.3. $\mathbb{C}^n$ là không gian Unita với tích vô hướng định bởi

$$\langle u, v \rangle = x_1 \overline{y_1} + \cdots + x_n \overline{y_n}$$

với mọi $u = (x_1, \ldots, x_n)$, $v = (y_1, \ldots, y_n) \in \mathbb{C}^n$. 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 8.1.4. Không gian các đa thức hệ số phức $\mathbb{C}[x]$ là một không gian Unita với tích vô hướng định bởi

$$\langle P, Q\rangle = \int_0^1 P(x)\overline{Q(x)}\, dx$$

với mọi $P, Q \in \mathbb{C}[x]$. Kết quả trên được chứng minh tương tự như trong Ví dụ 7.1.7, Chương 7.

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 chúng tôi chỉ xin nêu lại một số kết quả có một ít thay đổi so với không gian Euclid

Mệnh đề 8.1.5. *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*

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

Mệnh đề 8.1.6. *Cơ sở $\mathcal{B}$ của không gian Unita V là một 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 8.1.7. *Cho V là một không gian Unita, $\mathcal{B}$ và $\mathcal{B}'$ là hai cơ sở của V. Khi đó*

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

8.2 Toán tử tuyến tính liên hợp

Định lý 8.2.1. *Cho V là một không gian Unita và f là một phiếm hàm tuyến tính trên V. Khi đó tồn tại duy nhất một vectơ $v \in V$ sao cho $f(u) = \langle u, v\rangle$ với mọi $u \in V$.*

Chứng minh. Gọi $\mathcal{B} = (e_1, \ldots, e_n)$ là một cơ sở trực chuẩn của V. Đặt

$$v = \sum_{j=1}^{n} \overline{f(e_j)}\, e_j$$

và gọi f_v là phiếm hàm tuyến tính trên V định bởi $f_v(u) = \langle u, v\rangle$. Khi đó, với mọi $1 \leq i \leq n$, ta có

$$f_v(e_i) = \langle e_i, \sum_{j=1}^{n} \overline{f(e_j)}\, e_j\rangle = f(e_i)$$

nên $f = f_v$, nghĩa là $f(u) = \langle u, v\rangle$ với mọi $u \in V$. Bây giờ giả sử $w \in V$ cũng thỏa $f(u) = \langle u, w\rangle$ với mọi $u \in V$. Khi đó

$$\langle u, v - w\rangle = \langle u, v\rangle - \langle u, w\rangle = f(u) - f(u) = 0$$

với mọi $u \in V$ nên $v - w \in V^{\perp}$. Suy ra $v = w$. ■

Định lý 8.2.2. *Cho V là một 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 v là một vectơ bất kỳ trong V. Khi đó ánh xạ $u \mapsto \langle f(u), v\rangle$ là một toán tử tuyến tính trên V nên theo Định lý 8.2.1 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 $u \in V$. Ta chỉ cần kiểm chứng f^* là một toán tử tuyến tính trên V. Thật vậy, với mọi $u, v, w \in V$, $\alpha \in \mathbb{C}$, ta có

$$\begin{aligned}
\langle u, f^*(\alpha v + w)\rangle &= \langle f(u), \alpha v + w\rangle \\
&= \langle f(u), \alpha v + w\rangle \\
&= \bar{\alpha}\langle f(u), v\rangle + \langle f(u), w\rangle \\
&= \bar{\alpha}\langle u, f^*(v)\rangle + \langle u, f^*(w)\rangle \\
&= \langle u, \alpha f^*(v) + f^*(w)\rangle.
\end{aligned}$$

Suy ra $f^*(\alpha v + w) = \alpha f^*(v) + f^*(w)$ và f^* là một toán tử tuyến tính. Bây giờ g cũng là một toán tử tuyến tính trên V thỏa $\langle f(u), v\rangle = \langle u, g(v)\rangle$ với mọi $u, v \in V$. Khi đó

$$\begin{aligned}\langle u, f^*(v) - g(v)\rangle &= \langle u, f^*(v)\rangle - \langle u, g(v)\rangle \\ &= \langle f(u), v\rangle - \langle f(u), v\rangle \\ &= 0\end{aligned}$$

với mọi $u, v \in V$ nên $f^*(v) = g(v)$, nghĩa là $g = f^*$. ■

Định nghĩa 8.2.3. Cho $A = (a_{ij})$ là một ma trận loại $m \times n$ với hệ số phức. *Ma trận liên hợp chuyển vị* của A, ký hiệu là A^*, là ma trận (a'_{ji}) loại $n \times m$, trong đó $a'_{ji} = \overline{a_{ij}}$ với mọi $1 \le i \le m, 1 \le j \le n$.

Định lý 8.2.4. *Cho V là một không gian Unita và f là một toán tử tuyến tính trên V. Khi đó ma trận biểu diễn của toán tử tuyến tính liên hợp f^* trong một cơ sở trực chuẩn bất kỳ của V là ma trận liên hợp chuyển vị của ma trận biểu diễn của f trong cơ sở trực chuẩn đó.*

Chứng minh. Cho $\mathcal{B} = (e_1, \ldots, e_n)$ là một cơ sở trực chuẩn bất kỳ của V. Khi đó mỗi vectơ $u \in V$ được viết dưới dạng $u = \sum\limits_{j=1}^{n} \langle u, e_j\rangle e_j$ 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 8.2.5. *Cho f, g là hai toán tử tuyến tính trên một không gian Unita hữu hạn chiều và $\alpha \in \mathbb{C}$. Khi đó*

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

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

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

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

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

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

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

nên $(f^*)^* = f$.

$$\text{(ii) } \langle u, (\alpha f)^*(v)\rangle = \langle \alpha f(u), v\rangle = \alpha\langle f(u), v\rangle = \alpha\langle u, f^*(v)\rangle = \langle u, \bar{\alpha} f^*(v)\rangle,$$

từ đó $(\alpha f)^* = \bar{\alpha} f^*$.

$$\begin{aligned}\text{(iii) } \langle u, (f+g)^*(v)\rangle &= \langle (f+g)(u), v\rangle \\ &= \langle f(u), v\rangle + \langle g(u), v\rangle \\ &= \langle u, f^*(v)\rangle + \langle u, g^*(v)\rangle \\ &= \langle u, f^*(v) + g^*(v)\rangle.\end{aligned}$$

Suy ra $(f+g)^* = f^* + g^*$.

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

Từ đó suy ra $(fg)^* = g^* f^*$.

$$\text{(v) } Id_V = Id_V^* = (f^{-1}f)^* = (f^{-1})^* f^* = (f\,f^{-1})^* = f^*(f^{-1})^*.$$

Do đó $(f^{-1})^* = (f^*)^{-1}$. ■

Định lý 8.2.6. *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 $(A^{-1})^* = (A^*)^{-1}$.

8.3 Toán tử chuẩn tắc

Định nghĩa 8.3.1. Toán tử tuyến tính f trên một không gian Unita hữu hạn chiều được gọi là *toán tử chuẩn tắc* nếu f và f^* giao hoán nhau, nghĩa là nếu $f\,f^* = f^* f$.

Bổ đề 8.3.2. *Cho f, g là hai toán tử tuyến tính trên không gian vectơ hữu hạn chiều V trên trường $\mathbb{K}$, giao hoán nhau, nghĩa là $fg = gf$. 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 $\lambda \in \mathbb{K}$ là một trị riêng tùy ý của f. Với mọi $u \in E(\lambda)$ ta có $f(u) = \lambda u$ nên $f(g(u)) = (fg)(u) = (gf)(u) = g(f(u)) = g(\lambda u) = \lambda g(u)$, Nên $g(u) \in E(\lambda)$ và bổ đề được chứng minh. ■

Bổ đề 8.3.3. *Cho f là một toán tử tuyến tính trên không gian Unita hữu hạn chiều 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 $\langle u, f^*(v)\rangle = \langle f(u), v\rangle = 0$ (do $f(u) \in W$), do đó $f^*(v) \in W^{\perp}$ và bổ đề được chứng minh. ■

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

Định lý 8.3.4. *Toán tử tuyến tính f trên không gian Unita hữu hạn chiều V là toán tử chuẩn tắc khi và chỉ khi tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho ma trận biểu diễn $[f]_{\mathcal{B}}$ là một ma trận chéo.*

Chứng minh. ($\Leftarrow$) Giả sử tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho $[f]_{\mathcal{B}}$ là một 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 chất đóng đại số của trường số phức, 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 8.3.3 $W^{\perp}$ bất biến bởi f^*. Do tính chuẩn tắc của f, theo Bổ đề 8.3.2 W cũng bất biến bởi f^*. Từ đây, lại theo Bổ đề 8.3.3 $W^{\perp}$ bất biến bởi $(f^*)^* = f$. Như vậy, W, $W^{\perp}$ đều bất biến bởi f, f^* và do đó $f_{|W}$, $f_{|W^{\perp}}$ lần lượt là các toán tử chuẩn tắc trên các không gian Unita W, $W^{\perp}$. Vì $\dim W$, $\dim W^{\perp}$ đều nhỏ hơn n nên theo giả thiết quy nạp, tồn tại các cơ sở trực chuẩn $\mathcal{B}_1$ của W và $\mathcal{B}_2$ của $W^{\perp}$ sao cho $\left[f_{|W}\right]_{\mathcal{B}_1}$, $\left[f_{|W^{\perp}}\right]_{\mathcal{B}_2}$ là các ma trận chéo. Bây giờ ta chỉ cần đặt $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ thì $\mathcal{B}$ là một cơ sở trực chuẩn của V và $[f]_{\mathcal{B}}$ là một ma trận chéo. ■

8.4 Toán tử Unita

Định nghĩa 8.4.1. Toán tử tuyến tính f trên không gian Unita V được gọi là *toán tử Unita* nếu f bảo toàn tích vô hướng, nghĩa là nếu $\langle f(u), f(v)\rangle = \langle u, v\rangle$ với mọi $u, v \in V$.

Định lý 8.4.2. *Toán tử tuyến tính f trên không gian Unita hữu hạn chiều V là toán tử Unita khi và chỉ khi f bảo toàn độ dài của các vectơ, nghĩa là $\|f(u)\| = \|u\|$ với mọi $u, v \in V$. Nói riêng, mọi toán tử Unita đều là đẳng cấu tuyến tính.*

Chứng minh. Chiều thuận là hiển nhiên. Xét chiều đảo. Giả sử f bảo toàn độ dài của các vectơ. Khi đó với mọi $u, v \in V$, ta có

$$\begin{aligned}
&\|f(u+v)\|^2 = \|u+v\|^2 \\
\Leftrightarrow\ &\langle f(u+v), f(u+v)\rangle = \langle u+v, u+v\rangle \\
\Leftrightarrow\ &\langle f(u), f(u)\rangle + \langle f(u), f(v)\rangle + \langle f(v), f(u)\rangle + \langle f(v), f(v)\rangle \\
&\qquad = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
\Leftrightarrow\ &\|f(u)\|^2 + 2\,\mathrm{Re}\langle f(u), f(v)\rangle + \|f(v)\|^2 \\
&\qquad = \|u\|^2 + 2\,\mathrm{Re}\langle u, v\rangle + \|v\|^2 \\
\Leftrightarrow\ &\mathrm{Re}\langle f(u), f(v)\rangle = \mathrm{Re}\langle u, v\rangle.
\end{aligned}$$

Áp dụng kết quả trên cho các vectơ iu và v ta được

$$\mathrm{Re}\langle f(iu), f(v)\rangle = \mathrm{Re}\langle iu, v\rangle,$$

do đó

$$\mathrm{Im}\langle f(u), f(v)\rangle = \mathrm{Im}\langle u, v\rangle$$

Từ (1) và (2) ta suy ra $\langle f(u), f(v)\rangle = \langle u, v\rangle$ hay f là toán tử Unita.

Với f là một toán tử Unita và $u \in \mathrm{Ker} f$ ta có $\|u\| = \|f(u)\| = \|0\|$ nên $u = 0$. Vậy Kerf $= 0$, do đó f là đơn cấu. Mà V hữu hạn chiều nên f cũng là đẳng cấu. ■

Định lý 8.4.3. *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. Cho f là một toán tử tuyến tính trên không gian Unita V. Ta có

$$f \text{ là một toán tử Unita}$$

$$
\begin{aligned}
&\Leftrightarrow \forall u, v \in V, \langle f(u), f(v)\rangle = \langle u, v\rangle \\
&\Leftrightarrow \forall u, v \in V, \langle u, (f^* f)(v)\rangle = \langle u, v\rangle \\
&\Leftrightarrow \forall v \in V, (f^* f)(v) = v \\
&\Leftrightarrow f^* f = Id_V \\
&\Leftrightarrow f^* = f^{-1}.
\end{aligned}
$$

Khẳng định còn lại là hiển nhiên. ■

Mệnh đề 8.4.4. *(i) Nếu f là một toán tử Unita thì ánh xạ ngược f^{-1} cũng là toán tử Unita.*

(ii) Tích của hai toán tử Unita cũng là toán tử Unita.

Chứng minh. (i) Theo Định lý 8.4.2 mọi toán tử Unita f đều là đẳng cấu tuyến tính. Hơn nữa, theo Định lý 8.2.5 ta có

$$\left(f^{-1}\right)^* = (f^*)^{-1} = \left(f^{-1}\right)^{-1} = f$$

nên f^{-1} cũng là toán tử Unita.

(ii) Cho f, g là hai toán tử Unita trên không gian Unita V. Theo Định lý 8.2.5 ta có

$$(fg)^* = g^* f^* = g^{-1} f^{-1} = (fg)^{-1}$$

nên fg cũng là toán tử Unita. ■

Định nghĩa 8.4.5. 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 từ Định lý 8.2.6 ta có nhận xét sau.

Nhận xét 8.4.6. (i) Mọi ma trận Unita đều khả nghịch và ma trận nghịch đảo của chúng cũng là ma trận Unita.

(ii) Tích hai ma trận Unita cùng cấp cũng là ma trận Unita.

Bây giờ, từ các Định lý 8.2.4 và Định lý 8.4.3 ta suy ra ngay kết quả sau.

Định lý 8.4.7. *Toán tử tuyến tính f trên không gian Unita V là toán tử Unita khi và chỉ khi ma trận biểu diễn của f trong một cơ sở trực chuẩn bất kỳ của V là ma trận Unita.*

Nhận xét 8.4.8. Chỉ cần tồn tại một cơ sở trực chuẩn sao cho ma trận biểu diễn của f trong cơ sở đó là ma trận Unita thì f là toán tử Unita.

Định lý 8.4.9. *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.

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 8.4.2 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 một 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 8.1.6 ta có

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

và chiều đảo được chứng minh. ■

Nhận xét 8.4.10. Từ chứng minh Định lý 8.4.9 ta thấy chỉ cần toán tử tuyến tính f biến một cơ sở trực chuẩn nào đó của V thành cơ sở trực chuẩn thì f là toán tử Unita.

Bổ đề 8.4.11. *Mọi trị riêng của một toán tử Unita có môđun bằng 1.*

Chứng minh. Cho $\lambda \in \mathbb{C}$ là một trị riêng của toán tử Unita f trên không gian Unita V. Khi đó tồn tại $u \in V\backslash\{0\}$ sao cho $f(u) = \lambda u$ nên

$$\langle u, u\rangle = \langle f(u), f(u)\rangle = \langle \lambda u, \lambda u\rangle = \lambda\bar{\lambda}\langle u, u\rangle.$$

Vì $\langle u, u\rangle > 0$ nên $\lambda\bar{\lambda} = 1$. Suy ra $|\lambda| = 1$ và bổ đề được chứng minh. ■

Định lý 8.4.12. *Toán tử tuyến tính f trên không gian Unita hữu hạn chiều V là toán tử Unita khi và chỉ khi tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho ma trận biểu diễn $[f]_{\mathcal{B}}$ là một ma trận chéo với các hệ số trên đường chéo có môđun bằng 1.*

Chứng minh. ($\Rightarrow$) Giả sử f là toán tử Unita trên không gian Unita V. Khi đó theo Định lý 8.4.3 f là toán tử chuẩn tắc. Bây giờ chiều thuận của định lý được suy trực tiếp từ Định lý 8.3.4 và Định lý 8.4.7.

($\Leftarrow$) Giả sử tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho ma trận biểu diễn $[f]_{\mathcal{B}}$ là một 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 $\bar{\lambda} = \lambda^{-1}$, do đó $[f]_{\mathcal{B}}$ là ma trận Unita, do đó theo Nhận xét 8.4.8 f là toán tử Unita. ■

Định lý 8.4.13. *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. Vì $A^*A = I_n$ 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 $(AA^*)_{ij} = \sum\limits_{k=1}^{n} a_{ik}\overline{a_{jk}} = \left\langle \sum\limits_{k=1}^{n} a_{ik}u_k, \sum\limits_{k=1}^{n} a_{jk}v_k \right\rangle = \langle u_i, v_j \rangle = \delta_{ij}$, nghĩa là $AA^* = I_n$ và chiều đảo được chứng minh. ■

Bây giờ, từ các Định lý 8.4.9 và Định lý 8.4.13 ta suy ra ngay kết quả sau.

Hệ quả 8.4.14. *Toán tử tuyến tính f trên không gian Unita V là toán tử Unita khi và chỉ khi f ma trận biểu diễn của f trong các cặp cơ sở trực chuẩn của V là ma trận Unita.*

Định lý 8.4.15. *Ma trận $A \in M_n(\mathbb{C})$ là ma trận Unita khi và chỉ khi tồn tại một ma trận Unita P sao cho $P^{-1}AP$ là một ma trận chéo với các hệ số trên đường chéo có môđun bằng 1.*

Chứng minh. Vì 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 $\mathbb{C}^n$ 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 8.4.12 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 các hệ số trên đường chéo có môđun bằng 1. 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 8.4.13 và $P^{-1}AP = [f]_{\mathcal{B}}$. Định lý được chứng minh. ■

8.5 Toán tử Hermite

Định nghĩa 8.5.1. 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 8.5.2. *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. Cho f là một toán tử tuyến tính trên không gian Unita V. Ta có f là một toán tử Hermite

$$\begin{aligned}
&f \text{ là một toán tử Hermite}\\
\Leftrightarrow\, &\forall u, v \in V, \langle f(u), v\rangle = \langle u, f(v)\rangle\\
\Leftrightarrow\, &\forall u, v \in V, \langle u, f^*(v)\rangle = \langle u, f(v)\rangle\\
\Leftrightarrow\, &\forall v \in V, f^*(v) = f(v)\\
\Leftrightarrow\, &f^* = f.
\end{aligned}$$

Khẳng định còn lại là hiển nhiên. ■

Mệnh đề 8.5.3. *Nếu f là toán tử Hermite và là đẳng cấu thì ánh xạ ngược f^{-1} cũng là toán tử Hermite.*

Chứng minh. Theo Định lý 8.4.2 ta có $(f^{-1})^* = (f^*)^{-1} = f^{-1}$ nên f^{-1} cũng là toán tử Hermite. ■

Định nghĩa 8.5.4. 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 8.2.6.

Nhận xét 8.5.5. Nếu ma trận Hermite A khả nghịch thì ma trận nghịch đảo A^{-1} cũng là ma trận Hermite.

Bây giờ, từ các Định lý 8.2.4 và Định lý 8.5.2 ta suy ra ngay kết quả sau.

Định lý 8.5.6. *Toán tử tuyến tính f trên không gian Unita V là toán tử Hermite khi và chỉ khi ma trận biểu diễn của f trong một cơ sở trực chuẩn bất kỳ của V là ma trận Hermite.*

Nhận xét 8.5.7. Chỉ cần tồn tại một cơ sở trực chuẩn sao cho ma trận biểu diễn của f trong cơ sở đó là ma trận Hermite thì f là toán tử Hermite.

Bổ đề 8.5.8. *Mọi trị riêng của một toán tử Hermite đều là số thực.*

Chứng minh. Cho $\lambda \in \mathbb{C}$ là một trị riêng của toán tử Hermite f trên không gian Unita V. Khi đó tồn tại $u \in V\backslash\{0\}$ sao cho $f(u) = \lambda u$ nên

$$\begin{cases} \langle f(u), u\rangle = \langle \lambda u, u\rangle = \lambda\langle u, u\rangle\\ \langle u, f(u)\rangle = \langle u, \lambda u\rangle = \bar{\lambda}\langle u, u\rangle. \end{cases}$$

Mà $\langle f(u), u\rangle = \langle u, f(u)\rangle$ và $\langle u, u\rangle > 0$ nên $\lambda = \bar{\lambda}$, nghĩa là $\lambda \in \mathbb{R}$.

Định lý 8.5.9. *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 ma trận biểu diễn $[f]_\mathcal{B}$ là một ma trận chéo với hệ số thực.*

Chứng minh. ($\Rightarrow$) Giả sử f là toán tử Hermite trên không gian Unita V. Khi đó theo Mệnh đề 8.5.3 f là toán tử chuẩn tắc. Bây giờ chiều thuận của định lý được suy trực tiếp từ Định lý 8.3.4 và Bổ đề 8.5.8.

($\Leftarrow$) Giả sử tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho ma trận biểu diễn $[f]_\mathcal{B}$ là một ma trận chéo với hệ số thực. Khi đó $[f]_\mathcal{B}$ là ma trận Hermite nên theo Nhận xét 8.4.8 f là toán tử Hermite. ■

Định lý 8.5.10. *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 một ma trận chéo với hệ số thực.*

Chứng minh. Vì 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 $\mathbb{C}^n$ 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 8.5.9 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 B_0 sang $\mathcal{B}$ thì P là ma trận Unita theo Định lý 8.4.13 và $P^{-1}AP = [f]_\mathcal{B}$. Định lý được chứng minh. ■

8.6 Toán tử phản Hermite

Định nghĩa 8.6.1. 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 8.6.2. *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. Với f là một toán tử tuyến tính trên không gian Unita V, ta có

$$\begin{aligned}
&f \text{ là toán tử phản Hermite}\\
\Leftrightarrow\ &\forall u, v \in V, \langle f(u), v\rangle = -\langle u, f(v)\rangle\\
\Leftrightarrow\ &\forall u, v \in V, \langle u, f^*(v)\rangle = \langle u, -f(v)\rangle\\
\Leftrightarrow\ &\forall v \in V, f^*(v) = -f(v)
\end{aligned}$$

$$\Leftrightarrow f^* = -f.$$

Khẳng định còn lại là hiển nhiên. ■

Mệnh đề 8.6.3. *Nếu f là toán tử phản Hermite và là đẳng cấu thì ánh xạ ngược f^{-1} cũng là toán tử phản Hermite.*

Chứng minh. Theo Định lý 8.2.5 ta có

$$\left(f^{-1}\right)^* = (f^*)^{-1} = (-f)^{-1} = -f^{-1}$$

nên f^{-1} cũng là toán tử phản Hermite. ■

Định nghĩa 8.6.4. Ma trận $A \in M_n(\mathbb{C})$ được gọi là *ma trận phản Hermite* nếu $A^* = -A$.

Nhận xét sau được suy trực tiếp từ Định lý 8.2.6.

Nhận xét 8.6.5. Nếu ma trận phản Hermite A khả nghịch thì ma trận nghịch đảo A^{-1} cũng là ma trận phản Hermite.

Bây giờ, từ các Định lý 8.2.4 và Định lý 8.6.2 ta suy ra ngay kết quả sau.

Định lý 8.6.6. *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 trong một cơ sở trực chuẩn bất kỳ của V là một ma trận phản Hermite.*

Nhận xét 8.6.7. Chỉ cần tồn tại một cơ sở trực chuẩn sao cho ma trận biểu diễn của f trong cơ sở đó là ma trận phản Hermite thì f là toán tử phản Hermite.

Bổ đề 8.6.8. *Mọi trị riêng khác 0 của một toán tử phản Hermite đều thuần ảo.*

Chứng minh. Cho $\lambda \in \mathbb{C}\backslash\{0\}$ là một trị riêng của toán tử phản Hermite f trên không gian Unita V. Khi đó tồn tại $u \in V\backslash\{0\}$ sao cho $f(u) = \lambda u$ nên

$$\begin{cases} \langle f(u), u\rangle = \langle \lambda u, u\rangle = \lambda\langle u, u\rangle, \\ \langle u, f(u)\rangle = \langle u, \lambda u\rangle = \bar{\lambda}\langle u, u\rangle. \end{cases}$$

Mà $\langle f(u), u\rangle = -\langle u, f(u)\rangle$ và $\langle u, u\rangle > 0$ nên $\bar{\lambda} = -\lambda$ hay λ là số thuần ảo. ■

Định lý 8.6.9. *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 ma trận biểu diễn $[f]_{\mathcal{B}}$ là ma trận chéo với các hệ số trên đường chéo hoặc bằng 0 hoặc là thuần ảo.*

Chứng minh. ($\Rightarrow$) Giả sử f là toán tử phản Hermite trên không gian Unita

V. Khi đó theo Định lý 8.6.2 f là toán tử chuẩn tắc. Bây giờ chiều thuận của định lý được suy trực tiếp từ Định lý 8.3.4 và Bổ đề 8.6.8.

($\Leftarrow$) Giả sử tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của V sao cho ma trận biểu diễn $[f]_{\mathcal{B}}$ là một ma trận chéo với các hệ số trên đường chéo hoặc bằng 0 hoặc là thuần ảo. Khi đó $[f]_{\mathcal{B}}$ là ma trận phản Hermite nên theo Nhận xét 8.6.7 f là toán tử phản Hermite. ■

Định lý 8.6.10. *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 một ma trận chéo với các hệ số trên đường chéo hoặc bằng 0 hoặc là thuần ảo.*

Chứng minh. Vì mọi ma trận với các hệ số trên đường chéo hoặc bằng 0 hoặc là thuần ảo đều là ma trận phản Hermite nên chiều đảo là hiển nhiên. Ta chứng minh chiều thuận. Xét toán tử tuyến tính f_A trên $\mathbb{C}^n$ 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 phản Hermite nên f_A là toán tử phản Hermite. Theo Định lý 8.6.9 tồn tại một cơ sở trực chuẩn $\mathcal{B}$ của $\mathbb{C}^n$ sao cho ma trận biểu diễn $[f_A]_{\mathcal{B}}$ là ma trận với các hệ số trên đường chéo hoặc bằng 0 hoặc là thuần ảo. 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 8.4.13 và $P^{-1}AP = [f]_{\mathcal{B}}$. Định lý được chứng minh. ■

Chương 9

DẠNG SONG TUYẾN TÍNH

Trong chương này ký hiệu $\mathbb{K}$ 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 $\mathbb{K}$.

9.1 Dạng song tuyến tính

Định nghĩa 9.1.1. Một *dạng song tuyến tính* trên V là một ánh xạ có tính chất tuyến tính theo từng biến u, v, nghĩa là với mọi $u, u_1, u_2, v, v_1, v_2 \in V$ và $\alpha, \beta \in \mathbb{K}$ ta có

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

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

Dạng song tuyến tính f được gọi là *dạng song tuyến tính* nếu $f(u, v) = f(v, u)$ với mọi $u, v \in V$.

Ví dụ 9.1.2. 1) Với mỗi $u = (x_1, \ldots, x_n), v = (y_1, \ldots, y_n) \in \mathbb{R}^n$, đặt

$$f(u, v) = x_1 y_1 + \cdots + x_n y_n.$$

Khi đó f là một dạng song tuyến tính trên $\mathbb{R}^n$.

2) Một tích vô hướng trên không gian Euclide V là một dạng song tuyến tính trên V.

Định nghĩa 9.1.3. Giả sử $\mathcal{B} = (u_1, \ldots, u_n)$ là một cơ sở của V. *Ma trận của dạng song tuyến tính f* theo cơ sở $\mathcal{B}$, ký hiệu $[f]_{\mathcal{B}}$, là ma trận $A = (a_{ij})_{n \times n}$, trong đó $a_{ij} = f(u_i, u_j)$ với mọi $1 \leq i, j \leq n$.

Với mọi $u = x_1u_1 + \cdots + x_nu_n$ và $v = y_1u_1 + \cdots + y_nu_n$ thuộc V ta có

$$\begin{aligned} f(u,v) &= f(\sum_{i=1}^{n} x_iu_i, \sum_{j=1}^{n} y_ju_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} f(u_i,u_j)x_iy_j \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j \qquad (1) \end{aligned}$$

Đảo lại, (1) xác định dạng song tuyến tính f trên V có ma trận theo cơ sở $\mathcal{B}$ là $A = (a_{ij})_{n\times n}$. Chú ý rằng (1) còn được viết dưới dạng

$$\begin{aligned} f(u,v) &= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= [u]_{\mathcal{B}}^{\top}\,[f]_{\mathcal{B}}\,[v]_{\mathcal{B}}. \end{aligned}$$

Vậy với mọi $u, v \in V$,

$$f(u,v) = [u]_{\mathcal{B}}^{\top}\,[f]_{\mathcal{B}}\,[v]_{\mathcal{B}}. \qquad (1')$$

Ta gọi (1) và (1′) là *biểu thức toạ độ của dạng song tuyến tính* f theo cơ sở $\mathcal{B}$.

Nhận xét 9.1.4. 1) Với cơ sở $\mathcal{B} = (u_1, \ldots, u_n)$ cho trước, dạng song tuyến tính f được hoàn toàn xác định bởi ma trận $[f]_{\mathcal{B}}$.

2) Dạng song tuyến tính f trên V là đối xứng khi và chỉ khi $[f]_{\mathcal{B}}$ là ma trận đối xứng.

Xét $V = \mathbb{K}^n$ với cơ sở chính tắc $\mathcal{B}_0 = (e_1, \ldots, e_n)$. Đặt $A = [f]_{\mathcal{B}_0} = (a_{ij})_{n\times n}$. Khi đó với mọi $u = (x_1, \ldots, x_n), v = (y_1, \ldots, y_n)$ thuộc $\mathbb{K}^n$ ta có

$$f(u,v) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j \qquad (2)$$

Đảo lại, (2) xác định dạng song tuyến tính f trên $\mathbb{K}^n$ có ma trận theo cơ sở chính tắc là $A = (a_{ij})_{n\times n}$. Để đơn giản, trong trường hợp này ta gọi A là *ma trận* của f và (2) là *biểu thức của* f.

Ví dụ 9.1.5. Xét dạng song tuyến tính f trên $\mathbb{R}^3$ định bởi: Với mọi $u =$

$(x_1, x_2, x_3), v = (y_1, y_2, y_3)$,

$$f(u, v) = x_1y_1 + 2x_1y_2 - 4x_1y_3 - x_2y_2 + 3x_2y_3 + x_3y_1 + 9x_3y_2.$$

Ma trận của f là

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

9.2 Dạng toàn phương

Định nghĩa 9.2.1. Cho f là một dạng song tuyến tính đối xứng trên V. Khi đó ánh xạ

$$\begin{array}{rccc} Q: & V & \longrightarrow & \mathbb{K} \\ & u & \longmapsto & f(u, u) \end{array}$$

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. *phức*). Một dạng toàn phương trên $\mathbb{K}^n$ (t.ư. $\mathbb{R}^n, \mathbb{C}^n$) còn được gọi là một *dạng toàn phương n biến trên* $\mathbb{K}$ (t.ư. *n biến thực, n biến phức*).

Nhận xét 9.2.2. Dạng cực f của dạng toàn phương Q được hoàn toàn xác định bởi Q. Thật vậy,

$$\begin{aligned} f(u+v, u+v) &= f(u, u) + f(u, v) + f(v, u) + f(v, v) \\ &= f(u, u) + 2f(u, v) + f(v, v). \end{aligned}$$

Suy ra

$$f(u, v) = \frac{1}{2}[Q(u+v) - Q(u) - Q(v)].$$

Định nghĩa 9.2.3. Giả sử Q là một dạng toàn phương trên V ứng với dạng song tuyến tính đối xứng f. Với $\mathcal{B}$ là một cơ sở bất kỳ của V, ma trận $[f]_{\mathcal{B}}$ được gọi là *ma trận của dạng toàn phương* Q theo cơ sở $\mathcal{B}$, ký hiệu $[Q]_{\mathcal{B}}$.

Nhận xét rằng vì f đối xứng nên ma trận của dạng toàn phương Q trong 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 $u = x_1u_1 + \cdots + x_nu_n$ thuộc V ta có

$$Q(u) = [u]_{\mathcal{B}}^{\top} [Q]_{\mathcal{B}} [u]_{\mathcal{B}} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j. \tag{3}$$

trong đó $a_{ij} = a_{ji}$.

Đảo lại, (3) 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 (3) dưới dạng

$$Q(u) = [u]_{\mathcal{B}}^{\top} [Q]_{\mathcal{B}} [u]_{\mathcal{B}} = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1\le i<j\le n} 2a_{ij} x_i x_j \tag{3'}$$

Ta gọi (3) và (3′) là biểu thức toạ độ của dạng toàn phương Q theo cơ sở $\mathcal{B}$. Đặc biệt, xét $V = \mathbb{K}^n$ với cơ sở chính tắc $\mathcal{B}_0 = (e_1, \ldots, e_n)$. Ma trận $[Q]_{\mathcal{B}_0} = (a_{ij})_{n\times n}$ được gọi là ma trận của Q. Với mọi $u = (x_1, \ldots, x_n)$ thuộc $\mathbb{K}^n$ ta có

$$Q(u) = X^{\top} A X = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1\le i<j\le n} 2a_{ij} x_i x_j \tag{4}$$

trong đó $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, là một đa thức thuần nhất bậc 2 theo biến $x_1, \ldots, x_n$. Đảo lại, một đa thức thuần nhất bậc 2 theo n biến $x_1, \ldots, x_n$ như trong (4) luôn luôn xác định dạng toàn phương n biến trên $\mathbb{K}$ có ma trận theo cơ sở chính tắc là ma trận đối xứng $A = (a_{ij})_{n\times n}$. Ta gọi (4) là *biểu thức của dạng toàn phương Q*.

Ví dụ 9.2.4. Xét dạng toàn phương Q trên $\mathbb{R}^3$ có biểu thức định bởi

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

Khi đó ma trận của Q là

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

và $Q(x_1, x_2, x_3) = X^{\top} A X$, trong đó $X = \begin{pmatrix} x_1 \\ x_2 \\ x_n \end{pmatrix}$.

2) Giả sử dạng toàn phương Q 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 $Q(x_1, x_2, x_3) = X^\top AX$, trong đó $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, hay

$$Q(x_1, x_2, x_3) = 2x_1^2 - 4x_3^2 + 6x_1x_2 - 2x_1x_3 + 10x_2x_3.$$

Định lý 9.2.5. *Cho f là một dạng song tuyến tính trên V. Khi đó với $\mathcal{B}_1, \mathcal{B}_2$ là hai cơ sở bất kỳ của V, ta có*

$$[f]_{\mathcal{B}_2} = (\mathcal{B}_1 \to \mathcal{B}_2)^\top [f]_{\mathcal{B}_1} (\mathcal{B}_1 \to \mathcal{B}_2).$$

Chứng minh. Đặt $\mathcal{B}_1 = (u_1, \ldots, u_n), \mathcal{B}_2 = (v_1, \ldots, v_n)$. Khi đó với mọi u, v thuộc V ta có

$$\begin{aligned} f(u, v) &= [u]_{\mathcal{B}_1}^\top [f]_{\mathcal{B}_1} [v]_{\mathcal{B}_1} \\ &= ((\mathcal{B}_1 \to \mathcal{B}_2)[u]_{\mathcal{B}_2})^\top [f]_{\mathcal{B}_1} (\mathcal{B}_1 \to \mathcal{B}_2)[v]_{\mathcal{B}_2} \\ &= [u]_{\mathcal{B}_2}^\top (\mathcal{B}_1 \to \mathcal{B}_2)^\top [f]_{\mathcal{B}_1} (\mathcal{B}_1 \to \mathcal{B}_2)[v]_{\mathcal{B}_2}. \end{aligned}$$

Từ đó suy ra

$$[f]_{\mathcal{B}_2} = (\mathcal{B}_1 \to \mathcal{B}_2)^\top [f]_{\mathcal{B}_1} (\mathcal{B}_1 \to \mathcal{B}_2).$$

Hệ quả 9.2.6. *Cho Q là một dạng toàn phương trên V. Khi đó với $\mathcal{B}_1, \mathcal{B}_2$ là hai cơ sở bất kỳ của V, ta có*

$$[Q]_{\mathcal{B}_2} = (\mathcal{B}_1 \to \mathcal{B}_2)^\top [Q]_{\mathcal{B}_1} (\mathcal{B}_1 \to \mathcal{B}_2).$$

Nhận xét 9.2.7. Cho biểu thức tọa độ của dạng toàn phương Q theo cơ sở $\mathcal{B}_1$ là

$$Q(u) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \le i < j \le n} 2a_{ij} x_i x_j \qquad (*)$$

với $u = x_1u_1 + \cdots + x_nu_n$. Giả sử $\mathcal{B}_2 = (v_1, \ldots, v_n)$ là một cơ sở khác của V

và biểu thức tọa độ của Q theo cơ sở $\mathcal{B}_2$ là

$$Q(u) = \sum_{i=1}^{n} b_{ii}y_i^2 + \sum_{1\le i<j\le n} 2b_{ij}y_iy_j \qquad (**)$$

với $u = y_1u_1 + \cdots + y_nu_n$. Khi đó ma trận $(\mathcal{B}_1 \to \mathcal{B}_2)$ khả nghịch và phép biến đổi tọa độ không suy biến

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\mathcal{B}_1 \to \mathcal{B}_2)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

đưa biểu thức của dạng toàn phương Q từ $(*)$ về $(**)$. Đảo lại, ứng với mỗi phép biến đổi tọa độ không suy biến

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

(P là một ma trận vuông cấp n khả nghịch), gọi $\mathcal{B}_2$ là cơ sở của V sao cho $(\mathcal{B}_1 \to \mathcal{B}_2) = P$. Khi đó biểu thức tọa độ của Q theo cơ sở $\mathcal{B}_2$ có dạng $(**)$, trong đó

$$(b_{ij})_{n\times n} = P^\top (a_{ij})_{n\times n}P.$$

Định nghĩa 9.2.8. Cho Q là một dạng toàn phương trên không gian vectơ n chiều V và $\mathcal{B}$ là một cơ sở bất kỳ của V. Hạng của ma trận $[Q]_\mathcal{B}$ được gọi là *hạng* của Q, ký hiệu $\operatorname{rank}(Q)$ (hay $r(Q)$).

Từ Hệ quả 9.2.6 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 $rank(Q) < n$ thì Q *suy biến*.

Ví dụ 9.2.9. Xét Q là dạng toàn phương 3 biến thực 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

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

a) Ta có $\det A = -20$ nên $\operatorname{rank}(Q) = \operatorname{rank}(A) = 3$, 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

$$\begin{aligned}
[Q]_{\mathcal{B}} &= (\mathcal{B}_0 \to \mathcal{B})^\top [Q]_{\mathcal{B}_0} (\mathcal{B}_0 \to \mathcal{B}) \\
&= \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}^\top \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 4 \\ -2 & 4 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} \\
&= \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 4 \\ -2 & 4 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} \\
&= \begin{pmatrix} -4 & 2 & -18 \\ 2 & 5 & 28 \\ -18 & 28 & -20 \end{pmatrix}.
\end{aligned}$$

Vậy biểu thức tọa độ của Q theo cơ sở $\mathcal{B}$ định bởi: Với mọi $u = y_1u_1 + y_2u_2 + y_3u_3$ thuộc $\mathbb{R}^3$,

$$\begin{aligned}
Q(u) &= [u]_{\mathcal{B}}^\top \, [Q]_{\mathcal{B}} \, [u]_{\mathcal{B}} \\
&= \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} -4 & 2 & -18 \\ 2 & 5 & 28 \\ -18 & 28 & -20 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\
&= -4y_1^2 + 5y_2^2 - 20y_3^2 + 4y_1y_2 - 36y_1y_3 + 56y_2y_3
\end{aligned}$$

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

$$\begin{cases} x_1 &= y_1 - y_2 + 2y_3 \\ x_2 &= -y_1 + 2y_2 \\ x_3 &= y_2 + 3y_3 \end{cases}$$

Ví dụ 9.2.10. Xét không gian $\mathbb{R}^3$ với cơ sở $\mathcal{B} = (u_1, u_2, u_3)$, trong đó $u_1 = (1, -1, 0); u_2 = (-1, 2, 1); u_3 = (2, 0, 3)$. Cho Q là dạng toàn phương 3 biến thực có biểu thức toạ độ theo cơ sở $\mathcal{B}$ như sau:

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

với mọi $u = x_1u_1 + x_2u_2 + x_3u_3$ thuộc $\mathbb{R}^3$.

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

b) Tìm biểu thức của Q và chỉ ra phép biến đổi toạ độ không suy biến tương ứng.

Giải. Ma trận của Q theo cơ sở $\mathcal{B}$ là

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

a) Ta có $\det A = -20$ nên $\operatorname{rank}(Q) = \operatorname{rank}(A) = 3$, do đó Q không suy biến.

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

$$(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}.$$

Suy ra ma trận chuyển từ cơ sở $\mathcal{B}$ sang cơ sở chính tắc $\mathcal{B}_0$ là

$$(\mathcal{B} \to \mathcal{B}_0) = (\mathcal{B}_0 \to \mathcal{B})^{-1} = \begin{pmatrix} 6 & 5 & -4 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}.$$

Do đó ma trận của Q (theo cơ sở chính tắc $\mathcal{B}_0$) là

$$\begin{aligned}
[Q]_{\mathcal{B}_0} &= (\mathcal{B} \to \mathcal{B}_0)^\top [Q]_{\mathcal{B}} (\mathcal{B} \to \mathcal{B}_0) \\
&= \begin{pmatrix} 6 & 5 & -4 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}^\top \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 4 \\ -2 & 4 & 0 \end{pmatrix} \begin{pmatrix} 6 & 5 & -4 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 6 & 3 & -1 \\ 5 & 3 & -1 \\ -4 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 4 \\ -2 & 4 & 0 \end{pmatrix} \begin{pmatrix} 6 & 5 & -4 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 45 & 34 & -30 \\ 34 & 24 & -22 \\ -30 & -22 & 20 \end{pmatrix}
\end{aligned}$$

Suy ra biểu thức của Q 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

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (\mathcal{B} \to \mathcal{B}_0) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 6 & 5 & -4 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

hay

$$\begin{cases} x_1 &= 6y_1 + 5y_2 - 4y_3 \\ x_2 &= 3y_1 + 3y_2 - 2y_3 \\ x_3 &= -y_1 - y_2 + y_3 \end{cases}$$

9.3 Dạng chính tắc của dạng toàn phương

Định nghĩa 9.3.1. Cho Q là một dạng toàn phương trên V có dạng cực là f. Cơ sở $\mathcal{B} = (u_1, \dots, u_n)$ của V được gọi là một *cơ sở Q-chính tắc* nếu

$$f(u_i, u_j) = 0 \text{ với mọi } 1 \leq i \neq j \leq n.$$

Điều này tương đương với tính chất ma trận $[Q]_{\mathcal{B}}$ là một ma trận đường chéo, hay cũng vậy, biểu thức toạ độ của Q theo cơ sở $\mathcal{B}$ có dạng

$$Q(u) = \sum_{i=1}^{n} a_{ii} x_i^2. \tag{1}$$

với mọi $u = x_1u_1 + \cdots + x_nu_n$ thuộc V. Khi đó ta nói (1) là *dạng chính tắc* của Q.

Định lý 9.3.2. *Cho Q là một dạng toàn phương trên V. Khi đó trong V tồn tại một cơ sở Q-chính tắc.*

Chứng minh. Việc xây dựng một cơ sở Q-chính tắc được thực hiện thông qua thuật toán sau:

Thuật toán Lagrange đưa dạng toàn phương về dạng chính tắc

Giả sử biểu thức của dạng toàn phương Q theo cơ sở $\mathcal{B} = (u_1, \ldots, u_n)$ định bởi

$$Q(u) = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{1\leq i<j\leq n} 2a_{ij}x_ix_j \qquad (*)$$

Để đưa Q về dạng chính tắc ta chia bài toán thành 3 trường hợp:

Trường hợp 1. $a_{ii} \neq 0$ với một i nào đó. 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

$$\begin{aligned}
Q(u) &= a_{11}(x_1^2 + 2x_1\sum_{i=2}^{n}\frac{a_{1i}}{a_{11}}x_i) + (\text{những số hạng không chứa } x_1)\\
&= a_{11}(x_1 + \sum_{i=2}^{n}\frac{a_{1i}}{a_{11}}x_i)^2 + (\text{một dạng toàn phương của } x_2, \ldots, x_n)\\
&= a_{11}y_1^2 + \sum_{i,j=2}^{n} b_{ij}y_iy_j
\end{aligned}$$

trong đó

$$\begin{cases} y_1 &= x_1 + \sum\limits_{i=2}^{n}\dfrac{a_{1i}}{a_{11}}x_i \\ y_j &= x_j \qquad (j \geq 2) \end{cases}$$

là một phép biến đổi tọa độ không suy biến. Việc đưa Q về dạng chính tắc được qui về việc đưa dạng toàn phương $(n-1)$ biến $Q_1 = \sum\limits_{i,j=2}^{n} b_{ij}y_iy_j$ về dạng chính tắc. Điều này có thể thực hiện bằng qui nạp.

Trường hợp 2. $a_{ii} = 0$ với mọi i nhưng có $a_{ij} \neq 0$ với $i \neq j$ nào đó. 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

$$\begin{cases} x_1 &= y_1 + y_2 \\ x_2 &= y_1 - y_2 \\ x_j &= y_j \qquad (j \geq 3) \end{cases}$$

ta có

$$2a_{12}x_1x_2 = 2a_{12}(y_1^2 - y_2^2).$$

Từ đó

$$Q(u) = \sum_{i,j=1}^{n} a_{ij}x_ix_j = \sum_{i,j=1}^{n} b_{ij}y_iy_j$$

có hệ số của y_1^2 là $2a_{12} \neq 0$. Ta trở về trường hợp 1 đã xét.

Trường hợp 3. $a_{ij} = 0$ với mọi i, j. Khi đó $Q(u) = 0$ với mọi u nên Q có dạng chính tắc trong bất kỳ cơ sở nào của V.

Ví dụ 9.3.3. Đưa dạng toàn phương thực sau đây về dạng chính tắc:

$$Q(u) = x_1^2 + x_2^2 + x_3^2 - 2x_4^2 - 2x_1x_2 + 2x_1x_3 - 2x_1x_4 + x_2x_3 - 4x_2x_4$$

với $u = (x_1, x_2, x_3, x_4)$. Chỉ ra cơ sở Q-chính tắc và phép biến đổi tọa độ không suy biến tương ứng.

Giải. Ta có

$$\begin{aligned}
Q(u) &= x_1^2 + 2x_1(-x_2 + x_3 - x_4) + x_2^2 + x_3^2 - 2x_4^2 + x_2x_3 - 4x_2x_4 \\
&= (x_1 - x_2 + x_3 - x_4)^2 - (-x_2 + x_3 - x_4)^2 + x_2^2 + x_3^2 - 2x_4^2 + x_2x_3 - 4x_2x_4 \\
&= (x_1 - x_2 + x_3 - x_4)^2 - 3x_4^2 + 3x_2x_3 - 6x_2x_4 + 2x_3x_4 \\
&= (x_1 - x_2 + x_3 - x_4)^2 - 3[x_4^2 + 2x_4(x_2 - \frac{1}{3}x_3)] + 3x_2x_3 \\
&= (x_1 - x_2 + x_3 - x_4)^2 - 3(x_4 + x_2 - \frac{1}{3}x_3)^2 + 3(x_2 - \frac{1}{3}x_3)^2 + 3x_2x_3 \\
&= (x_1 - x_2 + x_3 - x_4)^2 - 3(x_4 + x_2 - \frac{1}{3}x_3)^2 + 3(x_2^2 + 2x_2\frac{1}{6}x_3) + \frac{1}{3}x_3^2 \\
&= (x_1 - x_2 + x_3 - x_4)^2 - 3(x_4 + x_2 - \frac{1}{3}x_3)^2 + 3(x_2 + \frac{1}{6}x_3)^2 + \frac{1}{4}x_3^2
\end{aligned}$$

Thực hiện phép biến đổi tọa độ không suy biến

$$\begin{cases} y_1 = x_1 - x_2 + x_3 - x_4 \\ y_2 = x_4 + x_2 - \frac{1}{3}x_3 \\ y_3 = x_2 + \frac{1}{6}x_3 \\ y_4 = x_3 \end{cases} \Leftrightarrow \begin{cases} x_1 = y_1 - y_2 + \frac{2}{3}y_4 \\ x_2 = y_3 - \frac{1}{6}y_4 \\ x_3 = y_4 \\ x_4 = y_2 - y_3 + \frac{1}{2}y_4 \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} x_1 \\ x_4 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 2/3 \\ 0 & 0 & 1 & -1/6 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 1/2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_4 \\ y_3 \\ y_4 \end{pmatrix}.$$

Ta đưa được Q về dạng chính tắc

$$Q(u) = y_1^2 - 3y_2^2 + 3y_3^2 + \frac{1}{4}y_4^2$$

với $u = y_1u_1 + y_2u_2 + y_3u_3 + y_4u_4$, trong đó cơ sở Q-chính tắc $\mathcal{B} = (u_1, u_2, u_3, u_4)$ của $\mathbb{R}^4$ thoả

$$(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} 1 & -1 & 0 & 2/3 \\ 0 & 0 & 1 & -1/6 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & 1/2 \end{pmatrix}$$

($\mathcal{B}_0$ là cơ sở chính tắc của $\mathbb{R}^4$) nghĩa là

$$u_1 = (1, 0, 0, 0); u_2 = (-1, 0, 0, 1); u_3 = (0, 1, 0, -1); u_4 = (2/3, -\frac{1}{6}, 1, \frac{1}{2}).$$

Ví dụ 9.3.4. Đưa dạng toàn phương thực sau đây về dạng chính tắc:

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

Chỉ ra cơ sở Q-chính tắc và phép biến đổi tọa độ không suy biến tương ứng.

Giải. Đổi biến

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3. \end{cases}$$

Ta có

$$\begin{aligned} Q(u) &= y_1^2 - y_2^2 + 2(y_1 + y_2)y_3 - 2(y_1 - y_2)y_3 \\ &= y_1^2 - y_2^2 + 4y_2y_3. \end{aligned}$$

Ta biến đổi

$$\begin{aligned} Q(u) &= y_1^2 - [y_2^2 - 2y_2(2y_3)] \\ &= y_1^2 - (y_2 - 2y_3)^2 + 4y_3^2. \end{aligned}$$

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 toạ độ không suy biến tương ứng là

$$\begin{cases} x_1 &= z_1 + z_2 + 2z_3 \\ x_2 &= z_1 - z_2 - 2z_3 \\ x_3 &= z_3 \end{cases}$$

Cơ sở Q-chính tắc tương ứng là $\mathcal{B} = (u_1, u_2, u_3)$ của $\mathbb{R}^3$ thoả

$$(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

($\mathcal{B}_0$ là cơ sở chính tắc của $\mathbb{R}^3$) nghĩa là

$$u_1 = (1, 1, 0); u_2 = (1, -1, 0); u_3 = (2, -2, 1).$$

9.4 Dạng chính tắc trực giao của dạng toàn phương trên không gian Euclid

Định nghĩa 9.4.1. Cho Q là một dạng toàn phương trên không gian Euclid V. Cơ sở $\mathcal{B}$ được gọi là một *cơ sở Q-chính tắc trực giao* nếu $\mathcal{B}$ là một cơ sở trực chuẩn đồng thời cũng là một 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 9.4.2. *Cho Q là một dạng toàn phương trên không gian Euclid V. Khi đó trong V tồn tại một cơ sở Q-chính tắc trực giao.*

Chứng minh. Xét $\mathcal{B}_0$ là một cơ sở trực chuẩn nào đó của V. Khi đó ma trận $[Q]_{\mathcal{B}_0}$ là ma trận đối xứng thực nên chéo hoá 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

$$[Q]_{\mathcal{B}} = (\mathcal{B}_0 \to \mathcal{B})^\top [Q]_{\mathcal{B}_0} (\mathcal{B}_0 \to \mathcal{B}) = P^\top [Q]_{\mathcal{B}_0} P = P^{-1} [Q]_{\mathcal{B}_0} P$$

là ma trận đường chéo. Vì $[Q]_{\mathcal{B}}$ là ma trận đường chéo nên $\mathcal{B}$ là cơ sở Q-chính tắc. Mặt khác, do $(\mathcal{B}_0 \to \mathcal{B}) = P$ là ma trận trực giao nên $\mathcal{B}$ là một cơ sở trực chuẩn. Suy ra $\mathcal{B}$ là một cơ sở Q-chính tắc trực giao của V.

Nhận xét 9.4.3. 1) Giả sử Q có dạng chính tắc trực giao

$$Q(u) = \sum_{i=1}^{n} a_i x_i^2 \tag{1}$$

với $u = x_1u_1 + \cdots + x_nu_n$, trong đó $\mathcal{B} = (u_1, \ldots, u_n)$ là cơ sở Q- chính tắc trực giao tương ứng. Khi đó dãy $a_1, \ldots, a_n$ gồm tất cả các trị riêng của $[Q]_{\mathcal{B}}$ (kể cả số bội) và không phụ thuộc vào việc chọn cơ sở Q-chính tắc trực giao $\mathcal{B}$. Thật vậy, từ (1) ta suy ra

$$[Q]_{\mathcal{B}} = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$$

nên hiển nhiên $a_1, \ldots, a_n$ là tất cả các trị riêng của $[Q]_{\mathcal{B}}$. Bây giờ cho $\mathcal{B}' = (u'_1, \ldots, u'_n)$ là một cơ sở Q-chính tắc trực giao khác của V. Khi đó

$$Q(u) = \sum_{i=1}^{n} {a'}_i y_i^2 \tag{1'}$$

với $u = y_1u'_1 + \cdots + y_nu'_n$. Theo chứng minh trên $a'_1, \ldots, a'_n$ là các trị riêng của $[Q]'_{\mathcal{B}}$ (kể cả số bội). Theo Hệ quả 9.2.6 ta có

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

Chú ý rằng do $\mathcal{B}, \mathcal{B}'$ là hai cơ sở trực chuẩn của V nên ma trận chuyển cơ sở $(\mathcal{B} \to \mathcal{B}')$ là một 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 hai ma trận $[Q]_{\mathcal{B}}$ và $[Q]'_{\mathcal{B}}$ đồng dạng nên cúng có cùng trị riêng (kể cả số bội), nghĩa là hai dãy $a_1, \ldots, a_n$ và $a'_1, \ldots, a'_n$ trùng nhau. Điều này chứng tỏ $a_1, \ldots, a_n$ không phụ thuộc vào việc chọn cơ sở trực chuẩn Q-chính tắc trực giao $\mathcal{B}$.

2) Từ chứng minh Định lý 9.3.2 ta thấy để đưa Q về dạng chính tắc trực giao ta dùng phép biến đổi tọa độ $X = (\mathcal{B}_0 \to \mathcal{B})Y$ với mọi $X = [u]_{\mathcal{B}_0}, Y = [u]_{\mathcal{B}}$. Vì $(\mathcal{B}_0 \to \mathcal{B})$ là ma trận trực giao nên ta nói phép biến đổi trên là một phép biến đổi tọa độ trực giao.

Thuật toán đưa dạng toàn phương trên không gian Euclide về dạng chính tắc trực giao

Cho Q là một dạng toàn phương trên không gian Euclide n chiều V. Khi đó ta đưa được Q về dạng chính tắc trực giao và chỉ ra cơ sở Q-chính tắc trực giao và phép biến đổi tọa độ trực giao tương ứng thông qua các bước sau:

Bước 1: Xác định $[Q]_{\mathcal{B}_0}$ với $\mathcal{B}_0$ là một cơ sở trực chuẩn nào đó của V.

Bước 2: Chéo hoá trực giao ma trận $[Q]_{\mathcal{B}_0}$ tìm ma trận trực giao P sao cho

$$P^{-1}[Q]_{\mathcal{B}_0}P = \operatorname{diag}(a_1, \dots, a_n).$$

Bước 3: Cơ sở Q-chính tắc trực giao $\mathcal{B} = (u_1, \dots, u_n)$ định bởi $(\mathcal{B}_0 \to \mathcal{B}) = P$ và phép biến đổi tọa độ trực giao là $X = PY$. Dạng chính tắc trực giao của Q là

$$Q(u) = \sum_{i=1}^{n} a_i x_i^2$$

với $u = x_1u_1 + \cdots + x_nu_n$.

Ví dụ 9.4.4. Đưa dạng toàn phương thực sau đây về dạng chính tắc trực giao:

$$Q(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

Chỉ ra cơ sở Q-chính tắc trực giao và phép biến đổi tọa độ trực giao tương ứng.

Giải. Bước 1: Ma trận của Q (theo cơ sở chính tắc $\mathcal{B}_0$) là

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

Bước 2: Chéo hoá trực giao ma trận A.

a) Đa thức đặc trưng của A là

$$P_A(\lambda) = |A - \lambda I_3| = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = (\lambda + 1)^2(\lambda - 2).$$

b) Trị riêng: A có 2 trị riêng là $\lambda_1 = -1$ (bội 2), $\lambda_2 = 2$ (bội 1)

c) Không gian riêng $E(\lambda_1)$ ứng với trị riêng $\lambda_1 = -1$ là không gian nghiệm của hệ

$$(A - \lambda_1 I_3)X = 0 \tag{1}$$

Ta có

$$A - \lambda_1 I_3 = A + I_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Như vậy

$$(1) \Leftrightarrow x_1 + x_2 + x_3 = 0 \Leftrightarrow \begin{cases} x_2 &= \alpha \\ x_3 &= \beta \\ x_1 &= -\alpha - \beta \end{cases}.$$

Suy ra (1) có vô số nghiệm

$$(x_1, x_2, x_3) = (-\alpha - \beta, \alpha, \beta).$$

Do đó

$$\begin{aligned} E(\lambda_1) &= \{(-\alpha - \beta, \alpha, \beta) \,|\, \alpha, \beta \in \mathbb{R}\} \\ &= \{\alpha(-1, 1, 0) + \beta(-1, 0, 1) \,|\, \alpha, \beta \in \mathbb{R}\} \\ &= \langle(-1, 1, 0); (-1, 0, 1)\rangle. \end{aligned}$$

Như vậy $E(\lambda_1)$ có $\dim E(\lambda_1) = 2$ với cơ sở (u_1, u_2) với $u_1 = (-1, 1, 0); u_2 = (-1, 0, 1)$. Ta xây dựng cơ sở trực chuẩn của $E(\lambda_1)$ qua quá trình trực chuẩn Gram–Schmidt:

$$\begin{aligned} &v_1 := u_1 = (-1, 1, 0); \\ &v_2 = u_2 - \frac{\langle u_1, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = (\frac{1}{2}, -\frac{1}{2}, 1). \end{aligned}$$

Ta đặt

$$\begin{aligned} &w_1 := \frac{v_1}{||v_1||} = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0); \\ &w_2 := \frac{v_2}{||v_2||} = (-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}). \end{aligned}$$

Khi đó (w_1, w_2) là cơ sơ trực chuẩn của $E(\lambda_1)$.

d) Không gian riêng $E(\lambda_2)$ ứng với trị riêng $\lambda_2 = 2$ là không gian nghiệm của hệ

$$(A - \lambda_2 I_3)X = 0. \tag{2}$$

Ta có

$$A - \lambda_2 I_3 = A - 2I_3 = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Như vậy

$$(2) \Leftrightarrow \left\{ \begin{array}{l} x_1 - 2x_2 + x_3 = 0 \\ x_2 - x_3 = 0 \end{array} \right. \Leftrightarrow x_1 = x_2 = x_3 = \alpha.$$

Suy ra (2) có vô số nghiệm $(x_1, x_2, x_3) = (\alpha, \alpha, \alpha)$. Do đó

$$\begin{aligned} E(\lambda_2) &= \{(\alpha, \alpha, \alpha) \,|\, \alpha \in \mathbb{R}\} \\ &= \{\alpha(1, 1, 1) \,|\, \alpha \in \mathbb{R}\} \\ &= \langle (1, 1, 1) \rangle \end{aligned}$$

Như vậy $E(\lambda_2)$ có $\dim E(\lambda_2) = 1$ với cơ sở (u_3) với $u_3 = (1, 1, 1)$. Ta xây dựng cơ sở trực chuẩn (w_3) của $E(\lambda_2)$ với

$$w_3 = \frac{u_3}{||u_3||} = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}).$$

e) Đặt $\mathcal{B} = (w_1, w_2, w_3)$. Ta có $\mathcal{B}$ là một cơ sở trực chuẩn của $\mathbb{R}^3$ và

$$P = (\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Bước 3: Từ kết quả tìm được ở bước 2, ta suy ra dạng chính tắc trực giao của Q là

$$Q(u) = -y_1^2 - y_2^2 + 2y_3^2$$

với $u = y_1 w_1 + y_2 w_2 + y_3 w_3$. 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 $X = PY$, nghĩa là

$$\left\{ \begin{array}{lcl} x_1 &=& -\frac{1}{\sqrt{2}}y_1 - \frac{1}{\sqrt{6}}y_2 + \frac{1}{\sqrt{3}}y_3 \\ x_2 &=& \frac{1}{\sqrt{2}}y_1 - \frac{1}{\sqrt{6}}y_2 + \frac{1}{\sqrt{3}}y_3 \\ x_3 &=& \frac{2}{\sqrt{6}}y_2 + \frac{1}{\sqrt{3}}y_3 \end{array} \right.$$

9.5 Dạng chính tắc - Luật quán tính của dạng toàn phương thực

Định nghĩa 9.5.1. Cho Q là một dạng toàn phương trên không gian vectơ thực n chiều V và $\mathcal{B}$ là một cơ sở của V. Giả sử biểu thức tọa độ của Q theo cơ sở $\mathcal{B}$ có dạng

$$Q(u) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 \qquad (1)$$

với $u = x_1u_1 + \cdots + x_nu_n$, trong đó r, s là các số nguyên thỏa $0 \leq s \leq r \leq n$. Khi đó ta nói $\mathcal{B}$ là một *cơ sở Q-chuẩn tắc* và (1) là dạng chuẩn tắc của Q.

Định lý 9.5.2. *Cho V là một không gian vectơ thực hữu hạn chiều và Q là một dạng toàn phương trên V. Khi đó trong V tồn tại một cơ sở Q-chuẩn tắc.*

Chứng minh. Theo Định lý 9.3.2 tồn tại một cơ sở Q-chính tắc của V. Đặt $r = \operatorname{rank}(Q)$. Bằng cách đánh số lại nếu cần ta có thể giả sử biểu thức tọa độ của Q theo cơ sở trên có dạng

$$Q(u) = a_1x_1^2 + \cdots + a_rx_r^2$$

và tồn tại số nguyên $0 \leq s \leq r$ sao cho $a_i > 0\,(i \in \overline{1,s})$; $a_i < 0\,(i \in \overline{s+1,r})$. Dùng phép biến đổi toạ độ không suy biến

$$x_j = \begin{cases} \frac{1}{\sqrt{a_j}}y_j & \text{nếu } \leq j \leq s \\ \frac{1}{\sqrt{-a_j}}y_j & \text{nếu } s+1 \leq j \leq r \\ y_j & \text{nếu } r+1 \leq j \leq n \end{cases}$$

ta thu được dạng chuẩn tắc của Q

$$Q(u) = y_1^2 + \cdots + y_s^2 - y_{s+1}^2 - \cdots - y_r^2.$$

Cơ sở tương ứng chính là cơ sở Q-chuẩn tắc cần tìm.

Định lý 9.5.3. *Cho Q là một dạng toàn phương trên V và $\mathcal{B}$ là một 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 $r = \operatorname{rank}(Q)$ và $0 \leq s \leq r$ không phụ thuộc vào cách chọn cơ sở $\mathcal{B}$.

Chứng minh. Hiển nhiên $r = \operatorname{rank}(Q)$ không phụ thuộc vào cơ sở $\mathcal{B}$. 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

$$Q(u) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 \qquad (1)$$

$$Q(u) = y_1^2 + \cdots + y_t^2 - y_{t+1}^2 - \cdots - y_r^2 \qquad (2)$$

Ta chứng minh $s = t$. Đặt

$$V_1 = \langle u_1, \ldots, u_s \rangle \text{ và } V_2 = \langle v_{t+1}, \ldots, v_n \rangle.$$

Trước hết ta chỉ ra rằng $V_1 \cap V_2 = \{0\}$. Thật vậy,

$$\text{Với } u \in V_1 \cap V_2 \Rightarrow \begin{cases} u &= \alpha_1 u_1 + \cdots + \alpha_s u_s \\ u &= \beta_{t+1} v_{t+1} + \cdots + \beta_n v_n \end{cases}$$

$$\Rightarrow \begin{cases} Q(u) &= \alpha_1^2 + \cdots + \alpha_s^2 \geq 0 \\ Q(u) &= -\beta_{t+1}^2 - \cdots - \beta_r^2 \leq 0 \end{cases}$$

$$\Rightarrow Q(u) = \alpha_1^2 + \cdots + \alpha_s^2 = 0$$

$$\Rightarrow \alpha_1 = \cdots = \alpha_s = 0 \text{ hay } u = 0.$$

nghĩa là $V_1 \cap V_2 = \{0\}$. Kết quả trên cho thấy

$$n \geq \dim (V_1 + V_2) = \dim V_1 + \dim V_2 = s + (n - t).$$

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 9.5.4. 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.

Nhận xét 9.5.5. Giả sử Q là dạng toàn phương thực có dạng chính tắc

$$Q(u) = a_1 x_1^2 + \cdots + a_n x_n^2.$$

Xét dãy $a_1, \ldots, a_n$ $(*)$. Ta có

(i) Chỉ số dương quán tính của Q bằng số các số hạng dương của $(*)$.

(ii) Chỉ số âm quán tính của Q bằng số các số hạng âm của $(*)$.

Ví dụ 9.5.6. Xét lại Ví dụ trong 9.3.3, ta thấy dạng toàn phương Q có dạng chính tắc là

$$Q(u) = y_1^2 - 3y_2^2 + 3y_3^2 + \frac{1}{4}y_4^2$$

Do đó Q có

- Chỉ số dương quán tính là 3.
- Chỉ số âm quán tính là 1.
- Cặp chỉ số quán tính là (3, 1).
- Ký số là 2.

9.6 Dạng toàn phương xác định

Định nghĩa 9.6.1. Cho Q là một dạng toàn phương trên không gian vectơ thực hữu hạn chiều V. Ta nói

(i) *Q xác định dương* nếu $Q(u) > 0$ với mọi $u \in V \setminus \{0\}$.

(ii) *Q xác định âm* nếu $Q(u) < 0$ với mọi $u \in V \setminus \{0\}$.

Nhận xét 9.6.2. Q xác định dương khi và chỉ khi dạng cực của Q là một tích vô hướng trên V.

Nhận xét 9.6.3. Cho Q là một dạng toàn phương trên không gian vectơ thực n chiều. Khi đó

(i) Q xác định dương $\Leftrightarrow$ Q có chỉ số dương quán tính bằng n.

(ii) Q xác định âm $\Leftrightarrow$ Q có chỉ số âm quán tính bằng n.

Chứng minh. (i) ($\Leftarrow$) Giả sử 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 $\mathcal{B}$ như sau:

$$Q(u) = x_1^2 + \cdots + x_n^2$$

với $u = x_1u_1 + \cdots + x_nu_n$. Nếu $u \neq 0$ thì tồn tại i sao cho $x_i \neq 0$, đưa đến $Q(u) > 0$. Vậy Q xác định dương.

($\Rightarrow$) 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

$$Q(u) = a_1x_1^2 + \cdots + a_nx_n^2,$$

trong đó có $a_i \leq 0$ với một i nào đó. Đặt $u = u_i$. Ta có $u \neq 0$ và $Q(u) = a_i \leq 0$. Mâu thuẫn với tính xác định dương của Q.

(ii) suy ra (i) cùng với nhận xét: Q xác định âm $\Leftrightarrow -Q$ xác định dương.

Hệ quả 9.6.4. *Mọi dạng toàn phương xác định dương hay xác định âm đều không suy biến.*

Định nghĩa 9.6.5. Cho $A = (a_{ij})_{n\times n}$ là một ma trận vuông cấp n. *Định thức con chính cấp k* $(1 \leq k \leq n)$ của A là định thức con sinh bởi các dòng $1, \ldots, k$ và các cột $1, \ldots, k$:

$$\Delta_k = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}.$$

Định lý 9.6.6 (Tiêu chuẩn Syster). *Cho Q là một dạng toàn phương trên không gian vectơ thực hữu hạn chiều V và A là ma trận biểu diễn của Q theo một cơ sở nào đó của V. Khi đó*

(i) Q xác định dương khi và chỉ khi mọi định thức con chính của A đều dương.

(ii) Q xác định âm khi và chỉ khi 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) Ta chỉ cần xét trường hợp Q không suy biến. Gọi f là dạng cực của Q. Giả sử $\mathcal{B} = (u_1, \ldots, u_n)$ là cơ sở V và $[Q]_{\mathcal{B}} = A$. Chú ý rằng với giả thiết Q xác định dương hoặc mọi định thức con chính của A đều dương, tương tự như quá trình trực chuẩn hoá Gram–Schmidt ta xây dựng được cơ sở f-trực giao $\mathcal{B}' = (v_1, \ldots, v_n)$. Điều này hiển nhiên nếu Q xác định dương vì khi đó f xác định một tích vô hướng trên V. Vậy ta chỉ cần xét trường hợp mọi định thức con chính của A đều dương. Ta chọn các hệ số $\alpha_{ij} \in \mathbb{R}$ thỏa hệ sau sao cho $(v_1, \ldots, v_n)$ là một cơ sở của V :

$$\begin{cases} v_1 = \alpha_{11}u_1 \\ v_2 = \alpha_{21}u_1 + \alpha_{22}u_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ v_n = \alpha_{n1}u_1 + \alpha_{n2}u_2 + \cdots + \alpha_{nn}u_n \\ \forall 1 \leq k \leq n, \forall 1 \leq j \leq k-1,\ f(u_j, v_k) = 0, f(u_k, v_k) = 1 \end{cases} \qquad (*)$$

Với $k = 1 : 1 = f(u_1, v_1) = f(u_1, \alpha_{11}u_1) = \alpha_{11}f(u_1, u_1) = \alpha_{11}a_{11} = \alpha_{11}\Delta_1$. Suy ra $\alpha_{11} = \frac{1}{\Delta_1}$.

Xét $1 < k \le n$. Giả sử đã xác định được các hệ số $\alpha_{ij}, 1 \le i \le k-1, 1 \le j \le i$. Khi đó

$$(*) \Leftrightarrow \begin{cases} a_{11}\alpha_{k1} + \quad a_{12}\alpha_{k2} \quad + \dots + \quad a_{1k}\alpha_{kk} & = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ a_{(k-1)1}\alpha_{k1} + a_{(k-1)2}\alpha_{k2} + \dots + a_{(k-1)k}\alpha_{kk} & = 0 \\ a_{k1}\alpha_{k1} + \quad a_{k2}\alpha_{k2} \quad + \dots + \quad a_{kk}\alpha_{kk} & = 1 \end{cases} \qquad (**)$$

Ma trận hệ số của hệ $(**)$ có định thức bằng $\Delta_k \neq 0$ nên các hệ số $\alpha_{kj}, 1 \le j \le k$ được xác định duy nhất, trong đó

$$\alpha_{kk} = \frac{1}{\Delta_k}\begin{vmatrix} a_{11} & \cdots & a_{1(k-1)} & 0 \\ \cdots & \cdots & \cdots & . \\ a_{(k-1)1} & \cdots & a_{(k-1)(k-1)} & 0 \\ a_{k1} & \cdots & a_{k(k-1)} & 1 \end{vmatrix} = \frac{\Delta_{k-1}}{\Delta_k} \neq 0.$$

Vì

$$\begin{aligned} \begin{vmatrix} \alpha_{11} & 0 & \cdots & 0 \\ \alpha_{21} & \alpha_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & 0 \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{vmatrix} &= \alpha_{11}\alpha_{22}\dots\alpha_{nn} \\ &= \frac{1}{\Delta_1}\frac{\Delta_1}{\Delta_2}\dots\frac{\Delta_{n-1}}{\Delta_n} = \frac{1}{\Delta_n} \neq 0 \end{aligned}$$

nên $(v_1, \dots, v_n)$ là một cơ sở của V. Mặt khác, từ $(*)$ ta suy ra

$$\forall 1 \le i \le n, \forall 1 \le j \le n, f(v_i, v_j) = \delta_{ij}\alpha_{ij}$$

nên $(v_1, \dots, v_n)$ là một cơ sở f-trực giao của V. Khẳng định trên được chứng minh. Trong cơ sở $\mathcal{B}' = (v_1, \dots, v_n)$, ma trận của Q có dạng chéo

$$[Q]_{\mathcal{B}'} = B = \begin{pmatrix} Q(v_1) & & 0 \\ & \ddots & \\ 0 & & Q(v_n) \end{pmatrix}.$$

Với mỗi $1 \le k \le n$, 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 $(u_1, \dots, u_k)$ và $(v_1, \dots, v_k)$. Gọi P_k là ma trận chuyển từ cơ sở trước sang cơ sở sau, ta có

$$\mathcal{B}_k = (P_k)^\top A_k P_k.$$

Chú ý rằng tương tự quá trình trực chuẩn hoá Gram–Schmidt ta có P_k là ma trận tam giác trên có 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 $k = 1,\dots,n$ $\Leftrightarrow$ $\Delta_k = \det(A_k) > 0$ với mọi $k = 1,\dots,n$. (ii) suy từ (i) cùng với nhận xét: Q xác định âm $\Leftrightarrow$ $-Q$ xác định dương.

Ví dụ 9.6.7. Đưa dạng toàn phương thực sau về dạng chuẩn tắc

$$Q(x,y,z) = 2x^2 + 9y^2 + 9z^2 + 8xy + 4xz + 12yz.$$

Chỉ ra cơ sở Q-chuẩn tắc và phép biến đổi toạ độ tương ứng. Từ đó xác định các chỉ số quán tính của Q. Xét xem Q có xác định dương hay xác định âm không.

Giải. Trước hết ta đưa Q về dạng chính tắc bằng thuật toán Lagrange:

$$\begin{aligned} Q(x,y,z) &= 2[x^2 + 2x(2y+z)] + 9y^2 + 9z^2 + 12yz \\ &= 2(x+2y+z)^2 - 2(2y+z)^2 + 9y^2 + 9z^2 + 12yz \\ &= 2(x+2y+z)^2 + y^2 + 4yz + 7z^2 \\ &= 2(x+2y+z)^2 + (y+2z)^2 + 3z^2 \\ &= [(x+2y+z)\sqrt{2}]^2 + (y+2z)^2 + (z\sqrt{3})^2 \end{aligned}$$

Thực hiện phép biến đổi tọa độ không suy biến

$$\begin{cases} x' = (x+2y+z)\sqrt{2} \\ y' = y + 2z \\ z' = z\sqrt{3} \end{cases} \Leftrightarrow \begin{cases} x = \frac{1}{\sqrt{2}}x' - 2y' + \sqrt{3}z' \\ y = y' - \frac{2}{\sqrt{3}}z' \\ z = \frac{1}{\sqrt{3}}z' \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -2 & \sqrt{3} \\ 0 & 1 & -\frac{2}{\sqrt{3}} \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

Ta đưa Q về dạng chuẩn tắc

$$Q(u) = x'^2 + y'^2 + z'^2 \tag{1}$$

với $u = x'u_1 + y'u_2 + z'u_3$, trong đó cơ sở Q-chuẩn tắc $\mathcal{B} = (u_1, u_2, u_3)$ định bởi

$$(\mathcal{B}_0 \to \mathcal{B}) = \begin{pmatrix} \frac{1}{\sqrt{2}} & -2 & \sqrt{3} \\ 0 & 1 & -\frac{2}{\sqrt{3}} \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

($\mathcal{B}_3$ là cơ sở chính tắc của $\mathbb{R}^3$) nghĩa là

$$u_1 = (\frac{1}{\sqrt{2}}, 0, 0); u_2 = (-2, 1, 0); u_3 = (\sqrt{3}, -\frac{2}{\sqrt{3}}, \frac{1}{\sqrt{3}}).$$

Từ (1) ta suy ra:

- Chỉ số dương quán tính của Q là 3.
- Chỉ số âm quán tính của Q là 0.
- Q xác định dương.

Ví dụ 9.6.8. Đưa dạng toàn phương sau về dạng chính tắc

$$Q(x, y, z) = 2x^2 + 9y^2 + \lambda z^2 + 8xy + 4xz + 12yz.$$

Xác định tham số $\lambda \in \mathbb{R}$ để Q không suy biến; Q xác định dương.

Giải. Biến đổi tương tự như trong Ví dụ 9.6.7 ta được

$$Q(x, y, z) = 2(x + 2y + z)^2 + (y + 2z)^2 + (\lambda - 6)z^2.$$

Thực hiện phép biến đổi tọa độ không suy biến

$$\begin{cases} x' = x + 2y + z \\ y' = y + 2z \\ z' = z \end{cases} \Leftrightarrow \begin{cases} x = x' - 2y' + 3z' \\ y = y' - 2z' \\ z = z' \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

ta đưa Q về dạng chuẩn tắc

$$Q(u) = 2x'^2 + y'^2 + (\lambda - 6)z'^2 \tag{2}$$

Từ (2) ta suy ra:

- Q không suy biến $\Leftrightarrow \lambda - 6 \neq 0 \Leftrightarrow \lambda \neq 6$.
- Q xác định dương $\Leftrightarrow \lambda - 6 > 0 \Leftrightarrow \lambda > 6$.

Ví dụ 9.6.9. Xác định tham số $\lambda \in \mathbb{R}$ để dạng toàn phương sau xác định dương

$$Q(x,y,z) = x^2 + \lambda y^2 + (\lambda+3)z^2 - 2xy + 4xz - 6yz.$$

Giải. Ma trận của dạng toàn phương Q là

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

Các định thức con chính của A là

$$\Delta_1 = 1;$$

$$\Delta_2 = \begin{vmatrix} 1 & -1 \\ -1 & \lambda \end{vmatrix} = \lambda - 1;$$

$$\Delta_3 = \begin{vmatrix} 1 & -1 & 2 \\ -1 & \lambda & -3 \\ 2 & -3 & \lambda+3 \end{vmatrix} = (\lambda-1)^2 - 1.$$

Theo tiêu chuẩn Sylvester ta có

$$Q \text{ xác định dương} \Leftrightarrow \begin{cases} \Delta_1 > 0 \\ \Delta_2 > 0 \\ \Delta_3 > 0 \end{cases} \Leftrightarrow \begin{cases} 1 > 0 \\ \lambda - 1 > 0 \\ (\lambda-1)^2 - 1 > 0 \end{cases} \Leftrightarrow \lambda > 2.$$

Từ chứng minh Định lý 9.6.6 ta có thuật toán tìm cơ sở Q-chính tắc

Thuật toán Jacobi đưa dạng toàn phương về dạng chính tắc

Giả sử dạng toàn phương Q có biểu thức của theo cơ sở $\mathcal{B} = (u_1, \ldots, u_n)$ định bởi

$$Q(u) = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{1 \le i < j \le n} 2a_{ij}x_ix_j$$

sao cho ma trận $A = (a_{ij})$ có các định thức con chính $\Delta_k \neq 0, \forall 1 \le k \le n$. Khi đó ta có thể xây dựng cơ sở Q-chính tắc $\mathcal{B}' = (v_1, \ldots, v_n)$ và xác định biểu thức của Q theo cơ sở $\mathcal{B}'$ như sau:

Bước 1: Tính các định thức con chính $\Delta_k\,(1 \le k \le n)$ và kiểm chứng $\Delta_k \neq 0, \forall 1 \le k \le n$.

Bước 2: Q có biểu thức của theo cơ sở $\mathcal{B}' = (v_1, \ldots, v_n)$ như sau:

$$Q(u) = \sum_{i=1}^{n} \frac{\Delta_{k-1}}{\Delta_k} y_i^2,$$

với $u = y_1v_1 + y_2v_2 + \cdots + y_nv_n, \Delta_0 = 1$.

Bước 3: Xây dựng cơ sở Q-chính tắc $\mathcal{B}' = (v_1, \ldots, v_n)$ như sau:

$$\begin{cases} v_1 &= \alpha_{11}u_1 \\ v_2 &= \alpha_{21}u_1 + \alpha_{22}u_2 \\ \multicolumn{2}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ v_n &= \alpha_{n1}u_1 + \alpha_{n2}u_2 + \cdots + \alpha_{nn}u_n \end{cases}$$

Trong đó $\alpha_{kk} = \dfrac{\Delta_{k-1}}{\Delta_k}$ và $A_k \begin{pmatrix} \alpha_{k1} \\ \vdots \\ \alpha_{kk} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$ $(1 \leq k \leq n)$.,

Ví dụ 9.6.10. Đưa dạng toàn phương thực sau về dạng chính tắc bằng thuật toán Jacobi:

$$Q(x, y, z) = 2x^2 + 9y^2 + 9z^2 + 8xy + 4xz + 12yz.$$

BÀI TẬP

Bài 9.1. Đưa các dạng toàn phương thực Q sau đây về dạng chính tắc. Chỉ ra cơ sở Q-chính tắc và phép biến đổi toạ độ không suy biến tương ứng. Xác định tập hợp $Q(\mathbb{R}^3)$:

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) $Q(u) = 3x_1x_2 + 4x_1x_3 + 12x_2x_3$

trong đó $u = (x_1, x_2, x_3) \in \mathbb{R}^3$

Bài 9.2. Đưa các dạng toàn phương thực Q sau đây về dạng chính tắc trực giao. Chỉ ra cơ sở Q-chính tắc trực giao và phép biến đổi toạ độ trực giao tương ứng. Xác định tập hợp $Q(\mathbb{R}^3)$:

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) $Q(u) = 6x_1^2 + 5x_2^2 + 7x_3^2 + 4x_1x_2 + 4x_1x_3$

trong đó $u = (x_1, x_2, x_3) \in \mathbb{R}^3$.

Bài 9.3. Xác định tham số thực m để dạng toàn phương thực Q sau đây xác định dương:

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) $Q(u) = x_1^2 + 4x_2^2 + x_3^2 + 2mx_1x_2 + 10x_1x_3 + 6x_2x_3$

trong đó $u = (x_1, x_2, x_3) \in \mathbb{R}^3$.

Bài 9.4. Cho Q là một dạng toàn phương trên không gian vectơ thực hữu hạn chiều V.

a) Giả sử $Q(u) \neq 0$ với mọi $u \in V \setminus \{0\}$. Chứng minh rằng Q hoặc xác định dương hoặc xác định âm.

b) Giả sử Q không suy biến và tồn tại $u \in V \setminus \{0\}$ sao cho $Q(u) = 0$. Chứng minh $Q(V \setminus \{0\}) = \mathbb{R}$.

c) Giả sử Q suy biến. Chứng minh rằng tồn tại $u \in V \setminus \{0\}$ sao cho $Q(u) = 0$. Cho ví dụ để thấy rằng trong trường hợp này $Q(V \setminus \{0\})$ có thể bằng $\mathbb{R}$, có thể khác $\mathbb{R}$.

Bài 9.5. Cho Q là dạng toàn phương thực định bởi

$$Q(x, y, z) = x^2 + 3y^2 - 4xy + 2xz + 2yz.$$

Tìm tất cả các không gian con W của $\mathbb{R}^3$ có $\dim W = 2$ và $Q(u) > 0$ với mọi $u \in W \setminus \{0\}$.

Bài 9.6. Cho Q là một dạng toàn phương xác định dương trên không gian vectơ thực hữu hạn chiều 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 9.7. 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 9.8. 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 (tức 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 9.9. Cho Q là một dạng toàn phương thực có biểu thức toạ độ theo cơ sở $\mathcal{B} = (u_1, \ldots, u_n)$ có dạng

$$Q(u) = f_1^2 + \cdots + f_s^2 - f_{s+1}^2 - \cdots - f_{s+t}^2$$

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 9.10. 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 9.11. 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