Universita' degli Studi di Roma "Tor Vergata"
Facolta' di Ingegneria: Elettronica

Corso di **Geometria ed Algebra**
Docente F. Flamini

Capitolo IV - § 3: Teorema Spettrale degli operatori autoaggiunti e Teorema di Sylvester

In questo paragrafo consideriamo il problema dell'esistenza di basi diagonalizzanti per le matrici di operatori simmetrici (equiv. autoaggiunti) o equivalentemente, da quanto visto in IV - § 1, per forme quadratiche reali.

Teorema. *(Teorema spettrale degli operatori simmetrici) Sia $\mathbb{R}^n$ munito del prodotto scalare standard $\cdot$. Sia $T : \mathbb{R}^n \to \mathbb{R}^n$ un operatore lineare simmetrico rispetto al prodotto scalare $\cdot$. Allora, esiste una base $\mathcal{F}$ di $\mathbb{R}^n$ che e' ortonormale e che e' costituita da autovettori di T.*

La dimostrazione del Teorema precedente si basa sui due seguenti risultati preliminari:

Lemma. 1 *Sia A una matrice simmetrica $n \times n$ a coefficienti reali. Allora A ha n autovalori reali contati con la rispettiva molteplicita'.*

Dimostrazione Lemma 1. Se A e' la matrice nulla, essa ha tutti gli autovalori uguali a zero, i.e. ha un unico autovalore nullo di molteplicita' n.

Assumiamo A matrice non identicamente nulla. Sia $P(t) = P_A(t) \in \mathbb{R}[t]$ il polinomio caratteristico di A. Ora $P(t)$ si puo' vedere come polinomio in $\mathbb{C}[t]$. Poiche' $\mathbb{C}$ e' algebricamente chiuso (cioe' ogni polinomio in $\mathbb{C}[t]$ di grado n ammette esattamente n radici in $\mathbb{C}$, contate con la relativa molteplicita'), allora $P(t)$ ha n radici in $\mathbb{C}$. Si tratta di dimostrare che, per ogni radice $\lambda \neq 0$ di $P(t)$, si ha $\lambda \in \mathbb{R}$.

Sia $T = T_A$ l'operatore lineare associato alla matrice A rispetto alla base canonica $\mathcal{E}_{\mathbb{C}}$ di $\mathbb{C}^n$. Poiche' λ e' un autovalore di T, esiste un autovettore $z \in \mathbb{C}^n$ di T tale che

$$T(z) = Az = \lambda z.$$

Moltiplichiamo a sinistra l'ultima eguaglianza per $\overline{z}^t$, dove $\overline{z} = \overline{x + iy} = x - iy$ e' il coniugato di z. Si ha

$$\overline{z}^t Az = \overline{z}^t \lambda z = \lambda \overline{z}^t z.$$

Ma $\overline{z}^t z = \sum_{i=1}^{n} \overline{z}_i z_i \in \mathbb{R} \setminus \{0\}$ per ogni $z \in \mathbb{C}^n$. Percio' poiche' $\lambda = \overline{z}^t Az / \overline{z}^t z$, e' sufficiente dimostrare che $\overline{z}^t Az \in \mathbb{R}$.

Ricordiamo che un numero complesso $w \in \mathbb{C}$ e' reale se, e solo se, $\overline{w} = w$. Percio', poiche' A e' una matrice reale e simmetrica,

$$\overline{\overline{z}^t A z} = z^t A \overline{z} = z^t A^t \overline{z} = (\overline{z}^t A z)^t = \overline{z}^t A z$$

dove l'ultima eguaglianza discende dal fatto che $\overline{z}^t A z$ e' uno scalare. $\square$

Lemma. 2 *Sia T un operatore autoaggiunto (equiv. simmetrico) e sia u un suo autovettore non nullo, relativo ad un autovalore $\lambda \in \mathbb{R}$. Allora, denotato con*

$$u^{\perp} := \{v \in \mathbb{R}^n | u \cdot v = 0\},$$

si ha

$$T(u^{\perp}) \subseteq u^{\perp}.$$

Dimostrazione Lemma 2. Per ipotesi, $T(u) = \lambda u$. Per ogni $v \in u^{\perp}$,

$$T(v) \cdot u = v \cdot T(u) = v \cdot \lambda u = \lambda v \cdot u = 0,$$

cioe' $T(v) \in u^{\perp}$.

$\square$

Veniamo alla dimostrazione del Teorema Spettrale.

Dimostrazione del Teorema Spettrale degli operatori simmetrici. Si procede per induzione su n.

Su $\mathbb{R}$, se prendiamo un qualsiasi $f \in \mathbb{R}$, tale che $|f| = 1$, si ha la base voluta.

Supponiamo $n \geq 2$ e assumiamo vero l'asserto per $n-1$. Se T e' l'operatore nullo, non c'e' niente da dimostrare perche' la sua matrice associata e' la matrice nulla. Sia T non identicamente nullo. In base al Lemma 1, T ha un autovalore reale non nullo. Sia λ_1 tale autovalore e sia f_1 l'autovettore ad esso associato. A meno di normalizzare, si puo' supporre che $||f_1|| = 1$.

Osserviamo che $f_1^{\perp}$ e' un sottospazio vettoriale di $\mathbb{R}^n$ di dimensione $n-1$ e che e' ortogonale a f_1. In base al Lemma 2, $T(f_1^{\perp}) \subseteq f_1^{\perp}$, cioe' la restrizione dell'operatore T al sottospazio $f_1^{\perp}$ definisce un operatore

$$T_1 : f_1^{\perp} \to f_1^{\perp}$$

che e' ovviamente ancora autoaggiunto, in quanto T_1 agisce come T sui vettori di $f_1^{\perp}$.

Per ipotesi induttiva, esiste una base ortonormale di $f_1^{\perp}$ che e' costituita da autovettori di T_1. Denotiamo tale base con $\{f_2, \dots f_n\}$. Per come e' costruito il tutto, $\mathcal{F} := \{f_1, f_2, \dots, f_n\}$ e' una base ortonormale di $\mathbb{R}^n$ perche' sono a due a due ortogonali fra loro e ciascuno di norma unitaria. Inoltre sono autovettori di T, quindi in base $\mathcal{F}$, l'operatore T ha matrice rappresentativa data da una matrice diagonale

$$D := \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

dove i λ_i, non tutti nulli, sono gli autovalori degli autovettori f_i, contati con la rispettiva molteplicita'. $\square$

Osservazione. a) Osserviamo che in base $\mathcal{E}$ l'operatore T aveva una matrice rappresentativa simmetrica A; nella base diagonalizzante $\mathcal{F}$, T ha matrice rappresentativa diagonale D. Se M e' la matrice cambiamento di base da $\mathcal{E}$ ad $\mathcal{F}$, i.e. $\mathcal{F} = \mathcal{E}M$, allora M e' ortogonale, dato che trasforma basi ortonormali in basi ortonormali. Percio' $M^t = M^{-1}$, cioe'

$$D = M^t A M.$$

Quest' ultima eguaglianza matriciale equivale a dire che la matrice simmetrica A e' *congruente* alla matrice diagonale D.
b) Non esiste un Teorema spettrale per gli operatori ortogonali. Ad esempio, la matrice

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

e' una matrice ortogonale perche' rappresenta, rispetto alla base canonica di $\mathbb{R}^2$, la rotazione di $\pi/2$ attorno all'origine; tale matrice non e' diagonalizzabile, dato che e' priva di autovalori reali.

Come diretta conseguenza della dimostrazione abbiamo:

Proposizione. *Sia $T : \mathbb{R}^n \to \mathbb{R}^n$ un operatore autoaggiunto (equiv. simmetrico). Se x e y sono due autovettori di T relativi a due autovalori distinti di T allora x e y sono ortogonali.*

Dimostrazione. Sia $T(x) = \lambda x$ e $T(y) = \mu y$, con $\lambda \neq \mu$. Ora

$$T(x) \cdot y = \lambda x \cdot y$$

e

$$x \cdot T(y) = \mu x \cdot y.$$

Poiche' T e' autoaggiunto, allora

$$\lambda x \cdot y = \mu x \cdot y.$$

Dal fatto che $\lambda \neq \mu$, segue che $x \cdot y = 0$.

$\square$

Abbiamo inoltre anche diverse formulazioni equivalenti del Teorema spettrale degli operatori autoaggiunti o simmetrici:

Proposizione. *(i) Sia Q una forma quadratica su $\mathbb{R}^n$. Allora esiste sempre una base ortonormale di $\mathbb{R}^n$ che diagonalizza la forma quadratica Q;*
(ii) Ogni matrice simmetrica reale $n \times n$ e' congruente ad uma matrice diagonale.

Sia T un operatore autoaggiunto e sia A la matrice simmetrica che rappresenta T rispetto alla base canonica di $\mathbb{R}^n$, considerato come munito del prodotto scalare standard. Sia $Q = Q_T$ la forma quadratica associata a T. Ricordando come abbiamo definito il *rango della forma quadratica* Q in IV - § 1, dalla dimostrazione precedente del Teorema Spettrale degli operatori autoaggiunti, notiamo che il rango di Q, e quindi di A, coincide con il numero dei coefficienti λ_i non nulli della matrice D congruente ad A, che e' l'espressione di Q (equivalentemente di T) in base $\mathcal{F}$.

Esercizio Sia $\mathbb{R}^3$ dotato del prodotto scalare standard e sia T l'operatore autoaggiunto che, rispetto alla base canonica $\mathcal{E}$, e' associato alla forma quadratica

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

per ogni $(x_1, x_2, x_3) \in \mathbb{R}^3$.
(i) Scrivere la matrice di T in base $\mathcal{E}$;
(ii) Determinare gli autovalori e la matrice di T rispetto ad una sua base $\mathcal{F}$ ortonormale e diagonalizzante;
(iii) Scrivere l'espressione della forma quadratica Q in base $\mathcal{F}$.

Svolgimento. (i) In base $\mathcal{E}$ la matrice di Q, e quindi di T, e' data da:

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

Sia A tale matrice.
(ii) Il polinomio caratteristico di A e' dato da $P(t) = det(A - tI_3) = -t(t - (1 + \sqrt{3}))(t - (1 - \sqrt{3}))$. Dunque A, e quindi T, ha tre autovalori distinti che sono

$$0, \ 1 + \sqrt{3}, \ 1 - \sqrt{3}.$$

Nella base $\mathcal{F}$ degli autovettori (normalizzati), T ha matrice associata data da:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 + \sqrt{3} & 0 \\ 0 & 0 & 1 + \sqrt{3} \end{pmatrix}.$$

(iii) In base $\mathcal{F}$, la forma quadratica si scrive

$$Q(y_1, y_2, y_3) = (1 + \sqrt{3})y_2^2 + (1 - \sqrt{3})y_3^2,$$

se (y_1, y_2, y_3) sono le coordinate di $\mathbb{R}^3$ rispetto alla base $\mathcal{F}$. Percio' la forma quadratica Q e' degenere, di rango 2.

Il Teorema Spettrale afferma l' esistenza di una base diagonalizzante per una forma quadratica reale. Pero', poiche' siamo in $\mathbb{R}$, si possono dare degli enunciati piu' precisi.

Teorema. *(Teorema di Sylvester) Sia $\mathbb{R}^n$ munito del prodotto scalare standard $\cdot$. Sia $Q : \mathbb{R}^n \to \mathbb{R}$ una forma quadratica reale di rango $r = rg(Q)$. Allora, esistono un numero intero non negativo $p \leq r$, dipendente solo da Q, ed una base $\mathcal{G}$ di $\mathbb{R}^n$, tali che rispetto a tale base, la matrice rappresentativa di Q e'*

$$\begin{pmatrix} I_p & O & O \\ O & -I_{r-p} & O \\ O & O & O \end{pmatrix},$$

dove il simbolo O denota matrici nulle degli ordini opportuni.

Equivalentemente, ogni matrice simmetrica reale A $n \times n$ e' congruente ad una matrice del tipo precedente, in cui $r = rg(A)$ e p dipendono solo da A.

Prima di vederne la dimostrazione, analizziamo alcune conseguenze.

Osservazione. a) Se in base $\mathcal{G}$ abbiamo coordinate $(y_1, y_2, \ldots, y_n)$, il Teorema di Sylvester asserisce che in base $\mathcal{G}$ la forma quadratica Q si scrive come

$$Q(y_1, y_2, \ldots, y_n) = y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2; \tag{1}$$

tale espressione si dice anche *forma canonica di Sylvester* della forma quadratica Q. Gli interi p e $r - p$ si dicono rispettivamente *indice di positivita'* e *indice di negativita'* della forma quadratica Q. La coppia $(p, r - p)$ e' la *segnatura* di Q.

Percio', le differenti forme canoniche di Sylvester sono le seguenti:

- *definita positiva*: ha segnatura $(n, 0)$ e quindi forma canonica di Sylvester $y_1^2 + \ldots + y_n^2$;
- *semidefinita positiva*: ha segnatura $(r, 0)$, con $r = rg(Q) \leq n$, e quindi forma canonica di Sylvester $y_1^2 + \ldots + y_r^2$;
- definita negativa: ha segnatura $(0, n)$ e quindi forma canonica di Sylvester $-y_1^2 - \ldots - y_n^2$;
- *semidefinita negativa*: ha segnatura $(0, r)$, con $r = rg(Q) \leq n$, e quindi forma canonica di Sylvester $-y_1^2 - \ldots - y_r^2$;
- *indefinita*: esiste $0 < p < r \leq n$ intero, tale che la segnatura e' $(p, r - p)$ e la forma canonica di Sylvester e' $y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2$.

Vedremo nel paragrafo successivo alcune formulazioni equivalenti delle precedenti definizioni.

Dimostrazione Teorema di Sylvester. L'equivalenza delle due affermazioni e' ovvia. Dimostriamo quindi la prima.

La forma quadratica Q corrisponde ad una matrice simmetrica $A = A_Q$ e quindi, nella base canonica di $\mathbb{R}^n$, ad un operatore simmetrico $T = T_Q$. Per il Teorema Spettrale, esiste una base ortonormale $\mathcal{F} = \{f_1, \ldots, f_n\}$ di autovettori di T; in tale base, la forma quadratica Q ha una matrice diagonale D, dove sulla diagonale principale ci sono gli autovalori dell'operatore T, i.e.

$$Q(z_1, z_2, \ldots, z_n) = \sum_{i=1}^{n} \lambda_i z_i^2.$$

Il numero dei coefficienti λ_i che sono diversi da zero, e' pari al rango di Q, e quindi dipende solo da Q e non dalla base prescelta.

Salvo riordinare la base, possiamo supporre che i primi r coefficienti siano diversi da zero e che tra essi figurino prima quelli positivi, che sono in numero $p \le r$. Si avra' dunque:

$$\lambda_1 = \alpha_1^2, \ldots, \lambda_p = \alpha_p^2, \lambda_{p+1} = -\alpha_{p+1}^2, \ldots, \lambda_r = -\alpha_r^2,$$

per opportuni numeri reali α_j, $1 \le j \le r$.

Prendiamo i vettori di $\mathbb{R}^n$ definiti da:

$$g_i = f_i/\alpha_i,\ 1 \le i \le r,\ g_j = f_j,\ r+1 \le j \le n.$$

Ovviamente essi formano una base ortogonale di $\mathbb{R}^n$; inoltre

$$g_i \cdot g_i = \lambda_i/\alpha_i^2,\ g_j \cdot g_j = 0.$$

Quindi, la forma quadratica ha l'espressione come in (1).

Resta da dimostrare che p dipende solo da Q e non dalla base $\mathcal{G}$. Supponiamo di avere in un'altra base $\mathcal{B} = \{b_1, \ldots, b_n\}$ che la forma quadratica Q si esprima come:

$$Q(w_1, w_2, \ldots, w_n) = w_1^2 + \ldots + w_s^2 - w_{s+1}^2 - \ldots - w_r^2,$$

per un opportuno intero $s \le r$. Dobbiamo verificare che $s = p$.

Supponiamo per assurdo che $s \ne p$; si puo' supporre cvhe $s < p$. Prendiamo i sottospazi di $\mathbb{R}^n$

$$U = < g_1, \ldots, g_p >,\ \ V = < b_{s+1}, \ldots, b_n > .$$

Poiche'

$$dim(U) + dim(V) = p + n - s > n$$

dalla formula di Grassmann segue che $U \cap V \neq \{0\}$. Esiste quindi un $z \in U \cap V$ non nullo. Ma allora si avrebbe contemporaneamente $Q(z) > 0$ e $Q(z) < 0$, che e' assurdo. Percio' $s = p$.

$\square$