Note per il corso di *Geometria e algebra lineare* 2024-25 LT in Informatica

9 Il teorema spettrale

Ricordiamo (cf. §2.9) il prodotto scalare euclideo di $\mathbb{R}^n$:

Definizione 1. Il *prodotto scalare euclideo (o canonico)* di $\mathbb{R}^n$ è la funzione che alla coppia $x, y \in \mathbb{R}^n$ associa il numero reale

$$x \cdot y = \sum_{i=1}^{n} x_i y_i = (x_1 \cdots x_n)(y_1 \cdots y_n)^T.$$

9.1 Basi ortonormali

Quando si studiano problemi che riguardano lunghezze di vettori, ortogonalità, angoli, conviene utilizzare delle basi di tipo particolare.

Definizione 2. Una base $\{u_1, \ldots, u_m\}$ di un sottospazio U di $\mathbb{R}^n$ è una *base ortonormale* di U se $u_i \cdot u_j = 0$ per $i \neq j$ e $u_i \cdot u_i = 1$ per $i = 1, \ldots, m$.

Ad esempio, la base canonica di $\mathbb{R}^n$ è ortonormale.

Se $\mathcal{B} = \{u_1, \ldots, u_m\}$ è una base ortonormale di U, le coordinate $x_1, \ldots, x_m$ di un vettore $v \in U$ rispetto a $\mathcal{B}$ si ottengono mediante il prodotto scalare

$$x_j = v \cdot u_j, \quad j = 1, \ldots, m.$$

Infatti, se $v = x_1 u_1 + \cdots + x_m u_m$, allora $v \cdot u_j = \sum_{i=1}^{m} x_i (u_i \cdot u_j) = x_j$.

Costruzione di basi ortonormali (Gram-Schmidt) A partire da una base $\{v_1, \ldots, v_m\}$ di U, si può costruire una base ortonormale di U. Si procede ricorsivamente, ponendo $u_1 = \frac{v_1}{\|v_1\|}$ e poi definendo in successione i vettori $u_2, \ldots, u_m$ con la formula

$$u_k = c \left(v_k - \sum_{i=1}^{k-1} (v_k \cdot u_i) u_i \right)$$

con $c \in \mathbb{R}$ scelto in modo che sia $\|u_k\| = 1$.

Esempio. Dalla base $\mathcal{B} = \{(1,0,1), (1,0,0), (2,1,0)\}$ di $\mathbb{R}^3$, con il prodotto canonico, mediante il procedimento di Gram-Schmidt si ottiene

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}(1,0,1) = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

$$v_2 - (v_2 \cdot u_1) u_1 = (1,0,0) - \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} (1,0,1) = \left(\frac{1}{2}, 0, -\frac{1}{2}\right) \text{ da cui}$$

$$u_2 = \frac{v_2 - (v_2 \cdot u_1) u_1}{\|v_2 - (v_2 \cdot u_1) u_1\|} = \sqrt{2}\left(\frac{1}{2}, 0, -\frac{1}{2}\right) = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right).$$

$$v_3 - (v_3 \cdot u_1) u_1 - (v_3 \cdot u_2) u_2 = (2,1,0) - \sqrt{2}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) - \sqrt{2}\left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right) = (0,1,0)$$

da cui

$$u_3 = (0,1,0).$$

Osservazione. Sia $P = M_{\mathcal{B}}^{\mathcal{E}}(Id)$ la matrice di transizione dalla base ortonormale $\mathcal{B}$ alla base canonica $\mathcal{E}$ di $\mathbb{R}^n$. Le colonne P^i di P sono gli elementi di $\mathcal{B}$. Quindi

$$P^i \cdot P^j = \begin{cases} 1 \text{ per } i = j \\ 0 \text{ per } i \neq j \end{cases} \quad \text{che equivale a} \quad P^T P = I_n.$$

Definizione 3. Una matrice quadrata $P \in M_n(\mathbb{R})$ è detta *ortogonale* se $P^T P = I_n$.

Ogni matrice ortogonale è invertibile, con inversa $P^{-1} = P^T$, poiché per il Teorema di Binet $\det(P^T)\det(P) = \det(I_n) = 1$, e quindi

$$(\det P)^2 = 1 \Rightarrow \det P = \pm 1.$$

Una matrice ortogonale di determinante 1 è anche detta *matrice di rotazione*, poiché le rotazioni del piano attorno all'origine e le rotazioni dello spazio attorno a una retta per l'origine sono rappresentate da matrici di questo tipo.

9.2 Il teorema spettrale

L'insieme degli autovalori di una matrice è anche chiamato *spettro* della matrice.

Proposizione 1. *Sia A una matrice reale simmetrica $n \times n$. Allora A ha n autovalori reali, contati con la loro molteplicità algebrica (cioè il polinomio caratteristico di A si scompone nel prodotto di n fattori lineari a coefficienti reali).*

Dimostrazione. Sia $T_A : \mathbb{C}^n \to \mathbb{C}^n$ l'endomorfismo definito da A. Sia $x \in \mathbb{C}^n$, $x \neq 0$, un autovettore di T_A con autovalore (complesso) λ. Per la simmetria di A

$$Ax = \lambda x \;\Rightarrow\; x^T A = \lambda x^T \;\Rightarrow\; \bar{x}^T A = \bar{\lambda}\bar{x}^T.$$

Dunque

$$\bar{x}^T A x = \bar{\lambda}\bar{x}^T x.$$

D'altra parte,

$$\bar{x}^T A x = \bar{x}^T \lambda x = \lambda \bar{x}^T x$$

e quindi $0 = (\bar{\lambda} - \lambda)\bar{x}^T x$, con $\bar{x}^T x = \sum_{i=1}^n |x_i|^2 > 0$. Ne segue che $\lambda = \bar{\lambda}$ è un numero reale. □

Teorema 1. (Teorema spettrale *o* Teorema degli assi principali) *Sia A una matrice reale simmetrica $n \times n$. Allora A è diagonalizzabile: esiste una base ortonormale di $\mathbb{R}^n$ costituita da autovettori di A, cioè esiste una matrice ortogonale P tale che $P^{-1}AP = D$ sia diagonale.*

Dimostrazione. Procediamo per induzione su n. Se $n = 1$, ogni vettore di $\mathbb{R}^n$ di norma uno è una base ortonormale di autovettori di A. Sia $n > 1$ e supponiamo l'enunciato vero per matrici di ordine $n - 1$. Per la Proposizione precedente, possiamo trovare un autovalore (reale) λ_1 di A, con autovettore u_1, che possiamo scegliere di norma uno. Sia $\mathcal{B}' = \{u_1, v_2, \ldots, v_n\}$ una base ortonormale di $\mathbb{R}^n$ con primo vettore u_1.

Sia $U = \langle u_1 \rangle$ e sia $\mathcal{P} = pr_{u_1}$ la proiezione ortogonale su u_1:

$$\mathcal{P}(v) = (v \cdot u_1)u_1.$$

$\mathcal{P}$ è una funzione lineare, con immagine U e nucleo il sottospazio $U^\perp$ di dimensione $n-1$ formato dai vettori ortogonali a u_1. Inoltre $T_A(v) = Av \in U^\perp$ per ogni $v \in U^\perp$, poiché

$$(Av) \cdot u_1 = (Av)^T u_1 = (v^T A^T)u_1 = v^T(Au_1) = v^T(\lambda_1 u_1) = \lambda_1(v \cdot u_1) = 0.$$

Dunque si può considerare la restrizione S dell'endomorfismo T_A al sottospazio $U^\perp$. La matrice A è simile alla matrice

$$A' = M_{\mathcal{B}'}(T_A) = \begin{bmatrix} \lambda_1 & w \\ 0 & B \end{bmatrix} = P'^{-1}AP'$$

con $P' = M^{\mathcal{E}}_{\mathcal{B}'}(id)$ matrice ortogonale. Dunque $A'^T = (P'^{-1}AP')^T = P'^{-1}A^TP' = P'^{-1}AP' = A'$ e anche A' è simmetrica, cioè $w = 0$ e $B = B^T$. Ma $B = M_{\mathcal{B}''}(S)$, con $\mathcal{B}'' = \{v_2, \ldots, v_n\}$.

Per l'ipotesi induttiva, essendo B simmetrica di ordine $n-1$, B è simile a una matrice diagonale, e si può trovare una base ortonormale $\{u_2, \ldots, u_n\}$ di $U^\perp$ costituita da autovettori di S, e quindi di T_A. I vettori $u_2, \ldots, u_n$ sono ortogonali a u_1, poiché appartengono a $U^\perp$. Quindi $\{u_1, u_2, \ldots, u_n\}$ è un insieme ortonormale, e dunque una base ortonormale di $\mathbb{R}^n$.

□

Osservazione. La relazione di similitudine $P^{-1}AP = D$ tra la matrice simmetrica A e la matrice diagonale

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

contiene la matrice di transizione $P = M^{\mathcal{E}}_{\mathcal{B}}(id)$ dalla base $\mathcal{B}$ alla base canonica $\mathcal{E}$ di $\mathbb{R}^n$, cioè una matrice ortogonale. Ne deriva che la relazione di similitudine tra A e D enunciata nel Teorema spettrale può essere riscritta nel modo seguente:

$$P^{-1}AP = P^TAP = D.$$

Il Teorema degli assi principali viene talvolta enunciato nel modo seguente: *ogni matrice reale simmetrica è "ortogonalmente simile" a una matrice reale diagonale*. La matrice ortogonale P del teorema può essere scelta di determinante 1. Infatti, se ha determinante -1, è sufficiente cambiare segno a una sua colonna (è ancora un autovettore di A) per ottenere $\det P = 1$.

Osservazione. Per ottenere una base ortonormale di autovettori di A, è sufficiente unire basi ortonormali degli autospazi di A. Infatti, autovettori relativi ad autovalori distinti di una matrice simmetrica sono sempre ortogonali: se $Ax = \lambda x$, $Ay = \mu y$, con $\lambda \neq \mu$, allora

$$\lambda(x \cdot y) = (\lambda x) \cdot y = (Ax) \cdot y = (Ax)^T y = x^T Ay = x^T \mu y = \mu(x \cdot y).$$

Dunque $(\lambda - \mu)(x \cdot y) = 0$, da cui $x \cdot y = 0$ e i due autovettori sono ortogonali.

Esempio. La matrice reale simmetrica $A = \begin{bmatrix} 7 & 1 & 1 \\ 1 & 7 & 1 \\ 1 & 1 & 7 \end{bmatrix}$ ha polinomio caratteristico

$$p_A(\lambda) = \det(A - \lambda I_4) = \det \begin{bmatrix} 7-\lambda & 1 & 1 \\ 1 & 7-\lambda & 1 \\ 1 & 1 & 7-\lambda \end{bmatrix} = \det \begin{bmatrix} 7-\lambda & 1 & 1 \\ 1 & 7-\lambda & 1 \\ 0 & \lambda-6 & 6-\lambda \end{bmatrix}$$

$$= (\lambda - 6) \det \begin{bmatrix} 7-\lambda & 1 & 1 \\ 1 & 7-\lambda & 1 \\ 0 & 1 & -1 \end{bmatrix} = (\lambda-6)(-(6-\lambda) - (7-\lambda)^2 + 1)$$

$$= -(\lambda - 6)(\lambda^2 - 15\lambda + 54) = -(\lambda - 6)^2(\lambda - 9).$$

Dunque ha autovalori 6 (doppio), e 9 (semplice), con autovettori

$$E(6) = N(A - 6I_4) = N \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = N \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \langle (1, 0, -1), (0, 1, -1) \rangle.$$

$$E(9) = N(A - 9I_4) = N \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} = N \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \langle (1, 1, 1) \rangle.$$

Per ottenere una base ortonormale di $E(6)$ basta applicare il procedimento di Gram-Schmidt: $\frac{1}{\sqrt{2}}(1, 0, -1)$ ha norma 1,

$$(0, 1, -1) - \frac{1}{2}(0, 1, -1)(1, 0, -1)^T(1, 0, -1) = (0, 1, -1) - \frac{1}{2}(1, 0, -1) = \left(-\frac{1}{2}, 1, -\frac{1}{2}\right)$$

che ha norma $\frac{\sqrt{6}}{2}$. Dunque $E(6)$ ha base ortonormale $\left\{\left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right)\right\}$, mentre $E(9)$ ha base ortonormale $\left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)\right\}$, che unite formano una base ortonormale costituita da autovettori di A.

La matrice

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix},$$

con colonne gli elementi della base ortonormale di $\mathbb{R}^3$, è una matrice ortogonale, di determinante 1, ed ha la proprietà

$$P^T AP = P^{-1}AP = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$

9.3 Risolvere sistemi lineari impossibili

Un sistema $Ax = b$ di m equazioni lineari in n incognite è risolubile se e solo se $b \in Im(T_A) = \langle A^1, \dots, A^n \rangle$. Se il sistema *non* è risolubile, si può cercare una n-upla $y \in \mathbb{R}^n$ che minimizzi la norma $\|Ax - b\|$, detta *soluzione ai minimi quadrati* del sistema:

$$\|Ay - b\| = \min_{x \in \mathbb{R}^n} \|Ax - b\|.$$

(Il minimo è zero se, e solo se, $Ay = b$, cioè y è soluzione)
Per semplicità, supponiamo che $rg(A) = n$, in modo che la soluzione, se esiste, sia unica. Per ottenere una soluzione ai minimi quadrati, consideriamo la proiezione ortogonale $b' = \pi(b)$ di b su $Im(T_A)$. Se $\mathcal{B} = \{u_1, \dots, u_n\}$ è una base ortonormale di $Im(T_A)$, la proiezione di $x \in \mathbb{R}^n$ sul sottospazio $Im(T_A)$ è data da

$$\pi(x) = \sum_{i=1}^{n} pr_{u_i}(x) = \sum_{i=1}^{n} (x \cdot u_i)u_i.$$

(Infatti $x - \pi(x)$ è ortogonale a ogni u_j: $(x - \pi(x)) \cdot u_j = x \cdot u_j - \sum_i (x \cdot u_i)(u_i \cdot u_j) = x \cdot u_j - x \cdot u_j = 0$ e quindi $x - \pi(x)$ è ortogonale ad ogni vettore di $Im(T_A)$.)

La proiezione ortogonale $b' = \pi(b)$ è il vettore del sottospazio $Im(T_A)$ più vicino a b: per il Teorema di Pitagora, se $z \in Im(T_A)$,

$$\|z - b\|^2 = \|(z - b') + (b' - b)\|^2 = \|z - b'\|^2 + \|b' - b\|^2 \geq \|b' - b\|^2,$$

con uguaglianza se e solo se $z = b'$.

L'unica soluzione y del sistema (risolubile) $Ax = b'$ è una soluzione ai minimi quadrati del sistema $Ax = b$:

$$\|Ay - b\| = \|b' - b\| = \|\pi(b) - b\| = \min_{z \in Im(T_A)} \|z - b\| = \min_{x \in \mathbb{R}^n} \|Ax - b\|.$$

Per trovare la soluzione ai minimi quadrati y di $Ax = b$, senza calcolare la proiezione ortogonale b', si può considerare il sistema quadrato associato (detto *sistema di equazioni normali* per y)

$$A^T Ax = A^T b.$$

Infatti la soluzione y di $Ax = b'$ risolve anche il sistema $A^T Ax = A^T b'$. Ma $A^T(b' - b) = 0$ poiché $b' - b$ è ortogonale alle colonne di A (righe di A^T).

Il sistema quadrato associato $n \times n$ ha matrice dei coefficienti simmetrica $A^T A$ di rango n:

$$A^T Az = 0 \Rightarrow (Az)^T Az = \|Az\|^2 = 0 \Rightarrow Az = 0 \Rightarrow z = 0$$

poiché $null(A) = n - rg(A) = 0$. Dunque $null(A^T A) = 0$ e $rg(A^T A) = n$. L'unica soluzione del sistema $A^T Ax = A^T b$ è dunque data da

$$y = (A^T A)^{-1} A^T b \quad (\textit{soluzione ai minimi quadrati di } Ax = b).$$

Esempio (Retta di regressione). Se $n = 1$, A e b sono vettori colonna $a = (a_1 \cdots a_m)^T \neq 0$ e $b = (b_1 \cdots b_m)^T$. Il sistema $Ax = b$ diventa

$$\begin{cases} a_1 x = b_1 \\ \vdots \\ a_m x = b_m \end{cases}$$

Il sistema ha una soluzione $y = t$ se e solo se gli m punti $(a_1, b_1), \ldots, (a_m, b_m)$ stanno sulla retta $y = tx$. La soluzione ai minimi quadrati fornisce il coefficiente angolare t della *retta di regressione*. Il sistema quadrato associato $A^T Ax = A^T b$ fornisce la soluzione

$$t = (a^T a)^{-1} a^T b = \frac{a \cdot b}{a \cdot a}.$$

(Se gli m punti non hanno baricentro 0, si applica il procedimento ai punti traslati

$$(a_1 - \bar{a}, b_1 - \bar{b}), \ldots, (a_m - \bar{a}, b_m - \bar{b})$$

dove $P = (\bar{a}, \bar{b})$ è il baricentro, cioè $\bar{a} = \frac{1}{m} \sum_i a_i$ e $\bar{b} = \frac{1}{m} \sum_i b_i$ e la retta di regressione è la retta per P con coefficiente angolare t).

Esercizio. Mostrare che la parabola il cui grafico approssima (non *interpola*!) i 4 punti $(1, 1)$, $(2, -6)$, $(-2, 4)$, $(-1, 1)$ ha equazione $p(x) = \frac{5}{3} - 2x - \frac{2}{3}x^2$.