MATRICI ORTOGONALI E FORME QUADRATICHE

Prodotto scalare

Introduciamo in $\mathbf{R}^n$, una generalizzazione del prodotto scalare studiato nel caso $n = 2, 3$, per poter definire distanza, modulo, misura angolare e in particolare ortogonalità, visto che questi concetti, per n qualunque, non hanno a priori un significato geometrico.

Definizione. Un **prodotto scalare** è una legge $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ che associa ad ogni coppia di vettori $\mathbf{v}, \mathbf{w}$ un numero reale che si indica con $\mathbf{v} \cdot \mathbf{w}$ o anche $< \mathbf{v}, \mathbf{w} >$, per cui valgono le seguenti proprietà:

1) $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$

2) $(m\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (m\mathbf{w}) = m(\mathbf{v} \cdot \mathbf{w})$

3) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

4) $\mathbf{v} \cdot \mathbf{v} \geq 0$, per ogni $\mathbf{v}$ e $\mathbf{v} \cdot \mathbf{v} = \mathbf{0}$ se e solo se $\mathbf{v} = \mathbf{0}$

Si verifica facilmente che il prodotto scalare ordinario, studiato nel caso $n = 2, 3$, soddisfa tali proprietà. Nello spazio vettoriale $\mathbf{R}^n$, con n qualsiasi, è utile estendere questa definizione di prodotto scalare (che d'altra parte non è l'unica definizione possibile di prodotto scalare).

Definizione. Nello spazio vettoriale $\mathbf{R}^n$, dati due vettori $\mathbf{v} = (x_1, \cdots, x_n)$ e $\mathbf{w} = (y_1, \cdots, y_n)$, diciamo loro **prodotto scalare euclideo** il numero reale $(x_1, \cdots, x_n) \cdot (y_1, \cdots, y_n) = \sum_{i=1}^n x_i y_i$. Ricordando il prodotto riga per colonna definito per le matrici, se identifichiamo $\mathbf{v}$ e $\mathbf{w}$ come vettori colonna di $\mathbf{R}^{n,1}$ e indichiamo con ${}^t\mathbf{v}$ il vettore trasposto di $\mathbf{v}$, possiamo scrivere

$$\mathbf{v} \cdot \mathbf{w} =^t \mathbf{v}\mathbf{w}$$

In analogia con quanto accade nello spazio della geometria ($n = 2, 3$) diremo che in $\mathbf{R}^n$ due vettori $\mathbf{v}, \mathbf{w}$ sono ortogonali se $\mathbf{v} \cdot \mathbf{w} = 0$; definiamo inoltre **modulo o norma** di un vettore $\mathbf{v}$ il numero $\|v\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = (\sum_{i=1}^n {x_i}^2)^{\frac{1}{2}}$.

Basi ortonormali

Consideriamo la base canonica di $\mathbf{R}^n$: $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$

Ogni vettore della base canonica ha norma 1:

$$\|\mathbf{e}_i\| = \sqrt{\mathbf{e}_i \cdot \mathbf{e}_i}$$

Inoltre i vettori della base canonica sono a due a due ortogonali:

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0, \ \textit{se } i \neq j$$

Tutto ciò si può riassumere con

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$$

dove δ_{ij} è la delta di Kronecker.

Ma esistono altre basi di $\mathbf{R}^n$ che godono di analoghe proprietà.

Definizione. Una base di $\mathbf{R}^n$ si dice **ortogonale** se è costituita da vettori a due a due ortogonali.

Definizione. Una base di $\mathbf{R}^n$ si dice **ortonormale** se è costituita da vettori a due a due ortogonali e di norma 1.

Supponiamo che $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ sia una base ortonormale di $\mathbf{R}^n$. Poichè $\mathcal{B}$ è una base, possiamo scrivere ogni vettore $\mathbf{v} \in \mathbf{R}^n$ in modo unico come combinazione lineare dei vettori di $\mathcal{B}$:

$$\mathbf{v} = c_1\mathbf{u}_1 + \ldots + c_n\mathbf{u}_n$$

Dato che $\mathcal{B}$ è ortonormale, è molto facile determinare i coefficienti di questa combinazione lineare; basta infatti calcolare il prodotto scalare di $\mathbf{v}$ con ciascun vettore di $\mathcal{B}$:

$$\mathbf{v} \cdot \mathbf{u}_i = c_1\mathbf{u}_1 \cdot \mathbf{u}_i + \ldots + c_i\mathbf{u}_i \cdot \mathbf{u}_i + \ldots + c_n\mathbf{u}_n \cdot \mathbf{u}_i = c_i$$

da cui segue il

Teorema. Rispetto a una base ortonormale $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ ogni vettore $\mathbf{v} \in \mathbf{R}^n$ si esprime come

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + \ldots + (\mathbf{v} \cdot \mathbf{u}_n)\mathbf{u}_n$$

.

Lemma. Se i vettori $\mathbf{v}_1, \cdots, \mathbf{v}_m \in \mathbf{R}^n$ sono non nulli e a due a due ortogonali, allora sono linearmente indipendenti.

Dimostrazione. Infatti moltiplicando scalarmente la relazione $a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m = \mathbf{0}$ per $\mathbf{v}_i$ al variare di $i = 1, \cdots, m$, si ottiene $a_i = 0$ per ogni i.

Matrici ortogonali

Definizione. Una matrice $P \in \mathbf{R}^{n,n}$ si dice **ortogonale** se i suoi vettori colonna sono una base ortonormale di $\mathbf{R}^n$.

Proposizione. Ogni matrice ortogonale $P \in \mathbf{R}^{n,n}$ è invertibile e l'inversa coincide con la trasposta: ${}^tP = P^{-1}$.

Osserviamo che da questo risultato segue anche che $I =^t PP = P^{-1}P = PP^{-1} = P^tP$, ossia in una matrice ortogonale anche le righe sono una base ortonormale di $\mathbf{R}^n$.

Esempio. Sono ortogonali le seguenti matrici:

$$P = \begin{pmatrix} cos\,\phi & -sin\,\phi \\ sin\,\phi & cos\,\phi \end{pmatrix} \in \mathbf{R}^{2,2}; \quad P^{-1} = \begin{pmatrix} cos\,\phi & sin\,\phi \\ -sin\,\phi & cos\,\phi \end{pmatrix} \in \mathbf{R}^{2,2}$$

$$Q = \begin{pmatrix} cos\,\phi & sin\,\phi \\ sin\,\phi & -cos\,\phi \end{pmatrix} \in \mathbf{R}^{2,2}; \quad R = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \in \mathbf{R}^{2,2}$$

Proprietà. Siano A e B matrici ortogonali di $\mathbf{R}^{n,n}$, allora A^{-1} e AB sono ortogonali.

Dimostrazione. È noto che ${}^t(AB) =^t B^tA$. Per ipotesi ${}^tAA = I, {}^tBB = I$, quindi ${}^tA = A^{-1}$ e ${}^t(A^{-1})A^{-1} = I$, cioe' A è ortogonale. Inoltre ${}^t(AB)(AB) =^t B^tAAB =^t BIB =^t BB = I$.

Teorema di Binet. Siano A e B matrici quadrate, allora $det(AB) = (detA)(detB)$

Corollario. Ogni matrice ortogonale ha determinante 1 oppure -1.

Dimostrazione. Se $P^tP = I$ si ha $1 = detI = det(P^tP) = det(^tP)detP = (detP)^2$, essendo $det(^tP) = detP$.

Definizione. Una matrice ortogonale con determinante 1 si dice **speciale** .

Osserviamo che ogni matrice di $\mathbf{R}^{2,2}$ del tipo $P = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ con la condizione $a^2 + b^2 = 1$, è ortogonale speciale e rappresenta una rotazione nel piano (con centro nell'origine) di un angolo ϕ che si ottiene dalle equazioni $cos\,\phi = a$, $sin\,\phi = b$. Analogamente si può mostrare che ogni matrice ortogonale speciale di $\mathbf{R}^{3,3}$ rappresenta una rotazione attorno a una retta fissata; per esempio una rotazione attorno all'asse delle z si ottiene con la matrice

$$\begin{pmatrix} cos\,\phi & -sin\,\phi & 0 \\ sin\,\phi & cos\,\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Più in generale si può dimostrare che:

Proposizione. Le matrici ortogonali speciali di $\mathbf{R}^{n,n}$ sono le matrici di passaggio tra basi ortonormali di $\mathbf{R}^n$ con lo stesso orientamento.

Matrici simmetriche reali

Tra le matrici quadrate a coefficienti reali, ne esiste un insieme particolare che risulta sempre diagonalizzabile: sono le matrici simmetriche. Ricordiamo che $S \in \mathbf{R}^{n,n}$ è una matrice simmetrica quando ${}^tS = S$.

Proposizione. Siano $\mathbf{v}_1, \mathbf{v}_2$ autovettori di una matrice simmetrica S corrispondenti a due distinti autovalori λ_1, λ_2; allora $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$

Dimostrazione. Tenendo presente che ${}^tS = S$, si ha: $(S\mathbf{v}_1)\cdot\mathbf{v}_2 =^t (S\mathbf{v}_1)\mathbf{v}_2 =^t \mathbf{v}_1^t S\mathbf{v}_2 =^t \mathbf{v}_1(S\mathbf{v}_2) = \mathbf{v}_1\cdot(S\mathbf{v}_2)$. Per ipotesi $S\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ e $S\mathbf{v}_2 = \lambda_2\mathbf{v}_2$, quindi $\lambda_1\mathbf{v}_1 \cdot \mathbf{v}_2 = \lambda_2\mathbf{v}_1 \cdot \mathbf{v}_2$ e $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, da cui la tesi, essendo $(\lambda_1 - \lambda_2) \neq 0$.

Si può dimostrare il seguente:

Teorema. Sia $S \in \mathbf{R}^{n,n}$ una matrice simmetrica, allora:

(a) le radici del polinomio caratteristico di S sono tutte reali;

(b) esiste una matrice ortogonale $P \in \mathbf{R}^{n,n}$ tale che $P^{-1}SP =^t PSP = D$, con D matrice diagonale, avente sulla diagonale principale gli autovalori di S scritti con la dovuta molteplicità. Segue in particolare che esiste una base ortonormale di $\mathbf{R}^n$ formata da autovettori di S.

Esempio. La matrice

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbf{R}^{3,3}$$

ha autovalori 1 e -1 di molteplicità rispettivamente 2 e 1; una base dell'autospazio V_{-1} è per esempio $(-1, 1, 0)$; una base per l'autospazio V_1 è per esempio $((0, 0, 1), (1, 1, 0))$. Questi vettori sono a due a due ortogonali, quindi per ottenere una base di $\mathbf{R}^3$ ortonormale, basta dividere ognuno di essi per il suo modulo. Una matrice ortogonale che diagonalizza A è perciò

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

e quindi (anche senza eseguire i calcoli) possiamo scrivere

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

.

Osservazione. Se per l'autospazio relativo a 1 si sceglie un'altra base, per esempio $((1,1,1),(1,1,0))$, che non è ortogonale, esistono procedimenti di ortonormalizzazione che permettono di ottenere da $((1,1,1),(1,1,0))$ una base ortonormale.

Metodo di ortonormalizzazione di Gram – Schmidt

Per ottenere una base ortonormale $(\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n)$ a partire da una base qualsiasi $(\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n)$ di $\mathbf{R}^n$, si pone innanzitutto

$$\mathbf{e}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$$

e si cerca un vettore $\mathbf{v}_2$ del tipo $\mathbf{v}_2 = \mathbf{u}_2 + \lambda\mathbf{e}_1$ che sia ortogonale a $\mathbf{e}_1$. Deve quindi essere

$$\mathbf{v}_2 \cdot \mathbf{e}_1 = \mathbf{u}_2 \cdot \mathbf{e}_1 + \lambda\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{u}_2 \cdot \mathbf{e}_1 + \lambda = 0$$

da cui

$$\mathbf{v}_2 = \mathbf{u}_2 - (\mathbf{u}_2 \cdot \mathbf{e}_1)\mathbf{e}_1$$

ma $\mathbf{u}_1, \mathbf{u}_2$ sono linearmente indipendenti, quindi pure $\mathbf{e}_1, \mathbf{u}_2$ lo sono, di conseguenza $\mathbf{v}_2 \neq \mathbf{0}$; si pone quindi

$$\mathbf{e}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}$$

Analogamente si cerca un vettore $\mathbf{v}_3$ del tipo $\mathbf{v}_3 = \mathbf{u}_3 + \lambda_1\mathbf{e}_1 + \lambda_2\mathbf{e}_2$ che sia ortogonale a $\mathbf{e}_1$ e $\mathbf{e}_2$; si trova $\mathbf{v}_3 = \mathbf{u}_3 - (\mathbf{u}_3 \cdot \mathbf{e}_1)\mathbf{e}_1 - (\mathbf{u}_3 \cdot \mathbf{e}_2)\mathbf{e}_2$ da cui si determina

$$\mathbf{e}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|}$$

e cosi' di seguito.

Teorema di Cayley – Hamilton

Sia $A \in \mathbf{K}^{n,n}$ è una matrice quadrata qualsiasi e sia $f(T) = a_0 + a_1T + \ldots + (-1)^mT^m$ un polinomio a coefficienti in $\mathbf{K}$. Poniamo per definizione $f(A) = a_0I + a_1A + \ldots + (-1)^mA^m$, dove I è la matrice identica di $\mathbf{K}^{n,n}$ e $A^2 = AA, A^3 = AAA$, ecc. Ne risulta che $f(A)$ è una matrice di $\mathbf{K}^{n,n}$.

Teorema (di Cayley-Hamilton). Sia $f(T)$ il polinomio caratteristico della matrice $A \in \mathbf{K}^{n,n}$, allora si ha che $f(A) = 0$.

Tale risultato è utilizzato per il calcolo, per esempio, della matrice inversa di una matrice. Sia infatti $f(T) = a_0 + a_1T + \ldots + (-1)^nT^n$ il polinomio caratteristico di una matrice $A \in \mathbf{K}^{n,n}$ invertibile. Si ha per il teorema:

$$f(A) = a_0I + a_1A + \ldots + (-1)^nA^n = 0$$

da cui

$$A(a_1I + \ldots + (-1)^{n-1}A^{n-1}) = -a_0I = -det(A)I$$

Essendo per ipotesi $det(A) \neq 0$ si ottiene

$$A^{-1} = \frac{-1}{det(A)}(a_1 I + \ldots + (-1)^{n-1} A^{n-1})$$

Forme quadratiche

Definizione. Una **forma lineare** è una applicazione lineare $f : \mathbf{R}^n \to \mathbf{R}$.

Segue che $f(x_1, \cdots, x_n) = a_1 x_1 + \cdots + a_n x_n = A\mathbf{v}$, dove $A \in \mathbf{R}^{1,n}$ è la matrice $(a_1 \cdots a_n)$ che rappresenta f e $\mathbf{v}$ il vettore colonna ${}^t(x_1, \cdots, x_n)$. Se A non è la matrice nulla, $Imf = \mathbf{R}$ e quindi $kerf$ sarà un sottospazio di dimensione $n-1$ di $\mathbf{R}^n$.

Definizione. Una **forma quadratica** è una funzione $q : \mathbf{R}^n \to \mathbf{R}$ del tipo $q(x_1 \cdots, x_n) = \sum_{i,j=1}^n a_{ij} x_i x_j$, cioè una combinazione lineare di tutti i monomi di secondo grado in $x_1, \cdots, x_n$ univocamente determinata da una matrice simmetrica $A = (a_{ij})$.

Per $n = 2$ otteniamo $q(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$, con $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$, ossia $q(x, y) = (x\ y)A \begin{pmatrix} x \\ y \end{pmatrix}$

Per $n = 3$, $q(x, y, z) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz$, con $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$, ossia

$q(x, y, z) = (x\ y\ z)A \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

In generale, $q(x_1 \cdots, x_n) =^t \mathbf{v}A\mathbf{v}$ e la funzione q può essere semplificata diagonalizzando A. Infatti, visto che A è una matrice simmetrica reale, esiste una matrice ortogonale P tale che $P^{-1}AP =^t PAP$ è una matrice diagonale D avente sulla diagonale principale gli autovalori di A: facciamo per esempio il caso $n = 2$, con il cambiamento di base $\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} X \\ Y \end{pmatrix}$:

$$q(x, y) = (x\ y)A \begin{pmatrix} x \\ y \end{pmatrix} =^t \begin{pmatrix} x \\ y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} =^t (P \begin{pmatrix} X \\ Y \end{pmatrix})A(P \begin{pmatrix} X \\ Y \end{pmatrix}) =^t \begin{pmatrix} X \\ Y \end{pmatrix} D \begin{pmatrix} X \\ Y \end{pmatrix} = (X\ Y)D \begin{pmatrix} X \\ Y \end{pmatrix}$$

Definizione. L'espressione $q(x_1 \cdots, x_n) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2 =^t \mathbf{v}D\mathbf{v}$ con D matrice diagonale, si dice **forma canonica** di q.

Per studiare il segno di una forma quadratica q, si può osservare che si ha sempre $q(0, \ldots, 0) = 0$, inoltre, utilizzando una forma canonica di q, si vede che il segno di q dipende dal segno degli autovalori della matrice A e dalla eventuale presenza dell'autovalore 0. Segue in particolare che:

Proposizione-Definizione.

Se gli autovalori di A sono tutti positivi, $q(x_1 \cdots, x_n) > 0$ per ogni $\mathbf{v} \neq \mathbf{0}$ e la forma quadratica si dice **definita positiva**.

Se gli autovalori di A sono tutti negativi, $q(x_1 \cdots, x_n) < 0$ per ogni $\mathbf{v} \neq \mathbf{0}$ e la forma quadratica si dice **definita negativa**.

Se gli autovalori di A sono tutti positivi o nulli, $q(x_1 \cdots, x_n) \geq 0$ per ogni $\mathbf{v}$ e la forma quadratica si dice **semidefinita positiva**.

Se gli autovalori di A sono tutti negativi o nulli, $q(x_1 \cdots, x_n) \leq 0$ per ogni $\mathbf{v}$ e la forma quadratica si dice **semidefinita negativa**.

La forma quadratica si dice **indefinita** o **non definita** se non vale nessuno dei casi precedenti.

Per lo studio del segno di una forma quadratica è anche utile la seguente:

Regola di Cartesio.

Dato un polinomio di grado n in una variabile scritto per potenze crescenti (o decrescenti), avente coefficienti reali e tutte le radici reali, il numero delle radici positive è uguale al numero di variazioni di segno dei coefficienti del polinomio, trascurando eventuali coefficienti nulli.

Esempio. Studiare il segno della forma quadratica $q(x, y, z) = x^2 + y^2 + 2z^2 + 2xz + 2yz$.

La forma quadratica $q(x, y, z) = x^2 + y^2 + 2z^2 + 2xz + 2yz$ è associata alla matrice

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \in \mathbf{R}^{3,3}$$

che ha polinomio caratteristico $-3T + 4T^2 - T^3$; per regola di Cartesio si può osservare che tale polinomio ha due radici positive, infatti, scritto per potenze di T crescenti, esso presenta due variazioni di segno; si vede inoltre che una radice vale 0; dunque si può concludere subito che la forma q è semidefinita positiva.

In alternativa si può procedere con uno studio più approfondito, cercando una forma canonica di q nel seguente modo: il polinomio caratteristico di A ha autovalori $1, 0, 3$; esiste quindi una base ortonormale formata da autovettori ad essi corrispondenti (per esempio quella che si costruisce normalizzando $\mathcal{B} = ((1, 1, -1), (-1, 1, 0), (1, 1, 2))$), dette allora x', y', z' le coordinate relative a tale base ortonormale, si ottiene la forma quadratica q nella forma canonica:

$$q(x', y', z') = (x')^2 + 3(z')^2$$

Ne segue quindi che la forma q è semidefinita positiva cioè è positiva dappertutto, tranne che nei punti della retta $x' = z' = 0$.