Forme bilineari e prodotti scalari

La generalizzazione dei concetti di prodotto scalare standard di due vettori e norma di un vettore di $\mathbb{R}^n$ è la nozione di forma bilineare simmetrica su uno spazio vettoriale V.

Definizione 1. *Una forma bilineare da V in $\mathbb{R}$ è un'applicazione $b : V \times V \to \mathbb{R}$ lineare in entrambe le variabili, ossia per ogni $\underline{u}, \underline{v}, \underline{w} \in V$, $k \in \mathbb{R}$*

$$b(\underline{u}+\underline{v},\underline{w}) = b(\underline{u},\underline{w}) + b(\underline{v},\underline{w}) \quad b(k\underline{u},\underline{v}) = kb(\underline{u},\underline{v})$$
$$b(\underline{u},\underline{v}+\underline{w}) = b(\underline{u},\underline{u}) + b(\underline{u},\underline{w}) \quad b(\underline{u},k\underline{v}) = kb(\underline{u},\underline{v})$$

La forma b si dice simmetrica se e solo se $b(\underline{u},\underline{v}) = b(\underline{v},\underline{u})$ $\forall \underline{u}$, $\underline{v} \in V$

In queste note considereremo solo il caso $V = \mathbb{R}^n$. Per le proprietà del prodotto di matrici, se $A \in \mathrm{Mat}(n)$ possiamo definire la forma ϕ_A indotta da A mediante la formula

$$\phi_A(\underline{x},\underline{y}) = \underline{x}^t A \underline{y}.$$

Se A è una matrice simmetrica, ϕ_A sarà una forma simmetrica, ad esempio il prodotto scalare standard di $\mathbb{R}^n$ può essere scritto come $\langle \underline{u},\underline{v}\rangle_{\mathbb{R}^n} = \underline{u}^t I \underline{v}$ dove I è la matrice identità, dunque è la forma simmetrica indotta dalla matrice identità. Viceversa, con una dimostrazione simile a quella vista per le applicazioni lineari si può dimostrare che tutte le forme bilineari sono indotte da una matrice A. Infatti scelta una base, ad esempio la base canonica, possiamo associare ad ogni forma bilineare una matrice poichè *ogni applicazione bilineare è determinata in modo unico dai suoi valori su tutte le possibili coppie ordinate di elementi di una base di $\mathbb{R}^n$.*

Osservazione 2. *Se $A = (a_{ij})$ allora $\phi_A(\underline{e}_i,\underline{e}_j) = a_{ij}$*

Proposizione 3. *Sia $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ una forma bilineare. Definiamo la matrice A con coefficienti $a_{ij} = \psi(\underline{e}_i,\underline{e}_j)$ allora $\psi = \phi_A$, ossia per ogni coppia di vettori $\underline{u},\underline{v} \in \mathbb{R}^n$ abbiamo $\psi(\underline{u},\underline{v}) = \phi_A(\underline{u},\underline{v})$. A si dice la matrice associata all'applicazione bilineare ψ rispetto alla base canonica di $\mathbb{R}^n$.*

dim: Scriviamo $\underline{u} = x_1\underline{e}_1 + \ldots x_n\underline{e}_n$, $\underline{v} = y_1\underline{e}_1 + \ldots + y_n\underline{e}_n$, allora, applicando la linearità prima in una variabile, poi nell'altra

$$\begin{aligned}
&\psi(\underline{u},\underline{v}) = \psi(x_1\underline{e}_1 + \ldots x_n\underline{e}_n,\underline{v}) = x_1\psi(\underline{e}_1,\underline{v}) + \ldots + x_n\psi(\underline{e}_n,\underline{v}) = \\
=&x_1\psi(\underline{e}_1, y_1\underline{e}_1 + \ldots + y_n\underline{e}_n) + \ldots + x_n\psi(\underline{e}_n, y_1\underline{e}_1 + \ldots + y_n\underline{e}_n) = \\
=&x_1(y_1\psi(\underline{e}_1,\underline{e}_1) + \ldots + y_n\psi(\underline{e}_1,\underline{e}_n)) + \ldots + x_n(y_1\psi(\underline{e}_n,\underline{e}_1) + \ldots + y_n\psi(\underline{e}_n,\underline{e}_n)) = \\
=&x_1y_1a_{11} + x_1y_2a_{12} + \ldots + x_1y_na_{1n} + \ldots x_ny_1a_{n1} + \ldots x_ny_na_{nn} = \\
=&\sum_{i,j=1}^{n} x_iy_ja_{ij} = \underline{x}^t A \underline{y} = \phi_A(\underline{u},\underline{v})
\end{aligned}$$

Sempre come nel caso delle applicazioni lineari, anche per le forme bilineari possiamo definire la matrice associata ad una forma bilineare rispetto ad una base qualsiasi.

Definizione 4. *Sia $\psi : \mathbb{R}^n \times \mathbb{R}^n \to R$ una forma bilineare, $\mathcal{B} = \{\underline{u}_1, \ldots, \underline{u}_n\}$ una base di $\mathbb{R}^n$. Allora la matrice di ψ rispetto alla base $\mathcal{B}$ è la matrice*

$$A = (a_{ij}) = \psi(\underline{u}_i, \underline{u}_j).$$

Se $\underline{\alpha}, \underline{\beta}$ sono le coordinate di $\underline{v}, \underline{w}$ rispetto alla base $\mathcal{B}$ allora si avrà

$$\psi(\underline{v}, \underline{w}) = \underline{\alpha}^t A \underline{\beta}. \tag{1}$$

Con la stessa dimostrazione della proposizione 3 possiamo dimostrare che ψ è determinata in maniera univoca dalla sua matrice associata rispetto ad una base qualsiasi, quindi se una matrice A ha la proprietà (1) allora è la matrice associata a ψ rispetto alla base $\mathcal{B}$

Esempio 1. Consideriamo la base di $\mathbb{R}^3$

$$\mathcal{B} = \left\{ \underline{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \underline{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \underline{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

La matrice con coefficienti $a_{ij} = \langle \underline{v}_i, \underline{v}_j \rangle_{\mathbb{R}^3}$ è la matrice

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

che è la matrice del prodotto scalare standard di $\mathbb{R}^3$ rispetto alla base $\mathcal{B}$ e ha questa proprietà: consideriamo i vettori $\underline{x}_i = (a_i, b_i, c_i)^t$ con $i = 1, 2$ di coordinate $\underline{\xi}_i = (a_i - b_i, b_i - c_i, c_i)^t$ rispetto alla base $\mathcal{B}$. Il prodotto

$$\xi_1^t A \xi_2 = a_1 a_2 + b_1 b_2 + c_1 c_2 = \langle \underline{x}_1, \underline{x}_2 \rangle_{\mathbb{R}^3}$$

da come risultato il loro prodotto scalare.

Proposizione 5. *Data una forma bilineare $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, sia A la sua matrice associata rispetto a una base $\mathcal{B}$ e sia $\mathcal{B}'$ un'altra base di V. Allora la matrice associata a ψ rispetto a $\mathcal{B}'$ è la matrice $A' = ({}_{\mathcal{B}}M_{\mathcal{B}'})^t A \, {}_{\mathcal{B}}M_{\mathcal{B}'}$, dove ${}_{\mathcal{B}}M_{\mathcal{B}'}$, qui per brevità M, è la matrice di passaggio da $\mathcal{B}$ a $\mathcal{B}'$*

dim: La dimostrazione è simile a quella vista per le applicazioni lineari, essenzialmente la formula del cambiamento di coordinate. Denotiamo con $\underline{u}_i$ i vettori della base $\mathcal{B}$, con $\underline{v}_i$ quelli della base $\mathcal{B}'$, consideriamo due vettori

$\underline{x}, \underline{y} \in \mathbb{R}^n$. Se $M = _{\mathcal{B}} M_{\mathcal{B}'}$ è la matrice di passaggio, denotando come al solito con $_{\mathcal{D}}(\underline{z})$ le coordinate di un vettore $\underline{z}$ rispetto ad una base $\mathcal{D}$ abbiamo

$$_{\mathcal{B}}(\underline{x}) = M\, _{\mathcal{B}'}(\underline{x})$$

e la stessa equazione per le coordinate di $\underline{y}$. Allora

$$\psi(\underline{x}, \underline{y}) = _{\mathcal{B}} (\underline{x})^t A _{\mathcal{B}}(\underline{x}) = (M\, _{\mathcal{B}'}(\underline{x}))^t AM\, _{\mathcal{B}'}(\underline{x}) = _{\mathcal{B}'} (\underline{x})^t (M^t AM) _{\mathcal{B}'}(\underline{x}).$$

In particolare, per $\underline{x} = \underline{v}_i$, $\underline{y} = \underline{v}_j$, le coordinate rispetto a $\mathcal{B}'$ sono i vettori della base canonica $\underline{e}_i, \underline{e}_j$, dunque il coefficiente i, j della matrice $M^t AM$ è proprio $\psi(\underline{v}_i, \underline{v}_j)$ dunque $M^t AM$ è la matrice associata a ψ rispetto alla base $\mathcal{B}'$.

La relazione tra matrici che rappresentano la stessa forma bilineare rispetto a basi diverse si chiama congruenza.

Definizione 6. *Siano $A, B \in \mathrm{Mat}(n)$ due matrici quadrate. B si dice congruente ad A se esiste una matrice invertibile M tale che*

$$B = M^t AM$$

Esercizio. Dimostrare che

a) La congruenza di matrici è una relazione di equivalenza.

b) Matrici congruenti hanno lo stesso rango

Osservazione 7. *Congruenza e similitudine sono relazioni diverse sulle matrici: ad esempio*

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

ma la matrice A non è simile alla matrice B. Però se la matrice M è ortogonale, la congruenza è anche una similitudine.

Esempio 2. Riprendendo l'Esempio 1, detta M la matrice di passaggio dalla base canonica alla base $\mathcal{B}$, abbiamo che la matrice A associata al prodotto standard di $\mathbb{R}^3$ rispetto a $\mathcal{B}$ è proprio $M^t IM = M^t M$, visto che I è la matrice del prodotto standard rispetto alla base canonica.

Forme quadratiche

Una forma bilineare simmetrica simmetrica si dice anche *prodotto scalare*, la generalizzazione del concetto di norma è

Definizione 8. *Sia $\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ un prodotto scalare su $\mathbb{R}^n$ (o più in generale su uno spazio V), la forma quadratica associata ad α è la funzione $Q_\alpha : \mathbb{R}^n \to \mathbb{R}$, $Q_\alpha(\underline{v}) = \alpha(\underline{v}, \underline{v})$*

Ad esempio, la forma quadratica associata al prodotto scalare standard di $\mathbb{R}^n$ è la norma al quadrato di un vettore.

Definizione 9. *Sia* $\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *una forma bilineare simmetrica. Diciamo che* α *è*

- *Degenere se* $\exists \underline{u} \neq \underline{0} \in \mathbb{R}^n |\ \forall v \in \mathbb{R}^n\ \alpha(\underline{u}, \underline{v}) = 0$
- *Definita positiva (negativa) se* $\forall \underline{u} \in \mathbb{R}^n\ \underline{u} \neq 0 \Rightarrow \alpha(\underline{u}, \underline{u}) > 0 \quad (< 0)$
- *Semidefinita positiva (negativa) se* $\forall \underline{u} \in \mathbb{R}^n\ \Rightarrow \alpha(\underline{u}, \underline{u}) \geq 0 \quad (\leq 0)$
- *Indefinita se esistono* $\underline{v}, \underline{w} \in \mathbb{R}^n$ *tali che* $\alpha(\underline{v}, \underline{v}) > 0$, $\alpha(\underline{w}, \underline{w}) < 0$, *ossia* B *non è definita positiva o negativa.*

La stessa terminologia si usa anche per le forme quadratiche associate.

Esempio 3. In $\mathbb{R}^3$, ponendo $\alpha(\underline{u}, \underline{v}) = \underline{u}^t A_i \underline{v}$, dove A_i è una delle matrici

$$A_1 = -I_3, \qquad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

otteniamo forme rispettivamente definita negativa, semidefinita positiva, indefinita su $\mathbb{R}^3$.

La forma A_2 è degenere e per questa forma abbiamo $(\mathbb{R}^3)^\perp = L[\underline{e}_3]$ e $L[\underline{e}_1, \underline{e}_3]^\perp = L[\underline{e}_2, \underline{e}_3]$, quindi l'intersezione tra un sottospazio E e il sottospazio dei vettori ortogonali a tutti i vettori di E può essere non vuota.

Anche quando la forma α è indefinita ci sono vettori non nulli tali che $\alpha(\underline{u}, \underline{u}) = 0$. Ad esempio se $\underline{u} = (1, 0, 1)^t$, $\underline{u}^t A_3 \underline{u} = 0$, e dunque $L[\underline{u}] \cap L[\underline{u}]^\perp$ contiene vettori non nulli (in particolare il vettore $\underline{u}$ stesso). Però, quando la forma è non degenere indefinita, possiamo sempre trovare un vettore $\underline{v}$ tale che $\alpha(\underline{u}, \underline{v}) \neq 0$.

Esercizio. Sia A una matrice simmetrica

- Mostrare che se A non ha rango massimo, esistono vettori non nulli $\underline{v} \in \mathbb{R}^n$ tali che $\forall \underline{u} \in \mathbb{R}^n \quad \underline{u}^t A \underline{v} = 0$, ossia che sono 'ortogonali' a tutti i vettori di $\mathbb{R}^n$, inclusi se stessi.
- Siano $\lambda > 0, \mu < 0$ autovalori di A, allora esiste un vettore $\underline{u} \in \mathbb{R}^n$ t.c. $\underline{u}^t A \underline{u} = 0$.

Il nome forma quadratica deriva dal fatto che possiamo rappresentare ogni polinomio omogeneo di grado 2 nelle coordinate $(x_1, \ldots, x_n)^t$ con la forma quadratica associata rispetto ad una forma simmetrica bilineare con una formula del tipo

$$p(\underline{x}) = \alpha(\underline{x}, \underline{x}) = \underline{x}^t A \underline{x}$$

dove A è la matrice di α rispetto alla base canonica.

Viceversa possiamo interpretare un qualsiasi polinomio omogeneo (ossia solo con termini di grado due) in n incognite come la forma quadratica associata ad una forma bilineare simmetrica, ad esempio abbiamo

$$x^2 - 2y^2 + z^2 + 3xz + 2yz = (x, y, z) \begin{pmatrix} 1 & 0 & 3/2 \\ 0 & -2 & 1 \\ 3/2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

In generale gli elementi a_{ii} della diagonale sono i coefficienti dei termini puri x_i^2, i coefficienti a_{ij} $i \neq j$ sono metà del coefficiente del prodotto misto $x_i x_j$.

Abbiamo visto che per risolvere i sistemi lineari è utile fare dei cambiamenti di coordinate in modo che le equazioni del sistema diventino più semplici: ad esempio se la matrice del sistema $A\underline{x} = \underline{b}$ è diagonalizzabile, passando ad una base di autovettori la risoluzione del sistema è immediata. Allo stesso modo, trovando una matrice congruente alla matrice di una forma quadratica diagonale, semplifichiamo lo studio dei polinomi omogenei di secondo grado e dei loro zeri, potendoci ricondurre con un cambiamento di coordinate, allo studio di poche equazioni 'standard'.

Esempio 4. Sia $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, la forma quadratica associata alla forma bilineare definita dalla matrice è

$$(x, y) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2xy$$

Con il cambiamento di variabili $X = \frac{1}{\sqrt{2}}(x + y)$, $Y = \frac{1}{\sqrt{2}}(x - y)$ abbiamo che $2xy = X^2 - Y^2$. Questo cambiamento corrisponde a cambiare base dalla base canonica alla base $\mathcal{B} = \{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^t, (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})^t\}$. Detta M la matrice di cambiamento abbiamo $M^t A M = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ che ha come forma quadratica associata $X^2 - Y^2$.

Abbiamo il seguente teorema

Teorema 10 (Matrice in forma di Sylvester)**.** *Data una forma bilineare simmetrica $\alpha : V \times V \to \mathbb{R}$ esiste una base $\mathcal{B} = \{\underline{u}_1, \ldots, \underline{u}_n\}$ di V tale che la matrice associata ad α rispetto a $\mathcal{B}$ sia diagonale con i primi p elementi della diagonale uguali a 1, i successivi q elementi della diagonale uguali a -1 e gli ultimi $n - p - q$ elementi nulli. Inoltre due matrici sono congruenti se e solo se hanno la stessa forma di Sylvester.*

La dimostrazione si basa sul teorema spettrale: diagonalizzando rispetto ad una base di autovettori ortonormali troviamo una matrice congruente diagonale, e 'normalizzando' questi autovettori dividendo per la radice del valore assoluto dell'autovalore corrispondente, troviamo una base rispetto alla quale la matrice è nella forma di Sylvester. Questa dimostrazione fa anche vedere che i numeri p, q corrispondono al numero di autovalori positivi e negativi.

Esempio 5. Consideriamo la matrice $A = \left(\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right)$, gli autovalori sono $4, -2$ corrispondenti agli autovettori di norma 1 $\underline{u}_1 = \frac{1}{\sqrt{2}}(1,1)^t, \underline{u}_2 = \frac{1}{\sqrt{2}}(1,-1)^t$ Se M è la matrice con colonne questi due vettori abbiamo che $M^{-1}AM = M^tAM$ è una matrice diagonale con gli autovalori $4, -2$ sulla diagonale. Ora consideriamo i vettori $\frac{1}{2}\underline{u}_1, -\frac{1}{\sqrt{2}}\underline{u}_2$, detta N la matrice con colonne $\underline{u}_1, \underline{u}_2$, N^tAN è in diagonale con $1, -1$ sulla diagonale.

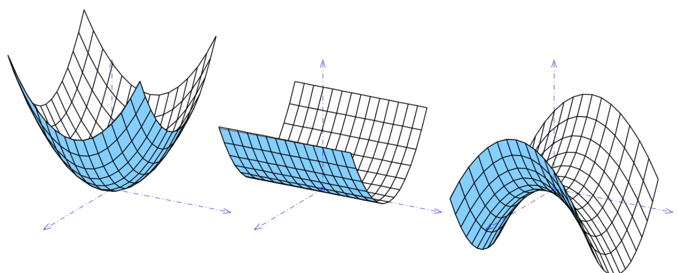
Definizione 11. *La segnatura di una matrice simmetrica è la coppia di numeri (p, q) ed è uguale al numero di autovalori positivi e negativi (eventualmente se la dimensione n non è sottintesa si può usare la terna (positivi,negativi,nulli).*

Applicazione all'analisi

Nello studio delle funzioni di più variabili, si usano le derivate parziali $\frac{\partial f}{\partial x_i}$. Se abbiamo n variabili tutte le possibili derivate seconde sono n^2 e quindi formano una matrice di ordine n con coefficienti $\frac{\partial^2 f}{\partial x_i \partial x_j}$, che nel caso di funzioni sufficientemente regolari è simmetrica perchè l'ordine in cui derivo non influenza il risultato. L'analogo dell'espansione di Taylor è (nel caso di una funzione in due incognite $f(x,y)$ nel punto $(0,0)$)

$$f(x,y) = f(0,0) + (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}) \begin{pmatrix} x \\ y \end{pmatrix} + (x,y) \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial x^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + R(x,y)$$

(tutte le derivate si intendono calcolate in (0,0)) con $R(x,y)$ che va a zero come una funzione cubica della distanza di (x,y) da (0,0). Se $(0,0)$ è un punto stazionario, le derivate prime si annullano e localmente il grafico $z = f(x,y)$ ha la forma di un paraboloide o di un cilindro parabolico. infatti scegliendo una base di autovettori abbiamo che se i due autovalori sono entrambi positivi o negativi abbiamo un massimo o un minimo ($z = \pm(\lambda x^2 + \mu y^2)$ $\lambda, \mu > 0$), se i due autovalori sono di segno opposto abbiamo un punto di sella ($z = \lambda x^2 - \mu y^2$ $\lambda, \mu > 0$) se uno è nullo abbiamo un cilindro ($z = \lambda x^2$).



$z = x^2 + y^2$ $\quad$ $z = x^2$ $\quad$ $z = x^2 - y^2$

Applicazione allo studio delle coniche e delle quadriche

Una quadrica in uno spazio di dimensione n è il luogo degli zeri di un polinomio di secondo grado in n incognite. Nel piano si parla di coniche. Abbiamo visto che possiamo descrivere polinomi di secondo grado *omogenei* usando la forma quadratica associata ad una forma bilineare simmetrica; per descrivere polinomi di grado 2 qualsiasi ricorriamo ad uno stratagemma: usiamo una matrice di ordine $n+1$ e dei vettori di $\mathbb{R}^{n+1}$ con l'ultima coordinata uguale a 1. Ad esempio nel caso delle coniche abbiamo che

$$p(x,y) = ax^2 + bxy + cy^2 + dx + ey + f = (x,y,1)\begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

In pratica, è come se aggiungessimo una terza incognita 'invisibile' e considerassimo il termine di grado zero come il coefficiente di questa incognita al quadrato, i termini di grado uno come termini di grado due dati dal prodotto di questa incognita con le altre. Il minore formato dalle prime due righe e due colonne si chiama parte principale della conica e rappresenta la parte omogenea di secondo grado di $p(x,y)$. La segnatura della parte principale $\mathcal{A}$ determina il tipo di conica: le coniche del piano si dividono in ellissi reali o immaginarie (includiamo anche le circonferenze) se $\mathcal{A}$ è definita positiva o negativa, iperboli se $\mathcal{A}$ è indefinita non degenere, e parabole se $\mathcal{A}$ è semidefinita (rango di $\mathcal{A}$ è uno). Se il determinante di tutta la matrice è nullo, avremo la stessa classificazione ma ma le coniche saranno degeneri, ossia coppie di rette (ad esempio $x^2 - y^2 = (x+y)(x-y) = 0$).