

Università degli Studi di Palermo
Facoltà di Economia
CdS Statistica per l'Analisi dei Dati

Appunti del corso di Matematica

**25 - Funzioni di più
Variabili – Introduzione**

Anno Accademico 2013/2014

M. Tumminello e A. Consiglio

1. Introduzione

Nella prima parte del corso abbiamo analizzato le proprietà delle funzioni scalari, $f : \mathbb{R} \rightarrow \mathbb{R}$. In questa seconda parte studieremo alcune delle proprietà delle funzioni che hanno come dominio e codominio lo spazio vettoriale \mathbb{R}^n . In generale tali funzioni sono indicate da $\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e il simbolo $\vec{f}(\cdot)$ denota, appunto, un vettore di funzioni che trasforma vettori di \mathbb{R}^n in vettori di \mathbb{R}^m . In particolare, la funzione $\vec{y} = \vec{f}(\vec{x})$ è detta *funzione vettoriale* (*vector-valued function* o semplicemente *vector function*). ed è descritta da un vettore m -dimensionale di funzioni che associa ad $\vec{x} \in \mathbb{R}^n$ un vettore $\vec{y} \in \mathbb{R}^m$.

Per esempio:

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \end{bmatrix} = \begin{bmatrix} \sin(x_1) \cdot \cos(x_2) \\ x_1^2 + x_2^2 + x_3^2 \end{bmatrix}$$

è una funzione che trasforma vettori di \mathbb{R}^3 in vettori di \mathbb{R}^2 , $\vec{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$.

Per gli obiettivi del nostro corso studieremo principalmente funzioni scalari di variabile vettoriale, $f : \mathbb{R}^n \rightarrow \mathbb{R}$, ossia funzioni che trasformano vettori di \mathbb{R}^n in scalari di \mathbb{R} . Da un punto di vista grafico, possiamo rappresentare, in un grafico 3D, funzioni $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ il cui dominio è il piano coordinato XY e il codominio l'asse Z . Per esempio la funzione $z = f(x, y) = 2x + 3y - 1$ definisce un piano la cui equazione è $z - 2x - 3y + 1 = 0$. In altre parole, la funzione $f(x, y)$ definisce in forma esplicita il piano di equazione data. Nella figura di seguito sono mostrate le funzioni $z = 2x + 3y - 1$ e $z = 2x - y$.

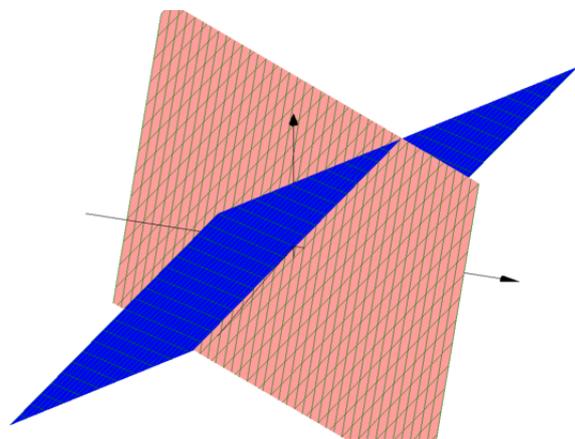


FIGURE 1. Rappresentazione delle funzioni $z = 2x + 3y - 1$ e $z = 2x - y$.

Nella figura di seguito sono invece rappresentate le funzioni $z = -2x^2 + xy - 3y^2$ e $z = 2x + 3y - 1$.

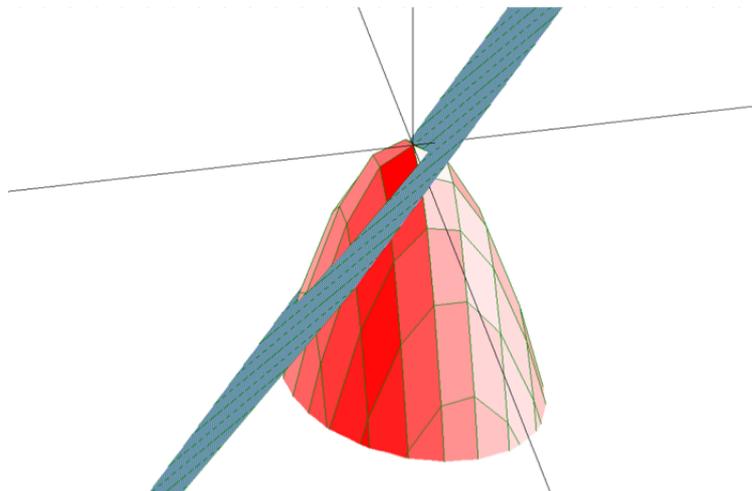


FIGURE 2. Rappresentazione delle funzioni $z = -2x^2 + xy - 3y^2$ e $z = 2x + 3y - 1$.

In molti casi è conveniente rappresentare una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ tramite le cosiddette curve di livello. Nella figura di seguito è mostrata la funzione $z = -2x^2 + xy - 3y^2$ e le funzioni $z = -2$, $z = -4$ e $z = -6$, ovvero le equazioni di tre piani paralleli al piano coordinato XY (di equazione $z = 0$).

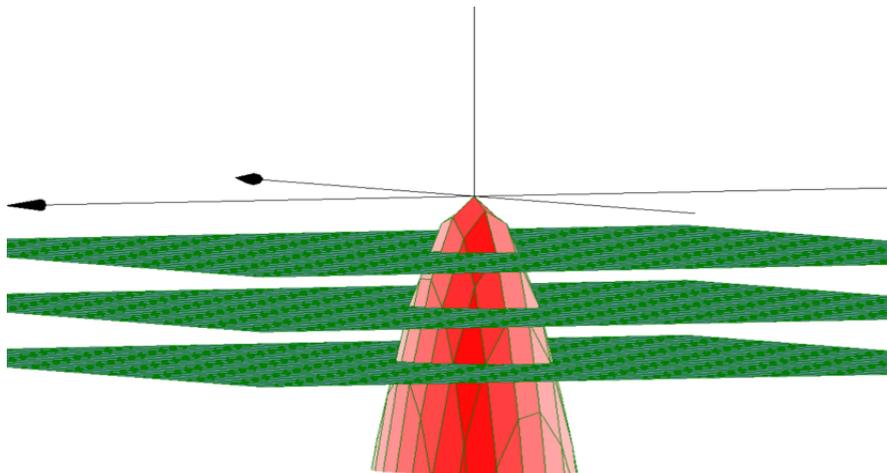


FIGURE 3. Rappresentazione delle funzioni $z = -2x^2 + xy - 3y^2$, $z = -2$, $z = -4$ e $z = -6$.

Le tre sezioni individuate in figura descrivono il luogo dei punti (x, y) tale che $-2x^2 + xy - 3y^2 = k$ con $k = -2, -4, -6$. Formalmente, fissato un certo k , il luogo dei punti (x, y) individuato si scrive:

$$c(k) = \{(x, y) : f(x, y) = k\}.$$

Queste coppie di valori (x, y) , che non sono altro che l'insieme delle soluzioni dell'equazione $f(x, y) = k$, descrivono delle curve sul piano XY . Ciascuna di esse è una *curva di livello*. Nelle figure di seguito sono riportate le curve di livello di $f(x, y) = -2x^2 + xy - 3y^2$, $g(x, y) = 3x - 2y - 1$ e $h(x, y) = 2xye^{x-y}$.

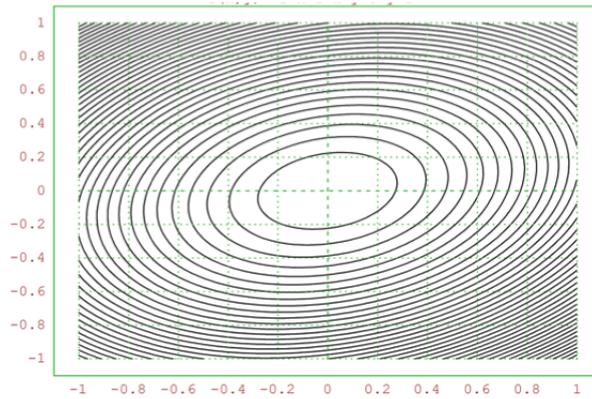


FIGURE 4. Curve di livello della funzione $f(x, y) = -2x^2 + xy - 3y^2$.

Si noti che maggiore è la vicinanza fra curve di livello, maggiore sarà la pendenza in quell'area. La funzione $f(x, y)$ (Fig.4) avrà una pendenza pari a zero nell'intorno della sua sommità.

Nel caso del piano $g(x, y)$ (Fig.5) le curve di livello sono equidistanti, poiché la pendenza di un piano è costante.

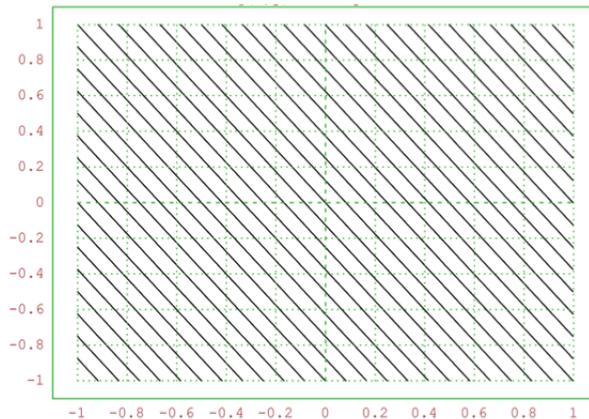


FIGURE 5. Curve di livello della funzione $g(x, y) = 3x - 2y - 1$.

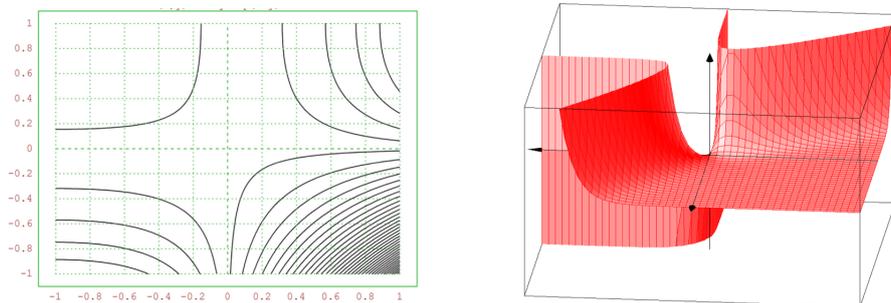


FIGURE 6. Pannello di sinistra: curve di livello della funzione $h(x, y) = 2xy e^{x-y}$. Pannello di destra: Rappresentazione della funzione $h(x, y) = 2xy e^{x-y}$.

Nel caso della funzione $h(x, y)$ (Fig.6), nell'intorno di $(x, y) = (0, 0)$ la funzione è praticamente piatta, ossia ha pendenza pari a zero.

Alcune classi di funzioni a più variabili hanno una valenza particolare e saranno analizzate più dettagliatamente. In particolare, sia $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ una funzione così definita:

$$T(x_1, x_2) = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - x_2 \\ -x_1 + 3x_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

La funzione $T(x_1, x_2)$ trasforma vettori di \mathbb{R}^2 in vettori di \mathbb{R}^3 e tale trasformazione avviene attraverso l'applicazione ad un vettore di \mathbb{R}^2 della matrice indicata. Una funzione come $T(x_1, x_2)$ (avvero lineare) è anche detta *trasformazione lineare*, o *mappa lineare*, o ancora *applicazione lineare* e gode di alcune interessanti proprietà.

1.1. Definizione di trasformazione lineare. Si definisce trasformazione lineare una relazione fra due spazi vettoriali, $T : U \rightarrow V$ che preserva le operazioni di somma di vettori e di moltiplicazione scalare-vettore, ovvero, $\forall \vec{x}, \vec{y} \in U$ e $\forall \alpha \in \mathbb{R}$ si ha:

- (a): $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$;
- (b): $T(\alpha \vec{x}) = \alpha T(\vec{x})$.

Nell'esempio sopraindicato, alla trasformazione lineare $T(\vec{x})$ è associata una matrice di dimensione 3×2 :

$$A_{3 \times 2} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ -1 & 3 \end{bmatrix}.$$

E' sempre così? In altri termini, ogni trasformazione lineare $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è in generale associata ad una matrice $A_{m \times n}$?

TEOREMA 1.1. *Sia $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una trasformazione lineare. Allora esiste una matrice $A_{m \times n}$ tale che $T(\vec{x}) = A \vec{x}$, $\forall \vec{x} \in \mathbb{R}^n$.*

PROOF. Come è noto, ogni vettore di \mathbb{R}^n può essere espresso come la combinazione lineare dei vettori della base canonica. Quindi, $\forall \vec{x} \in \mathbb{R}^n$ possiamo scrivere:

$$\vec{x} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \cdots + x_n \vec{e}_n,$$

dove x_1, \dots, x_n sono le componenti del vettore \vec{x} rispetto alla base canonica. Per ipotesi, la trasformazione $T(\vec{x})$ è una applicazione lineare che associa ad ogni vettore di \mathbb{R}^n un vettore di \mathbb{R}^m . Indichiamo con $\vec{a}_i = (a_{1i}, a_{2i}, \dots, a_{mi})$ il vettore di \mathbb{R}^m ottenuto trasformando l' i -esimo vettore \vec{e}_i della base canonica di \mathbb{R}^n : $\vec{a}_i = T(\vec{e}_i)$, $\forall i = 1, \dots, n$. Per le proprietà (a) e (b) di una trasformazione lineare abbiamo:

$$\begin{aligned} T(\vec{x}) &= x_1 T(\vec{e}_1) + x_2 T(\vec{e}_2) + \cdots + x_n T(\vec{e}_n) = \\ &= x_1 \vec{a}_1 + x_2 \vec{a}_2 + \cdots + x_n \vec{a}_n = \\ &= [\vec{a}_1 \quad \vec{a}_2 \quad \cdots \quad \vec{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \end{aligned}$$

□

Quindi le matrici non sono semplici tabelle di numeri, o un modo per rappresentare un sistema lineare. Bensì, ad ogni matrice A corrisponde una relazione tra due spazi vettoriali. In altri termini, le trasformazioni lineari sono trasformazioni che conservano le proprietà degli spazi e sottospazi vettoriali.

Un'altra classe di funzioni che studieremo da vicino sono le cosiddette *forme quadratiche* che sono funzioni da $\mathbb{R}^n \rightarrow \mathbb{R}$. Le forme quadratiche possono essere viste come generalizzazioni delle funzioni $f(x) = bx^2$. Nel caso di due variabili, una forma quadratica si esprime come:

$$Q(x_1, x_2) = a_{11} x_1^2 + a_{12} x_1 x_2 + a_{22} x_2^2,$$

dove a_{11}, a_{12} e a_{22} sono parametri che definiscono univocamente la forma quadratica. Si noti che $Q(x_1, x_2)$ è un polinomio di secondo grado in x_1 e x_2 . In generale, una forma quadratica $Q: \mathbb{R}^n \rightarrow \mathbb{R}$ è rappresentata da un polinomio omogeneo di grado 2. Più precisamente, si ha la seguente definizione di forma quadratica.

1.2. Definizione di Forma Quadratica. Una forma quadratica in \mathbb{R}^n è una funzione $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ definita nel modo seguente:

$$Q(\vec{x}) = Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j,$$

dove q_{ij} , $\forall i, j = 1, \dots, n$ sono parametri. Si consideri il coefficiente di $x_i x_j$. Dall'espressione precedente, risulta che questo coefficiente è pari a $q_{ij} + q_{ji}$ per $i \neq j$. Quindi, ponendo $a_{ij} = q_{ij} + q_{ji}$ per $i \neq j$, possiamo limitare la seconda sommatoria (quella su j) agli indici $j \geq i$:

$$Q(\vec{x}) = Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=i}^n a_{ij} x_i x_j,$$

In alcuni casi è conveniente esprimere una forma quadratica in forma matriciale. In particolare, una forma quadratica in \mathbb{R}^2 , $Q(x_1, x_2) = a_{11} x_1^2 + a_{12} x_1 x_2 + a_{22} x_2^2$, può essere rappresentata come

$$Q(x_1, x_2) = [x_1 \quad x_2] \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{x}^T A \vec{x}.$$

Come il seguente esempio mostra, la scelta della matrice A non è unica.

Esempio 1.1

Si scriva in forma matriciale la seguente forma quadratica:

$$Q(x_1, x_2) = 2x_1^2 - x_1 x_2 + 3x_2^2.$$

Usando la rappresentazione matriciale sopra indicata, si ha che $a_{11} = 2$, $a_{22} = 3$ e $a_{12} = -1$, $a_{21} = 0$. Si osservi che un risultato identico si ottiene utilizzando, al posto della matrice A , una matrice B di componenti $b_{11} = 1$, $b_{12} = 0$, $b_{21} = -1$ e $b_{22} = 3$. Si lascia allo studente di verificare che:

$$Q(x_1, x_2) = 2x_1^2 - x_1 x_2 + 3x_2^2 = \vec{x}^T A \vec{x} = \vec{x}^T B \vec{x},$$

dove A e B sono le matrici di cui abbiamo già indicato le componenti:

$$A = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix}.$$

In generale vi sono infinite matrici che rappresentano una forma quadratica. Basterà, infatti, assegnare agli elementi a_{ij} e a_{ji} dei valori tali che $a_{ij} + a_{ji}$ sia uguale al coefficiente di $x_i x_j$ nella forma quadratica. Nel caso specifico il coefficiente di $x_1 x_2$ è pari a -1 ; una possibile (altra) alternativa per rappresentare la forma quadratica in termini matriciali è dunque quella di porre $a_{12} = a_{21} = -1/2$.

In questo modo otteniamo la matrice $C = \begin{bmatrix} 2 & -1/2 \\ -1/2 & 3 \end{bmatrix}$. Questa è una matrice simmetrica che, come abbiamo imparato, gode di

proprietà particolari (è sempre diagonalizzabile ed ha autovalori reali). Questa dunque risulta una rappresentazione conveniente della forma quadratica. Vale il seguente teorema.

TEOREMA 1.2. Sia $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ una forma quadratica in \mathbb{R}^n :

$$Q(\vec{x}) = Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=i}^n a_{ij} x_i x_j.$$

Allora $Q(\vec{x})$ ammette una rappresentazione matriciale del tipo $Q(\vec{x}) = \vec{x}^T A \vec{x}$, dove A è una matrice simmetrica di ordine n così definita:

$$A = \begin{bmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{13}}{2} & \cdots & \frac{a_{1n}}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} & \cdots & \frac{a_{2n}}{2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{a_{1n}}{2} & \frac{a_{2n}}{2} & \frac{a_{3n}}{2} & \cdots & a_{nn} \end{bmatrix}.$$

Di converso, se A è una matrice simmetrica di ordine n , allora la funzione $Q(\vec{x}) = \vec{x}^T A \vec{x}$ è una forma quadratica. Come vedremo, le forme quadratiche hanno un ruolo fondamentale nella nostra analisi. Le forme quadratiche possono essere classificate sulla base del segno che assume $Q(\vec{x})$. Nelle figure che seguono sono raffigurate delle forme quadratiche che differiscono soltanto nei valori dei parametri a_{11} , a_{12} e a_{22} .

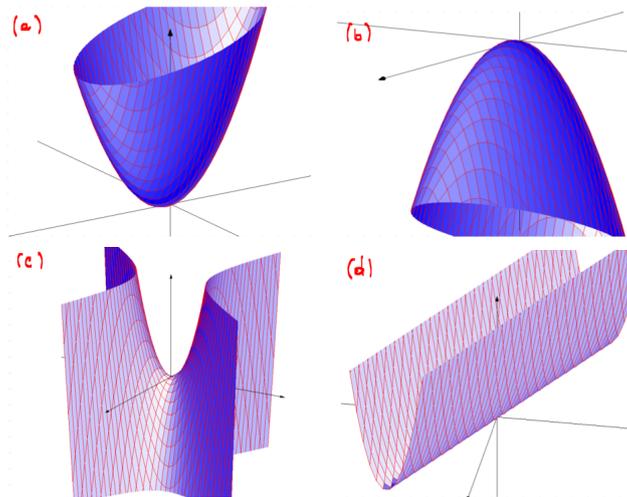


FIGURE 7. Rappresentazione 3D delle seguenti forme quadratiche. Pannello (a): $Q(x_1, x_2) = 1/2 x_1^2 - x_1 x_2 + x_2^2$; pannello (b): $Q(x_1, x_2) = -1/2 x_1^2 + x_1 x_2 - x_2^2$; pannello (c): $Q(x_1, x_2) = -1/2 x_1^2 - x_1 x_2 + x_2^2$; pannello (d): $Q(x_1, x_2) = 1/2 x_1^2 + 2 x_1 x_2 + 2 x_2^2$.

Si osservi che, nel caso (a), comunque presa la coppia (x_1, x_2) si ha che $Q(x_1, x_2) \geq 0$ e $Q(x_1, x_2) = 0$ SOLO per $(x_1, x_2) = (0, 0)$, ovvero per $\vec{x} = \vec{0}$. Nel caso (b) si verifica il contrario: $Q(x_1, x_2) \leq 0$ e $Q(x_1, x_2) = 0$ SOLO per $\vec{x} = \vec{0}$. Per la forma quadratica mostrata nel pannello (c) non è possibile stabilire un segno $Q(x_1, x_2)$ comunque preso (x_1, x_2) . Nel quarto ed ultimo caso, pannello (d), la forma quadratica $Q(x_1, x_2)$ è strettamente maggiore di zero in gran parte del dominio di \mathbb{R}^2 , mentre esiste un sottoinsieme di \mathbb{R}^2 , una retta, in cui $Q(x_1, x_2) = 0$.

1.3. Definizione di Forma Quadratica definita e semidefinita.

Sia $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ una forma quadratica. $Q(\vec{x})$ si dice *definita positiva* (*definita negativa*) se, $\forall \vec{x} \in \mathbb{R}^n$, tale che $\vec{x} \neq \vec{0}$ allora $Q(\vec{x}) > 0$ ($Q(\vec{x}) < 0$).

Una Forma quadratica $Q(\vec{x})$ si dice *semidefinita positiva* (*semidefinita negativa*) se, $\forall \vec{x} \in \mathbb{R}^n$, tale che $\vec{x} \neq \vec{0}$ allora $Q(\vec{x}) \geq 0$ ($Q(\vec{x}) \leq 0$).

Il segno di una forma quadratica dipende dalla matrice A . Vediamo alcuni teoremi che permettono di stabilire il segno di una forma quadratica.

TEOREMA 1.3. *Sia $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ una forma quadratica definita come $Q(\vec{x}) = \vec{x}^T A \vec{x}$, dove A è una matrice simmetrica di ordine n . La matrice A avrà n autovalori reali, non necessariamente tutti distinti. Si ha che:*

- (1) *La forma quadratica è definita positiva se e solo se $\lambda_i > 0$, $i = 1, \dots, n$;*
- (2) *La forma quadratica è definita negativa se e solo se $\lambda_i < 0$, $i = 1, \dots, n$;*
- (3) *La forma quadratica è semidefinita positiva (negativa) se e solo se $\lambda_i > 0$ ($\lambda_i < 0$) ed esiste almeno un indice h tale che $\lambda_h = 0$, $i = 1, \dots, n$;*
- (4) *In tutti gli altri casi la forma quadratica è indefinita.*

PROOF. Per ipotesi la matrice A è simmetrica. Quindi esiste una matrice ortogonale U (si ricorda che per una matrice ortogonale $U^{-1} = U^T$) che diagonalizza A :

$$U^T A U = D,$$

dove D è una matrice diagonale i cui elementi sono gli autovalori λ_i ($i = 1, \dots, n$) di A . Risolvendo la precedente equazione per A , ovvero premoltiplicando ambo i membri dell'equazione per U e postmoltiplicando per U^T , si ottiene:

$$A = U D U^T.$$

Sostituendo questa espressione di A nella definizione della forma quadratica $Q = \vec{x}^T A \vec{x}$ si ha che:

$$\vec{x}^T A \vec{x} = \vec{x}^T U D U^T \vec{x} = (\vec{x}^T U) D (U^T \vec{x}).$$

Se si pone $\vec{y} = U^T \vec{x}$ allora (ricordando che $(AB)^T = B^T A^T$ e che $(A^T)^T = A$) $\vec{y}^T = (U^T \vec{x})^T = \vec{x}^T U$. Quindi, rispetto a \vec{y} , otteniamo la seguente espressione della forma quadratica:

$$Q(\vec{y}) = \vec{y}^T D \vec{y} = \sum_{i=1}^n \lambda_i y_i^2$$

Come si può notare il segno di $Q(\vec{y})$ dipende soltanto dal segno degli autovalori λ_i , da cui la tesi. \square

E' possibile dimostrare che il teorema suvvisto vale anche per matrici non simmetriche. Un metodo alternativo per stabilire il carattere di una forma quadratica (definita o semidefinita) si basa sull'analisi del segno dei minori principali della matrice associata alla forma quadratica. Vale infatti il seguente teorema:

TEOREMA 1.4. *Sia $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ una forma quadratica definita come $Q(\vec{x}) = \vec{x}^T A \vec{x}$, dove A è una matrice simmetrica di ordine n . Allora:*

- (1) *La forma quadratica è definita positiva se e solo se tutti i minori principali dominanti di A sono strettamente maggiori di 0: $MPD_k > 0, \forall k = 1, \dots, n$.*
- (2) *La forma quadratica è definita negativa se e solo se i minori principali dominanti di ordine dispari sono strettamente negativi e quelli pari strettamente positivi: $(-1)^k MPD_k > 0, \forall k = 1, \dots, n$.*
- (3) *La forma quadratica è semidefinita positiva se e solo se tutti i minori principali di A sono non negativi: $MP_k \geq 0, \forall k = 1, \dots, n$.*
- (4) *La forma quadratica è semidefinita negativa se e solo se i minori principali di ordine dispari sono non-positivi e quelli pari non negativi: $(-1)^k MP_k \geq 0, \forall k = 1, \dots, n$.*
- (5) *La forma quadratica è indefinita in tutti gli altri casi.*

1

Lo studente verifichi che le forme quadratiche raffigurate nei pannelli (a), (b), (c) e (d) di Fig.7 sono, rispettivamente, definite positiva, negativa, indefinita e semidefinita positiva.

¹Si noti che, mentre nei casi 1) e 2) basta considerare i minori principali dominanti, nei casi 3) e 4) si devono considerare tutti i minori principali. Lo studente risponda alle seguenti domande: considerata una matrice 3×3 , quanti sono i minori principali dominanti? Quanti i minori principali?

Esercizi 1.1

Determinare il segno delle (forme quadratiche associate alle) seguenti matrici:

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & -3 & 2 \\ 0 & 2 & -4 \end{bmatrix} ; \quad \begin{bmatrix} 4 & -1 & 2 \\ -1 & 6 & 4 \\ 2 & 4 & 7 \end{bmatrix} ; \quad \begin{bmatrix} -1 & 1 & 0 \\ 1 & -3 & 2 \\ 0 & 2 & -2 \end{bmatrix}$$

$$\begin{bmatrix} -2 & 1 & -1 \\ 1 & -2 & -1 \\ -1 & -1 & -2 \end{bmatrix} ; \quad \begin{bmatrix} 2 & 1 & -1 \\ 1 & 4 & -2 \\ -1 & -2 & 4 \end{bmatrix} ; \quad \begin{bmatrix} 2 & -1 & 3 \\ -1 & 5 & 3 \\ 3 & 3 & 9 \end{bmatrix}$$