# FORME QUADRATICHE E DIAGONALIZZAZIONE

In queste note si vuole affrontare, in modo semplificato, lo studio del segno delle forme quadratiche.

In $\mathbb{R}^n$ una forma quadratica è semplicemente un polinomio omogeneo di secondo grado $H(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$.

Una prima osservazione da svolgere, elementare ai limiti della banalità ma vitale per quanto seguirà, è che i coefficienti $a_{ij}$ possono essere sostituiti con altri verificanti $a_{ij} = a_{ji}$, $\forall i,j = 1..n$.

Infatti, $H\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1^2 + x_1 x_2 + x_2^2 = 2x_1^2 + \frac{1}{2} x_1 x_2 + \frac{1}{2} x_2 x_1 + x_2^2$

e, scritta in quest'ultima forma, $a_{11} = 2$, $a_{12} = a_{21} = \frac{1}{2}$, $a_{22} = 1$. In generale, basta porre $\tilde{a}_{ij} = \frac{1}{2}(a_{ij} + a_{ji})$ per ottenere dalla matrice originale $a_{ij}$, anche non simmetrica, un'altra simmetrica $\tilde{a}_{ij}$ che definisce la stessa funzione, perché $x_i x_j = x_j x_i$:

$$H(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \sum_{i,j=1}^{n} \frac{1}{2}(a_{ij} + a_{ji}) x_i x_j = \sum_{i,j=1}^{n} \tilde{a}_{ij} x_i x_j$$

Lo studio del segno di H sarebbe enormemente semplificato se la matrice $a_{ij}$ fosse diagonale. Se infatti fosse $a_{ii} = \lambda_i$ e $a_{ij} = 0$ se $i \neq j$, allora per H si avrebbe

$$H(x) = \sum_{i=1}^{n} \lambda_i x_i^2$$

Poiché $x_i^2 \geq 0$ $\forall i = 1..n$ ed inoltre, per ogni vettore $e_i$ della base canonica, vale $H(e_i) = \lambda_i$ ne segue subito il

**TEOREMA**: Sia H una forma quadratica definita da una matrice diagonale $a_{ij} = \begin{cases} \lambda_i & \text{se } j = i \\ 0 & \text{se } j \neq i \end{cases}$.

Allora, detta $e_1, e_2, \dots e_n$ la base canonica, si ha:

1) Se $\lambda_i > 0$ $\forall i = 1..n$, $H(x) \geq 0$ e si annulla se e solo se $x_1 = x_2 = \dots = x_n = 0$.

   In tal caso la forma si dirà DEFINITA POSITIVA.

2) Se $\lambda_i < 0$ $\forall i = 1..n$, $H(x) \leq 0$ e si annulla se e solo se $x_1 = x_2 = \dots = x_n = 0$.

   In tal caso la forma si dirà DEFINITA NEGATIVA

3) Se $\lambda_i = 0$ per $i = i_1, i_2, \dots, i_h$ e $\lambda_i > 0$ per gli altri, allora $H(x) \geq 0$ $\forall x \in \mathbb{R}^n$. Inoltre, $H(x) = 0$ $\forall x \in \langle e_{i_1}, e_{i_2}, \dots, e_{i_h} \rangle$. In tal caso, la forma si dirà SEMIDEFINITA POSITIVA

4) Se $\lambda_i = 0$ per $i = i_1, i_2, \dots, i_h$ e $\lambda_i < 0$ per gli altri, allora $H(x) \leq 0$ $\forall x \in \mathbb{R}^n$. Inoltre $H(x) = 0$ $\forall x \in \langle e_{i_1}, e_{i_2}, \dots, e_{i_h} \rangle$. In tal caso, la forma si dirà SEMIDEFINITA NEGATIVA

5) Se $\exists i, j : \lambda_i > 0$ e $\lambda_j < 0$ allora $H$ cambia segno, e verrà detta INDEFINITA. $\boxed{/\!/}$

Lo studio del segno di una forma quadratica si ottiene cambiando coordinate, in modo da trasformare la forma originale in una diagonale, e studiando poi la forma diagonale in accordo col teorema precedente.

In vista di impiegare la teoria spettrale è utile un ulteriore passo: se si scrive l'espressione della forma $H$ associando (e cioè eseguendo per prima) la somma rispetto all'indice $j$ si ottiene, essendo $e_1 \dots e_n$ ortonormale

$$H(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} x_j = \left(\sum_{h=1}^{m} x_h e_h\right)\left(\sum_{i,j=1}^{n} a_{ij} x_j e_i\right) = x\, A(x)$$

ove

$$A(x) = \sum_{i,j=1}^{n} a_{ij} x_j e_i$$

e cioè l'operatore che ha come matrice associata rispetto alla base canonica (sia nel dominio, sia nell'immagine) la stessa matrice $a_{ij}$ che definisce la forma quadratica, che si è visto può sempre essere scelta simmetrica.

Poiché $A$ è reale e simmetrica, $\mathbb{R}^n$ ha una base ortonormale spettrale per $A$ e dunque, con un cambio di base dalla base canonica a quella spettrale, si ottiene

$$H(x) = \underbrace{\left(\sum_{h=1}^{n} x'_h e_h\right)}_{x} \underbrace{\left(\sum_{i,j=1}^{n} a'_{ij} x'_j e_i\right)}_{A(x)} = \sum_{i,h} a'_{ii} x'_i x'_h e_h e_i = \sum_{1}^{n} \lambda_i x_i'^2$$

ove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di $A$, $x'_1, \dots, x'_n$ le coordinate del vettore $x$ rispetto ai relativi autovettori della base spettrale e $a'_{ij}$ è la matrice associata alla base spettrale.

<u>TEOREMA</u>: <u>Ogni forma quadratica $\sum_{i,j=1}^{n} a_{ij} x_i x_j$ può essere ridotta alla forma diagonale $\sum_{i=1}^{n} \lambda_i x_i'^2$ effettuando il cambio di base dalla base canonica alla base spettrale dell'operatore $A(x) = \sum_{1}^{n} a_{ij} x_j e_i$; ogni coefficiente $\lambda_i$ coincide con l'autovalore relativo allo autovettore corrispondente alla coordinata $x'_i$.</u>

<u>La forma originale $\sum_{i,j} a_{ij} x_i x_j$ ha lo stesso segno di quella diagonale $\sum_{i=1}^{n} \lambda_i x_i'^2$ e verrà dunque detta</u>

- DEFINITA POSITIVA se ha tutti gli autovalori (i coefficienti) strettamente positivi. Si annulla solo nell'origine.

- DEFINITA NEGATIVA se ha tutti gli autovalori strettamente negativi. Si annulla solo nell'origine.

- SEMIDEFINITA POSITIVA se ha l'autovalore 0, e quelli non nulli sono positivi. Si annulla solo su tutto l'autospazio dell'autovalore 0.

- SEMIDEFINITA NEGATIVA se ha l'autovalore 0, e quelli non nulli sono negativi. Si annulla solo su tutto l'autospazio dell'autovalore 0.

- INDEFINITA, se possiede due autovalori (non nulli) discordi. Sarà negativa, positiva, o nulla sugli autospazi degli autovalori rispettivamente negativi, positivi e dello zero. Si annulla fuori dell'origine. (Es. $x^2 - y^2$ si annulla per $y = \pm x$).

NOTA: Se $\alpha(x,y)$ è un prodotto scalare su $\mathbb{R}^n$, dagli assiomi segue che $\alpha(x,x) \geq 0$ e $\alpha(x,x) = 0$ se e solo se $x=0$, e che

$$a(x,x) = a\left(\sum_1^n x_i e_i, \sum_1^n x_j e_j\right) = \sum_{i,j=1}^n x_i x_j \, a(e_i, e_j) = \sum_{i,j=1}^n x_i x_j a_{ij}.$$

Dunque $\alpha(x,x)$ è una forma quadratica definita positiva.

E' facile ricordare gli assiomi del prodotto scalare (REALE) pensando ad esso come ad una forma bilineare, simmetrica, e definita positiva. Nel caso complesso, è conveniente rinunciare (parzialmente) alle prime due proprietà pur di conservare la terza: un prodotto scalare COMPLESSO (detto anche hermitiano) è una forma sesquilineare, emisimmetrica e definita positiva: basta ricordare che è lineare rispetto al primo argomento, e scambiando l'ordine dei fattori il prodotto (invece di non cambiare, come abbiamo nel cervello sin da piccoli) viene CONIUGATO. $\boxed{/\!/}$

Lo studio del segno di una forma quadratica è dunque completo non appena si riesce a stabilire il <u>SEGNO</u> degli autovalori della sua matrice (simmetrica) e <u>**NON**</u> il loro <u>VALORE</u>. Due tecniche efficienti per effettuare tale studio sono la regola dei segni di Cartesio, che richiede solo

il calcolo del polinomio caratteristico e non la determinazione delle sue radici, e l'algoritmo di Gauss usato in congiunzione col teorema di Sylvester, che non richiede neppure di calcolare il polinomio caratteristico.
Un'altra via percorribile se la dimensione è molto bassa è quella di calcolare gli invarianti. In dimensione due, ad esempio, si ha (nel caso simmetrico, $a_{12}=a_{21}$)

$$p(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}^2 =$$

$$= \lambda^2 - (a_{11}+a_{22})\lambda + a_{11}a_{22} - a_{12}^2 =$$

$$= \lambda^2 - (\operatorname{tr} A)\lambda + \det A$$

ove $$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Poiché il polinomio caratteristico è invariante per cambio di base, dal principio di identità dei polinomi tali sono anche i suoi coefficienti $\operatorname{tr} A$ e $\det A$. Scegliendo la base spettrale ne segue subito che, detti $\lambda_1$ e $\lambda_2$ gli autovalori, la matrice associata è $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, e dunque la traccia vale $\lambda_1+\lambda_2$ mentre il determinante vale $\lambda_1\lambda_2$. Allora:

1) Se $\det A < 0$ gli autovalori sono discordi e la forma è <u>indefinita</u>.

2) Se $\det A = 0$ Almeno uno degli autovalori è nullo e l'altro ha lo stesso segno di $\operatorname{tr} A$.

La forma è <u>semidefinita positiva</u> se $\det A = 0$ e $\operatorname{tr} A > 0$, è <u>semidefinita negativa</u> se $\det A = 0$ e $\operatorname{tr} A < 0$. <u>E'</u> <u>identicamente nulla</u> se $\det A = 0$ e $\operatorname{tr} A = 0$.

3) Se $\det A > 0$ gli autovalori sono non nulli e concordi con $\operatorname{tr} A$.

La forma è <u>definita positiva</u> se $\det A > 0$ e $\operatorname{tr} A > 0$, <u>definita negativa</u> se $\det A > 0$ e $\operatorname{tr} A < 0$.

Tutto dipende esclusivamente dai segni di $\operatorname{tr} A$ e $\det A$, che possono essere calcolati direttamente su $A$, senza determinare gli autovalori, sfruttando l'invariante: $\operatorname{tr} A = a_{11} + a_{12}$ e $\det A = a_{11}a_{22} - a_{12}^2$.
La teoria degli invarianti è ancora ragionevolmente efficiente per $n=3$ dove, per classificare le quadriche, occorrono solo $\operatorname{tr} A$ (l'invariante lineare), $\det A$ (l'invariante cubico), e l'invariante quadratico $J = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$.

Nel caso generale di un polinomio caratteristico di grado $n$

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = (-\lambda)^n + a_1 (-\lambda)^{n-1} + \dots + a_{n-1}(-\lambda) + a_0$$

i coefficienti (e così gli invarianti) hanno delle espressioni intollerabilmente complesse: infatti $a_k$ è la somma di tutti i determinanti che si possono formare dalla matrice originale soppri= mendo, in tutti i modi possibili, $n-k$ righe e le colonne di indici uguali: quelle nelle quali, sulla diagonale $a_{ii}-\lambda$, si sceglie il fattore $-\lambda$ invece del termine della matrice originale $a_{ii}$ nei prodotti corrispondenti.

Si tratta dunque di calcolare $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$ determinanti di matrici $k \times k$. Anche ammettendo di saper calcolare in un battito di ciglia i determinanti $k \times k$ (il che è pura illusione!) calcolare il quinto invariante di una forma in $\mathbb{R}^{10}$ richiederebbe

$$\binom{10}{5} = \frac{10\cdot 9\cdot 8\cdot 7\cdot 6\cdot}{5\cdot 4\cdot 3\cdot 2\cdot 1} = \frac{10}{5}\cdot\frac{6\cdot 8}{4\cdot 3\cdot 2\cdot 1}\cdot 9\cdot 7 = 2\cdot 2\cdot 9\cdot 7 = 36\cdot 7 = 252$$

252 "battiti di ciglia": ... TROPPI!

Il "metodo degli invarianti" è dunque limitato a $n=2$ o $n=3$, dove è ragionevolmente efficiente; in tali casi nulla vieta di impiegarlo, in luogo degli altri metodi più generali (Segni di Cartesio o Algoritmo di Gauss-Sylvester) reperibili in altre note.

**NOTA**: si può dare una definizione astratta di forma quadratica, come restrizione alla "diagonale" $v=u$ di una funzione bilineare $\alpha(u,v)$. Gli aspetti elementari della teoria sono stati già presentati in un altro contributo.

**NOTA**: Lo studio del segno delle forme quadratiche è utilissimo per la risoluzione di problemi di estremo (max o min) per le funzioni di più variabili.

In effetti, poiché per una forma quadratica $H$ vale $H(0)=0$, ne segue subito che

0 è di minimo per $H$ se $H$ è definita o semidefinita positiva

0 è di massimo per $H$ se $H$ è definita o semidefinita negativa

0 non è né di massimo né di minimo se $H$ è indefinita.

Così come accade per le funzioni di una variabile, le condizioni sufficienti di estremo si ottengono studiando il segno di $f(x)-f(x_0)$ attraverso l'impiego della formula di Taylor.

I dettagli sono reperibili sui testi di Analisi Matematica, e presentano differenze con le condizioni precedenti per l'effetto dei termini di ordine superiore sulle direzioni nelle quali il complesso dei termini di secondo grado (cioè la forma quadratica hessiana) si annulla: è proprio l'analogo di quanto si verifica in una variabile quando $f''(x_0)=0$: il caso più ostico!

La simmetria della matrice è essenziale per poter applicare il teorema spettrale e diagonalizzare la forma quadratica.
In pratica, qualunque studio del segno ha come passo preliminare la scrittura della forma quadratica in forma simmetrica.

**ESEMPIO**: Studiare il segno di $x^2+6xy-4xz+z^2$.
La matrice originale è $\begin{pmatrix} 1 & 6 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Per renderla simmetrica basta "ridistribuire in parti uguali" la somma dei coefficienti simmetrici rispetto alla diagonale, ottenendo $\begin{pmatrix} 1 & 3 & -2 \\ 3 & 0 & 0 \\ -2 & 0 & 1 \end{pmatrix}$, che in realtà è la matrice di $x^2+3xy+3yx-2xz-2zx+z^2$, coincidente con l'originale per ogni $x,y,z$. Determiniamo ora il **segno** degli autovalori di quest'ultima matrice. Le due alternative sono

1) Segni di Cartesio

$$\begin{vmatrix} 1-\lambda & 3 & -2 \\ 3 & -\lambda & 0 \\ -2 & 0 & 1-\lambda \end{vmatrix} = -\lambda(\lambda^2-2\lambda+1) - (-4\lambda+9-9\lambda) =$$

$$= -\lambda^3+2\lambda^2+12\lambda-9$$

I coefficienti sono tutti non nulli, ma non sono né concordi, né a segni alterni: la forma è INDEFINITA!

2) Algoritmo di Gauss (col teorema di Sylvester)

$$\begin{pmatrix} 1 & 3 & -2 \\ 3 & 0 & 0 \\ -2 & 0 & 1 \end{pmatrix} \xrightarrow{II - 3I} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -9 & 6 \\ -2 & 0 & 1 \end{pmatrix} \xrightarrow[\text{riga. Le altre sono inalterate}]{\text{La II colonna diventa uguale alla II}} \begin{pmatrix} 1 & 0 & -2 \\ 0 & -9 & 6 \\ -2 & 6 & 1 \end{pmatrix}$$

(L'ultima matrice è di nuovo simmetrica, ed è stata ottenuta applicando alle colonne la stessa operazione applicata alle righe)

$$\xrightarrow{III + 2I} \begin{pmatrix} 1 & 0 & -2 \\ 0 & -9 & 6 \\ 0 & 6 & -3 \end{pmatrix} \xrightarrow[\text{sulle colonne: III colonna = III riga}]{\text{Stessa operazione}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -9 & 6 \\ 0 & 6 & -3 \end{pmatrix} \xrightarrow{III + \frac{2}{3} II}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & -9 & 6 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[\text{uguale alla III riga}]{\text{III colonna}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$\boxed{\text{Due termini diagonali positivi, uno negativo, nessuno nullo}}$

In conclusione: non ci sono autovalori nulli (0 non appare sulla diagonale) ed i tre autovalori (reali per la simmetria) sono due positivi ed uno negativo: la forma è indefinita.

Il calcolo del polinomio caratteristico (che è un determinante $n \times n$, dipendente da un parametro $\lambda$) diventa gravosissimo al crescere della dimensione $n$, rendendo l'algoritmo di Gauss più vantaggioso rispetto all'impiego della regola dei segni di Cartesio. Alcune altre "astuzie" per l'uso dell'algoritmo di Gauss anche in questo contesto, sono reperibili nel contributo ad esso dedicato.

La "bizzarria" di applicare alle colonne le stesse operazioni appena effettuate sulle righe viene dal teorema di Sylvester: il numero (e non il valore) degli autovalori nulli, positivi e negativi non cambia se si moltiplica una matrice simmetrica a destra per una matrice e a sinistra per la sua trasposta: in effetti, la trasposta della matrice che somma ad una riga un multiplo di un'altra (ad esempio), se moltiplicata a sinistra, esegue la stessa operazione sulle colonne se moltiplicata a destra. Ridotta a forma diagonale, il numero di elementi nulli, positivi e negativi è uguale alla molteplicità dell'autovalore nullo e al numero degli autovalori positivi o negativi, rispettivamente.