Laboratorio di Matematica, 28.10.2003

PROBLEMA *Studiare il segno del polinomio omogeneo di II grado*

$$f(x, y, z) = 2x^2 - 2xy - 2xz + y^2 + z^2$$

nelle variabili reali x, y, z.

I dati che caratterizzano f sono i coefficienti delle variabili, che possono essere rappresentati dalla seguente tabella

	x	y	z
x	2	−1	−1
y	−1	1	0
z	−1	0	1

dove nella prima riga compaiono il coefficiente di x^2, la meta' del coefficiente di xy, la meta' del coefficiente di xz, mentre nella seconda riga compaiono la meta' del coefficiente di $yx = xy$, il coefficiente di y^2, la meta' del coefficiente di yz, ed infine nella terza ... Si osservi che

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2x^2 - 2xy - 2xz + y^2 + z^2.$$

Ora, indicata con A la matrice ottenuta dai coefficienti di f ed indicato con v il vettore colonna delle variabili x, y, z, si puo' scrivere sinteticamente

$$f(v) = v^T Av.$$

Il polinomio f viene dunque riguardato come una funzione $f : R^3 \to R$, di un'unica variabile vettoriale $v = (x, y, z)$. In quest'ottica, si preferisce parlare di *forma quadratica su* R^3 al posto di "polinomio omogeneo di II grado nelle variabili reali x, y, z."

In generale, i dati che caratterizzano un polinomio φ omogeneo di II grado in n variabili reali $x_1, x_2, \dots, x_n$ sono i coefficienti delle variabili, che possono essere rappresentati dalla matrice

$$A = [a_{ij}], \qquad a_{ij} = \begin{cases} coeff.\ di\ x_i^2 & per\ i = j \\ \frac{1}{2} coeff.\ di\ x_i x_j & per\ i \neq j \end{cases}.$$

Si osservi che A e' una matrice simmetrica:

$$a_{ij} = a_{ji}, \quad per\ ogni\ i, j = 1, 2, \dots, n; \qquad in\ sintesi \quad A^T = A.$$

Ora, indicato con v il vettore colonna delle variabili $x_1, x_2, \dots, x_n$ si puo' scrivere sinteticamente

$$\varphi(v) = v^T Av.$$

Il polinomio φ viene dunque riguardato come una funzione $\varphi : R^n \to R$, di un'unica variabile vettoriale $v = (x_1, x_2, \dots, x_n)$. In quest'ottica, si preferisce parlare di *forma quadratica su* R^n al posto di "polinomio omogeneo di II grado nelle variabili reali $x_1, x_2, \dots, x_n$."

Il senso in cui intendiamo "studiare il segno" e' espresso nei termini della seguente definizione.

Def *Una forma quadratica $\varphi : R^n \to R$ su R^n si dice*

$$\begin{array}{rccll} definita & positiva & se & \varphi(v) > 0 & per\ ogni\ v \in R^n,\ v \neq \underline{0} \\ semidefinita & positiva & se & \varphi(v) \geq 0 & per\ ogni\ v \in R^n \\ definita & negativa & se & \varphi(v) < 0 & per\ ogni\ v \in R^n,\ v \neq \underline{0} \\ semidefinita & negativa & se & \varphi(v) \leq 0 & per\ ogni\ v \in R^n \end{array}.$$

Infine, φ di dice indefinita se non e' ne' semidefinita positiva ne' semidefinita negativa.

Soluzione del problema. Un test significativo per determinare la categoria cui appartiene una data forma quadratica si ottiene valutandola sugli autovettori della matrice corrispondente.
Nel nostro esempio

$$f(x, y, z) = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = v^T A v = f(v),$$

gli autovalori e gli autovettori della matrice A sono descritti dalla seguente tabella

autovalore	*base dell'autospazio corrispondente*
0	$v_1 = (1, 1, 1)$
1	$v_2 = (0, 1, -1)$
3	$v_3 = (-2, 1, 1)$

e si ha

$$\begin{array}{rrrrl} f(v_1) = & v_1^T A v_1 = & v_1^T 0 v_1 = & 0 & \\ f(v_2) = & v_2^T A v_2 = & v_2^T 1 v_2 = & \|v_2\|^2 & > 0 \\ f(v_3) = & v_3^T A v_3 = & v_3^T 3 v_3 = & 3\|v_3\|^2 & > 0 \end{array}$$

dunque la forma quadratica f si comporta su questi vettori come una forma quadratica semidefinita positiva. E sugli altri vettori?

Osserviamo che gli autovettori v_1, v_2, v_3 costituiscono una base ortogonale di R^3, che, dividendo ciascuno di essi per la sua norma, porge una base ortonormale w_1, w_2, w_3 di R^3 :

autovalore	*base normalizzata dell'autospazio corrispondente*
0	$w_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$
1	$w_2 = \frac{1}{\sqrt{2}}(0, 1, -1)$
3	$w_3 = \frac{1}{\sqrt{6}}(-2, 1, 1)$

Ora, ogni vettore v di R^3 si puo' scrivere (in uno ed un solo modo) come combinazione lineare

$$v = a\ w_1 + b\ w_2 + c\ w_3$$

dei vettori w_1, w_2, w_3 e si ha

$$\begin{aligned}
f(v) = &\ v^T\ A\ v \\
= &\ (a\ w_1 + b\ w_2 + c\ w_3)^T\ A\ (a\ w_1 + b\ w_2 + c\ w_3) \\
= &\ (a\ w_1^T + b\ w_2^T + c\ w_3^T)\ A\ (a\ w_1 + b\ w_2 + c\ w_3) \\
= &\ (a\ w_1^T + b\ w_2^T + c\ w_3^T)\ (a\ A\ w_1 + b\ A\ w_2 + c\ A\ w_3) \\
= &\ (a\ w_1^T + b\ w_2^T + c\ w_3^T)\ (b\ w_2 + 3\ c\ w_3) \\
= &\ ab\ w_1^T w_2 + 3\ ac\ w_1^T w_3 + \\
&\ b^2\ w_2^T w_2 + 3\ bc\ w_2^T w_3 + \\
&\ cb\ w_3^T w_2 + 3\ c^2\ w_3^T w_3
\end{aligned}$$

Tenendo presente che i vettori w_1, w_2, w_3 sono a due a due ortogonali, questa espressione si riduce a

$$b^2\ w_2^T w_2 + 3\ c^2\ w_3^T w_3,$$

tenendo poi presente che ciascun w_i e' un versore, si ottiene infine

$$b^2 + 3\ c^2 \geq 0 \quad \textit{per ogni}\ a, b, c \in R.$$

Possiamo dunque affermare che la forma quadratica f e' semidefinita positiva, ma non e' definita positiva, in quanto $f(w_1) = 0$.

Matrici ortogonali Rappresentando i vettori w_1, w_2, w_3 come colonne ed affiancando queste colonne si ottiene una matrice

$$P = [w_1\ w_2\ w_3].$$

Poiche' questi vettori formano una base ortonormale di R^3, si ha che la tabella dei loro prodotti scalari e' la matrice unita'

$$P^T P = I_3.$$

Ora, questa uguaglianza puo' essere letta dicendo che la matrice P^T e' un'inversa sinistra della matrice P; essendo queste matrici quadrate, si ha che P e' invertibile e P^T e' l'inversa, sinistra e destra, della matrice P. Dunque si ha pure

$$PP^T = I_3,$$

che significa che la tabella dei prodotti scalari delle righe di P e' la matrice unita', cioe' anche le righe di P formano una base ortonormale di R^3.

In generale, si dice che una matrice P quadrata di ordine n e' ortogonale se

$$PP^T = I_n = P^T P,$$

oppure, equivalentemente, se le colonne di P formano una base ortonormale di R^n, cosi' come le righe di P formano una base ortonormale di R^n.

L'approccio seguito nella soluzione di questo particolare problema sostanzialmente e' consistito nel ricondurre la valutazione del segno della forma quadratica f su un qualsiasi vettore alla valutazione del segno di f sugli autovettori della matrice simmetrica A associata ad f; il punto focale in questa riconduzione e' stata la possibilita' di costruire una base ortogonale, in realta' ortonormale, dello spazio R^3 costituita da autovettori di A. Questa possibilita' e' garantita sempre dal seguente importante Teorema, che non dimostriamo.

Teorema Spettrale *Sia A una matrice quadrata di ordine n, ad elementi reali. Se A e' simmetrica, allora:*

- *tutti gli autovalori di A sono reali;*
- *autovettori in autospazi distinti sono ortogonali;*
- *esiste una base ortonormale di R^n costituita da autovettori di A.*

Data una matrice A quadrata di ordine n, reale simmetrica, sia $w_1, w_2, \ldots, w_n$, una base ortonormale di R^n costituita da autovettori di A, con autovalori associati $\lambda_1, \lambda_2, \ldots, \lambda_n$, rispettivamente. Posto

$$P = [w_1\ w_2\ \ldots w_n], \qquad L = diag[\lambda_1, \lambda_2, \ldots, \lambda_n],$$

si hanno le relazioni

$$AP = PL, \qquad P^T P = I_n = PP^T,$$

dalle quali possiamo ricavare A in funzione di L e P, cosi' come possiamo ricavare L in funzione di A e P :

$$A = PLP^T \qquad P^T AP = L.$$

Dunque il Teorema spettrale si puo' esprimere anche nella forma

Sia A una matrice quadrata di ordine n, ad elementi reali. Se A e' simmetrica, allora esistono una matrice diagonale L ed una matrice ortogonale P tali che

$$A = PLP^T \qquad P^T AP = L.$$

Soluzione del problema generale. Sia ora $\varphi : R^n \to R$ una forma quadratica su R^n, e sia A la matrice dei suoi coefficienti:

$$\varphi(v) = v^T\ A\ v \qquad \textit{per ogni } v \in R^n.$$

Per il teorema spettrale, esistono una matrice diagonale L ed una matrice ortogonale P tali che

$$P^T AP = L.$$

Le colonne della matrice P sono costituite da autovettori $w_1, w_2, \ldots, w_n$, di A, mentre sulla diagonale della matrice L compaiono i corrispondenti autovalori $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Ora, il generico vettore $(x_1, x_2, \ldots, x_n) = v$ di R^n si puo' scrivere (in uno ed un solo modo) come combinazione lineare

$$v = w_1\overline{x}_1 + w_2\overline{x}_2 + \ldots + w_n\overline{x}_n$$

dei vettori $w_1, w_2, \ldots, w_n$. Sinteticamente, si puo' scrivere

$$v = P\overline{v},$$

dove $\overline{v} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_n)$ e' il vettore colonna delle coordinate del vettore v rispetto alla base dei vettori w_i.

La valutazione della forma quadratica φ sul vettore v puo' essere calcolata come

$$\begin{aligned}
\varphi(v) = & \quad v^T \ A \ v \\
= & \quad (P\overline{v})^T \ A \ (P\overline{v}) \\
= & \quad \overline{v}^T P^T \ A \ P\overline{v} \\
= & \quad \overline{v}^T L \overline{v} \qquad = \lambda_1 \overline{v}_1^2 + \lambda_2 \overline{v}_2^2 + \ldots + \lambda_n \overline{v}_n^2.
\end{aligned}$$

Allora e' facile decidere in quale categoria la forma quadratica φ rientra:

$$\begin{array}{rllll}
definita & positiva & sse & tutti\ gli\ autovalori\ di\ A\ sono & positivi \\
semidefinita & positiva & sse & tutti\ gli\ autovalori\ di\ A\ sono & positivi\ o\ nulli \\
\ldots & & & &
\end{array}.$$

Infine, φ risulta essere indefinita sse ci sono due autovalori di A aventi segno opposto.