Complementi di Algebra e Fondamenti di Geometria

Capitolo 5

Forme quadratiche in $\mathbb{R}^n$

M. Ciampa

Ingegneria Elettrica, a.a. 2009/2010

Capitolo 5

Forme quadratiche in $\mathbb{R}^n$

In questo capitolo si definisce la nozione di *forma quadratica* associata ad una matrice simmetrica di $\mathbb{R}^{n\times n}$ e si definisce il *problema della classificazione* di una forma quadratica. Si introduce la nozione di *matrici congruenti* e la si utilizza per descrivere come una forma quadratica possa essere classificata in base al segno degli autovalori della matrice ad essa associata. Infine, si descrive una procedura elementare che, per ciascuna matrice di $\mathbb{R}^{n\times n}$, consente di classificare la forma quadratica ad essa associata.

Il problema della classificazione di una forma quadratica trova applicazione nello *studio dei punti stazionari* di funzioni regolari da $\mathbb{R}^n$ in $\mathbb{R}$. La forma quadratica che interessa classificare, in tal caso, è quella associata alla *matrice hessiana*.

Si ricordi che se a, b sono vettori di $\mathbb{R}^n$, si indica con $a \bullet b = b^{\mathsf{T}} a$ il loro *prodotto scalare* (vedere il Capitolo 2).

5.1 Forme quadratiche: il problema della classificazione

In questa sezione si introduce la nozione di *forma quadratica associata ad una matrice simmetrica* di $\mathbb{R}^{n\times n}$ e si definisce il *problema della classificazione di una forma quadratica.*

Sia $A \in \mathbb{R}^{n\times n}$ una matrice *simmetrica.* La funzione $F : \mathbb{R}^n \to \mathbb{R}$ definita da

$$F(x) = Ax \bullet x$$

si chiama *forma quadratica associata ad A.*

5.1 Esempio

Sia:

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 4 & 0 \\ 2 & 0 & -1 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

Per ogni $x \in \mathbb{R}^3$ si ha:

$$Ax = \begin{bmatrix} x_1 + x_2 + 2x_3 \\ x_1 + 4x_2 \\ 2x_1 - x_3 \end{bmatrix}$$

e:

$$Ax \bullet x = x_1(x_1 + x_2 + 2x_3) + x_2(x_1 + 4x_2) + x_3(2x_1 - x_3)$$

e quindi:

$$F(x) = x_1^2 + 4x_2^2 - x_3^2 + 2x_1x_2 + 4x_1x_3$$

Ripetendo il procedimento, si constata che se $A \in \mathbb{R}^{3\times 3}$ è simmetrica di elementi a_{ij}, la forma quadratica associata ad A è:

$$F(x) = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3$$

Se *per ogni vettore non nullo* $x \in \mathbb{R}^n$ si ha:

$$F(x) \begin{cases} > 0 \\ \geqslant 0 \\ \leqslant 0 \\ < 0 \end{cases}, \text{ la forma quadratica } F \text{ si dice } \begin{cases} \textit{definita positiva} \\ \textit{semidefinita positiva} \\ \textit{semidefinita negativa} \\ \textit{definita negativa} \end{cases}$$

altrimenti (ovvero: se esistono vettori non nulli $x, y \in \mathbb{R}^n$ tali che $F(x) > 0$ e $F(y) < 0$) si dice *indefinita*. Analoga terminologia si adotta per la matrice A.

5.2 Esempio

La forma quadratica associata alla matrice $I \in \mathbb{R}^{n\times n}$ è *definita positiva*. Infatti, per ogni vettore non nullo $x \in \mathbb{R}^n$ si ha:

$$F(x) = Ix \bullet x = x \bullet x > 0$$

La forma quadratica associata alla matrice $-I \in \mathbb{R}^{n\times n}$ è *definita negativa*. Infatti, per ogni vettore non nullo $x \in \mathbb{R}^n$ si ha:

$$F(x) = (-I)x \bullet x = -(x \bullet x) < 0$$

La forma quadratica associata alla matrice nulla $0 \in \mathbb{R}^{n\times n}$ è sia *semidefinita positiva* che *semidefinita negativa*. Infatti, per ogni vettore non nullo $x \in \mathbb{R}^n$ si ha:

$$F(x) = 0x \bullet x = 0 \bullet x = 0$$

In ciascun caso, la matrice si classifica allo stesso modo della forma quadratica ad essa associata: I è definita positiva, $-I$ è definita negativa, 0 è sia semidefinita positiva che semidefinita negativa.

5.3 Esempio

Siano $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^{n\times n}$, ed F la forma quadratica associata ad A.
Per ogni $x \in \mathbb{R}^n$ si ha:

$$F(x) = Ax \bullet x = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2$$

Se $\lambda_1 > 0, \ldots, \lambda_n > 0$ allora per ogni $x \in \mathbb{R}^n$ tutti gli addendi di F hanno valore non negativo e, se x non è il vettore nullo, almeno un addendo ha valore positivo. Dunque F è definita positiva.

Viceversa, se F è definita positiva, ricordando che $e_1, \ldots, e_n$ è la base canonica di $\mathbb{R}^n$, si ha:

$$0 < F(e_1) = \lambda_1, \ldots, 0 < F(e_n) = \lambda_n$$

ovvero: $\lambda_1 > 0, \ldots, \lambda_n > 0$.

I due asserti si riassumono in:

$$F \textit{ è definita positiva se e solo se } \lambda_1 > 0, \ldots, \lambda_n > 0$$

Con ragionamenti analoghi si prova che:

$$F \textit{ è semidefinita positiva se e solo se } \lambda_1 \geqslant 0, \ldots, \lambda_n \geqslant 0$$

Allo stesso modo si classifica la matrice A.

5.4 Problema

Siano $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^{n\times n}$, ed F la forma quadratica associata ad A.

(1) Formulare la condizione necessaria e sufficiente su $\lambda_1, \dots, \lambda_n$ affinché F sia *definita negativa.*

(2) Formulare la condizione necessaria e sufficiente su $\lambda_1, \dots, \lambda_n$ affinché F sia *semidefinita negativa.*

(3) Indicare valori di $\lambda_1, \dots, \lambda_n$ affinché F risulti *sia semidefinita positiva che semidefinita negativa.*

(4) Indicare valori di $\lambda_1, \dots, \lambda_n$ affinché F risulti *semidefinita positiva ma non definita positiva.*

(5) Indicare valori di $\lambda_1, \dots, \lambda_n$ affinché F risulti *indefinita*, e poi formulare la condizione necessaria e sufficiente su $\lambda_1, \dots, \lambda_n$ affinché F sia *indefinita.*

Il *problema della classificazione della forma quadratica F* consiste nel *decidere se F sia definita (positiva o negativa), semidefinita (positiva o negativa) o indefinita.* Analogamente si pone il *problema della classificazione della matrice simmetrica A.*

5.5 Problema

Siano $A = \operatorname{diag}(-1, 0, 1, 2) \in \mathbb{R}^{4\times 4}$, ed F la forma quadratica associata ad A.

(1) Classificare F in base al *segno* degli elementi sulla diagonale.

(2) Determinare vettori non nulli $x, y, z \in \mathbb{R}^4$ tali che:

$$F(x) > 0 \quad , \quad F(y) = 0 \quad , \quad F(z) < 0$$

5.6 Osservazione (classificazione nel caso di matrice diagonale)

L'Esempio 5.3 ed il Problema 5.4 mostrano che il problema della classificazione di una forma quadratica è facilmente risolto nel caso sia associata ad una *matrice diagonale*, e che tale classificazione dipende *solo dal segno* degli elementi sulla diagonale. Precisamente, se la forma quadratica è associata alla matrice $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ si ha:

(D) la forma quadratica è *definita positiva* (rispettivamente: *negativa*) se e solo se *per ogni k si ha* $\lambda_k > 0$ (rispettivamente: $\lambda_k < 0$)

(S) la forma quadratica è *semidefinita positiva* (rispettivamente: *negativa*) se e solo se *per ogni k si ha* $\lambda_k \geqslant 0$ (rispettivamente: $\lambda_k \leqslant 0$)

(I) la forma quadratica è *indefinita* se e solo se *esistono λ_i e λ_j tali che* $\lambda_i \lambda_j < 0$.

Allo stesso modo si classifica la matrice $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

Il seguito di questo capitolo è dedicato alla soluzione del problema della classificazione nel caso generale.

5.2 Classificazione e autovalori

In questa sezione si introduce la nozione di *matrici congruenti* e si descrive come sia possibile classificare una forma quadratica in base al *segno degli autovalori* della matrice a cui la forma quadratica è associata.

Siano $A \in \mathbb{R}^{n\times n}$ una matrice simmetrica, F la forma quadratica ad essa associata e $C \in \mathbb{R}^{n\times n}$ una matrice invertibile (ovvero una matrice le cui colonne costituiscono *una base* di $\mathbb{R}^n$). Per ogni vettore $x \in \mathbb{R}^n$, sia ξ il vettore delle coordinate di x rispetto alla base di $\mathbb{R}^n$ costituita dalle colonne di C, ovvero:

$$x = C\xi$$

Si osservi che C definisce un'applicazione lineare invertibile tra i vettori e le coordinate. Si ha inoltre:

$$F(x) = Ax \bullet x = x^\mathsf{T} Ax = \xi^\mathsf{T} C^\mathsf{T} AC\xi = (C^\mathsf{T} AC)\xi \bullet \xi$$

La matrice $B = C^\mathsf{T} AC$ risulta *simmetrica* e, detta G la forma quadratica associata a B, l'uguaglianza ottenuta significa che *F in x ha lo stesso valore* (e quindi *lo stesso segno*) *di G in ξ*. Da questo si deduce che *F e G hanno la stessa classificazione* (risultano *entrambe* definite positive, oppure *entrambe* definite negative ecc.)

5.7 Definizione (matrici congruenti)

Matrici simmetriche $A, B \in \mathbb{R}^{n\times n}$ si dicono *congruenti* se esiste $C \in \mathbb{R}^{n\times n}$ invertibile tale che

$$B = C^\mathsf{T} AC$$

La matrice C *realizza la congruenza* tra A e B.

La congruenza è una relazione di equivalenza tra le matrici simmetriche di $\mathbb{R}^{n\times n}$.

Con la terminologia introdotta, il risultato precedente si enuncia: *se A e B sono matrici congruenti, la forma quadratica associata ad A e quella associata a B hanno la stessa classificazione.*

5.8 Osservazione

Per quanto mostrato nel Capitolo 2, la matrice simmetrica $A \in \mathbb{R}^{n\times n}$ è diagonalizzabile ed esiste una matrice *ortogonale* $Q \in \mathbb{R}^{n\times n}$ che realizza la similitudine, ovvero, detti $\lambda_1, \ldots, \lambda_n$ gli autovalori di A e ricordando che $Q^{-1} = Q^\mathsf{T}$, si ha:

$$A = Q\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)Q^\mathsf{T}$$

Le matrici A e $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ *sono dunque,* oltre che simili*, congruenti* (la matrice Q^T realizza la congruenza) e quindi la forma quadratica associata ad A e quella associata a $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ hanno la stessa classificazione. Quest'ultima si classifica come specificato nell'Osservazione 5.6.

Se ne deduce che:

5.9 Teorema (classificazione mediante gli autovalori)

Siano $A \in \mathbb{R}^{n\times n}$ una matrice simmetrica ed F la forma quadratica ad essa associata.

La forma quadratica F si classifica in base al segno *degli autovalori $\lambda_1, \ldots, \lambda_n$ di A come specificato nei punti* (`D`), (`S`) *ed* (`I`) *dell'Osservazione 5.6.*

5.10 Esempio

Sia:

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

ed F la forma quadratica ad essa associata. Il polinomio caratteristico di A è: $p_A(x) = (1-x)^2(-1-x)$, dunque gli autovalori sono $\lambda_1 = \lambda_2 = 1$ e $\lambda_3 = -1$.

Poichè $\lambda_1\lambda_3 < 0$, in base al punto (`I`) dell'Osservazione 5.6, la forma quadratica F è *indefinita.*

Dalla procedura di diagonalizzazione (Capitolo 1) applicata ad A, e scegliendo per ciascun autospazio una base ortonormale (certamente possibile perché A è simmetrica), si ottiene: $\mathrm{diag}(1, 1, -1) = Q^{\mathsf{T}} A\, Q$ con

$$Q = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Dunque, una matrice che realizza la congruenza è Q e, detta G la forma quadratica associata a $\mathrm{diag}(1, 1, -1)$, ovvero:

$$G(\xi) = \xi_1^2 + \xi_2^2 - \xi_3^2$$

si ha $F(x) = G(\xi)$ per tutti gli $x, \xi \in \mathbb{R}^3$ tali che $x = Q\xi$. Dunque, poiché:

$$G(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}) = 1 \qquad \text{e} \qquad G(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}) = -1$$

si ottiene:

$$F(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}) = 1 \qquad \text{e} \qquad F(\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}) = -1$$

a conferma dell'essere F indefinita.

Il teorema suggerisce la seguente procedura di classificazione della forma quadratica associata ad A: (a) calcolare gli autovalori di A, (b) di ciascuno guardare il segno e (c) classificare F di conseguenza. Tale procedura, però, non è *elementare* (il termine "procedura elementare" significa qui *che richiede solo un numero finito di operazioni aritmetiche per essere portata a termine*): basti pensare che calcolare gli autovalori è equivalente a determinare le radici del polinomio caratteristico di A.

Rileggendo le considerazioni che hanno portato al Teorema 5.9, si constata che la classificazione della forma quadratica associata alla matrice A è ottenuta in due passaggi: (1) determinare una matrice *diagonale* congruente ad A, (2) utilizzare l'Osservazione 5.6 per classificare la forma quadratica associata alla matrice diagonale. Il "passaggio critico" è il primo: le considerazioni che hanno portato al teorema suggeriscono di utilizzare come matrice diagonale congruente ad A la matrice $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ *simile* ad A, e questo obbliga a calcolare gli autovalori di A.

Ci si domanda se esistano matrici *diagonali* e *congruenti* ma *non necessariamente simili* ad A e, in caso affermativo, se sia possibile determinarne una *con una procedura elementare*.

La risposta a tale domanda è affermativa e nella prossima sezione si descrive una procedura elementare che per ciascuna matrice $A \in \mathbb{R}^{n\times n}$ simmetrica consente di determinare una matrice *diagonale e congruente* (ma non necessariamente simile) ad A, ed una matrice che realizza la congruenza.

5.3 Procedura di classificazione

In questa sezione si descrive una *procedura elementare* che, per ciascuna matrice $A \in \mathbb{R}^{n\times n}$ simmetrica determina una matrice diagonale e congruente ad A, ed una matrice che realizza la congruenza. La procedura viene descritta mediante alcuni esempi significativi.

Sia $A \in \mathbb{R}^{n\times n}$ simmetrica. Si ricordi che se H è una matrice di $\mathbb{R}^{n\times n}$ di elementi h_{ij}:

- *le righe* della matrice HA sono combinazioni lineari di quelle di A a coefficienti gli elementi h_{ij}. Precisamente, dette $r_1, \ldots, r_n$ le righe della matrice A e $\hat{r}_k$ la riga k-esima di HA si ha:
$$\hat{r}_k = h_{k1}r_1 + \cdots + h_{kn}r_n$$
- *le colonne* della matrice AH^T sono combinazioni lineari di quelle di A a coefficienti gli elementi h_{ij}. Precisamente, dette $c_1, \ldots, c_n$ le colonne della matrice A e $\hat{c}_k$ la colonna k-esima di AH^T si ha:
$$\hat{c}_k = h_{k1}c_1 + \cdots + h_{kn}c_n$$
Si osservi che i coefficienti della combinazione lineare delle colonne di A che determina $\hat{c}_k$ *sono gli stessi* della combinazione lineare delle righe di A che determina $\hat{r}_k$.

In questa sezione, per ogni intero k, si utilizza la sigla $a_{ij}^{(k)}$ per indicare l'elemento di posto i, j della matrice A_k.

5.11 Esempio

Sia:
$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{3\times3}$$
Essendo $a_{11} \neq 0$, si pone:
$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$
e:
$$A_1 = H_1AH_1^\mathsf{T} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} H_1^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$
Si osservi che:

(1) La matrice H_1 è invertibile (è triangolare con elementi non nulli sulla diagonale).

(2) *La prima riga di H_1A è la prima riga di A inalterata, gli elementi di H_1A al di sotto della diagonale nella prima colonna sono tutti nulli.* La scelta degli elementi di H_1 ha esattamente questo scopo: la seconda riga di H_1A è la seconda riga di A inalterata ($\hat{r}_2 = r_2$) perché $a_{21} = 0$; la terza riga di H_1A è invece $\hat{r}_3 = r_3 - r_1$ perché $a_{31} - a_{11} = 0$.

(3) A_1 è simmetrica, congruente ad A e le prime riga e colonna "sono diagonali".

Essendo $a_{22}^{(1)} = 0$, si pone:
$$H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$
e:
$$A_2 = H_2A_1H_2^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} H_2^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$
Si osservi che:

(1) La matrice H_2 è invertibile.

(2) *Le prime riga e colonna di A_2 coincidono con quelle di A_1, e l'elemento $a_{22}^{(2)}$ è non nullo.* La scelta degli elementi di H_2 ha esattamente questo scopo: la prima riga di H_2A_1 è la prima riga di A_1 inalterata, la seconda riga di H_2A_1 è invece $\hat{r}_2 = r_2 + r_3$ perché $a_{22}^{(1)} + a_{32}^{(1)} \neq 0$.

(3) A_2 è simmetrica e congruente ad A_1.

Essendo $a_{22}^{(2)} \neq 0$, si pone:

$$H_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 2 \end{bmatrix}$$

e, infine:

$$A_3 = H_3 A_2 H_3^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix} H_3^\mathsf{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Si osservi che:

(1) La matrice H_3 è invertibile.

(2) *Le prime due righe di H_3A_2 coincidono con quelle di A_2, e gli elementi di H_3A_2 al di sotto della diagonale nella seconda colonna sono tutti nulli.* La scelta degli elementi di H_3 ha esattamente questo scopo: la terza riga di H_3A_2 è $\hat{r}_3 = 2r_3 - r_2$ perché $2a_{32}^{(2)} - a_{22}^{(2)} = 0$.

(3) A_3 è diagonale e congruente ad A_2.

Per la proprietà transitiva della relazione di congruenza, la matrice diagonale A_3 è congruente ad A, e poiché:

$$A_3 = (H_3H_2H_1)A(H_1^\mathsf{T} H_2^\mathsf{T} H_3^\mathsf{T})$$

una matrice che realizza la congruenza (invertibile perché prodotto di matrici invertibili) è:

$$C = H_1^\mathsf{T} H_2^\mathsf{T} H_3^\mathsf{T} = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Secondo quanto stabilito nel punto (I) dell'Osservazione 5.6, la matrice $\operatorname{diag}(1, 2, -2)$ è *indefinita*, e quindi anche A e la forma quadratica F ad essa associata lo sono.

Si osservi che, detta G la forma quadratica associata a $\operatorname{diag}(1, 2, -2)$, ovvero:

$$G(\xi) = \xi_1^2 + 2\xi_2^2 - 2\xi_3^2$$

si ha $G(\xi) = F(x)$ per ogni $x, \xi \in \mathbb{R}^3$ tali che $\xi = Cx$. Dunque, essendo $G(e_2) = 2$ e $G(e_3) = -2$,* dette c_1, c_2 e c_3 le colonne della matrice C si ha $F(c_2) = 2$ e $F(c_3) = -2$.

5.12 Esempio

Sia:

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

*Si ricordi che con $e_1, \ldots, e_n$ si indica la *base canonica* di $\mathbb{R}^n$.

Essendo $a_{11} = 0$, si pone:

$$H_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

e:

$$A_1 = H_1 A H_1^{\mathsf{T}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} H_1^{\mathsf{T}} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Si osservi che:

(1) La matrice H_1 è invertibile.

(2) *L'elemento di $a_{1,1}^{(1)}$ è non nullo.* La scelta degli elementi di H_1 ha esattamente questo scopo.

(3) A_1 è simmetrica e congruente ad A.

Essendo $a_{11}^{(1)} \neq 0$, si pone:

$$H_2 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

e:

$$A_2 = H_2 A_1 H_2^{\mathsf{T}} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} H_2^{\mathsf{T}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 2 \\ 0 & 2 & -2 \end{bmatrix}$$

Si osservi che:

(1) La matrice H_2 è invertibile;

(2) *La prima riga di H_2A_1 è la prima riga di A_1 inalterata, gli elementi di H_2A_1 al di sotto della diagonale nella prima colonna sono tutti nulli.* La scelta degli elementi di H_2 ha esattamente questo scopo: la seconda riga di H_2A_1 è $\hat{r}_2 = 2r_2 - r_1$ perché $2a_{21}^{(1)} - a_{11}^{(1)} = 0$, la terza riga di H_2A_1 è $\hat{r}_3 = 2r_3 - r_1$ perché $2a_{31}^{(1)} - a_{11}^{(1)} = 0$.

(3) A_2 è simmetrica, congruente ad A_1 e le prime riga e colonna "sono diagonali".

Essendo $a_{22}^{(2)} \neq 0$, si pone:

$$H_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

e, infine:

$$A_3 = H_3 A_2 H_3^{\mathsf{T}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & 0 \end{bmatrix} H_3^{\mathsf{T}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Si osservi che:

(1) La matrice H_3 è invertibile;

(2) *Le prime due righe di H_3A_2 coincidono con quelle di A_2, e gli elementi di H_3A_2 al di sotto della diagonale nella seconda colonna sono tutti nulli.* La scelta degli elementi di H_3 ha esattamente questo scopo.

(3) A_3 è diagonale e congruente ad A_2.

Per la proprietà transitiva della relazione di congruenza, la matrice diagonale A_3 è congruente ad A, e poiché:

$$A_3 = (H_3 H_2 H_1) A (H_1^\mathsf{T} H_2^\mathsf{T} H_3^\mathsf{T})$$

una matrice che realizza la congruenza (invertibile perché prodotto di matrici invertibili) è:

$$C = H_1^\mathsf{T} H_2^\mathsf{T} H_3^\mathsf{T} = \begin{bmatrix} 1 & -1 & -2 \\ 0 & 2 & 2 \\ 1 & -1 & 0 \end{bmatrix}$$

Secondo quanto stabilito nel punto (I) dell'Osservazione 5.6, la matrice $\mathrm{diag}(2, -2, 0)$ è *indefinita*, e quindi anche A e la forma quadratica ad essa associata lo sono.

5.13 Problema

Sia F la forma quadratica dell'esempio precedente. Indicare $x, y \in \mathbb{R}^3$ tali che $F(x) > 0$ e $F(y) < 0$.

5.14 Esempio

Sia:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

Le prime riga e colonna di A "sono diagonali". Essendo $a_{22} = 0$, si pone:

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

e:

$$A_1 = H_1 A H_1^\mathsf{T} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} H_1^\mathsf{T} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Si osservi che:

(1) La matrice H_1 è invertibile.

(2) *L'elemento* $a_{22}^{(1)}$ *è non nullo.* La scelta degli elementi di H_1 ha esattamente questo scopo.

(3) A_1 è simmetrica e congruente ad A.

Essendo $a_{22}^{(1)} \neq 0$, si pone:

$$H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 2 \end{bmatrix}$$

e:

$$A_2 = H_2 A_1 H_2^\mathsf{T} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix} H_2^\mathsf{T} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Si osservi che:

(1) La matrice H_2 è invertibile;

(2) *Le prime due righe di H_2A_1 sono le prime due righe di A_1 inalterate, gli elementi di H_2A_1 al di sotto della diagonale nella seconda colonna sono tutti nulli.* La scelta degli elementi di H_2 ha esattamente questo scopo.

(3) A_2 è diagonale e congruente ad A_1.

Per la proprietà transitiva della relazione di congruenza, la matrice diagonale A_2 è congruente ad A, e poiché:

$$A_2 = (H_2H_1)A(H_1^\top H_2^\top)$$

una matrice che realizza la congruenza (invertibile perché prodotto di matrici invertibili) è:

$$C = H_1^\top H_2^\top = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

Secondo quanto stabilito nel punto (I) dell'Osservazione 5.6, la matrice $\mathrm{diag}(0, 2, -2)$ è *indefinita*, e quindi anche A e la forma quadratica ad essa associata lo sono.

5.15 Osservazione (invarianti per congruenza: inerzia di una matrice)

Si osservi che le matrici A e A_2 dell'ultimo esempio sono congruenti *ma hanno autovalori diversi*: gli autovalori di A sono $0, 1$ e -1, quelli di A_2 sono $0, 2$ e -2. *Gli autovalori non sono invarianti per congruenza.*

Invece, *è invariante per congruenza il numero di autovalori positivi, nulli e negativi* della matrice. Precisamente: Sia $A \in \mathbb{R}^{n\times n}$ una matrice *simmetrica* simile a $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$ (dunque $\lambda_1, \dots, \lambda_n$ sono gli autovalori di A) e siano N, Z e P il numero di elementi negativi, nulli e positivi dell'elenco $\lambda_1, \dots, \lambda_n$. La terna (N, Z, P) si chiama *inerzia* di A e: *Se A e B sono matrici congruenti, l'inerzia di A e quella di B coincidono.*†

Ad esempio, le matrici:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

e $\mathrm{diag}(0, 1, 0)$ sono congruenti. La prima ha autovalori $0, 0, 2$ e la seconda $0, 0, 1$. Entrambe hanno inerzia $(0, 2, 1)$.

Questo esempio evidenzia come sia banale determinare l'inerzia di una matrice diagonale. Questa constatazione, insieme all'invarianza dell'inerzia per congruenza, mostra che la procedura di classificazione descritta in questa sezione costituisce anche una *procedura elementare per determinare l'inerzia di una matrice.*

5.16 Problema

Siano $A, B \in \mathbb{R}^{n\times n}$ matrici diagonali e congruenti. Dimostrare che A e B hanno lo stesso numero di elementi negativi, nulli e positivi sulla diagonale.

5.4 Esercizi

(1) Classificare ciascuna delle seguenti matrici simmetriche, indicarne l'inerzia ed una matrice che realizza la congruenza. Nella soluzione si riportano solo l'inerzia e la classificazione della matrice: per la matrice che realizza la congruenza, si raccomanda di *verificare la risposta*!

(a) $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ [Soluzione: inerzia di $A = (0, 1, 1)$, semidefinita positiva.]

† Questo risultato è noto come *Legge d'inerzia di Sylvester.*

(b) $A = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$ [Soluzione: inerzia di $A = (1,1,1)$, indefinita.]

(c) $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ [Soluzione: inerzia di $A = (1,0,2)$, indefinita.]

(d) $A = \begin{bmatrix} 0 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$ [Soluzione: inerzia di $A = (1,1,1)$, indefinita.]

(e) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ [Soluzione: inerzia di $A = (0,2,1)$, semidefinita positiva.]

(f) $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 2 \\ -1 & 2 & 0 \end{bmatrix}$ [Soluzione: inerzia di $A = (1,0,2)$, indefinita.]

(2) Determinare la forma quadratica $F(x)$ associata alla matrice A definita nel Problema 1, parte (d), ed indicare $a, b, c \in \mathbb{R}^3$ tali che: $F(a) > 0$, $F(b) < 0$, $c \neq 0$ e $F(c) = 0$.

(3) Siano $A, B \in \mathbb{R}^{n\times n}$, simmetriche. Per ciascuno dei seguenti asserti, decidere se sia *vero* o *falso*:

(a) Se A è simile a B, allora A e B hanno la stessa inerzia.
(b) Se A è congruente a B, allora A e B sono simili.
(c) Se A è congruente a B, allora A e B hanno lo stesso determinante.
(d) Se A è congruente a B, allora $\det A = 0$ se e solo se $\det B = 0$.

(4) Siano $A = \mathrm{diag}(4, -9, 1, -1)$ e $B = \mathrm{diag}(1, 1, -1, -1)$.

(a) Determinare $P \in \mathbb{R}^{4\times 4}$ *di permutazione* tale che $PAP^\mathsf{T} = \mathrm{diag}(4, 1, -9, -1)$
(b) Determinare $S \in \mathbb{R}^{4\times 4}$ *diagonale* invertibile tale che $SBS^\mathsf{T} = \mathrm{diag}(4, 1, -9, -1)$
(c) Dimostrare che A e B sono congruenti.

(5) Determinare una matrice simmetrica associata alla forma quadratica $F(x)$ definita, per ogni $x \in \mathbb{R}^3$, da:

$$F(x) = -2x_2^2 + 3x_3^2 + 3x_1x_2 - 6x_2x_3$$

Decidere se la matrice determinata è l'unica che verifica la richiesta o se esiste qualche altra matrice, diversa da quella trovata, che definisce F.

(6) Siano $A \in \mathbb{R}^{n\times n}$ una matrice simmetrica, ed F la forma quadratica ad essa associata. Detti a_{ij} gli elementi di A e ricordando che $e_1, \ldots, e_n$ sono gli elementi della base canonica di $\mathbb{R}^n$, constatare che:

$$F(e_1) = a_{11} \quad , \quad \ldots \quad , \quad F(e_n) = a_{nn}$$

Se ne deduce che:

(a) *condizione necessaria* affinché la forma quadratica sia *definita positiva* è che:

$$a_{11} > 0 \quad , \quad \ldots \quad , \quad a_{nn} > 0$$

(b) *condizione sufficiente* affinché la forma quadratica sia *indefinita* è che esistano a_{ii} e a_{kk} tali che $a_{ii}a_{kk} < 0$.

Infine, verificare che la forma quadratica associata alla matrice:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

è *indefinita*. Questo esempio mostra che la condizione (a) *non è sufficiente*, e la condizione (b) *non è necessaria*.

(7) Sia $A \in \mathbb{R}^{n \times n}$ una matrice simmetrica. Per ciascuno dei seguenti asserti, decidere se sia *vero* o *falso*:

(a) Se A è definita positiva, allora A è invertibile.

(b) Se A è definita positiva, allora per ogni intero k la matrice A^k è definita positiva. (Suggerimento: ricordare il legame tra gli autovalori di A e quelli di A^k, ed utilizzare il Teorema 5.9.)

(c) Se A è definita negativa, allora A^2 è definita positiva.

(d) Se A è indefinita, allora A^2 è semidefinita positiva.

(8) Siano $A \in \mathbb{R}^{k \times n}$ ed F la forma quadratica associata alla matrice $A^{\mathsf{T}}A \in \mathbb{R}^{n \times n}$, simmetrica. Per ogni $x \in \mathbb{R}^n$ si ha:

$$F(x) = (A^{\mathsf{T}}A)x \bullet x = x^{\mathsf{T}}A^{\mathsf{T}}Ax = Ax \bullet Ax = \|Ax\|^2 \geqslant 0$$

dunque: *F è semidefinita positiva.*

Dimostrare che: *F è definita positiva se e solo se le colonne di A sono elementi linearmente indipendenti di $\mathbb{R}^k$.*