NOTA BENE. Questi appunti non sono esaustivi, non contengono tutto quanto detto a lezione/esercitazioni. Non bastano come materiale per preparare l'esame. Possono essere utili per un veloce ripasso.

Matrici reali simmetriche, teorema spettrale, forme quadratiche (Capitolo 9 dello Schlesinger e capitoli 8/9 del Bernardi Gimigliano)

Def.1. Una matrice quadrata e reale si dice *ortogonalmente diagonalizzabile* se esiste una matrice ortogonale U tale che $U^{-1}AU$ sia diagonale. Poiché U è ortogonale si ha $U^{-1}=U^T$ e quindi una matrice è ortogonalmente diagonalizzabile se esiste una matrice ortogonale U tale che U^TAU sia diagonale .

Una matrice reale ortogonalmente diagonalizzabile è simmetrica, infatti da $U^TAU=D$ con D diagonale abbiamo $(U^TAU)^T=D^T=D= U^TAU$.

Studiamo allora le matrici reali simmetriche.

Ci viene utile la seguente

Def.2. Dati due vettori $\underline{v}=[v_1,v_2,...,v_n]^T$, $\underline{w}=[w_1,w_2,...,w_n]^T$ in C^n si dice *prodotto hermitiano* $\underline{v}\times\underline{w}$ il numero complesso $\sum_{i=1}^{n} v_i\overline{w}_i = \underline{v}^T\overline{\underline{w}}$, dove $\overline{w}_i$ indica il coniugato di w_i e $\overline{\underline{w}}$ indica il vettore le cui componenti sono i coniugati delle componenti di $\underline{w}$.

Proposizione 1. Gli autovalori di una matrice reale simmetrica sono reali.

*Dim.*Sia λ un autovalore di A. Abbiamo $A\underline{v}\times\underline{v} = (A\underline{v})^T\overline{\underline{v}} =\underline{v}^TA^T\overline{\underline{v}}= \underline{v}^TA\overline{\underline{v}}$ (perché A è simmetrica) e $\underline{v}\times A\underline{v} = (\underline{v})^T\,\overline{A\underline{v}} =\underline{v}^TA\overline{\underline{v}}$ (perché A è reale) ; ma $A\underline{v}\times\underline{v}=\lambda\underline{v}\times\underline{v}=\lambda(\underline{v}\times\underline{v})$ e $\underline{v}\times A\underline{v}=\underline{v}\times\lambda\underline{v}=\overline{\lambda}(\underline{v}\times\underline{v})$ e quindi $\lambda(\underline{v}\times\underline{v})= \overline{\lambda}(\underline{v}\times\underline{v})$ da cui $\lambda= \overline{\lambda}$, quindi λ è reale.

Proposizione 2. Sia A una matrice reale simmetrica. Allora autovettori associati ad autovalori distinti sono ortogonali.

Dim. Siano λ,μ due autovalori distinti di A a cui sono associati rispettivamente gli autovettori $\underline{v}$ e $\underline{w}$. Poiché A è reale e simmetrica, λ e μ sono reali per la proposizione 1 e quindi $\underline{v}$, $\underline{w} \in R^n$. Si ha $(A\underline{v})^T\underline{w}=\underline{v}^TA^T\underline{w}= \underline{v}^TA\underline{w}$, ma $(A\underline{v})^T\underline{w}=\lambda\underline{v}^T\underline{w}$ e $\underline{v}^TA\underline{w} =\mu\underline{v}^T\underline{w}$ e quindi $=\lambda\underline{v}^T\underline{w} = \mu\underline{v}^T\underline{w}$, da cui $\underline{v}^T\underline{w} =0$.

Possiamo ora dimostrare il seguente

Teorema 1 (Teorema spettrale o degli assi principali). Una matrice reale simmetrica è ortogonalmente diagonalizzabile.

Dim. Il teorema si dimostra per induzione sull'ordine n di A. Il teorema vale ovviamente se n=1. Supponiamo ora che il teorema valga per matrici reali simmetriche di ordine n-1. Sia A una matrice reale simmetrica di ordine n. Per il teorema fondamentale dell'algebra il polinomio caratteristico di A ha almeno una radice λ_1 in C che per la Proposizione 1, essendo un auto valore di A, è reale. Sia allora $\underline{q}_1$ un autovettore di A associato all'autovalore λ_1 di norma 1 (ed ovviamente appartenente a R^n). Estendiamo $\underline{q}_1$ ad una base ortonormale $B=\{\underline{q}_1, \underline{q}_2,..., \underline{q}_n\}$ di R^n (cosa che si può fare estendendo $\underline{q}_1$ ad una base di R^n e poi applicando l'algoritmo di Gram Schmidt e la normalizzazione). Sia $P=[\,\underline{q}_1|\,\underline{q}_2|...|\underline{q}_n]$ la matrice ortogonale formata dall'accostamento dei vettori colonna della base ortonormale B e sia $C=P^TAP=P^{-1}AP$. La matrice C rappresenta l'applicazione lineare f_A associata ad A rispetto alla nuova base B . La prima colonna di C è allora il vettore delle coordinate di $f_A(\underline{q}_1)$ rispetto alla base

B. Poiché $f_A(\underline{q}_1) = \lambda_1\underline{q}_1$ si ha $\underline{q}_1=\begin{bmatrix}\lambda_1\\0\\\vdots\\0\end{bmatrix}$. La matrice C è reale simmetrica, infatti $C^T=(P^TAP)^T=P^TA^T(P^T)^T=P^TAP=C$, dunque $C=\left[\begin{array}{c|c}\lambda_1 & \underline{0}^T\\ \hline \underline{0} & A_1\end{array}\right]$, dove A_1 è a sua volta essere reale simmetrica ed ha ordine n-1, quindi per ipotesi di induzione esiste una matrice ortogonale Q tale che $Q^TA_1Q=D_1$ con D_1 matrice diagonale. Consideriamo allora la matrice $R=\left[\begin{array}{c|c}1 & \underline{0}^T\\ \hline \underline{0} & Q\end{array}\right]$. Si verifica subito che $R^TCR=$

$\left[\begin{array}{c|c}1 & \underline{0}^T\\ \hline \underline{0} & Q\end{array}\right]^T\left[\begin{array}{c|c}\lambda_1 & \underline{0}^T\\ \hline \underline{0} & A_1\end{array}\right]\left[\begin{array}{c|c}1 & \underline{0}^T\\ \hline \underline{0} & Q\end{array}\right]=\left[\begin{array}{c|c}1 & \underline{0}^T\\ \hline \underline{0} & Q^T\end{array}\right]\left[\begin{array}{c|c}\lambda_1 & \underline{0}^T\\ \hline \underline{0} & A_1\end{array}\right]\left[\begin{array}{c|c}1 & \underline{0}^T\\ \hline \underline{0} & Q\end{array}\right]=\left[\begin{array}{c|c}\lambda_1 & \underline{0}^T\\ \hline \underline{0} & Q^TA_1Q\end{array}\right]=\left[\begin{array}{c|c}\lambda_1 & \underline{0}^T\\ \hline \underline{0} & D_1\end{array}\right]$ è una matrice diagonale D. Quindi sostituendo a C la sua espressione P^TAP in funzione di A si ottiene $R^TP^TAP\,R=(PR)^TAPR=D$. La matrice U=PR è ortogonale, essendo prodotto di due matrici ortogonali, e quindi A è ortogonalmente diagonalizzabile.

Ne segue

<u>Corollario 1</u>. Sia A una matrice reale simmetrica di ordine n e siano $\lambda_1, \lambda_2,..., \lambda_s$ i suoi autovalori distinti . Sia $V(\lambda_i)$ l'autospazio associato all'autovalore λ_i e sia B_i una base ortonormale di $V(\lambda_i)$. Allora

- $B_1\cup B_2\cup...\cup B_s$ è una base ortonormale di R^n.
- R^n è la somma diretta degli autospazi $V(\lambda_i)$
- Gli autospazi $V(\lambda_i)$ sono a due a due ortogonali.

Dim. Dal teorema spettrale sappiamo che A è ortogonalmente diagonalizzabile, quindi ogni autovalore ha molteplicità geometrica uguale alla molteplicità algebrica, quindi la somma delle dimensioni degli autospazi associati agli autovalori di A è n. Dunque $B_1\cup B_2\cup...\cup B_s$ è formata da n elementi, ogni coppia di vettori in uno stesso B_i è formata per definizione da vettori ortonormali ed ogni vettore di B_i è per la Proposizione 2 ortogonale ad ogni vettore di B_j se $i\neq j$. Abbiamo così provato che $B_1\cup B_2\cup...\cup B_s$ è una base ortonormale di R^n. Da questo segue che R^n è somma degli autospazi $V(\lambda_i)$. Inoltre è noto che due autospazi associati ad autovalori distinti hanno in comune solo il vettore $\underline{0}$. La Proposizione 2 dice che gli autospazi $V(\lambda_i)$. sono a due a due ortogonali.

<u>Corollario 2</u>. Sia A una matrice reale simmetrica di ordine n e siano $\lambda_1, \lambda_2,..., \lambda_s$ i suoi autovalori distinti . Sia P_i la matrice proiezione sull'autospazio $V(\lambda_i)$ associato all'autovalore λ_i. Allora

- $I=P_1+P_2+...+P_s$
- $A=\lambda_1P_1+\lambda_2P_2+...+\lambda_sP_s$ (decomposizione spettrale di A)
- Ogni P_i è simmetrica, $P_i^2= P_i$, $P_iP_j=0$.

Dim. Poiché gli autospazi $V(\lambda_i)$ sono a due a due ortogonali ogni vettore $\underline{v}$ di R^n si può scrivere come somma delle sue proiezioni ortogonali su tali autospazi. Avviamo dunque $\underline{v}=\underline{v}_1+\underline{v}_2+\ldots+\underline{v}_s$ dove, per ogni i, $\underline{v}_i=P_i\underline{v}$, dunque per ogni $\underline{v}$ si ha $I\underline{v}= P_1\underline{v} + P_2\underline{v} +\ldots+ P_s\underline{v}$ da cui si ricava la prima asserzione. Moltiplicando a sinistra per A entrambi i membri di $I\underline{v}= P_1\underline{v} + P_2\underline{v} +\ldots+ P_s\underline{v}$ si ottiene $A\underline{v}= AP_1\underline{v} + AP_2\underline{v} +\ldots+A\,P_s\underline{v}$ da cui essendo $P_i\underline{v}$ un elemento dell'autospazio $V(\lambda_i)$ si ricava ottiene $A\underline{v}= \lambda_1 P_1\underline{v} + \lambda_2 P_2\underline{v} +\ldots+\lambda_s P_s\underline{v}$ da cui si ricava la seconda asserzione.
Da quanto osservato nel capitoletto sugli spazi euclidei per le matrici proiezione, sappiamo che ogni P_i è simmetrica e che $P_i^2= P_i$. Ora per ogni vettore $\underline{v}$, $P_j(\underline{v})\in V(\lambda_j)$ e se $j\neq i$ $P_i(P_j\underline{v})=\underline{0}$ essendo la proiezione ortogonale su $V(\lambda_i)$ di un vettore ortogonale a $V(\lambda_i)$ per cui P_iP_j è la matrice nulla.

Dalla decomposizione spettrale $A=\lambda_1P_1+\lambda_2P_2+...+\lambda_sP_s$ ricaviamo, tenuto conto delle proprietà degli autovalori e autovettori associati, che per ogni n>0 si ha $A^n=\lambda_1^nP_1+\lambda_2^nP_2+...+\lambda_s^nP_s$; inoltre se tutti gli autovalori di A sono diversi da 0 si ha $A^{-1}=\lambda_1^{-1}P_1+\lambda_2^{-1}P_2+...+\lambda_s^{-1}P_s$; se f(x) è una funzione nel cui campo di esistenza stanno gli autovalori abbiamo $f(A)= f(\lambda_1)P_1+ f(\lambda_2)P_2+...+ f(\lambda_s)Ps$, in particolare se gli autovalori di A sono tutti non negativi possiamo definire $\sqrt{A}=\sqrt{\lambda_1}P_1+...+\sqrt{\lambda_s}P_s$. Si verifica che la matrice così ottenuta è l'unica matrice il cui quadrato è uguale ad A.

Endomorfismi simmetrici

Def 3. Sia V uno spazio eucldeo. Un endomorfismo $f:V\rightarrow V$ si dice *simmetrico* se per ogni $\underline{v},\underline{w}\in V$ si ha $<\underline{v},f(\underline{w})>=<f(\underline{v}),\underline{w}>$.

- Un endomorfismo è simmetrico se e solo se rispetto ad ogni base ortonormale di V è rappresentato da una matrice simmetrica.
  Osserviamo che lavorando con base ortonormale la matrice di Gram è la matrice identica e quindi il prodotto scalare diventa il prodotto scalare standard. Se $f(\underline{v})=A\underline{v}$, con A matrice simmetrica, si ha allora $<\underline{v},f(\underline{w})>= \underline{v}^TAf(\underline{w})=\underline{v}^TA^Tf(\underline{w})=(A\underline{v})^Tf(w)=<f(\underline{v}),\underline{w}>$.. Viceversa se per ogni $\underline{v},\underline{w}\in V$ è $\underline{v}^TAf(\underline{w})=\underline{v}^TA^Tf(\underline{w})$ si ricava subito $A=A^T$.
- Un endomorfismo è simmetrico se e solo se rispetti ad una base ortonormale di V è rappresentato da una matrice simmetrica.
  Ovviamente se f è simmetrico è rappresentato da una matrice simmetrica rispetto ad una base ortonormale di V. Viceversa sia f rappresentato da una matrice simmetrica A rispetto ad una base ortonormale *B* di V. Sia *B'* una qualsiasi altra base ortonormale di V. Rispetto a *B'* f è rappresentato da $U^{-1}AU$ con U matrice di passaggio tra le due basi ortonormali e quindi ortogonale. Pertanto $U^{-1}AU = U^TAU$ e anche rispetto alla nuova base f è rappresentata da una matrice simmetrica.

.Forme quadratiche.

Def. 4. Una *forma quadratica reale* $q(\underline{x})$ nelle variabili $x_1,x_2,...,x_n$ è un polinomio di secondo grado (quadratica) omogeneo (forma) a coefficienti reali nelle variabili $x_1,x_2,...,x_n$; quindi $q(\underline{x})=a_{11}x_1^2+2a_{12}x_1x_2++a_{22}x_2^2+2a_{13}x_1x_3+...+2a_{n-1n}\,x_{n-1}x_n+a_{nn}x_n^2=\underline{x}^TA\underline{x}$ dove $\underline{x}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$, ed A è la matrice reale simmetrica $A=\begin{bmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{12} & a_{22} & \dots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{1n} & a_{2n} & \dots & a_{nn}\end{bmatrix}$. La forma quadratica si dice allora rappresentata da A.

Esempio 1. La forma quadratica $x^2+4xy-y^2+xz-3z^2$ è rappresentata dalla matrice $A=\begin{bmatrix}1 & 2 & \frac{1}{2}\\ 2 & -1 & 0\\ \frac{1}{2} & 0 & -3\end{bmatrix}$.

Per ogni forma quadratica $q(\underline{x})$ si ha ovviamente $q(\underline{0})=0$ e $q(t\underline{x})=t^2q(\underline{x})$, per ogni $t\in R$.

Def. 5. Una forma quadratica reale nelle variabili $x_1,x_2,...,x_n$ si dice

- *definita positiva* (*negativa*) se $q(\underline{x})>0$ (<0) per ogni $\underline{x}\in R^n$, $\underline{x}\neq\underline{0}$,
- *semidefinita positiva* (*negativa*) se $q(\underline{x})\geq 0$ (≤ 0) per ogni $\underline{x}\in R^n$, ed esiste un $\underline{x}\neq\underline{0}$ per cui $q(\underline{x})=0$,

- *indefinita* se esistono $\underline{x}_1,\underline{x}_2\in R^n$ tali che $q(\underline{x}_1)<0<q(\underline{x}_2)$.

Una matrice simmetrica A si dice (semi)definita positiva (negativa), indefinita se la forma quadratica $\underline{x}^T A\underline{x}$ è (semi)definita positiva (negativa), indefinita.

Le matrici reali simmetriche definite positive sono legate al prodotto scalare. Infatti abbiamo visto che dato uno spazio euclideo V di dimensione finita n e considerata una sua base $B=\{\underline{b}_1,\underline{b}_2,\ldots,\underline{b}_n\}$ il prodotto scalare di due vettori $\underline{v},\underline{w}\in V$ si può scrivere nella forma $(\underline{v}|_B)^T G\underline{w}|_B$ dove G è la matrice iil cui generico elemento di posto (i,j) è il prodotto scalare $<\underline{b}_i,\underline{b}_j>$, G si chiama matrice di Gram del prodotto scalare e per le proprietà del prodotto scalare deve essere una matrice reale simmetrica definita positiva. Viceversa se abbiamo una matrice G reale simmetrica definita positiva $<\underline{x},\underline{y}>=\underline{x}^T G\underline{y}$ è un prodotto scalare su R^n..

Lo studio del segno di una forma quadratica ha molte applicazioni ed è facilitato se la forma quadratica è trasformata in una forma che contiene solo i termini coi quadrati delle variabili.
Consideriamo la forma quadratica $q(\underline{x}) =\underline{x}^T A\underline{x}$, con un cambiamento di base $\underline{x}=S\underline{X}$ la forma diventa $q(\underline{X})=\underline{X}^T(S^TAS)\underline{X}$.

Poiché A è reale simmetrica, per il teorema spettrale esiste sempre una trasformazione ortogonale $\underline{x}=Q\underline{X}$ tale che $q(\underline{x})$ diventi $q(\underline{X})= \lambda_1X_1^2+\lambda_2 X_2^2+...+\lambda_n X_n^2$ dove $\lambda_1, \lambda_2,..., \lambda_n$ sono gli autovalori di A. Siano λ_{min}, λ_{max} rispettivamente il minimo ed il massimo fra gli autovalori di A, dalla espressione $q(\underline{X})= \lambda_1X_1^2+\lambda_2 X_2^2+...+\lambda_n X_n^2$ si ricava subito che $\lambda_{min}\|\underline{X}\|^2\leq q(X)\leq \lambda_{max}\|\underline{X}\|^2$. Poiché ovviamente $q(\underline{x})$ e $q(\underline{X})$ hanno lo stesso segno, si ha

<u>Proposizione 3</u>. Sia $q(\underline{x})=\underline{x}^T A\underline{x}$. Allora
- $q(\underline{x})$ è definita positiva (negativa) se A ha tutti gli autovalori positivi (negativi)
- $q(\underline{x})$ è semidefinita positiva (negativa) se A ha tutti gli autovalori non negativi (non positivi) ed uno almeno nullo (quindi A ha determinante 0)
- $q(\underline{x})$ è indefinita se ha un autovalore negativo ed uno positivo.

<u>Def.5.</u> Si chiama *minore principale* di una matrice quadrata A di ordine n il determinante di una qualsiasi sottomatrice quadrata di A la cui diagonale stia sulla diagonale di A. Si chiama *minore principale di Nord Ovest (NO)* di ordine k di A, il determinante della sottomatrice ottenuta da A eliminando le ultime n-k righe e le ultime n-k colonne di A.

Si può verificare che

<u>Proposizione 4.</u> Una matrice reale simmetrica A è definita positiva se e solo se i minori principali di nord ovest di A sono positivi, è semidefinita positiva se e solo se det A=0 e i minori principali di NO di A sono non negativi, è definita negativa se e solo se i minori principali di nord ovest di A di ordine k hanno il segno di $(-1)^k$, è semidefinita negativa se e solo se det A=0 e i minori principali di NO di A o sono nulli o hanno il segno di $(-1)^k$, altrimenti è indefinta.

Questa proposizione è di aiuto nel calcolare il segno di una forma quadratica, in quanto è in generale difficile calcolare esplicitamente gli autovalori di una matrice.

<u>Def .6.</u> Una matrice B si dice *congruente* ad una matrice A se esiste una matrice non singolare S tale che $B=S^TAS$.

Si può verificare che
- Se B è congruente ad una matrice simmetrica A, anche B è simmetrica

- La relazione di congruenza fra matrici simmetriche è una relazione di equivalenza.
- Sia $q(x)=\underline{x}^T A\underline{x}$ una forma quadratica, allora A è congruente ad una matrice diagonale (basta infatti usare il teorema spettrale, visto che A è reale e simmetrica)
- Due matrici congruenti non sono sempre simili (infatti sono simili se e solo se S è una matrice ortogonale)
- Due matrici sono congruenti se e solo se rappresentano a stessa forma quadratica rispettoa coordinate distinte
- Due matrici sono congruenti se e solo se hanno lo stesso numero di autovalori positivi e lo stesso numero di autovalori negativi (legge di inerzia di Sylvester);
- Sia A una matrice quadrata e simmetrica di ordine con s autovalori positivi, t autovalori negativi (ovviamente $s+t\leq n$), allora esiste un cambiamento di coordinate $\underline{x}=S\underline{X}$ tale che la forma quadratica associata ad A diventa $q(\underline{X})=X_1^2+X_2^2+...+X_t^2-X_{t+1}^2-...-X_{t+s}^2$.

La terna (t, n-t-s,s) si dice *segnatura delle forma quadratica*.

Per estensione a quanto si fa in R^2 ed R^3 introduciamo le seguenti definizioni

Def.7. Chiamiamo *rototraslazione in* R^n una qualsiasi trasformazione del tipo $\underline{x}=U\underline{X}+\underline{v}$,

dove $\underline{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\underline{v}=\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ ed U è una matrice ortogonale di ordine n con det U=1.

Quadriche in R^n.

Def.8. Chiamiamo *quadrica* di R^n l'insieme dei vettori $\underline{x}$ di R^n tali che $\underline{x}^T A\underline{x}+2\underline{b}^T\underline{x}+c=0$
dove A è una matrice reale simmetrica di ordine n, $\underline{b}$ è un vettore di R^n, c è un numero reale.
Se si identificano i vettori di R^n con i punti di uno spazio ad n dimensioni le cui coordinate sono le componenti del vettore rispetto alla base canonica possiamo chiamare quadrica di R^n il luogo dei punti di R^n tali le cui coordinate soddisfano una equazione $p(\underline{x})=0$ con $p(\underline{x})=\underline{x}^T A\underline{x}+2\underline{b}^T\underline{x}+c$ polinomio di secondo grado nelle variabili $x_1,x_2,...,x_n$. $\underline{x}^T A\underline{x}$ è la parte dei termini quadratici del polinomio (quindi è una forma quadratica), $2\underline{b}^T\underline{x}$ è il complesso dei termini lineari, c è il termine noto.

Lemma 1. Sotto l'azione di una rototraslazione $\underline{x}= U\underline{X}+\underline{v}$, un polinomio di secondo grado $p(\underline{x})= \underline{x}^T A\underline{x}+2\underline{b}^T\underline{x}+c$ viene portato nella forma $\tilde{p}(\underline{X}) =\underline{X}^T\tilde{A}\underline{X}+2\tilde{\underline{b}}^T\underline{X}+ \tilde{c}$ dove :

- $\tilde{A} =U^TAU$,
- $\tilde{\underline{b}} =U^T(A\underline{v}+\underline{b})$,
- $\tilde{c}=\underline{v}^T A\underline{v}+2\underline{b}^T\underline{v}+c$.

Dim. Sostituendo $U\underline{X}+\underline{v}$ ad $\underline{x}$ in $p(\underline{x})$ si ha $\tilde{p}(\underline{X}) =(\underline{X}^TU^T+\underline{v}^T)A(U\underline{X}+\underline{v})+2\underline{b}^T(U\underline{X}+\underline{v})+c= \underline{X}^TU^TAU\underline{X}+ \underline{X}^TU^TA\underline{v}+\underline{v}^TAU\underline{X}+ \underline{v}^TA\underline{v}+2\underline{b}^TU\underline{X}+2\underline{b}^T\underline{v}+c$, da cui tenuto conto che $\underline{X}^TU^TA\underline{v}=\underline{v}^TAU\underline{X}$ (perché $(\underline{X}^TU^TA\underline{v})^T=\underline{v}^TAU\underline{X}$ essendo A simmetrica ed una matrice di tipo (1,1) coincide con la sua trasposta), si ottiene $\tilde{p}(\underline{X}) = \underline{X}^TU^TAU\underline{X}+2\underline{v}^TAU\underline{X}+ \underline{v}^TA\underline{v}+2\underline{b}^TU\underline{X}+2\underline{b}^T\underline{v}+c +2\underline{b}^TUX+v^TAv+2bTv+c= \underline{X}^T(U^TAU)\underline{X}+2[U^T(A\underline{v}+\underline{b})]^T\underline{X}+(\underline{v}^TA\underline{v}+2\underline{b}^T\underline{v}+\underline{c})$, in cui ancora si è tenuto conto che A è simmetrica.

Osservazione 1. L'equazione $\underline{x}^T A\underline{x}+2\underline{b}^T\underline{x}+c=0$ può essere scritta nella forma più compatta $\underline{z}^TB\underline{z} =0$,

dove $\underline{z}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{bmatrix}$ e $B=\left[\begin{array}{c|c} A & \underline{b} \\ \hline \underline{b}^T & c \end{array}\right]$. La rototraslazione $\underline{x}= U\underline{X}+\underline{v}$ può essere scritta nella forma $\underline{z}=F\underline{Z}$ dove

$\underline{Z}==\begin{bmatrix} X_1 \\ \vdots \\ X_n \\ 1 \end{bmatrix}$, F= $\left[\begin{array}{c|c} U & \underline{v} \\ \hline \underline{0}^T & 1 \end{array}\right]$. Usando queste notazioni, sotto l'azione della rototraslazione, l'equazione della quadrica diventa $\underline{Z}^T\widetilde{B}\,\underline{Z}=0$ con $\widetilde{B}=F^TBF$, B e $\widetilde{B}$ quindi sono matrici congruenti ma non simili in quanto F non è ortogonale, hanno tuttavia lo stesso determinante in quanto det F=1.

Possiamo quindi dimostrare il seguente

<u>Teorema 2</u> Sia $p(\underline{x})=\underline{z}^TB\underline{z}=\underline{x}^TA\underline{x}+2\underline{b}^T\underline{x}+c$ un polinomio di secondo grado nelle variabili $x_1,x_2,\ldots,x_n$ (ovvero sia $p(\underline{x})=0$ l'equazione di una quadrica in R^n) Allora:

- gli autovalori e quindi il rango e la traccia di A, il determinante ed il rango di B sono invarianti per rototraslazioni;
- se r=rk A e $\lambda_1,\lambda_2,\ldots,\lambda_r$ sono gli autovalori non nulli di A, il polinomio $p(\underline{x})$ può essere portato tramite una isometria (rototraslazione) $\underline{x}=U\underline{X}+\underline{v}$ in una delle seguenti forme
  - $\overline{p}(\underline{X})=\lambda_1X_1^2+\lambda_2X_2^2+\ldots+\lambda_rX_r^2+c'$, con c'=0 se rk B=r c'≠0 se rk B=r+1
  - $\overline{p}(\underline{X})=\lambda_1X_1^2+\lambda_2X_2^2+\ldots+\lambda_rX_r^2++2pX_{r+1}$ con p≠0, se rk B=r+2.

Dim. Come visto nel lemma 1 la matrice $\widetilde{A}$ ottenuta da A dopo aver applicato una rototraslazione è ortogonalmente simile ad A, e quindi ha gli stessi autovalori di A e di conseguenza lo stesso determinante (che è il prodotto degli autovalori) e la stessa traccia (che è la somma degli autovalori). Come notato nell'osservazione 1 $\widetilde{B}$ non è simile a B ma ha lo stesso determinante di B, ed inoltre essendo F ha anche lo stesso rango di B. Abbiamo quindi dimostrato il primo punto del teorema. Portiamo ora $p(\underline{x})$ a forma canonica, mediante un'opportuna rototraslazione. Osserviamo che se agiamo su $p(\underline{x})$ con una traslazione del tipo $\underline{x}=\underline{X}+\underline{v}$, non modifichiamo i termini quadratici (ovvero la matrice A), mentre trasformiamo il vettore $\underline{b}$ in $A\underline{v}+\underline{b}$; se invece agiamo su $p(\underline{x})$ con una rotazione del tipo $\underline{x}=U\underline{X}$ trasformiamo A in U^TAU e $\underline{b}$ in $U^T\underline{b}$. Useremo quindi le traslazioni per eliminare, dove possibile, i termini lineari nell'equazione della quadrica. Distinguiamo due casi
1) Sia rk A=rk [A| $\underline{b}$], allora esiste un vettore $\underline{w}$ soluzione del sistema lineare $A\underline{x}=-\underline{b}$, per cui, con la traslazione $\underline{x}=\underline{x}'+\underline{w}$, otteniamo $p_1(\underline{x}')=p(\underline{x}'+\underline{w})=(\underline{x}')^TA\underline{x}'+2(A\underline{w}+\underline{b})^T\underline{x}'+c'=(\underline{x}')^TA\underline{x}'+c'$ dove $c'=\underline{w}^TA\underline{w}+2\underline{b}^T\underline{w}+c$ (tenuto conto del Lemma 1). Poiché A è reale simmetrica per il teorema spettrale esiste una matrice ortogonale U, tale che $U^TAU=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_r,0,\ldots,0)$. U è la matrice formata dall'accostamento dei vettori di una base di autovettori di A. Si può assumere che U abbia determinante 1, perché in caso il determinante fosse -1, basta cambiare il segno di un autovettore, dunque la trasformazione $\underline{x}'=U\underline{X}$ è una isometria lineare del tipo richiesto. Applicando tale isometria a $p_1(\underline{x}')$ otteniamo $\overline{p}(\underline{X})=p_1(U\underline{X})=\lambda_1X_1^2+\lambda_2X_2^2+\ldots+\lambda_rX_r^2+c'$. Ora l'equazione della quadrica in forma compatta diventa $\underline{Z}^TB'\underline{Z}=0$ ove se c'=0 $B'=\left[\begin{array}{c|c} U^TAU & 0 \\ \hline \underline{0}^T & 0 \end{array}\right]$ e dunque rk A=rk (U^TAU) = rk B'=rk B; se invece c'≠0 $B'=\left[\begin{array}{c|c} U^TAU & 0 \\ \hline \underline{0}^T & c' \end{array}\right]$ e quindi rk B=rk B'= rk (U^TAU) +1=rk A+1 ed abbiamo quindi trovato il primo caso del secondo punto del teorema.
2) Sia allora rk A<rk [A| $\underline{b}$]. In tal caso $\underline{b}\notin Col(A)$. Allora possiamo scrivere $\underline{b}=\underline{b}_0+\underline{b}_1$, dove $\underline{b}_0\in Col(A)$ e $\underline{b}_1\in Col(A)^{\perp}$ ed ovviamente $\underline{b}_1\neq\underline{0}$. Essendo A simmetrica, Col(A)=Row(A), e quindi $\underline{b}_1\in Row(A)^{\perp}=\ker A$ (infatti un vettore $\underline{v}$ appartiene a $Row(A)^{\perp}$ se e solo se è ortogonale ad ogni generatore di Row(A), cioè se e solo se il prodotto (righe per colonne) di ogni riga di A per $\underline{v}$ è 0, quindi se e solo se $A\underline{v}=\underline{0}$. Poiché $\underline{b}_0\in Col(A)$ esiste un vettore $\underline{w}$ soluzione del sistema $A\underline{x}=-\underline{b}_0$, per cui con la traslazione $\underline{x}=\underline{x}'+\underline{w}$, otteniamo $p_1(\underline{x}')=p(\underline{x}'+\underline{w})=(\underline{x}')^TA\underline{x}'+2\underline{b}_1^T\underline{x}'+c'$ dove $\underline{b}_1$ è la proiezione ortogonale di $\underline{b}$ su Col(A), appartiene a ker A e quindi è un autovettore di A associato all'autovalore 0 e c'=

$\underline{w}^T A\underline{w}+2\underline{b}^T\underline{w}+c$. Ora la matrice reale simmetrica A può essere diagonalizzata da una qualsiasi matrice le cui colonne sono una base ortonormale di autovettori di A. Tale base si ottiene come unione delle basi ortonormali degli autospazi associati ai vari autovalori di A. Avremo quindi una base ortonormale composta da $\underline{v}_1,\underline{v}_2,\dots,\underline{v}_r$ che sono autovettori di norma 1 a due a due ortogonali appartenenti agli autospazi associati agli autovalori non nulli $\lambda_1,\lambda_2,\dots,\lambda_r$ di A (che possono essere non tutti distinti) e da n-r vettori che sono una base ortonormale dell'autospazio associato all'autovalore 0 (che coincide con ker A). Usando l'algoritmo di Gram-Schmidt costruiamo una base ortonormale $\underline{v}_{r+1}, \underline{v}_{r+2},\dots,\underline{v}_n$ di ker A usando come primo vettore $\underline{v}_{r+1}=\frac{\underline{b}_1}{\|\underline{b}_1\|}$ Sia U la matrice, che (come già notato) possiamo pensare di determinante 1, formata dall'accostamento dei vettori ortonormali $\underline{v}_1,\underline{v}_2,\dots,\underline{v}_n$. La isometria lineare $\underline{x}'=U\underline{X}'$ agisce sul polinomio $p_1(\underline{x}')$ portandolo nella forma $p_2(\underline{X}')= p_1(U\underline{x}') = (\underline{X}')^T U^T A U\underline{X}'+2\underline{b}_1^T U\underline{X}'+c'$ dove $U^T AU=\mathrm{diag}(\lambda_1,\lambda_2,\dots,\lambda_r, 0,\dots,0)$. $\underline{b}_1^T U$ è un vettore riga la cui componente i-esima si ottiene moltiplicando $\underline{b}_1^T$ per la i-esima colonna di U. Tutte le colonne di U (eccetto la r+1 esima) sono ortogonali a $\underline{b}_1$ e quindi turtte le componenti i-esime con $i\neq r+1$ del vettore $\underline{b}_1^T U$ sono nulle. La r+1 esima componente è invece $\underline{b}_1^T \frac{\underline{b}_1}{\|\underline{b}_1\|} =\|\underline{b}_1\|$. Da qui si ha $p_2(\underline{X}')= (\underline{X}')^T\mathrm{diag}(\lambda_1,\dots,\lambda_r ,0,\dots,0)\underline{X}'+ +2\|\underline{b}_1\|X'_{r+1} + c'$. Facendo la traslazione $X_i=X'_i$ per ogni $i\neq r+1$, e $X'_{r+1}= X_{r+1}- \|b1\|/2$ si ha $\overline{p}(\underline{X})= \lambda_1 X_1^2+\lambda_2 X_2^2+\dots+\lambda_r X_r^2+2pX_{r+1}$ con $p\neq 0$ ed abbiamo quindi trovato il terzo caso del secondo punto del teorema. L'equazione della quadrica in forma compatta diventa $\underline{Z}^T B'\underline{Z}=0$ dove la sottomatrice formata dalle prime r+2 righe e colonne di B' è $\left[\begin{array}{c|c|c} U^T AU & \underline{0} & \underline{0} \\ \hline \underline{0}^T & 0 & p \\ \hline \underline{0}^T & p & 0 \end{array}\right]$ e le restanti n-(r+2) righe e colonne sono tutte nulle. Pertanto rk B=rk B'= rk $(U^T AU)+2=$rk A+2.