Autovalori, Autovettori, Diagonalizzazione.

1. Autovalori e Autovettori

Definizione 1.1 Sia V uno spazio vettoriale sul campo $\mathbb{K}\,(=\mathbb{R}\,\mathrm{o}\,\mathbb{C})$ e sia $T\colon V\to V$ un endomorfismo. Un vettore non nullo $v\in V\setminus\{O\}$ si dice *autovettore* di T relativo all'autovalore $\lambda\in\mathbb{K}$ se si ha

$$T(v)=\lambda v \tag{1.1}$$

Proposizione 1.1: *Sia V uno spazio vettoriale sul campo $\mathbb{K}\,(=\mathbb{R}\,\mathrm{o}\,\mathbb{C})$ e sia $T\colon V\to V$ un endomorfismo. Allora $\lambda\in\mathbb{K}$ è autovalore di $T \iff Ker(T-\lambda Id_V)\neq\{O\}$. Inoltre $v\in V$ è una autovettore di T relativo all'autovalore $\lambda\in\mathbb{K}\iff v\in Ker(T-\lambda Id_V)\setminus\{O\}$.*

Dimostrazione: Basta osservare che (1.1) equivale a $(T-\lambda Id_V)(v)=O$. □

Definizione 1.2 Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $T\colon V\to V$ un endomorfismo. Se $\lambda\in\mathbb{K}$ è un autovalore di T, il sottospazio $V_\lambda=Ker(T-\lambda Id_V)$ si dice *autospazio* di T relativo all'autovalore λ e la sua dimensione $dim(V_\lambda)$ si dice *molteplicità geometrica* di λ.

Proposizione 1.2: *Sia $T\colon V\to V$ un endomorfismo di uno spazio vettoriale V e sia $\mathcal{B}$ una base di V. Allora la matrice di T relativa a $\mathcal{B}$ è diagonale $\iff \mathcal{B}$ è composta di autovettori di T.*

Dimostrazione: Sia $\mathcal{B}=\{v_1,\dots,v_n\}$. Se i v_i sono tutti autovettori relativi a autovalori λ_i, allora $T(v_i)=\lambda_i v_i$ per ogni i e quindi la matrice di T relativa a $\mathcal{B}$ è diagonale con i $\lambda_1,\dots,\lambda_n$ sulla diagonale. Viceversa se la matrice di T relativa a $\mathcal{B}$ è diagonale con i $\lambda_1,\dots,\lambda_n$ sulla diagonale, allora $T(v_i)=\lambda_i v_i$ per ogni i e quindi $\mathcal{B}$ è composta di autovettori di T. □

Corollario 1.3: *Una matrice quadrata A è simile a una matrice diagonale $\iff$ l'applicazione $L_A\colon\mathbb{K}^n\to\mathbb{K}^n$ ammette una base di autovettori.*

Definizione 1.3 Un endomorfismo $T\colon V\to V$ di uno spazio vettoriale V si dice *diagonalizzabile* se esiste una base di V rispetto alla quale T ha matrice diagonale. Questo equivale a richiedere che esiste una base di autovettori di V. Una matrice A quadrata di ordine n si dice *diagonalizzabile* se è simile a una matrice diagonale, ossia se esiste una matrice non singolare B tale che $B^{-1}AB$ è diagonale.

Teorema 1.4: *Sia $T\colon V\to V$ un endomorfismo di uno spazio vettoriale V di dimensione n sul campo $\mathbb{K}$. Siano $\mathcal{B}$ una base di V e sia A la matrice di T relativa a $\mathcal{B}$. Allora*

(i) La funzione $p_T\colon\mathbb{K}\to\mathbb{K}$ definita da

$$p_T(\lambda)=det(A-\lambda I_n)$$

è un polinomio di grado n che non dipende dalla scelta della base $\mathcal{B}$.

(ii) Uno scalare $\lambda_0\in\mathbb{K}$ è un autovalore di $T\iff p_T(\lambda_0)=0$.

Dimostrazione: (i) Dimostriamo che $p_T(\lambda)$ non dipende dalla scelta della base. Sia A' la matrice di T relativa a un'altra scelta di base $\mathcal{B}'$ di V. Se B è la matrice del cambio di base da $\mathcal{B}$ a $\mathcal{B}'$ allora $A'=B^{-1}AB$. Occorre provare che $det(A'-\lambda I_n)=det(A-\lambda I_n)$. Questo segue utilizzando il Teorema di Binet:

$$\begin{aligned}det(A'-\lambda I_n)&=det(B^{-1}AB-\lambda I_n)=det(B^{-1}AB-B^{-1}\lambda I_nB)\\&=det\left(B^{-1}(A-\lambda I_n)B\right)=detB^{-1}det(A-\lambda I_n)detB\\&=(detB)^{-1}det(A-\lambda I_n)detB=det(A-\lambda I_n).\end{aligned}$$

Per dimostrare che $p_T(\lambda)$ è un polinomio di grado n si può utilizzare la formula di Leibnitz per il determinante di una matrice B quadrata di ordine n:

$$det(B)=\sum_{\sigma\in S_n}\epsilon(\sigma)b_{1\sigma(1)}b_{2\sigma(2)}\dots b_{n\sigma(n)}.$$

Per $B = A - \lambda I_n$ segue immediatamente che $p_T(\lambda)$ è un polinomio di grado n in λ e la potenza massima λ^n si ottiene moltiplicando i termini sulla diagonale.

In alternativa si può procedere per induzione sulla dimensione n di V. Se $n = 1$ il fatto è ovvio. Supponiamo che sia vero per $n-1$. Allora, utilizzando la formula di Laplace per la prima riga (o per la prima colonna), si osserva che

$$\begin{aligned} p_T(\lambda) = det(A - \lambda I_n) = det \begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n1}-\lambda \end{pmatrix} \\ = (a_{11}-\lambda) det \begin{pmatrix} a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{n1}-\lambda \end{pmatrix} + \text{ monomi di grado } \leq n-2 \text{ in } \lambda \end{aligned}.$$

Quindi per l'ipotesi induttiva, $p_T(\lambda)$ è un polinomio di grado n.

(ii) Lo scalare $\lambda_0 \in \mathbb{K}$ è autovalore di $T \iff Ker(T - \lambda_0 Id_V) \neq \{O\} \iff (A - \lambda_0 I_n)$ non è invertibile $\iff det(A - \lambda_0 I_n) = 0$. □

Definizione 1.4 Sia $T\colon V \to V$ un endomorfismo di uno spazio vettoriale V. Il polinomio $p_T(\lambda)$ si dice *polinomio caratteristico* di T.

Osservazione. Calcolando per esempio con la formula di Leibnitz il polinomio caratteristico e osservando che, se $A = (a_{ij}) \in M_n$ è una matrice associata a un endomorfismo T si ha $p_T(0) = det(A)$, si ottiene:

$$\begin{aligned} p_T(\lambda) &= (-1)^n \lambda^n + (-1)^{n-1}(a_{11} + \ldots + a_{nn})\lambda^{n-1} + \ldots + det(A) \\ &= (-1)^n \lambda^n + (-1)^{n-1} Traccia(A) \lambda^{n-1} + \ldots + det(A). \end{aligned}$$

Definizione 1.5 Sia λ_0 un autovalore di un endomorfismo di uno spazio vettoriale V la *molteplicità algebrica* di T è la molteplicità di λ_0 come radice del polinomio $p_T(\lambda)$

Come conseguenza immediata del Teorema 1.4 e del Teorema Fondamentale dell'Algebra si ha il seguente

Corollario 1.5: *Ogni endomorfismo di uno spazio vettoriale su $\mathbb{C}$ di dimensione n ha esattamente n autovalori (contati con la molteplicità).*

Spesso invece che con endomorfismi, si ha a che fare solo con matrici. Ricapitoliamo brevemente le ovvie nozioni di autovalore, di autovettore, autospazio e molteplicità nel caso di matrici:

Definizione 1.6 Se $A \in M_n(\mathbb{K})$ è una matrice quadrata di ordine n. Un vettore non nullo $x \in \mathbb{K}^n \setminus \{O\}$ si dice *autovettore* di A relativo all'*autovalore* $\lambda_0 \in \mathbb{K}$ se si ha $Ax = \lambda_0 x$. Inoltre $p_A(\lambda) = det(A - \lambda I_n)$ (che secondo la Definizione (1.4) è il polinomio caratteristico dell'applicazione L_A) si dice *polinomio caratteristico* di A e le sue radici sono gli autovalori di A. Si dice *autospazio* relativo a un autovalore λ_0 di A il sottospazio $V_{\lambda_0} = Ker(A - \lambda_0 I_n) = Ker(L_A - \lambda_0 Id_{\mathbb{K}^n})$. La *molteplicità geometrica* di un autovalore λ_0 di A è la dimensione dell'autospazio V_{λ_0}. La *molteplicità algebrica* di un autovalore è la sua molteplicità come radice del polinomio caratteristica.

2. Diagonalizzabilità

Nel paragrafo precedente abbiamo dato due nozioni di molteplicità per un autovalore di un endomorfismo di uno spazio vettoriale di dimensione finita. È possibile dare un criterio di diagonalizzabilità in termini di molteplicità. Si dimostra infatti che affinchè un endomorfismo di uno spazio di dimensione n su un campo $\mathbb{K}$ sia diagonalizzabile occorre e basta che, contati con la loro molteplicità algebrica, ci siano n autovalori nel campo $\mathbb{K}$ e che per tutti gli autovalori le molteplicità algebriche e geometriche coincidano. L'idea è che affinchè un endomorfismo sia diagonalizzabile debbano esserci “abbastanza” autovalori e che per ogni autovalore ci siano “abbastanza” autovettori linearmente indipendenti. Per capire cosa succede cominciamo con due esempi tipic di endomorfimi (e matrici) non diagonalizzabili. Il primo illustra la situazione in cui non ci sono “abbastanza” autovalori, il secondo il caso in cui non ci sono “abbastanza” autovettori linearmente indipendenti.

Esempio 1. Sia $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ e si consideri $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ definita come al solito da $L_A(x) = Ax$. Il polinomio caratteristico di A è $\lambda^2 + 1$ e quindi non vi sono autovalori in $\mathbb{R}$ di A. Si conclude allora che $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ non è diagonalizzabile e che A non è simile a una matrice diagonale reale ossia non esiste una matrice invertibile B reale tale che $B^{-1}AB$ sia diagonale.

Esempio 2. Sia $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ e si consideri $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ definita come al solito da $L_A(x) = Ax$. Dato che il polinomio caratteristico di di L_A (e di A) è $(\lambda - 1)^2$ Si vede subito che l'unico autovalore di L_A (e di A) è 1. D'altra parte l'autospazio corrispondente è $\{x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1 = 0\}$ che ha dimensione 1. Dunque non può esistere una base di $\mathbb{R}^2$ costituita da autovettori di L_A (o di A) e quindi L_A (e A) non sono diagonalizzabili.

Cominciamo con la seguente

Proposizione 2.1: *Sia $T\colon V \to V$ un endomorfismo di uno spazio vettoriale V su un campo $\mathbb{K}$. Se $v_1, \ldots, v_k$ sono autovettori di T corrispondenti a autovalori distinti $\lambda_1, \ldots, \lambda_k$, allora $v_1, \ldots, v_k$ sono linearmente indipendenti.*

Dimostrazione: Procediamo per induzione su k. Se $k = 1$, il risultato è banale. Suppo-niamo il risultato vero per $k - 1$. Siano $\alpha_1, \ldots, \alpha_k$ scalari tali che

$$O = \alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1} \alpha_k v_k. \tag{2.1}$$

Allora

$$\begin{aligned} O &= T(\alpha_1 v_1 + \ldots + \alpha_k v_k) = \alpha_1 T(v_1) + \ldots + \alpha_k T(v_k) \\ &= \alpha_1 \lambda_1 v_1 + \ldots + \alpha_{k-1} \lambda_{k-1} v_{k-1} + \alpha_k \lambda_k v_k \end{aligned} \tag{2.2}$$

Moltiplicando la (2.1) per λ_k e sottraendo alla (2.2) otteniamo:

$$\begin{aligned} O &= \alpha_1(\lambda_1 - \lambda_k) v_1 + \ldots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} + \alpha_k(\lambda_k - \lambda_k) v_k \\ &= \alpha_1(\lambda_1 - \lambda_k) v_1 + \ldots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} \end{aligned}.$$

Per l'ipotesi induttiva $\alpha_1(\lambda_1 - \lambda_k) = \ldots = \alpha_{k-1}(\lambda_{k-1} - \lambda_k) = 0$ e quindi $\alpha_1 = \ldots = \alpha_{k-1} = 0$, dato che $\lambda_i \neq \lambda_j$ se $i \neq j$. Ma allora da (2.1) segue anche $\alpha_k = 0$. □

Proposizione 2.2: *Sia V spazio vettoriale di dimensione n su un campo $\mathbb{K}$. Se un endomorfismo $T\colon V \to V$ ha n autovalori distinti in $\mathbb{K}$, T è diagonalizzabile.*

Dimostrazione: Il risultato è immediata conseguenza della Proposizione 2.1: se $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ sono autovalori a due a due distinti di T e $v_1, \ldots, v_n$ sono autovettori corrispondenti, allora $v_1, \ldots, v_n$ sono linearmente indipendenti e quindi, dato che $dim(V) = n$, formano una base per V. □

Proposizione 2.3: *Siano V spazio vettoriale di dimensione n su un campo $\mathbb{K}$ e $T\colon V \to V$ un endomorfismo. Per ogni autovalore $\lambda_0 \in \mathbb{K}$ di T la molteplicità geometrica è minore o uguale della molteplicità algebrica.*

Dimostrazione: Siano r la molteplicità geometrica e s la molteplicità algebrica di λ_0. Sia $\{v_1, \ldots, v_r\}$ una base dell'autospazio V_{λ_0} e sia $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ un suo completamento a una base di V. La matrice A di T relativa a questa base ha questa forma:

$$A = \left(\begin{array}{cccc|cccc} \lambda_0 & 0 & \ldots & 0 & & & & \\ 0 & \lambda_0 & \ldots & 0 & & & & \\ \vdots & \vdots & \ddots & \vdots & & C & & \\ 0 & 0 & \ldots & \lambda_0 & & & & \\ -- & -- & -- & -- & -- & -- & -- & -- \\ 0 & 0 & \ldots & 0 & & & & \\ 0 & 0 & \ldots & 0 & & B & & \\ 0 & 0 & \ldots & 0 & & & & \end{array} \right)$$

dopo C è una matrice r righe e $n - r$ colonne e B è una matrice $n - r$ righe e $n - r$ colonne. Allora se P_T, P_A e P_B sono rispettivamente il polinomio caratteristico di T, della matrice A e della matrice B, si ha

$$P_T(\lambda) = P_A(\lambda) = (\lambda - \lambda_0)^r P_B(\lambda)$$

e quindi la moteplicità algebrica di λ_0 è $s \geq r$. □

Siamo pronti per enunciare il criterio di diagonalizzabilità annunciato:

Teorema 2.4: *Sia $T: V \to V$ un endomorfismo di uno spazio vettoriale V su un campo $\mathbb{K}$ di dimensione n. Allora T è diagonalizzabile se e solo se T ha n autovalori (contati con la molteplicità algebrica) in $\mathbb{K}$ e per ogni autovalore la molteplicità geometrica coincide con la molteplicità algebrica.*

Dimostrazione: Se T è diagonalizzabile allora esiste una base di autovalori, la matrice di T relativa è diagonale con sulla diagonale gli autovalori di T. Quindi è immediato osservare che T ha n autovalori (contati con la molteplicità algebrica) in $\mathbb{K}$ e per ogni autovalore la molteplicità geometrica coincide con la molteplicità algebrica.

Viceversa siano $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ gli autovalori distinti di T, $m_1, \ldots, m_r$ le rispettive molteplicità algebriche e $V_{\lambda_1}, \ldots, V_{\lambda_r}$ i rispettivi autospazi. Supponiamo che $m_1 + \ldots + m_r = n$ e che $dim(V_{\lambda_j}) = m_j$ per ogni $j = 1, \ldots, r$. Poniamo inoltre $\mu_j = \sum_{k=1}^{j} m_k$ per ogni $j \geq 1$ in modo che in particolare risulti $\mu_r = n$. Siano rispettivamente

$$\mathcal{B}_1 = \{v_1, \ldots, v_{\mu_1}\} \quad \text{una base di} \quad V_{\lambda_1},$$
$$\mathcal{B}_2 = \{v_{\mu_1+1}, \ldots, v_{\mu_2}\} \quad \text{una base di} \quad V_{\lambda_2},$$
$$\vdots$$
$$\mathcal{B}_r = \{v_{\mu_{r-1}+1}, \ldots, v_n\} \quad \text{una base di} \quad V_{\lambda_r}.$$

Allora $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$ è costituita da n autovettori di T. Per concludere la dimostrazione, basta provare che $\mathcal{B}$ è una base di V e a tal fine basta far vedere che è un sistema libero. Supponiamo che per $\alpha_1, \ldots, \alpha_{\mu_1}, \ldots, \alpha_{\mu_{r-1}+1}, \ldots, \alpha_{\mu_n} \in \mathbb{K}$ si abbia

$$\alpha_1 v_1 + \ldots + \alpha_{\mu_1} v_{\mu_1} + \ldots + \alpha_{\mu_{r-1}+1} v_{\mu_{r-1}+1} + \ldots + \alpha_{\mu_n} v_{\mu_r} = O_V$$

e per ogni $j = 1, \ldots, r$ si ponga $w_j = \alpha_{\mu_{j-1}+1} v_{\mu_{j-1}+1} + \ldots + \alpha_{\mu_j} v_{\mu_j} \in V_{\lambda_j}$. Dunque si ha

$$w_1 + \ldots + w_r = O_V$$

e questo, dato che i w_j appartengono a autospazi corrispondenti a autovalori distinti, per la Proposizione 2.1 può succedere solo se $w_1 = \ldots = w_r = O_V$. Allora per ogni $j = 1, \ldots, r$ si ha $O_V = w_j = \alpha_{\mu_{j-1}+1} v_{\mu_{j-1}+1} + \ldots + \alpha_{\mu_j} v_{\mu_j}$ e quindi, dato che $\mathcal{B}_j$ è una base, allora $\alpha_{\mu_{j-1}+1} = \ldots = \alpha_{\mu_j} = 0$. Dunque $\mathcal{B}$ è libero.

$\square$

Il Teorema 2.4 si può riformulare come criterio di diagonalizzabilità per matrici nel modo seguente:

Corollario 2.5: *Una matrice $A \in M_n(\mathbb{K})$ è simile a una matrice diagonale se e solo se A ha n autovalori (contati con la molteplicità algebrica) in $\mathbb{K}$ e per ogni autovalore la molteplicità geometrica coincide con la molteplicità algebrica.*

Concludiamo illustrando la procedura che si deve seguire per verificare se un endomorfismo è diagonalizzabile e nel caso per trovare una base composta di autovettori.

Algoritmo di diagonalizzazione per endomorfismi: Sia $T: V \to V$ un endomorfismo di uno spazio vettoriale V su $\mathbb{K}$ con $dim(V) = n$

I passo: Si fissi una base $\mathcal{B}$ di V e si trovi la matrice A associata a T relativa alla base $\mathcal{B}$.

II passo: Trovare il polinomio caratteristico $P_T(\lambda) = det(A - \lambda I_n)$ e determinarne le radici. Se, contate con la loro molteplicità, ha n radici in $\mathbb{K}$ si può proseguire, altrimenti si conclude che T non è diagonalizzabile (perchè non ha abbastanza autovalori in $\mathbb{K}$!).

III passo: Se $\lambda_1, \ldots, \lambda_r$ sono gli autovalori distinti di T, per ogni $j = 1, \ldots, r$ si determnini una base $\mathcal{C}_j$ per l'autospazio V_{λ_j} relativo all'autovalore λ_j.

IV passo: Si consideri $\mathcal{C} = \mathcal{C}_1 \cup \ldots \cup \mathcal{C}_r$. Allora $\mathcal{C}$ è un sistema di vettori linearmente indipendenti di V. L'endomorfismo T è diagonalizzabile se e solo se $\mathcal{C}$ contiene n vettori perchè in questo caso $\mathcal{C}$ è una base di V (contiene il numero giusto di vettori linearmente indipendenti) costituita da autovettori di T. Se $\mathcal{C}$ è una base, la matrice di T relativa a $\mathcal{C}$ è la matrice diagonale Δ con sulla diagonale gli autovalori di T. La matrice B del cambio di base dalla base $\mathcal{B}$ alla base $\mathcal{C}$ (quella che ha per colonne le coordinate rispetto a $\mathcal{B}$ dei vettori di $\mathcal{C}$, è la matrice tale che $\Delta = B^{-1}AB$.

Naturalmente l'algoritmo si può adattare per diagonalizzare matrici.

Algoritmo di diagonalizzazione matrici: Sia $A \in M_n(\mathbb{K})$ una matrice quadrata di ordine n.

I passo: Trovare il polinomio caratteristico $P_A(\lambda) = det(A - \lambda I_n)$ e determinarne le radici. Se, contate con la loro molteplicità, ha n radici in $\mathbb{K}$ si può proseguire, altrimenti si conclude che A non è diagonalizzabile (perché non ha abbastanza autovalori in $\mathbb{K}$!).

II passo: Se $\lambda_1, \dots, \lambda_r$ sono gli autovalori distinti di A, per ogni $j = 1, \dots, r$ si determnini una base $\mathcal{C}_j$ per l'autospazio $V_{\lambda_j} = Ker(A - \lambda_j I_n)$ relativo all'autovalore λ_j (che non è altro che l'insieme delle soluzioni del sistema omogeneo $(A - \lambda_j I_n)x = 0$).

III passo: Si consideri $\mathcal{C} = \mathcal{C}_1 \cup \ldots \cup \mathcal{C}_r$. Allora $\mathcal{C}$ è un sistema di vettori linearmente indipendenti di R^n. La matrice A è diagonalizzabile se e solo se $\mathcal{C}$ contiene n vettori perchè in questo caso $\mathcal{C}$ è una base di R^n (contiene il numero giusto di vettori linearmente indipendenti) costituita da autovettori di A. Se $\mathcal{C}$ è una base, la matrice di L_A relativa a $\mathcal{C}$ è la matrice diagonale Δ con sulla diagonale gli autovalori di A. Dunque se B è la matrice del cambio di base dalla base canonica alla base $\mathcal{C}$ (quella che ha per colonne i vettori di $\mathcal{C}$, è la matrice tale che $\Delta = B^{-1}AB$.

Esempio. Se $M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ e $N = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, consideriamo l'endomorfismo $T \colon M_2(\mathbb{R}) \to M_2(\mathbb{R})$ definito da $T(X) = XM - NX$. Dimostreremo ora che T è diagonalizzabile e troveremo una base diagonalizzante. Se $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$, allora si vede subito che

$$T(X) = \begin{pmatrix} 2x_{11} + x_{12} & x_{11} + 2x_{12} \\ x_{21} + x_{22} & x_{21} + x_{22} \end{pmatrix}.$$

Allora rispetto alla base

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$$

di $M_2(\mathbb{R})$, la matrice di T è

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Il polinomio caratteristico di T è dunque $p_T(\lambda) = det(A - \lambda I_4) = (\lambda^2 - 4\lambda + 3)(\lambda^2 - 2\lambda)$ che ha 4 radici distinte $0, 1, 2, 3$ che sono gli autovalori di T (e di A). Cercare una base per gli autospazi corrispondenti a ciascun autovalore corrisponde a trovare basi rispettivamente per

$$KerT, Ker(T - Id), Ker(T - 2Id), Ker(T - 2Id).$$

Dato che

$$KerA = Span \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad Ker(A - I_4) = Span \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix},$$

$$Ker(A - 2I_4) = Span \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad Ker(A - 3I_4) = Span \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

allora

$$KerT = Span \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix}, \quad Ker(T - Id) = Span \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

$$Ker(T - 2Id) = Span \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad Ker(T - 2Id) = Span \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

e quindi una base di $M_2(\mathbb{R})$ rispetto alla quale T ha matrice diagonale è data da

$$\left\{ \begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right\}.$$

La matrice di T rispetto a questa base è

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Si osservi che la matrice

$$B = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

che ha per colonne i vettori che formano basi rispettivamente di $KerA$, $Ker(A - I_4)$, $Ker(A - 2I_4)$, $Ker(A - 3I_4)$ ha inversa

$$B^{-1} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix}$$

e con un semplice calcolo si verifica che $D = B^{-1}AB$.

Concludiamo queste note esaminando il seguente problema. Sia V uno spazio vettoriale di dimensione n e siano $T_i\colon V \to V$, $i = 1, 2$ due endomorfismi diagonalizzabili. Sotto quale condizione esiste una base di V rispetto alla quale sia T_1 sia T_2 ha matrice diagonale? In altre parole sotto quale ipotesi T_1 e T_2 sono "simultaneamente diagonalizzabili'?'

È abbastanza facile trovare una condizione necessaria:

Proposizione 2.6: *Per $i = 1, 2$ siano $T_i\colon V \to V$ endomorfismi di uno spazio vettoriale V su un campo $\mathbb{K}$ di dimensione n simultaneamente diagonalizzabili. Allora T_1 e T_2 commutano ossia sono tali che $T_1 \circ T_2 = T_2 \circ T_1$*

Dimostrazione: Sia $\mathcal{B} = \{v_1, \ldots, v_1\}$ una base rispetto alla quale sia T_1 sia T_2 hanno matrice diagonale. Allora per ogni $i = 1, \ldots, n$ esistono scalari λ_i e μ_i tali che $T_1(v_i) = \lambda_i v_i$ e $T_2(v_i) = \mu_i v_i$. Ma allora per ogni $i = 1, \ldots, n$

$$T_1(T_2((v_i)) = T_1(\mu_i v_i) = \mu_i T_1(v_i) = \mu_i \lambda_i v_i = \lambda_i \mu_i v_i = \lambda_i T_2(v_i) = T_2(\lambda_i v_i) = T_2(T_1(v_i))$$

e quindi $T_1 \circ T_2((v_i) = T_2 \circ T_1((v_i)$. Dato che assumono gli stessi valori sugli elementi di una base, segue che $T_1 \circ T_2 = T_2 \circ T_1$ □

In effetti dimostreremo ora che la condizione è anche sufficiente. Cominciamo con una ossevazione che vale anche per spazi non finitamente generati.

Lemma 2.7: *Per $i = 1, 2$, siano $T_i\colon V \to V$ endomorfismi di uno spazio vettoriale V su un campo $\mathbb{K}$ che commutano ossia tali che $T_1 \circ T_2 = T_2 \circ T_1$. Se U è un autospazio di T_1 allora $T_2(U) \subset U$ e se W è un autospazio di T_2 allora $T_1(W) \subset W$.*

Dimostrazione: Naturalmente basta dimostrare una delle due conclusioni, l'altra si ottiene scambiando i ruoli di T_1 e T_2. Sia dunque U un autospazio di T_1 relativo all'autovalore λ. Allora $U = Ker(T_1 - \lambda Id)$. Se $u \in U$, allora $T_1(T_2((u)) = T_2(T_1(u)) = T_2(\lambda u) = \lambda T_2(u)$ e quindi $T_2(u) \in U$. □

Il risultato conclusivo è il seguente:

Teorema 2.8: *Per $i = 1, 2$, siano $T_i\colon V \to V$ endomorfismi di uno spazio vettoriale V su un campo $\mathbb{K}$ di dimensione n. Allora T_1 e T_2 sono simultaneamente diagonalizzabili se e solo se commutano.*

Dimostrazione: Grazie alla Proposizione 2.6 ci basta dimostrare che se T_1 e T_2 commutano allora sono simultaneamente diagonalizzabili. Supponiamo prima di tutto che per uno fra T_1 e T_2, per esempio T_1, tutto V sia un autospazio. Allora $T_1 = \lambda Id$ per qualche scalare λ e quindi T_1 ha matrice diagonale rispetto qualunque base di V e quindi T_1 e T_2 sono simultaneamente diagonalizzabili. Se invece gli autospazi di T_1 e T_2 sono tutti propri (ossia diversi da V), procediamo per induzione per ottenere la conclusione. La conclusione del Teorema vale banalmente per spazi vettoriai di dimensione 1. Supponiamo che il Teorema valga per ogni spazio vettoriale di dimensione $\leq n - 1$. Siano $\lambda_1, \ldots, \lambda_k$ gli autovalori distinti di T_1 e $V_1, \ldots, V_k$ i corrispondenti autospazi. Se $U = V_2 \oplus \ldots \oplus V_k$ allora, sia V_1 sia U sono sottospazi propri di V e quindi $dim(V_1) \leq n - 1$ e $dim(U) \leq n - 1$. Inoltre dal Lemma 2.7 segue che $T_2(V_1) \subset V_1$ e $T_2(U) \subset U$. Dunque le restrizioni di T_1 e T_2 sia a V_1 sia a U commutano e quindi, per l'ipotesi induttiva, sono simultaneamente diagonalizzabili. Dato che $V = V_1 \oplus U$, segue che T_1 e T_2 sono simultaneamente diagonalizzabili su tutto V.

□

2. Diagonalizzazione ortogonale di endomorfismi autoaggiunti

In questo paragrafo dimostreremo che ogni endomorfismo autoaggiunto di uno spazio vettoriale metrico è diagonalizzabile e che la base rispetto alla quale ha matrice diagonale si può scegliere ortonormale. Cominciamo con la seguente cruciale osservazione:

Teorema 2.1: *Sia $T: V \to V$ un endomorfismo autoaggiunto di uno spazio vettoriale metrico V di dimensione n su $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Allora T ha n autovalori (contati con la molteplicità) e sono tutti in $\mathbb{R}$.*

Dimostrazione: Rispetto a una base ortonormale T ha una matrice A simmetrica nel caso reale, hermitiana nel caso complesso. Gli autovalori di T sono esattamente le radici del suo polinomio caratteristico di A che sono nel campo $\mathbb{K}$. Per il teorema fondamentale dell'algebra, il polinomio caratteristico di A ha esattamente n radici (contate con la molteplicità) in $\mathbb{C}$. Dobbiamo allora solo dimostrare che le radici del polinomio caratteristico di A sono reali. Si consideri $L = L_A: \mathbb{C}^n \to \mathbb{C}^n$. Dato che $A = A^H$ (sia nel caso reale, sia nel caso complesso), allora L è un endomorfismo autoaggiunto di $\mathbb{C}^n$ con il prodotto hermitiano canonico. Sia λ una radice del polinomio caratteristico di A. Allora λ è un autovalore di L. Sia $u \in \mathbb{C}^n \setminus \{0\}$ tale che $L(u) = \lambda u$. Si ha

$$\lambda \|u\|^2 = \lambda < u, u >=< \lambda u, u >=< Lu, u >=< u, Lu >= \overline{\lambda} < u, u >= \overline{\lambda} \|u\|^2$$

e quindi $(\lambda - \overline{\lambda})\|u\|^2 = 0$. Dato che $\|u\|^2 \neq 0$, allora deve essere $\lambda - \overline{\lambda} = 0$ ossia $\lambda \in \mathbb{R}$. $\square$

Il risultato principale che vogliamo dimostrare è il seguente

Teorema 2.2: (Teorema spettrale) *Sia V uno spazio vettoriale metrico di dimensione n e sia $T: V \to V$ un endomorfismo. Le seguenti affermazioni sono equivalenti:*

(i) T è autoaggiunto.

(ii) V ha una base ortonormale formata da autovettori di T ciascuno con autovalore reale rispetto alla quale, quindi, T ha matrice diagonale reale.

(iii) Esiste una base ortonormale $\{v_1, \ldots, v_n\}$ di V e numeri reali $\lambda_1, \ldots, \lambda_n$ tali che se $u = \sum_{i=1}^n x_i v_i$ allora $T(u) = \sum_{i=1}^n \lambda_i x_i v_i$.

Dimostrazione: $(i) \iff (ii)$: Se vale (ii) allora T ha matrice diagonale reale rispetto a una base ortonormale. Dato che una matrice diagonale reale coincide con la sua trasposta e con la sua aggiunta, necessariamente $T = T^*$. Supponiamo viceversa che valga (i). Procederemo per induzione sulla dimensione n di V. Se $n = 1$ allora (ii) è banale. Supponiamo che se T è autoaggiunta valga (ii) per spazi di dimensione $n-1$. Per il Teorema 2.1, esiste un autovalore $\lambda \in \mathbb{R}$ di T. Sia v_1 un autovettore con autovalore λ e con $\|v_1\| = 1$. Sia $W = (Span(v_1))^\perp$. Allora W è un sottospazio di V di dimensione $n-1$ e, per costruzione, si ha che $w \in W$ se e solo se $< w, v_1 >= 0$ e quindi $W = \{w \in V | < w, v_1 >= 0\}$.

Si ha inoltre $T(W) \subset W$. Infatti $u \in T(W) \Leftrightarrow u = T(w)$ per qualche $w \in W$. Allora

$$< u, v_1 >=< T(w), v_1 >=< w, T(v_1) >=< w, \lambda v_1 >= \overline{\lambda} < w, v_1 >= 0$$

e quindi $u \in W$. Il prodotto definito su V, ristretto a W, definisce una struttura di spazio vettoriale metrico per W. Dato che $T(W) \subset W$ allora $T: W \to W$ è un endomorfismo autaggiunto di uno spazio vettoriale metrico di dimensione $n-1$. Esiste allora, per l'ipotesi induttiva, una base ortonormale $\{v_2, \ldots, v_n\}$ formata da autovettori di di T ciascuno con autovalore reale. Dato che $< v_j, v_1 >= 0$ per $j = 2, \ldots, n$, allora $\{v_1, v_2, \ldots, v_n\}$ è una base ortonormale di V formata da autovettori di di T ciascuno con autovalore reale e quindi (ii) vale.

$(ii) \iff (iii)$: Se vale (iii) allora si ha $T(v_i) = \lambda_i v_i$ e quindi vale (ii). D'altro canto se $\{v_1, \ldots, v_n\}$ è una base ortonormale formata da autovettori di T ciascuno con autovalore reale λ_i corrispondente a v_i allora la (iii) segue immediatamente per la linearità di T.

$\square$

Conseguenza immediata del Teorema 2.2 è il seguente

Corollario 2.3: (Teorema spettrale reale) *Sia $T: V \to V$ un endomorfismo di uno spazio vettoriale metrico reale. Esiste una base ortonormale formata da autovettori di T se e solo se T è autoaggiunto.*

In termini di matrici il Teorema 2.2 si traduce in questo modo:

Teorema 2.4: *(i) Sia $A \in M_{n,n}(\mathbb{R})$ una matrice simmetrica reale. Allora esiste una matrice ortogonale $B \in \mathcal{O}(n)$ tale che $D = B^{-1}AB = B^TAB$ è una matrice diagonale.*

(ii) Sia $A \in M_{n,n}(\mathbb{C})$ una matrice hermitiana. Allora esiste una matrice unitaria $B \in \mathcal{U}(n)$ tale che $D = B^{-1}AB = B^HAB$ è una matrice diagonale reale.

Dimostrazione: (i) Si consideri $L_A: \mathbb{R}^n \to \mathbb{R}^n$. Allora L_A è un endomorfismo autoaggiunto di $\mathbb{R}^n$ con il prodotto scalare canonico. Per il Teorema 2.2 esiste una base ortonormale $\mathcal{B}$ rispetto alla quale L_A ha matrice diagonale. La matrice B del cambio di base dalla base canonica alla base $\mathcal{B}$, che ha per colonne le coordinate dei vettori di $\mathcal{B}$, è la matrice cercata.

(ii) La dimostrazione, analoga al caso reale, è lasciata per esercizio. □

Vogliamo illustrare ora come in pratica si possa diagonalizzare un endomorfismo autoaggiunto. Il teorema spettrale ci assicura che un endomorfismo autoaggiunto è sempre diagonalizzabile rispetto a una base ortogonale. Anche il procedimento per ottenere una base diagonalizzante si semplifica perché autovettori di un endomorfismi autoaggiunti relativi a autovalori distinti sono ortogonali come dimostriamo nel seguente utile e importante risultato:

Teorema 2.5: *Siano V uno spazio vettoriale metrico e $T: V \to V$ un endomorfismo autoaggiunto. Autovettori di T relativi a autovalori distinti sono ortogonali.*

Dimostrazione: Siano $\mu \neq \nu$ autovalori di T e v, w autovettori relativi rispettivamente a μ e ν. Allora

$$\mu < v, w >=< \mu v, w >=< T(v), w >=< v, T(w) >=< v, \nu w >= \nu < v, w >$$

e quindi, dato che $\mu \neq \nu$ deve essere $< v, w >= 0$. □

Il Teorema 2.5 ci dice che uno spazio vettoriale metrico si decompone in somma diretta ortogonale di autospazi di un endomorfismo autoaggiunto. Suggerisce quindi di procedere per diagonalizzare un endomorfismo autaggiunto $T: V \to V$ di uno spazio metrico V seguendo il seguente schema:

I Passo: Fissata una base, si scrive la matrice associata a T, si calcola il polinomio caratteristico sfruttando la matrice e si trovano gli autovalori di T

II Passo: Per ciascun autovalore si trova una base ortonormale per il corrispondente autospazio. L'unione delle basi degli autospazi determinate in questo modo, grazie alla Proposizione 2.5 è una base ortonormale diagonalizzante per T.

Data una matrice simmetrica reale A si può adattare il procedimento per trovare una matrice ortogonale B tale che $D = B^TAB = B^{-1}AB$ sia diagonale:

I Passo: Si scrive il polinomio caratteristico e si calcolano gli autovalori di A.

II Passo: Per ciascun autovalore si trova una base ortonormale per il corrispondente autospazio. L'unione delle basi degli autospazi determinate in questo modo è una base ortonormale diagonalizzante per L_A. La matrice B che ha per colonne i vettori della base ottenuta in questo modo è la matrice cercata.

In tutti e due i casi nel I Passo occorre trovare le radici di un polinomio. Mentre sotto le ipotesi imposte si è sicuri che esistono il numero "giusto" di radici reali, trovarle effettivamente è tutta un'altra faccenda! Per diagonalizzare mediante matrici unitarie matrici hermitiane si procede in modo analogo.

Esempio Si consideri la matrice $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$. Il suo polinomio caratteristico è $p_A(\lambda) = (\lambda^2 - 1)^2$.
Dunque gli autovalori di A sono $1, -1$ ciascuno con moteplicità 2. Se V_1, V_{-1} sono i rispettivi autospazi si ha

$$V_1 = Ker(A - I_4) = Ker \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} = Span \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right)$$

$$V_{-1} = Ker(A + I_4) = Ker \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = Span \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right)$$

Dunque basi ortonormali per V_1 e V_{-1} sono date ad esempio rispettivamente da:

$$\left\{\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}\right\} \quad \left\{\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}\right\}$$

e la matrice ortonormale diagonalizzante B è data da

$$B = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

Concludiamo il paragrafo con un'osservazione che è alle volte utile. Abbiamo ripetutamente osservato che un endomorfismo arbitrario di uno spazio vettoriale non è in genere diagonalizzabile. D'altra parte si dimostra facilmente che se un endomorfismo di uno spazio vettoriale metrico di dimensione n su $\mathbb{K}$ ha n autovalori (se contati con la molteplicità) in $\mathbb{K}$, esiste una base ortonormale rispetto alla quale l'endomorfismo ha matrice triangolare superiore: dunque se non si può sempre diagonalizzare un endomorfismo, se c'è il numero giusto di autovalori, almeno lo si può triangolarizzare! Diamo di seguito la dimostrazione di questo risultato nella versione per matrici:

Teorema 2.6: *Per ogni $A \in M_n(\mathbb{C})$ esiste una matrice unitaria U tale che $S = U^{-1}AU = U^H AU$ è una matrice triangolare superiore. Analogamente, per ogni $A \in M_n(\mathbb{R})$ che ha n autovalori (contati con la loro molteplicità algebrica) reali esiste una matrice ortogonale O tale che $S = O^{-1}AO = O^T AU$ è una matrice triangolare superiore.*

Dimostrazione:

Dimostriamo il caso complesso procedendo per induzione. Il caso reale è del tutto analogo. Per $n = 1$ il risultato è ovvio. Supponiamo sia vero per matrici quadrate di ordine $n-1$. Sia λ_1 un autovalore dell'endomorfismo $L_A: \mathbb{C}^n \to \mathbb{C}^n$ e sia z_1 un autovettore relativo a λ_1 con $\|z_1\| = 1$. Sia infine $\{z_1, z_2, \ldots, z_n\}$ una base ortonormale di $\mathbb{C}^n$. La matrice B di L_A relativa a questa base ha questa forma:

$$B = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

dove A_1 è una matrice quadrata di ordine $n-1$. Dato che A e B sono matrici dello stesso endomorfismo L_A relative a basi ortonormali diverse, esiste una matrice unitaria U_0 di ordine n tale che $B = U_0^{-1}AU_0 = U_0^H AU_0$.

Per l'ipotesi induttiva esiste una matrice unitaria U_1' di ordine $n-1$ tale che $S' = (U_1')^{-1}AU_1' = U_1'^H AU_1'$ è triangolare superiore. Se definiamo

$$U_1 = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & U_1' & \\ 0 & & & \end{pmatrix},$$

allora U_1 è una matrice unitaria di ordine n tale che

$$\begin{aligned} S =& U_1^{-1}BU_1 = U_1^H BU_1 \\ =& \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & (U_1')^H & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & U_1' & \\ 0 & & & \end{pmatrix} \\ =& \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & (U_1')^H A_1 U_1' & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & S' & \\ 0 & & & \end{pmatrix} \end{aligned}$$

è triangolare superiore. Ma allora la matrice unitaria $U = U_0U_1$ è la matrice cercata dato che, rimettendo insieme i calcoli fatti, si ha $S = U^{-1}AU = U^H AU$. $\square$

Appendice: Il teorema di Hamilton-Cayley

Sia $T: V \to V$ un endomorfismo di uno spazio vettoriale V di dimensione n su $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Denotiamo $T^0 = Id_V$, $T^1 = T$ e per ogni intero $k \geq 2$

$$T^k = \underbrace{T \circ \ldots \circ T}_{k\ volte}$$

la composizione k volte di T. Dato un polinomio $p(\lambda) = a_0 + a_1\lambda + \ldots + a_m\lambda^m \in \mathbb{K}[\lambda]$, allora per un endomorfismo $T \in \mathcal{L}(V, V)$ definiamo

$$p(T) = a_0T^0 + a_1T^1 + \ldots + a_mT^m = a_0Id_V + a_1T + \ldots + a_mT^m \in \mathcal{L}(V, V).$$

Allo stesso per una matrice quadrata $A \in M_n(\mathbb{K})$ di ordine n, posto $A^0 = I_n$, $A^1 = A$ e per ogni intero $k \geq 2$

$$A^k = \underbrace{A \cdot \ldots \cdot A}_{k\ volte}$$

il prodotto k volte di A. Dato un polinomio $p(\lambda) = a_0 + a_1\lambda + \ldots + a_m\lambda^m \in \mathbb{K}[\lambda]$, allora per una matrice quadrata $A \in M_n(\mathbb{K})$ di ordine $n$definiamo

$$p(A) = a_0A^0 + a_1A^1 + \ldots + a_mA^m = a_0I_n + a_1A + \ldots + a_mA^m \in M_n(\mathbb{K}).$$

Dato che $dim(\mathcal{L}(V, V)) = n^2$, necessariamente gli endomorfismi $T^0, T^1, \ldots, T^{n^2}$ sono linearmente dipendenti e quindi esistono scalari $a_0, \ldots, a_{n^2} \in \mathbb{K}$ tali che $a_0T^0 + a_1T^1 + \ldots + a_{n^2}T^{n^2} = 0$. Questo fatto si può esprimere dicendo che esiste un polinomio $p(\lambda)$ di grado n^2 tale che $p(T) = 0$. Analogamente si osserva per una matrice $A \in M_n(\mathbb{K})$ di ordine n esiste un polinomio di grado $p(\lambda)$ di grado n^2 tale che $p(A) = 0$.

In effetti, usando il Teorema 2.6 si può dimostrare molto di più:

Teorema 2.7: (di Hamilton-Cayley)
(i) Sia $p_T(\lambda)$ il polinomio caratteristico di un endomorfismo $T \in \mathcal{L}(V, V)$. Allora $p_T(T) = O$.
(ii) Sia p_A il polinomio caratteristico di una matrice $A \in M_n(\mathbb{K})$. Allora $p_A(A) = 0$.

Dimostrazione: Dato che un endomorfismo è l'endomorfismo nullo se e solo se la sua matrice rispetto a una scelta qualunque di base è la matrice nulla i due punti dell'enunciato del teorema sono equivalenti. Inoltre se $T \in \mathcal{L}(V, V)$ è un automorfismo di uno spazio vettoriale reale e A è la matrice associata a T relativamente a una qualunque base fissata, dato che $p_T(\lambda) = p_A(\lambda) = p_{L_A}(\lambda)$ dove consideriamo $L_A: \mathbb{C}^n \to \mathbb{C}^n$, ci basta dimostrare il teorema per $L_A: \mathbb{C}^n \to \mathbb{C}^n$ per una qualunque matrice $A \in M_n(\mathbb{C})$. Osserviamo che per il Teorema 2.6 la matrice A è simile a una matrice triangolare superiore con sulla diagonale gli autovalori $\lambda_1, \ldots, \lambda_n$ di A (non necessariamente distinti). Questo equivale ad affermare che esiste una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ rispetto alla quale L_A ha matrice triangolare superiore con sulla diagonale gli autovalori $\lambda_1, \ldots, \lambda_n$. Il risultato desiderato seguirà allora una volta dimostrato il seguente

FATTO: Sia $T: V \to V$ un endomorfismo di uno spazio vettoriale di dimensione n su $\mathbb{K}$ tale che ha esiste una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ rispetto alla quale T ha matrice triangolare superiore con sulla diagonale gli autovalori $\lambda_1, \ldots, \lambda_n \in K$. Allora $p_T(T) = O$

Dimostrazione del Fatto: Dimostriamo l'affermazione per induzione su $n = dim(V)$. Per $n = 1$ non vi è molto da dimostrare. Infatti in questo caso $T(v) = \lambda_1 v$ per ogni $v \in V$ e $p_T(\lambda) = \lambda_1 - \lambda$. Dunque $p_T(T)(v) = (\lambda_1 Id_V - T)(v) = O$. Supponiamo ora che l'affermazione sia verificata per gli spazi vettoriali di dimensione $n - 1$. Per $j = 1, \ldots, n$ poniamo $V_j = Span(v_1, \ldots, v_j)$, in modo che $V = V_n$. Si ha che per ogni $j = 1, \ldots, n$

$$T(v_j) = \lambda_j v_j + w$$

per qualche $w \in V_{j-1}$. Segue allora che per ogni $j = 1, \ldots, n$ si ha $T(V_j) \subset V_j$. Dunque, in particolare, $T(V_{n-1}) \subset V_{n-1}$ e quindi $S = T_{|V_{n-1}}: V_{n-1} \to V_{n-1}$ è un endomorfismo dello spazio vettoriale $n-1$–dimensionale V_{n-1}. Per costruzione, relativamente alla base $\tilde{\mathcal{B}} = \{v_1, \ldots, v_{n-1}\}$, l'endomorfismo la restrizione S dell'endomorfismo T a V_{n-1} ha matrice triangolare superiore con sulla diagonale gli autovalori $\lambda_1, \ldots, \lambda_{n-1}$. Dunque, per l'ipotesi induttiva $p_S(S) = O$. D'altra parte, è immediato verificare che

$$p_S(\lambda) = (\lambda_1 - \lambda) \ldots (\lambda_{n-1} - \lambda) \qquad \text{e} \qquad p_T(\lambda) = (\lambda_1 - \lambda) \ldots (\lambda_{n-1} - \lambda)(\lambda_n - \lambda)$$

e quindi che

$$p_S(S) = (\lambda_1 Id - S) \circ \ldots \circ (\lambda_{n-1} Id - S) \qquad \text{e} \qquad p_T(\lambda) = (\lambda_1 Id - T) \circ \ldots \circ (\lambda_{n-1} Id - T) \circ (\lambda_n Id - T).$$

Sia allora $v \in V = V_n$. Allora si scrive in modo unico come $v = tv_n + w$ per uno scalare t e per $w \in V_{n-1}$. Come già osservato $T(w) \in V_{n-1}$ e che per ogni $u \in V_{n-1}$ si ha $T(w) = S(w)$ e

$$O = [p_S(S)](u) = [(\lambda_1 Id - S) \circ \ldots \circ (\lambda_{n-1} Id - S)]\,(u) = [(\lambda_1 Id - T) \circ \ldots \circ (\lambda_{n-1} Id - T))]\,(u).$$

Pertanto:

$$\begin{aligned}
[p_T(T)](v) &= [(\lambda_1 Id - T) \circ \ldots \circ (\lambda_{n-1} Id - T) \circ (\lambda_n Id - T)]\,(v) \\
&= [(\lambda_1 Id - T) \circ \ldots \circ (\lambda_{n-1} Id - T)]\,(\lambda_n(tv_n + w) - T(tv_n + w)) \\
&= [(\lambda_1 Id - T) \circ \ldots \circ (\lambda_{n-1} Id - T)]\,(\lambda_n w - T(w)) = p_S(\lambda_n w - T(w)) = O
\end{aligned}$$

dove si è usato il fatto che $\lambda_n w - T(w) \in V_{n-1}$ e la tesi segue. □

Il Teorema di Hamilton-Cayley ha molte conseguenze. Ne segnaliamo una immediata per il calcolo dell'inversa di una matrice:

Corollario 2.8: *Sia $p_A(\lambda) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1 \lambda + det(A)$ il polinomio caratteristico di una matrice invertibile $A \in M_n(\mathbb{K})$. Allora*

$$A^{-1} = -\frac{1}{det(A)} \left((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \ldots + a_1 I_n\right).$$

Dimostrazione: Dato che $P_A(A) = 0$, allora si ha

$$det(A) I_n = -\left((-1)^n A^n + a_{n-1} A^{n-1} + \ldots + a_1 A\right)$$

e quindi la formula desiderata si ottiene moltiplicando a destra e sinistra dell'eguaglianza per A^{-1}. □

3. Forme quadratiche su $\mathbb{R}^n$

Definizione 3.1 Una forma quadratica su $\mathbb{R}^n$ è una funzione $\phi: \mathbb{R}^n \to \mathbb{R}$ definita da $\phi(x) =< x, x >_\phi$ dove $< \bullet, \bullet >_\phi$ un prodotto scalare (non necessariamente definito positivo) su $\mathbb{R}^n$.

Sappiamo che dato un prodotto scalare $< \bullet, \bullet >_\phi$, se $A = (a_{ij}) \in M_{n,n}$ è la matrice simmetrica definita da

$$a_{ij} =< e_j, e_i >_\phi \tag{3.1}$$

si ha $< x, y >_\phi = y^T AT$ per ogni $x, y \in \mathbb{R}^n$. Dunque, in particolare si ha

$$\phi(x) =< x, x >_\phi = x^T Ax = \sum_{i,j=1}^{n} a_{ij} x_i x_j. \tag{3.2}$$

Chiameremo la matrice A definita in (3.1) la *matrice associata a ϕ rispetto alla base canonica*

Osservazione. Dalla (3.2) si vede subito che una forma quadratica su $\mathbb{R}^n$ non è altro che un polinomio reale omogeneo di grado 2 in n variabili ossia un polinomio reale $P(x_1, \ldots, x_n)$ in n variabili tale che per ogni $t \in \mathbb{R}^n$ risulta:

$$P(tx_1, \ldots, tx_n) = t^2 P(x_1, \ldots, x_n). \tag{3.3}$$

Viceversa dato un polinomio reale omogeneo di grado 2 in n variabili $p(x_1, \ldots, x_n)$, si vede subito che

$$P(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \sum_{i=1}^{n} a_i x_i^2 + \sum_{1 \le i < j \le n} 2a_{ij} x_i x_j$$

dove $A = (a_{ij}) \in M_{n,n}$ è una matrice simmetrica. Dunque il polinomio $p(x_1, \ldots, x_n)$ definisce una forma quadratica con matrice associata A.

Esempio. Si consideri il polinomio $P(x, y, z) = 5y^2 - 3z^2 + 8xy + 8xz - 2yz$. Il polinomio $P(x, y, z)$ è omogeneo di grado 2 e quindi definisce una forma quadratica:

$$P(x, y, z) = 5y^2 - 3z^2 + 8xy + 8xz - 2yz = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 0 & 4 & 4 \\ 4 & 5 & -1 \\ 4 & -1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

e $A = \begin{pmatrix} 0 & 4 & 4 \\ 4 & 5 & -1 \\ 4 & -1 & -3 \end{pmatrix}$ è la sua matrice associata rispetto alla base canonica.

Definizione 3.2 Sia $\phi: \mathbb{R}^n \to \mathbb{R}$ la forma quadratica associata alla matrice simmetrica $A = (a_{ij})$. La forma ϕ e la matrice A si dicono

(i) *definite positive* se

$$\phi(x) = x^T Ax = \sum_{i,j=1}^{n} a_{ij} x_i x_j > 0 \quad \forall x \neq 0; \tag{3.4}$$

(ii) *definite negative* se

$$\phi(x) = x^T Ax = \sum_{i,j=1}^{n} a_{ij} x_i x_j < 0 \quad \forall x \neq 0; \tag{3.5}$$

(iii) *semidefinite positive* se

$$\phi(x) = x^T Ax = \sum_{i,j=1}^{n} a_{ij} x_i x_j \geq 0 \quad \forall x \,\mathrm{e}\, \phi(x_0) = x_0^T Ax_0 = 0 \text{ per qualche } x_0 \neq 0; \tag{3.6}$$

(iv) *semidefinite negative* se

$$\phi(x) = x^T Ax = \sum_{i,j=1}^{n} a_{ij} x_i x_j \leq 0 \quad \forall x \,\mathrm{e}\, \phi(x_0) = x_0^T Ax_0 = 0 \text{ per qualche } x_0 \neq 0. \tag{3.7}$$

La forma ϕ e la matrice A si dicono *indefinite* se ϕ assume sia valori negativi sia positivi.

È naturale chiedersi se esistono invarianti per le forme quadratiche che ci permettono di classificarle e se è possibile determinare un cambio di base per $\mathbb{R}^n$ tale che nelle nuove coordinate una forma quadratica prende una espressione così semplice che le sue proprietà (ad esempio quelle di positività viste sopra) siano evidenti. Per cominciare, utilizzando quanto già osservato su come i cambi di base influiscono sull'espressione dei prodotti scalari, abbiamo immediatamente l'analogo per forme quadratiche:

Proposizione 3.1: *Sia $\phi: \mathbb{R}^n \to \mathbb{R}$ una forma quadratica e sia $A = (a_{ij})$ la matrice associata a ϕ rispetto alla base canonica. Sia $\mathcal{B} = \{\epsilon_1 \ldots, \epsilon_n\}$ una base di $\mathbb{R}^n$ e sia B la matrice del cambio di base dalla base canonica alla base $\mathcal{B}$. Se $A' = B^T AB$, allora per ogni $v \in \mathbb{R}^n$, se x' il vettore delle coordinate di v relative a $\mathcal{B}$, si ha*

$$\phi(v) = (x')^T A' x'. \tag{3.8}$$

La matrice $A' = (a'_{ij})$ si dice matrice associata a ϕ rispetto alla base $\mathcal{B}$ e si ha $a'_{ij} =< \epsilon_j, \epsilon_i >_\phi$.

Come applicazione dei risultati precedenti, abbiamo la seguente:

Proposizione 3.2: *Sia $\phi: \mathbb{R}^n \to \mathbb{R}$ la forma quadratica associata alla matrice simmetrica $A = (a_{ij})$. Siano $\lambda_1, \ldots, \lambda_n$ gli autovalori (eventualmente con ripetizione) di A. Esiste una base ortonormale $\mathcal{B}$ di $\mathbb{R}^n$ tale che se x' è il vettore delle coordinate rispetto a $\mathcal{B}$ di un vettore $v \in \mathbb{R}^n$, allora ϕ è data da*

$$\phi(v) = \lambda_1 (x'_1)^2 + \ldots \lambda_n (x'_n)^2. \tag{3.9}$$

Inoltre ϕ e A sono definite positive (negative) se e solo se tutti gli autovalori di A sono positivi (negativi); ϕ e A sono semidefinite positive (negative) se e solo se tutti gli autovalori di A sono non negativi (nonpositivi)e almeno uno nullo; ϕ e A sono indefinite se e solo se A ha autovalori sia positivi sia negativi.

Dimostrazione: Sia $\mathcal{B}$ una base ortonormale di $\mathbb{R}^n$ formata da autovettori di L_A e sia B la matrice del cambio di base dalla base canonica alla base $\mathcal{B}$. Allora, se x designa il vettore delle coordinate di $v \in \mathbb{R}^n$ relative alla base canonica e x' quello delle coordinate relative a $\mathcal{B}$, si ha $x = Bx'$. Dunque, ricordando che $D = B^T AB$ è una matrice diagonale con sulla diagonale gli autovalori di A, otteniamo immediatamente dalla Proposizione 3.1:

$$\phi(v) = \lambda_1 (x'_1)^2 + \ldots \lambda_n (x'_n)^2 \tag{3.10}$$

Dalla (3.10) è semplice studiare il segno della forma quadratica ϕ Ad esempio è immediato concludere che se tutti gli autovalori di A sono positivi allora ϕ è definita positiva. D'altra parte se ϕ è definita positiva, e $\mathcal{B} = \{v_1, \ldots, v_n\}$ allora per ogni j si ha $0 < \phi(v_j) = \lambda_j$. Le altre affermazioni si dimostrano in modo analogo.
□

Definizione 4.2 L'espressione (3.9) si dice *forma canonica metrica* della forma quadratica ϕ.

Si osservi che mentre per calcolare la forma canonica metrica di una forma quadratica occorre riuscire a calcolare tutti gli autovalori della matrice associata A, per studiarne il segno basta conoscere il segno degli autovalori. A tal fine viene in aiuto il Criterio di Cartesio che, pur di non facile dimostrazione per polinomi di grado arbitrario, è molto semplice da enunciare e da utilizzare. Lo ricordiamo rapidamente.

Criterio di Cartesio Sia $P(t) = a_k t^k + \ldots + a_d t^d$, dove $0 \leq d \leq k$ e $a_d \neq 0$, un polinomio a coefficienti reali con tutte le radici reali. Allora:

(i) 0 è radice di $P(t)$ se e solo se $d \geq 1$ e in questo caso è radice di molteplicità d.

(ii) il numero p delle radici positive del polinomio $P(t)$, contate con la loro molteplicità, è uguale al numero delle variazioni di segno nella successione dei coefficienti non nulli di $P(t)$.

(iii) il numero delle radici negative del polinomio $P(t)$, contate con la loro molteplicità, è $n = k - d - p$.

Esercizio. Si consideri la forma quadratica su $\mathbb{R}^3$ definita da $\phi(x) = 25x_1^2 - 7x_2^2 + 7x_3^2 + 48x_2x_3$. Trovare la forma canonica metrica di ϕ, decidere se è definita positiva e determinare una base ortonormale rispetto alla quale ϕ è in forma canonica metrica.

Si vede subito che ϕ è la forma quadratica associata alla matrice $A = \begin{pmatrix} 25 & 0 & 0 \\ 0 & -7 & 24 \\ 0 & 24 & 7 \end{pmatrix}$. Il polinomio caratteristico di A è $p(\lambda) = -(\lambda - 25)^2(\lambda + 25)$ e quindi A ha autovalori 25 con molteplicità 2 e -25 con molteplicità 1. Dunque ϕ ha forma canonica metrica $\phi(x') = 25x'^2_1 + 25x'^2_2 - 25x'^2_3$. Calcolando gli autospazi, si vede facilmente che una base ortonormale dell'autospazio relativo all'autovalore 25 è data da $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/5 \\ 4/5 \end{pmatrix} \right\}$ e una base ortonormale per l'autospazio relativo all'autovalore -25 è data da

$\left\{\begin{pmatrix} 0 \\ -4/5 \\ 3/5 \end{pmatrix}\right\}$. Si può allora concludere che una base ortonormale rispetto alla quale ϕ è in forma canonica metrica è $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/5 \\ 4/5 \end{pmatrix}, \begin{pmatrix} 0 \\ -4/5 \\ 3/5 \end{pmatrix}\right\}$.

Definizione 3.3 Sia $A = (a_{ij})$ una matrice simmetrica. Se p e q sono rispettivamente il numero di autovalori positivi e il numero di autovalori negativi di A, la coppia di numeri (p, q) si dice *segnatura* di A.

È evidente che se A' è la matrice associata a ϕ ripetto a un'altra base ortonormale, allora il numero di autovalori positivi e il numero di negativi di A' sono uguali al numero di autovalori positivi e il numero di negativi di A. Infatti le matrici A e A', per un'opportuna matrice ortogonale O, sono legate dalla relazione

$$A' = O^T AO = O^{-1}AO$$

e quindi, essendo simili, hanno gli stessi autovalori. Dunque la segnatura è un invariante della forma quadratica rispetto a cambiamenti di coordinate ortonormali. In realtà molto di più è vero. Si può dimostrare che la segnatura di una matrice simmetrica è invariante per congruenza ossia che due matrici congruenti (che in generale hanno autovalori diversi), hanno la stessa segnatura. Questo fatto porta a sua volta a classificare mediane la segnatura le forme quadratiche reali. Se (p, q) è la segnatura di una matrice simmetrica A, allora $r = p + q$ è il rango di A. Cominciamo osservando che il rango di una matrice è invariante per congruenza:

Lemma 3.3: *Matrici congruenti hanno lo stesso rango.*

Dimostrazione: Siano $A, A' \in M_n(\mathbb{R})$ tali che $A' = B^T AB$ per qualche $B \in GL_n(\mathbb{R})$. Dato che B è invertibile, si ha che

$$Im(L_{AB}) = L_A(L_B(\mathbb{R}^n)) = L_A(\mathbb{R}^n) = Im(L_A).$$

D'altra parte anche B^T è invertibile e quindi:

$$L_{B^T} \colon Im(L_{AB}) \to L_{B^T}(Im(L_{AB})) = Im(L_{B^T AB})$$

è lineare, iniettiva e suriettiva. Dunque $Im(L_A) = Im(L_{AB})$ e $Im(L_{B^T AB})$ sono isomorfi e quindi hanno la stessa dimensione e la tesi è dimostrata. □

Il secondo fatto che ci serve è osservare che una matrice simmetrica con assegnata segnatura è congruente a matrici molto speciali:

Lemma 3.4: *Sia A una matrice reale simmetrica di ordine n con segnatura (p, q). Allora A è congruente alla matrice $M_{p,q} = (m_{ij})$ diagonale di ordine n tale che*

$$m_{ii} = \begin{cases} 1 \text{ se } 1 \le i \le p \\ -1 \text{ se } p + 1 \le i \le p + q \\ 0 \text{ se } p + q + 1 \le i \le n \end{cases}$$

ossia

$$M_{p,q} = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0_{n-p-q} \end{pmatrix} \tag{3.11}$$

dove I_p, I_q sono le matrici identiche rispettivamente di ordine p e q, la matrice 0_{n-p-q} è la matrice quadrata nulla di ordine n e le gli altri 0 denotano matrici nulle delle appropriate dimensioni.

Dimostrazione: Se A è una matrice reale simmetrica di ordine n con segnatura (p, q), siano $\lambda_1, \dots, \lambda_p$ gli autovalori positivi di A e $\lambda_{p+1}, \dots, \lambda_{p+q}$ quelli negativi. Allora, come nella dimostrazione della Proposizione 3.2, usando il teorema spettrale, abbiamo che esiste una matrice ortogonale B tale che la matrice $(\alpha_{ij}) = A' = B^T AB$ `e diagonale con sulla diagonale esattamente gli autovalori di A, ossia $\alpha_{ij} = 0$ se $i \ne j$ e:

$$\alpha_{ii} = \begin{cases} \lambda_i \text{ se } 1 \le i \le p + q \\ \\ 0 \text{ se } p + q + 1 \le i \le n. \end{cases}$$

Sia ora $S = (s_{ij})$ la matrice diagonale, ossia tale che $s_{ij} = 0$ se $i \neq j$, e diagonale definita da:

$$s_{ii} = \begin{cases} |\lambda|^{\frac{1}{2}} \text{ se } 1 \leq i \leq p+q \\ \\ 1 \text{ se } p+q+1 \leq i \leq n. \end{cases}$$

Si vede subito che $M = S^T A' S = S^T B^T ABS = (BS)^T A(BS)$ è la desiderata matrice diagonale congruente ad A. □

Come conseguenza immediata per le forme quadratiche abbiamo

Corollario 3.5: *Se $\phi: \mathbb{R}^n \to \mathbb{R}$ è la forma quadratica che ha matrice $A = (a_{ij})$ rispetto alla base canonica, allora esiste una base $\mathcal{B}$ di $\mathbb{R}^n$, tale che, se $X_1, \dots, X_n$ sono le coordinate di $v \in \mathbb{R}^n$ rispetto alla base $\mathcal{B}$, si ha*

$$\phi(v) = X_1^2 + \ldots + X_p^2 - X_{p+1}^2 - \ldots - X_{p+q}^2.$$

Affinché si possa concludere che la segnatura è effettivamente un invariante per le matrici simmetriche rispetto alla congruenza e per le forme quadratiche reali rispetto ai cambi di base, occorre il seguente:

Lemma 3.6: *Per interi non negativi p, q, p', q' con $0 \leq p+q \leq n$ e $0 \leq p'+q' \leq n$ le matrici $M_{p,q}$ e $M_{p',q'}$ definite come in (3.11) sono congruenti se e solo se $(p,q) = (p',q')$ (ossia $p = p'$ e $q = q'$).*

Dimostrazione: Ovviamente se $(p,q) = (p',q')$ allora $M_{p,q} = M_{p',q'}$ e quindi le due matrici sono congruenti. Supponiamo che $M_{p,q}$ e $M_{p',q'}$ siano congruenti e sia $D \in GL_n(\mathbb{R})$ la matrice invertibile tale che $M_{p',q'} = D^T M_{p,q} D$. Dato i ranghi di $M_{p,q}$ e $M_{p',q'}$ devono essere uguali, si ha $p+q = p'+q'$. Dunque basterà dimostrare che $p = p'$. Se $p \neq p'$, allora $p > p'$ oppure $p < p'$. Supponiamo allora $p > p'$. Denotiamo con $\mathcal{C}$ la base canonica di $\mathbb{R}^n$ e con $\mathcal{D}$ la base di $\mathbb{R}^n$ che ha per colonne la matrice D. Per $v \in \mathbb{R}^n$, denotiamo con $x = F_{\mathcal{C}}(v)$ e $y = F_{\mathcal{D}}(v)$ rispettivamente le colonne delle coordinate di v relative rispettivamente alla base canonica $\mathcal{C}$ e alla base $\mathcal{D}$ e che sono quindi legate dalla relazione $x = Dy$. Se ϕ è la forma quadratica che ha matrice $M_{p,q}$ rispetto alla base canonica, allora si ha per ogni $v \in \mathbb{R}^n$:

$$\sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{p+q} x_i^2 = \phi(v) = \sum_{i=1}^{p'} y_i^2 - \sum_{i=p'+1}^{p'+q'} y_i^2. \tag{3.12}$$

Sia $D^{-1} = (u_{ij})$ la matrice inversa di D e si consideri il sistema:

$$\begin{cases} \sum u_{ij} x_j \qquad \text{per } i = 1, \dots, p' \\ \\ x_i = 0 \qquad \text{per } i = p+1, \dots, n. \end{cases} \tag{3.13}$$

Il sistema (3.13) ha n incognite e $p' + n - p < n$ equazioni e quindi, come conseguenza del teorema della dimensione, ammette una soluzione $\overline{x} \neq O$. Allora dato che $\overline{x} \neq 0$, vale (3.13) e D^{-1} è una matrice invertibile, si ha:

$$O \neq \overline{x} = \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{e} \qquad O \neq \overline{y} = D^{-1}(\overline{x}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \overline{y}_{p'+1} \\ \vdots \\ \overline{y}_n \end{pmatrix}$$

dove per almeno un indice i_0 con $0 \leq i_0 \leq p$ si ha $\overline{x}_{i_0} \neq 0$. Dunque se $\overline{v} \in \mathbb{R}^n$ è tale che $\overline{x} = F_{\mathcal{C}}(\overline{v})$ e $\overline{y} = F_{\mathcal{D}}(\overline{v})$, allora da (3.12) abbiamo

$$\sum_{i=1}^{p} \overline{x}_i^2 - \sum_{i=p+1}^{p+q} \overline{x}_i^2 = \phi(\overline{v}) = \sum_{i=1}^{p'} \overline{y}_i^2 - \sum_{i=p'+1}^{p'+q'} \overline{y}_i^2$$

e quindi l'assurdo seguente:

$$0 < x_{i_0}^2 \leq \sum_{i=1}^{p} \overline{x}_i^2 + \sum_{i=p'+1}^{p'+q'} \overline{y}_i^2 = \sum_{i=p+1}^{p+q} \overline{x}_i^2 + \sum_{i=1}^{p'} \overline{y}_i^2 = 0.$$

Dunque deve essere $p \leq p'$. Scambiando i ruoli di p e p' e ripetendo l'argomento con i cambiamenti ovvi, si ottiene che deve essere $p' \leq p$ e quindi la conclusione $p = p'$ segue. □

Raccogliendo tutte le informazioni che abbiamo, possiamo ora riassumere quello che abbiamo dimostrato nel seguente importante risultato:

Teorema 3.7: (di Sylvester) (i) *Sia A una matrice reale simmetrica di ordine n con segnatura (p, q). Allora A è congruente alla matrice $M_{p,q} = (m_{ij})$ diagonale di ordine n tale che*

$$m_{ii} = \begin{cases} 1 \text{ se } 1 \leq i \leq p \\ -1 \text{ se } p+1 \leq i \leq p+q \\ 0 \text{ se } p+q+1 \leq i \leq n. \end{cases}$$

(ii) *Due matrici reali simmetriche A, A' di ordine n sono congruenti (ossia $A' = B^T AB$ per qualche matrice invertibile $B \in GL(n, \mathbb{R})$) se e solo se A e A' hanno la stessa segnatura.*

(iii) *Se $\phi: \mathbb{R}^n \to \mathbb{R}$ è la forma quadratica che ha matrice $A = (a_{ij})$ rispetto a una base di $\mathbb{R}^n$ e A ha segnatura (p, n), allora la matrice associata a ϕ rispetto a una qualunque base ha segnatura (p, n) e esiste una base $\mathcal{B}$ di $\mathbb{R}^n$, tale che, se $X_1, \ldots, X_n$ sono le coordinate di $p \in \mathbb{R}^n$ rispetto alla base $\mathcal{B}$, si ha*

$$\phi(p) = X_1^2 + \ldots + X_p^2 - X_{p+1}^2 - \ldots - X_{p+n}^2.$$

Dimostrazione: Il punto (i) è il Lemma 3.4. Il punto (ii) segue dal fatto che matrici A e A' di segnatura rispettivamente (p, q) e (p', q') sono congruenti rispettivamente a $M_{p,q}$ e $M_{p',q'}$ e che, per il Lemma 3.6, $M_{p,q}$ e $M_{p',q'}$ sono congruenti se e solo se $(p, q) = (p', q')$. Il punto (iii) segue dal 3.5 e dal punto (ii). □

Il punto (iii) del Teorema di Sylvester rende ben posta la seguente

Definizione Si dice segnatura di una forma quadratica $\phi: \mathbb{R}^n \to \mathbb{R}$ la segnatura di una matrice associata a ϕ rispetto a una scelta di base di $\mathbb{R}^n$.

Il significato intrinseco della segnatura di una forma quadratica è illustrato nel risultato finale di questo paragrafo:

Teorema 3.8: *Se (p, q) è la segnatura della forma quadratica $\phi: \mathbb{R}^n \to \mathbb{R}$, allora p è la massima dimensione per un un sottospazio di $\mathbb{R}^n$ sul quale ϕ è definita positiva e q è la massima dimensione per un un sottospazio di $\mathbb{R}^n$ sul quale ϕ è definita negativa.*

Dimostrazione: Se (p, q) è la segnatura della forma quadratica $\phi: \mathbb{R}^n \to \mathbb{R}$, esiste una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di $\mathbb{R}^n$, tale che, se $X_1, \ldots, X_n$ sono le coordinate di $v \in \mathbb{R}^n$ rispetto alla base $\mathcal{B}$, si ha

$$\phi(v) = X_1^2 + \ldots + X_p^2 - X_{p+1}^2 - \ldots - X_{p+q}^2. \tag{3.14}$$

Definiamo $V_p = Span(v_1, \ldots, v_p)$ e $W = Span(v_{p+1}, \ldots, v_n)$. Allora, per costruzione, $V = V_p \oplus W$. Sia U un qualunque sottospazio di $\mathbb{R}^n$ tale che la restrizione ϕ_U di ϕ a U è definita positiva tale che per ogni $v \in U \setminus \{O\}$ si ha $\phi(v) > 0$. Allora $U \cap W = \{O\}$. Infatti se $v \in U \cap W$ allora, dato che $v \in U$ si ha $\phi(v) \geq 0$, mentre dato che $v \in W$, da (3.14), si ha $\phi(v) \leq 0$. Dunque deve essere $\phi(v) = 0$ e quindi, necessariamente, $v = O$. Ma allora $U + W = U \oplus W \subset V$ e pertanto

$$dim(U) + dim(W) = dim(U \oplus W) \leq dim(V) = n = dim(V_p) + dim(W) = p + dim(W)$$

ossia $dim(U) \leq p$ e la prima parte dell'enunciato è dimostrata. La seconda si dimostra con analogo argomento.

□