Corso di Laurea in Fisica. Geometria. a.a. 2023-24.
Prof. P. Piazza
Diagonalizzazione delle forme bilineari simmetriche

1. Forma bilineari e matrici.

Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Sia

$$b : V \times V \to \mathbb{K}$$

una forma bilineare. Sia V di dimensione n e $\mathcal{B} := \{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ una base di V. Rimane allora definita la matrice associata a $b(\ ,\ )$ nella base fissata; per definizione questa è la matrice $n \times n$

$$A_b^{\mathcal{B}} \equiv (a_{ij}) := (b(\underline{v}_j, \underline{v}_i)) .$$

L'espressione della forma bilineare nelle coordinate associate a $\mathcal{B} = \{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ è data da

$$(1)\quad b(x_1\underline{v}_1 + \cdots + x_n\underline{v}_n\ ,\ y_1\underline{v}_1 + \cdots + y_n\underline{v}_n) = |y_1, \dots, y_n| \cdot A_b^{\mathcal{B}} \cdot \begin{vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{vmatrix} = \underline{y}^T \cdot A_b^{\mathcal{B}} \cdot \underline{x}$$

La verifica di quest'espressione è diretta, usando la bilinearità di b e ricordando la definizione di prodotto righe per colonne. Per semplificare la notazione scriveremo semplicemente $\underline{y}^T\, A_b^{\mathcal{B}}\, \underline{x}$, senza i punti centrali.
Quindi:

$$(2)\qquad b(\underline{v}, \underline{w}) = \underline{y}^T\, A_b^{\mathcal{B}}\, \underline{x}$$

se $\underline{v}$ ha coordinate $\underline{x}$ e $\underline{w}$ ha coordinate $\underline{y}$ nella base fissata. La (2) è l'*espressione della forma bilineare $b(\ ,\ )$ nelle coordinate associate a $\mathcal{B}$* .
Prima di procedere vogliamo richiamare la seguente utile osservazione (l'abbiamo già fatta quando abbiamo visto il Teorema Spettrale).
Osservazione. Se A e B sono due matrici in $M_{n\times n}(\mathbb{K})$, allora

$$\underline{y}^T A\underline{x} = \underline{y}^T B\underline{x}\ \ \forall \underline{x}, \forall \underline{y} \in \mathbb{K}^n \quad \Leftrightarrow \quad A = B\,.$$

Notiamo che per ogni matrice quadrata $A \in M_{n\times n}(\mathbb{K})$ si ha

$$(3)\qquad \underline{y}^T A\underline{x} = \underline{x}^T\, A^T\, \underline{y}\ ;$$

infatti $\underline{y}^T A\underline{x}$ è una matrice 1×1 e quindi $\underline{y}^T A\underline{x} = (\underline{y}^T A\underline{x})^T$ da cui

$$\underline{y}^T A\underline{x} = \underline{x}^T A^T (\underline{y}^T)^T = \underline{x}^T A^T \underline{y}$$

come si voleva. Vediamo allora che

$$b \text{ è simmetrica se e solo se } A_b^{\mathcal{B}} \text{ è simmetrica :}$$

infatti, scrivendo l'espressione in coordinate di $b(\underline{w}, \underline{v})$ otteniamo

$$b(\underline{w}, \underline{v}) = \underline{x}^T\, A_b^{\mathcal{B}}\, \underline{y}$$

e questa espressione è uguale a $\underline{x}^T (A_b^{\mathcal{B}})^T \underline{y} = b(\underline{v}, \underline{w})$ (dove abbiamo usato (3)) se e solo se $A_b^{\mathcal{B}} = (A_b^{\mathcal{B}})^T$ dove nell'ultimo passaggio abbiamo utilizzato l'Osservazione.

Una forma bilineare simmetrica è detta un *prodotto scalare* e spesso denotata con $<\,,\,>$.

D'ora in avanti ci concentreremo sulle forme bilineari simmetriche, dette anche prodotti scalari. (Attenzione: in molti testi un prodotto scalare è una forma bilineare simmetrica *definita positiva*, tale cioè che $b(\underline{v}, \underline{v}) > 0$ per ogni $\underline{v}$ non-nullo.)

Vi ricordo, Definizione 11.5 nel libro di testo, che il nucleo (o *radicale*) di una forma bilineare simmetrica è il sottospazio $V^{\perp}$ di V costituito dai vettori $\underline{w} \in V$ tali che $b(\underline{w}, \underline{v}) = 0 \ \forall \underline{v} \in V$. Sia $\mathcal{B} := \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ una base di V e sia $F_{\mathcal{B}} : V \to \mathbb{R}^n$ l'isomorfismo dato dalle coordinate in questa base. Sia $A_b^{\mathcal{B}}$ la matrice associata a questa forma bilineare simmetrica in questa base $\mathcal{B}$.

Proposizione. *L'immagine del nucleo di $b(\,,\,)$ tramite $F_{\mathcal{B}}$ è uguale al nucleo della matrice $A_b^{\mathcal{B}}$.*
Dimostrazione. Prima di dare l'argomento facciamo un'altra osservazione elementare. Se $\underline{z} \in \mathbb{R}^n$ e $\underline{y}^T \cdot \underline{z} = 0 \ \forall \underline{y} \in \mathbb{R}^n$ allora $\underline{z} = \underline{0}$. Infatti $\underline{y}^T \cdot \underline{z} = \underline{z} \bullet \underline{y}$, il prodotto scalare canonico. Sappiamo che il prodotto scalare canonico è non degenere e quindi il suo nucleo, e cioè $\{\underline{z} \in \mathbb{R}^n \,|\, \underline{z} \bullet \underline{y} = 0 \ \forall \underline{y} \in \mathbb{R}^n\}$ è costituito dal solo vettore nullo. E infatti se $\underline{y}^T \cdot \underline{z} = 0 \ \forall \underline{y} \in \mathbb{R}^n$ possiamo scegliere $\underline{y} = \underline{z}$ ed allora $\underline{z}^T \cdot \underline{z} = 0$ che vuol dire $z_1^2 + z_2^2 + \cdots + z_n^2 = 0$ che vuol dire $\underline{z} = \underline{0}$. Torniamo alla dimostrazione della Proposizione. Il vettore $\underline{v}$ di coordinate $\underline{x}$ è un elemento in $V^{\perp}$ se e solo se $b(\underline{v}, \underline{w}) = 0 \ \forall \underline{w} \in V$, se e solo se $\underline{y}^T \cdot A_b^{\mathcal{B}} \cdot \underline{x} = 0 \ \forall \underline{y} \in \mathbb{R}^n$ se e solo se (per quanto appena osservato) $A_b^{\mathcal{B}} \cdot \underline{x} = \underline{0}$ se e solo se $\underline{x} \in \mathrm{Ker} A_b^{\mathcal{B}}$, che era quello che dovevamo dimostrare.

In particolare:

Proposizione. *$b(\,,\,)$ è non-degenere se e solo se $A_b^{\mathcal{B}}$ è non-singolare se e solo se $\det(A_b^{\mathcal{B}}) \neq 0$.*

2. Matrici congruenti

Sia $\mathcal{F} = \{\underline{f}_1, \ldots, \underline{f}_n\}$ un'altra base di V. Sia $A_b^{\mathcal{F}}$ la matrice associata a $b(\,,\,)$ in questa base. Il vettore generico $\underline{v}$ ha coordinate $(x_1, \ldots, x_n)$ nella base $\mathcal{B}$ e coordinate $(x'_1, \ldots, x'_n)$ nella base $\mathcal{F}$. Analogamente $\underline{w}$ ha coordinate $(y_1, \ldots, y_n)$ nella base $\mathcal{B}$ e coordinate $(y'_1, \ldots, y'_n)$ nella base $\mathcal{F}$. Sia C la matrice del cambiamento di base, dalla base $\mathcal{B}$ alla base $\mathcal{F}$. Vi ricordo, che questa è la matrice invertibile che ha come j-ma colonna le coordinate di $\underline{f}_j$ nella base $\mathcal{B}$; abbiamo denotato questa matrice con il simbolo $M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$. Sappiamo che $\underline{x} = C\underline{x}'$, $\underline{y} = C\underline{y}'$ e quindi

$$b(\underline{v}, \underline{w}) = \underline{y}^T A_b^{\mathcal{B}} \underline{x} = (C\underline{y}')^T A_b^{\mathcal{B}} (C\underline{x}') = (\underline{y}')^T (C^T A_b^{\mathcal{B}} C)\underline{x}'$$

da cui deduciamo

$$(\underline{y}')^T (C^T A_b^{\mathcal{B}} C)\underline{x}' = (\underline{y}')^T A_b^{\mathcal{F}} \underline{x}' \,, \quad \forall \underline{x}', \underline{y}'$$

da cui, infine, per la usuale osservazione

$$A_b^{\mathcal{F}} = C^T \, A_b^{\mathcal{B}} \, C \,. \qquad \text{con } C = M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)\,. \tag{4}$$

Questa è la **formula magica** per le matrici associate ad una forma bilineare in due basi diverse.
Diamo ora la seguente

Definizione. *Due matrici A e D sono dette congruenti se esiste una matrice invertibile C tale che $D = C^T AC$.*

Da quanto appena visto concludiamo che vale la seguente

Proposizione. *Le matrici associate ad una forma bilineare simmetrica in due basi diverse sono* **congruenti**.

Proposizione. *Due matrici congruenti hanno lo stesso rango.*
Dimostrazione. Sia $D = C^T AC$, con C invertibile. Osserviamo che se A è una matrice e M è una matrice invertibile, allora $\mathrm{rg}(MA) = \mathrm{rg}(A)$, perché gli operatori $L_{MA} = L_M \circ L_A$ e L_A hanno immagini della stessa dimensione (L_M è un isomorfismo e gli isomorfismi preservano la dimensione dei sottospazi). Analogamente, se N è invertibile

$$\mathrm{rg}(AN) = \mathrm{rg}((AN)^T) = \mathrm{rg}(N^T A^T) = \mathrm{rg}(A^T) = \mathrm{rg}A\,,$$

dove abbiamo utilizzato noti risultati sul rango e il risultato appena dimostrato per giustificare che $\mathrm{rg}(N^T A^T) = \mathrm{rg}(A^T)$. In definitiva

$$\mathrm{rg}(C^T AC) = \mathrm{rg}(AC) = \mathrm{rg}A$$

come si voleva.

Definizione. *Il rango di una forma bilineare $b(\,,\,)$ è il rango di $A_b^{\mathcal{B}}$ con $\mathcal{B}$ una qualsiasi base di V. La definizione è ben posta per la Proposizione.*

3. Prodotti scalari definiti positivi.

Supponiamo che V sia uno spazio vettoriale **reale**. Un prodotto scalare definito positivo è una forma bilineare simmetrica $b(\,,\,)$ tale che $b(\underline{v}, \underline{v}) > 0$ per ogni vettore non-nullo $\underline{v}$. Qui abbiamo utilizzato il fatto che il campo reale è un campo totalmente ordinato (questa definizione non ha senso per un campo qualsiasi). Si utilizza spesso la notazione $<\,,\,>$ invece di $b(\,,\,)$.

La coppia $(V, <\,,\,>)$ è per definizione uno **spazio vettoriale metrico.**

Cambiamento di base ortonormale. Sia $\mathcal{B}$ una base ortonormale in uno spazio vettoriale metrico $(V, <\,,\,>)$. Si ha allora $A^{\mathcal{B}}_{<\,,\,>} = I_n$. Sia $\mathcal{F}$ una seconda base. Assumiamo che $\mathcal{F}$ sia anche ortonormale. Si ha allora $A^{\mathcal{F}}_{<\,,\,>} = I_n$. Riassumendo,

$$A^{\mathcal{B}}_{<\,,\,>} = I_n\,, \quad A^{\mathcal{F}}_{<\,,\,>} = I_n\,.$$

La formula magica (4) ci dice che

$$A^{\mathcal{F}}_{<\,,\,>} = C^T\, A^{\mathcal{B}}_{<\,,\,>}\, C\,. \qquad \text{con } C = M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V).$$

Quindi $M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$, la matrice del cambiamento di base, verifica

$$I_n = ((M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V))^T\, I_n\, M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$$

e cioè

$$I_n = (M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V))^T M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$$

Riassumendo:

Proposizione. *Sia $\mathcal{B}$ una base ortonormale di $(V, <\,,\,>)$. Sia $\mathcal{F}$ una seconda base ortonormale Sia B la matrice del cambiamento di base, da $\mathcal{B}$ a $\mathcal{F}$: $B = M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$. Allora $B^T B = I_n$*

4. Gruppo Ortogonale

Definizione. Sia $O(n) := \{A \in M_{n,n}(\mathbb{R}) \,|\, A^T A = I_n\}$. Le matrici di $O(n)$ sono dette *ortogonali.*

Osserviamo innanzitutto che le matrici ortogonali sono invertibili. Infatti, dalla definizione, si ha $\det(A^T A) = 1$. Ma allora, per Binet, $1 = \det(A^T)\det(A) = (\det A)^2$ e quindi $\det A = \pm 1$. Quindi, riassumendo, $O(n) \subset GL_n(\mathbb{R})$
Notiamo anche che

$$O(n) = \{A \in GL_n(\mathbb{R}) \,|\, A^{-1} = A^T\}\,. \tag{5}$$

Infatti da $A^T A = I_n$ si deve avere, moltiplicando a destra ambo i membri per A^{-1}, $A^T = A^{-1}$ che è quello che stiamo enunciando in (5).
Possiamo **rienunciare** la Proposizione della sezione precedente nel modo seguente:

Proposizione. *Sia $\mathcal{B}$ una base ortonormale di $(V, <\,,\,>)$. Sia $\mathcal{F}$ una seconda base ortonormale. Sia B la matrice del cambiamento di base, da $\mathcal{B}$ a $\mathcal{F}$: $B = M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$. Allora $B \in O(n)$.*

Proposizione 4. *$O(n)$ con il prodotto righe per colonne è un sottogruppo [1] di $GL_n(\mathbb{R})$, detto gruppo ortogonale.*
Dimostrazione. Dobbiamo dimostrare che il prodotto righe per colonne di due matrici ortogonali è ancora ortogonale e che l'inversa di una matrice ortogonale è ortogonale. Se $A, B \in O(n)$ allora $(AB)^T AB = B^T A^T AB = B^T I_n B = B^T B = I_n$; quindi $AB \in O(n)$ come volevasi. Inoltre, se $A \in O(n)$ allora $A^T = A^{-1}$ e quindi $(A^T)^{-1} = (A^{-1})^{-1} (= A)$; dato che $(A^T)^{-1} = (A^{-1})^T$ abbiamo in definitiva $(A^{-1})^T = (A^{-1})^{-1}$ e cioè $A^{-1} \in O(n)$ come dovevamo dimostrare.

Il sottoinsieme $SO(n) = \{A \in O(n) \,|\, \det A = 1\}$ è anche un sottogruppo di $GL(n, \mathbb{R})$, (il prodotto di due elementi in $SO(n)$ è ancora in $SO(n)$ e l'inverso di un elemento in $SO(n)$ è ancora in $SO(n)$).
$SO(n)$ è detto *gruppo ortogonale speciale.*

Esempio. Non è difficile dimostrare (potete farlo per esercizio o consultare il libro di testo) che

$$O(2) = \{\begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix}, \theta \in \mathbb{R}\} \cup \{\begin{vmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{vmatrix}, \theta \in \mathbb{R}\}$$

$$SO(2) = \{\begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix}, \theta \in \mathbb{R}\}$$

5. Il teorema di Sylvester

In questa sezione utilizzeremo il teorema spettrale per dimostrare il seguente importante risultato di diagonalizzazione per le forme bilineari simmetriche.

Teorema (Sylvester) *Sia $b(\,,\,)$ una forma bilineare simmetrica su uno spazio vettoriale reale V di dimensione n. Sia r il rango di $b(\,,\,)$. Allora esiste un intero*

[1] Se $(G, \cdot)$ è un gruppo e H è un sottoinsieme di G allora diremo che H è un sottogruppo se
(1) $h_1 \cdot h_2 \in H \ \forall h_1, h_2 \in H$
(2) $\forall h \in H$ si ha che $h^{-1} \in H$.

positivo ρ, dipendente solo da $b(\ ,\ )$, ed una base $\mathcal{F}$ tali che

$$A_b^{\mathcal{F}} = \begin{vmatrix} I_\rho & 0 & 0 \\ 0 & -I_{r-\rho} & 0 \\ 0 & 0 & 0_{n-r} \end{vmatrix}.$$

I numeri ρ, $r-\rho$ e $n-r$ sono detti rispettivamente, indice di positività, indice di negatività ed indice di nullità della forma bilineare simmetrica $b(\ ,\ )$. La coppia di numeri $(\rho, r-\rho)$ è detta segnatura di b. La base $\mathcal{F}$ è, per definizione, **una base di Sylvester.**

Dimostrazione. Consideriamo una qualsiasi base $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ e la matrice associata a $b(\ ,\ )$ in questa base: $A_b^{\mathcal{B}}$. Questa matrice ha rango r per ipotesi. La denotiamo più semplicemente con A ed i suoi coefficienti li denotiamo a_{ij}. Utilizzando la base $\mathcal{B}$ introduciamo una struttura di spazio vettoriale metrico in V definendo

$$< \underline{v}, \underline{w} > := \underline{x} \bullet \underline{y}$$

se

$$\underline{v} = x_1\underline{v}_1 + \cdots + x_n\underline{v}_n\,, \quad \underline{w} = y_1\underline{v}_1 + \cdots + y_n\underline{v}_n.$$

È immediato verificare che questo è un prodotto scalare definito positivo; inoltre la base $\mathcal{B}$ è ortonormale per $<\ ,\ >$ perché

$$< \underline{v}_i, \underline{w}_j > := \underline{e}_i \bullet \underline{e}_j$$

e sappiamo che la base canonica è ortonormale per il prodotto scalare canonico $\bullet$.
Consideriamo l'operatore $T : V \to V$ definito da

$$T\underline{v}_j := \sum_k a_{kj}\underline{v}_k \tag{6}$$

Per costruzione la matrice associata all'operatore T nella base $\mathcal{B}$, $M_{\mathcal{B},\mathcal{B}}(T)$, è proprio A. In particolare T ha un nucleo di dimensione $n-r$ perché la dimensione del nucleo è uguale alla dimensione di n meno il rango di A e sappiamo che A ha rango r. Dato che $A = A^T$, perché $b(\ ,\ )$ è simmetrica, e dato che $\mathcal{B}$ è ortonormale per il prodotto scalare $<\ ,\ >$ che abbiamo definito in V, ne segue che T è un operatore simmetrico nello spazio vettoriale metrico $(V, <\ ,\ >)$ [2].

Abbiamo quindi un operatore simmetrico in uno spazio vettoriale metrico; per il teorema spettrale esiste una base *ortonormale* $\mathcal{W} = \{\underline{w}_1, \dots, \underline{w}_n\}$ costituita da autovettori per T. Siano $\lambda_1, \dots, \lambda_\rho \in \mathbb{R}$ gli autovalori positivi di T. Ci sono allora $r-\rho$ autovalori negativi $\lambda_{\rho+1}, \dots, \lambda_r$; inoltre, come già osservato, l'autovalore 0 ha molteplicità $n-r$. Ordiniamo gli autovalori come sopra, in modo tale cioè che i primi ρ siano positivi, i seguenti $r-\rho$ siano negativi ed i rimanenti $n-r$ siano uguali a zero. Penseremo la base $\mathcal{W}$ ordinata di conseguenza, quindi $\underline{w}_1$ è associato al primo autovalore della nostra lista ordinata e cosi' via. Sia C la matrice del cambiamento di base, dalla base $\mathcal{B}$ alla base $\mathcal{W}$: questa è la matrice che ha come colonne le coordinate degli autovettori $\underline{w}_j$ nella base $\mathcal{B}$ ed è denotata, come è ben noto, con $M_{\mathcal{B},\mathcal{W}}(\mathrm{Id}_V)$. Sappiamo che C è ortogonale, $C \in O(n)$, perché entrambe le basi sono ortonormali. Inoltre, sappiamo dalla formula magica per operatori che

$$M_{\mathcal{W},\mathcal{B}}(\mathrm{Id}_V) \cdot M_{\mathcal{B},\mathcal{B}}(T) \cdot M_{\mathcal{B},\mathcal{W}}(\mathrm{Id}_V) = M_{\mathcal{W},\mathcal{W}}(T) = \Lambda$$

---

[2] Abbiamo visto nelle **Note sul Teorema Spettrale** che T è simmetrico se e solo se data una qualsiasi base ortonormale $\mathcal{B}$ si ha che $M_{\mathcal{B},\mathcal{B}}(T)$ è simmetrica.

con Λ la matrice diagonale che ha gli autovalori (ordinati) sulla diagonale principale. Di conseguenza

$$(M_{\mathcal{B},\mathcal{W}}(\mathrm{Id}_V))^{-1} \cdot M_{\mathcal{B},\mathcal{B}}(T) \cdot M_{\mathcal{B},\mathcal{W}}(\mathrm{Id}_V) = \Lambda$$

che riscriviamo brevemente come

$$C^{-1}AC = \Lambda\,.$$

Dato che $C \in O(n)$ si ha che $C^{-1} = C^T$ e quindi la cruciale relazione:

$$C^T AC = \Lambda\,.$$

Ora, da quanto visto nella sezione precedente,

$$A_b^{\mathcal{W}} = C^T A_b^{\mathcal{B}} C \equiv C^T AC$$

e quindi, da quanto appena dimostrato,

$$A_b^{\mathcal{W}} = \Lambda,$$

che è diagonale. Abbiamo quindi diagonalizzato la forma bilineare; abbiamo cioè trovato una base $\mathcal{W}$ tale che

$$b(\underline{w}_i, \underline{w}_j) = 0 \;\text{ se }\; i \neq j \quad \text{e} \quad b(\underline{w}_i, \underline{w}_i) = \lambda_i$$

Notare che la base *ortonormale* di autovettori diagonalizza **simultaneamente** l'operatore T e la forma bilineare simmetrica b.
Consideriamo ora la base $\mathcal{F}$ ottenuta riscalando opportunamente gli autovettori ortonormali; poniamo quindi

$\underline{f}_j = \frac{1}{\sqrt{\lambda_j}} \underline{w}_j$ se $1 \leq j \leq \rho$,
$\underline{f}_j = \frac{1}{\sqrt{-\lambda_j}} \underline{w}_j$ se $j \in \{\rho+1, \ldots, r\}$ e
$\underline{f} = \underline{w}_j$ se $r+1 \leq j \leq n$.
(Questi sono ovviamente ancora autovettori, ma non sono di lunghezza unitaria.)
Si ha:

$b(\underline{f}_\ell, \underline{f}_k) = 0$ per ogni $k \neq \ell$
$b(\underline{f}_j, \underline{f}_j) = 1$ se $1 \leq j \leq \rho$
$b(\underline{f}_j, \underline{f}_j) = -1$ se $j \in \{\rho+1, \ldots, r\}$ e
$b(\underline{f}_j, \underline{f}_j) = 0$ se $r+1 \leq j \leq n$.

È a questo punto chiaro che

$$A_b^{\mathcal{F}} = \begin{vmatrix} I_\rho & 0 & 0 \\ 0 & -I_{r-\rho} & 0 \\ 0 & 0 & 0_{n-r} \end{vmatrix}$$

come si voleva.

Rimane da dimostrare che in questa matrice r e ρ dipendono solo da $b(\,,\,)$ e non dalla particolare base scelta. Già sappiamo che r dipende solo da $b(\,,\,)$. Sia $\mathcal{A} := \{\underline{a}_1, \ldots, \underline{a}_n\}$ un'altra base di V rispetto alla quale $b(\,,\,)$ si scriva in forma diagonale. Sia quindi

$$A_b^{\mathcal{A}} = \begin{vmatrix} I_\sigma & 0 & 0 \\ 0 & -I_{r-\sigma} & 0 \\ 0 & 0 & 0_{n-r} \end{vmatrix}\,.$$

Dobbiamo dimostrare che $\rho = \sigma$. Per assurdo $\rho \neq \sigma$ e supponiamo che sia $\sigma < \rho$. Sia $U = \mathrm{Span}(\underline{f}_1, \ldots, \underline{f}_\rho)$ e sia $W = \mathrm{Span}(\underline{a}_{\sigma+1}, \ldots, \underline{a}_n)$. Per ipotesi $\dim U + \dim W = \rho + (n - \sigma) \equiv (\rho - \sigma) + n > n$ e quindi, per Grassmann, l'intersezione di questi due sottospazi ha dimensione > 0. Sia $\underline{h}$ un vettore non nullo in $U \cap W$; quindi

$$\underline{h} = b_1 \underline{f}_1 + \cdots + b_\rho \underline{f}_\rho \quad \text{perché} \quad \underline{h} \in U$$

e

$$\underline{h} = \beta_{\sigma+1} \underline{a}_{\sigma+1} + \cdots + \beta_n \underline{a}_n \quad \text{perché} \quad \underline{h} \in W.$$

Consideriamo il numero reale $b(\underline{h}, \underline{h})$. Nel primo sistema di coordinate $b(\underline{h}, \underline{h}) = b_1^2 + \cdots + b_\rho^2 > 0$; nel secondo sistema di coordinate $b(\underline{h}, \underline{h}) = -\beta_{\sigma+1}^2 - \cdots - \beta_n^2 \leq 0$ e questo è chiaramente assurdo. Il teorema di Sylvester è dimostrato.

6. Conseguenze del teorema di Sylvester

Innanzitutto, direttamente dal teorema di Sylvester capiamo che una forma bilineare è definita positiva se e solo se $\rho = n$ nella matrice di Sylvester. Infatti:

- se $b(\, , \,)$ è definita positiva allora per Grahm-Schmidt esiste una base ortonormale, nella quale quindi la forma bilineare ha matrice associata uguale all'identità;
- viceversa, se $\rho = n$ allora nella base $\mathcal{F}$ dell'enunciato del teorema si ha per $\underline{v}$ di coordinate $\underline{x}$, $b(\underline{v}, \underline{v}) = x_1^2 + \cdots + x_n^2$ che è maggiore di zero se $\underline{v}$ è non nullo.

In particolare, questo vuol dire che $b(\, , \,)$ è definito positivo se e solo se gli autovalori della matrice simmetrica A, $A = A_b^{\mathcal{B}}$, con $\mathcal{B}$ una qualsiasi base, sono tutti positivi. Analogamente, $b(\, , \,)$ è definita negativa se e solo se gli autovalori di A sono tutti minori di zero. È inoltre chiaro che se il rango di A non è n allora la forma bilineare è degenere e che se esistono due autovalori di segno discorde allora la forma è indefinita: infatti se $\underline{v}_\lambda$ è un autovettore associato a $\lambda > 0$ per l'operatore simmetrico definito da A e se $\underline{v}_\mu$ è un autovettore associato all'autovalore $\mu < 0$ allora si ha

$$b(\underline{v}_\lambda, \underline{v}_\lambda) = \lambda > 0 \quad b(\underline{v}_\mu, \underline{v}_\mu) = \mu < 0.$$

Il messaggio qui è che **leggete tutte le proprietà di $b(\, , \,)$ dalla matrice Λ o, equivalentemente dalla forma canonica di Sylvester.**

Importante: per capire il segno degli autovalori (senza calcolarli !) potete applicare l'utilissimo *criterio di Cartesio* che enuncio qui sotto e la cui dimostrazione trovate nel libro di Abate *Geometria* (complementi al Capitolo 16).

Criterio di Cartesio.
Sia $p(t) = a^n t^n + a_{n-1} t^{n-1} + \cdots + a_d t^d$ un polinomio di grado n a coefficienti reali, con $0 \leq d \leq n$ e $a_d \neq 0$. Supponiamo che tutte le radici di p siano reali [3]*. Allora:*
(i) 0 è una radice di p se e solo se $d \geq 1$ ed in tal caso è una radice di molteplicità esattamente d.
(ii) p ha tante radici positive, contate con relativa molteplicità, quante sono le variazioni di segno nella successione dei coefficienti non nulli di p.

[3] ad esempio, p è il polinomio caratteristico di una matrice reale simmetrica

7. Il caso $V = \mathbb{R}^n$ con prodotto scalare canonico.

Prima di trattare in maniera dettagliata $V = \mathbb{R}^n$ con il prodotto scalare canonico facciamo un' **osservazione generale**.
Sia $(V, <\,,\,>)$ uno spazio vettoriale metrico e sia b una forma bilineare simmetrica. A differenza del caso trattato nel Teorema di Sylvester ci viene ora assegnato *ab initio* un prodotto scalare. Quanto visto si applica senza alcun cambiamento anche a questo caso. Infatti: fissiamo una base ortonormale $\mathcal{B}$, consideriamo $A \equiv A_b^{\mathcal{B}}$ e definiamo T come sopra, tramite la (6). In particolare

$$M_{\mathcal{B},\mathcal{B}}(T) = A \equiv A_b^{\mathcal{B}}$$

che è simmetrica. Tutto procede come prima: l'operatore T è simmetrico, perché ha matrice associata in una base ortonormale che è una matrice simmetrica, ed otteniamo quindi una base ortonormale di autovettori, $\mathcal{W}$, ed una matrice $C = M_{\mathcal{B},\mathcal{W}}(\mathrm{Id}_V)$ che è ortogonale: $C \in O(n)$. Il resto dell'argomento è identico.
Fine **osservazione generale**.

Consideriamo quindi $V = \mathbb{R}^n$ con prodotto scalare canonico $\bullet$. Fissiamo la base canonica $\mathcal{E} = \{\underline{e}_1, \dots, \underline{e}_n\}$; in tal caso le coordinate di $\underline{x}$ rispetto a $\mathcal{E}$ sono proprio $\underline{x}$. Inoltre la base canonica è ortonormale rispetto a $\bullet$, come già osservato. Data una forma bilineare $b(\,,\,)$ otteniamo una matrice simmetrica A con $A := A_b^{\mathcal{E}}$ e si ha $b(\underline{x}, \underline{y}) = \underline{y}^T A \underline{x}$. Viceversa, abbiamo visto in classe che possiamo dare una forma bilineare simmetrica semplicemente assegnando una matrice simmetrica A e considerando

$$b_A(\underline{x}, \underline{y}) := \underline{y}^T A \underline{x}\,.$$

Conclusione: *le forme bilineari di $\mathbb{R}^n$ sono tutte e sole quelle del tipo $\underline{y}^T A \underline{x}$, con A simmetrica.*

Sia $b(\underline{x}, \underline{y}) = \underline{y}^T A \underline{x}$, $A = A^T$, una forma bilineare simmetrica. L'operatore T associato a $b(\,,\,)$ è in questo caso uguale a L_A (per costruzione, $T(\underline{e}_j) = L_A(\underline{e}_j)$ e questo basta). Abbiamo appena visto che una base **ortonormale** di autovettori $\mathcal{W}$ per L_A diagonalizza simultaneamente l'operatore L_A e la forma bilineare $b(\underline{x}, \underline{y}) = \underline{y}^T A \underline{x}$. Quindi $b(\,,\,)$ è definita positiva se e solo se gli autovalori di A sono tutti positivi. Analogamente si procede per definita negativa. Più in generale: se esiste un autovalore nullo la forma è degenere, se esistono autovalori non-nulli di segni discordi la forma è indefinita. Tutto si legge dalla matrice $A_b^{\mathcal{W}} = \Lambda$. Ancora una volta: *per determinare il segno degli autovalori senza calcolarli potete applicare il criterio di Cartesio al polinomio caratteristico di A e ciò è lecito perché sappiamo che quel polinomio reale ha tutte le radici reali.*

Forme bilineari simmetriche e forme quadratiche.
Una forma bilineare $b_A(\underline{x}, \underline{y}) = \underline{y}^T A \underline{x}$ in $\mathbb{R}^n$ definisce un polinomio omogeneo di grado 2 nelle variabili $x_1, \dots, x_n$; questo è il polinomio

$$q_A(\underline{x}) := \underline{x}^T A \underline{x} = \sum_j a_{jj} x_j^2 + \sum_{i<j} 2 a_{ij} x_i x_j\,.$$

Un tale polinomio è detto *una forma quadratica.* Viceversa, se $q(\underline{x})$ è una forma quadratica, cioè un polinomio omogeneo di grado 2 nelle variabili $\underline{x}_1, \ldots, \underline{x}_n$,

$$q(\underline{x}) = \sum_j \alpha_{jj} x_j^2 + \sum_{i<j} \alpha_{ij} x_i x_j \,,$$

allora possiamo definire una forma bilineare simmetrica considerando la matrice simmetrica A_q con coefficienti

$$a_{ii} = \alpha_{ii} \,, \quad a_{ij} = a_{ji} = \frac{\alpha_{ij}}{2} \quad \text{se} \quad i < j$$

e la forma bilineare $b_{A_q}(\underline{x}, \underline{y}) := \underline{y}^T A_q \underline{x}$. È allora chiaro che la forma bilineare $b_{A_q}(\,,\,)$ ha forma quadratica associata precisamente uguale a q; $b_{A_q}(\,,\,)$ è detta la forma bilineare simmetrica **polare** di q.

Esempio 1. In $\mathbb{R}^4$ la forma bilineare simmetrica polare della forma quadratica $x_1^2 - x_2^2 + 2x_3x_4$ è la forma bilineare simmetrica

$$b(\underline{x}, \underline{y}) = x_1y_1 - x_2y_2 + x_3y_4 + x_4y_3 \,.$$

Esempio 2. In $V = \mathbb{R}^4$ si consideri la forma quadratica definita da

$$q(\underline{x}) = 2x_1x_3 + 2x_2x_3.$$

La forma bilineare simmetrica polare di q è $b(\underline{x}, \underline{y}) = x_1y_3 + x_3y_1 + x_2y_3 + x_3y_2$.

Quindi *è equivalente parlare di forme quadratiche o di forme bilineari simmetriche*; spesso in alcuni testi si parla quindi di *diagonalizzazione delle forme quadratiche.*
Vi faccio notare che nelle coordinate $\underline{z}$ associate ad una base *ortonormale* di autovettori di A con autovalori

$$\{\lambda_1, \cdots, \lambda_\rho, -\mu_{\rho+1}, -\mu_r, 0, \cdots, 0\} \,, \quad \lambda_j > 0, \mu_\ell > 0$$

la forma quadratica q si "diagonalizza" e si scrive nella forma

$$\lambda_1 z_1^2 + \cdots + \lambda_\rho z_\rho^2 - \mu_{\rho+1} z_{\rho+1}^2 - \cdots - \mu_r z_r^2 \,, \quad \lambda_j > 0, \mu_\ell > 0$$

con r uguale al rango di A.
Infatti, se $C \in O(n)$ è la matrice del cambiamento di base, dalla base canonica ad una base ortonormale di autovettori per L_A allora,

$$q(\underline{x}) = \underline{x}^T A \underline{x} = (C\underline{z})^T A C \underline{z} = (\underline{z})^T (C^T A C) \underline{z} = \underline{z}^T \Lambda \underline{z}$$

e, come già visto,

$$\underline{z}^T \Lambda \underline{z} = \lambda_1 z_1^2 + \cdots + \lambda_\rho z_\rho^2 - \mu_{\rho+1} z_{\rho+1}^2 - \cdots - \mu_r z_r^2$$

Questa è la *forma canonica metrica* della forma quadratica; metrica perché è ottenuta facendo un cambiamento di basi ortonormali (da $\mathcal{E}$ alla base ortonormale di autovettori per L_A), cioè tramite una matrice che è ortogonale. Se vogliamo ammettere cambiamenti di base non-ortogonali allora possiamo ulteriormente modificare la base degli autovettori, come spiegato nella dimostrazione del teorema di Sylvester, riscalandoli opportunamente. Otteniamo in tal modo la *forma canonica affine* della forma quadratica:

$$\zeta_1^2 + \cdots + \zeta_\rho^2 - \zeta_{\rho+1}^2 - \cdots - \zeta_r^2 \,.$$

Un esempio di forma quadratica, e quindi di forma bilineare, è dato dal complesso dei termini di grado 2 nello sviluppo di Taylor di una funzione $F \in C^\infty(U)$, U aperto di $\mathbb{R}^n$, attorno ad un punto $(x_1^0, \ldots, x_n^0)$. I risultati che abbiamo presentato sono particolarmente importanti nello studio del compartamento locale di F in un intorno di un suo punto stazionario. Vedrete tutto ciò nel corso di Analisi ma registrate sin d'ora che quanto visto in queste note, ed in particolare il criterio di Cartesio, vi permetterà di decidere se la forma quadratica associata alla matrice Hessiana calcolata in un punto stazionario $(x_1^0, \ldots, x_n^0)$ è definita positiva/negativa e quindi se è possibile applicare il noto criterio sufficiente per decidere se un tale punto è un punto di minimo/massimo relativo. Analogamente, se scoprite che la forma quadratica associata alla matrice Hessiana è indefinita, allora potete affermare che il punto stazionario in esame è un punto di sella.

È bene osservare che esistono molte altre applicazioni del teorema di diagonalizzazione delle forme quadratiche, sia in Matematica che in Fisica.

Nei libri si parla spesso di *matrice simmetrica reale* definita positiva/negativa. Con ciò si sottintende la seguente

Definizione. Una *matrice simmetrica reale* A è definita positiva/negativa se la forma bilineare simmetrica $b_A(\underline{x}, \underline{y}) := \underline{y}^T A\underline{x}$, o equivalentemente la forma quadratica $q_A(\underline{x}) = \underline{x}^T A\underline{x}$, è definita positiva/negativa.

Abbiamo visto che il criterio di Cartesio ci permette di determinare il segno degli autovalori non-nulli del polinomio caratteristico senza calcolarli esplicitamente. Se il nostro obiettivo è di decidere se una matrice simmetrica è definita positiva/negativa, senza calcolare gli autovalori, allora possiamo anche utilizzare il seguente risultato [4].
Prima, una definizione. Sia A una qualsiasi matrice $n \times n$. Sia $1 \leq k \leq n$ e consideriamo la sottomatrice $k \times k$, denotata A_k, ottenuta intersecando le prime k righe e le prime k colonne. Poniamo $D_k := \det(A_k) \in \mathbb{R}$. D_k è il k-mo *minore principale* della matrice A.

Teorema. Sia $A \in M_{nn}(\mathbb{R})$ una matrice simmetrica.
1. A è definita positiva se e solo se $D_k > 0$ $\forall k \in \{1, \ldots, n\}$.
1. A è definita negativa se e solo se

$$D_1 < 0, \quad D_2 > 0, \quad D_3 < 0, \quad D_4 > 0 \ldots\ldots$$

Esercizio. (i) Determinare indici di positività, negatività e nullità di $b(\underline{x}, \underline{y}) = x_1y_3 + x_3y_1 + x_2y_3 + x_3y_2$.
(ii) Determinare la forma quadratica associata a b e la sua forma canonica metrica.
(iii) Determinare una base di Sylvester per $b(\ ,\ )$.

Soluzione esercizio. La matrice simmetrica associata alla forma bilineare nella base canonica è la matrice

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

[4] per la dimostrazione potete consultare Sernesi, *Geometria 1*

Consideriamo $\mathbb{R}^4$ con il prodotto scalare canonico; la base canonica è ortonormale rispetto a questo prodotto scalare. Come già osservato, l'operatore simmetrico associato a $b(\, , \,)$ nella base canonica è proprio L_A. Per rispondere alle due domande cominciamo con il determinare gli autovalori di L_A. Per determinare una base di Sylvester occorrerà inoltre determinare una base *ortonormale* di autovettori per L_A [5]. È subito visto che $P_A(\lambda) = \lambda^2(\lambda^2 - 2)$ e quindi A ammette gli autovalori $\lambda_1 = \sqrt{2}$, con $m_a(\sqrt{2}) = 1$, $\lambda_2 = -\sqrt{2}$, con $m_a(-\sqrt{2}) = 1$, $\lambda_3 = 0$, con $m_a(0) = 2$. Da Sylvester+Spettrale deduciamo che esiste una base ortonormale, chiamiamola $\mathcal{W}$, tale che

$$A_b^{\mathcal{W}} = \begin{vmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & -\sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix}$$

ed una base $\mathcal{F}$, ottenuta riscalando i vettori di $\mathcal{W}$, tale che

$$A_b^{\mathcal{F}} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix}$$

In particolare vediamo che b è degenere (perché il rango di A è 2 e non 4) e che è indefinita, perché $b(\underline{f}_1, \underline{f}_1) = 1 > 0$ e $b(\underline{f}_2, \underline{f}_2) = -1 < 0$.
La forma quadratica associata a b è la forma $q(\underline{x}) = 2x_1x_3 + 2x_2x_3$ ed ha forma canonica metrica $\sqrt{2}z_1^2 - \sqrt{2}z_2^2$.
Osservazione. in questo caso siamo riusciti a determinare esplicitamente gli autovalori; in generale questo non è possibile, ma dalla struttura del polinomio caratteristico e dal criterio di Cartesio saremo in grado di determinare comunque indici di positività, negatività e nullità. *Fine osservazione.*
Determiniamo ora una base ortonormale di autovettori. Si verifica senza difficoltà che $V_A(\sqrt{2}) = \mathbb{R}(1/\sqrt{2}, 1/\sqrt{2}, 1, 0)$, $V_A(-\sqrt{2}) = \mathbb{R}(-1/\sqrt{2}, -1/\sqrt{2}, 1, 0)$ e $V_A(0) = \mathrm{Span}((-1, 1, 0, 0), (0, 0, 0, 1))$.
Questi 4 vettori costituiscono una base ortogonale di autovettori (infatti, autovettori associati ad autovalori distinti sono ortogonali perché L_A è simmetrico; inoltre siamo stati bravi ed abbiamo direttamente scelto una base ortogonale all'interno di $V_A(0)$). Una base ortonormale di autovettori è quindi

$$\underline{w}_1 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}, 0)\,,\ \underline{w}_2 = (-\frac{1}{2}, -\frac{1}{2}, \frac{1}{\sqrt{2}}, 0)\,,\ \underline{w}_3 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0)\,,\ \underline{w}_4 = (0, 0, 0, 1)\,.$$

Denotiamo con $\mathcal{W}$ questa base ortonormale. Questa base diagonalizza simultaneamente l'operatore L_A e la forma bilineare $b(\underline{x}, \underline{y}) = \underline{y}^T A \underline{x}$ ed è la base la cui esistenza è stata enunciata sopra, nella prima parte della soluzione. La base di Sylvester $\mathcal{F}$ si ottiene semplicemente riscalando gli autovettori associati agli autovalori non-nulli.

[5] **Attenzione:** se prendete gli autovettori non-ortonormali allora diagonalizzate l'operatore L_A ma **non** diagonalizzate la forma bilineare $(\underline{x}, \underline{y}) \to \underline{y}^T A \underline{x}$