Corso di Laurea in Fisica. Geometria. a.a. 2020-21.
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 $\mathcal{B} := \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ 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 $\mathcal{B} = \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ è data da

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

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 $\underline{y}^T\, A_b^{\mathcal{B}}\, \underline{x}$.
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 fare la seguente utile osservazione (l'abbiamo già fatta nelle note sul Teorema Spettrale).
Osservazione. Se A e B sono due matrici in $M_{n\times n}(\mathbb{K})$, allora

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

In una direzione ($\Leftarrow$) è ovvio; per l'altra basta scegliere $\underline{y} = \underline{e}_j$ e $\underline{x} = \underline{e}_k$, con $\underline{e}_i$ l' i-mo vettore della base canonica di $\mathbb{K}^n$. Con questa particolare scelta è facile verificare che $\underline{y}^T A\underline{x}$ dà il coefficiente al posto (jk) della matrice A; dall'uguaglianza $\underline{x}^T A\underline{y} = \underline{x}^T B\underline{y}$ possiamo concludere che A e B hanno i coefficienti al posto (jk) uguali; dato che k e j erano arbitrari, possiamo concludere che le matrici sono uguali, $A = B$. **Fine Osservazione.**

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

$$\underline{y}^T A\underline{x} = \underline{x}^T\, A^T\, \underline{y}\ ;$$

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 *è simmetrica se e solo se* $A_b^{\mathcal{B}}$ *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 $\underline{x}^T A^T \underline{y} = b(\underline{v}, \underline{w})$ 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 $< , >$.

Viceversa, data una matrice $A = (a_{i\,j})$ ed una base $\mathcal{B}$ possiamo definire una forma bilineare $b_A^{\mathcal{B}} : V \times V \to \mathbb{K}$ ponendo

$$b_A^{\mathcal{B}}(x_1\underline{v}_1 + \cdots + x_n\underline{v}_n\,,\, y_1\underline{v}_1 + \cdots + y_n\underline{v}_n) := \underline{y}^T A\, \underline{x} \tag{3}$$

Brevemente:

$$b_A^{\mathcal{B}}(\underline{v}, \underline{w}) := \underline{y}^T A\, \underline{x} \tag{4}$$

se $\underline{v}$ ha coordinate $\underline{x}$ e $\underline{w}$ ha coordinate $\underline{y}$. È facile verificare che $b_A^{\mathcal{B}}$ è una forma bilineare. Ciò segue da tre osservazioni:
(i) il prodotto righe per colonne è distributivo rispetto alla somma;
(ii) l'operazione di trasposizioneè lineare;
(iii) l'applicazione $F_{\mathcal{B}} : V \to \mathbb{K}^n$ che associa ad un vettore le sue coordinate rispetto a $\mathcal{B}$ è un isomorfismo di spazi vettoriali [1].
Si ha quindi, ad esempio per l'additività nel primo argomento,

$$b_A^{\mathcal{B}}(\underline{v}+\underline{v}', \underline{w}) = (\underline{x}+\underline{x}')^T A\underline{y} = (\underline{x}^T+(\underline{x}')^T)A\underline{y} = \underline{x}^T A\underline{y}+(\underline{x}')^T A\underline{y} = b_A^{\mathcal{B}}(\underline{v}, \underline{w})+b_A^{\mathcal{B}}(\underline{v}', \underline{w}).$$

Analogamente si procede per dimostrare che

$$b_A^{\mathcal{B}}(\alpha\underline{v}, \underline{w}) = \alpha b_A^{\mathcal{B}}(\underline{v}, \underline{w})$$

La dimostrazione della linearità nel secondo argomento è identica.
Denoteremo a volte questa forma bilineare con b_A, omettendo quindi la base $\mathcal{B}$ dalla notazione.
Ragionando come prima vediamo che $b_A^{\mathcal{B}}$ è simmetrica se e solo se A è simmetrica.

Osservazione. Abbiamo quindi definito una forma bilineare $b_A^{\mathcal{B}}(\ ,\ )$ ed è immediato verificare che la matrice associata a $b_A^{\mathcal{B}}(\ ,\ )$ nella base $\mathcal{B}$ è proprio A. Analogamente, è chiaro che se prendiamo $A_b^{\mathcal{B}}$ e ne prendiamo la forma bilineare associata tramite (3) allora riotteniamo proprio b. [2]

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 $\mathcal{B}$. Non è difficile verificare la validità della seguente

---

[1] in particolare, le coordinate di $\underline{v} + \underline{v}'$ sono $\underline{x} + \underline{x}'$ e le coordinate di $\alpha\underline{v}$ sono $\alpha\underline{x}$.

[2] Fissata una base $\mathcal{B}$ esiste quindi una bigezione fra l'insieme delle forme bilineari su V, $\mathrm{Bil}(V)$, e l'insieme $M_{n\times n}(\mathbb{K})$. Infatti l'applicazione $\Phi_{\mathcal{B}} : \mathrm{Bil}(V) \to M_{n\times n}(\mathbb{K})$ definita da

$$\Phi_{\mathcal{B}}(b) := A_b^{\mathcal{B}}$$

ha inversa data da $\Psi_{\mathcal{B}} : M_{n\times n}(\mathbb{K}) \to \mathrm{Bil}(V)$

$$\Psi_{\mathcal{B}}(A) = b_A^{\mathcal{B}}\,.$$

Il fatto che $\Psi_{\mathcal{B}}(\Phi_{\mathcal{B}}(b)) = b$ e $\Phi_{\mathcal{B}}(\Psi_{\mathcal{B}}(A)) = A$ segue dall'ultima osservazione. (Di fatto $\mathrm{Bil}(V)$ ha una naturale struttura di spazio vettoriale e $\Phi_{\mathcal{B}}$ è un isomorfismo di spazi vettoriali.)

Proposizione. *L'immagine del nucleo di $b(,)$ tramite $F_{\mathcal{B}}$ è uguale al nucleo della matrice $A_b^{\mathcal{B}}$.*
Per la dimostrazione, vedere il libro di testo, Lemma 11.16.

In particolare: $b(,)$ *è non-degenere se e solo se* $A_b^{\mathcal{B}}$ *è non-singolare.*

2. Matrici congruenti

Sia $\mathcal{F} = \{\underline{f}_1, \dots, \underline{f}_n\}$ 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 $(x_1, \dots, x_n)$ nella base $\mathcal{B}$ e coordinate $(x'_1, \dots, x'_n)$ nella base $\mathcal{F}$. Analogamente $\underline{w}$ ha coordinate $(y_1, \dots, y_n)$ nella base $\mathcal{B}$ e coordinate $(y'_1, \dots, y'_n)$ 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,

$$A_b^{\mathcal{F}} = C^T A_b^{\mathcal{B}} C . \qquad \text{con } C = M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V) . \tag{5}$$

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 *le matrici associate ad una forma bilineare in due basi diverse sono* **congruenti**.

Vale anche il *viceversa*: siano A e D due matrici congruenti, $D = C^T AC$, con C invertibile. Sia $b := b_A^{\mathcal{B}}$ la forma bilineare definita dalla formula (3) nella base $\mathcal{B}$. Sia $\mathcal{F}$ la base che ha come j-mo vettore quello che ha coordinate, nella base $\mathcal{B}$, uguali alla j-ma colonna di C; in formule

$$\underline{f}_j := \sum_k c_{kj} \underline{v}_k .$$

Abbiamo quindi scelto $\mathcal{F}$ in modo tale che $C = M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$. Sia $b_D^{\mathcal{F}}$ la forma bilineare simmetrica definita dalla formula (3) (ma, ovviamente, nelle coordinate associate alla base $\mathcal{F}$). Si ha allora

$$b_A^{\mathcal{B}}(\underline{v}, \underline{w}) = \underline{x}^T A \, \underline{y} = (C\underline{x}')^T A (C\underline{y}') = \underline{x}' D \underline{y}' = b_D^{\mathcal{F}}(\underline{v}, \underline{w})$$

e quindi $b_A^{\mathcal{B}}$ e $b_D^{\mathcal{F}}$ definiscono la stessa forma bilineare: $b_A^{\mathcal{B}} = b_D^{\mathcal{F}}$.
In conclusione abbiamo dimostrato la seguente :

Proposizione. *A e D sono congruenti se e solo se rappresentano la stessa forma bilineare in basi diverse.*

Questa Proposizione è l'analoga di quella che abbiamo visto per gli endomorfismi di V:

due matrici A e D sono simili se e solo se rappresentano lo stesso endomorfismo in basi diverse.

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). 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.**

Esistenza di un prodotto scalare definito positivo. Abbiamo definito uno spazio vettoriale metrico come uno spazio vettoriale dotato di un prodotto scalare $<\ ,\ >$ definito positivo.
Domanda: ma i prodotti scalari definiti positivi esistono sempre ?
La risposta è affermativa e segue da quanto appena visto; vediamo i dettagli. Fissiamo una qualsiasi base $\mathcal{B}$ di V e la matrice identità I_n che è certamente simmetrica; consideriamo la forma bilineare (3) con $A = I_n$. Otteniamo una forma bilineare simmetrica $b_{I_n}^{\mathcal{B}}$ che è chiaramente definita positiva, perché se $\underline{v}$ è un vettore non-nullo di coordinate $\underline{x}$ nella base $\mathcal{B}$ allora

$$b_{I_n}^{\mathcal{B}}(\underline{v}, \underline{v}) = \underline{x}^T\, I_n\, \underline{x} = \underline{x}^T\, \underline{x} = (x_1)^2 + \cdots + (x_n)^2 > 0.$$

Per costruzione i vettori della base $\mathcal{B}$ sono ortonormali rispetto a questo prodotto scalare definito positivo; infatti le coordinate dei vettori $\underline{v}_i$ e $\underline{v}_j$ nella base $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ sono dati dai vettori della base canonica $\underline{e}_i$, $\underline{e}_j$. Quindi

$$b_{I_n}^{\mathcal{B}}(\underline{v}_i, \underline{v}_j) = \underline{e}_j^T\, \underline{e}_i$$

che è uguale a 1 se $i = j$ e 0 se $i \neq j$, come si voleva.
In parole: *abbiamo introdotto un prodotto scalare definito positivo in V dichiarando ortonormali i vettori della base $\mathcal{B}$ che avevamo precedentemente fissato ed estendendo per bilinearità.*

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 (5) ci dice che

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

e 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

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

Abbiamo quindi dimostrato una direzione della seguente

Proposizione. Sia $\mathcal{B}$ una base ortonormale di $(V, <\,,\,>)$. Sia $\mathcal{F}$ una seconda base. Sia B la matrice del cambiamento di base, da $\mathcal{B}$ a $\mathcal{F}$: $B = M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$. Allora $\mathcal{F}$ è ortonormale se e solo se $B^T B = I_n$
Dimostrazione. Se $\mathcal{F}$ è ortonormale, allora abbiamo appena visto che $B^T B = I_n$. Supponiamo ora che $B = M_{\mathcal{B},\mathcal{F}}(\mathrm{Id}_V)$ verifichi questa equazione. Quindi il prodotto della i-ma riga di B^T con la j-ma colonna di B è uguale a δ_{ij}, con $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$. Ma la i-ma riga di B^T è per definizione uguale alla i-ma colonna di B. Quindi

$$\begin{vmatrix} b_{1i} & b_{2i} & \cdots & b_{ni} \end{vmatrix} \cdot \begin{vmatrix} b_{1j} \\ b_{2j} \\ \cdots \\ b_{nj} \end{vmatrix} = \delta_{ij}$$

che scriviamo in forma compatta come

$$(B^i)^T B^j = \delta_{ij}$$

Possiamo ovviamente riscrivere questa espressione come

$$(B^i)^T\, I_n\, B^j = \delta_{ij}$$

Ora, a sinistra c'e' l'espressione nelle coordinate associate a $\mathcal{B}$ di $< \underline{f}_j, \underline{f}_i >$, perché la k-ma colonna di B è data, appunto, dalle coordinate di $\underline{f}_k$ nella base $\mathcal{B}$ e, inoltre, $A^{<\,,\,>}_{\mathcal{B}} = I_n$ perché $\mathcal{B}$ è ortonormale. Riassumendo,

$$(6) \qquad < \underline{f}_i, \underline{f}_j >= 0 \quad \text{se} \quad i \neq j\,; \quad < \underline{f}_i, \underline{f}_i >= 1\,.$$

La Proposizione è dimostrata.

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*. Notiamo che $O(n) = \{A \in GL_n(\mathbb{R}) \,|\, A^{-1} = A^T\}$, infatti $\det A = \pm 1$ (dato che $\det A^2 = 1$) da cui $O(n) \subset GL_n(\mathbb{R})$; quindi le matrici ortogonali sono invertibili. Inoltre essendo per definizione $A^T A = I_n$ si deve avere, moltiplicando a destra ambo i membri per A^{-1}, $A^T = A^{-1}$ che è quello che avevamo enunciato.
Possiamo rienunciare la Proposizione precedente nel modo seguente:

Proposizione. Sia $\mathcal{B}$ una base ortonormale di $(V, <\,,\,>)$. Sia $\mathcal{F}$ una seconda base. Sia B la matrice del cambiamento di base, da $\mathcal{B}$ a $\mathcal{F}$. Allora $\mathcal{F}$ è ortonormale se e solo se $B \in O(n)$.

Proposizione 4. $O(n)$ con il prodotto righe per colonne è un gruppo, 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 $SO(n) = \{A \in O(n) \mid \det A = 1\}$ è anche un gruppo (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 $SO(n)$) detto *gruppo ortogonale speciale.*

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

$$O(2) = \{ \left| \begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array} \right| , \theta \in \mathbb{R}\} \cup \{ \left| \begin{array}{cc} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{array} \right| , \theta \in \mathbb{R}\}$$

$$SO(2) = \{ \left| \begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array} \right| , \theta \in \mathbb{R}\}$$

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

$$A_b^{\mathcal{F}} = \left| \begin{array}{ccc} I_\rho & 0 & 0 \\ 0 & -I_{r-\rho} & 0 \\ 0 & 0 & 0_{n-r} \end{array} \right| .$$

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 $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ 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, si veda l'inizio della Sezione 3. Abbiamo quindi un prodotto scalare definito positivo $<\,,\,>$ e la base $\mathcal{B}$ è ortonormale per $<\,,\,>$.
Consideriamo l'operatore $T : V \to V$ definito da

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

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 $(V, < \, , \, >)$ [3].

Abbiamo quindi un operatore simmetrico in uno spazio vettoriale metrico; per il teorema spettrale esiste una base *ortonormale* $\mathcal{W} = \{\underline{w}_1, \ldots, \underline{w}_n\}$ costituita da autovettori per T. Siano $\lambda_1, \ldots, \lambda_\rho \in \mathbb{R}$ gli autovalori positivi di T. Ci sono allora $r - \rho$ autovalori negativi $\lambda_{\rho+1}, \ldots, \lambda_r$; 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,

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

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

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

Dato che $C \in O(n)$ si ha allora la cruciale relazione:

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

Ora, da quanto visto nella sezione precedente,

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

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 $k \neq \ell$

[3] 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.

$b(\underline{f}_j, \underline{f}_j) = 1$ 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 $\dim U + \dim W = \rho + (n - \sigma) > n$ 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 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.

Osservazione. Il teorema di Sylvester può anche essere dimostrato direttamente, senza utilizzare il teorema spettrale. Vedere i complementi alla fine di queste note.

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 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. 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* [4]*. 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*.*

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. Fissiamo una base ortonormale $\mathcal{B}$ e definiamo T come sopra. In particolare $M_{\mathcal{B},\mathcal{B}}(T) = A \equiv A_b^{\mathcal{B}}$. Otteniamo una base ortonormale di autovettori, $\mathcal{W}$, ed una matrice $C = M_{\mathcal{B},\mathcal{W}}(\mathrm{Id}_V)$ che è ortogonale: $C \in O(n)$. Sappiamo che

$$\Lambda = C^{-1}AC = C^T AC$$

con Λ la matrice diagonale che ha gli autovalori di T sulla diagonale. In particolare,

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

che è diagonale.

Consideriamo ora $V = \mathbb{R}^n$. Fissiamo la base canonica $\mathcal{E} = \{\underline{e}_1, \ldots, \underline{e}_n\}$; in tal caso le coordinate di $\underline{x}$ rispetto a $\mathcal{E}$ sono proprio $\underline{x}$. 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, sappiamo 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.*

Consideriamo quindi $b(\underline{x}, \underline{y}) = \underline{y}^T A\underline{x}$, con $A = A^T$. Consideriamo ora il prodotto scalare canonico in $\mathbb{R}^n$. $\mathbb{R}^n$ diviene uno spazio vettoriale metrico e la base canonica $\mathcal{E}$ è ortonormale. 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 (ovvio). 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.*

---

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

Forme bilineari simmetriche e forme quadratiche.
Una forma bilineare in $\mathbb{R}^n$ definisce un polinomio omogeneo di grado 2 nelle variabili $x_1, \ldots, x_n$; questo è il polinomio

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

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

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

allora possiamo definire una forma bilineare considerando la matrice simmetrica A_ϕ 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 $b_{A_\phi}(\underline{x}, \underline{y}) := \underline{y}^T A_\phi \underline{x}$. È allora chiaro che la forma bilineare $b_{A_\phi}(\,,\,)$ ha forma quadratica associata precisamente uguale a ϕ; $b_{A_\phi}(\,,\,)$ è detta forma **polare** di ϕ.

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

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

La forma bilineare simmetrica polare di ϕ 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 ϕ 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. Questa forma canonica, detta **forma canonica metrica della forma quadratica**, gioca un ruolo fondamentale in molte questioni di Matematica e Fisica

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 una base di Sylvester per $b(\,,\,)$.

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

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

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 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 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.
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.

7. Diagonalizzare le forme bilineari simmetriche senza il teorema spettrale

Questa sezione è parzialmente facoltativa. È interessante osservare che il teorema di Sylvester può essere dimostrato senza far uso del teorema spettrale. Questo approccio, che è quello originale di Sylvester, permette di costruire induttivamente

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

una base di Sylvester senza calcolare alcun autovalore e senza calcolare alcun autospazio.
Il passo fondamentale è il seguente:

Teorema *Sia* $b : V \times V \to \mathbb{R}$ *una forma bilineare simmetrica. Allora esiste una base di* V, $\{\underline{f}_1, \ldots, \underline{f}_n\}$, *che diagonalizza* $b(\ ,\ )$, *tale cioè che* $b(\underline{f}_i, \underline{f}_j) = 0$ *per* $i \neq j$.

Prima di passare alla dimostrazione, premettiamo una definizione ed un' *osservazione fondamentale.*
Definizione: un vettore $\underline{v}$ è **isotropo** per $b(\ ,\ )$ se $b(\underline{v}, \underline{v}) = 0$.
Osserviamo che se $\underline{f}$ è un vettore *non isotropo* di V allora vale la decomposizione

$$V = \mathbb{R}\underline{f} \oplus (\mathbb{R}\underline{f})^{\perp_b} \tag{7}$$

dove

$$(\mathbb{R}\underline{f})^{\perp_b} = \{\underline{v} \in V \mid b(\underline{v}, \underline{f}) = 0\,\}\,.$$

In generale, per un qualsiasi sottospazio U di V si pone

$$U^{\perp_b} = \{\underline{v} \in V \mid b(\underline{v}, \underline{u}) = 0\,,\, \forall \underline{u} \in U\}\,.$$

$U^{\perp_b}$ è detto il b-ortogonale di U. La dimostrazione della decomposizione (7) utilizza una tecnica già vista per i prodotti scalari definiti positivi. Vediamo.
Dimostrazione di (7) (facoltativa). Per ipotesi $\underline{f}$ è non isotropo; quindi, per definizione, $b(\underline{f}, \underline{f}) \neq 0$. Scriviamo allora

$$\underline{v} = \frac{b(\underline{v}, \underline{f})}{b(\underline{f}, \underline{f})}\underline{f} \,+\, (\underline{v} - \frac{b(\underline{v}, \underline{f})}{b(\underline{f}, \underline{f})}\underline{f})$$

Il primo addendo a destra appartiene sicuramente a $\mathbb{R}\underline{f}$. Verifichiamo che il secondo addendo appartiene a $(\mathbb{R}\underline{f})^{\perp_b}$; utilizzando la bilinearità abbiamo:

$$b(\underline{v} - \frac{b(\underline{v}, \underline{f})}{b(\underline{f}, \underline{f})}\underline{f}\,,\, \underline{f}) = b(\underline{v}, \underline{f}) - b(\frac{b(\underline{v}, \underline{f})}{b(\underline{f}, \underline{f})}\underline{f}, \underline{f}) = b(\underline{v}, \underline{f}) - (\frac{b(\underline{v}, \underline{f})}{b(\underline{f}, \underline{f})})b(\underline{f}, \underline{f}) = 0$$

Quindi

$$V = \mathbb{R}\underline{f} + (\mathbb{R}\underline{f})^{\perp_b}$$

Verifichiamo anche che $\mathbb{R}\underline{f} \cap (\mathbb{R}\underline{f})^{\perp_b} = \underline{0}$; se $\underline{w} \in \mathbb{R}\underline{f} \cap (\mathbb{R}\underline{f})^{\perp_b}$ allora $\underline{w} = \alpha\underline{f}$ (perché $\underline{w} \in \mathbb{R}\underline{f}$) **e** $b(\underline{w}, \underline{f}) = 0$ (perché $\underline{w} \in (\mathbb{R}\underline{f})^{\perp_b}$). Ma allora deve essere $\alpha b(\underline{f}, \underline{f}) = 0$ ed essendo $\underline{f}$ non isotropo deve necessariamente essere $\alpha = 0$ cioè $\underline{w} = \underline{0}$. Abbiamo dimostrato che $V = \mathbb{R}\underline{f} \oplus (\mathbb{R}\underline{f})^{\perp_b}$.
Fine osservazione fondamentale.

Dimostrazione teorema (facoltativa). Procediamo per induzione su $\dim V$. Se $\dim V = 1$ non c'è nulla da dimostrare. Supponiamo vero il teorema per spazi vettoriali di dimensione $n-1$ e dimostriamolo per spazi vettoriali di dimensione n. Se $b(\ ,\ )$ è la forma bilineare identicamente uguale a zero, allora non c'è nulla da dimostrare, dato che ogni base è diagonalizzante.
Se $b(\ ,\ )$ non è identicamente nulla allora $\exists\, \underline{v}\,,\, \underline{w} \in V$ non nulli tali che $b(\underline{v}, \underline{w}) \neq 0$. Ora, fra i tre vettori $\underline{v}, \underline{w}, \underline{v} + \underline{w}$ ne esiste almeno uno che è *non isotropo.* Infatti se $\underline{v}$ e $\underline{w}$ sono isotropi allora

$$\begin{aligned} b(\underline{v} + \underline{w}, \underline{v} + \underline{w}) &= b(\underline{v}, \underline{v}) + b(\underline{v}, \underline{w}) + b(\underline{w}, \underline{v}) + b(\underline{w}, \underline{w}) = \\ & b(\underline{v}, \underline{v}) + 2b(\underline{v}, \underline{w}) + b(\underline{w}, \underline{w}) = 0 + 2b(\underline{v}, \underline{w}) + 0 \neq 0 \end{aligned}$$

come si voleva. Riassumendo: se $b(\, , \,)$ non è identicamente nulla allora esiste un vettore non isotropo $\underline{f}_1$. Ma allora, per l'*osservazione fondamentale*

$$V = \mathbb{R}\underline{f}_1 \oplus (\mathbb{R}\underline{f}_1)^{\perp_b} .$$

Sia $b'(\, , \,)$ la restrizione di $b(\, , \,)$ al sottospazio $(n-1)$-dimensionale $(\mathbb{R}\underline{f}_1)^{\perp_b}$. $b'(\, , \,)$ è ovviamente una forma bilineare simmetrica su uno spazio vettoriale di dimensione $(n-1)$. Per ipotesi induttiva esiste una base $\{\underline{f}_2, \ldots, \underline{f}_n\}$ che diagonalizza $b'(\, , \,)$. Ma allora è immediato verificare che $\{\underline{f}_1, \underline{f}_2, \ldots, \underline{f}_n\}$ è una base diagonalizzante per $b(\, , \,)$. Il teorema è dimostrato.

Quindi possiamo diagonalizzare le forme bilineari simmetriche senza utilizzare il teorema spettrale. Sia $\{\underline{f}_1, \ldots, \underline{f}_n\}$ la base diagonalizzante ottenuta tramite il procedimento induttivo appena spiegato. Se $\alpha_j = b(\underline{f}_j, \underline{f}_j)$ allora possiamo assumere, a meno di riordinare i vettori,

$\alpha_i > 0$ se $i \leq \rho_+$
$\alpha_i < 0$ se $(\rho_+ + 1) \leq i \leq (\rho_+ + \rho_-)$
$\alpha_i = 0$ se $((\rho_+ + \rho_-) + 1) \leq i \leq (\rho_+ + \rho_- + \rho_0) = n$.

Poniamo
$\underline{v}_i = \frac{1}{\sqrt{\alpha_i}}\underline{f}_i$ se $i \leq \rho_+$
$\underline{v}_i = \frac{1}{\sqrt{-\alpha_i}}\underline{f}_i$ se $(\rho_+ + 1) \leq i \leq (\rho_+ + \rho_-)$
$\underline{v}_i = \underline{f}_i$ se $((\rho_+ + \rho_-) + 1) \leq i \leq (\rho_+ + \rho_- + \rho_0)$.

La matrice associata a $b(\, , \,)$ in questa base è

$$\begin{vmatrix} I_{\rho_+} & 0 & 0 \\ 0 & -I_{\rho_-} & 0 \\ 0 & 0 & 0_{\rho_0} \end{vmatrix} .$$

Per gli interi ρ_+, ρ_- e ρ_0 possiamo ragionare come segue: ρ_0 deve essere $n - r$ perché questa è in particolare una matrice associata a $b(\, , \,)$ e quindi deve avere rango r. Quindi $\rho_- = r - \rho_+$ e ragionando come nella dimostrazione originale del Teorema di Sylvester (quella con il Teorema Spettrale) possiamo dimostrare, procedendo per assurdo, che ρ_+ dipende solo da $b(\, , \,)$.

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 una base di Sylvester per $b(\, , \,)$.
Soluzione. Abbiamo già risolto l'esercizio utilizzando il teorema spettrale. Vediamo ora come risolverlo *à la* Sylvester.
Il metodo si basa sulla costruzione induttiva di una base diagonalizzante, prendendo in esame vettori non-isotropi. Ovviamente i vettori nel nucleo di $b(\, , \,)$, $V^{\perp_b}$, non sono mai non-isotropi, dato che sono b-ortogonali a *tutti* i vettori di V. Il primo passo in questa soluzione dell'esercizio consiste quindi nel trovare il nucleo di $b(\, , \,)$ e nel fissare una base di questo nucleo; gli altri vettori della base di Sylvester, quelli non-isotropi, andranno cercati *fuori* dal nucleo. Sappiamo che il nucleo $V^{\perp_b}$ è dato

da Ker$A$, con

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

la matrice associata a $b(\, , \,)$ nella base canonica. Chiaramente Ker$A$ ha equazioni cartesiane

$$\begin{cases} x_3 = 0 \\ x_1 + x_2 = 0 \end{cases}$$

e quindi

$$\mathrm{Ker}A = \mathrm{Span}((-1, 1, 0, 0), (0, 0, 0, 1)).$$

Poniamo $\underline{g}_3 = (-1, 1, 0, 0)\,, \underline{g}_4 = (0, 0, 0, 1)$ [6]. Che questi vettori siano ortogonali nel prodotto scalare canonico (quello definito positivo) non ci interessa; in questa soluzione non si utilizzano nozioni metriche. A questo punto possiamo iniziare il procedimento induttivo, fissando un vettore non-isotropo fuori da Ker$A$. Il vettore $(1, 1, 1, 0)$ non appartiene al nucleo, perché non ne soddisfa le equazioni, ed è non isotropo, perché

$$b((1, 1, 1, 0), (1, 1, 1, 0)) = 4 \neq 0$$

Poniamo $\underline{g}_1 = (1, 1, 1, 0)$. Ora cerchiamo $\underline{g}_2$ non-isotropo in $(\mathbb{R}\underline{g}_1)^{\perp_b}$. Si ha,

$$(\mathbb{R}\underline{g}_1)^{\perp_b} = \{\underline{x} \in \mathbb{R}^4 \mid b((\underline{x}, \underline{g}_1,) = 0\} = \{\underline{x} \in \mathbb{R}^4 \mid b(\underline{x}, (1, 1, 1, 0)) = 0\} =$$

$$\{\underline{x} \in \mathbb{R}^4 \mid (1, 1, 1, 0)^T \cdot A \cdot \underline{x} = 0\} = \{\underline{x} \in \mathbb{R}^4 \mid (1, 1, 1, 0)^T \cdot \begin{vmatrix} x_3 \\ x_3 \\ x_1 + x_2 \\ 0 \end{vmatrix} = 0\}$$

Quindi $(\mathbb{R}\underline{g}_1)^{\perp_b}$ è costituito dai vettori $\underline{x} \in \mathbb{R}^4$ che soddisano l' equazione

$$2x_3 + x_1 + x_2 = 0$$

Il vettore $(1, 1, -1, 0)$ soddisfa quest'equazione; inoltre non soddisfa le equazioni cartesiane del nucleo e quindi non appartiene al nucleo; infine, non è un vettore isotropo perché

$$b((1, 1, -1, 0), (1, 1, -1, 0)) = -4 \neq 0\,.$$

Conclusione: la base

$$\mathcal{G} = \{\underline{g}_1, \underline{g}_2, \underline{g}_3, \underline{g}_4\}$$

è una base diagonalizzante, con matrice associata

$$A_b^{\mathcal{G}} = \begin{vmatrix} 4 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix}$$

Una base di Sylvester è data da

$$\mathcal{F} = \{\underline{f}_1 := \frac{1}{2}\underline{g}_1, \underline{f}_2 := \frac{1}{2}\underline{g}_2, \underline{f}_3 := \underline{g}_3, \underline{f}_4 := \underline{g}_4\}$$

---

[6] li ordiniamo cosi' perché usualmente i vettori del nucleo sono messi in ultima pozione

ed ha matrice associata

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

come deve essere.