Corso di Laurea in Fisica. Geometria. a.a. 2013-14.
Prof. P. Piazza
Diagonalizzazione delle forme bilineari simmetriche

1. Osservazioni preliminari.

Sia V uno spazio vettoriale reale. Sia

$$< \, , \, >: V \times V \to \mathbb{R}$$

un prodotto scalare; per definizione questo vuol dire che $< \, , \, >$ è una forma bilineare simmetrica. Per ragioni tipografiche in queste note preferisco utilizzare la notazione $b : V \times V \to \mathbb{R}$ per un arbitrario prodotto scalare $< \, , \, >: V \times V \to \mathbb{R}$: quindi $b(\, , \,) := < \, , \, >$.

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 *simmetrica*

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

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

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

Questa formula segue facilmente dalla bilinearità; quindi se $\underline{v}$ ha coordinate $\underline{x}$ nella base $\mathcal{B}$ e $\underline{w}$ ha coordinate $\underline{y}$ nella base $\mathcal{B}$ allora $b(\underline{v}, \underline{w}) = \underline{y}^T \, A \, \underline{x}$.
Chiamiamo il membro a destra della (1), *l'espressione in coordinate della forma bilineare* $b(\, , \,)$ *nella base* $\mathcal{B}$.

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

$$(2) \quad b_A^{\mathcal{B}}(x_1\underline{v}_1 + \cdots + x_n\underline{v}_n \, , \, y_1\underline{v}_1 + \cdots + y_n\underline{v}_n) := |y_1, \dots, y_n| \, A \begin{vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{vmatrix} = \underline{y}^T \, A \, \underline{x}$$

Equivalentemente, definiamo la forma bilineare simmetrica $b_A^{\mathcal{B}}$ sulla base $\mathcal{B}$ ponendo $b_A^{\mathcal{B}}(\underline{v}_i, \underline{v}_j) := a_{ij}$ e poi estendiamo per bilinearità a tutti i vettori $\underline{v}$, $\underline{w}$ di V.
Denoteremo quasi sempre questa forma bilineare simmerica con b_A, omettendo quindi la base $\mathcal{B}$ dalla notazione. La verifica che l' applicazione (2) è effettivamente una forma bilineare simmetrica è molto semplice ma per completezza la riporto.
Per la simmetria potete fare la verifica in maniera brutale oppure ragionare sul fatto che $\underline{y}^T \, A \, \underline{x}$ è una matrice 1×1 e che quindi $\underline{y}^T \, A \, \underline{x} = (\underline{y}^T \, A \, \underline{x})^T$ da cui

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

come si voleva. Nell'ultimo passaggio abbiamo usato l'ipotesi ($A = A^T$) ed il fatto generale che la trasposta della trasposta di una matrice è la matrice di partenza.
La bilinearità segue da due osservazioni elementari:

1) se $\underline{v}$ ha coordinate $\underline{x}$ e $\underline{v}'$ ha coordinate $\underline{x}'$ allora $\underline{v}+\underline{v}'$ ha coordinate $\underline{x}+\underline{x}'$ e $\lambda\underline{v}$ ha coordinate $\lambda\underline{x}$, $\lambda \in \mathbb{R}$.
2) le proprietà del prodotto righe per colonne, si veda la Proposizione 7.5.
Lascio a voi i dettagli che sono a questo punto banali.

Abbiamo quindi definito una forma bilineare $b_A^{\mathcal{B}}(\ ,\ )$ ed è ovvio che la matrice associata a $b_A^{\mathcal{B}}(\ ,\ )$ nella base fissata $\mathcal{B}$ è proprio A.

Osservazione. (Esistenza di un prodotto scalare definito positivo). Abbiamo definito uno spazio vettoriale metrico come uno spazio vettoriale dotato di un prodotto scalare $<\ ,\ >$ definito positivo, tale cioè che $<\underline{v},\underline{v}>\,> 0$ per ogni $\underline{v}$ non-nullo. Domanda: ma i prodotti scalari definiti positivi esistono sempre ? La risposta segue da quanto appena visto; fissiamo una qualsiasi base $\mathcal{B}$ di V e la matrice identità I_n che è certamente simmetrica; consideriamo la forma bilineare (2) 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

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

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.

Osservazione. Vi ricordo [1] che il nucleo di una forma bilineare simmetrica è il sottospazio 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 che l'immagine del nucleo di $b(\ ,\ )$ tramite $F_{\mathcal{B}}$ è proprio in nucleo della matrice $A_b^{\mathcal{B}}$ (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. (Il teorema di Sylvester, che ci accingiamo ad enunciare e dimostrare, darà informazioni molto più dettagliate su $b(\ ,\ )$ a partire da $A_b^{\mathcal{B}}$.)

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 $\mathcal{B}$; 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

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

[1] Definizione 11.5 nel libro di testo

da cui, infine,

$$A_b^{\mathcal{F}} = C^T A_b^{\mathcal{B}} C \,.$$

Quest'ultimo passaggio è vero perché possiamo scegliere $\underline{x}', \underline{y}'$ arbitrariamente; se scegliete $\underline{x}' = \underline{e}_j$ e $\underline{y}' = \underline{e}_k$, con $\underline{e}_i$ l' i-mo vettore della base canonica di $\mathbb{R}^n$, potete concludere che queste due matrici hanno i coefficienti al posto (kj) uguali; dato che k e j erano arbitrari, potete concludere che le matrici sono uguali.

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 simmetrica in due basi diverse sono* **congruenti**.

Viceversa, siano ora A e D due matrici *simmetriche* congruenti, $D = C^T AC$, con C invertibile. Sia $b_A^{\mathcal{B}}$ la forma bilineare simmetrica definita dalla formula (2) 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

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

sia $b_D^{\mathcal{F}}$ la forma bilineare simmetrica definita dalla formula (2) ma nelle coordinate associate alla base $\mathcal{F}$. Si ha allora

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

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

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.

Possiamo allora definire il rango di una forma bilineare $b(\, , \,)$ come il rango di $A_b^{\mathcal{B}}$ con $\mathcal{B}$ una qualsiasi base di V. La definizione è ben posta per la Proposizione. Inoltre, per quanto abbiamo detto circa il nucleo di $b(\, , \,)$ abbiamo che la dimensione del nucleo è precisamente uguale a $\dim V - \mathrm{rango}(b)$.

3. Il teorema di Sylvester

In questa sezione utilizzeremo il teorema spettrale per dimostrare il seguente importante

Teorema (Sylvester) *Sia $b(\, , \,)$ una forma bilineare simmetrica su uno spazio vettoriale 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*

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

I numeri ρ, $r-\rho$ e $n-r$ sono detti rispetttivamente, indice di positività, indice di negatività ed indice di nullità della forma bilineare simmetrica $b(\, , \,)$

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 (vedere la prima Osservazione della pagina 2). 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 $\mathcal{B}$ è proprio A. In particolare T ha un nucleo di dimensione $n-r$. 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 $(V, < \, , \, >)$.

Verifichiamolo: se $\underline{v}$ ha coordinate $\underline{x}$ in $\mathcal{B}$ e $\underline{w}$ ha coordinate $\underline{y}$ in $\mathcal{B}$ allora

$$< T\underline{v}, \underline{w} > = < \sum_j x_j T(\underline{v}_j), \sum_\ell y_\ell \underline{v}_\ell > = \sum_{j,\ell} x_j y_\ell < T(\underline{v}_j), \underline{v}_\ell >$$

Ma dalla definizione di T e l'ortonormalità di $\mathcal{B}$ segue che $< T(\underline{v}_j), \underline{v}_\ell > = a_{\ell j}$. Quindi $< T\underline{v}, \underline{w} > = \sum_{j,\ell} x_j y_\ell a_{\ell j}$. Scrivendo l'analoga espressione per $< \underline{v}, T\underline{w} >$ e sfruttando il fatto che $A = A^T$ vediamo che $< T\underline{v}, \underline{w} > = < \underline{v}, T\underline{w} >$. Lascio a voi i conti perché sono a questo punto elementari.

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 $\mathcal{B}$. Sappiamo che C è ortogonale, $C \in O(n)$, perché entrambe le basi sono ortonormali. Inoltre, $C^{-1}AC =$

Λ con Λ la matrice diagonale che ha gli autovalori sulla diagonale principale. Dato che $C \in O(n)$ si ha allora l'importante relazione:

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

Ora, da quanto visto nella sezione precedente,

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

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

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

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 $U = \text{Span}(\underline{f}_1, \ldots, \underline{f}_\rho)$ e sia $W = \text{Span}(\underline{a}_{\sigma+1}, \ldots, \underline{a}_n)$. 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.

4. 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.

Importante: per capire il segno degli autovalori (senza calcolarli !) potete applicare il magico *criterio di Cartesio*, che trovate nel libro di testo (Teorema 14.8; notate che il polinomio caratteristico di una matrice simmetrica ha tutte le radici reali, quindi le ipotesi del criterio di Cartesio sono effettivamente soddisfatte).

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.

Il caso $V = \mathbb{R}^n$. Consideriamo in particolare $V = \mathbb{R}^n$. Per quanto appena visto, data una forma bilineare $b(\ ,\ )$ otteniamo una matrice simmetrica fissando una base. Fissiamo allora 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 $\underline{x}$ e quindi $b(\underline{x}, \underline{y}) = \underline{y}^T A \underline{x}$, con $A = A_b^{\mathcal{E}}$. Viceversa possiamo dare una forma bilineare simmetrica semplicemente assegnando una matrice simmetrica A e considerando

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

ed è ovvio che la matrice associata a questa forma bilineare nella base canonica è proprio A.
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 allora $b(\underline{x}, \underline{y}) = \underline{y}^T A \underline{x}$, con $A = A^T$.

Osservazione: 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

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

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. 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

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

Quindi *è equivalente parlare di forme quadratiche o di forme bilineari simmetriche*; spesso in alcuni testi si parla quindi di *diagonalizzazione delle forme quadratiche.*

Fissiamo ora il prodotto scalare canonico. $\mathbb{R}^n$ diviene uno spazio vettoriale metrico e la base canonica $\mathcal{E}$ è ortonormale. Scegliamo questa come base iniziale, quindi $\mathcal{B}$ nella dimostrazione del teorema di Sylvester è proprio $\mathcal{E}$. L'operatore T in questo caso è uguale a L_A (ovvio); da quanto visto $b(\ ,\ )$ 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. Ancora una volta: *per determinare il segno degli autovalori potete applicare il criterio di Cartesio al polinomio caratteristico di A.*

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 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 critico. Vedrete tutto ciò nel corso di Analisi.

Altre applicazioni della forma canonica metrica ed affine sono alla classificazione delle coniche e delle quadriche nel piano e nello spazio affine e nel piano e nello spazio euclideo. La diagonalizzazione delle forme bilineari simmetriche gioca un ruolo fondamentale in molte altre questioni di Matematica e Fisica

5. Complementi: diagonalizzare senza il teorema spettrale

Questa sezione è del tutto facoltativa. È interessante osservare che il teorema di Sylvester può essere dimostrato senza far uso del teorema spettrale. 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, $\{\underline{f}_1, \dots, \underline{f}_n\}$, *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.* 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

$$V = \mathbb{R}\underline{f} \oplus (\mathbb{R}\underline{f})^{\perp_b} \tag{3}$$

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 (3) utilizza una tecnica già vista per i prodotti scalari definiti positivi (che sono ovviamente particolari forme bilineari simmetriche). [2]

Dimostrazione teorema. 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 $b(\underline{v}+\underline{w}, \underline{v}+\underline{w}) = b(\underline{v}, \underline{v}) + 2b(\underline{v}, \underline{w}) + b(\underline{w}, \underline{w}) = 0 + 2b(\underline{v}, \underline{w}) + 0 \neq 0$ come si voleva. Riassumendo: se $b(\ ,\ )$ non è identicaente 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 $(\mathbb{R}\underline{g}_1)^{\perp_b}$. $b'(\ ,\ )$ è ovviamente una forma bilineare simmetrica su uno spazio vettoriale di dimensione

---

[2] **Dimostrazione di (3)**. 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 $V = \mathbb{R}\underline{f} \oplus (\mathbb{R}\underline{f})^{\perp_b}$.

$(n-1)$. 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,

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

Poniamo
$\underline{v}_i = \frac{1}{\sqrt{\alpha_i}} \underline{f}_i \quad$ se $i \leq \rho_+$
$\underline{v}_i = \frac{1}{\sqrt{-\alpha_i}} \underline{f}_i \quad$ se $(\rho_+ + 1) \leq i \leq (\rho_+ + \rho_-)$
$\underline{v}_i = \underline{f}_i \quad$ 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 sopra dimostriamo che ρ_+ dipende solo da $b(\,,\,)$.