Corso di Laurea in Matematica. Geometria 1. a.a. 2014-15.
Prof. P. Piazza
Diagonalizzazione delle forme bilineari simmetriche (caso reale)

1. Osservazioni preliminari.

Sia V uno spazio vettoriale **reale**. Sia

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

una forma bilineare simmetrica. 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*

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

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

$$b(x_1\underline{v}_1 + \cdots + x_n\underline{v}_n\,,\, y_1\underline{v}_1 + \cdots + y_n\underline{v}_n) = \underline{y}^T A\,\underline{x} \tag{1}$$

Notiamo che

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

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

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

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. Possiamo quindi concludere che l'espressione a destra in (1) è precisamente uguale a quella considerata a lezione. 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 $b^{\mathcal{B}}_A : V \times V \to \mathbb{R}$ ponendo

$$b^{\mathcal{B}}_A(x_1\underline{v}_1 + \cdots + x_n\underline{v}_n\,,\, y_1\underline{v}_1 + \cdots + y_n\underline{v}_n) := \underline{y}^T A\,\underline{x} \equiv \underline{x}^T A \underline{y}\,. \tag{2}$$

Equivalentemente, definiamo la forma bilineare simmetrica $b^{\mathcal{B}}_A$ sulla base $\mathcal{B}$ ponendo $b^{\mathcal{B}}_A(\underline{v}_i, \underline{v}_j) := a_{ij}$ e poi estendiamo per bilinearità a tutti i vettori $\underline{v}$, $\underline{w}$ di V.
È facile verificare che $b^{\mathcal{B}}_A$ è una forma bilineare simmetrica. Denoteremo spesso questa forma bilineare simmerica con b_A, omettendo quindi la base $\mathcal{B}$ dalla notazione.

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

Osservazione. (Esistenza di un prodotto scalare). Sernesi definisce uno spazio vettoriale euclideo come uno spazio vettoriale dotato di un prodotto scalare, e cioè di una forma bilineare simmetrica $<\,,\,>$ definita positiva, tale cioè che

$$< \underline{v}, \underline{v} > \geq 0 \quad \text{e} < \underline{v}, \underline{v} >= 0 \quad \text{se e solo se} \quad \underline{v} = \underline{0}$$

[1]

Domanda: ma i prodotti scalari esistono sempre ?

[1] Attenzione: per Abate-De Fabritiis un prodotto scalare è invece una qualsiasi forma bilineare simmetrica $<\,,\,>$ non-degenere.

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 $b^{\mathcal{B}}_{I_n}$ che è chiaramente definita positiva, perché se $\underline{v}$ è un vettore non-nullo di coordinate $\underline{x}$ nella base $\mathcal{B}$ allora

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

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 $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ sono dati dai vettori della base canonica $\underline{e}_i$, $\underline{e}_j$. Quindi

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

che è uguale a 1 se $i = j$ e 0 se $i \neq j$, come si voleva.

2. Matrici congruenti

Sia $\mathcal{F} = \{\underline{f}_1, \dots, \underline{f}_n\}$ un'altra base di V. Sia $A^{\mathcal{F}}_b$ 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

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

da cui deduciamo

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

da cui, infine,

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

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**.

Vale anche il *viceversa*: siano A e D due matrici *simmetriche* congruenti, $D = C^T AC$, con C invertibile. Sia $b := b^{\mathcal{B}}_A$ la forma bilineare simmetrica definita dalla formula (2) nella base $\mathcal{B}$; per costruzione la matrice associata a b nella base $\mathcal{B}$ è A. 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

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

Allora la matrice associata a b nella base $\mathcal{F}$ è proprio D e quindi:

A e D simmetriche sono congruenti se e solo se rappresentano la stessa forma bilineare simmetrica 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.

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.

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 reale 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 rispetttivamente, indice di positività, indice di negatività ed indice di nullità della forma bilineare simmetrica $b(\, , \,)$
Dimostrazione. Consideriamo una qualsiasi base $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$ 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 euclideo in V (vedere l'Osservazione a pagina 1). Abbiamo quindi un prodotto scalare $< \, , \, >$ 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 $n-r$ 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 $(V, < \, , \, >)$.

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 euclideo; 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 $\mathcal{W}$: $M_{\mathcal{B},\mathcal{W}}(\mathrm{Id}_V)$. 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, $C^{-1}AC = \Lambda$ 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 $U = \mathrm{Span}(\underline{f}_1, \ldots, \underline{f}_\rho)$ 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.

Importante: per capire il segno degli autovalori (senza calcolarli !) potete applicare il magico *criterio di Cartesio*, che trovate nel libro di Abate-De Fabritiis (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).