Corso di Laurea in Matematica. Geometria 1. a.a. 2016-17.
Prof. P. Piazza

Teorema spettrale e diagonalizzazione delle forme bilineari simmetriche reali

Sia V uno spazio vettoriale reale di dimensione finita. Vogliamo vedere come sia possibile diagonalizzare le forme bilineari simmetriche utilizzando il teorema spettrale per gli operatori simmetrici.

Osservazione preliminare. (esistenza di un prodotto scalare in V) Vi ricordo che data una matrice *simmetrica* $A = |a_{i\,j}|$ ed una base $\mathcal{B}$ di V possiamo definire una forma bilineare simmetrica $b^{\mathcal{B}}_A : V \times V \to \mathbb{R}$ ponendo

$$(1) \qquad b^{\mathcal{B}}_A(x_1\underline{v}_1 + \cdots + x_n\underline{v}_n\,,\, y_1\underline{v}_1 + \cdots + y_n\underline{v}_n) := \underline{x}^T A\underline{y}\,.$$

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 estendendo per bilinearità a tutti i vettori* $\underline{v}$, $\underline{w}$ *di* V.
Applichiamo questo procedimento ad una matrice simmetrica molto particolare: la matrice identità I_n. Fissiamo quindi una base $\mathcal{B}$ di V e la matrice identità I_n; consideriamo la forma bilineare (1) 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

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

Per costruzione i vettori della base $\mathcal{B}$ sono ortonormali rispetto a a questo prodotto scalare; 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

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

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

1. Teorema di Sylvester e Teorema spettrale

In questa sezione utilizzeremo il teorema spettrale per dimostrare il

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*

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

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* $b(\,,\,)$. *La coppia* $(\rho, r-\rho)$ *è detta segnatura di* b.
Dimostrazione. Consideriamo una qualsiasi base $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$ e la matrice associata a $b(\,,\,)$ in questa base: $A^{\mathcal{B}}_b$. 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 precedente). 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 $\mathcal{B}$, $M_{\mathcal{B},\mathcal{B}}(T)$, 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 euclideo $(V, <\, , \,>)$.

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 $< T\underline{v}, \underline{w} > = \sum_{j,\ell} x_j y_\ell a_{\ell j}$. Scriviamo l'analoga espressione per $< \underline{v}, T\underline{w} >$:

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

Ma $< \underline{v}_j, T(\underline{v}_\ell) > = \sum_k a_{k\ell} < \underline{v}_j, \underline{v}_k > = a_{j\ell}$ e quindi sfruttando il fatto che $A = A^T$ vediamo che $< T\underline{v}, \underline{w} > = \sum_{j,\ell} x_j y_\ell a_{\ell j} = \sum_{j,\ell} x_j y_\ell a_{j\ell} = < \underline{v}, T\underline{w} >$ che è quello che dovevamo dimostrare.

Abbiamo quindi un operatore simmetrico in uno spazio vettoriale euclideo; 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. A meno di un riordinamento dei vettori possiamo pensare alla base $\mathcal{W}$ ordinata di conseguenza, quindi $\underline{w}_1$ è associato al primo autovalore della nostra lista ordinata di autovalori e cosi' via. Sia C la matrice del cambiamento di base: $M_{\mathcal{B},\mathcal{W}}(\mathrm{Id}_V)$. 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 nelle note sulle forme bilineari (oppure in Sernesi),

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

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

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

Quindi C diagonalizza **simultaneamente** l'operatore T e la forma bilineare 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 $j \in \{\rho+1, \dots, r\}$ 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 $j \in \{\rho+1, \dots, r\}$ 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.

La dimostrazione che in questa matrice r e ρ dipendono solo da $b(\ ,\ )$ e non dalla particolare base scelta procede come nella dimostrazione del Teorema di Sylvester che non fa uso del teorema spettrale. Il teorema di Sylvester è dimostrato.

Importante: per capire il segno degli autovalori (senza calcolarli !) potete applicare il magico *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 [1]*. 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.

Esempio. Il polinomio $p(t) = -t^3 + 38t - 31$ è il polinomio caratteristico di una matrice reale; per il criterio di Cartesio 0 non è radice, p ha 2 radici positive (perché ci sono due variazioni di segno) e quindi 1 radice negativa (perché 1=3-2).

Osservazione.
Sia $(V, <\ ,\ >)$ uno spazio vettoriale euclideo 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,

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

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

Equivalentemente, la forma quadratica q associata a b si scrive nelle coordinate $\underline{y}$ associate a $\mathcal{W}$ come

$$\lambda_1 y_1^2 + \cdots \lambda_r y_r^2 \tag{2}$$

dove λ_j sono gli autovalori non nulli di T.
Osserviamo che se $\mathcal{U}$ è un'altra base **ortonormale** con la proprietà che

$$A_b^{\mathcal{U}} = \Delta$$

con Δ diagonale, allora Λ e Δ possono differire solo nell'ordine degli elementi diagonali.
Dimostrazione. Sappiamo che $A_b^{\mathcal{U}} = D^T AD$ con $A = A_b^{\mathcal{B}}$ e $D = M_{\mathcal{B},\mathcal{U}}(\mathrm{Id}_V)$. Inoltre A è anche uguale, per definizione, a $M_{\mathcal{B},\mathcal{B}}(T)$. Per ipotesi $D \in O(n)$ perché $\mathcal{U}$ e $\mathcal{B}$ sono ortonormali. In particolare $D^T = D^{-1}$. Ne segue che la matrice Δ è anche uguale alla matrice associata a T nella base $\mathcal{U}$ perché

$$\begin{aligned} \Delta = D^T AD = D^{-1}AD = (M_{\mathcal{B},\mathcal{U}}(\mathrm{Id}_V))^{-1}\, A\, M_{\mathcal{B},\mathcal{U}}(\mathrm{Id}_V) = \\ (M_{\mathcal{B},\mathcal{U}}(\mathrm{Id}_V))^{-1}\, M_{\mathcal{B},\mathcal{B}}(T)\, M_{\mathcal{B},\mathcal{U}}(\mathrm{Id}_V) = M_{\mathcal{U},\mathcal{B}}(\mathrm{Id}_V)\, M_{\mathcal{B},\mathcal{B}}(T)\, M_{\mathcal{U},\mathcal{B}}(\mathrm{Id}_V) \\ = M_{\mathcal{U},\mathcal{U}}(\mathrm{Id}_V \circ T \circ \mathrm{Id}_V) = M_{\mathcal{U},\mathcal{U}}(T) \end{aligned}$$

Quindi $\Delta = M_{\mathcal{U},\mathcal{U}}(T)$. Ma Δ ̀diagonale; ne segue che $\mathcal{U}$ è una base di autovettori e Δ ha sulla diagonale principale gli autovalori di T. Dato che gli autovalori non dipendono dalla base scelta per rappresentare T, ne deduciamo che Λ e Δ differiscono solo nell'ordine degli elementi sulla diagonale principale, che era quello che dovevamo giustificare.
La (2) è per definizione *la forma canonica euclidea* della forma quadratica q. È ben definita a meno dell'ordine. (Se ordiniamo gli autovalori non nulli in ordine decrescente e, a meno di riordinamento, assumiamo la base ortonormale diagonalizzante ordinata di conseguenza (con gli ultimi $n-r$ vettori associati all'autovalore nullo) allora otteniamo un'unica espressione.)