Geometria 1. II^0 Modulo. a.a. 01/02. Gruppo A-L (Prof. P. Piazza)
Complementi sulla diagonalizzazione delle forme quadratiche

Sia ϕ una forma quadratica: $\phi : \mathbb{R}^n \to \mathbb{R}$. Esiste quindi una matrice simmetrica A tale che $\phi(\underline{x}) = \underline{x}^T A \underline{x}$. Scriveremo anche $z = \phi(\underline{x})$ oppure $z = \underline{x}^T A \underline{x}$.
Siano $\lambda_1, \dots, \lambda_\rho \in \mathbb{R}$ gli autovalori positivi di A e sia r il rango di A. Ci sono allora $r - \rho$ autovalori negativi $\lambda_{\rho+1}, \dots, \lambda_r$; inoltre l'autovalore 0 ha molteplicità $n - r$. Poniamo $\nu = n - r$. Ordiniamo gli autovalori come sopra, in modo tale cioè che i primi ρ siano positivi, i seguenti $r - \rho$ siano negativi ed i rimanenti $\nu = n - r$ siano uguali a zero.
Per il teorema spettrale sappiamo che esiste una base ortonormale $\underline{w}_1, \dots, \underline{w}_n$ di $\mathbb{R}^n$ costituita da autotovettori di A. Equivalentemente, detta C la matrice che ha come colonne le coordinate di questi autovettori si ha $C \in O(n)$ e $C^T A C = \Lambda$ con Λ la matrice diagonale che ha gli autovalori sulla diagonale principale. Siano $\underline{y}$ le coordinate associate alla base $\underline{w}_1, \dots, \underline{w}_n$; sappiamo che $\underline{x} = C\underline{y}$. Possiamo scrivere

$$z = \phi(\underline{x}) = \underline{x}^T A \underline{x} = (C\underline{y})^T A C \underline{y} = \underline{y}^T \Lambda \underline{y}.$$

Ma

$$\underline{y}^T \Lambda \underline{y} = \lambda_1 y_1^2 + \dots + \lambda_\rho y_\rho^2 - (-\lambda_{\rho+1}) y_{\rho+1}^2 - \dots - (-\lambda_r) y_r^2 + 0 y_{r+1}^2 + \dots + 0 y_n^2$$

e quindi, in definitiva,

$$z = \lambda_1 y_1^2 + \dots + \lambda_\rho y_\rho^2 - (-\lambda_{\rho+1}) y_{\rho+1}^2 - \dots - (-\lambda_r) y_r^2.$$

Questa è la forma canonica *metrica* della forma quadratica.
Poniamo ora
$z_j = \sqrt{\lambda_j} y_j$ se $1 \le j \le \rho$,
$z_j = \sqrt{-\lambda_j} y_j$ se $j \in \{\rho+1, \dots, r\}$ e
$z_j = y_j$ se $r + 1 \le j \le n$.
Queste coordinate $\underline{z}$ sono indotte dalla nuova base $\{\underline{v}_1, \dots, \underline{v}_n\}$ definita come segue:
$\underline{v}_j = \frac{1}{\sqrt{\lambda_j}} \underline{w}_j$ se $1 \le j \le \rho$,

$\underline{v}_j = \frac{1}{\sqrt{-\lambda_j}} \underline{w}_j$ se $j \in \{\rho+1, \dots, r\}$
$\underline{v}_j = \underline{w}_j$ se $r + 1 \le j \le n$.
In queste coordinate la forma quadratica si scrive

$$z_1^2 + \dots + z_\rho^2 - z_{\rho+1}^2 - \dots - z_r^2.$$

Concludendo: *esiste un base di $\mathbb{R}^n$ con coordinate associate $\underline{z}$ tale che $\phi(\,)$ si scriva in queste coordinate nella forma*

$$z_1^2 + \dots + z_\rho^2 - z_{\rho+1}^2 - \dots - z_r^2.$$

La matrice associata alla forma quadratica in questa base è quindi

$$\begin{vmatrix} I_\rho & 0 & 0 \\ 0 & -I_{r-\rho} & 0 \\ 0 & 0 & 0_{n-r} \end{vmatrix}$$

In parole diverse: ogni matrice simmetrica è *congruente* ad una tale matrice.
Proposizione. Gli interi r e ρ dipendono solo da $\phi(\,)$ e non dalla particolare base scelta.

Dimostrazione. Sia $\underline{a}_1, \dots, \underline{a}_n$ un'altra base di $\mathbb{R}^n$ rispetto alla quale ϕ si scriva in forma diagonale. Siano $\underline{t}$ le coordinate associate a questa base e sia quindi

$$t_1^2 + \cdots + t_\sigma^2 - t_{\sigma+1}^2 - \cdots - t_s^2$$

tale forma diagonale. Nel libro di testo è dimostrato che ν, o equivalentemente r, è un invariante per congruenza della matrice A. Quindi dovrà essere necessariamente

$$z = t_1^2 + \cdots + t_\sigma^2 - t_{\sigma+1}^2 - \cdots - t_r^2$$

perché altrimenti le matrici A e

$$\begin{vmatrix} I_\sigma & O & O \\ 0 & -I_{s-\sigma} & 0 \\ 0 & 0 & 0_{n-s} \end{vmatrix}$$

sarebbero congruenti ma con ranghi diversi. Rimane da dimostrare che $\rho = \sigma$. Per assurdo $\rho \neq \sigma$ e supponiamo che sia $\sigma < \rho$. Sia $U = \mathrm{Span}(\underline{v}_1, \dots, \underline{v}_\rho)$ e sia $W = \mathrm{Span}(\underline{a}_{\sigma+1}, \dots, \underline{a}_n)$. Per ipotesi $\dim U + \dim W = \rho + n - \sigma > n$ e quindi, per Grassmann, l'intersezione di questi due sottospazi ha dimensione > 0. Sia $\underline{b}$ un vettore in non nullo in $U \cap W$; quindi

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

e

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

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

La forma canonica ottenuta tramite il procedimento appena illustrato, e cioè

$$z_1^2 + \cdots + z_\rho^2 - z_{\rho+1}^2 - \cdots - z_r^2$$

è quindi *unica*; essa è detta *forma canonica affine.* Mettendo insieme quanto sopra potete ridimostare il teorema 4.4 pag 28.

Esercizio. Sia $z = \phi(\underline{x})$ la forma quadratica $\phi(\underline{x}) = 2x_1^2 + 2x_2^2 + 2x_3^2 - 4x_1x_2 + 4x_1x_3 - 4x_2x_3$. Determinare una base di $\mathbb{R}^3$, con coordinate associate $\underline{z}$, che porti la forma quadratica nella sua forma canonica affine.

Soluzione. La matrice che definisce ϕ è

$$\begin{vmatrix} 2 & -2 & 2 \\ -2 & 2 & -2 \\ 2 & -2 & 2 \end{vmatrix}.$$

Determiniamo il polinomio caratteristico di A; è subito visto che $P_A(\lambda) = (\lambda - 6)\lambda^2$. C'è quindi un solo autovalore positivo, 6, e l'autovalore 0 con molteplicità $\nu = 2$. Possiamo già concludere, per il teorema pag. 28, che la forma canonica affine di ϕ è

$$z = z_1^2$$

per qualche sistema di coordinate $\underline{z}$. Per determinare tale sistema di coordinate troviamo dapprima una base **ortonormale** di autovettori per A [1]. È chiaro che

[1] **Attenzione:** se prendete gli autovettori non-ortonormali allora diagonalizzate l'operatore $A : \mathbb{R}^n \to \mathbb{R}^n$ ma **non** diagonalizzate la forma quadratica $\underline{x}^T A\underline{x}$, essendo $C^T \neq C^{-1}$

$V_A(0) = \mathrm{Ker}A = \{\underline{x} \in \mathbb{R}^3 \mid x_1 - x_2 + x_3 = 0\}$ e quindi che $V_A(6) = \mathbb{R}(1,-1,1)$ (perché sappiamo già che $V_A(6) \perp V_A(0)$). Una base **ortogonale** di $V_A(0)$ è

$$(1,1,0)\,, \quad (1,-1,-2)\,.$$

Quindi una base ortonormale (ordinata) di autovettori per A è data da

$$\underline{w}_1 = (\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}})\,, \quad \underline{w}_2 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)\,, \quad \underline{w}_3(\frac{1}{\sqrt{6}}, \frac{-1}{\sqrt{3}}, \frac{-2}{\sqrt{3}})\ ;$$

se $\underline{y}$ sono le coordinate associate a questa base ortonormale allora

$$z = 6y_1^2 \quad \text{(forma canonica metrica)}$$

La base che porta la forma quadratica nella sua forma canonica affine è allora $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ con

$$\underline{v}_1 = \frac{1}{\sqrt{6}}\underline{w}_1\,, \quad \underline{v}_2 = \underline{w}_2\,, \quad \underline{v}_3 = \underline{w}_3\ ;$$

se $\underline{z}$ sono le coordinate associate a questa base allora la forma quadratica si esprime in queste coordiante come:

$$z = z_1^2 \quad \text{(forma canonica affine)}$$

La forma quadratica è quindi semidefinita positiva. La soluzione è completa.

Per completezza diamo una dimostrazione della seguente

Proposizione. Sia $\phi(\underline{x}) = \underline{x}^T A \underline{x}$, $A = A^T$, una forma quadratica. Supponiamo che i determinanti delle sottomatrici principali $\mathcal{A}_k$, $k = 1, \ldots, n$, siano tutti positivi. Allora $\phi(\,)$ è definita positiva.

Dimostrazione. Prima di iniziare facciamo una
Osservazione preliminare. Se $B = C^T AC$ allora

$$b_{ij} = b_{ji} = (\underline{c}^j)^T A \underline{c}^j\,.$$

La dimostrazione è del tutto elementare.
Passiamo alla dimostrazione. Procediamo per induzione su n. Per $n = 1$ non c'è nulla da dimostrare. Supponiamo la proposizione vera per spazi di dimensione $n-1$ e dimostriamola per quelli di dimensione n. Sia $W = \mathrm{Span}(\underline{e}_1, \ldots \underline{e}_{n-1})$ con $\{\underline{e}_j\}$ i vettori della base canonica. È ovvio che $\underline{x} \in W$ sse $\underline{x} = (\underline{x}', 0)$ con $\underline{x}' \in \mathbb{R}^{n-1}$. Consideriamo $\mathcal{A}_{n-1}$. Questa è una matrice simmetrica ed è tale che tutte le sue sottomatrici principali hanno determinante positivo. Quindi, per ipotesi induttiva,

$$(\underline{x}')^T \mathcal{A}_{n-1} \underline{x}' > 0 \ \text{ se } \ \underline{x}' \neq \underline{0}\,.$$

La forma bilineare simmetrica $W \times W \to \mathbb{R}$ definita da

$$\underline{x}'\,,\, \underline{y}' \ \to < \underline{x}', \underline{y}' >_W \ := \ (\underline{y}')^T \mathcal{A}_{n-1} \underline{x}'$$

è quindi definita positiva e definisce un prodotto scalare in W. Sia $\{\underline{v}_1, \ldots, \underline{v}_{n-1}\}$ una base ortonormale di W rispetto a questo prodotto scalare. Sia

$$\underline{v}_n = \underline{e}_n - \sum_1^{n-1} (\underline{e}_n^T A \underline{v}_j) \underline{v}_j\,.$$

È chiaro che $\{\underline{v}_1, \ldots, \underline{v}_n\}$ è una base di $\mathbb{R}^n$. Siano $\underline{y}$ le coordinate associate a questa base; quindi $\underline{x} = C\underline{y}$ con C la matrice che ha come colonne le coordinate dei vettori

$\underline{v}_j$ nella base canonica. Consideriamo $B = C^T AC$; per l'*osservazione preliminare* $b_{ij} = \underline{v}_j^T A\underline{v}_i$. Se $i, j \in \{1, \dots, n-1\}$ allora è chiaro per costruzione che

$$b_{ij} = \underline{v}_j^T A\underline{v}_i = \underline{v}_j^T \mathcal{A}_{n-1}\underline{v}_i \ \ ;$$

quindi, utilizzando il fatto che $\{\underline{v}_1, \cdots, \underline{v}_{n-1}\}$ è una base ortonormale rispetto al prodotto scalare $<, >_W$, otteniamo immediatamente che

$$b_{ii} = 1 \quad \text{per } i \in \{1, \dots, n-1\}\ ; \quad b_{ij} = 0 \quad \text{se } i \neq j \quad \text{e} \quad i, j \in \{1, \dots, n-1\}\,.$$

Un semplice calcolo mostra anche che con la nostra definizione $b_{jn} = b_{nj} = \underline{v}_n^T A\underline{v}_i = 0$. Lascio a voi questa semplice verifica. Ma allora

$$B = \begin{vmatrix} I & \underline{0} \\ \underline{0}^T & b_{nn} \end{vmatrix} \quad \text{con} \quad \underline{0}^T = (0, \dots, 0) \in \mathbb{R}^{n-1}\,.$$

Quindi $\det B = b_{nn}$. D'altra parte $\det B = \det(C^T AC) = (\det C)^2 \det A > 0$ per ipotesi e quindi $b_{nn} > 0$. Ma allora

$$\underline{x}^T A\underline{x} = \underline{y}^T C^T AC\underline{y} = \underline{y}^T B\underline{y} = y_1^2 + \cdots + y_{n-1}^2 + b_{nn}y_n^2$$

che è > 0 se $\underline{y} \neq \underline{0}$. Ma $\underline{x} \neq \underline{0}$ sse $\underline{y} \neq \underline{0}$ e quindi $\underline{x}^T A\underline{x} > 0$ se $\underline{x} \neq \underline{0}$ che è quello che dovevamo dimostrare.