Autovalori e autovettori, matrici simmetriche e forme quadratiche (cenni)
(prof. M. Salvetti)

April 14, 2011

(alcune note non complete sugli argomenti trattati: eventuali completamenti saranno aggiunti)

1 Autovalori e autovettori

(sarebbe meglio discutere l'argomento partendo da un punto di vita piu' generale, ma per brevita' parleremo subito di matrici).

Sia A una matrice quadrata di ordine n. Si consideri l'applicazione lineare di $\mathbb{R}^n$ in sé definita da

$$\mathbf{x} \longrightarrow A\mathbf{x}, \ \mathbf{x} \in \mathbb{R}^n.$$

Definizione 1.1 *Un vettore $\mathbf{x} \in \mathbb{R}^n$ per il quale esiste un numero λ tale che*

$$A\mathbf{x} \ = \ \lambda\mathbf{x}$$

*si dice **autovettore** della matrice A. Il numero λ si chiama **autovalore** di A.*

Un caso particolare é quello in cui l'autovalore é nullo, cio**e**' $A\mathbf{x} = \mathbf{0}$. Dire che $\mathbf{x}$ é autovettore con autovalore 0 equivale completamente a dire che $\mathbf{x}$ sta nel nucleo di A.

E' chiaro che se $\mathbf{x}$ e' autovettore con autovalore λ, anche tutti i vettori $\mathbf{y}$ multipli di $\mathbf{x}$ sono autovettori con lo stesso autovalore: in altri termini tutta la retta direzione di $\mathbf{x}$ e' una retta di autovettori con autovalore λ. Quindi gli autovettori sono quei vettori che, o stanno nel nucleo, oppure "conservano la direzione" quando trasformati tramite A.

Vale qualcosa in piú:

Lemma 1.2 *Tutti gli autovettori aventi uno stesso autovalore λ costituiscono un sottospazio vettoriale di $\mathbb{R}^n$ (cioé sono un insieme chiuso per somma e moltiplicazione per scalari).*

(Tale sottospazio **e** detto l'*autospazio* relativo all'autovalore λ).

Dimostrazione. Se $A\mathbf{x} = \lambda\mathbf{x},\ A\mathbf{y} = \lambda\mathbf{y}$, allora

$$A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha A(\mathbf{x}) + \beta A(\mathbf{y}) = \alpha\lambda\mathbf{x} + \beta\lambda\mathbf{y} = \lambda(\alpha\mathbf{x} + \beta\mathbf{y})$$

per linearitá, il che dimostra che qualunque $\alpha\mathbf{x} + \beta\mathbf{y}$ é ancora autovettore con autovalore λ. $\square$

Definizione 1.3 *La* ***molteplicitá geometrica*** *dell'autovalore λ é la dimensione dell'autospazio relativo a λ.*

Il seguente lemma ci dá un modo per calcolare gli autovalori di una matrice.

Lemma 1.4 *Gli autovalori di A sono le radici dell'equazione in λ :*

$$det(A - \lambda I) = 0$$

dove I é la matrice identitá di ordine n.

L'espressione in λ data da $P(A, \lambda) := det(A - \lambda I)$ é un polinomio di grado n in λ, detto il *polinomio caratteristico* di A. Per un teorema generale sulle equazioni polinomiali, $P(A, \lambda)$ ha sempre esattamente n radici (contate con molteplicitá), e quindi ogni matrice ha sempre n autovalori (con molteplicitá). Si noti tuttavia che in generale gli autovalori saranno complessi (non reali), anche se si parte da una matrice reale.

Dimostrazione del lemma 1.4. Riscriviamo $A\mathbf{x} = \lambda\mathbf{x}$ come $(A - \lambda I)\mathbf{x} = 0$. Questo é un sistema quadrato, per cui i numeri λ per cui esiste $x \neq 0$ che lo verifica sono quelli che annullano il determinante associato. $\square$

Esempio. Gli autovalori di una matrice diagonale D sono gli elementi della diagonale principale, come é ovvio dal modo in cui si calcola il determinante. La stessa proprietá si ha per le matrici triangolari.

Definizione 1.5 *Si dice* ***molteplicitá algebrica*** *dell'autovalore λ_0 la massima potenza per cui il polinomio caratteristico $P(A, \lambda)$ é divisibile per il fattore lineare $\lambda - \lambda_0$ (si ricorda che per il teorema di Ruffini x_0 é radice di un polinomio $p(x)$ se e solo se il polinomio é divisibile per $x - x_0$).*

Citiamo senza dimostrazione la seguente proprietá.

Lemma 1.6 *La molteplicitá geometrica di un autovalore λ é un numero (≥ 1) che vale al massimo quanto la sua molteplicitá algebrica, ma puó essere inferiore.*

Esempio. La matrice $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ ha polinomio caratteristico λ^2 e quindi l'autovalore 0 di molteplicitá algebrica 2. La molteplicitá geometrica di 0 é la dimensione del nucleo di A, che é 1.

Definizione 1.7 *La matrice A é diagonalizzabile se e solo se é simile a una matrice diagonale. In altri termini, esiste un cambiamento di base in $\mathbb{C}^n$, dato da una matrice invertibile P, per cui $P^{-1}AP$ é diagonale.*

Osservazione. Matrici simili hanno gli stessi autovalori. Infatti:

$$det(P^{-1}AP - \lambda I) = det(P^{-1}(A - \lambda I)P) = det(P^{-1})det(A - \lambda I)det(P) =$$

$$= det(P)^{-1}det(A - \lambda I)det(P) = det(A - \lambda I).$$

Quindi se $P^{-1}AP = D$ **e**' diagonale, sulla diagonale di D compaiono gli autovalori di A. Ne segue che se A é diagonalizzabile tramite una matrice *reale* P, allora tutti gli autovalori devono essere reali.

Esempio. La matrice $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ ha autovalori i e $-i$ quindi non puó essere diagonalizzata con una matrice reale P. Lo é pero' con una matrice complessa.

Osservazione. Riscrivendo $P^{-1}AP = D$ come $AP = PD$, si deduce la seguente importante caratterizzazione delle matrici diagonalizzabili: *A é diagonalizzabile se e solo se esiste una base di $\mathbb{C}^n$ costituita da autovettori di A.*

Inoltre: *A é diagonalizzabile su $\mathbb{R}$ (cioé utilzzando una P reale) se e solo se esiste una base di $\mathbb{R}^n$ costituita da autovettori di A.*

Citiamo anche senza dimostrazione:

Teorema 1 *La matrice A é diagonalizzabile su $\mathbb{C}$ se per ogni suo autovalore le due molteplicitá (geometrica e algebrica) coincidono. La matrice A é diagonalizzabile su $\mathbb{R}$ se tutti i suoi autovalori sono reali e per ognuno di essi le due molteplicitá coincidono.*

É chiaro che se gli autovalori di A sono tutti semplici (cioé hanno molteplicitá algebrica 1) allora le due molteplicitá coincidono, per cui si ha dal teorema il seguente corollario

Corollario 1.8 *Una matrice A di ordine n che abbia tutti autovalori semplici (equivalentemente: A ha n autovalori tutti distinti fra loro) é diagonalizzabile.*

2 Matrici simmetriche

Useremo il prodotto hermitiano canonico in $\mathbb{C}^n$ dato da

$$< \mathbf{x}, \mathbf{y} >= \overline{\mathbf{x}}^T \mathbf{y} = \sum_{i=1}^{n} \overline{x}_i y_i$$

dove la sopralineatura é il coniugato complesso e i vettori si pensano come colonne con n componenti. Se i vettori sono reali il prodotto si riduce al prodotto scalare canonico in $\mathbb{R}^n$

$$< \mathbf{x}, \mathbf{y} >= \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

Definizione 2.1 *Una marice quadrata $A = (a_{ij})_{i=1,\ldots,n; j=1,\ldots,n}$ a coefficienti reali é* simmetrica *se $A^T = A$, cioé se $a_{ij} = a_{ji}$ per ogni elemento di matrice. Qui la matrice A^T é la matrice trasposta di A.*

Per ogni matrice reale A e per ogni coppia di vettori reali $\mathbf{x}, \mathbf{y}$ vale

$$< A\mathbf{x}, \mathbf{y} >= (A\mathbf{x})^T \mathbf{y} = \mathbf{x}^T A^T \mathbf{y} =< \mathbf{x}, A^T \mathbf{y} > .$$

Ne segue che se A é simmetrica allora

$$< A\mathbf{x}, \mathbf{y} >=< \mathbf{x}, A\mathbf{y} >, \text{ per ogni } \mathbf{x}, \mathbf{y} \tag{1}$$

La condizione (1) in realtá caratterizza le matrici simmetriche perché se $\mathbf{x} = \mathbf{e}_i$, $\mathbf{y} = \mathbf{e}_j$ sono l'$i-$simo e il $j-$simo vettore della base canonica, si verifica subito che

$$< A\mathbf{e}_i, \mathbf{e}_j >= (A\mathbf{e}_i)^T \mathbf{e}_j = (\mathbf{e}_i)^T A^T \mathbf{e}_j = a_{ji}, \ < \mathbf{e}_i, A\mathbf{e}_j >= (\mathbf{e}_i)^T A\mathbf{e}_j = a_{ij}.$$

Teorema 2 *Una matrice simmetrica ha tutti gli autovalori reali.*

Dimostrazione. Se $A\mathbf{x} = \lambda\mathbf{x}$, allora, usando il prodotto hermitiano standard:

$$\overline{\lambda} < \mathbf{x}, \mathbf{x} >=< \lambda\mathbf{x}, \mathbf{x} >=< A\mathbf{x}, \mathbf{x} >=< \mathbf{x}, A\mathbf{x} >=< \mathbf{x}, \lambda\mathbf{x} >= \lambda < \mathbf{x}, \mathbf{x} >$$

da cui $\lambda = \overline{\lambda}$ perché $< \mathbf{x}, \mathbf{x} >$ é reale positivo per ogni vettore $\mathbf{x} \neq \mathbf{0}$. □

Teorema 3 *Se A é simmetrica, autovettori (reali) $\mathbf{x}_1, \ldots, \mathbf{x}_k$ relativi ad autovalori distinti rispettivamente $\lambda_1, \ldots, \lambda_k$ sono a due a due ortogonali.*

Dimostrazione. Sia $A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \ i = 1, \ldots, k$. Si ha

$$\lambda_i < \mathbf{x}_i, \mathbf{x}_j >=< A\mathbf{x}_i, \mathbf{x}_j >=< \mathbf{x}_i, A\mathbf{x}_j >= \lambda_j < \mathbf{x}_i, \mathbf{x}_j >$$

per cui se $i \neq j$ allora $< \mathbf{x}_i, \mathbf{x}_j >= 0$. □

Siano $\lambda_1, \ldots, \lambda_k$ tutti i diversi autovalori di A simmetrica (alcuni possono essere multipli, e quindi puo' essere $k < n$). Siano $V_1, \ldots, V_k$ i relativi autospazi.

Teorema 4 *$V_1 + \cdots + V_k = \mathbb{R}^n$, cioé gli autovettori generano tutto $\mathbb{R}^n$. Ne segue che A é diagonalizzabile.*

Dimostrazione. Sia V il sottospazio generato dagli autovettori. Se $V \neq \mathbb{R}^n$, allora V ha un ortogonale W in $\mathbb{R}^n$. Se $\mathbf{x}$ é un autovettore di autovalore λ e $\mathbf{w} \in W$,

$$0 =< \mathbf{x}, \mathbf{w} >= \lambda < \mathbf{x}, \mathbf{w} >=< A\mathbf{x}, \mathbf{w} >=< \mathbf{x}, A\mathbf{w} >$$

per cui anche $A\mathbf{w} \in W$. Ma allora ci sarebbe un ulteriore autovettore in W, contro l'ipotesi. □

Dai teoremi 3.3 e 4 segue subito

Corollario 2.2 *Se A é simmetrica, esiste una base ortonormale di autovettori per A.*

Si ricirda che una matrice O tale che $O^T O = OO^T = I$ si dice *ortogonale.* Equivalentemente, O é ortogonale se $O^T = O^{-1}$.

La proprietá 2.2 si puó dire in modo equivalente come

Corollario 2.3 *Esiste una matrice ortogonale O tale che*

$$O^{-1} AO = O^T AO = D \tag{2}$$

é diagonale.

3 Forme quadratiche

Una *forma quadratica* nelle variabili $x_1, \ldots, x_n$ é un polinomio omogeneo di grado 2:

$$Q(x_1, \ldots, x_n) \;=\; \sum_{1 \le i,j \le n} \alpha_{ij} x_i x_j.$$

Ovviamente, se $i < j$

$$\alpha_{ij} x_i x_j + \alpha_{ji} x_j x_i \;=\; (\alpha_{ij} + \alpha_{ji}) x_i x_j$$

(cioé conta solo la somma dei due coefficienti) per cui il polinomio si puo' riscrivere in maniera simmetrica:

$$Q(x_1, \ldots, x_n) \;=\; \sum_{1 \le i,j \le n} a_{ij} x_i x_j \tag{3}$$

dove $a_{ij} = a_{ji} = (\alpha_{ij} + \alpha_{ji})/2$. Introducendo il vettore $\mathbf{x}$ delle coordinate e la matrice simmetrica $A = (a_{ij})$ la (3) si riscrive come prodotto

$$Q(\mathbf{x}) \;=\; \mathbf{x}^T A \mathbf{x} \tag{4}$$

Proposizione 3.1 (Cambiamento di base) *Se si esprimono le* $\mathbf{x}$ *tramite altre variabili* $\mathbf{y}$ *con un cambiamento espresso da una matrice invertibile* P,

$$\mathbf{x} = P\mathbf{y}$$

la forma quadratica si trasforma nella

$$Q'(\mathbf{y}) = \mathbf{y}^T A' \mathbf{y}$$

dove

$$A' = P^T A P$$

Dimostrazione. Infatti

$$\mathbf{x}^T A \mathbf{x} = (P\mathbf{y})^T A (P\mathbf{y}) = \mathbf{y}^T P^T A P \mathbf{y}$$

□

Definizione 3.2 *Una forma quadratica diagonale é una forma dove la matrice* A *é diagonale, quindi del tipo*

$$Q(\mathbf{x}) = \sum_{i=1}^{n} a_{ii} x_i^2$$

Corollario 3.3 *Ogni forma quadratica é diagonalizzabile tramite un cambiamento di variabili dato da una matrce ortogonale.*

Dimostrazione. Per il corollario 2.3 esiste O ortogonale tale che $O^T A O$ sia diagonale. Allora cambiando coordinate tramite $\mathbf{x} = O\mathbf{y}$ la nuova forma é diagonale. □

Osservazione. Poiché $O^T = O^{-1}$ per matrici ortogonali, la trasformazione di matrici data dal cambiamento di base fatto con una matrice ortogonale é una similitudine, quindi la matrice diagonale che si ottiene ha gli stessi autovalori di A. Ma per una matrice diagonale gli autovalori coincidono con gli elementi della diagonale principale. Si ottiene quindi: cambiando base, la forma diventa:

$$\lambda_1 y_1^2 + \cdots + \lambda_n \mathbf{y}_n^2,$$

dove i λ_j sono gli autovalori di A. $\square$

Ovviamente vale sempre $Q(\mathbf{0}) = 0$.

Definizione 3.4

1. *La forma Q si dice* **definita positiva** *se assume sempre valori positivi (se $\mathbf{x} \neq \mathbf{0}$).*
2. *La forma Q si dice* **definita negativa** *se assume sempre valori negativi (se $\mathbf{x} \neq \mathbf{0}$).*
3. *La forma Q si dice* **semidefinita positiva** *se assume sempre valori* ≥ 0.
4. *La forma Q si dice* **semidefinita negativa** *se assume sempre valori* ≤ 0.
5. *La forma Q si dice* **non-degenere** *se la matrice A é non singolare.*
6. *La forma Q si dice* **indefinita** *se é non-degenere e se assume valori sia positivi che negativi.*

Osservazione. Le proprietá della definizione 3.4 si conservano per cambiamento di variabili del tipo della prop. 3.1

Proposizione 3.5 *Per la forma Q vale:*
Q é definita positiva se e solo se tutti gli autovalori di A sono positivi;
Q é definita negativa se e solo se tutti gli autovalori di A sono negativi;
Q é semi-definita positiva se e solo se tutti gli autovalori di A sono ≥ 0;
Q é semi-definita negativa se e solo se tutti gli autovalori di A sono ≤ 0;
Q é indefinita se e solo se A ha autovalori sia positivi che negativi, ma non nulli.

Dimostrazione. Si puó guardare alla forma diagonalizzata, come risulta dall' osservazione 3, da cui tutte le condizioni scritte sono evidenti. $\square$

Osservazione. Una forma non degenere ha matrice associata data con tutti autovalori non nulli (perché il determinante di una matrice é il prodotto degli autovalori). Ne segue che una forma non degenere o é definita (positiva o negativa) o é indefinita.

Definizione 3.6 *Data la matrice quadrata A di ordine n, il* minore principale *a_k di A é il minore formato con le prime k righe e colonne di A ($k = 1, \ldots, n$).*

Definizione 3.7 *Una matrice A di ordine n si dice* regolare *se tutti i minori principali a_k sono non nulli (in particolare, $a_n = det(A) \neq 0$ e quindi A sará invertibile).*

Teorema 5 (forma normale di Gauss) . *Data una matrice A regolare, esistono uniche matrici L, D, R, di ordine n, con:*
L triangolare inferiore, con tutti 1 sulla diagonale principale;
D diagonale, invertibile;
R triangolare superiore, con tutti 1 sulla diagonale;
tali che

$$A = LDR$$

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ * & 1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ * & * & \cdots & 1 \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 & \cdots \\ 0 & d_2 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} 1 & * & * & \cdots \\ 0 & 1 & * & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Inoltre si ha

$$d_1 = a_1, \; d_i = a_i/a_{i-1}, \; i = 2, \ldots, n \tag{$*$}$$

Dimostrazione (basata sulla riduzione a scala col metodo di Gauss). Prima si determina una matrice L_1, triangolare inferiore con tutti 1 sulla diagonale, tale che la matrice $L_1 A$ sia triangolare superiore. I coefficienti della $i-$sima riga di L_1 si scelgono (al crescere di i da 1 a n) in modo da azzerare i termini a_{ij} di A, con $1 \leq j < i$ (riduzione di Gauss). Denotiamo i coefficienti di L_1 con λ_{ij}, $i, j = 1, \ldots, n$. Poniamo $\lambda_{ii} = 1$, $i = 1, \ldots, n$, e $\lambda_{ij} = 0$ se $i < j$. La condizione detta si traduce per i rimanenti elementi della riga $i - sima$ di L_1 $(i > 1)$ come

$$\sum_{r=1}^{i-1} a_{rs}\lambda_{ir} + a_{is} = 0, \quad per \; s < i$$

$$\sum_{r=1}^{i-1} a_{ri}\lambda_{ir} + a_{ii} = 1.$$

Questo sistema lineare quadrato con incognite λ_{ir}, $r = 1, \ldots, i-1$ ha determinante esattamente a_{i-1}, e quindi ha un unica soluzione. Si trova quindi

$$L_1 A = R_1$$

con R_1 triangolare superiore. Siano $d_1, \ldots, d_n$ sono gli elementi diagonali di R_1 (i pivot). Il minore principale di R_1 formato con le prime i righe e colonne, che ha determinante $d_1 \ldots d_i$, **e**' il prodotto (per il teorema di Binet) dell'analogo minore di L_1, che ha determinante 1, e di a_i. Quindi

$$a_i = d_1 \ldots d_i \tag{$**$}.$$

Sia D la matrice diagonale tale che che $D_{ii} = d_i$, si puó scrivere

$$R_1 = DR$$

dove R é triangolare superiore e con tutti gli elementi diagonali uguali a 1. Ponendo $L = L_1^{-1}$ si trova la decomposizione sopra scritta.

L'unicitá deriva dal fatto che se

$$LDR \;=\; L'D'R'$$

allora

$$DR(R')^{-1} \;=\; L^{-1}L'D'$$

ma a sinistra abbiamo una matrice triangolare superiore i cui elementi diagonali sono dati da D, a destra una matrice triangolare inferiore con elementi diagonali dati da D', quindi si deduce subito $D = D'$, e $R(R')^{-1} = Id$, $L^{-1}L' = Id$.

Inoltre (**) é equivalente ovviamente alla (*). □

Corollario 3.8 *Se A é simmetrica, le matrici L e R del teorema precedente sono una la trasposta dell'altra.*

Dimostrazione. Deriva dalla

$$A = A^T = (LDR)^T = R^T D^T L^T$$

e dall'unicitá della decomposizione di Gauss. □

Teorema 6 *La forma Q, con matrice associata A,*

(i) é definita positiva se e solo se $a_k > 0,\ k = 1, \dots, n$;

(ii) é definita negativa se e solo se $a_k > 0$ per k pari e $a_k < 0$ per k dispari.

Dimostrazione. (i) Q é definita positiva se e solo se la sua restrizione allo spazio delle prime k coordinate lo é ($k = 1, \dots, n$). La matrice di tale restrizione é la sottomatrice A_k di A formata con le prime k righe e colonne, quindi Q é definita positiva se e solo se A_k ha tutti gli autovalori positivi, $k \geq 1$. Ma a_k é il prodotto di tali autovalori, per cui se Q é definita positiva allora $a_k > 0$.

Viceversa, se $a_k > 0,\ k = 1, \dots, n$, per il corollario precedente si puó scrivere

$$A \;=\; R^T\ D\ R$$

e per la formula (*) D ha tutti gli elementi diagonali positivi. Quindi D, che é diagonale, ha tutti i suoi autovalori positivi e definisce quindi una forma positiva. Poiché A e D sono collegate da un cambiamento di base (prop. 3.1) si conclude che Q é definita positiva.
(ii) Q é definita negativa se e solo se la sua restrizione allo spazio delle prime k coordinate lo é ($k = 1, \dots, n$). La matrice di tale restrizione é la sottomatrice di A formata con le prime k righe e colonne, e deve avere quindi tutti gli autovalori negativi. Poiché a_k é il prodotto degli autovalori, se la forma **e**' definita negativa allora segue subito che gli a_i hanno segni alterni. Per il viceversa, se gli a_i hanno segno alterno, si procede in maniera simile alla parte (i). □

Osservazione. In generale gli autovalori della sottomatrice di A formata dalle prime k righe e colonne non sono anche autovalori di A : ad esempio l'elemento a_{11} costituisce il primo minore principale di A, ma in generale a_{11} non é autovalore di tutta A.