Matematica per Analisi dei Dati, 19.03.09

1. Consideriamo una matrice A quadrata di ordine due; per ogni vettore x in $\mathbb{R}^2$, possiamo moltiplicare la matrice A per x, ottenendo il vettore Ax in $\mathbb{R}^2$.

   Abbiamo cosi' una funzione

   $$f_A : \mathbb{R}^2 \to \mathbb{R}^2$$
   $$f_A(x) = Ax, \qquad x \in \mathbb{R}^2.$$

   Osserviamo che questa funzione si comporta bene rispetto alle operazioni di addizione fra vettori e di moltiplicazione di vettori per scalari:

   $f_A(u + v) = f_A(u) + f_A(v)$,

   $f_A(r\ u) = r\ f_A(u)$,

   per ogni u, v in $\mathbb{R}^2$ ed ogni r in $\mathbb{R}$, in quanto

   $A(u + v) = Au + Av$,

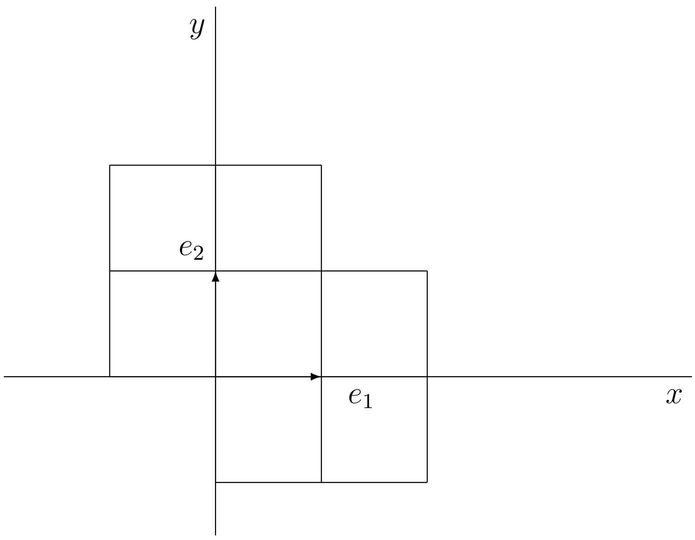
   $A(r\ u) = r\ Au$.

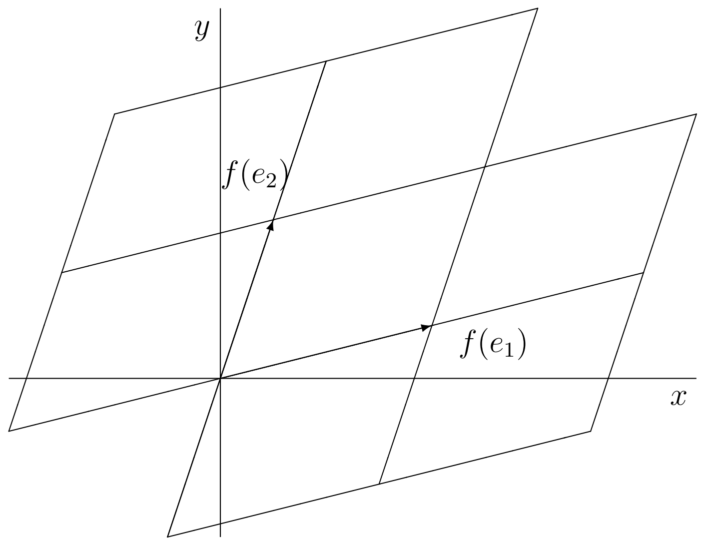
   Cosi' la funzione si comporta bene anche rispetto alla operazione di prendere combinazioni lineari, cioe'

   $f_A(r\ u + s\ v) = r\ f_A(u) + s\ f_A(v)$,

   per ogni u, v in $\mathbb{R}^2$ ed ogni r, s in $\mathbb{R}$.

   Da cio' segue che, se conosciamo i valori di f_A su due vettori che formano una base di $\mathbb{R}^2$, allora conosciamo i valori di f_A su ogni altro vettore di $\mathbb{R}^2$. Di seguito sono mostrate una figura e la sua immagine tramite una funzione f della quale conosciamo i valori sui vettori della base canonica.





Osserviamo che, se la matrice A e' diagonale

$$A = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix},$$

allora la funzione corrispondente e' particolaremte semplice:

$$f_A\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \beta x_2 \end{bmatrix}.$$

Possiamo dire che la funzione f_A e' la composizione di una dilatazione nella direzione dell'asse x, con coefficiente di dilatazione α e di una dilatazione nella direzione dell'asse y, con coefficiente di dilatazione β.

Nei punti successivi vedremo come si possa sempre cercare di ricondursi a questo caso. Vedremo inoltre sotto quali condizioni cio' sia possibile.

2. Consideriamo la matrice

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}.$$

Cerchiamo dei vettori v non nulli in $\mathbb{R}^2$ che vengano mandati dalla funzione f_A in un loro multiplo scalare, cioe' tali che

$$Av = \lambda v,$$

per un opportuno scalare $\lambda \in \mathbb{R}$. Un tale vettore v si dice *autovettore* della matrice A, ed il corrispondente scalare λ si dice *autovalore* della matrice A associato all'autovettore v.

Conviene prima cercare gli autovalori. Uno scalare λ e' un autovalore della matrice A se e' associato a un qualche autovettore di A, cioe' se esiste un vettore non nullo $v = \begin{bmatrix} x \\ y \end{bmatrix}$ di $\mathbb{R}^2$ tale che

$$Av = \lambda v, \qquad \textit{nel nostro caso} \qquad \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}.$$

Ora, questo e' un sistema lineare omogeneo di due equazioni nelle incognite x, y, che dipende dal parametro λ. Lo possiamo riscrivere nella forma

$$(A - \lambda I_2)v = 0_2, \qquad \textit{nel nostro caso} \qquad \begin{bmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0_2.$$

Ora, dire che lo scalare λ e' un autovalore di A significa dire che questo sistema lineare possiede una soluzione non banale. Cio' capita se e solo se il determinante della matrice dei coefficienti e' nullo:

$$\det(A - \lambda I_2) = 0, \qquad \textit{nel nostro caso} \qquad \det \begin{bmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix} = 0.$$

Questa equazione viene detta *equazione caratteristica* della matrice A, e il suo primo membro viene detto *polinomio caratteristico* di A.

Sviluppando il determinante abbiamo l'equazione di secondo grado

$$(2-\lambda)(3-\lambda) - 2 = 0; \qquad \lambda^2 - 5\lambda + 4 = 0$$

le cui soluzioni sono date da

$$\lambda_{1,2} = \frac{5 \pm \sqrt{9}}{2} = \begin{matrix} \frac{5+3}{2} = 4 \\ \frac{5-3}{2} = 1 \end{matrix}.$$

Dunque la matrice A possiede esattamente due autovalori: $\lambda = 4$ e $\lambda = 1$.

Cerchiamo ora gli autovettori di A. Gli autovettori di A cui e' associato un autovalore λ sono vettori v di $\mathbb{R}^2$ tali che

$$A\lambda = \lambda v, \qquad o \qquad (A - \lambda I_2)v = 0_2,$$

cioe'

$$\begin{bmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0_2.$$

Gli autovettori di A associati all'autovalore $\lambda = 4$ sono vettori v di $\mathbb{R}^2$ tali che

$$\begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0_2.$$

Questo sistema di due equazioni si riduce all'unica equazione

$$x - y = 0,$$

le cui soluzioni costituiscono una retta V_4, che possiamo pensare come la retta $V_4 = \mathit{span}\{a\}$ generata dal vettore

$$a = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Gli autovettori di A associati all'autovalore $\lambda = 1$ sono vettori v di $\mathbb{R}^2$ tali che

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0_2.$$

Questo sistema di due equazioni si riduce all'unica equazione

$$x + 2y = 0,$$

le cui soluzioni costituiscono una retta V_1, che possiamo pensare come la retta $V_1 = span\{b\}$ generata dal vettore

$$b = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Ora, l'insieme $\{a, b\}$ e' una base di $\mathbb{R}^2$, e possiamo rappresentare ogni vettore x di $\mathbb{R}^2$ come combinazione lineare

$$x = r\ a + s\ b$$

dei vettori a, b, dove r, s sono le coordinate di x rispetto a questa base.

L'azione di A su x sara' data da

$$Ax = A(r\ a + s\ b) = r\ (Aa) + s\ (Ab) = 4r\ a + s\ b.$$

Possiamo dire che la funzione f_A associata alla matrice A e' la composizione di una dilatazione nella direzione della retta V_4, con coefficiente di dilatazione 4 e di una dilatazione nella direzione della retta V_1, con coefficiente di dilatazione 1.

3. Possiamo riassumere le informazioni ottenute sulla matrice A nel modo seguente. Le uguaglianze

$$Aa = 4a, \qquad Ab = b$$

possono essere riassunte nell'unica uguaglianza

$$\left[\ Aa \mid Ab\ \right] = \left[\ 4a \mid \ \ b\ \right],$$

i cui membri si possono scomporre nella forma

$$A \left[\ a \mid b\ \right] = \left[\ a \mid b\ \right] \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}.$$

Indicata con P la matrice che ha per colonne gli autovettori a, b e indicata con D la matrice diagonale che ha per elementi diagonali gli autovalori associati ad a, b, possiamo sitetizzare l'uguaglianza nella forma

$$AP = PD.$$

Ora, l'insieme $\{a, b\}$ e' una base di $\mathbb{R}^2$, cosi' la matrice P e' invertibile e possiamo ricavare A in funzione di P e D :

$$A = PDP^{-1}.$$

4. Ci sono matrici che non posseggono alcun autovalore in $\mathbb{R}$, e dunque non poseggono alcun autovettore in $\mathbb{R}^2$. Ad esempio, la matrice

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

ha polinomio caratteristico

$$\det(A - \lambda I_2) = \det \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = \lambda^2 + 1,$$

che non ha alcuna radice reale.

5. Ci sono matrici che poseggono autovettori in $\mathbb{R}^2$, ma non abbastanza per formare una base di $\mathbb{R}^2$. Ad esempio, la matrice

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

ha polinomio caratteristico

$$\det(A - \lambda I_2) = \det \begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix} = \lambda^2,$$

che ha due radici reali coincidenti $\lambda_1 = \lambda_2 = 0$.

Gli autovettori di A sono dunque i vettori v di $\mathbb{R}^2$ tali che

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0_2.$$

Questo sistema di due equazioni si riduce all'unica equazione

$$y = 0,$$

le cui soluzioni costituiscono l'asse x.

Chiaramente, con vettori che giacciono tutti su una stessa retta non si puo' formare una base di $\mathbb{R}^2$.

6. Sia ora A una matrice quadrata di ordine n :

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Uno scalare λ e' un autovalore della matrice A se esistono dei vettori non nulli $x \neq 0_n$, che soddisfano la condizione

$$Ax = \lambda x, \qquad cioe' \qquad (A - \lambda I_n)x = 0_n.$$

Questi vettori sono gli autovettori di A associati all'autovalore λ.

Abbiamo cosi' trovato il sistema lineare omogeneo di n equazioni in n incognite

$$\begin{bmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0_n.$$

Ora, questo sistema lineare omogeneo avra' una soluzione non banale $x \neq 0_n$. se e solo se λ e' soluzione dell'equazione

$$Det(A-\lambda I_n) = Det \begin{bmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{bmatrix} = 0.$$

Il polinomio che compare al primo membro di questa equazione e' detto *polinomio caratteristico* della matrice A; e' un polinomio di grado n pari all'ordine della matrice. Abbiamo dunque che gli autovalori di una matrice A di ordine n sono le radici del polinomio caratteristico di A.

Supponiamo che la matrice A possegga n autovalori $\lambda_1, \dots, \lambda_n$ non necessariamente distinti, e siano $v_1, \dots, v_n$ autovettori corrispondenti. Ora, le uguaglianze

$$Av_1 = \lambda_1 v_1, \quad \dots, \quad Av_n = \lambda_n v_n$$

possono essere riassunte nell'unica uguaglianza

$$\left[\ Av_1 \mid \dots \mid Av_n\ \right] = \left[\ \lambda_1 v_1 \mid \dots \mid \lambda_n v_n\ \right],$$

i cui membri si possono scomporre nella forma

$$A \left[\ v_1 \mid \dots \mid v_n\ \right] = \left[\ v_1 \mid \dots \mid v_n\ \right] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Indicata con P la matrice che ha per colonne gli autovettori $v_1, \dots, v_n$ e indicata con D la matrice diagonale che ha per elementi diagonali gli autovalori $\lambda_1, \dots, \lambda_n$ ad essi associati, possiamo sintetizzare l'uguaglianza nella forma

$$AP = PD.$$

Ora, se l'insieme $\{v_1, \dots, v_n\}$ e' una base di $\mathbb{R}^n$, cosi' che la matrice P sia invertibile, possiamo ricavare A in funzione di P e D :

$$A = PDP^{-1}.$$

Forme quadratiche

1. Consideriamo una matrice simmetrica A quadrata di ordine n; per ogni vettore x in $\mathbb{R}^n$, possiamo moltiplicare la matrice A a sinistra per x^T e a destra per x, ottenendo lo scalare x^TAx in $\mathbb{R}$.

   Abbiamo cosi' una funzione

$$\varphi_A : \mathbb{R}^n \to \mathbb{R}$$
$$\varphi_A(x) = x^TAx, \qquad x \in \mathbb{R}^n.$$

   Ora, posto

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix},$$

   si ha

$$\varphi_A\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

   cioe' $\varphi_A(x)$ e' un polinomio omogeneo di secondo grado nelle coponenti di x. Per questa ragione, φ_A si dice *forma quadratica* su $\mathbb{R}^n$.

   Diciamo che una forma quadratica φ e'

   - *semidefinita positiva* se $\varphi(x) \geq 0$ per ogni $x \in \mathbb{R}^n$;
   - *semidefinita negativa* se $\varphi(x) \leq 0$ per ogni $x \in \mathbb{R}^n$;
   - *indefinita* altrimenti.

   Ora, se la matrice A e' diagonale

$$A = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix},$$

   si ha

$$\varphi_A\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} \alpha_i x_i^2.$$

   E' allora facile riconoscere se φ_A e' semidefinita positiva, negativa o indefinita. Precisamente, si ha:

   - φ_A e' semidefinita positiva se e solo se $\alpha_1, \dots, \alpha_n \geq 0$;
   - φ_A e' semidefinita negativa se e solo se $\alpha_1, \dots, \alpha_n \leq 0$;
   - φ_A e' indefinita se e solo se ci sono un α_i e un α_j non nulli e di segno opposto.

   Nello studio delle forme quadratiche ci si puo' sempre ricondurre al caso delle matrici diagonali. Si usano ancora autovettori ed autovalori.