Matematica II 05.05.06

1. Nella lezione scorsa abbiamo descritto e dimostrato, tra l'altro, i seguenti risultati.

   Nello spazio R^n consideriamo un vettore $a \neq 0$, e la retta $l = \{ar;\ r \in R\}$ da esso generata. Per ogni vettore b di R^n esiste uno ed un solo vettore b' della retta l, detto proiezione ortogonale di b su a, tale che

   $$b - b' \perp a,$$

   e si ha

   $$b' = a\ \frac{\langle a, b\rangle}{\langle a, a\rangle};$$

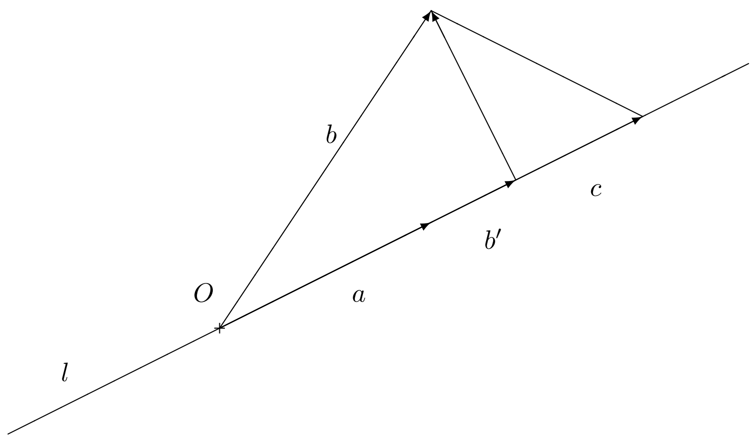
   lo scalare $\frac{\langle a,b\rangle}{\langle a,a\rangle}$ viene detto coefficiente di Fourier di b rispetto ad a.

   Per ogni vettore c sulle retta l si ha

   $d(b, b') \leq d(b, c)$,

   e si pone

   $d(b, l) = d(b, b')$.



   Ci apprestiamo a descrivere una generalizzazione di questi risultati nella quale, al posto del vettore a e dalla retta l da esso generata, compaiono piu' vettori $a_1, a_2, \ldots, a_m$ e lo spazio S da essi generato.

2. Sia

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix}$$

una matrice $n \times m$, e siano

$$a_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} \in R^n, \qquad i = 1, 2, \ldots, m$$

le sue m colonne.

Lo spazio S generato dai vettori $a_1, a_2, \ldots, a_m \in R^n$ e', per definizione, il sottinsieme dello spazio R^n costituito dalle combinazioni lineari

$$a_1 r_1 + a_2 r_2 + \cdots + a_m r_m,$$

al variare dei pesi $r_1, r_2, \ldots, r_m \in R$.

Sostituendo ad ogni a_i la sua scrittura per esteso, queste combinazioni lineari vengono descritte dall'espressione

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} r_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} r_2 + \cdots + \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} r_m,$$

che, sviluppando le operazioni, diventa

$$\begin{bmatrix} a_{11}r_1 + a_{12}r_2 + \cdots + a_{1m}r_m \\ a_{21}r_1 + a_{22}r_2 + \cdots + a_{2m}r_m \\ \vdots \\ a_{n1}r_1 + a_{n2}r_2 + \cdots + a_{nm}r_m \end{bmatrix},$$

che a sua volta, per la definizione di prodotto di matrici, si puo' scrivere nella forma

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix},$$

e, sinteticamente, come

$$Ar,$$

dove

$$r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}$$

e' la colonna degli m pesi.

Riassumendo, possiamo dire che

- *Data una matrice*
$$A = [a_1 \ a_2 \ \dots \ a_m]$$
di tipo $n \times m$, *lo spazio* S *generato dalle sue* m *colonne* $a_i \in R^n$ *e' l'insieme dei vettori di* R^n *del tipo*
$$Ar,$$
dove r *varia fra i vettori di* R^m.

Per definizione, diciamo che un vettore $p \in R^n$ e' ortogonale allo spazio S generato dalle colonne della matrice A, e scriviamo
$$p \perp S$$
se e solo se p e' ortogonale ad ogni vettore di S; per esercizio, si verifichi che cio' capita se e solo se p e' ortrogonale a ciascuna delle colonne a_i di A.

3. **Teorema**

Nello spazio R^n *consideriamo*

- *una matrice*
$A = [a_1 \ a_2 \ \dots \ a_m]$
di tipo $n \times m$, *avente* m *colonne* $a_i \in R^n$;
- *lo spazio*
$S = \{a_1 r_1 + a_2 r_2 + \cdots + a_m r_m; \quad r_1, r_2, \dots, r_m \in R\}$
$= \{Ar; \quad r \in R^m\}$
generato dalle colonne della matrice A;
- *un vettore*
$b \in R^n$.

Allora:

- *esiste uno ed un solo vettore* $b' \in S$ *tale che*
$$b - b' \perp S;$$
- *se le colonne di* A *sono linearmente indipendenti, allora la matrice* $A^T A$ *e' invertiblie e si ha*
$$b' = A \ (A^T A)^{-1} \ A^T b.$$

Il vettore b' si dice *proiezione ortogonale* del vettore b sullo spazio S, e si indica col simbolo

$$pr_S\ b;$$

la colonna

$$(A^T A)^{-1}\ A^T b$$

viene detta *coefficiente di Fourier generalizzato* del vettore b rispetto alla matrice A.

4. Nel caso $m = 1$ il teorema precedente diventa

   Nello spazio R^n consideriamo

   - *un vettore a;*
   - *la retta $S = \{ar; \quad r \in R\}$ generata dal vettore a;*
   - *un vettore $b \in R^n$.*

   Allora:

   - *esiste uno ed un solo vettore $b' \in S$ tale che*

     $$b - b' \perp S;$$

   - *se il vettore $a \neq 0$, allora lo scalare $a^T a$ e' $\neq 0$ e si ha*

     $$b' = a\ (a^T a)^{-1}\ a^T b.$$

   Osserviamo che

   $$(a^T a)^{-1}\ a^T b = \frac{\langle a, b\rangle}{\langle a, a\rangle},$$

   il coefficiente di Fourier di b rispetto ad a.

5. **Esempio per $n = 3$ e $m = 2$.**

   Siano

   $$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix}, \qquad a_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \\ 9 \end{bmatrix} \in R^3.$$

   Lo spazio S generato dalle colonne $a_1, a_2 \in R^3$ e', per definizione, il sottinsieme dello spazio R^3 costituito dalle combinazioni lineari

   $$a_1 r_1 + a_2 r_2,$$

   al variare dei pesi $r_1, r_2 \in R$.

Notiamo che

$$\begin{aligned}
a_1 r_1 + a_2 r_2 &= \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} r_1 + \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} r_2 \\
&= \begin{bmatrix} r_1 + 2r_2 \\ 2r_1 + 3r_2 \\ 2r_1 + 3r_2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \\
&= Ar,
\end{aligned}$$

dove $r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$ e' la colonna dei 2 pesi.

Per il teorema di proiezione ortogonale, possiamo affermare che esiste uno ed un solo vettore

$$b' = a_1 r_1' + a_2 r_2' = Ar'$$

nello spazio S tale che

$$b - b' \perp c$$

per ogni vettore $c = a_1 r_1 + a_2 r_2 = Ar$ nello spazio S; inoltre, poiche' le colonne a_1, a_2 della matrice A sono linearmente indipendenti, la matrice

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

e' invertibile, e si ha

$$\begin{aligned}
pr_S\, b &= A\ (A^T A)^{-1}\ A^T b \\
&= \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix} \left(\begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 9 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 6 & 11 \\ 11 & 22 \end{bmatrix}^{-1} \begin{bmatrix} 20 \\ 32 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & \frac{6}{11} \end{bmatrix} \begin{bmatrix} 20 \\ 32 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 8 \\ -\frac{28}{11} \end{bmatrix}
\end{aligned}$$

$$= \begin{bmatrix} \frac{32}{11} \\ \frac{4}{11} \\ \frac{92}{11} \end{bmatrix}$$

Per esercizio, si verifichi che il vettore $b - pr_S\ b$ e' ortogonale allo spazio S.

Osserviamo che la matrice

$$A^T A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 6 & 11 \\ 11 & 22 \end{bmatrix}$$

e' la tabella dei prodotti scalari delle due colonne della matrice A, mentre la matrice

$$A^T b = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 9 \end{bmatrix} = \begin{bmatrix} 20 \\ 32 \end{bmatrix}$$

e' la colonna dei prodotti scalari delle due colonne della matrice A con il vettore b; infine, i coefficiente di Fourier generalizzato

$$(A^T A)^{-1}\ A^T b = \begin{bmatrix} 8 \\ -\frac{28}{11} \end{bmatrix}$$

fornisce i pesi da assegnare ai vettori a_1 ed a_2 per ottenere la proiezione ortogonale di b sullo spazio S da essi generato:

$$pr_S\ b = 8a_1 - \frac{28}{11}a_2.$$

6. Nello spazio ordinario, consideriamo un sistema di riferimento cartesiano ortogonale monometrico di origine O, e rappresentiamo i vettori di R^3 come segmenti orientati uscenti da O.

   Data una matrice

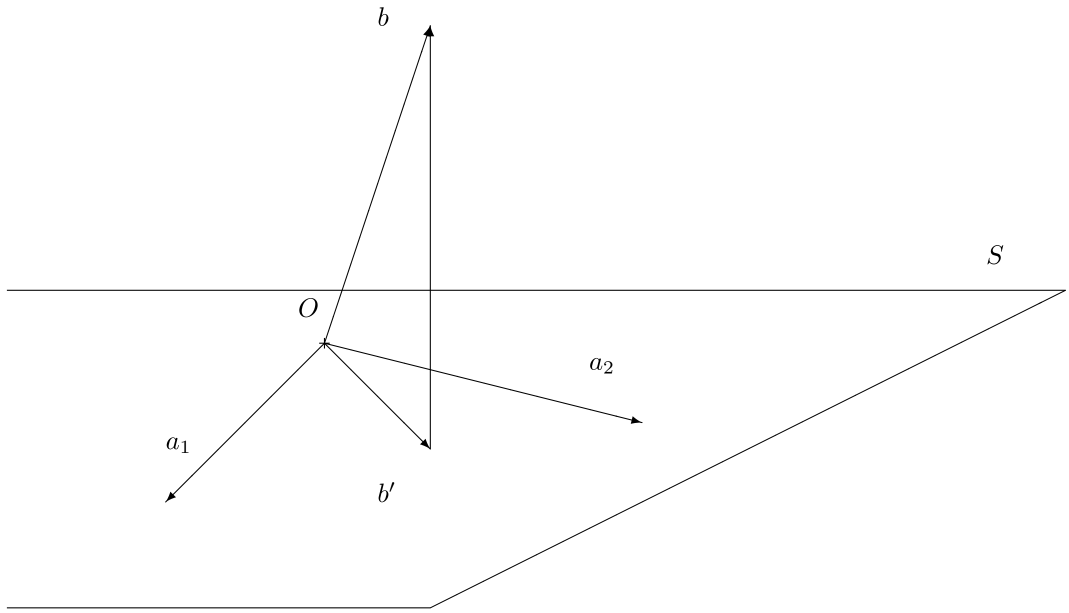
   $A = [a_1\ a_2], \qquad a_1, a_2 \in R^3$

   si ha che le due colonne di A, se sono linearmente indipendenti, generano un insieme di vettori

   $S = \{a_1 r_1 + a_2 r_2;\ r_1, r_2 \in R\}$

   che e' rappresentato da un piano nello spazio ordinario.

   La proiezione ortogonale b' di un vettore b di R^3 su S viene allora rappresentata dall'ordinaria proiezione ortogonale di un vettore su un piano.



7. Nello spazio ordinario, si ha che la proiezione ortogonale di un vettore b su un piano e', fra i vettori del piano, quello piu' vicino al vettore b. Piu' in generale, si ha:

   Teorema *Nello spazio R^n consideriamo un vettore b, lo spazio*

   $S = \{Ar;\ r \in R^m\}$

   generato dalle colonne di una matrice A di tipo $n \times m$, e il vettore $b' = pr_S\ b$ proiezione ortogonale di b su S. Allora b' e', fra i vettori dello spazio S, quello piu' vicino al vettore b, nel senso che

   $d(b, b') \leq d(b, c)$, *per ogni $c \in S$,*

   $d(b, b') = d(b, c)$, *solo se $c = b'$.*

   Siamo cosi' portati a definire la distanza di un vettore b da uno spazio S come la distanza del vettore dalla sua proiezione ortogonale $pr_S\ b$ sullo spazio:

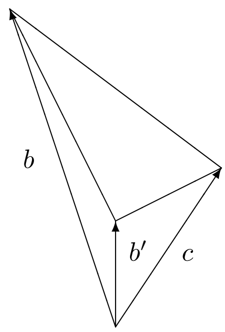
   $d(b, S) = d(b, pr_S\ b)$.

   Dim. Osserviamo che, per ogni vettore c nello spazio S, si ha

$b - b' \perp b' - c$;

infatti, essendo $b - b'$ ortogonale a tutti i vettori di S, in particolare a b' e c, si ha

$\langle b - b', b' - c \rangle = \langle b - b', b' \rangle - \langle b - b', c \rangle = 0 - 0 = 0.$



Allora, per il teorema di Pitagora, si ha

$d(b,c)^2 = d(b,b')^2 + d(b',c)^2$

da cui si puo' dedurre

$d(b,c)^2 \geq d(b,b')^2, \quad cioe' \quad d(b,c) \geq d(b,b').$

Inoltre, se $d(b,c) = d(b,b')$, sempre dal teorema di Pitagora segue $d(b',c)^2 = 0$, cioe' $d(b',c) = 0$, cioe' $b' = c$.