Matematica per Analisi dei Dati, 12.03.09

1. Nello spazio abbiamo visto che, preso un piano α passante per un punto O, i vettori applicati in O ortogonali ad α formano una retta l, e che ogni vettore b applicato in O si puo' scrivere in uno ed un solo modo come somma di un vettore applicato in O che sta sul piano α, la proiezione ortogonale di b su α, e un vettore applicato in O che sta sulla retta l, la proiezione ortogonale di b su l.

   Associando a ciascun vettore la sua proiezione ortogonale sul piano α otteniamo una funzione $\mathrm{pr}_\alpha : \mathbb{R}^3 \to \mathbb{R}^3$ dallo spazio in se'. Associando a ciascun vettore la sua proiezione ortogonale sulla retta l otteniamo un'altra funzione $\mathrm{pr}_l : \mathbb{R}^3 \to \mathbb{R}^3$ dallo spazio in se'. Osserviamo che per ogni vettore b si ha

$$\begin{aligned}
&\mathrm{pr}_\alpha\left(\mathrm{pr}_\alpha(b)\right) = \mathrm{pr}_\alpha(b),\\
&\mathrm{pr}_l\left(\mathrm{pr}_l(b)\right) = \mathrm{pr}_l(b),\\
&\mathrm{pr}_l\left(\mathrm{pr}_\alpha(b)\right) = 0_3,\\
&\mathrm{pr}_\alpha\left(\mathrm{pr}_l(b)\right) = 0_3;
\end{aligned}$$

   queste uguaglianze possono essere espresse sinteticamente nella forma

$$\begin{aligned}
&\mathrm{pr}_\alpha^2 = \mathrm{pr}_\alpha,\\
&\mathrm{pr}_l^2 = \mathrm{pr}_l,\\
&\mathrm{pr}_l \circ \mathrm{pr}_\alpha = 0,\\
&\mathrm{pr}_\alpha \circ \mathrm{pr}_l = 0,
\end{aligned}$$

   dove $\circ$ e' la composizione di funzioni, il quadrato e' inteso nel senso della composizione di funzioni, e 0 indica la funzione $\mathbb{R}^3 \to \mathbb{R}^3$ che manda ogni vettore nel vettore nullo.

2. Abbiamo anche dato la descrizione algebrica di questa costruzione. Ora, per certi piani si ha che questa descrizione risulta particolarmente semplice. Ad esempio, preso come α piano xy, si ha che la retta l e' l'asse z, ed e' facile scomporre il generico vettore $b \in \mathbb{R}^3$ come somma di un vettore che sta sul piano xy e di un vettore che sta sull'asse z :

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ b_3 \end{bmatrix}.$$

   Osserviamo che

$$\mathrm{pr}_\alpha(b) = \mathrm{pr}_\alpha\left(\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\right) = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = Pb.$$

   Cosi', la funzione "proiezione sul piano xy" viene rappresentata dalla matrice P.

   Analogamente, si ha che

$$\mathrm{pr}_l(b) = \mathrm{pr}_l\left(\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\right) = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = Qb.$$

Cosi', la funzione "proiezione sull'asse z" viene rappresentata dalla matrice Q.

Osserviamo che

$$\begin{aligned}
&P + Q = I_3\\
&P^2 = P, \quad Q^2 = Q\\
&PQ = QP = 0_{3\times 3},
\end{aligned}$$

dove I_3 e' la matrice unita' di ordine 3 e $0_{3\times 3}$ e' la matrice nulla di ordine 3.

Una matrice quadrata che coincide con la sua seconda potenza viene detta *idempotente,* in quanto coincidera' con tutte le sue potenze.

3. **Ortogonalita' in $\mathbb{R}^n$.**

   Nello spazio vettoriale $\mathbb{R}^n$, per definizione diciamo unu vettore $u = [u_i]_{i=1}^n$ e' *ortogonale* ad un vettore $v = [v_i]_{i=1}^n$, e scriviamo $u \perp v$, se

   $$\sum_{i=1} u_i v_i = 0.$$

   Questa relazione si puo' scrivere

   $$\begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = 0,$$

   e sintetizzare nella forma

   $$u^T v = 0.$$

4. **Prodotto interno in $\mathbb{R}^n$.**

   Per ogni $u, v \in \mathbb{R}^n$, il numero reale ottenuto da

   $$u^T v$$

   viene detto *prodotto interno* dei vettori u, v. L'operazione che associa a due vettori di $\mathbb{R}^n$ il loro prodotto interno eredita dalle proprieta' delle operazioni sulle matrici le seguenti proprieta':

   $$\begin{aligned}
   &u^T(v_1 + v_2) = u^T v_1 + u^T v_2,\\
   &(u_1 + u_2)^T v = u_1^T v + u_2^T v,\\
   &(ru)^T v = r\left(u^T v\right) = u^T(rv),\\
   &u^T v = v^T u,
   \end{aligned}$$

   dove u, u_i, u_2, v, v_1, v_2 indicano vettori in $\mathbb{R}^n$ ed r indica uno scalare. Inoltre si ha

   $$v^T v = 0 \qquad \textit{se e solo se} \qquad v = 0_n.$$

5. **Complemento ortogonale di un sottospazio**

Nello spazio vettoriale $\mathbb{R}^n$ sia dato un sottospazio V di dimensione r, e sia

$$\{a_1, a_2, \ldots, a_r\} \qquad a_i \in V$$

una base di V. Dunque i vettori $a_1, \ldots, a_n$ sono liearmente indipendenti e V e' costituito dalle loro combinazioni lineari

$$a_1 t_1 + a_2 t_2 + \cdots + a_r t_r, \qquad t_i \in \mathbb{R}.$$

Ora, quest'espressione si puo' riscrivere nella forma

$$\left[\begin{array}{c|c|c|c} & & & \\ a_1 & a_2 & \cdots & a_r \\ & & & \end{array}\right] \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_r \end{bmatrix} = At,$$

dove A e' la matrice di tipo $n \times r$ che ha per colonne i vettori a_i e $t \in BR^r$ e' il vettore che ha per componenti gli scalari t_i. Dunque V e' lo spazio delle colonne della matrice A :

$$V = \mathcal{C}(A),$$

ed A ha rango r :

$$r(A) = r.$$

Cerchiamo ora di esprimere le condizioni sotto le quali un vettore $x \in \mathbb{R}^n$ e' ortogonale ad ogni vettore di V. In particolare, x dovra' essere ortogonale ai vettori $a_1, a_2, \ldots, a_r$, cioe' dovra' soddisfare le condizioni

$$\begin{aligned} x^T a_1 &= 0 \\ x^T a_2 &= 0 \\ &\vdots \\ x^T a_r &= 0 \end{aligned}$$

Ci chiediamo se questo basta per assicurare che x sia ortogonale a tutti i vettori di V. Consideriamo dunque il prodotto interno di x per il generico vettore di V

$$x^T \left(a_1 t_1 + a_2 t_2 + \cdots + a_r t_r\right),$$

dove $t_1, \ldots, t_r$ e' una qualsiasi sequenza di scalari. Questa espressione, in base alle proprieta' del prodotto scalare, si puo' sviluppare in

$$\left(x^T a_1\right) t_1 + \left(x^T a_2\right) t_2 + \cdots + \left(x^T a_r\right) t_r = 0 t_1 + 0 t_2 + \cdots + 0 t_r = 0.$$

La risposta e' dunque affermativa.

Le condizioni sopra riportate si possono riassumere nell'unica condizione

$$x^T \left[\begin{array}{c|c|c|c} & & & \\ a_1 & a_2 & \cdots & a_r \\ & & & \end{array} \right] = 0_r^T,$$

che a sua volta si puo' esprimere sinteticamente nella forma

$$x^T A = 0_r^T,$$

oppure

$$A^T x = 0_r.$$

Abbiamo cosi' trovato che i vettori ortogonali ad ogni vettore di V formano un sottospazio di $\mathbb{R}^n$; questo sottospazio viene detto *complemento ortogonale* di V e viene indicato con $V^\perp$. Inoltre, questo sottospazio risulta essere lo spazio nullo della matrice A^T :

$$V^\perp = \mathcal{N}(A^T),$$

che ha dimensione

$$dim\left(V^\perp\right) = dim\left(\mathcal{N}(A^T)\right) = n - r(A^T) = n - r(A) = n - r.$$

6. **Proiezioni ortogonali su sottospazi**

In analogia con quanto stabilito nel caso dello spazio ordinario, si ha il seguente

Teorema *Sia V un sottospazio di $\mathbb{R}^n$. Ogni vettore $b \in \mathbb{R}^n$ si puo' scrivere in uno ed un solo modo come somma*

$$b = b' + b''$$

di un vettore $b' \in V$ e di un vettore $b'' \in V^\perp$.

Il vettore b' si dice proiezione ortogonale del vettore b sullo spazio V. Associando ad ogni vettore la sua proiezione ortogonale su V si ottiene una funzione

$$\mathrm{pr}_V : \mathbb{R}^n \to \mathbb{R}^n.$$

Il vettore b'' si dice proiezione ortogonale del vettore b sullo spazio $V^\perp$. Associando ad ogni vettore la sua proiezione ortogonale su $V^\perp$ si ottiene una funzione

$$\mathrm{pr}_{V\perp} : \mathbb{R}^n \to \mathbb{R}^n.$$

Osserviamo che

$$\begin{aligned} &\mathrm{pr}_V(v) = v, && \textit{per ogni } v \in V \\ &\mathrm{pr}_V(v) = 0_n, && \textit{per ogni } v \in V^\perp \\ &\mathrm{pr}_{V^\perp}(v) = v, && \textit{per ogni } v \in V^\perp \\ &\mathrm{pr}_{V^\perp}(v) = 0_n, && \textit{per ogni } v \in V. \end{aligned}$$

7. **Formula per la proiezione ortogonale su un sottospazio**

Determiniamo ora una formula per la proiezione ortogonale di un vettore su un sottospazio. Sia V un sottospazio di $\mathbb{R}^n$ avente dimensione r. Comunque presa una base $\{a_1, \ldots, a_r\}$ di V, possiamo rappresentare V come lo spazio delle colonne

$$V = \mathcal{C}(A)$$

della matrice A, di tipo $n \times r$ e rango r, avente per colonne i vettori $a_1, \ldots, a_r$, e possiamo rappresentare il complemento ortogonale $V^\perp$ di V come lo spazio nullo

$$V^\perp = \mathcal{N}(A^T)$$

della matrice A^T, di tipo $r \times n$ e rango r, avente per righe i vettori $a_1, \ldots, a_r$.

Dato un qualsiasi vettore $b \in \mathbb{R}^n$, vogliamo scriverlo come somma

$$b = b' + b''.$$

di un vettore $b' \in V$ e di un vettore $b'' \in V^\perp$. Ora, b' deve potersi scrivere come $b' = At$, per un opportuno $t \in \mathbb{R}^r$; sostituendo nella scomposizione At per b' si ha

$$b = At + b''.$$

D'altro canto, si deve avere $A^T b'' = 0_r$; moltiplicando entrambe i membri della scomposizione a sinistra per A^T si ha

$$A^T b = A^T (At + b''),$$

da cui

$$A^T b = A^T A t,$$

o

$$A^T A t = A^T b.$$

Questa e' un'equazione matriciale nell'incognita $t \in \mathbb{R}^r$. Vedremo in seguito che

> il prodotto della trasposta di una matrice per la matrice e' invertibile,
> a patto che la matrice abbia le colonne indipendenti.

Dunque possiamo ricavare univocamente

$$t = \left(A^T A\right)^{-1} A^T b,$$

e cosi'

$$\begin{aligned} b' &= A\left(A^T A\right)^{-1} A^T b \\ b'' &= b - b' = b - A\left(A^T A\right)^{-1} A^T b = \left(I_n - A\left(A^T A\right)^{-1} A^T\right) b. \end{aligned}$$

Abbiamo cosi' che, per ogni $b \in \mathbb{R}^n$,

$$\begin{aligned} \mathrm{pr}_V(b) &= A\left(A^T A\right)^{-1} A^T b \\ \mathrm{pr}_{V^\perp}(b) &= \left(I_n - A\left(A^T A\right)^{-1} A^T\right) b, \end{aligned}$$

o piu' semplicemente

$$\begin{aligned} \mathrm{pr}_V(b) &= Pb \\ \mathrm{pr}_{V^\perp}(b) &= (I_n - P)\, b, \end{aligned}$$

dove

$$P = A\left(A^T A\right)^{-1} A^T.$$

Diciamo che P e' la matrice che rappresenta la funzione di proiezione ortogonale sulllo spazio V.

8. **Esempio**

Sia $V = \mathcal{C}(A)$ lo spazio delle colonne della matrice di rango 2

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

Allora la matrice che rappresenta la funzione di proiezione ortogonale sulllo spazio V e' data da

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} =$$

$$= \frac{1}{8} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} =$$

$$= \frac{1}{8} \begin{bmatrix} 4 & 2 & 2 & -2 & -2 \\ 2 & 3 & 3 & 1 & 1 \\ 2 & 3 & 3 & 1 & 1 \\ -2 & 1 & 1 & 3 & 3 \\ -2 & 1 & 1 & 3 & 3 \end{bmatrix}.$$

Dunque la funzione proiezione ortogonale sul sottospazio V

$$\mathrm{pr}_V : \mathbb{R}^5 \to \mathbb{R}^5$$

e' data da

$$\mathrm{pr}_V\left(\begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_5 \end{bmatrix}\right) = \frac{1}{8}\begin{bmatrix} 4 & 2 & 2 & -2 & -2 \\ 2 & 3 & 3 & 1 & 1 \\ 2 & 3 & 3 & 1 & 1 \\ -2 & 1 & 1 & 3 & 3 \\ -2 & 1 & 1 & 3 & 3 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_5 \end{bmatrix}.$$

9. **Proprieta' delle matrici di proiezione ortogonale**

   Sia A una matrice di tipo $n \times r$ avente rango r, e sia $V = \mathcal{C}(A)$ lo spazio delle colonne di A. La funzione $\mathrm{pr}_V : \mathbb{R}^n \to \mathbb{R}^n$ di proiezione ortogonale sullo spazio delle colonne di A e' rappresentata dalla matrice

   $$P = A\left(A^T A\right)^{-1} A^T,$$

   che e' quadrata di ordine n.

   Viene naturale chiedersi se P sia idempotente. La risposta e' affermativa:

   $$\begin{aligned} P^2 = PP &= A\left(A^T A\right)^{-1} A^T A \left(A^T A\right)^{-1} A^T \\ &= A\left(A^T A\right)^{-1}\left(A^T A\right)\left(A^T A\right)^{-1} A^T \\ &= A\left(A^T A\right)^{-1} A^T \\ &= P \end{aligned}$$

   Inoltre, P e' simmetrica in quanto

   $$\begin{aligned} P^T &= \left(A\left(A^T A\right)^{-1} A^T\right)^T \\ &= \left(A^T\right)^T\left(\left(A^T A\right)^{-1}\right)^T A^T \\ &= A\left(\left(A^T A\right)^T\right)^{-1} A^T \\ &= A\left(A^T A\right)^{-1} A^T \\ &= P. \end{aligned}$$