

Elementi di Algebra Lineare

Corso di Calcolo Numerico, a.a. 2009/2010

Francesca Mazzia

Dipartimento di Matematica
Università di Bari

13 Marzo 2006

Prodotto matrice vettore

Sia $A \in \mathbb{R}^{m \times n}$, $\mathbf{x} \in \mathbb{R}^n$ di n :

$A\mathbf{x} = \mathbf{b}$ è un vettore di m componenti.

$$a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1$$

$$a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2$$

...

...

$$a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = b_m$$

Numero di operazioni per calcolare b :

- mn moltiplicazioni
- $m(n - 1)$ addizioni.

Le matrici trasformano un vettore \mathbf{x} di dimensione n in un vettore \mathbf{b} di dimensione m .

esempio

$$A = \begin{pmatrix} 1 & 1/2 & 3/2 \\ 1/3 & 2 & 1/2 \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

$$\mathbf{b} = A\mathbf{x} = \begin{pmatrix} 1 & 1/2 & 3/2 \\ 1/3 & 2 & 1/2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} =$$

$$\begin{pmatrix} 1 * 1 + (1/2) * 1 + (3/2) * 2 \\ (1/3) * 1 + 2 * 1 + (1/2) * 2 \end{pmatrix} = \begin{pmatrix} 9/2 \\ 10/3 \end{pmatrix}$$

Posso rappresentare una matrice utilizzando le sue colonne:

$$A = (a_1, a_2, \dots, a_n)$$

Il prodotto $A\mathbf{x} = \mathbf{b}$ può anche scriversi:

$$(a_1, a_2, \dots, a_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{b}$$

$$x_1 a_1 + x_2 a_2 + \dots + x_n a_n = \mathbf{b}$$

cioè il vettore \mathbf{b} si ottiene come combinazione lineare dei vettori colonna della matrice A .

Esempio

$$A = \begin{pmatrix} 1 & 1/2 & 3/2 \\ 1/3 & 2 & 1/2 \end{pmatrix}$$

$$a_1 = \begin{pmatrix} 1 \\ 1/3 \end{pmatrix} \quad a_2 = \begin{pmatrix} 1/2 \\ 2 \end{pmatrix} \quad a_3 = \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix}$$

$$A\mathbf{x} = (a_1, a_2, a_3) \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = a_1 + a_2 + 2a_3 = \mathbf{b}$$

Partizionamento

Il partizionamento è una decomposizione della matrice in sottomatrici.

Esempio

$$\left(\begin{array}{cc|ccc|cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} \end{array} \right)$$

Possiamo scrivere

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}$$

con

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad A_{12} = \begin{pmatrix} a_{13} & a_{14} & a_{15} \\ a_{23} & a_{24} & a_{25} \end{pmatrix}$$

Esempio

$$\left(\begin{array}{cc|ccc|cc} 4 & 5 & 6 & 4 & 3 & 8 & 8 \\ 9 & 2 & 3 & 3 & 5 & 8 & 6 \\ \hline 4 & 6 & 8 & 1 & 1 & 5 & 8 \\ 4 & 8 & 5 & 1 & 6 & 4 & 6 \\ 8 & 0 & 7 & 6 & 3 & 8 & 3 \end{array} \right)$$

con

$$A_{11} = \begin{pmatrix} 4 & 5 \\ 9 & 2 \end{pmatrix} \quad A_{12} = \begin{pmatrix} 6 & 4 & 3 \\ 3 & 3 & 5 \end{pmatrix} \quad A_{13} = \begin{pmatrix} 8 & 8 \\ 8 & 6 \end{pmatrix}$$

$$A_{21} = \begin{pmatrix} 4 & 6 \\ 4 & 8 \\ 8 & 0 \end{pmatrix} \quad A_{22} = \begin{pmatrix} 8 & 1 & 1 \\ 5 & 1 & 6 \\ 7 & 6 & 3 \end{pmatrix} \quad A_{23} = \begin{pmatrix} 5 & 8 \\ 4 & 6 \\ 8 & 3 \end{pmatrix}$$

Se le partizioni sono compatibili possiamo trattare le sottomatrici come scalari nell'eseguire le operazioni tra matrici.

Per esempio:

$$\begin{aligned} & \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \\ & \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix} \\ & \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \\ & \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix} \end{aligned}$$

Prodotto esterno fra vettori

Se $\mathbf{a} \in \mathbb{R}^m$ e $\mathbf{b} \in \mathbb{R}^n$, si ha:

$$W = \mathbf{a}\mathbf{b}^T = \begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_n \\ a_2b_1 & a_2b_2 & \dots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_mb_1 & a_mb_2 & \dots & a_mb_n \end{pmatrix}.$$

Le matrici $W = \mathbf{a}\mathbf{b}^T$ non vengono mai memorizzate completamente ma si utilizza sempre la loro rappresentazione con i vettori \mathbf{a} e \mathbf{b} .

Come eseguiamo il prodotto Wx ?

$$\mathbf{c} = W\mathbf{x} = (\mathbf{a}\mathbf{b}^T)\mathbf{x} = \mathbf{a}(\mathbf{b}^T\mathbf{x}) = (\mathbf{b}^T\mathbf{x})\mathbf{a}$$

$\mathbf{b}^T\mathbf{x}$ è uno scalare quindi calcoliamo prima

$$\mu = \mathbf{b}^T\mathbf{x}$$

e poi

$$\mathbf{c} = \mu\mathbf{a}$$

Numero di operazioni:

- $2n$ moltiplicazioni
- $n - 1$ addizioni.

Esempio

$$W = \mathbf{ab}^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} (2 \ 1) = \begin{pmatrix} 2 & 1 \\ 4 & 2 \\ 6 & 3 \end{pmatrix}$$

Dato

$$x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

Calcoliamo prima $b^T x = 10$ e poi $Wx = 10a$.

Inversa di un matrice

Sia $A \in \mathbb{R}^{n \times n}$.

Se esiste una matrice $A^{-1} \in \mathbb{R}^{n \times n}$ tale che

$$A^{-1}A = AA^{-1} = I$$

allora A^{-1} viene chiamata inversa della matrice A .

Se l'inversa esiste, essa è unica.

Infatti sia B tale che $BA = AB = I$, allora:

$$B = BI = BAA^{-1} = IA^{-1} = A^{-1}$$

.

Valgono le seguenti proprietà



$$(AB)^{-1} = B^{-1}A^{-1},$$

infatti:

$$(AB)B^{-1}A^{-1} = A^{-1}B^{-1}BA = I.$$



$$(A^T)^{-1} = (A^{-1})^T.$$

Una matrice che non ammette l'inversa è detta singolare.

Se A è non singolare, lo è anche A^{-1} .

Infatti

$$(A^{-1})^{-1} = A$$

.

Dipendenza lineare di vettori

Dati r vettori $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r$, diversi dal vettore nullo, se

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_r \mathbf{x}_r = \mathbf{0}$$

con gli $\alpha_i \neq 0$, i vettori si dicono

linearmente dipendenti.

Se

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_r \mathbf{x}_r = \mathbf{0}$$

solo se $\alpha_i = 0$ per $i = 1, 2, \dots, r$ allora i vettori \mathbf{x}_i , $i = 1, 2, \dots, r$ sono

linearmente indipendenti.

Se vi è un $\alpha_i \neq 0$, ve ne deve essere almeno un altro perchè nessuno degli \mathbf{x}_i è nullo. Quindi alcuni dei vettori \mathbf{x}_i possono esprimersi come combinazione lineare degli altri.

Infatti se $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r$ sono **l.d.** con $\alpha_r \neq 0$ si ha:

$$\mathbf{x}_r = -\frac{\alpha_1}{\alpha_r}\mathbf{x}_1 - \frac{\alpha_2}{\alpha_r}\mathbf{x}_2 - \dots - \frac{\alpha_{r-1}}{\alpha_r}\mathbf{x}_{r-1}.$$

Se i vettori \mathbf{x}_i sono **l.i.**, un qualunque sottoinsieme di essi è ancora **l.i.**.
Una base è l'insieme minimo di vettori **l.i.** che generano un insieme dato.
La base canonica di \mathbb{R}^m è costituita dai vettori

$$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m$$

con elementi $(\mathbf{e}_i)_j = 0$ per $i \neq j$ e $(\mathbf{e}_i)_i = 1$.

Rango

Dato un insieme di vettori

$$\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}, \quad \mathbf{x}_i \in \mathbb{R}^m$$

si dice **rango** dell'insieme il numero massimo di vettori linearmente indipendenti che è possibile estrarre dall'insieme.

Poichè il massimo numero di vettori **l.i.** in tutto \mathbb{R}^m è m , il rango di un insieme di vettori di \mathbb{R}^m non può superare m .

Data una matrice A , si dice rango della matrice, e si indica con $\text{rank}(A)$, il rango dell'insieme dei suoi vettori colonna.

Vettori Ortogonali e Unitari

Siano \mathbf{x} e $\mathbf{y} \in \mathbb{R}^n$ se $\mathbf{x}^T \mathbf{y} = 0$ i due vettori si dicono ortogonali.

$$\mathbf{x}^T \mathbf{x} = \sum_{i=1}^n x_i^2,$$

quindi

$$\mathbf{x}^T \mathbf{x} > 0$$

Se $\mathbf{x}^T \mathbf{x} = 1$, il vettore \mathbf{x} si dice unitario.

Vettori ortogonali

Un insieme di vettori tra loro ortogonali sono **l.i.**.

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_r \mathbf{x}_r = \mathbf{0}$$

moltiplichiamo, ad esempio, per \mathbf{x}_1^T

$$\alpha_1 \mathbf{x}_1^T \mathbf{x}_1 + \alpha_2 \mathbf{x}_1^T \mathbf{x}_2 + \cdots + \alpha_r \mathbf{x}_1^T \mathbf{x}_r = 0$$

Essendo ortogonali $\mathbf{x}_1^T \mathbf{x}_2 = 0, \dots, \mathbf{x}_1^T \mathbf{x}_r = 0$ quindi

$$\alpha_1 \mathbf{x}_1^T \mathbf{x}_1 + \alpha_2 \mathbf{x}_1^T \mathbf{x}_2 + \cdots + \alpha_r \mathbf{x}_1^T \mathbf{x}_r = \alpha_1 \mathbf{x}_1^T \mathbf{x}_1$$

questo implica che $\alpha_1 = 0$.

Sistemi lineari

Non è sempre vero che fissato un vettore \mathbf{b} di dimensione m , si possa sempre trovare un vettore \mathbf{x} per cui $A\mathbf{x} = \mathbf{b}$ sia soddisfatta.

Trovare \mathbf{x} tale che:

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

è il problema della soluzione di sistemi lineari.

Se \mathbf{b} è il vettore nullo, il sistema lineare $A\mathbf{x} = 0$ si dice omogeneo.

Consideriamo le due matrici A ed

$$\hat{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{pmatrix} = (A, \mathbf{b})$$

detta matrice ampliata, avente in comune con A le prime n colonne, mentre l'ultima colonna è il vettore \mathbf{b} .

Si ha il seguente Teorema, noto come teorema di *Rouché-Capelli*:

Teorema

Condizione necessaria e sufficiente affinché il sistema $Ax = b$ ammetta soluzioni è che il rango di A e quello di \hat{A} coincidano.

Se A è una matrice quadrata di ordine n allora le seguenti proposizioni sono equivalenti:

- 1 Per ogni vettore $b \in \mathbb{R}^n$, il sistema $Ax = b$ ammette soluzione
- 2 Se una soluzione del sistema esiste essa è unica
- 3 Per ogni $x \in \mathbb{R}^n$, $Ax = 0$ implica $x = 0$
- 4 Le colonne (righe) di A sono linearmente indipendenti
- 5 La matrice A è invertibile
- 6 $\det(A) \neq 0$

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j}),$$

L'ultima relazione può essere fuorviante, perchè il determinante cambia scalando la matrice. Infatti

$$\det(\sigma A) = \sigma^n \det(A)$$

Se $n = 30$ e $\det(A) = 1$ allora

$$\det(0.1A) = 10^{-30}$$

Non è facile verificare al calcolatore quando questa quantità è veramente 0. Il calcolo del determinante con la regola di Laplace richiede più di $n!$ moltiplicazioni,

$$n = 16$$

il nostro computer esegue una moltiplicazione ogni 10^{-10} secondi
tempo di calcolo : $16! \cdot 10^{-10} / 60 \approx 35$ minuti

Autovalori ed autovettori

Data una matrice quadrata $A \in \mathbb{R}^{n \times n}$, un numero complesso λ si dice autovalore di A se esiste un vettore $u \neq 0$ a componenti reali o complessi tale che:

$$Au = \lambda u,$$

o

$$(A - \lambda I)u = 0,$$

u si dice autovettore di A ed è soluzione non nulla di un sistema lineare omogeneo.

Tale soluzione esiste se $\det(A - \lambda I) = 0$. Si dimostra che $\det(A - \lambda I) = p(\lambda)$ è un polinomio di grado n in λ che ha n radici nel campo complesso (polinomio caratteristico).

Si dice raggio spettrale della matrice A , e lo si indica con $\rho(A)$, il massimo modulo degli autovalori di A .

Se u è un autovettore di A , anche αu , con $\alpha \neq 0$ è autovettore di A .
Gli autovettori unitari o normalizzati sono tali che $u^T u = 1$.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

autovalori di A : $\lambda_1 = 3$ e $\lambda_2 = -1$, autovettori:

$$u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{e} \quad u_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

autovettori normalizzati di A :

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{e} \quad u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Norme

Abbiamo definito i concetti di errore assoluto e relativo per gli scalari e visto come usarli per verificare la qualità dei nostri algoritmi.

Gli errori misurano la distanza fra la soluzione vera e quella calcolata.

Se lavoriamo con matrici e vettori abbiamo bisogno di generalizzare il concetto di valore assoluto, per poter misurare la distanza fra vettori e matrici.

Norme vettoriali

Sia $x \in \mathbb{R}^n$, si chiama norma di x , e si indica con $\|x\|$, un numero non negativo che soddisfa le seguenti condizioni:

- N1) se x è il vettore nullo, $\|x\| = 0$ e viceversa;
- N2) se λ è uno scalare, $\|\lambda x\| = |\lambda| \|x\|$;
- N3) se x ed y sono due vettori, $\|x + y\| \leq \|x\| + \|y\|$
(disuguaglianza triangolare)

Sia $x = (x_1, x_2, \dots, x_n)^T$

Norma ∞

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Norma 1.

$$\|x\|_1 = \sum_{i=1}^n |x_i|$$

Norma 2.

$$\|x\|_2 = (x^T x)^{1/2}$$

Norme matriciali

Sia $A \in \mathbb{R}^{m \times n}$, si chiama norma di A , e si indica con $\|A\|$, un numero non negativo che soddisfa le seguenti condizioni:

M1) $\|A\| = 0$ se e solo se A è la matrice nulla;

M2) Se λ è uno scalare allora $\|\lambda A\| = |\lambda| \|A\|$;

M3) Se B è una matrice avente le stesse dimensioni di A allora:

$$\|A + B\| \leq \|A\| + \|B\|;$$

M4) Se B è una matrice il cui numero di righe è uguale al numero di colonne di A , allora:

$$\|AB\| \leq \|A\| \|B\|.$$

Norma matriciale indotta da una norma vettoriale

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

Una norma matriciale è compatibile con una norma su vettori se dato $y \in \mathbb{R}^n$,

$$\|Ay\| \leq \|A\| \|y\|.$$

Le norme matriciali indotte sono compatibili con le associate norme vettoriali.

Norma ∞

$$\|A\|_{\infty} = \max \sum_{j=1}^n |a_{i,j}|.$$

Norma 1

$$\|A\|_1 = \max \sum_{i=1}^m |a_{i,j}|.$$

Norma 2

$$\|A\|_2 = \sqrt{\rho(A^T A)}$$