Richiami di Algebra Lineare

Fabrizio Silvestri

December 14, 2010

Matrice

Sia $\mathbb{R}$ il campo dei numeri reali. Si indica con $\mathbb{R}^{m\times n}$ l'insieme delle matrici ad elementi reali con m righe ed n colonne. Se $A \in \mathbb{R}^{n\times n}$, si dice che A è una *matrice quadrata* di ordine n.

Generalmente le matrici vengono indicate con la lettera maiuscola, mentre i loro elementi sono indicati con la lettera minuscola seguita dai due indici (indice di rga e indice di colonna): ad esempio a_{ij} è elemento della matrice A. Si usa scrivere

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

Gli elementi a_{ij} tali che $i = j$ vengono detti *elementi diagonali* o *principali* di A e formano la *diagonale principale* di A.

Vettore

Un vettore $v \in \mathbb{R}^n$ è definito come la matrice colonna formata da n righe e una sola colonna.

Generalmente i vettori vengono indicati con la lettera minuscola, si usa scrivere

$$v = [v_1, v_2, \ldots, v_n]$$

per indicare un vettore in $\mathbb{R}^n$ e gli elementi si indicano con il simbolo v_i, con $1 \le i \le n$.

La matrice trasposta

Sia $A \in \mathbb{R}^{m\times n}$, si definisce la matrice trasposta $B = A^T$ con $B \in \mathbb{R}^{n\times m}$ come segue: $b_{ij} = a_{ji}\ \forall 1 \le i \le m \quad 1 \le j \le n$.

Semplicemente, quindi, la matrice trasposta corrisponde alla matrice di partenza dove le righe sono state invertite con le colonne. Ad esempio

$$A = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 2 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \qquad B = A^T = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}$$

Tipi speciali di matrici

- *Matrice diagonale* è una matrice quadrata di ordine n in cui tutti gli elementi diversi dalla diagonale principale sono uguali a zero, cioè $A \in \mathbb{R}^{n\times n}$ è diagonale se $a_{ij} = 0$ quando $i \neq j$.
- *Matrice identità* di ordine n $I_n \in \mathbb{R}^{n\times n}$ è la matrice diagonale in cui tutti gli elementi sulla diagonale sono uguali ad 1, cioè $a_{ii} = 1$, per $1 \leq i \leq n$. Quando la dimensione n è implicita la matrice identità sarà indicata semplicemente con I.
- *Matrice simmetrica.* Una matrice $A \in \mathbb{R}^{n\times n}$ si dice simmetrica quando $A = A^T$.
- *Matrice unitaria.* Una matrice $A \in \mathbb{R}^{n\times n}$ si dice unitaria quando $AA^T = A^T A = I$.

Determinante di una matrice

Il determinante è una funzione che associa ad ogni matrice quadrata A uno scalare che ne sintetizza alcune proprietà algebriche.
Il significato geometrico principale del determinante si ottiene interpretando la matrice quadrata A di ordine n come trasformazione lineare di uno spazio vettoriale a n dimensioni: con questa interpretazione, il valore assoluto di $\det(A)$ è il fattore con cui vengono modificati i volumi degli oggetti contenuti nello spazio. Se è diverso da zero, il segno del determinante indica inoltre se la trasformazione A preserva o cambia l'orientazione dello spazio.
Il determinante di una matrice può essere espresso utilizzando la *regola di Laplace*. Indicata con A_{ij} la sottomatrice quadrata di ordine $n-1$ ottenuta dalla matrice A eliminando la i-esima riga e la j-esima colonna, per un qualunque indice di riga i si ha:

$$\det(A) = \begin{cases} a_{11} & \text{se } n = 1, \\ \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) & \text{se } n > 1. \end{cases}$$

Calcolare il determinante di una matrice è un compito molto oneroso in termini di tempo di calcolo. Esistono, però, in pratica dei metodi semplici da applicare. Supponiamo di voler calcolare il determinante della matrice

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

Si puó utilizzare lo sviluppo di Laplace secondo una riga o una colonna. Conviene scegliere una riga o una colonna con molti zeri, in modo da ridurre gli addendi dello sviluppo; nel nostro caso sviluppiamo secondo la seconda colonna:
$\det(A) = (-1)^{1+2} \cdot 2 \cdot \det \begin{bmatrix} -1 & 3 \\ 2 & -1 \end{bmatrix} + (-1)^{2+2} \cdot 1 \cdot \det \begin{bmatrix} -2 & -3 \\ 2 & -1 \end{bmatrix} + (-1)^{3+2} \cdot 0 \cdot \det \begin{bmatrix} -2 & -3 \\ -1 & 3 \end{bmatrix} =$
$= (-2) \cdot ((-1)(-1) - 2 \cdot 3) + 1 \cdot ((-2) \cdot (-1) - 2 \cdot (-3)) = (-2)(-5) + 8 = 18$

Proprietà elementari del determinante

- Se tutti gli elementi di una riga (o colonna) di A sono nulli, allora $\det(A) = 0$.
- Se A ha due righe (o colonne) eguali, o proporzionali, allora $\det(A) = 0$.
- $\det(AB) = \det(A)\det(B)$.
- Se $A \in \mathbb{R}^n$ allora $\det(\lambda A) = \lambda^n \det(A)$, dove λ è uno scalare.

Matrice inversa

Una matrice quadrata $A \in \mathbb{R}^{n \times n}$ è detta invertibile o regolare se esiste una matrice B tale che:

$$AB = BA = I$$

dove I denota la matrice identità di ordine n e la moltiplicazione usata è l'ordinaria moltiplicazione di matrici. Se è questo il caso, allora la matrice B è univocamente determinata da A ed è chiamata l'inversa di A, indicata con A^{-1}.

Come prima osservazione notiamo che se A è una matrice unitaria allora $A^T = A^{-1}$. Inoltre, una matrice è invertibile se, e solo se, il suo determinante è non nullo, cioè data $A \in \mathbb{R}^{n \times n}$, A è invertibile se, e solo se, $\det(A) \neq 0$.

Proprietà elementari delle matrici invertibili

- $\det(A^{-1}) = \det(A)^{-1}$.
- $(AB)^{-1} = B^{-1}A^{-1}$

L'algoritmo di *Gauss-Jordan* può essere usato per trovare (quando esiste) l'inversa di una matrice. Funziona nel modo seguente: sia A una matrice invertibile. Si costruisce la matrice $B = (A|I)$ con n righe e $2n$ colonne affiancando A e la matrice identità I. A questo punto si applica l'algoritmo di Gauss-Jordan alla nuova B. Questo algoritmo trasforma la matrice B in una matrice a scalini, che sarà del tipo $(I|C)$. La matrice C così trovata è proprio l'inversa di A.

L'algoritmo di Gauss-Jordan esula dagli scopi di questo corso e non sarà presentato in queste note. Per completezza se ne riportano qui i passi:

1. Se la prima riga ha il primo elemento nullo, scambiala con una riga che ha il primo elemento non nullo. Se tutte le righe hanno il primo elemento nullo, vai al punto 3.

2. Per ogni riga A_i con primo elemento non nullo, eccetto la prima ($i > 1$), moltiplica la prima riga per un coefficiente scelto in maniera tale che la somma tra la prima riga e A_i abbia il primo elemento nullo (quindi coefficiente $= -\frac{A_{i1}}{A_{11}}$). Sostituisci A_i con la somma appena ricavata.

3. Adesso sulla prima colonna tutte le cifre, eccetto forse la prima, sono nulle. A questo punto ritorna al punto 1 considerando la sottomatrice che ottieni cancellando la prima riga e la prima colonna.

Prodotto scalare di due vettori

Siano $a \in \mathbb{R}^n$ e $B \in \mathbb{R}^n$ due vettori di reali. Si definisce il prodotto scalare di a per b come lo scalare $c \in \mathbb{R}$, $c = \langle a, b \rangle$ definito come segue:

$$c = \langle a, b \rangle = \sum_{k=1}^{n} a_k b_k$$

Ad esempio il prodotto scalare tra $a = [1, 3, 5, -1]^T$ e $b = [1, 2, 1, 4]^T$, è data da $\langle a, b \rangle = 1 \cdot 1 + 3 \cdot 2 + 5 \cdot 1 + (-1) \cdot 4 = 8$.

Norma 2 (L_2) di un vettore

Sia $v \in \mathbb{R}^n$ un vettore di numeri reali. Si definisce la norma due di v come:

$$\|v\|_2 = \left(\sum_{i=1}^{n} |v_i|^2 \right)^{\frac{1}{2}} = \sqrt{\sum_{i=1}^{n} (v_i)^2}$$

Ad esempio la norma due del vettore $v = [1, 3, 5, -1]^T$ è data da $\|v\|_2 = \sqrt{1^2 + 3^2 + 5^2 + (-1)^2} = 6$.

La norma L_2 al quadrato di un vettore v è anche definibile come il prodotto scalare di v per sè stesso, $\|v\|_2^2 = \langle v, v \rangle$.

Prodotto tra matrici

Siano $A \in \mathbb{R}^{m \times n}$ e $B \in \mathbb{R}^{n \times r}$ due matrici di numeri reali. Si definisce il prodotto di A per B come la matrice $C \in \mathbb{R}^{m \times r}$, $C = AB$, dove ogni elemento di C, c_{ij}, $1 \leq i \leq m$ e $1 \leq j \leq r$ viene definito come

$$c_{ij} = (AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Tale prodotto viene denominato anche *prodotto riga-colonna* perché, in pratica, l'elemento c_{ij} si ottiene calcolando il prodotto scalare tra la riga i-esima di A e la colonna j-esima di B.

Facciamo un esempio. Si considerino le seguenti due matrici:

$$A = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 2 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

La matrice prodotto C viene calcolata in questo modo:

$$\begin{aligned}
c_{11} &= 1 \cdot 1 + 0 \cdot 2 + \tfrac{1}{2} \cdot 1 = \tfrac{3}{2} \\
c_{12} &= 1 \cdot 0 + 0 \cdot 0 + \tfrac{1}{2} \cdot 2 = 1 \\
c_{13} &= 1 \cdot 0 + 0 \cdot 0 + \tfrac{1}{2} \cdot 1 = \tfrac{1}{2} \\
c_{21} &= 2 \cdot 1 + 0 \cdot 2 + 0 \cdot 1 = 2 \\
c_{22} &= 2 \cdot 0 + 0 \cdot 0 + 0 \cdot 2 = 0 \\
c_{23} &= 2 \cdot 1 + 0 \cdot 0 + 0 \cdot 1 = 2 \\
c_{31} &= 1 \cdot 1 + 2 \cdot 2 + 1 \cdot 1 = 6 \\
c_{32} &= 1 \cdot 0 + 2 \cdot 0 + 1 \cdot 2 = 2 \\
c_{33} &= 1 \cdot 1 + 2 \cdot 0 + 1 \cdot 1 = 2
\end{aligned}$$

Quindi, la matrice C corrisponde alla seguente

$$\begin{bmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 2 & 0 & 2 \\ 6 & 2 & 2 \end{bmatrix}$$

NOTA! Attenzione, è importante notare che nel caso della moltiplicazione tra matrici non vale la proprietà commutativa. Non è vero, infatti, in generale che $AB = BA$. Prima di tutto, infatti, non è detto che le dimensioni delle due matrici siano compatibili. Ad esempio se $A \in \mathbb{R}^{n \times m}$ e $B \in \mathbb{R}^{m \times r}$ con $r \neq m$ il prodotto AB può essere calcolato mentre BA no.

Inoltre, anche in caso di compatibilità, non è sempre detto che questo dia lo stesso risultato, come dimostra il seguente esempio:

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

Abbiamo che

$$AB = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \qquad BA = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

e le due matrici sono chiaramente diverse.

Un caso particolare, ma che vedremo molto frequentemente, è quello della moltiplicazione tra un vettore ed una matrice.

Prodotto vettore matrice

Siano $A \in \mathbb{R}^{m\times n}$ una matrice di numeri reali e sia $v \in \mathbb{R}^n$ un vettore n-dimensionale. Si definisce il prodotto di A per v come il vettore $p \in \mathbb{R}^{m\times 1}$, $p = Av$, dove ogni elemento di p, c_i, $1 \leq i \leq m$ viene definito come

$$p_i = (Av)_i = \sum_{k=1}^{n} a_{ik}v_k$$

Ad esempio, siano date la seguente matrice e il seguente vettore.

$$A = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 2 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \qquad v = [1, 0, 1]^T$$

Il prodotto Av viene definito come il vettore $c = Av = [\frac{3}{2}, 1, \frac{1}{2}]$, ottenuto tramite:

$$\begin{aligned} c_{11} &= 1 \cdot 1 + 0 \cdot 2 + \tfrac{1}{2} \cdot 1 = \tfrac{3}{2} \\ c_{12} &= 1 \cdot 0 + 0 \cdot 0 + \tfrac{1}{2} \cdot 2 = 1 \\ c_{13} &= 1 \cdot 0 + 0 \cdot 0 + \tfrac{1}{2} \cdot 1 = \tfrac{1}{2} \end{aligned}$$

Autovalori e Autovettori

Siano $A \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{R}$ e $x \in \mathbb{R}^n$, $x \neq 0$, tali che valga la relazione $Ax = \lambda x$. Allora λ è detto *autovalore* di A ed x è detto *autovettore* corrispondente a λ. L'insieme degli autovalori di una matrice A costituisce lo *spettro* di A e il modulo massimo $\rho(A)$ degli autovalori di A è detto *raggio spettrale* di A.

Il sistema $Ax = \lambda x$ si può riscrivere come $(A - \lambda I)x = 0$, il quale ammette soluzioni non nulle se e solo se $\det(A - \lambda I) = 0$. Sviluppando $\det(A - \lambda I)$ risulta $\det(A - \lambda I) = P(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \ldots + a_n$, in cui $a_0 = (-1)^n$, $a_i = (-1)^{n-1}\sigma_i, \quad i = 1, \ldots, n$, dove σ_i è la somma dei dterminanti delle $\binom{n}{i}$ sottomatrici principali di A di ordine i. In particolare risulta: $a_1 = (-1)^{n-1}\mathrm{tr(A)}$, $a_n = \det(A)$, in cui si è indicato con $\mathrm{tr}(A)$ la *traccia* di A, ossia la somma degli elementi principali di A.

Il polinomio $P(\lambda)$ detto *polinomio caratteristico* di A e l'equazione $P(\lambda) = 0$ è detta equazione caratteristica.

Matrice Diagonalizzabile

Una matrice $A \in \mathbb{R}^n$ si dice diagonalizzabile se è *simile* ad una matrice diagonale. In altre parole $A \in \mathbb{R}^n$ è diagonalizzabile se e solo se esistono due matrici $D, S \in \mathbb{R}^n$ con D diagonale e S non singolare tale per cui $A = SDS^{-1}$. Inoltre, se una matrice e' diagonalizzabile allora esiste una *base* formata dagli autovettori di A. In altre parole, e' possibile scrivere un qualsiasi vettore x_0 come $x_0 = \sum_{i=1}^{n} \alpha_i v_i$, dove $v_1, v_2, \ldots, v_n$ sono gli autovettori di A.

Il Power Method per il calcolo dell'autovettore principale

Sia $A \in \mathbb{R}^{n\times n}$ diagonalizzabile con autovalori $\lambda_1, \lambda_2, \dots, \lambda_n$ tali che $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.
Fissato un vettore arbitrario $x_0 \in \mathbb{R}^n$, si genera la successione $x_k = Ax_{k-1}$ per $k = 1, 2, \dots$
Si può allora dimostrare che la successione degli x_k tende all'autovettore relativo all'autovalore di modulo massimo.

La dimostrazione della correttezza del metodo è banale.

Siccome A è per ipotesi diagonalizzabile, allora esiste una base di autovettori $v_1, v_2, \dots, v_n$. Allora il vettore x_0 può essere riscritto come $x_0 = \sum_{i=1}^{n} \alpha_i v_i$ e conseguentemente, per la definizione di autovalore di una matrice, il risultato della k-esima iterazione come:

$$x_k = A^k x_0 = \sum_{i=1}^{n} \alpha_i A^k v_i = \sum_{i=1}^{n} \alpha_i \lambda^k v_i$$

Adesso, per comodità denotazionale, spostiamo gli indici da pedice in apice: $x_k = x^{(k)}$ e $x^{(k)} = \sum i = 1^n \alpha_i \lambda^k v^{(i)}$.

Indicando con $x_r^{(k)}$ e $v_r^{(i)}$ le componenti r-esime dei vettori $x^{(k)}$ e $v^{(i)}$, per gli indici j per cui $x_j^{(k)} \neq 0$ e $v_j^{(1)} \neq 0$ si ha:

$$\frac{x_j^{(k+1)}}{x_j^{(k)}} = \lambda_1 \frac{\alpha_1 v_j^{(1)} + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k+1} v_j^{(i)}}{\alpha_1 v_j^{(1)} + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k} v_j^{(i)}}$$

E poiché $\frac{\lambda_i}{\lambda_1} < 1$ per $i > 1$ si ha:

$$\lim_{k\to\infty} \frac{x_j^{(k+1)}}{x_j^{(k)}} = \lambda_1$$

Quindi da un certo indice k in poi l'autovalore λ_1 può essere approssimato con il rapporto indicato.

Con questo metodo si può approssimare anche l'autovettore v_1. Infatti si ha:

$$\frac{x^{(k+1)}}{x_j^{(k)}} = \lambda_1 \frac{\alpha_1 v^{(1)} + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k+1} v^{(i)}}{\alpha_1 v_j^{(1)} + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k} v_j^{(i)}}$$

E passando al limite

$$\lim_{k\to\infty} \frac{x^{(k+1)}}{x_j^{(k)}} = \frac{v^{(1)}}{v_j^{(1)}}$$

Che è l'autovettore $v^{(1)}$ opportunamente normalizzato.

NOTA. Il metodo delle potenze non viene praticamente mai implementato nella formulazione data, poiché dopo pochi passi si potrebbero avere problemi di underflow o overflow. Per evitare questi problemi è necessario eseguire ad ogni passo una normalizzazione del vettore ottenuto, costruendo una successione $x_k = Ax_{k-1}$ per $k = 1, 2, \dots$ definita nel seguente modo:

Il Power Method con normalizzazione ad ogni passo

Sia $A \in \mathbb{R}^{n \times n}$ diagonalizzabile con autovalori $\lambda_1, \lambda_2, \dots, \lambda_n$ tali che $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.
Fissato un vettore arbitrario $x_0 \in \mathbb{R}^n$, si genera la successione x_k nel seguente modo

$$\begin{cases} y_k = Ax_{k-1} \\ x_k = \frac{y_k}{\beta_k} \end{cases}$$

Dove β_k è uno scalare tale che $\|x\| = 1$. Tipicamente viene utilizzata la norma due di y_k per cui l'equazione definita sopra diventa:

$$\begin{cases} y_k = Ax_{k-1} \\ x_k = \frac{y_k}{\|y_k\|_2} \end{cases}$$