MATRICI E SISTEMI LINEARI

Matrici ridotte a scala, indipendenza lineare

Definizione. Una riga di una matrice M si dice non nulla se possiede almeno un elemento diverso da 0. Il primo elemento diverso da 0 (da sinistra) si chiama **indicatore** o **elemento speciale** della riga.

Definizione. Una matrice M si dice **ridotta per righe** se in ogni sua riga non nulla esiste almeno un elemento non nullo che ha al di sotto solo zeri oppure nessun elemento. Una matrice si dice **ridotta per colonne** se la sua trasposta è ridotta per righe.

Definizione. Una matrice M si dice **ridotta a scala** se in ogni sua riga non nulla esiste un elemento non nullo che ha alla sua sinistra e al di sotto solo zeri oppure nessun elemento, ossia:

- non ci sono due indicatori nella stessa colonna
- percorrendola dall' alto verso il basso, gli indicatori si spostano da sinistra a destra
- eventuali righe nulle sono tutte in basso.

Esempio. La matrice $A = \begin{pmatrix} 1 & 2 & 3 & 0 & 1 \\ 0 & 0 & 9 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ è ridotta per righe ed anche ridotta a scala.

La matrice $B = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 9 \end{pmatrix}$ è ridotta per righe, non è ridotta a scala.

Definizioni.

(a) Dati i vettori riga $\mathbf{r}_1, \cdots, \mathbf{r}_k$ appartenenti a $\mathbf{R}^{1,n}$, una loro **combinazione lineare** è un vettore riga della forma

$$a_1\mathbf{r}_1 + \cdots + a_k\mathbf{r}_k, \quad \textit{con } a_1, \cdots, a_k \in \mathbf{R}$$

Analoga definizione se $\mathbf{r}_1, \cdots, \mathbf{r}_k$ sono vettori colonna appartenenti a $\mathbf{R}^{n,1}$.

(b) I vettori riga $\mathbf{r}_1, \cdots, \mathbf{r}_k$ si dicono **linearmente indipendenti** se l'equazione:

$$x_1\mathbf{r}_1 + \cdots + x_k\mathbf{r}_k = \mathbf{0}$$

(dove $\mathbf{0}$ è il vettore riga di $\mathbf{R}^{1,n}$ che ha coefficienti tutti nulli) ha solo la soluzione $x_1 = \cdots = x_k = 0$.

(c) I vettori riga $\mathbf{r}_1, \cdots, \mathbf{r}_k$ si dicono **linearmente dipendenti** se non sono linearmente indipendenti, ossia se uno di essi si può esprimere come combinazione lineare degli altri (basta osservare che, per esempio, scrivere $\mathbf{r}_k = a_1\mathbf{r}_1 + \cdots + a_{k-1}\mathbf{r}_{k-1}$ equivale a scrivere $a_1\mathbf{r}_1 + \cdots + a_{k-1}\mathbf{r}_{k-1} + (-1)\mathbf{r}_k = \mathbf{0}$.

Proposizione.

(a) Le righe non nulle di una matrice M ridotta a scala sono linearmente indipendenti; le colonne in generale no, ma sono linearmente indipendenti le colonne dotate di indicatore;

(b) una colonna di M è priva di indicatore se e solo se è combinazione lineare di quelle che la precedono.

Esempio.

$$M = \begin{pmatrix} 0 & 2 & 3 & 0 & 4 & 6 \\ 0 & 0 & 7 & 3 & 8 & 1 \\ 0 & 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Supponiamo che esista una relazione $a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + a_3\mathbf{r}_3 = \mathbf{0}$ tra le righe non nulle di M; sviluppando i calcoli si ottiene: $2a_1 = 0$, $3a_1 + 7a_2 = 0$, $3a_2 = 0$, $aa_1 + 8a_2 + 2a_3 = 0$, $6a_1 + a_2 + 4a_3 = 0$, da cui $a_1 = a_2 = a_3 = 0$. Le colonne di M invece non sono linearmente indipendenti: per esempio $\mathbf{c}_4 = -\frac{9}{14}\mathbf{c}_2 + \frac{3}{7}\mathbf{c}_3$. Il fatto è che riusciamo sempre a esprimere una colonna priva di indicatori come combinazione lineare delle precedenti, trovando i coefficienti con il procedimento usato prima. Allo stesso modo si vede che le colonne con indicatore sono linearmente indipendenti.

Definizione. Sia M una matrice ridotta a scala; si dice **rango** di M, $rg(M)$, il numero delle sue righe non nulle o equivalentemente il numero degli indicatori.

La matrice a scala M considerata nell'esempio precedente ha rango 3.

Sistemi lineari

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots a_{mn}x_n = b_m \end{cases}$$

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & |b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & |b_2 \\ \ldots & & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} & |b_m \end{pmatrix}$$

Un sistema lineare di m equazioni in n incognite a coefficienti in $\mathbf{R}$ si può scrivere come equazione matriciale $AX = B$, dove $A = (a_{ij}) \in \mathbf{R}^{m,n}$ è la matrice dei coefficienti, $X = (x_j) \in \mathbf{R}^{n,1}$ è la colonna delle incognite, $B = (b_i) \in \mathbf{R}^{m,1}$ è la colonna dei termini noti. Ogni riga della matrice completa $(A|B) \in \mathbf{R}^{m,n+1}$ corrisponde a un'equazione del sistema.

Una **soluzione** del sistema è una n-upla di numeri $(a_1, \cdots, a_n) \in \mathbf{R}$ che soddisfano contemporaneamente tutte le sue equazioni; un sistema può avere una sola o infinite soluzioni (sistema compatibile) oppure nessuna soluzione (sistema incompatibile).

In particolare un **sistema omogeneo** (cioè con tutti i temini noti nulli) è sempre compatibile, infatti ha sempre almeno la soluzione che si ottiene assegnando a tutte le incognite il valore 0 (soluzione banale).

Esempi.

$$(1) \begin{cases} x - 2y = 0 \\ 2x + y = 0 \end{cases} \qquad (2) \begin{cases} x - 2y = 0 \\ 2x - 4y = 0 \end{cases}$$

$$(3) \begin{cases} x - 2y = 2 \\ 2x + y = -1 \end{cases} \qquad (4) \begin{cases} x - 2y = 2 \\ 2x - 4y = -1 \end{cases}$$

I sistemi (1), (2) sono omogenei; si vede facilmente che (1) ha un'unica soluzione $(x, y) = (0, 0)$, mentre (2) ha infinite soluzioni del tipo $(x, y) = (2y, y)$ che dipendono dall'incognita libera y; il sistema (3) non è omogeneo e ha l'unica soluzione $(x, y) = (0, -1)$, mentre il sistema non omogeneo (4) è incompatibile, cioè non ha soluzione.

Un sistema lineare $AX = B$ può essere riscritto come: $x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = B$, dove $\mathbf{c}_1, \cdots, \mathbf{c}_n$ sono le colonne della matrice dei coefficienti A; perciò risolvere un sistema lineare equivale a cercare di esprimere la colonna dei termini noti B come combinazione lineare delle colonne di A.

Per riconoscere se un sistema lineare è compatibile e, in caso affermativo, per trovarne tutte le soluzioni, si può procedere in molti modi, per esempio con il ben noto metodo di sostituzione. Per i sistemi lineari in cui la matrice dei coefficienti è a scala, risulta particolarmente facile applicare il metodo di sostituzione dal basso verso l'alto.

Proposizione. Un sistema lineare $AX = B$, dove A è una matrice a scala di tipo $r \times n$ di rango r (cioè tutte le r righe non nulle) è compatibile qualunque sia B. Se $r < n$ si hanno ∞^{n-r} soluzioni, se $r = n$ la soluzione è unica.

Un altro metodo per risolvere i sistemi lineari è il procedimento di eliminazione di Gauss, che ha lo scopo di trasformare il sistema dato in un sistema più semplice da risolvere, ma con le stesse soluzioni.

Metodo di eliminazione di Gauss

Definizione. Data una matrice M, diciamo **operazione elementare sulle righe** di M una delle seguenti operazioni:

- scambiare due righe: $R_i \to R_j$
- moltiplicare una riga per un coefficiente non nullo $k \in \mathbf{R}$: $R_i \to kR_i$
- aggiungere ad una riga un multiplo di un'altra riga : $R_i \to R_i + kR_j$

Eseguendo operazioni elementari sulle righe di una matrice M, si ottiene una nuova matrice che ha lo stesso numero di righe e di colonne di M.

Definizione. Due sistemi lineari nelle incognite $x_1, x_2, \ldots, x_n$ si dicono **equivalenti** se hanno esattamente le stesse soluzioni.

Lemma. Se $M = (A|B)$ è la matrice di un sistema lineare, la matrice $M' = (A'|B')$ ottenuta eseguendo operazioni elementari sulle righe di M, corrisponde a un sistema lineare equivalente.

Dimostrazione. Permutando l'ordine delle equazioni di un sistema, le soluzioni ovviamente non cambiano. È sufficiente quindi mostrare che il sistema

$$(1) \quad \begin{cases} a_1x_1 + \ldots + a_nx_n = \alpha \\ b_1x_1 + \ldots + b_nx_n = \beta \end{cases}$$

è equivalente al sistema

$$(2) \quad \begin{cases} a_1x_1 + \ldots + a_nx_n = \alpha \\ a_1x_1 + \ldots + a_nx_n + k(b_1x_1 + \ldots + b_nx_n) = \alpha + k\beta \end{cases}$$

con $k \neq 0$. È chiaro che ogni soluzione di (1) è anche soluzione di (2) (basta sostituire); viceversa se $(c_1, \ldots, c_n)$ è soluzione di (2), vuol dire che $a_1c_1 + \ldots + a_nc_n = \alpha$ e $a_1c_1 + \ldots + a_nc_n + k(b_1c_1 + \ldots + b_nc_n) = \alpha + k\beta$ da cui semplificando $k(b_1c_1 + \ldots + b_nc_n) = k\beta$ e quindi $b_1c_1 + \ldots + b_nc_n = \beta$, da cui la tesi.

Dato un sistema lineare $AX = B$, il metodo di Gauss consiste nell' applicare una opportuna successione di operazioni elementari sulle righe della matrice $M = (A|B) \in \mathbf{R}^{m,n+1}$, in modo da ottenere alla fine del procedimento una matrice $M' = (A'|B')$ ridotta a scala. Per il Lemma, il sistema cosi' ottenuto è equivalente al sistema $AX = B$. Successivamente si risolve (se è possibile) il sistema di matrice $M' = (A'|B')$, per esempio con il metodo di sostituzione dal basso verso l'alto.

Attenzione: il procedimento deve essere applicato sull'intera riga, compreso il termine noto!

Proposizione. Un sistema lineare $AX = B$ di m equazioni in n incognite è compatibile se e solo se in una sua qualunque riduzione a scala $A'X = B'$, detto r il rango di A', gli ultimi $m - r$ coefficienti di B sono nulli.

Osservazione. Gli indicatori di M' dipendono dalle scelte arbitrarie che si effettuano nel procedimento di riduzione a scala, ma il loro numero $r = rg(M)$ (cioè il numero delle righe non nulle di M') non dipende della scelta delle operazioni, (si confronti con il successivo studio degli spazi vettoriali $\mathbf{R}^n$) .

Le considerazioni precedenti ci permettono di definire il rango di una matrice qualsiasi:

Definizione. Si dice **rango** di una matrice M il rango di una matrice M' ridotta a scala ottenuta da M con una opportuna successione di operazioni elementari sulle righe.

Osserviamo che se $(A'|B')$ è a scala, lo è in particolare la matrice A', si ha quindi per definizione:

$$rg(A) = rg(A'), \quad rg(A|B) = rg(A'|B')$$

Teorema di Rouchè-Capelli

Utilizzando la nozione di rango per una matrice qualsiasi, possiamo formulare un risultato fondamentale che ci permette di:

- capire se un sistema ha soluzioni o no (anche senza calcolarle)
- dire quante sono le soluzioni nel caso in cui il sistema sia compatibile .

Teorema. Il sistema lineare $AX = B$, con $A \in \mathbf{R}^{m,n}$, $X \in \mathbf{R}^{n,1}$, $B \in \mathbf{R}^{m,1}$

(a) è incompatibile se e solo se $r(A) < r(A|B)$.

(b) se è compatibile, poniamo $r = rg(A) = rg(A|B)$; allora ci sono ∞^{n-r} soluzioni, che si possono scrivere in funzione di $n - r$ incognite libere. Se $n = r$ la soluzione è unica.

Dimostrazione. Riducendo a scala la matrice completa del sistema dato, si ottiene la matrice $M' = (A'|B')$; visto che anche A' è ridotta a scala si vede che ci sono due possibilità:

(a) $r(A) < r(A|B)$. Poichè il rango è uguale al numero degli indicatori, $rg(A|B) = rg(A) + 1$. In questo caso $(A'|B')$ ha una riga del tipo $(0 \cdots 0|c)$ con $c \neq 0$, il sistema è quindi incompatibile.

(b) $r(A) = r(A|B)$. In questo caso basta considerare il sistema ridotto a scala $A'X = B'$ e risolverlo rispetto alle incognite corrispondenti alle colonne prive di indicatore (che sono $n - r$); infatti, supponendo che le r colonne linearmente indipendenti di A' (quelle con indicatore) siano le prime r, data una soluzione $(a_1, \cdots a_n)$, si ha

$$a_1\mathbf{c}_1 + \cdots a_n\mathbf{c}_n = B$$

$$a_1\mathbf{c}_1 + \cdots a_r\mathbf{c}_r = B - a_{r+1}\mathbf{c}_{r+1} + \cdots - a_n\mathbf{c}_n$$

dove le incognite che compaiono a secondo membro si dicono libere.

Corollario. Se $B = \mathbf{0}$, siamo nel caso (b), cioè un sistema omogeneo ha sempre soluzione; se inoltre $n = r$, esiste l'unica soluzione $X = \mathbf{0}$ (soluzione banale).

Osservazione. È chiaro che il numero m delle equazioni del sistema dato è in definitiva irrilevante; quello che conta è il numero r delle equazioni linearmente indipendenti. Ognuna di queste permette di esprimere una delle n incognite in funzione delle altre.

Definizione. Dato un sistema lineare non omogeneo $AX = B$, il sistema $AX = \mathbf{0}$ si dice **sistema omogeneo associato**.

Supponiamo di conoscere una soluzione Y del sistema $AX = B$ e una soluzione Z del sistema omogeneo associato $AX = \mathbf{0}$, allora $A(Y + Z) = AY + AZ = B + \mathbf{0} = B$ cioè $Y + Z$ è un'altra soluzione del sistema $AX = B$; viceversa, trovata in qualche modo una soluzione particolare Y di $AX = B$, per ottenerle tutte basta sommarla con le tutte soluzioni del sistema $AX = \mathbf{0}$.

Matrice inversa

Operando formalmente nello stesso modo, si possono applicare il metodo di Gauss e il teorema di Rouchè-Capelli a sistemi lineari del tipo $AX = B$, con $A \in \mathbf{R}^{m,n}$, $X \in \mathbf{R}^{n,p}$, $B \in \mathbf{R}^{m,p}$; in questo caso si può interpretare $AX = B$ come un sistema le cui incognite sono le righe di X. Consideriamo il caso particolare in cui il sistema è

$$AX = I_n, \qquad A, X \in \mathbf{R}^{n,n}$$

Per definizione la soluzione (se esiste) è la matrice A^{-1}, inversa di A. Inoltre, essendo $rg(I_n) = n$, si ha $rg(A|I_n) = n$, quindi per il teorema di Rouchè-Capelli esistono soluzioni solo se $rg(A) = n$; di conseguenza se il sistema è compatibile, necessariamente ha una sola soluzione.

Esempi.

1) Data

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix}$$

trovare (se esiste) A^{-1}.

Applicando il metodo di eliminazione di Gauss al sistema $(A|I)$ si ottiene:

$$\begin{pmatrix} 2 & 1 & 1 & | & 1 & 0 & 0 \\ 4 & 1 & 0 & | & 0 & 1 & 0 \\ -2 & 2 & 1 & | & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 2 & 1 & 1 & | & 1 & 0 & 0 \\ 0 & -1 & -2 & | & -2 & 1 & 0 \\ 0 & 0 & -4 & | & -5 & 3 & 1 \end{pmatrix}$$

Nel sistema ridotto a scala si vede che $rg(A') = 3$, quindi A è invertibile. Risolvendo poi per sostituzione e tenendo presente che le incognite sono le righe di X, si ottiene

$$A^{-1} = \begin{pmatrix} 1/8 & 1/8 & -1/8 \\ -1/2 & 1/2 & 1/2 \\ 5/4 & -3/4 & -1/4 \end{pmatrix}$$

2) Calcoliamo l'inversa di una matrice 2×2 con un metodo alternativo rispetto a quanto visto (cfr. 00-matrici, pag.4). Come sappiamo la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ è invertibile se e solo se $ad - bc \neq 0$; in tal caso

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Risolviamo il sistema $AX = I$, dove $X \in \mathbf{R}^{2,2}$, $I \in \mathbf{R}^{2,2}$ è la matrice identità e le incognite sono le righe X_1, X_2 della matrice X. Applichiamo il metodo di eliminazione di Gauss al sistema $(A|I)$; supponendo $a \neq 0, c \neq 0$, possiamo eseguire sulla seconda riga l'operazione elementare $R_2 \to aR_2 - cR_1$:

$$\begin{pmatrix} a & b & | & 1 & 0 \\ c & d & | & 0 & 1 \end{pmatrix} \to \begin{pmatrix} a & b & | & 1 & 0 \\ 0 & ad - bc & | & -c & a \end{pmatrix}$$

Se $ad - bc \neq 0$, ossia se $det(A) \neq 0$, si ha $rg(A) = rg(A|I) = 2$, quindi il sistema è compatibile e per sostituzione si ottiene:

$$\begin{cases} aX_1 + b\dfrac{(-c,a)}{ad-bc} = (1,0) \\ (ad-bc)X_2 = (-c,a) \end{cases}$$

quindi

$$X_1 = -\frac{b}{a}\frac{(-c,a)}{(ad-bc)} + \frac{1}{a}(1,0) = \frac{1}{ad-bc}(d,-b)$$

$$X_2 = \frac{1}{ad-bc}(-c,a)$$

da cui

$$X = A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Se invece $ad - bc = 0$, ossia se $det(A) = 0$, si ha $rg(A) < rg(A|I)$, quindi il sistema è incompatibile e la matrice A non è invertibile.
Analogamente, se $a = c = 0$, si ottiene $ad - bc = 0$ e la matrice A non è invertibile.

Corollario. (Regola di Cramer) Il sistema lineare $AX = B$, con $A \in \mathbf{R}^{n,n}$ invertibile ha una sola soluzione data da $X = A^{-1}B$

Esempio. Risolvere il sistema lineare $AX = B$, con $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertibile e $B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

Moltiplicando a sinistra ambo i membri di $AX = B$ per A^{-1}, si ottiene $A^{-1}AX = A^{-1}B$ ossia $X = A^{-1}B$; dal precedente esempio 2)

$$X = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} b_1d - b_2b \\ -b_1c + b_2a \end{pmatrix}$$

Dunque il sistema ha una sola soluzione data da

$$\frac{1}{det(A)}(det\begin{pmatrix} b_1 & b \\ b_2 & d \end{pmatrix}, det\begin{pmatrix} a & b_1 \\ c & b_2 \end{pmatrix})$$

Cenno sul determinante di una matrice $\mathbf{n \times n}$

Abbiamo visto come si calcola il determinante di una matrice $n \times n$ quando $n = 2$ oppure $n = 3$. In generale vale la seguente:

Definizione. Sia A una matrice quadrata $n \times n$, si dice **determinante** di A il numero cosi' ottenuto:

$$det(A) = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

dove σ varia nell'insieme S_n di tutte le $n!$ permutazioni di n elementi.

Calcolare a mano $det(A)$ utilizzando questa formula può risultare estremamente pesante. Un procedimento alternativo consiste nel riscrivere la formula, mettendo in evidenza gli elementi di una riga (o di una colonna) fissata di A; ciò permette di ottenere lo **sviluppo di Laplace** di $det(A)$ relativo alla riga o colonna fissata, come segue:

Definizione. Data la matrice quadrata $A = (a_{ij})$, si dice **complemento algebrico** del coefficiente a_{ij} il numero definito dal prodotto di $(-1)^{i+j}$ per il determinante della matrice che si ottiene da A cancellando la riga i-esima e la colonna j-esima.

Teorema di Laplace. Data una matrice A quadrata $n \times n$, $det(A)$ è uguale alla somma dei prodotti degli elementi di una qualsiasi riga (o colonna) fissata per i rispettivi complementi algebrici:

$$det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$$

$$det(A) = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}$$

È facile verificare l'enunciato nei casi $n = 2, 3$. In generale il teorema permette il calcolo di un determinante a partire da determinanti con una riga (o colonna) in meno. La riga o colonna rispetto alla quale iniziare lo sviluppo può essere scelta a piacere; ovviamente conviene scegliere una riga (o colonna) che abbia tanti coefficienti nulli. Un caso particolare è quello delle matrici triangolari, il cui determinante è sempre il prodotto degli elementi della diagonale principale.

Sono conseguenza del teorema le seguenti proprietà (di facile verifica nel caso $n = 2, 3$):

Proprieta'. 1) Se nella matrice A si scambiano tra loro due righe, il determinante cambia segno.

2) Se nella matrice A si sostituisce a un vettore riga R_i il vettore riga $R_i + kR_j$, dove R_j è un'altra riga e k è un qualsiasi coefficiente numerico, il determinante non cambia.

3) Se si applica il metodo di eliminazione di Gauss alla matrice A, si ottiene una matrice S ridotta a scala e dai risultati precedenti si ha che $det(A) = \pm det(S)$; in particolare se $rg(A) = n$, $det(S)$ è il prodotto degli indicatori perciò $det(A) \neq 0$, mentre se $rg(A) < n$, $det(A) = \pm det(S) = 0$ e viceversa.