Lezione, Sistema di equazioni (parte 1)

Sistemi lineari e mappe lineari

Ricordiamo quanto segue dalla lezione 1: Siano $f_1,\dots,f_k:\mathbb{R}^n\to\mathbb{R}$ funzioni. Allora,

$$V(f_1,\dots,f_k)=\{\,x\in\mathbb{R}^n \mid f_1(x)=f_2(x)=\cdots=f_k(x)=0\,\}$$

Il sistema di equazioni $V(f_1,\dots,f_k)$ si dice lineare se (e sole se) ogni funzione ha la forma

$$f_i(x_1,\dots,x_n)=\left(\sum_{\ell=1}^{n} a_{i\ell}x_\ell\right)-b_\ell$$

dove i coefficienti $a_{i\ell}$ e b_ℓ sono costanti. Un sistema lineare si dice omogeneo se (e sole se) ogni $b_\ell=0$. Altrimenti, si dice che il sistema lineare è disomogeneo.

Se $V(f_1,\dots,f_k)$ è un sistema lineare allora

$$V(h_1,\dots,h_k),\qquad h_i(x_1,\dots,x_n)=f_i(x_1,....,x_n)-f_i(0,....,0)$$

è un sistema lineare omogeneo, che è chiamato il sistema associato di equazioni omogenee. Ogni sistema lineare omogeneo ha la soluzione

$$(0,\dots,0) \tag{E1}$$

Avendo introdotto il concetto di mappa lineare, possiamo ora riformulare la nozione di sistema di equazioni lineari come segue: La funzione,

$$L(x_1,\dots,x_n)=(h_1(x_1,\dots,x_n),\dots,h_k(x_1,\dots,x_n)):\mathbb{R}^n\to\mathbb{R}^k$$

è una mappa lineare da $\mathbb{R}^n$ a $\mathbb{R}^k$. Questo segue dal fatto che ogni funzione componente

$$h_i(x_1,\dots,x_n)=a_{i1}x_1+\cdots+a_{in}x_n:\mathbb{R}^n\to\mathbb{R}$$

è una mappa lineare. Poiché $f_i(x)=h_i(x)+b_i$ abbiamo che

$$V(f_1,\dots,f_n)=\{\,x\in\mathbb{R}^n \mid L(x)=b\,\},\qquad b=(b_1,\dots,b_k) \tag{E2}$$

<u>In generale</u>: Siano U e W spazi vettoriali e $f:U\to W$ sia una funzione. Allora

$$V(f)=\{u\in U \mid f(u)=0\} \tag{E3}$$

Allora V si dice che è un sistema di equazioni lineari se

$$f(u)=L(u)-b$$

dove L è una mappa lineare e b è una costante, quindi

$$V(f)=\{u\in U \mid L(u)=b\} \tag{E3'}$$

Esempio: Siano $U = \mathbb{R}[x], \quad W = \mathbb{R}^3, \quad L : U \to \mathbb{R}^3, \quad L(p) = (p(1), p(2), p(3))$

Ricordiamo che L è una mappa lineare perché:

(a) $p,\ q \in \mathbb{R}[x] \implies L(p+q) = ((p+q)(1), (p+q)(2), (p+q)(3))$
$$= (p(1)+q(1), p(2)+q(2), p(3)+q(3))$$
$$= (p(1), p(2), p(3)) + (q(1), q(2), q(2)) = L(p) + L(q)$$

(b) $p \in \mathbb{R}[x],\ c \in \mathbb{R} \implies L(cp) = ((cp)(1), (cp)(2), (cp)(3))$
$$= (cp(1), cp(2), cp(3)) = c(p(1), p(2), p(3)) = cL(p)$$

Il sistema lineare
$$V(L) = \{\, p \in \mathbb{R}[x] \mid L(x) = 0 \,\}$$
è solo l'insieme dei polinomi che sono nulli in $x = 1,\ x = 2,\ x = 3$. In altre parole, $p \in V(L)$ se e solo se $(x-1),\ (x-2),\ (x-3)$ sono fattori di p.
Quindi,
$$V(L) = \{\, (x-1)(x-2)(x-3)q(x) \mid q(x) \in \mathbb{R}[x] \,\}$$

La struttura delle soluzioni di un sistema lineare

In analogia con i sistemi lineari di equazioni in $\mathbb{R}^n$, un sistema lineare
$$L(u) = b \tag{E4}$$
si dice omogeneo se $b = 0$. Altrimenti, il sistema lineare è disomogeneo

Il sistema lineare
$$L(u) = 0 \tag{E5}$$
è chiamato il sistema omogeneo associato di (E4).

Proposizione: Se u' è una soluzione di (E4) allora ogni soluzione di (E4) ha la forma $u + u'$ dove u è una soluzione di (E5).
Dimostrare: Supponiamo che u' sia una soluzione di (E4) e che u sia una soluzione di (E5). Allora,
$$L(u') = b, \quad L(u) = 0 \implies L(u'+u) = L(u') + L(u) = b + 0 = b$$
Quindi, $u' + u$ è una soluzione di (E4). Viceversa, supponiamo che $u',\ u''$ sono soluzione di (E4). Allora,
$$L(u'' - u') = L(u'') - L(u') = b - b = 0$$
Quindi, $u = u'' - u'$ è una soluzione di (E5) tale che $u'' = u' + u$.

Esempio: Siano $U = \mathbb{R}[x], \quad W = \mathbb{R}^3, \quad L : U \to \mathbb{R}^3, \quad L(p) = (p(1), p(2), p(3))$

Trova tutte le soluzioni del sistema lineare $L(p) = (1,3,6)$.

Nell'esempio precedente, abbiamo risolto il sistema associato $L(p) = 0$.
Abbiamo trovato che le soluzioni erano:

$$\{(x-1)(x-2)(x-3)q(x) \mid q(x) \in \mathbb{R}[x]\}$$

Per trovare una soluzione di $L(p) = (1,3,6)$, possiamo usare l'interpolazione di Lagrange.

Nel nostro caso particolare, questo è stato fatto alle pagine 4 e 5 della lezione 5.
Il risultato fu:

$$p(x) = \frac{1}{2}x(x+1) \implies p(1) = 1, \quad p(2) = 2, \quad p(3) = 6$$

Perciò,

$$L(f) = (1,3,6) \iff f(x) = \frac{1}{2}x(x+1) + (x-1)(x-2)(x-3)q(x), \quad q(x) \in \mathbb{R}[x]$$

Matrici:

Una matrice A n per m (di solito scritto $n \times m$) è una tabella rettangolare di numeri che ha n righe e m colonne:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Esempio:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix}$$ (A è una matrice 3x4) (E6)

Una matrice quadrata è una matrice che ha lo stesso numero di righe e colonne.

Una matrice Vandermonde è una matrice quadrata della forma:

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^{n-1} \end{bmatrix}$$ (V è una matrice quadrata) (E7)

dove $x_1, \dots, x_m$ sono numeri.

Spesso scriviamo $A = (a_{ij})$ dove a_{ij} è la voce che appare nella i'th riga e j'th colonna di A.

In particolare, questo ci permette di specificare una matrice dando una formula per a_{ij}.

Esempio: La matrice (E6) è data dalla formula

$$A = (a_{ij}), \qquad a_{ij} = i^{j-1}, \quad 1 \le i \le 3, \quad 1 \le j \le 4$$

La matrice (E7) è data dalla formula

$$V = (v_{ij}), \quad v_{ij} = x_i^{j-1}, \quad 1 \le i, j \le m$$

Esempio: Determinare la matrice 3x4 data dalla formula $A = (a_{ij}), \quad a_{ij} = ij$

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{pmatrix}$$

Vettori di riga e colonna

$\mathbb{R}^n$ è il prodotto cartesiano $\mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in \mathbb{R}\}$

Un vettore riga di dimensione n è una matrice $n \times 1$:

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

Esempio: (1 2 3)

Un vettore colonna di dimensione m è una matrice $1 \times m$:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

Esempio: $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

Naturalmente possiamo pensare a un elemento di $\mathbb{R}^n$ come a un vettore riga di dimensione n o a un vettore colonna di dimensione n usando le mappe:

$$(x_1, \ldots, x_n) \mapsto \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}, \qquad (x_1, \ldots, x_n) \mapsto \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Tuttavia, come matrici, i vettori riga e colonna sono cose diverse, tranne nel caso 1x1.

Una matrice 1x1 non è la stessa cosa di uno scalare. Questo diventerà chiaro quando definiremo la moltiplicazione di matrici in generale.

Moltiplicazione sinistra di una matrice nxm e di un vettore a colonne di dimensione m

Sia A una matrice $n \times m$ e v un vettore colonna di dimensione m. Allora, Av è il vettore colonna n dimensionale ottenuto come segue:

$$Av = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \qquad c_i = \sum_{k=1}^{m} a_{ik} v_k$$

Un modo di pensare a questa formula è che

$$Av = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}, \qquad c_i = (\vec{a}_i, \vec{v})$$

Dove $\vec{a}_i$ è l'iesima riga di A, pensata come un elemento di $\mathbb{R}^m$ e $\vec{v}$ è v pensato come un elemento di $\mathbb{R}^m$.

Esempio:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1-1+1-1 \\ 1-2+4-8 \\ 1-3+9-27 \end{pmatrix} = \begin{pmatrix} 0 \\ -5 \\ -20 \end{pmatrix}$$

La forma matriciale di un sistema lineare

Sia

$$L\underbrace{(x_1, \ldots, x_m)}_{\text{anche un vettore di colonne in } \mathbb{R}^m} = \underbrace{(a_{11}x_1 + \cdots + a_{1m}x_m, a_{21}x_1 + \cdots + a_{2m}x_m, \cdots, a_{n1}x_1 + \cdots + a_{nm}x_m)}_{\text{anche un vettore di colonne in } \mathbb{R}^n}$$

Allora,

$$L(x_1, \ldots, x_m) = \underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}}_{A \text{ : matrice nxm}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{x \text{ : vettore collonne}}$$

Quindi, pensando agli elementi di $\mathbb{R}^n$ come vettori colonne, il sistema lineare $L(x) = b$ può essere scritto come

$$Ax = b$$

Questa è chiamata la forma matriciale del sistema lineare.

Esempio: Scrivi il seguente sistema di equazioni in forma di matrice.

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 + 2x_2 + 4x_3 + 8x_4 &= 2 \\ x_1 + 3x_2 + 9x_3 + 27x_4 &= 3 \end{aligned} \longrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \text{(E8)}$$

Sistemi lineari, forma a matrice aumentata:

Il sistema lineare $Ax = b$ è spesso scritto nella forma abbreviata $(A \mid b)$ dove la barra verticale è usata per separare la matrice A dal vettore di costanti b.

Esempio: L'equazione (E8) diventa

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 2 \\ 1 & 3 & 9 & 27 & 3 \end{array}\right)$$

$$\underbrace{\hphantom{1 \quad 1 \quad 1 \quad 1}}_{A} \qquad \underbrace{\hphantom{1}}_{b}$$

Matrici Saclini:

Sia A una matrice qualsiasi. Per qualsiasi riga R di A, chiamiamo pivot il primo elemento non nullo della riga R. Una matrice a scalini è una matrice che ha la seguente proprietà: il pivot di ogni riga è sempre strettamente a destra del pivot della riga precedente.

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad\qquad \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

i pivot — i pivot

Non è una matrice a scalini — Una matrice a scalini

In particolare, se A è una matrice a scalini, allora ogni riga zero di A si verifica nella parte inferiore di A.

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Una matrice a scalini.
Ha due pivot e le righe zero nella parte inferiore della matrice.

Questo algoritmo è facile da usare e da capire, ma un po' disordinato da scrivere.

Risolvere Ax=b per A una matrice a scalini.

Se A è una matrice a scalini, possiamo risolvere $Ax = b$ lavorando all'indietro dal fondo della matrice aumentata $(A \mid b)$.

Passo 1: Se $(A \mid b)$ contiene una riga della forma $(0 \cdots 0 \mid b \neq 0)$, fermatevi.
Non c'è soluzione.

Infatti, tale riga corrisponde a un'equazione della forma

$$0x_1 + 0x_2 + \cdots + 0x_n = b, \qquad b \neq 0$$

Notazione: Sia $A = (a_{ij})$.
Siano $A_1 = (a_{11} \ \cdots \ a_{1n}), \ldots, A_k = (a_{k1} \ \cdots \ a_{kn})$ le righe non nulle di A.
Sia $\alpha_j = a_{jp_j}$ il pivot di A_j

Passo 2: Usando l'ultima riga A_k, si ha

$$x_{p_k} = (\alpha_k)^{-1} \left(b_k - \sum_{\ell > p_k} a_{k\ell} x_\ell \right)$$

Passo 3: Allo stesso modo, usando la riga A_{k-1}, si ha

$$x_{p_{k-1}} = (\alpha_{k-1})^{-1} \left(b_{k-1} - \sum_{\ell > p_{k-1}} a_{(k-1)\ell} x_\ell \right)$$

Elimina x_{p_k} usando l'equazione trovata nel passo precedente.

Passo 4: Applicare lo stesso metodo in sequenza alle righe $A_{k-2}, A_{k-3}, \ldots, A_1$

Il risultato dell'algoritmo:

Chiamiamo $\{x_{p_1}, \ldots, x_{p_k}\}$ le variabili dipendenti.
Chiamiamo le variabili rimanenti variabili indipendenti.

L'algoritmo descritto sopra produce una formula per le variabili dipendenti in termini di variabili indipendenti.

Esempio: $A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \qquad (A \mid b) = \left(\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 1 \\ 0 & 0 & 3 & 4 & 2 \\ 0 & 0 & 0 & 2 & 4 \end{array}\right)$

Passo 1: $(A \mid b)$ non ha righe della forma $(0 \ \cdots \ 0 \mid b \neq 0)$.
Passo 2: L'ultima riga diversa da zero di A dà $2x_4 = 4 \implies x_4 = 2$.
Passo 3: $3x_3 + 4x_4 = 2, \ x_4 = 2 \implies x_3 = (1/3)(2 - (4)(2)) = -2$.
Passo 4: $x_1 + +x_2 + 2x_3 + x_4 = 1, \ x_3 = -2, \ x_4 = 2 \implies x_1 = 1 - x_2 - 2(-2) - (2) = 3 - x_2$

Allora: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 - x_2 \\ x_2 \\ -2 \\ 2 \end{pmatrix}$ dove x_2 è la variabile indipendente e x_1, x_3, x_4 sono le variabili dipendenti

Pagina 7.