

Sistemi lineari

1.15 Rango di matrici

Consideriamo una matrice $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. Denotiamo con $A_1, \dots, A_m \in \mathbb{R}^n$ le righe di A , e con $A^1, \dots, A^n \in \mathbb{R}^m$ le colonne di A . Poniamo la seguente definizione.

Definizione 1.15.1. *Il rango per righe di A è il massimo numero di righe di A linearmente indipendenti, vale a dire $\dim(L(A_1, \dots, A_m))$. Il rango per colonne di A è il massimo numero di colonne di A linearmente indipendenti, cioè $\dim(L(A^1, \dots, A^n))$.*

Vediamo ora che le due nozioni appena definite coincidono.

Proposizione 1.15.2. *Il rango per righe di una matrice $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ è uguale al rango per colonne di A .*

Dimostrazione. Denotiamo con r il massimo numero di righe di A linearmente indipendenti e siano A_{i_1}, \dots, A_{i_r} tali righe. Allora $\{A_{i_1}, \dots, A_{i_r}\}$ è una base del sottospazio vettoriale $L(A_1, \dots, A_m)$, e quindi per ogni $i \in \{1, \dots, m\}$ esistono $b_{i1}, \dots, b_{ir} \in \mathbb{R}$ tali che

$$A_i = b_{i1}A_{i_1} + \dots + b_{ir}A_{i_r}. \quad (1.49)$$

Se denotiamo con $B = (b_{ij}) \in \mathcal{M}_{m \times r}(\mathbb{R})$ e con $C \in \mathcal{M}_{r \times n}(\mathbb{R})$ la matrice che ha come righe proprio A_{i_1}, \dots, A_{i_r} , allora, espandendo la (1.49), è facile convincersi che

$$A = BC.$$

Allora

$$A^t = (BC)^t = C^t B^t.$$

Espandendo la precedente uguaglianza, analogamente alla (1.49), si ha che le colonne di A (cioè le righe di A^t) possono essere espresse come combinazione

lineare delle righe di B^t , che sono r . Segue che $L(A^1, \dots, A^n)$ ha r generatori, e quindi, per il Lemma di Steinitz, il numero s di colonne linearmente indipendenti di A è minore o uguale a r . Abbiamo dunque dimostrato che

$$s \leq r.$$

D'altra parte, possiamo applicare la precedente uguaglianza a A^t . Poichè le righe di A^t corrispondono alle colonne di A , si ha dunque

$$r \leq s$$

e questo prova l'enunciato del teorema. \square

La precedente proposizione giustifica la seguente definizione.

Definizione 1.15.3. *Sia $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. Il rango di A è il massimo numero $\text{rg}(A)$ di righe o, equivalentemente, di colonne di A linearmente indipendenti.*

Dalla definizione segue immediatamente che $0 \leq \text{rg}(A) \leq \min\{m, n\}$, e $\text{rg}(A) = 0$ se e solo se non ci sono righe (e colonne) linearmente indipendenti, cioè A è la matrice nulla.

Esempio 1.15.4. *Consideriamo la matrice*

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

Il rango di A è, in questo caso, almeno 1 (poichè vi è almeno una riga o una colonna non nulla) e può essere al massimo 3 (non vi sono più di 3 righe). Calcoliamo dapprima il rango per righe. Poichè la prima e la seconda riga non sono proporzionali, esse sono linearmente indipendenti. È dato che $A_3 = -A_1 + A_2$, concludiamo che il massimo numero di righe lineamente indipendenti è 2. Pertanto $\text{rg}(A) = 2$. Allo stesso risultato si poteva giungere considerando le colonne. Infatti la prima e terza colonna sono linearmente indipendenti, mentre $A^2 = 2 \cdot A^1 + 0 \cdot A^3$ e, chiaramente, $A^4 = 0 \cdot A^1 + 0 \cdot A^3$.

Vediamo ora alcuni criteri alternativi per calcolare il rango di una matrice. Partiamo dal caso delle matrici quadrate. Il seguente teorema mostra che l'invertibilità di una matrice equivale alla proprietà che tale matrice abbia rango massimo.

Teorema 1.15.5. *Sia $A \in \mathcal{M}_n(\mathbb{R})$. Allora $\text{rg}(A) = n$ se e solo se $\det(A) \neq 0$.*

Dimostrazione. Supponiamo che A abbia rango n , e quindi le sue n righe $A_1, \dots, A_n \in \mathbb{R}^n$ sono linearmente indipendenti. Dato che $\dim(\mathbb{R}^n) = n$, $\{A_1, \dots, A_n\}$ è una base di \mathbb{R}^n . In particolare ciò significa che ciascun elemento della base canonica $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di \mathbb{R}^n può essere espresso come combinazione lineare di A_1, \dots, A_n . Dunque per ogni $i \in \{1, \dots, n\}$ esistono $b_{i1}, \dots, b_{in} \in \mathbb{R}$ tali che

$$\mathbf{e}_i = b_{i1}A_1 + \dots + b_{in}A_n$$

cioè

$$\begin{aligned} (0, \dots, 1, \dots, 0) = \mathbf{e}_i &= b_{i1}(a_{11}, \dots, a_{1n}) + \dots + b_{in}(a_{n1}, \dots, a_{nn}) \\ &= (b_{i1}a_{11} + \dots + b_{in}a_{n1}, \dots, b_{i1}a_{1n} + \dots + b_{in}a_{nn}). \end{aligned}$$

In particolare, le j -esime componenti dei vettori nella precedente uguaglianza sono uguali, e quindi si ottiene

$$b_{i1}a_{1j} + \dots + b_{in}a_{nj} = \delta_{ij}. \quad (1.50)$$

Si noti che al variare di $i \in \{1, \dots, n\}$ e di $j \in \{1, \dots, n\}$, gli scalari b_{ij} possono essere pensati come entrate di una matrice $B = (b_{ij}) \in M_n(\mathbb{R})$. La (1.50) equivale dunque a

$$[B \cdot A]_{ij} = [I_n]_{ij}$$

per ogni $i, j \in \{1, \dots, n\}$, cioè

$$BA = I_n,$$

da cui, applicando la formula di Binet, si ottiene $\det(A) \neq 0$. (Si noti che la matrice B che abbiamo costruito è proprio l'inversa di A).

Viceversa, supponiamo che $\det(A) \neq 0$. Dimostriamo che le colonne di A sono linearmente indipendenti. Consideriamo una loro qualsiasi combinazione lineare posta uguale al vettore nullo di \mathbb{R}^n

$$x_1A^1 + \dots + x_nA^n = \mathbf{0} \quad (1.51)$$

con $x_1, \dots, x_n \in \mathbb{R}$. Vogliamo dimostrare che $x_1 = \dots = x_n = 0$. Espandendo la (1.51) si ha

$$\begin{aligned} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} &= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 & \cdots & a_{1n}x_n \\ \vdots & \ddots & \vdots \\ a_{n1}x_1 & \cdots & a_{nn}x_n \end{pmatrix} \end{aligned}$$

cioè

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases} \quad (1.52)$$

Si noti che (1.52) è un sistema lineare omogeneo di n equazioni in n incognite e, soprattutto, è un sistema di Cramer, in quanto la matrice dei coefficienti del sistema è proprio la matrice A che per ipotesi ha determinante diverso da 0. Di conseguenza l'unica soluzione di tale sistema è la soluzione nulla, e quindi $x_1 = \cdots = x_n = 0$, che è ciò che volevamo dimostrare. \square

Vediamo ora un modo alternativo per calcolare il rango di una matrice A con m righe ed n colonne. Fissiamo qualche notazione. Ricordiamo che una sottomatrice di A è una matrice $B \in \mathcal{M}_{p \times q}(\mathbb{R})$, con $1 \leq p \leq m$ e $1 \leq q \leq n$ le cui entrate sono gli elementi che hanno in comune p righe e q colonne fissate di A . Più precisamente, fissiamo le righe di A di posto

$$1 \leq i_1 < i_2 < \cdots < i_p \leq m$$

e le colonne di A di posto

$$1 \leq j_1 < j_2 < \cdots < j_q \leq n.$$

Allora denoteremo con

$$B = A(i_1, \dots, i_p | j_1, \dots, j_q)$$

la matrice con p righe e q colonne le cui entrate sono

$$[B]_{rs} := [A]_{i_r j_s}$$

per ogni $r \in \{1, \dots, p\}$ e $s \in \{1, \dots, q\}$. Per esempio, se

$$A = \begin{pmatrix} 1 & 3 & -1 & 2 & 5 & \frac{1}{3} \\ -1 & 0 & -2 & 0 & 2 & -2 \\ 0 & \sqrt{2} & -2 & 2 & \pi & 0 \\ 2 & -\frac{3}{4} & 0 & -2 & 2 & 1 \end{pmatrix}$$

allora

$$A(2, 3, 4 | 3, 5) = \begin{pmatrix} -2 & 2 \\ -2 & \pi \\ 0 & 2 \end{pmatrix}.$$

Con queste notazioni dimostreremo ora che il rango di A può essere facilmente calcolato a partire dal determinante delle sue sottomatrici quadrate.

Partiamo da questo risultato preliminare:

Proposizione 1.15.6. *Il rango di una qualsiasi sottomatrice di A è sempre minore o uguale al rango di A .*

Dimostrazione. Sia $A \in \mathcal{M}_{m \times n}(\mathbb{R})$ e consideriamo la sua sottomatrice $A(i_1, \dots, i_p | j_1, \dots, j_q) \in \mathcal{M}_{p \times q}(\mathbb{R})$. Dobbiamo dimostrare che

$$\text{rg}(A(i_1, \dots, i_p | j_1, \dots, j_q)) \leq \text{rg}(A).$$

Andiamo per gradi. Consideriamo dapprima la sottomatrice

$$B = A(i_1, \dots, i_p | 1, \dots, n)$$

le cui entrate sono gli elementi delle righe A_{i_1}, \dots, A_{i_p} della matrice A . Per definizione di rango si ha

$$\begin{aligned} \text{rg}(A(i_1, \dots, i_p | 1, \dots, n)) &= \dim(L(A_{i_1}, \dots, A_{i_p})) \\ &\leq \dim(L(A_1, \dots, A_m)) = \text{rg}(A). \end{aligned} \quad (1.53)$$

Allora, usando la (1.53), si ha

$$\begin{aligned} \text{rg}(A(i_1, \dots, i_p | j_1, \dots, j_q)) &= \text{rg}(B(1, \dots, p | j_1, \dots, j_q)) \\ &= \dim(L(B^{j_1}, \dots, B^{j_q})) \\ &\leq \dim(L(B^1, \dots, B^n)) = \text{rg}(B) \\ &= \text{rg}(A(i_1, \dots, i_p | 1, \dots, n)) \\ &\leq \text{rg}(A). \end{aligned}$$

□

Siamo finalmente in grado di fornire una caratterizzazione del rango, di fatto una definizione equivalente, molto utile nella pratica.

Teorema 1.15.7. *Il rango di una matrice è il massimo ordine di un suo minore avente determinante non nullo.*

Dimostrazione. Dobbiamo dimostrare che, data una matrice $A \in \mathcal{M}_{m \times n}(\mathbb{R})$,

$$\text{rg}(A) = r \iff \begin{cases} \text{esiste un minore } M_0 \text{ di } A \text{ di ordine } r \text{ tale che } \det(M_0) \neq 0 \\ \text{per ogni minore } M \text{ di } A \text{ di ordine } r + 1 \text{ si ha } \det(M) = 0 \end{cases}$$

Supponiamo che $\text{rg}(A) = r$ e siano A_{i_1}, \dots, A_{i_r} righe di A linearmente indipendenti. Allora

$$\text{rg}(A(i_1, \dots, i_r | 1, \dots, n)) = r$$

e quindi esistono r colonne B^{j_1}, \dots, B^{j_r} della sottomatrice

$$B := A(i_1, \dots, i_r | 1, \dots, n)$$

linearmente indipendenti. Allora $M_0 := A(i_1, \dots, i_r | j_1, \dots, j_r)$ è un minore di A di ordine r tale che

$$\operatorname{rg}(M_0) = \operatorname{rg}(A(i_1, \dots, i_r | j_1, \dots, j_r)) = \operatorname{rg}(B(1, \dots, r | j_1, \dots, j_r)) = r$$

e quindi, per il Teorema 1.15.5, $\det(M_0) \neq 0$. Sia ora M un qualsiasi altro minore di A di ordine $p + 1$. Se per assurdo $\det(M) \neq 0$ si avrebbe allora, ancora per il Teorema 1.15.5, che $\operatorname{rg}(M) = p + 1 > p = \operatorname{rg}(A)$. Questo però è assurdo, dato che, essendo M sottomatrice di A , per la Proposizione 1.15.6, si deve avere $\operatorname{rg}(M) \leq \operatorname{rg}(A)$.

Viceversa supponiamo che esista un minore M_0 di ordine r tale che $\det(M_0) \neq 0$ e ogni altro minore di ordine maggiore abbia determinante nullo. Poniamo $p := \operatorname{rg}(A)$. Dato che M_0 è sottomatrice di A , dalla Proposizione 1.15.6 segue subito che

$$r \leq p.$$

Vogliamo ora dimostrare che si ha anche $r \geq p$. Dato che $p := \operatorname{rg}(A)$ esistono p righe di A , diciamo A_{i_1}, \dots, A_{i_p} , linearmente indipendenti. Allora la sottomatrice

$$B := A(i_1, \dots, i_p | 1, \dots, n)$$

ha rango p , e quindi esistono p colonne di B , diciamo B^{j_1}, \dots, B^{j_p} linearmente indipendenti. In particolare si ha che

$$\operatorname{rg}(A(i_1, \dots, i_p | j_1, \dots, j_p)) = \operatorname{rg}(B(1, \dots, p | j_1, \dots, j_p)) = p.$$

Allora $M_0 := A(i_1, \dots, i_p | j_1, \dots, j_p)$ è un minore di A di ordine p tale che $\det(M_0) \neq 0$. Poichè r è il *massimo* ordine di un minore di A avente determinante diverso da 0, si deve quindi avere

$$p \leq r.$$

Di conseguenza $\operatorname{rg}(A) = p$. □

Il precedente teorema fornisce un metodo molto utile, e più rapido, per determinare il rango di una matrice. Per esempio, supponiamo di dover stabilire il rango della matrice

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 3 & -1 & -3 & -1 \\ 6 & 2 & 0 & 2 \end{pmatrix}. \quad (1.54)$$

Poichè il minore di ordine 2 di A

$$M_0 = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}$$

ha determinante diverso da 0, possiamo già dire che il rango di A è almeno 2. Il rango è 3 se riusciamo a trovare un minore di ordine 3 di A avente determinante diverso da 0. Ebbene, il lettore potrà verificare che tutti i minori di ordine 3 di A , che sono i seguenti

$$\begin{pmatrix} 1 & 1 & 1 \\ 3 & -1 & -3 \\ 6 & 2 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 \\ 3 & -1 & -1 \\ 6 & 2 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 \\ 3 & -3 & -1 \\ 6 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 \\ -1 & -3 & -1 \\ 2 & 0 & 2 \end{pmatrix}$$

hanno determinante uguale a zero. Concludiamo che $\text{rg}(A) = 2$.

Dal precedente esempio emerge chiaramente come può essere laborioso doversi calcolare il determinante di tutti i minori di una matrice, specialmente quelli di ordine elevato, almeno fintanto che si ha la fortuna di trovarne uno avente determinante non nullo. Una semplificazione da questo punto di vista è data dal Teorema di Kronecker, che andiamo ad enunciare.

Premettiamo che, dato un minore M di A , un altro minore M' di A si dice *minore orlato* di M se M è anche minore di M' . Per esempio, con riferimento alla matrice (1.54) incontrata poco fa, un minore orlato di M_0 è

$$\begin{pmatrix} 1 & 1 & 1 \\ 3 & -1 & -1 \\ 6 & 2 & 2 \end{pmatrix}.$$

Invece il minore di A

$$\begin{pmatrix} 1 & 1 & 1 \\ 3 & -3 & -1 \\ 6 & 0 & 2 \end{pmatrix}$$

non è un orlato di M_0 .

Enunciamo ora il Teorema di Kronecker, di cui si omette la dimostrazione.

Teorema 1.15.8 (Kronecker). *Sia $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. Se esiste un minore M_0 di ordine r tale che $\det(M_0) \neq 0$ e tutti i minori orlati di M_0 di ordine $r + 1$ hanno determinante nullo, allora $\text{rg}(A) = r$.*

Sempre con riferimento all'esempio precedente, una volta individuato il minore

$$M_0 = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix}$$

che ha determinante diverso da 0, in virtù del Teorema di Kronecker non è più necessario controllare *tutti* i minori di ordine 3 di A , ma solamente quelli orlati di M_0 , precisamente

$$\begin{pmatrix} 1 & 1 & 1 \\ 3 & -1 & -3 \\ 6 & 2 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 3 & -1 & -1 \\ 6 & 2 & 2 \end{pmatrix}.$$

Ricordiamo che, dato uno spazio vettoriale V di dimensione n e data una base $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di V possiamo considerare l'applicazione

$$\Phi_B : V \longrightarrow \mathbb{R}^n$$

tale che $\Phi_B(\mathbf{v}) = (x_1, \dots, x_n)$, dove $x_1, \dots, x_n \in \mathbb{R}$ sono le componenti di \mathbf{v} rispetto alla base B . Tale applicazione, come è facile verificare (si invita il lettore a farlo!), è un isomorfismo esplicito tra V e \mathbb{R}^n . Usiamo tale isomorfismo per dimostrare un legame tra rango e applicazioni lineari:

Teorema 1.15.9. *Sia $f : V \longrightarrow W$ una applicazione lineare, B una base di V e B' una base di W . Allora il rango della matrice associata ad f rispetto alle basi B e B' è proprio uguale alla dimensione di $\text{Im}(f)$.*

Dimostrazione. Sia $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ base di V , $B' = \{\mathbf{e}'_1, \dots, \mathbf{e}'_m\}$ base di W e $A = M_{B'B}(f)$ la matrice associata ad f rispetto alle basi B e B' . Per ogni $j \in \{1, \dots, n\}$ si ha

$$f(\mathbf{e}_j) = a_{1j}\mathbf{e}'_1 + a_{2j}\mathbf{e}'_2 + \dots + a_{mj}\mathbf{e}'_m.$$

Si noti che la m -pla $\Phi_{B'}(f(\mathbf{e}_j)) = (a_{1j}, a_{2j}, \dots, a_{mj})$ rappresenta la j -esima colonna A^j della matrice $A = M_{B'B}(f)$, per come è definita la matrice associata ad una applicazione lineare. Allora

$$\begin{aligned} \Phi_{B'}(\text{Im}(f)) &= \Phi_{B'}(L(f(\mathbf{e}_1), f(\mathbf{e}_2), \dots, f(\mathbf{e}_n))) \\ &= L(\Phi_{B'}(f(\mathbf{e}_1)), \Phi_{B'}(f(\mathbf{e}_2)), \dots, \Phi_{B'}(f(\mathbf{e}_n))) \\ &= L(A^1, A^2, \dots, A^n). \end{aligned}$$

Di conseguenza, la restrizione di $\Phi_{B'}$ a $\text{Im}(f)$

$$\Phi_{B'}|_{\text{Im}(f)} : \text{Im}(f) \longrightarrow L(A^1, A^2, \dots, A^n)$$

è ancora un isomorfismo. Poichè spazi isomorfi hanno la stessa dimensione, si ha

$$\dim(\text{Im}(f)) = \dim(L(A^1, A^2, \dots, A^n)) = \text{rg}(M_{B'B}(f)).$$

□

1.16 Teorema di Rouché - Capelli

In questa sezione, utilizzando la teoria degli spazi vettoriali, dimostreremo un criterio per stabilire se un sistema lineare ammette soluzioni, e un modo per trovare esplicitamente tali soluzioni.

Consideriamo un sistema lineare di m equazioni in n incognite

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases} \quad (1.55)$$

che può essere scritto in forma matriciale come

$$AX = B,$$

dove

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathcal{M}_{m \times n}(\mathbb{R})$$

è la matrice dei coefficienti del sistema,

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathcal{M}_{m \times 1}(\mathbb{R})$$

è la matrice dei termini noti e

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n \times 1}(\mathbb{R})$$

è la matrice delle incognite. Indichiamo con

$$(A|B) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \in \mathcal{M}_{m \times (n+1)}(\mathbb{R})$$

la matrice completa del sistema.

Ricordiamo che un sistema lineare si dice *compatibile* se esiste almeno una soluzione, cioè una n -pla $X_0 \in \mathbb{R}^n$ tale che $AX_0 = B$. Il prossimo teorema ci fornisce una esplicita condizione per stabilire se un sistema è compatibile e per trovare l'insieme delle sue soluzioni.

Teorema 1.16.1 (Teorema di Rouché - Capelli). *Il sistema lineare*

$$AX = B \quad (1.56)$$

è compatibile se e solo se $\text{rg}(A) = \text{rg}(A|B)$. Inoltre, se il sistema è compatibile, e $X_0 \in \mathbb{R}^n$ è una soluzione, allora l'insieme S delle soluzioni è dato da

$$S = \{X + X_0 : X \in V_A\} \quad (1.57)$$

dove $V_A := \{X \in \mathbb{R}^n : AX = \mathbf{0}\}$ è lo spazio vettoriale delle soluzioni del sistema lineare omogeneo

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

Infine,

$$\dim(V_A) = n - \text{rg}(A).$$

Dimostrazione. Il sistema (1.56) è compatibile se e solo se esiste $X_0 = (x_1^0, \dots, x_n^0) \in \mathbb{R}^n$ tale che $AX_0 = B$, vale a dire

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Quindi, sviluppando i calcoli nell'uguaglianza precedente, abbiamo che il sistema è compatibile se e solo se

$$\begin{pmatrix} a_{11}x_1^0 + a_{12}x_2^0 + \cdots + a_{1n}x_n^0 \\ a_{21}x_1^0 + a_{22}x_2^0 + \cdots + a_{2n}x_n^0 \\ \vdots \\ a_{m1}x_1^0 + a_{m2}x_2^0 + \cdots + a_{mn}x_n^0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

cioè

$$\begin{pmatrix} a_{11}x_1^0 \\ a_{21}x_1^0 \\ \vdots \\ a_{m1}x_1^0 \end{pmatrix} + \begin{pmatrix} a_{12}x_2^0 \\ a_{22}x_2^0 \\ \vdots \\ a_{m2}x_2^0 \end{pmatrix} + \cdots + \begin{pmatrix} a_{1n}x_n^0 \\ a_{2n}x_n^0 \\ \vdots \\ a_{mn}x_n^0 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

che è equivalente a

$$x_1^0 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2^0 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n^0 \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

In conclusione, il sistema lineare (1.56) ammette almeno una soluzione se e solo se esistono $x_1^0, x_2^0, \dots, x_n^0 \in \mathbb{R}$ tali che

$$x_1^0 A^1 + \cdots + x_n^0 A^n = B.$$

Questo è vero se e solo se

$$L(A^1, \dots, A^n) = L(A^1, \dots, A^n, B).$$

E, dato che $L(A^1, \dots, A^n)$ è un sottospazio di $L(A^1, \dots, A^n, B)$, l'ultima uguaglianza equivale a dire che $\dim(L(A^1, \dots, A^n)) = \dim(L(A^1, \dots, A^n, B))$ e quindi il sistema è compatibile se e solo se $\text{rg}(A) = \text{rg}(A|B)$.

Supponiamo ora che il sistema sia compatibile. Allora esiste una soluzione $X_0 \in \mathbb{R}^n$. Allora, se $X \in V_A$ (cioè è una soluzione del sistema lineare omogeneo $AX = \mathbf{0}$), si ha

$$A(X_0 + X) = AX_0 + AX = B + \mathbf{0} = B.$$

Viceversa, se $Y \in \mathbb{R}^n$ è una soluzione del sistema (1.56), allora $AY = B = AX_0$, da cui, posto $X := Y - X_0$, abbiamo $AX = \mathbf{0}$ e quindi $Y = X_0 + X$, con $X \in V_A$. Questo prova la (1.57).

Rimane da dimostrare l'ultima parte dell'enunciato. Consideriamo l'applicazione lineare

$$\begin{aligned} f_A : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ X &\mapsto AX \end{aligned}$$

che ha come matrice associata rispetto alle basi canoniche B_C e B'_C , rispettivamente di \mathbb{R}^n e \mathbb{R}^m , proprio la matrice A . Allora

$$\begin{aligned}\ker(f_A) &= \{X \in \mathbb{R}^n : f_A(X) = \mathbf{0}\} \\ &= \{X \in \mathbb{R}^n : AX = \mathbf{0}\} \\ &= V_A.\end{aligned}$$

Allora usando l'equazione dimensionale si ha

$$\begin{aligned}\dim(V_A) &= \dim(\ker(f_A)) \\ &= \dim(\mathbb{R}^n) - \dim(\text{Im}(f_A)) \\ &= n - \text{rg}(M_{B_C B'_C}(f)) \\ &= n - \text{rg}(A).\end{aligned}$$

□

Osservazione 1.16.2. *Dal teorema precedente abbiamo quindi se un sistema lineare (1.56) è compatibile, ogni soluzione del sistema può essere scritta come somma di una soluzione particolare X_0 e di un vettore di V_A . Quest'ultimo, a sua volta, è dato da una combinazione lineare di $n - \text{rg}(A)$ soluzioni del sistema omogeneo $AX = \mathbf{0}$ linearmente indipendenti.*

Un modo alternativo, e molto comune, per esprimere il fatto che l'insieme di un sistema lineare compatibile, dato dalla (1.57), dipenda da $n - \text{rg}(A)$ vettori di \mathbb{R}^n (la base di V_A) è dire che il sistema ammette $\infty^{n-\text{rg}(A)}$ soluzioni.

Al fine di determinare esplicitamente le soluzioni di un sistema lineare, è utile dare la seguente definizione. Per ogni $i \in \{1, \dots, m\}$, denotiamo con

$$P_i(x) := a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n$$

la i -esima equazione del sistema lineare (1.55), che quindi può essere riscritto in forma più compatta come

$$\begin{cases} P_1(x) &= b_1 \\ P_2(x) &= b_2 \\ \vdots & \vdots \\ P_m(x) &= b_m \end{cases} \quad (1.58)$$

Chiameremo combinazione lineare delle equazioni $P_1(x) = b_1, P_2(x) = b_2, \dots, P_m(x) = b_m$ con coefficienti $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}$ l'equazione

$$\lambda_1 P_1(x) + \lambda_2 P_2(x) + \dots + \lambda_m P_m(x) = \lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_m b_m.$$

Per esempio, la combinazione lineare delle equazioni

$$2x + 4y + 3z = 2, \quad 3x + 6y - z = -4$$

di coefficienti 3 e -2 è l'equazione

$$3 \cdot (2x + 4y + 3z) + (-2) \cdot (3x + 6y - z) = 3 \cdot 2 + (-2) \cdot (-4)$$

cioè

$$4y + 11z = 14.$$

Ricordiamo che due sistemi lineari si dicono *equivalenti* se hanno lo stesso insieme delle soluzioni. Ebbene, abbiamo il seguente risultato.

Proposizione 1.16.3. *Con le notazioni precedenti, supponiamo che l'equazione $P_i(x) = b_i$ sia combinazione lineare delle rimanenti equazioni. Allora il sistema lineare (1.58) è equivalente al sistema di $(m - 1)$ equazioni in n incognite*

$$\begin{cases} P_1(x) &= b_1 \\ \vdots & \vdots \\ P_{i-1}(x) &= b_{i-1} \\ P_{i+1}(x) &= b_{i+1} \\ \vdots & \vdots \\ P_m(x) &= b_m \end{cases} \quad (1.59)$$

Dimostrazione. Chiamiamo S_1 e S_2 gli insiemi delle soluzioni dei sistemi (1.58) e (1.59), rispettivamente. Vogliamo dimostrare che $S_1 = S_2$. Dato che, chiaramente, $S_1 \subseteq S_2$, dobbiamo dimostrare che $S_2 \subseteq S_1$, cioè, visto che i due sistemi hanno le stesse equazioni tranne una, che se $X_0 \in S_2$ allora $P_i(X_0) = b_i$. Per ipotesi sappiamo che la i -esima equazione del sistema (1.58) è combinazione lineare delle rimanenti, e quindi esistono $\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_m \in \mathbb{R}$ tali che

$$P_i = \lambda_1 P_1 + \dots + \lambda_{i-1} P_{i-1} + \lambda_{i+1} P_{i+1} + \dots + \lambda_m P_m$$

e

$$b_i = \lambda_1 b_1 + \dots + \lambda_{i-1} b_{i-1} + \lambda_{i+1} b_{i+1} + \dots + \lambda_m b_m.$$

Allora

$$\begin{aligned} P_i(X_0) &= \lambda_1 P_1(X_0) + \dots + \lambda_{i-1} P_{i-1}(X_0) + \lambda_{i+1} P_{i+1}(X_0) + \dots + \lambda_m P_m(X_0) \\ &= \lambda_1 b_1 + \dots + \lambda_{i-1} b_{i-1} + \lambda_{i+1} b_{i+1} + \dots + \lambda_m b_m \\ &= b_i \end{aligned}$$

□

Sulla scorta della Proposizione 1.16.3 vediamo con un esempio come determinare concretamente l'insieme delle soluzioni di un sistema lineare usando il Teorema di Rouché-Capelli.

Consideriamo il sistema lineare

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ -2x_1 + x_3 = 0 \\ 6x_1 + 2x_2 = 2 \end{cases} \quad (1.60)$$

La matrice dei coefficienti e la matrice completa del sistema sono, rispettivamente,

$$A = \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ 6 & 2 & 0 \end{pmatrix}, \quad (A|B) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -2 & 0 & 1 & 0 \\ 6 & 2 & 0 & 2 \end{pmatrix}.$$

Dato che $\det(A) = 0$, dal Teorema 1.15.5 si ha subito che $\text{rg}(A) < 3$. E, dato che

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad (1.61)$$

è un minore di A di ordine 2 avente determinante non nullo, per il Teorema 1.15.7 concludiamo che $\text{rg}(A) = 2$. Si noti che (1.61) è un minore anche di $(A|B)$ e quindi $\text{rg}(A|B) \geq 2$. Dobbiamo ora controllare se $\text{rg}(A|B) = 3$. Per farlo dobbiamo controllare se ci sono minori di ordine 3 di $(A|B)$ con il determinante diverso da 0. Usando il Teorema di Kronecker, possiamo restringere la nostra ricerca ai soli orlati del minore (1.61), che sono

$$\begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ 6 & 2 & 0 \end{pmatrix} = A$$

e

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

Entrambi questi minore hanno determinante nullo (in particolare, non valeva nemmeno la pena di calcolare il determinante del primo minore, visto che esso coincide con A e avevamo già calcolato in precedenza il determinante di A). Concludiamo quindi che

$$\text{rg}(A|B) = 2 = \text{rg}(A)$$

e per il Teorema di Rouché-Capelli il sistema è compatibile ed ammette $\infty^{3-\text{rg}(A)} = \infty^1$ soluzioni, cioè infinite soluzioni dipendenti da un parametro reale. Troviamo esplicitamente tali soluzioni. Si noti che, dato che $\text{rg}(A|B) = 2$, ed il minore di $(A|B)$ di ordine 2 con il determinante diverso da 0, cioè il minore (1.61), corrisponde alle prime due righe di $(A|B)$ (e anche di A), la terza riga di $(A|B)$ è combinazione lineare delle prime due (infatti se così non fosse, avremmo 3 righe di $(A|B)$ linearmente indipendenti e concluderemmo che $\text{rg}(A|B) = 3$). Allora, per la Proposizione 1.16.3, il sistema è equivalente al sistema di due equazioni

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ -2x_1 + x_3 = 0 \end{cases}$$

Ora riscriviamo tale sistema, isolando a sinistra del simbolo di uguaglianza i coefficienti corrispondenti al minore (1.61), e portando a secondo membro tutto il resto:

$$\begin{cases} x_2 + x_3 = 1 - x_1 \\ x_3 = 2x_1 \end{cases}$$

Trattiamo, formalmente, l'incognita che abbiamo portato a secondo membro (x_1) come se fosse un termine noto: di fatto, per ogni fissato valore che assume x_1 esso *sarà* a tutti gli effetti un termine noto; per ricordarci di questo, poniamo $t := x_1$. Il sistema diventa

$$\begin{cases} x_2 + x_3 = 1 - t \\ x_3 = 2t \end{cases}$$

che possiamo trattare come un sistema di 2 equazioni nelle 2 incognite x_2 e x_3 . Si noti che, *per costruzione*, tale sistema è di Cramer. Infatti la matrice dei coefficienti del sistema è proprio il minore (1.61), che già sappiamo avere il determinante diverso da 0. Quindi tale sistema è compatibile ed ammette un'unica soluzione, che si può trovare con la formula di Cramer, oppure semplicemente sostituendo il valore di x_3 nella prima equazione. Si ottiene così $x_2 = 1 - 3t$ e $x_3 = 2t$, per ogni fissato valore di $t \in \mathbb{R}$. Ricordando che avevamo posto $x_1 = t$, concludiamo che l'insieme S delle soluzioni del sistema (1.60) è

$$\begin{aligned} S &= \{(t, 1 - 3t, 2t) : t \in \mathbb{R}\} \\ &= \{(0, 1, 0) + (t, -3t, 2t) : t \in \mathbb{R}\} \\ &= \{(0, 1, 0) + t(1, -3, 2) : t \in \mathbb{R}\}. \end{aligned}$$

Si noti la somiglianza dell'ultima uguaglianza con la (1.57) del Teorema di Rouché-Capelli. In questo caso $X_0 = (0, 1, 0)$ è una soluzione particolare

del sistema (quella ottenuta per $t = 0$), mentre $V_A = \{t(1, -3, 2) : t \in \mathbb{R}\} = L(1, -3, 2)$. Ora dovrebbe essere anche più chiaro il significato della terminologia ∞^1 *soluzioni*: il sistema ammette infatti infinite soluzioni, che si ottengono tutte assegnando un arbitrario valore reale al parametro t .

Arrivati a questo punto, il lettore potrebbe domandarsi perché abbiamo dedicato tanta attenzione alla risoluzione di sistemi lineari. L'importanza dei sistemi lineari sta soprattutto nello studio della Geometria Affine e della Geometria Proiettiva. Ma, anche all'interno dell'Algebra Lineare, i sistemi lineari giocano un ruolo importante. Dimostriamo infatti ora che i sottospazi vettoriali di \mathbb{R}^n sono tutti e soli gli spazi delle soluzioni di qualche sistema lineare omogeneo in n incognite.

Proposizione 1.16.4. *Sia W un sottospazio vettoriale di \mathbb{R}^n . Esiste un sistema lineare*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

il cui spazio delle soluzioni coincide con W .

Dimostrazione. Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ una base di W . Per il Teorema di completamento ad una base esistono $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n \in \mathbb{R}^n$ tali che

$$B = \{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$$

è una base di V . Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ l'applicazione lineare tale che

$$\begin{aligned} f(\mathbf{v}_1) &= \mathbf{0} \\ &\vdots \\ f(\mathbf{v}_k) &= \mathbf{0} \\ f(\mathbf{v}_{k+1}) &= \mathbf{v}_{k+1} \\ &\vdots \\ f(\mathbf{v}_n) &= \mathbf{v}_n \end{aligned}$$

Sia $A := M_{B_C B_C}(f)$, dove B_C denota la base canonica di \mathbb{R}^n ed S lo spazio delle soluzioni del sistema lineare omogeneo

$$AX = \mathbf{0}.$$

Ebbene dimostriamo che $W = S$. Intanto osserviamo che $W = \ker(f)$. Infatti per ogni $\mathbf{w} \in W$ esistono $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tali che $\mathbf{w} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k$. Allora $f(\mathbf{w}) = \alpha_1 f(\mathbf{v}_1) + \dots + \alpha_k f(\mathbf{v}_k) = \mathbf{0}$, e quindi $\mathbf{w} \in \ker(f)$. Questo prova che $W \subseteq \ker(f)$. Inoltre i due spazi hanno la stessa dimensione. Infatti

$$\begin{aligned} \dim(\ker(f)) &= \dim(\mathbb{R}^n) - \dim(\text{Im}(f)) \\ &= n - \text{rg}(M_{BB}(f)) \\ &= n - \text{rg} \begin{pmatrix} O_k & O_{n-k,k} \\ O_{k,n-k} & I_{n-k} \end{pmatrix} \\ &= n - (n - k) \\ &= k \\ &= \dim(W). \end{aligned}$$

Conseguentemente $W = \ker(f)$. Allora, tenendo conto che A è la matrice associata ad f rispetto alla base canonica di \mathbb{R}^n , si ha che $\mathbf{v} = (x_1, \dots, x_n) \in W$ se e solo se

$$\mathbf{0} = f(\mathbf{v}) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

cioè se e solo se $\mathbf{v} = (x_1, \dots, x_n)$ è soluzione del sistema lineare omogeneo $AX = \mathbf{0}$. \square

Esercizi proposti

1. Trovare il rango della seguente matrice

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \\ 4 & 1 & -2 & -5 & -8 & -11 \\ 0 & 1 & 0 & -1 & 0 & 0 \end{pmatrix}$$

2. Siano dati i seguenti sei vettori di \mathbb{R}^4

$$\mathbf{v}_1 = (1, 6, 4, 0)$$

$$\mathbf{v}_2 = (2, 5, 1, 1)$$

$$\mathbf{v}_3 = (3, 4, -2, 0)$$

$$\mathbf{v}_4 = (4, 3, -5, -1)$$

$$\mathbf{v}_5 = (5, 2, -8, 0)$$

$$\mathbf{v}_6 = (6, 1, -11, 0)$$

Senza utilizzare il metodo degli scarti successivi, estrarre il massimo numero di vettori linearmente indipendenti tra i sei vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_6$. (Suggerimento: usare i risultati ottenuti nell'esercizio precedente)

3. Al variare di $k \in \mathbb{R}$ determinare il rango della seguente matrice

$$\begin{pmatrix} 1 & k & 3 \\ 2k & 2 & 6 \\ -k & -1 & -3 \end{pmatrix}$$

4. Utilizzando il Teorema di Rouché-Capelli, trovare lo spazio delle soluzioni, e una sua base, del seguente sistema lineare omogeneo

$$\begin{cases} 2x_1 - 3x_2 + x_3 + 2x_4 = 0 \\ x_1 + 5x_2 - 2x_3 - 3x_4 = 0 \\ 5x_1 - x_2 + x_4 = 0 \\ 3x_1 + 2x_2 - x_3 - x_4 = 0 \\ x_1 - 8x_2 + 3x_3 + 5x_4 = 0 \end{cases}$$

5. Utilizzando il Teorema di Rouché-Capelli, dire se il seguente sistema è compatibile ed in caso affermativo trovare l'insieme delle soluzioni

$$\begin{cases} x_1 + 3x_2 - x_3 + x_4 = 1 \\ x_2 + x_3 - x_4 = 2 \\ 2x_3 + 2x_4 = 3 \end{cases}$$

6. Utilizzando il Teorema di Rouché-Capelli, al variare di $k \in \mathbb{R}$ stabilire se il seguente sistema lineare è compatibile ed in caso affermativo trovarne le soluzioni

$$\begin{cases} x + 2y - z & = 1 \\ -x + y + 2z & = 0 \\ 2x - y & = 0 \\ x - y - z & = k \end{cases}$$

7. Utilizzando il Teorema di Rouché-Capelli, al variare di $k \in \mathbb{R}$ stabilire se il seguente sistema lineare è compatibile ed in caso affermativo trovarne le soluzioni

$$\begin{cases} kx_1 + kx_2 + 2kx_3 + kx_4 & = 0 \\ kx_1 & + (k - 1)x_3 + kx_4 & = 0 \\ & kx_2 + (k - 2)x_3 & = 0 \end{cases}$$

8. Utilizzando il Teorema di Rouché-Capelli, al variare di $k \in \mathbb{R}$ stabilire se il seguente sistema lineare è compatibile ed in caso affermativo trovarne le soluzioni

$$\begin{cases} x + 2y - 2z & = 0 \\ x - y - z & = -2h \\ 2x + ky - 2z & = -1 \end{cases}$$