Sistemi di equazioni lineari

maurizio candilera

October 1, 2018

Sistemi lineari

Raccogliamo le notazioni e la terminologia che useremo per parlare di sistemi lineari.

Un sistema di m equazioni lineari, a coefficenti in K, nelle incognite $x_1, \dots, x_n$, si scrive nella forma

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

ove

$$A = (a_{ij})_{\substack{1\le i\le m \\ 1\le j\le n}} \in M_{m\times n}(K)$$

e $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m$.

Il sistema si dice omogeneo se la colonna dei termini noti $b = 0 \in K^m$. Le soluzioni del sistema lineare sono le n-uple $\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n$ che, sostituite nelle equazioni in luogo di $x_1, \dots, x_n$, le rendono tutte delle identità.

La matrice A è la matrice dei coefficenti del sistema, la colonna b è la colonna dei termini noti e la matrice (a blocchi) $(A|b)$ è detta la matrice completa del sistema. Introducendo la colonna delle incognite, $x = {}^t(x_1, \dots, x_n)$, e utilizzando il prodotto riga per colonna, il sistema si scrive più brevemente nella forma $\boxed{Ax = b}$.

Sia $Ax = b$ il sistema lineare di matrice completa $(A|b) \in M_{m\times(n+1)}(K)$. La matrice A definisce un'applicazione lineare $\phi : K^n \to K^m$, ponendo $A = \alpha_{\mathcal{E}_n,\mathcal{E}_m}(\phi)$; cioè l'applicazione $v \mapsto Av$ per ogni $v \in K^n$. Il vettore b appartiene a K^m e si ha

$$\boxed{\mathrm{Sol}(A|b) = \{\, \xi \in K^n \mid A\xi = b \,\} = \{\, v \in K^n \mid \phi(v) = b \,\} = \phi^{-1}(b)}$$

ovvero risolvere il sistema equivale a trovare la controimmagine di un vettore di K^m tramite ϕ.

Sia $Ax = b$ il sistema lineare di matrice completa $(A|b) \in M_{m\times(n+1)}(K)$. La matrice A definisce un'applicazione lineare $\phi : K^n \to K^m$, ponendo $A = \alpha_{\mathcal{E}_n,\mathcal{E}_m}(\phi)$; cioè l'applicazione $v \mapsto Av$ per ogni $v \in K^n$. Il vettore b appartiene a K^m e si ha

$$\boxed{\mathrm{Sol}(A|b) = \{\, \xi \in K^n \mid A\xi = b \,\} = \{\, v \in K^n \mid \phi(v) = b \,\} = \phi^{-1}(b)}$$

ovvero risolvere il sistema equivale a trovare la controimmagine di un vettore di K^m tramite ϕ. Per quanto visto nella sezione sulle applicazioni lineari

$$\phi^{-1}(b) = \begin{cases} \varnothing & \text{se } b \notin \operatorname{im}\phi \\ v_0 + \ker\phi & \text{se } \exists v_0 \mid \phi(v_0) = b \end{cases}.$$

Dunque, $\mathrm{Sol}(A|b) \neq \varnothing$ se, e solo se, b appartiene al sottospazio di K^m generato dalle colonne di A, ovvero se, e solo se, $\mathrm{rk}(A|b) = \mathrm{rk}\, A$. In tal caso, preso un qualsiasi $v_0 \in \mathrm{Sol}(A|b)$, si ha $\mathrm{Sol}(A|b) = v_0 + \mathrm{Sol}(A|0)$ e, per la formula delle dimensioni, $\mathrm{Sol}(A|0) = \ker\phi$ è un sottospazio vettoriale di K^n di dimensione $n - \mathrm{rk}\, A$.

Teorema di Rouché-Capelli

Possiamo riassumere il contenuto della slide precedente nel seguente

Teorema (Rouché-Capelli)

Sia K un campo. Il sistema lineare $Ax = b$, di matrice completa $(A|b) \in M_{m\times(n+1)}(K)$, ha soluzione se, e solo se, $\mathrm{rk}(A|b) = \mathrm{rk}\, A$.
In tal caso si ha $\mathrm{Sol}(A|b) = v_0 + \mathrm{Sol}(A|0)$, ove $v_0 \in \mathrm{Sol}(A|b)$ e $Ax = 0$ è detto il sistema omogeneo associato, che ha come soluzione un sottospazio vettoriale di K^n di dimensione $n - \mathrm{rk}\, A$.
In particolare, il sistema $Ax = b$ ha come soluzione un sottospazio vettoriale di K^n se, e solo se, è omogeneo, ovvero se, e solo se, $b = 0$.

Dunque aver interpretato la matrice dei coefficienti del sistema come la matrice di un'applicazione lineare, ci ha permesso di dare una condizione necessaria e sufficiente all'esistenza di soluzioni per un sistema lineare e di descrivere completamente tali soluzioni quando esistono.

Tecnica di eliminazione (Gauss)

L'interpretazione delle soluzioni di un sistema come controimmagine di un vettore tramite un'applicazione lineare, permette di usare i cambiamenti di base per determinare più facilmente l'insieme delle soluzioni. Precisamente, faremo cambiamenti di base nel codominio dell'applicazione lineare, in modo da lasciare invariate le soluzioni (che sono un sottoinsieme del dominio) e rendere più evidente l'esistenza di soluzioni e la loro determinazione.

Tecnica di eliminazione (Gauss)

L'interpretazione delle soluzioni di un sistema come controimmagine di un vettore tramite un'applicazione lineare, permette di usare i cambiamenti di base per determinare più facilmente l'insieme delle soluzioni. Precisamente, faremo cambiamenti di base nel codominio dell'applicazione lineare, in modo da lasciare invariate le soluzioni (che sono un sottoinsieme del dominio) e rendere più evidente l'esistenza di soluzioni e la loro determinazione.
Abbiamo visto alla fine della sezione sugli spazi vettoriali come ogni cambiamento di base si possa ottenere iterando un numero finito di operazioni elementari sui vettori di una base data.

Precisamente, dato $\phi : K^n \to K^m$ e data una base $\mathcal{W} = \{w_1, \dots, w_m\}$ di K^m le operazioni elementari che faremo sono

- scambiare di posto due dei vettori: $w_i \mapsto w_j$, $w_j \mapsto w_i$;
- moltiplicare uno dei vettori per uno scalare $c \neq 0$: $w_i \mapsto cw_i$;
- aggiungere ad uno dei vettori un multiplo di un altro: $w_i \mapsto w_i + aw_j$, con $j \neq i$ e $a \in K$.

Tali operazioni cambiano la matrice di ϕ e le coordinate dei vettori di K^m, ma non cambiano le coordinate in K^n e quindi le controimmagini dei vettori.

Sia quindi $Ax = b$ un sistema lineare, con $A \in M_{m\times n}(K)$. Detta $\phi : K^n \to K^m$ l'applicazione lineare di matrice A nelle basi canoniche, fare operazioni elementari sulla base di K^m corrisponde a moltiplicare A a sinistra per le corrispondenti matrici di cambiamento di base (*matrici elementari*).

Le matrici delle operazioni elementari sono le seguenti:

- scambio dei vettori w_i e w_j: $\boxed{H(i,j) = \mathbf{1}_m + (\varepsilon(i,j) + \varepsilon(j,i) - \varepsilon(i,i) - \varepsilon(j,j))}$;
- moltiplicazione di w_i per $c \neq 0$: $\boxed{D(j,c) = \mathbf{1}_m + (c-1)\varepsilon(i,i)}$;
- sommare a w_i il vettore aw_j con $j \neq i$ e $a \in K$: $\boxed{E(i,j,a) = \mathbf{1}_m + a\varepsilon(i,j)}$;

ove $\mathcal{B} = \{\ \varepsilon(i,j) \mid 1 \leq i,j \leq m\ \}$ ̀la base canonica di $M_m(K)$.

Esempio : Sia $m = 4$. Allora

$$H(1,3) = \begin{pmatrix} 0&0&1&0\\ 0&1&0&0\\ 1&0&0&0\\ 0&0&0&1 \end{pmatrix}, \quad D(2,c) = \begin{pmatrix} 1&0&0&0\\ 0&c&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}, \quad E(2,4,a) = \begin{pmatrix} 1&0&0&0\\ 0&1&0&a\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}.$$

Queste operazioni si traducono in analoghe operazioni sulle *righe* della matrice (completa) del sistema e hanno lo scopo di ottenere una matrice a scala; ovvero una matrice in cui

(a) le eventuali righe nulle sono tutte al di sotto delle righe non nulle;
(b) l'elemento non nullo più a sinistra in una riga (non nulla) è detto il *pivot* e ha

In una matrice a scala tutte le entrate al di sotto di un pivot sono nulle. Inoltre, sempre con operazioni elementari sulle righe, possiamo fare in modo che tutti i pivot siano uguali a 1 e che siano nulle anche le entrate della matrice poste al di sopra dei pivot. Una matrice di questa forma si dice una matrice a scala ridotta.

Osservazione

Con operazioni elementari sulle righe ogni matrice può essere portata in forma a scala ridotta (e tale forma è unica).

Diamo un esempio di questo procedimento lasciando al lettore il compito di formalizzarlo per ottenere una dimostrazione dell'Osservazione qui sopra.

Esempio: Sia $A = \begin{pmatrix} 1 & 0 & 2 & 2 & 4 & -5 & 0 \\ 0 & 1 & -1 & 0 & 1 & 1 & 0 \\ -1 & 1 & -3 & 1 & 3 & 0 & -2 \\ 0 & 1 & -1 & 0 & 1 & 1 & -2 \\ 2 & -1 & 5 & 0 & -1 & -3 & 2 \end{pmatrix}$. Con operazioni di Gauss sulle righe possiamo portare la matrice A in forma a scala ridotta:

$$A \sim \begin{array}{r} I \\ II \\ III+I-II \\ IV-II \\ V-2I+II \end{array} \begin{pmatrix} 1 & 0 & 2 & 2 & 4 & -5 & 0 \\ 0 & 1 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 6 & -6 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -4 & -8 & 8 & 2 \end{pmatrix} \sim \begin{array}{r} I-2(III-IV)/3 \\ II \\ (III-IV)/3 \\ -IV/2 \\ 3V+4III-IV \end{array} \begin{pmatrix} 1 & 0 & 2 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ad ogni passo le operazioni indicate si riferiscono alle righe della matrice al passo precedente.

Interpretiamo l'esempio precedente, in termini di sistemi lineari; ovvero supponiamo che $A \in M_{5\times 7}(\mathbb{Q})$ sia la matrice dei coefficenti di un sistema lineare omogeneo. Il sistema ha le stesse soluzioni del sistema corrispondente alla matrice a scala ridotta. Ovvero

$$\begin{cases} x_1 + 2x_3 + 2x_4 + 4x_5 - 5x_6 = 0 \\ x_2 - x_3 + x_5 + x_6 = 0 \\ -x_1 + x_2 - 3x_3 + x_4 + 3x_5 - 2x_7 = 0 \\ x_2 - x_3 + x_5 + x_6 - 2x_7 = 0 \\ 2x_1 - x_2 + 5x_3 - x_5 - 3x_6 + 2x_7 = 0 \end{cases} \quad \text{ha le stesse soluzioni di} \quad \begin{cases} x_1 + 2x_3 - x_6 = 0 \\ x_2 - x_3 + x_5 + x_6 = 0 \\ x_4 + 2x_5 - 2x_6 = 0 \\ x_7 = 0 \end{cases}$$

Nella forma ridotta si vede chiaramente che la matrice del sistema ha rango 4 (il numero dei pivot). Dunque le soluzioni formano un sottospazio vettoriale di dimensione 3 di $\mathbb{Q}^7$ una cui base è data da tre soluzioni linearmente indipendenti.

$$\mathrm{Sol}(A|0) = \left\langle \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \quad \text{ovvero} \quad \begin{cases} x_1 = -2x_3 + x_6 \\ x_2 = x_3 - x_5 - x_6 \\ x_4 = -2x_5 + 2x_6 \\ x_7 = 0 \end{cases}$$

Cioè, dal sistema ridotto si ottengono le variabili corrispondenti ai pivot in funzione delle rimanenti, che assumono il ruolo di parametri. In tal modo abbiamo ricavato le soluzioni indicate.

Esempio: Trovare le soluzioni, se esistono, del sistema lineare

$$\begin{cases} x_1 - 2x_2 + 2x_3 + x_4 - 3x_5 = 3 \\ 2x_3 + x_4 - x_5 = 0 \\ 3x_1 - 6x_2 - 2x_4 = 13 \\ 2x_1 - 4x_2 + x_3 - x_4 = 9 \\ -x_1 + 2x_2 + x_4 - x_5 = -5 \end{cases} .$$

Svolg.: Portiamo in forma a scala (ridotta) la matrice del sistema, ovvero:

$$\left(\begin{array}{rrrrr|r} 1 & -2 & 2 & 1 & -3 & 3 \\ 0 & 0 & 2 & 1 & -1 & 0 \\ 3 & -6 & 0 & -2 & 0 & 13 \\ 2 & -4 & 1 & -1 & 0 & 9 \\ -1 & 2 & 0 & 1 & -1 & -5 \end{array}\right) \sim \begin{array}{r} -V \\ IV+2V \\ III+3V \\ II \\ I+V \end{array} \left(\begin{array}{rrrrr|r} 1 & -2 & 0 & -1 & 1 & 5 \\ 0 & 0 & 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & 1 & -3 & -2 \\ 0 & 0 & 2 & 1 & -1 & 0 \\ 0 & 0 & 2 & 2 & -4 & -2 \end{array}\right) \sim \begin{array}{r} I+III \\ II+III \\ III \\ V-2II \\ IV-2II+III \end{array} \left(\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & -2 & 3 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

La matrice ha rango 3 (sia la matrice dei coefficenti che quella completa); quindi, per il Teorema di Rouché-Capelli, il sistema ha soluzione e il sistema omogeneo associato ha un sottospazio di dimensione 2 di soluzioni, per cui

$$\mathrm{Sol}(A|b) = \begin{pmatrix} 3 \\ 0 \\ 1 \\ -2 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} 2 \\ 0 \\ -1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle .$$

Esempio: Dire se la matrice $A = \begin{pmatrix} -1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & -2 & -3 \\ -1 & -1 & 0 & 0 \end{pmatrix}$ è invertibile e, in caso affermativo, trovare l'inversa.

Svolg.: Le colonne della matrice inversa, A^{-1}, sono soluzioni del sistema lineare di matrice dei coefficenti A e colonna dei termini noti la corrispondente colonna della matrice identità. I sistemi hanno soluzione se, e solo se, lo spazio generato dalle colonne di A contiene la base canonica, ovvero se, e solo se, A ha rango massimo.
Mettiamo quindi a destra le colonne dei termini noti e procediamo portando la matrice A in forma a scala ridotta.

$$\left(\begin{array}{rrrr|rrrr} -1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & -2 & -3 & 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \sim \begin{array}{r} -I \\ II \\ III+2I \\ IV-I+II \end{array} \left(\begin{array}{rrrr|rrrr} 1 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 & 0 & 1 \end{array}\right)$$

$$\sim \begin{array}{r} I-IV \\ II \\ III-IV \\ -III \end{array} \left(\begin{array}{rrrr|rrrr} 1 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -2 & 0 & -1 & 0 \end{array}\right)$$

La matrice A ha rango 4; quindi è invertibile e la sua inversa è la matrice ottenuta dal procedimento di eliminazione, ovvero $A^{-1} = \begin{pmatrix} 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 3 & -1 & 1 & -1 \\ -2 & 0 & -1 & 0 \end{pmatrix}$. $\square$

N.B. Se la matrice di partenza ha rango massimo, allora la sua forma a scala ridotta è la matrice identità e col procedimento qui delineato troviamo la sua inversa; altrimenti la matrice inversa non esiste.

Esempio: Dire se la matrice reale $A = \begin{pmatrix} 2 & 2 & 2 & 1 & 0 \\ 0 & -2 & 2 & 1 & 2 \\ -2 & -4 & 0 & 1 & 2 \end{pmatrix}$ è invertibile a destra e, in caso affermativo, trovare le inverse destre di A.

Svolg.: Le inverse destre di A sono matrici $B \in M_{5\times 3}(\mathbb{R})$ tali che $AB = \mathbf{1}_3$; dunque le colonne di B sono soluzioni del sistema lineare di matrice dei coefficenti A e colonna dei termini noti la corrispondente colonna di $\mathbf{1}_3$.
I sistemi hanno soluzione se, e solo se, lo spazio generato dalle colonne di A ha dimensione 3, ovvero se, e solo se, $\operatorname{rk} A = 3$.
Mettiamo quindi a destra le colonne dei termini noti e procediamo portando la matrice A in forma a scala ridotta.

$$\left(\begin{array}{rrrrr|rrr} 2 & 2 & 2 & 1 & 0 & 1 & 0 & 0 \\ 0 & -2 & 2 & 1 & 2 & 0 & 1 & 0 \\ -2 & -4 & 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \sim \begin{array}{r} (I+II)/2 \\ II \\ III+I-II \end{array} \left(\begin{array}{rrrrr|rrr} 1 & 0 & 2 & 1 & 1 & 1/2 & 1/2 & 0 \\ 0 & -2 & 2 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 & 1 \end{array}\right)$$

$$\sim \begin{array}{r} I-III \\ (III-II)/2 \\ III \end{array} \left(\begin{array}{rrrrr|rrr} 1 & 0 & 2 & 0 & 1 & -1/2 & 3/2 & -1 \\ 0 & 1 & -1 & 0 & -1 & 1/2 & -1 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 & 1 \end{array}\right)$$

La matrice A ha rango 3; quindi è invertibile a destra e le sue inverse destre sono le matrici che hanno come colonne le soluzioni dei sistemi, ovvero $B = \begin{pmatrix} -1/2 & 3/2 & -1 \\ 1/2 & -1 & 1/2 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 2a+b & 2c+d & 2e+f \\ -a-b & -c-d & -e-f \\ -a & -c & -e \\ 0 & 0 & 0 \\ -b & -d & -f \end{pmatrix}$, al variare dei sei parametri reali $(a, b, c, d, e, f) \in \mathbb{R}^6$. $\square$

N.B. Se la matrice A è invertibile a sinistra, il problema di trovare le inverse sinistre si riconduce a quanto appena visto, considerando la trasposta ${}^tA$ in luogo della matrice A e trasponendo le inverse destre di ${}^tA$.

Equivalenza per riga e colonna

Le operazioni elementari fatte sulle matrici sono il primo esempio dello studio di una relazione di equivalenza tra matrici.

Definizione

Due matrici A e B in $M_{m\times n}(K)$ si dicono *riga-equivalenti* (risp. *colonna-equivalenti*) se esiste una matrice invertibile $P \in \mathrm{GL}(m, K)$ (risp. $Q \in \mathrm{GL}(n, K)$) tale che $B = PA$ (risp. $B = AQ$).
Due matrici A e B in $M_{m\times n}(K)$ si dicono *equivalenti* se esistono due matrici invertibili $P \in \mathrm{GL}(m, K)$ e $Q \in \mathrm{GL}(n, K)$ tali che $B = PAQ$.

Le relazioni appena definite sono relazioni di equivalenza tra matrici. Verifichiamolo per l'equivalenza per riga e scriviamo $A \sim B$ per dire che A e B sono riga-equivalenti. Si ha

- (riflessiva) $\boxed{A \sim A}$, perché $A = \mathbf{1}_m A$ e $\mathbf{1}_m \in \mathrm{GL}(m, K)$;
- (simmetrica) $\boxed{A \sim B \;\Rightarrow\; B \sim A}$, perché se $B = PA$, allora $A = P^{-1}B$ e $P, P^{-1} \in \mathrm{GL}(m, K)$;
- (transitiva) $\boxed{(A \sim B,\ B \sim C) \;\Rightarrow\; A \sim C}$, perché se $B = PA$ e $C = RB$, allora $C = (RP)A$ e $P, R \in \mathrm{GL}(m, K) \Rightarrow RP \in \mathrm{GL}(m, K)$.

Osserviamo che *due matrici sono colonna-equivalenti se, e solo se, le matrici trasposte sono riga-equivalenti*, per cui anche in questo caso si ha una relazione di equivalenza.

La relazione di equivalenza tra matrici è, in un certo senso, la "composizione" delle due relazioni di equivalenza per riga e colonna, perché *due matrici* A *e* B *sono equivalenti se, e solo se,* A *è riga-equivalente a una matrice colonna-equivalente a* B.

Osserviamo che *due matrici sono colonna-equivalenti se, e solo se, le matrici trasposte sono riga-equivalenti*, per cui anche in questo caso si ha una relazione di equivalenza.

La relazione di equivalenza tra matrici è, in un certo senso, la "composizione" delle due relazioni di equivalenza per riga e colonna, perché *due matrici A e B sono equivalenti se, e solo se, A è riga-equivalente a una matrice colonna-equivalente a B.*

Le matrici invertibili in $\mathrm{GL}(m, K)$ sono tutte e sole le matrici di cambiamento di base in K^m, per cui si ottengono tutte come prodotto finito di matrici elementari, ovvero di matrici di operazioni di Gauss (ogni cambiamento di base si ottiene con un numero finito di operazioni di Gauss). Quindi, operando sulle righe delle matrici, le abbiamo cambiate ottenendo matrici riga-equivalenti a quelle di partenza e, in particolare, la forma a scala ridotta è una forma canonica all'interno di ogni classe di equivalenza per riga. Ovvero:

- ogni matrice in $M_{m\times n}(K)$ è riga-equivalente a una matrice a scala ridotta;
- due matrici a scala ridotta in $M_{m\times n}(K)$ che siano riga-equivalenti, sono necessariamente uguali.

Dimostriamo le affermazioni precedenti "traducendole" nel linguaggio delle applicazioni lineari.

Proposizione

Sia $\phi : K^n \to K^m$ un'applicazione lineare. Esiste una base $\mathcal{W} = \{w_1, \dots, w_m\}$ di K^m tale che la matrice $\alpha_{\mathcal{E}_n,\mathcal{W}}(\phi)$ sia una matrice a scala ridotta. In particolare, se $r = \operatorname{rk}\phi$, i primi r vettori di $\mathcal{W}$ sono una base di $\operatorname{im}\phi$ e, se $\mathcal{W}' = \{w'_1, \dots, w'_m\}$ è base di K^m tale che la matrice $\alpha_{\mathcal{E}_n,\mathcal{W}'}(\phi)$ sia una matrice a scala ridotta, allora $w_i = w'_i$ per $i = 1, \dots, r$.

dim. Sia $\mathcal{E}_n = \{e_1, \dots, e_n\}$ la base canonica di K^n e consideriamo i vettori $\phi(e_1), \dots, \phi(e_n)$, generatori di $\operatorname{im}\phi$. Se questi vettori sono tutti nulli $\operatorname{rk}\phi = 0$, la matrice di ϕ (in qualunque base di K^m) è la matrice nulla, e tutte le affermazioni sono banalmente vere.
Altrimenti sia $i_1 = \min\{\, i \geq 1 \mid \phi(e_i) \neq 0 \,\}$ e determiniamo i successivi indici $i_2, \dots, i_r$ con la condizione

$$i_k = \min\left\{\, i \geq i_{k-1} \;\middle|\; \phi(e_i) \notin \left\langle \phi(e_{i_1}), \dots, \phi(e_{i_{k-1}}) \right\rangle \right\} \quad \text{per } 2 \leq k \leq r.$$

I vettori $w_1 = \phi(e_{i_1}), \dots, w_r = \phi(e_{i_r})$ così ottenuti sono una base di $\operatorname{im}\phi$ perché sono $r = \dim \operatorname{im}\phi$ e linearmente indipendenti (per costruzione). Se li completiamo a una base $\mathcal{W} = \{w_1, \dots, w_m\}$ di K^m otteniamo una base rispetto a cui la matrice $\alpha_{\mathcal{E}_n,\mathcal{W}}(\phi)$ è a scala ridotta perché le ultime $m - r$ righe sono nulle e le colonne di indice $i_1, \dots, i_r$ corrispondono ai pivot; infatti, per $j = 1, \dots, r$, si ha: $a_{j,i_j} = 1$ e $a_{h,i_j} = 0$ per $h \neq j$; inoltre, per costruzione, ogni colonna è combinazione lineare delle colonne che la precedono.
Dalla costruzione è chiaro che ogni altra base $\mathcal{W}'$ con le condizioni date differisce da questa per la scelta della base di un complementare di $\operatorname{im}\phi$ e non cambia perciò la matrice di ϕ. $\square$

Sia A una matrice e D la matrice a scala ridotta riga-equivalente ad A. Abbiamo visto nella dimostrazione precedente che le colonne di D corrispondenti ai pivot sono una base dell'immagine dell'omomorfismo di matrice A, quindi le colonne di A con gli stessi indici sono anch'esse una base dell'immagine (essendo le coordinate degli stessi vettori rispetto ad un'altra base).

Inoltre, matrici equivalenti hanno lo stesso rango, perché il rango di un prodotto è il minimo tra i ranghi dei fattori e i fattori invertibili hanno rango maggiore o uguale a tale minimo. Per una matrice a scala ridotta, il rango della matrice è il numero dei pivot ed è uguale al numero di righe non nulle (e necessariamente linearmente indipendenti, data la posizione dei pivot). Per cui il rango di tale matrice è uguale al rango della sua trasposta e lo stesso vale per ogni matrice, essendo equivalente a una matrice a scala. Ovvero, $\boxed{\text{per ogni } A \in M_{m\times n}(K) \text{ si ha } \operatorname{rk} A = \operatorname{rk} {}^tA}$.

Possiamo essere ancora più precisi nella descrizione delle classi per la relazione di equivalenza tra matrici. Ovvero, due matrici a elementi in un campo sono equivalenti se, e solo se, hanno lo stesso rango. La dimostrazione la scriviamo nella prossima slide "traducendo", come al solito, l'affermazione nel linguaggio delle applicazioni lineari.

Proposizione

Sia K un campo e $A \in M_{m\times n}(K)$. Allora A ha rango r se, e solo se, è equivalente alla matrice (a blocchi) $D = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix}$.

dim. Se esistono due matrici $P \in \mathrm{GL}(m, K)$ e $Q \in \mathrm{GL}(n, K)$ tali che $A = PDQ$, allora, $\mathrm{rk}\, P = m$, $\mathrm{rk}\, Q = n$ e $\mathrm{rk}\, A$ è il minimo tra i ranghi dei fattori, ovvero

$$\mathrm{rk}\, A = r = \mathrm{rk} \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} \le \min \{ m, n \} .$$

Viceversa, sia $\mathrm{rk}\, A = r$ e sia $\phi : K^n \to K^m$ l'applicazione lineare di matrice A nelle basi canoniche. Per la formula delle dimensioni, $\dim \ker \phi = n - \dim \mathrm{im}\, \phi = n - r$, e sia $v_{r+1}, \dots, v_n$ una base di $\ker \phi$, che completiamo a una base $\mathcal{V} = \{v_1, \dots, v_n\}$ di K^n. Nella dimostrazione della citata formula delle dimensioni, abbiamo visto che i vettori $w_i = \phi(v_i)$, per $i = 1, \dots, r$ sono una base di $\mathrm{im}\, \phi$ e possiamo completare questi vettori a una base $\mathcal{W} = \{w_1, \dots, w_m\}$ di K^m. Per costruzione

$$D = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} = \alpha_{\mathcal{V},\mathcal{W}}(\phi) \qquad \text{e} \qquad A = \alpha_{\mathcal{E}_n,\mathcal{E}_m}(\phi) = \alpha_{\mathcal{W},\mathcal{E}_m}(\mathrm{id}_{K^m})\alpha_{\mathcal{V},\mathcal{W}}(\phi)\alpha_{\mathcal{E}_n,\mathcal{V}}(\mathrm{id}_{K^n}).$$

Le matrici di cambiamento di base $P = \alpha_{\mathcal{W},\mathcal{E}_m}(\mathrm{id}_{K^m})$ e $Q = \alpha_{\mathcal{E}_n,\mathcal{V}}(\mathrm{id}_{K^n})$ sono invertibili e perciò A è equivalente a D e si conclude la dimostrazione. $\square$

Dalla Proposizione precedente si può ottenere un'altra dimostrazione del fatto che A e ${}^tA$ hanno ambedue rango r (fatto evidente per D e ${}^tD$).

Data una matrice A, possiamo usare il procedimento descritto nella dimostrazione della Proposizione per trovare due matrici invertibili che la portino alla forma canonica per equivalenza. Facciamo vedere su un esempio come si possa usare la tecnica di eliminazione di Gauss anche per affrontare questo problema.

Esempio: Sia $\phi : \mathbb{Q}^7 \to \mathbb{Q}^5$ l'applicazione lineare di matrice A nelle basi canoniche, ove

$$\alpha_{\mathcal{E}_7,\mathcal{E}_5}(\phi) = A = \begin{pmatrix} 2 & -6 & 1 & 0 & 2 & -1 & 3 \\ -1 & 3 & 3 & -14 & -1 & 0 & 10 \\ 0 & 0 & 1 & -4 & 0 & 1 & 1 \\ 1 & -3 & -1 & 6 & 1 & 0 & -4 \\ 1 & -3 & 0 & 2 & 1 & -3 & 5 \end{pmatrix} \in M_{5\times 7}(\mathbb{Q}).$$

(a) Si determinino la dimensione e una base per il nucleo e l'immagine di ϕ.

(b) Si determinino due matrici invertibili, P e Q, tali che $A = P \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} Q$, ove $r = \operatorname{rk}\phi$.

Svolg.: (*a*) Portiamo A in forma a scala ridotta.

$$A \sim \begin{array}{r} V \\ III \\ I-2V \\ II+V \\ IV-V \end{array} \begin{pmatrix} 1 & -3 & 0 & 2 & 1 & -3 & 5 \\ 0 & 0 & 1 & -4 & 0 & 1 & 1 \\ 0 & 0 & 1 & -4 & 0 & 5 & -7 \\ 0 & 0 & 3 & -12 & 0 & -3 & 15 \\ 0 & 0 & -1 & 4 & 0 & 3 & -9 \end{pmatrix} \sim \begin{array}{r} I \\ II \\ (III-II)/4 \\ IV-3II \\ V+II \end{array} \begin{pmatrix} 1 & -3 & 0 & 2 & 1 & -3 & 5 \\ 0 & 0 & 1 & -4 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & -6 & 12 \\ 0 & 0 & 0 & 0 & 0 & 4 & -8 \end{pmatrix}$$

$$\sim \begin{array}{r} I+3III \\ II-III \\ III \\ IV+6III \\ V-4III \end{array} \begin{pmatrix} 1 & -3 & 0 & 2 & 1 & 0 & -1 \\ 0 & 0 & 1 & -4 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

[segue]

Dalla matrice in forma a scala ridotta si vede che $\phi(e_1), \phi(e_3), \phi(e_6)$ sono una base di $\operatorname{im}\phi$; ovvero

$$\operatorname{im}\phi = \left\langle \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ -3 \end{pmatrix} \right\rangle \quad \text{e i generatori sono una base.}$$

Inoltre, $\phi(e_2) = -3\phi(e_1)$, $\phi(e_4) = 2\phi(e_1) - 4\phi(e_3)$, $\phi(e_5) = \phi(e_1)$ e $\phi(e_7) = -\phi(e_1) + 3\phi(e_3) - 2\phi(e_6)$; per cui $\dim \ker\phi = 7 - 3 = 4$ e

$$\ker\phi = \langle 3e_1 + e_2, 2e_1 - 4e_3 - e_4, e_1 - e_5, e_1 - 3e_3 + 2e_6 + e_7 \rangle\,.$$

(*b*) Le colonne della matrice in forma a scala ridotta (e quindi le colonne di A) sono combinazione lineare delle tre colonne contenenti i pivot. Quindi, prese le colonne di A corrispondenti ai pivot, si ha

$$A = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 3 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -3 \end{pmatrix} \begin{pmatrix} 1 & -3 & 0 & 2 & 1 & 0 & -1 \\ 0 & 0 & 1 & -4 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 3 & 0 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -3 & 0 & 2 & 1 & 0 & -1 \\ 0 & 0 & 1 & -4 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix}.$$

La forma canonica per equivalenza di A è $D = \begin{pmatrix} \mathbf{1}_3 & 0 \\ 0 & 0 \end{pmatrix} \in M_{5\times 7}(\mathbb{Q})$. Se aggiungiamo colonne al fattore di sinistra del prodotto qui sopra per poterlo moltiplicare per D, non cambia il prodotto, avendo aggiunto due righe nulle al di sotto di $\mathbf{1}_3$ per ottenere D. Allo stesso modo possiamo aggiungere righe al fattore di destra senza cambiare il prodotto. Dobbiamo quindi aggiungere righe e colonne ai due fattori in modo di ottenere matrici invertibili. Ad esempio, possiamo prendere

$$P = \begin{pmatrix} 2 & 1 & -1 & 0 & 1 \\ -1 & 3 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -3 & 0 & 0 \end{pmatrix} \quad \text{e} \quad Q = \begin{pmatrix} 1 & -3 & 0 & 2 & 1 & 0 & -1 \\ 0 & 0 & 1 & -4 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

e $A = PDQ$. □