Algebra lineare e geometria AA. 2019-2020

Appunti sul cambio di base in uno spazio vettoriale

1 Matrice di un'applicazione lineare

Siano V e W due spazi vettoriali su un campo $K \in \{\mathbb{R}, \mathbb{C}\}$, entrambi finitamente generati, $\dim_K(V) = n$, $\dim_K(W) = m$, con basi $\mathcal{B} = (v_1, \dots, v_n)$ e $\mathcal{D} = (w_1, \dots, w_m)$ rispettivamente. Sia

$$f: V \to W$$

un'applicazione lineare. La matrice associata ad f rispetto alle basi \mathcal{B} e \mathcal{D} si denota $M_{\mathcal{D}}^{\mathcal{B}}(f)$ ed è la matrice $m \times n$ la cui j-esima colonna è costituita dalle coordinate del vettore $f(v_j) \in W$ rispetto alla base \mathcal{D} , cioè $[f(v_j)]_{\mathcal{D}}$.

Esplicitamente, se $f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \ldots + a_{mj}w_m$, allora:

$$M_{\mathcal{D}}^{\mathcal{B}}(f) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Una volta definita la matrice associata ad f, dato un qualsiasi $v = x_1v_1 + \ldots + x_nv_n \in V$, se $f(v) = y_1w_1 + \ldots + y_mw_m \in W$, allora:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = M_{\mathcal{D}}^{\mathcal{B}}(f) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

ovvero

$$[f(v)]_{\mathcal{D}} = M_{\mathcal{D}}^{\mathcal{B}}(f)[v]_{\mathcal{B}}.$$

Se U è un altro spazio vettoriale su K, finitamente generato di $\dim_K(U) = s$ e con base $\mathcal{U} = (u_1, \dots, u_s)$, e

$$q:U\to V$$

è un'applicazione lineare componibile con f, in modo che sia ben definita l'applicazione

$$U \xrightarrow{g} V \xrightarrow{f \circ g} W$$
,

allora:

$$M_{\mathcal{D}}^{\mathcal{U}}(f \circ g) = M_{\mathcal{D}}^{\mathcal{B}}(f)M_{\mathcal{B}}^{\mathcal{U}}(g).$$

Come conseguenza, se $f:V\to V$ è un endomorfismo invertibile, e $\mathcal B$ e $\mathcal D$ sono due basi di V, allora:

$$M_{\mathcal{B}}^{\mathcal{D}}(f^{-1}) = (M_{\mathcal{D}}^{\mathcal{B}}(f))^{-1}$$

Inoltre se l'applicazione f coincide con la moltiplicazione a sinistra per una matrice A, cioè se f è del tipo μ_A , allora $M_D^B(\mu_A) = A$, esattamente come ci aspettiamo.

Esempio 1. Siano $\mathcal{B} = (v_1, v_2, v_3, v_4)$ e $\mathcal{D} = (w_1, w_2, w_3)$ basi di \mathbb{R}^4 ed \mathbb{R}^3 rispettivamente, e sia $f : \mathbb{R}^4 \to \mathbb{R}^3$ l'applicazione lineare definita dalle equazioni:

$$f(v_1) = 2w_1 - w_2 + 5w_3$$

$$f(v_2) = -w_2 - w_3$$

$$f(v_3) = w_1 + 3w_2 + 2w_3$$

$$f(v_4) = -w_1 + w_2$$

Determiniamo la matrice associata $M_{\mathcal{D}}^{\mathcal{B}}(f)$.

Per definizione, la matrice $M_{\mathcal{D}}^{\mathcal{B}}(f)$ è la matrice 3×4 le cui colonne sono le coordinate dei vettori $f(v_j)$ rispetto ai w_i , che abbiamo già esplicitamente. Infatti $f(v_1) = 2w_1 - w_2 + 5w_3$ significa che le coordinate di $f(v_1)$ nella base \mathcal{D} sono $[f(v_1)]_{\mathcal{D}} = (2, -1, 5)$, e così per gli altri 3 vettori della base \mathcal{B} . Quindi:

$$M_{\mathcal{D}}^{\mathcal{B}}(f) = \begin{pmatrix} 2 & 0 & 1 & -1 \\ -1 & -1 & 3 & 1 \\ 5 & -1 & 2 & 0 \end{pmatrix}$$

2 Matrice del cambiamento di coordinate (o del cambiamento di base)

Un caso particolarmente importante di matrice associata ad un'applicazione lineare si ha quando abbiamo due basi distinte $\mathcal{B} = (v_1, \ldots, v_n)$ e $\mathcal{D} = (w_1, \ldots, w_n)$ di uno stesso spazio vettoriale V, e $f: V \to V$ è la funzione identità $f = \mathrm{id}_V$. In questo caso la matrice $M_{\mathcal{D}}^{\mathcal{B}}(id_V)$ si dice matrice del cambiamento di coordinate dalla base \mathcal{B} alla base \mathcal{D} . Per definizione, $M_{\mathcal{D}}^{\mathcal{B}}(id_V)$ è la matrice la cui colonna j-esima è costituita dalle coordinate del vettore $id_V(v_j) = v_j$ rispetto alla base \mathcal{D} , cioè $[v_j]_{\mathcal{D}}$.

Esempio 2. Riprendiamo l'esempio di prima, e supponiamo di sapere le coordinate dei vettori della base \mathcal{D} di \mathbb{R}^3 rispetto alla base canonica \mathcal{C} :

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}, \quad w_1 = \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}$$

Scriviamo le due matrici di cambiamento di base: $M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3})$ che porta dalla base \mathcal{D} alla base canonica \mathcal{C} , e $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3})$ che viceversa porta da \mathcal{C} a \mathcal{D} . La prima l'abbiamo praticamente già scritta; infatti $M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3})$ è la matrice le cui colonne sono le coordinate dei w_j rispetto alla base canonica, cioè:

$$M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}) = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & 5 \end{pmatrix}$$

Per determinare $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3})$ possiamo procedere in due modi, sostanzialmente equivalenti: possiamo esplicitare gli e_i rispetto ai w_j , cioè trovare esplicitamente le coordinate $[e_j]_{\mathcal{D}}$, che formeranno le colonne di $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3})$. Oppure possiamo ricordarci che $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3}) = (M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}))^{-1}$, e invertire la matrice sopra, ottenendo:

$$M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3}) = (M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}))^{-1} = \frac{1}{3} \begin{pmatrix} -2 & 5 & -1 \\ -5 & 5 & -1 \\ 2 & -2 & 1 \end{pmatrix}$$

3 Mettiamo tutto insieme

Torniamo al caso generale dei due spazi vettoriali V e W con basi \mathcal{B} e \mathcal{D} rispettivamente, e supponiamo di avere una applicazione lineare $f:(V,\mathcal{B})\to (W,\mathcal{D})$, dove con questa scrittura si intende che f manda un elemento dello spazio V con la base \mathcal{B} in un elemento dello spazio V con la base V. Cosa succede alla matrice associata ad V se cambio base in V? e se la cambio in V? Per esprimere la matrice associata ad V rispetto alle nuove basi, ad esempio V0 e V1, procediamo componendo varie applicazioni, come nella sequenza:

$$M_{\mathcal{F}}^{\mathcal{E}}(f): (V, \mathcal{E}) \xrightarrow{M_{\mathcal{B}}^{\mathcal{E}}(id_V)} (V, \mathcal{B}) \xrightarrow{M_{\mathcal{D}}^{\mathcal{B}}(f)} (W, \mathcal{D}) \xrightarrow{M_{\mathcal{F}}^{\mathcal{D}}(id_W)} (W, \mathcal{F})$$

o equivalentemente nel diagramma:

$$(V, \mathcal{B}) \xrightarrow{M_{\mathcal{D}}^{\mathcal{B}}(f)} (W, \mathcal{D})$$

$$\downarrow^{M_{\mathcal{B}}^{\mathcal{E}}(id_{V})} \qquad \qquad \downarrow^{M_{\mathcal{F}}^{\mathcal{D}}(id_{W})}$$

$$(V, \mathcal{E}) \xrightarrow{M_{\mathcal{F}}^{\mathcal{E}}(f)} (W, \mathcal{F})$$

Ora basta ricordarsi che la composizione delle applicazioni corrisponde al prodotto tra matrici, e siamo a posto:

$$M_{\mathcal{F}}^{\mathcal{E}}(f) = M_{\mathcal{F}}^{\mathcal{D}}(id_W) \ M_{\mathcal{D}}^{\mathcal{B}}(f) \ M_{\mathcal{B}}^{\mathcal{E}}(id_V)$$

<u>Esempio 3.</u> Continuiamo con lo stesso esempio di prima. Nell'*Esempio 1* abbiamo calcolato la matrice $M_D^B(f)$ associata all'applicazione lineare

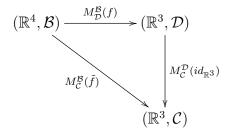
$$f: (\mathbb{R}^4, \mathcal{B}) \to (\mathbb{R}^3, \mathcal{D}).$$

Nell'*Esempio 2* invece abbiamo calcolato le matrici di cambiamento di base in \mathbb{R}^3 $M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3})$ e $M_{\mathcal{D}}^{\mathcal{C}}(id_{\mathbb{R}^3})$.

Adesso calcoliamo la matrice $M_{\mathcal{C}}^{\mathcal{B}}(\tilde{f})$ associata all'applicazione

$$\tilde{f}: (\mathbb{R}^4, \mathcal{B}) \to (\mathbb{R}^3, \mathcal{C}).$$

Il diagramma in questo caso è un triangolo invece che un quadrato, perchè facciamo un cambio di base solo nello spazio di arrivo:



Di nuovo, la composizione delle applicazioni corrisponde al prodotto tra matrici, e quindi:

$$M_{\mathcal{C}}^{\mathcal{B}}(\tilde{f}) = M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}) \ M_{\mathcal{D}}^{\mathcal{B}}(f) = \begin{pmatrix} 3 & 1 & -2 & -2 \\ 7 & -1 & 3 & -1 \\ 23 & -7 & 16 & 2 \end{pmatrix}$$

Se per caso passare dal diagramma quadrato al diagramma triangolare vi fa confusione, potete sempre guardare il secondo come un quadrato dove un lato è dato dalla matrice identità, visto che su \mathbb{R}^4 non facciamo nessun cambio di base:

$$(\mathbb{R}^{4},\mathcal{B}) \xrightarrow{M_{\mathcal{D}}^{\mathcal{B}}(f)} (\mathbb{R}^{3},\mathcal{D}) \longleftrightarrow (\mathbb{R}^{4},\mathcal{B}) \xrightarrow{M_{\mathcal{D}}^{\mathcal{B}}(f)} (\mathbb{R}^{3},\mathcal{D})$$

$$\downarrow^{M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^{3}})} (\mathbb{R}^{3},\mathcal{C}) \longleftrightarrow (\mathbb{R}^{4},\mathcal{B}) \xrightarrow{M_{\mathcal{D}}^{\mathcal{B}}(id_{\mathbb{R}^{4}}) = I_{4}} (\mathbb{R}^{3},\mathcal{C})$$

Ovviamente $M_{\mathcal{C}}^{\mathcal{B}}(\tilde{f}) = M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}) \ M_{\mathcal{D}}^{\mathcal{B}}(f) = M_{\mathcal{C}}^{\mathcal{D}}(id_{\mathbb{R}^3}) \ M_{\mathcal{D}}^{\mathcal{B}}(f) \ I_4.$

4 Un caso speciale: la diagonalizzazione

Il processo di diagonalizzazione di una matrice (quando possibile) è esattamente lo stesso visto sopra. Partiamo da una matrice $A \in K^{n,n}$, o se preferite, dall'applicazione lineare

$$f = \mu_A : K^n \to K^n$$
,

e quindi usando le notazioni sopra: $V \cong W \cong K^n$, le basi \mathcal{B} e \mathcal{D} sono entrambe uguali alla base canonica \mathcal{C} , e infine la matrice $M_{\mathcal{C}}^{\mathcal{C}}(f) = A$. Cerchiamo una base \mathcal{B} tale che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ sia una matrice diagonale D.

Supponiamo che la matrice A sia diagonalizzabile, e che quindi si riesca a trovare una base $\mathcal{B} = (v_1, \ldots, v_n)$ di autovettori relativi agli autovalori $\lambda_1, \ldots, \lambda_n$ (non necessariamente distinti). Se applichiamo il cambio di base il diagramma sopra ora si scrive così:

$$(K^{n}, \mathcal{C}) \xrightarrow{A} (K^{n}, \mathcal{C})$$

$$\downarrow^{M_{\mathcal{C}}^{\mathcal{C}}(id_{K^{n}})} \qquad \qquad \downarrow^{M_{\mathcal{B}}^{\mathcal{C}}(id_{K^{n}})}$$

$$(K^{n}, \mathcal{B}) \xrightarrow{D} (K^{n}, \mathcal{B})$$

Per definizione, la matrice $M_{\mathcal{C}}^{\mathcal{B}}(id_{K^n})$ è semplicemente la matrice le cui colonne sono gli autovettori v_j , mentre la matrice $M_{\mathcal{B}}^{\mathcal{C}}(id_{K^n})$ è la sua inversa: $M_{\mathcal{B}}^{\mathcal{C}}(id_{K^n}) = (M_{\mathcal{C}}^{\mathcal{B}}(id_{K^n}))^{-1}$.

Se chiamiamo $P=M_{\mathcal{C}}^{\mathcal{B}}(id_{K^n})$ allora il diagramma diventa ancora più semplice:

$$(K^{n}, \mathcal{C}) \xrightarrow{A} (K^{n}, \mathcal{C})$$

$$\downarrow^{P} \qquad \qquad \downarrow^{P-1}$$

$$(K^{n}, \mathcal{B}) \xrightarrow{D} (K^{n}, \mathcal{B})$$

e quindi

$$D = P^{-1}AP.$$

Osserviamo che

$$D = P^{-1}AP \quad \Leftrightarrow \quad PD = AP \quad \Leftrightarrow \quad PDP^{-1} = A$$

Esempio 4. Diagonalizziamo la matrice

$$A = \left(\begin{array}{ccc} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{array}\right).$$

Il suo polinomio caratteristico è (sviluppando rispetto alla seconda riga):

$$p_A(t) = \det(A - tI_3) = \det\begin{pmatrix} 1 - t & 0 & 4 \\ 0 & 2 - t & 0 \\ 1 & 0 & 1 - t \end{pmatrix} = (2 - t)[(1 - t)^2 - 4],$$

che ha tre radici distinte $\lambda_1=2,\ \lambda_2=3,\ \lambda_3=-1,$ quindi la matrice A è sicuramente diagonalizzabile.

Troviamo una base di autovettori ed effettuiamo il cambio di base. I tre autospazi sono:

$$E_{A}(2) = \operatorname{Ker}(A - 2I_{3}) = \operatorname{Ker}\begin{pmatrix} -1 & 0 & 4 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix} = \mathcal{L}((0, 1, 0)),$$

$$E_{A}(3) = \operatorname{Ker}(A - 3I_{3}) = \operatorname{Ker}\begin{pmatrix} -2 & 0 & 4 \\ 0 & -1 & 0 \\ 1 & 0 & -2 \end{pmatrix} = \mathcal{L}((2, 0, 1)),$$

$$E_{A}(-1) = \operatorname{Ker}(A + I_{3}) = \operatorname{Ker}\begin{pmatrix} 2 & 0 & 4 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix} = \mathcal{L}((-2, 0, 1)).$$

Per concludere, verifichiamo che:

$$P^{-1}AP = \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} \\ -\frac{1}{4} & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix} = D \qquad \checkmark$$

5 Un caso ancora più speciale: le matrici simmetriche

Se dobbiamo diagonalizzare una matrice simmetrica il Teorema Spettrale ci garantisce che non solo possiamo sempre farlo, ma che è possibile trovare una matrice diagonalizzante P ortogonale, cioè tale che $P^{-1} = {}^{t}P$. Questo semplifica ulteriormente il nostro compito, visto che non dobbiamo nemmeno fare la fatica di invertire P.

Attenzione però che se avete autovettori relativi ad autovalori distinti (di una matrice simmetrica) essi saranno automaticamente ortogonali, come abbiamo visto in classe; se però avete un autovalore con molteplicità geometrica > 1 allora dovete stare attenti a prendere una base ortogonale dell'autospazio. Solitamente si riesce a fare "a occhio"; altrimenti si usa l'algoritmo di Gram-Schmidt.

Esempio 5. Diagonalizziamo la matrice simmetrica

$$A = \left(\begin{array}{ccc} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{array}\right)$$

e troviamone una base diagonalizzante ortonormale.

Il suo polinomio caratteristico è:

$$p_A(t) = \det(A - tI_3) = \det\begin{pmatrix} 1 - t & 2 & 2 \\ 2 & 1 - t & 2 \\ 2 & 2 & 1 - t \end{pmatrix} = \dots = -(t - 5)(t + 1)^2,$$

che ha due radici distinte $\lambda_1 = 5$, e $\lambda_2 = -1$, la seconda con molteplicità algebrica 2 (ma questo non ci preoccupa perchè per il Teorema Spettrale sappiamo già che anche la sua molteplicità geometrica è 2).

Troviamo una base ortonormale di autovettori ed effettuiamo il cambio di base. I due autospazi sono:

$$E_A(5) = \text{Ker}(A - 5I_3) = \text{Ker}\begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} = \mathcal{L}((1, 1, 1)),$$

$$E_A(-1) = \text{Ker}(A + I_3) = \text{Ker}\begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} = \mathcal{L}((-1, 1, 0), (-1, 0, 1)).$$

Come previsto, abbiamo 3 autovettori indipendenti che formano una base. Ancora però non abbiamo una base ortonormale! L'unica cosa che abbiamo "gratis" è che l'autovettore v_1 relativo all'autovalore $\lambda_1 = 5$ è ortogonale a entrambi gli autovettori v_2 e v_3 dell'autovalore $\lambda_2 = -1$. Trasformiamo quindi la base $\mathcal{B} = (v_1, v_2, v_3)$ in una base ortonormale $\mathcal{B}' = (w_1, w_2, w_3)$.

Per w_1 ci è sufficiente normalizzare v_1 , quindi

$$w_1 = \frac{v_1}{|v_1|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

Per gli altri due applichiamo l'algoritmo di Gram-Schmidt alla base (v_1, v_2) di $E_A(-1)$:

$$u_2 = v_2 = (-1, 1, 0),$$

 $w_3 = v_3 - \frac{v_3 \cdot u_2}{|u_2|^2} u_2 = v_3 - \frac{1}{2} u_2 = \left(-\frac{1}{2}, -\frac{1}{2}, 1\right),$

da cui finalmente ricaviamo

$$w_2 = \frac{u_2}{|u_2|} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right),$$

$$w_3 = \frac{u_3}{|u_3|} = \left(-\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right).$$

Per concludere, verifichiamo che $P^{-1}AP={}^tPAP=D$: si ha infatti che

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

N.B. Se qualcosa non vi è chiaro, e/o se avete trovato un errore di stampa, fatemi sapere!