7.5 Forma canonica per Applicazioni Lineari

Il prossimo teorema ci dice che, purché le basi del dominio e del codominio siano scelte in maniera opportuna, la matrice di un'applicazione lineare può assumere una forma particolarmente semplice, detta " forma canonica".

Teorema 7.5.1. *Sia $f : V \to W$ un'applicazione lineare fra $\mathbb{K}$-spazi vettoriali di dimensioni finite n, m, sia $r = \operatorname{rg} f = \dim \operatorname{Im} f$ il rango di f. Esistono basi $\mathcal{B}$ di V e $\mathcal{B}'$ di W tali che $M_{\mathcal{B}}^{\mathcal{B}'}(f)$ è una matrice a blocchi della forma*

$$\begin{array}{cc} & \begin{array}{ccc} \quad r \quad & & \quad n-r \quad \end{array} \\ \begin{array}{c} r \\ \\ m-r \end{array} & \left(\begin{array}{ccc} \mathbb{I}_r & | & 0 \\ - - - & | & - - - \\ 0 & | & 0 \end{array}\right) \end{array} \tag{7.4}$$

dove $\mathbb{I}_r$ è la matrice unità $r \times r$, e gli zeri denotano delle matrici nulle rispettivamente di tipo $r \times (n - r),\ (m-r) \times r,\ (m-r) \times (n-r)$.

Dimostrazione. Cerchiamo di capire quali proprietà devono avere le due basi perchè la matrice di f rispetto a tali basi abbia la forma voluta.

Se la matrice $M_{\mathcal{B}}^{\mathcal{B}'}(f)$ ha la forma (7.4) rispetto a basi $\mathcal{B} = (v_1, \dots, v_n)$, $\mathcal{B}' = (w_1, \dots, w_m)$ significa che

$$\begin{aligned} f(v_1) &= w_1 &&= 1w_1 + 0w_2 + \cdots + 0w_m \\ &\vdots \\ f(v_r) &= w_r &&= 0w_1 + \cdots + 1w_r + \cdots + 0w_m \\ f(v_{r+1}) &= 0 \\ &\vdots \\ f(v_n) &= 0, \end{aligned}$$

perciò $v_{r+1}, \dots, v_n \in \ker f$, e sono linearmente indipendenti perchè fanno parte della base $\mathcal{B}$ di V. Siccome $r = \operatorname{rg} f$ e $\dim V = n$, per il Teorema della dimensione **??** $\dim \ker f = n - r$, quindi $v_{r+1}, \dots, v_n$ **formano una base di** $\ker f$. D'altra parte **i vettori** $w_1, \dots, w_r$ **generano** $\operatorname{Im} f$, perchè sono le immagini non nulle dei vettori di una base del dominio (Proposizione **??**) e quindi ne formano una base, perchè sono in numero di r. Ora abbiamo gli elementi necessari per costruire le basi volute.

Partiamo da una base di $\ker f$: $v_{r+1}, \dots, v_n$, e la completiamo a una base $\mathcal{B} = (v_1, \dots, v_n)$ di V. Allora $f(v_1), \dots, f(v_n)$ generano $\operatorname{Im} f$. Siccome $f(v_{r+1}) = \cdots = f(v_n) = 0$, non contribuiscono a generare $\operatorname{Im} f$, allora gli r vettori $f(v_1), \dots, f(v_r)$ sono un sistema di r generatori, e quindi una base di $\operatorname{Im} f$. Pongo $w_1 = f(v_1), \dots, w_r = f(v_r)$, e poi li completo a una base $\mathcal{B}'$ di W aggiungendo opportuni vettori $w_{r+1}, \dots, w_m$. Rispetto a $\mathcal{B}$ e $\mathcal{B}'$ la matrice di f ha la forma (7.4). □

7.6 Spazio vettoriale duale.

Sia V un $\mathbb{K}$-spazio vettoriale di dimensione finita n e $\mathcal{B} = (v_1, \dots, v_n)$ una sua base. Definiamo lo **spazio vettoriale duale**

$$V^* = \{f : V \to \mathbb{K} \mid f \text{ lineare}\} = \mathrm{Hom}(V, \mathbb{K}),$$

i cui elementi sono le **forme lineari** su V. Per ogni $i = 1, \dots, n$ definiamo la forma lineare $v_i^* : V \to K$ ponendo $v_i^*(v_j) = \delta_{ij}$, il simbolo di Kronecker, ossia

$$v_i^*(v_j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Per il Teorema di Struttura per Applicazioni Lineari, v_i^* lineare risulta determinata. Se $v = x_1v_1 + \cdots + x_nv_n$, si ha $v_i^*(v) = x_1\delta_{i1} + \cdots + x_n\delta_{ni} = x_i$: è la i-esima coordinata di v rispetto a $\mathcal{B}$. Per questo motivo le funzioni $v_1^*, \dots, v_n^*$ sono dette **funzioni coordinate rispetto alla base** $\mathcal{B}$.

Teorema 7.6.1 (Base duale)**.** *Le forme lineari $v_1^*, \dots, v_n^*$ formano una base di V^*, detta* **base duale di** *$\mathcal{B}$ e denotata $\mathcal{B}^*$.*

Dimostrazione. Proviamo che $v_1^*, \dots, v_n^*$ sono linearmente indipendenti. Supponiamo che $\lambda_1v_1^* + \cdots + \lambda_nv_n^* = 0$: ciò significa che per ogni $v \in V$ si ha $(\lambda_1v_1^* + \cdots + \lambda_nv_n^*)(v) = 0$. Questo vale in particolare per ogni vettore di $\mathcal{B}$, dunque per ogni i si ha

$$\begin{aligned} 0 = (\lambda_1v_1^* + \cdots + \lambda_nv_n^*)(v_i) &= \text{usando la definizione delle operazioni in } V^* \\ &= \lambda_1v_1^*(v_i) + \cdots + \lambda_nv_n^*(v_i) = \lambda_1\delta_{1i} + \cdots + \lambda_n\delta_{ni} = \lambda_i. \end{aligned}$$

Quindi $\lambda_1 = \cdots = \lambda_n = 0$.

Proviamo ora che $v_1^*, \dots, v_n^*$ generano V^*. Sia $f : V \to K$ una forma lineare; il nostro scopo è trovare degli scalari $c_1, \dots, c_n$ tali che $f = c_1v_1^* + \cdots + c_nv_n^*$. Cerchiamo di capire quali possono essere. Se esistono scalari per cui una tale uguaglianza vale, le due applicazioni f e $c_1v_1^* + \cdots + c_nv_n^*$ devono coincidere su ogni vettore di V, e in particolare sui vettori di $\mathcal{B}$, cioè dev'essere $f(v_1) = (c_1v_1^* + \cdots + c_nv_n^*)(v_1) = c_1, \dots, f(v_n) = (c_1v_1^* + \cdots + c_nv_n^*)(v_n) = c_n$. Quindi l'unica possibilità è che $c_1 = f(v_1), \dots, c_n = f(v_n)$, ossia che $f = f(v_1)v_1^* + \cdots + f(v_n)v_n^*$. E in effetti questa uguaglianza di forme lineari è vera per il Teorema di Struttura, poiché, per ogni i, $f(v_i) = (f(v_1)v_1^* + \cdots + f(v_n)v_n^*)(v_i)$: assumono lo stesso valore sui vettori di $\mathcal{B}$. $\square$

Osservazione 31. Abbiamo dunque dimostrato che $f = f(v_1)v_1^* + \cdots + f(v_n)v_n^*$. Osserviamo che vale anche un'uguaglianza "duale": per ogni vettore $v \in V$ si ha $v = v_1^*(v)v_1 + \cdots + v_n^*(v)v_n$. In altre parole, data la base $\mathcal{B}$ di V e la base duale $\mathcal{B}^*$ di V^*, le coordinate di f rispetto a $\mathcal{B}^*$ sono $f(v_1), \dots, f(v_n)$, i valori assunti da f sui vettori di $\mathcal{B}$, mentre le coordinate di v rispetto a $\mathcal{B}$ sono $v_1^*(v), \dots v_n^*(v)$, i valori assunti su v dai "vettori" di $\mathcal{B}^*$.

Capitolo 8

Determinante

In questo capitolo introdurremo la nozione di determinante, su uno spazio vettoriale di dimensione finita o per matrici. Premettiamo alcune proprietà dei gruppi di permutazioni che ci saranno utili per dare le definizioni.

8.1 Gruppi di permutazioni

Sia $I_n = \{1, 2, \dots, n\}$ l'insieme dei primi n numeri naturali.

Definizione 8.1.1. Una **permutazione** dell'insieme I_n è una biiezione

$$\begin{array}{rll} \sigma: & I_n \to & I_n \\ & 1 \to & \sigma(1) \\ & 2 \to & \sigma(2) \\ & \vdots & \\ & n \to & \sigma(n) \end{array}$$

Per esempio, per $n = 3$, $I_3 = \{1, 2, 3\}$ ha sei permutazioni:

$$1\,2\,3,\ 1\,3\,2,\ 2\,1\,3,\ 2\,3\,1,\ 3\,1\,2,\ 3\,2\,1.$$

Denoteremo $\mathcal{S}_n$ l'insieme delle permutazioni di I_n: ha $n!$ elementi, in quanto ci sono n scelte per $\sigma(1)$, $n - 1$ scelte per $\sigma(2)$, ecc.

Osserviamo che $\mathcal{S}_n$ è un gruppo rispetto alla composizione di applicazioni, e che non è abeliano per $n \geq 3$. È detto **gruppo simmetrico su** n **oggetti**.

Esempi 8.1.2.

- Per $n = 1$ $\mathcal{S}_1 = \{*\}$ si riduce a un elemento;
- Per $n = 2$, il gruppo $\mathcal{S}_2$ ha due elementi: $1\,2,\ 2\,1$; è isomorfo a $\mathbb{Z}_2$.

Introduciamo la seguente notazione per indicare come opera una permutazione σ:

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

Per esempio

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},\ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix};\ \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},\ \tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$

che conferma che $\mathcal{S}_3$ non è abeliano.

8.2 Cicli

Definizione 8.2.1. Un **ciclo o permutazione ciclica** di lunghezza k in $\mathcal{S}_n$ è una permutazione che verifica

$$n_1 \to n_2 \to \cdots \to n_k \to n_1,$$

dove $n_1, n_2, \ldots, n_k$ sono k elementi distinti di I_n, mentre gli altri elementi di I_n vengono lasciati fissi. Lo si denota $(n_1\ n_2\ \ldots n_k)$.

Vediamo alcuni esempi.

Esempi 8.2.2.

1. Sia $\tau \in \mathcal{S}_7$ la permutazione

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 6 & 2 & 1 & 7 & 3 & 4 \end{pmatrix}$$

   Osserviamo che $1 \to 5 \to 7 \to 4 \to 1$, e $2 \to 6 \to 3 \to 2$: il primo è un **ciclo o permutazione ciclica di lunghezza** 4 e il secondo un ciclo di lunghezza 3. Si usa la notazione $\tau = (1\,5\,7\,4)(2\,6\,3)$.

2. Sia $\sigma \in \mathcal{S}_7$ la seguente permutazione:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 3 & 1 & 2 & 7 & 6 & 4 \end{pmatrix}$$

   Osserviamo che $1 \to 5 \to 7 \to 4 \to 2 \to 3 \to 1$: questo è un **ciclo di lunghezza** 6, invece 6 rimane fisso in σ, è un ciclo di lunghezza 1. Si scrive $\sigma = (1\,5\,7\,4\,2\,3)$, ossia l'elemento che rimane fisso non si scrive.

3.
$$\sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 1 & 5 & 4 & 7 & 3 & 6 \end{pmatrix}$$

Questa volta $1 \to 2 \to 1$: ho un ciclo di lunghezza 2 $(1\ 2)$; $3 \to 5 \to 7 \to 6 \to 3$: un ciclo di lunghezza 4 $(3\ 5\ 7\ 6)$; $4 \to 4$: ciclo di lunghezza 1. La permutazione data σ' si indica $(1\ 2)(3\ 5\ 7\ 6)$: è prodotto di due cicli disgiunti di lunghezza 2 e 4: sono permutabili.

Per la composizione di permutazioni si usa una notazione diversa rispetto a quella usata di solito per la composizione di applicazioni: anzichè scrivere per esempio $\sigma \circ \tau$ si usa $\tau\sigma$, per intendere che si applica prima τ e poi σ. Si parla di **prodotto di permutazioni**.

Per esempio: $(1\ 3\ 5)(3\ 1\ 7\ 4)(6\ 7\ 2)$ è un prodotto di tre cicli non disgiunti; è la permutazione
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 6 & 5 & 3 & 2 & 7 & 4 \end{pmatrix} = (2\ 6\ 7\ 4\ 3\ 5)\ :$$
è un 6-ciclo.

Definizione 8.2.3. Un 2-ciclo è chiamato **trasposizione**. È del tipo $(n_1\ n_2)$, ossia n_1 e n_2 vengono scambiati e tutti gli altri elementi rimangono fissi.

Osservazione 32.
1. Osserviamo che $(3\ 5\ 1) = (5\ 1\ 3) = (1\ 3\ 5)$, ossia un ciclo può essere "iniziato" in una posizione qualunque.
2. Due cicli disgiunti commutano.

Proposizione 8.2.4. *1. Ogni permutazione è prodotto di cicli disgiunti.*

2. Ogni ciclo è prodotto di trasposizioni, in generale non disgiunte.

3. Ogni permutazione è prodotto di trasposizioni, in generale non disgiunte.

Dimostrazione. La 1. segue subito dalla definizione di ciclo.

Per dimostrare la 2. si ragiona per induzione sulla lunghezza $k \geq 2$ del ciclo: se $k = 2$ il ciclo è una trasposizione; supponiamo vera l'affermazione per cicli di lunghezza $k - 1$, e sia $\sigma = (n_1\ n_2\ \dots\ n_k)$ un ciclo di lunghezza k. L'osservazione cruciale è che σ si può scrivere come prodotto di un $(k-1)$-ciclo τ e di una trasposizione ρ, nella forma $\sigma = \tau\rho = (n_1\ n_2\ \dots\ n_{k-1})(n_k\ n_1)$. Per vederlo meglio scriviamo per esteso come operano queste permutazioni, gli elementi non indicati rimangono fissi:
$$\tau\rho = \begin{pmatrix} n_1 & n_2 & \dots & n_{k-2} & n_{k-1} & n_k \\ n_2 & n_3 & \dots & n_{k-1} & n_1 & n_k \end{pmatrix} \begin{pmatrix} n_1 & n_k \\ n_k & n_1 \end{pmatrix}.$$
Quindi concludiamo usando l'ipotesi induttiva.

La 3. segue dalle precedenti. □

8.3 Segno di una permutazione e gruppo alternante

L'espressione di una permutazione come prodotto di trasposizioni non è unica. Introdurremo ora la nozione di inversione, che permetterà di associare a ogni permutazione un segno.

Definizione 8.3.1. Sia $\sigma \in \mathcal{S}_n$ una permutazione. **Un'inversione di** σ è una coppia di indici $i < j$ compresi fra 1 e n tali che $\sigma(i) > \sigma(j)$.

Per esempio il ciclo $\sigma = (1\,2)$ ha una inversione, in quanto $1 < 2$ ma $\sigma(1) = 2 > 1 = \sigma(2)$. Il ciclo $\tau = (1\ 2\ 3)$ ha due inversioni: $\tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$; le coppie da considerare sono 3 e precisamente:

$$1 < 2 \text{ e } \sigma(1) = 2 < 3 = \sigma(2); 1 < 3 \text{ ma } \sigma(1) = 2 > 1 = \sigma(3); 2 < 3 \text{ ma } \sigma(2) = 3 > 1 = \sigma(3).$$

Definizione 8.3.2. Sia $\sigma \in \mathcal{S}_n$ una permutazione. **Il segno di** σ è $\operatorname{sgn} \sigma = (-1)^a$, dove a è il numero di inversioni di σ, ossia

$$\operatorname{sgn} \sigma = \begin{cases} 1 & \text{se } a \text{ è pari;} \\ -1 & \text{se } a \text{ è dispari.} \end{cases}$$

Una permutazione è detta pari se ha segno 1, dispari se ha segno -1. Le permutazioni pari sono quelle che presentano un numero pari di inversioni, mentre le permutazioni dispari presentano un numero dispari di inversioni.

Proposizione 8.3.3. *Ogni trasposizione ha segno -1, cioè è dispari.*

Dimostrazione. Sia $\sigma = (n_1\, n_2)$ una trasposizione. Scriviamo σ per esteso:

$$\begin{pmatrix} 1 & 2 & \dots & n_1 - 1 & n_1 & n_1 + 1 & \dots & n_2 - 1 & n_2 & n_2 + 1 & \dots & n \\ 1 & 2 & \dots & n_1 - 1 & n_2 & n_1 + 1 & \dots & n_2 - 1 & n_1 & n_2 + 1 & \dots & n \end{pmatrix}.$$

Se $i < j$ è un'inversione di σ, notiamo subito che nessuno fra i e j può essere compreso fra 1 e $n_1 - 1$ nè fra $n_2 + 1$ e n. Invece per ogni n_k con $n_1 < n_k < n_2$, n_k compare in due inversioni: $\sigma(n_1) = n_2 > \sigma(n_k) = n_k$; $\sigma(n_k) = n_k > \sigma(n_2) = n_1$. Infine $n_1 < n_2$ ma $\sigma(n_1) = n_2 > \sigma(n_2) = n_1$. Quindi complessivamente c'è un numero dispari di inversioni e $\operatorname{sgn} \sigma = -1$. $\square$

Il prossimo teorema viene enunciato senza dimostrazione.

Teorema 8.3.4. *Date due permutazioni* $\sigma, \tau \in \mathcal{S}_n$*, si ha* $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$*. Quindi l'applicazione* $\operatorname{sgn} : \mathcal{S}_n \to \{-1, 1\}$*, dove* $\{-1, 1\}$ *è il gruppo moltiplicativo degli elementi invertibili di* $\mathbb{Z}$*, conserva il prodotto, ed è dunque un omomorfismo di gruppi.*

La dimostrazione segue da un lemma che afferma che $\operatorname{sgn}\sigma = \prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}$.

Corollario 8.3.5. *1. Se* $\sigma = \tau_1 \cdots \tau_k$ *è un prodotto di* k *trasposizioni, allora* $\operatorname{sgn}\sigma = \prod_{i=1}^{k} \operatorname{sgn}(\tau_i) = (-1)^k$*.*

2. Se σ *è un* k*-ciclo, allora* $\operatorname{sgn}\sigma = (-1)^{k-1}$*.*

3. Le permutazioni σ *e* σ^{-1} *hanno lo stesso segno.*

Dimostrazione. La 1. segue immediatamente dalla Proposizione 8.3.3 e dal Teorema 8.3.4. Per dimostrare la 2. procediamo per induzione su $k \geq 2$. Se $k = 2$, un 2-ciclo è una trasposizione e ha dunque segno -1 per la Proposizione 8.3.3. Supponiamo vera la tesi per i $(k-1)$-cicli e consideriamo un k-ciclo $\sigma = (n_1, \ldots, n_k)$. Come nella dimostrazione della Proposizione 8.2.4, σ si scrive come prodotto di un $(k-1)$-ciclo τ e di una trasposizione ρ. Dunque per ipotesi induttiva $\operatorname{sgn}(\sigma) = (-1)^{k-2}(-1) = (-1)^{k-1}$. Per la 3., notiamo che $\operatorname{sgn}(\sigma\sigma^{-1}) = \operatorname{sgn}(\mathrm{id}_{I_n}) = 1 = \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma^{-1})$. $\square$

Per esempio, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} = (1\ 5)(2\ 4)$ è prodotto di due trasposizioni, dunque $\operatorname{sgn}(\sigma) = (-1)^2 = 1$.

Il Corollario 8.3.5 permette di trovare facilmente il segno di una permutazione: la si decompone come prodotto di cicli disgiunti e poi si applica la parte 2 del Corollario 8.3.5.

Definizione 8.3.6. Sia $\mathcal{A}_n = \{\sigma \in \mathcal{S}_n \mid \sigma \text{ è pari}\}$: è un sottogruppo di $\mathcal{S}_n$, detto **sottogruppo alternante** di $\mathcal{S}_n$, o semplicemente **gruppo alternante**.

Infatti dal Teorema 8.3.4 segue che $\mathcal{A}_n$ è chiuso rispetto al prodotto e che se σ è pari anche σ^{-1} lo è. Inoltre ovviamente la permutazione identica è pari.

Invece l'insieme delle permutazioni dispari non è chiuso rispetto al prodotto: il prodotto di due permutazioni dispari è pari, mentre il prodotto di una pari e una dispari è dispari.

Osservazione 33. Se τ è una permutazione dispari fissata, l'insieme $\tau\mathcal{A}_n = \{\tau\sigma \mid \sigma \in \mathcal{A}_n\}$ coincide con l'insieme di **tutte** le permutazioni dispari. Infatti se α è una qualunque permutazione dispari, la si può scrivere nella forma $\alpha = \tau(\tau^{-1}\alpha)$, dove chiaramente $\tau^{-1}\alpha$ è pari. Osserviamo che di conseguenza l'insieme delle permutazioni dispari è in biiezione con $\mathcal{A}_n$, tramite l'applicazione $\mathcal{A}_n \to \tau\mathcal{A}_n$, tale che $\sigma \to \tau\sigma$.

Quindi $\mathcal{S}_n = \mathcal{A}_n \cup \tau\mathcal{A}_n$ e i due sottinsiemi sono fra loro in biiezione, dunque hanno lo stesso numero di elementi $\frac{n!}{2}$.

Esempi 8.3.7.

$$\mathcal{S}_3 = \{\mathrm{id}_{I_3}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\} \supset \mathcal{A}_3 = \{\mathrm{id}_{I_3}, (1\ 2\ 3), (1\ 3\ 2)\} :$$

$\mathcal{A}_3$ è un gruppo abeliano con 3 elementi.

Elenchiamo ora i 24 elementi di $\mathcal{S}_4$. Le permutazioni dispari sono le trasposizioni e i 4-cicli:

$$(1\ 2), (1\ 3), (1\ 4), (2\ 3), (2\ 4), (3\ 4), (1\ 2\ 3\ 4), (1\ 2\ 4\ 3), (1\ 3\ 2\ 4), (1\ 3\ 4\ 2), (1\ 4\ 2\ 3)(1\ 4\ 3\ 2).$$

Scriviamo ora gli elementi di $\mathcal{A}_4$:

$$\mathrm{id}_{I_4}, (1\ 2\ 3), (1\ 3\ 2), (1\ 2\ 4), (1\ 4\ 2), (1\ 3\ 4), (1\ 4\ 3), (2\ 3\ 4), (2\ 4\ 3), (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3);$$

vi sono dunque, oltre alla permutazione identica, i 3-cicli e i prodotti di 2 2-cicli disgiunti.

8.4 Funzioni determinante

Iniziamo con un esempio. Consideriamo una matrice quadrata di ordine 2: $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$; per definizione il determinante di A è $\det(A) = a_{11}a_{22} - a_{12}a_{21}$. Come si è visto nel Foglio 2, Es. 1, $\det(A) \neq 0$ se e solo se le righe di A sono linearmente indipendenti, che equivale a dire che il rango di A è 2 o che A è invertibile. Possiamo interpretare det come un'applicazione che associa alla coppia delle righe di una matrice A lo scalare $\det(A)$, ossia come una funzione di due "variabili", che sono vettori di K^2. Questa funzione det è lineare rispetto a ogni riga ed è nulla se e solo se le righe sono linearmente dipendenti.

Generalizzeremo questo esempio, definendo che cosa si intende per una funzione determinante su uno spazio vettoriale V.

Definizione 8.4.1. Sia V un K-spazio vettoriale di dimensione n. Fissiamo un intero $k \geq 1$. Un'applicazione

$$D : V \times \cdots \times V = V^k \to K$$

è detta **multilineare o k-lineare** se è lineare in ogni argomento, cioè se per ogni $i = 1, \ldots, k$ si ha

$$D(v_1, \ldots, v_{i-1}, \lambda v_i + \mu v_i', v_{i+1} \ldots, v_k) = \lambda D(v_1, \ldots, v_i, \ldots, v_k) + \mu D(v_1, \ldots, v_i', \ldots, v_k).$$

D è detta **alternante** se $D(v_1, \ldots, v_k) = 0$ ogni volta che due degli argomenti sono uguali, ossia esistono indici $i \neq j$ tali che $v_i = v_j$.

Definizione 8.4.2. Una funzione determinante su V, K-spazio vettoriale di dimensione n, è un'applicazione multilineare alternante $D : V^n \to K$: in questo caso il dominio è la potenza n-esima di V dove $n = \dim V$.

Proposizione 8.4.3. *Sia D una funzione determinante su V. Allora:*

1. *D è* **antisimmetrica***: $D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -D(v_1, \ldots, v_j, \ldots, v_i \ldots, v_n)$: se scambiamo due degli argomenti la funzione cambia segno;*

2. *se $\sigma \in \mathcal{S}_n$, allora $D(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \operatorname{sgn}\sigma\ D(v_1, \ldots, v_n)$;*

3. *se $v_1, \ldots, v_n$ sono linearmente dipendenti, allora $D(v_1, \ldots, v_n) = 0$. In particolare se uno degli argomenti $v_i = 0$, allora $D(v_1, \ldots, v_n) = 0$.*

Dimostrazione. 1. Consideriamo $D(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_n) = 0$ perchè D è alternante. Per la multilinearità questo coincide con

$$D(\ldots, v_i, \ldots, v_i, \ldots) + D(\ldots, v_i, \ldots, v_j, \ldots) + D(\ldots, v_j, \ldots, v_i, \ldots) + D(\ldots, v_j \ldots, v_j, \ldots) =$$

$$\text{per l'alternanza} = D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) = 0.$$

2. Se σ è una trasposizione ha segno -1 e la proprietà segue dal punto 1. Altrimenti applichiamo la Proposizione 8.2.4 e il Corollario 8.3.5: se σ è un prodotto di k trasposizioni il suo segno è $(-1)^k$; ogni volta che operiamo con una trasposizione sugli argomenti, D cambia segno, dunque alla fine avremo la tesi.

3. Se $v_1, \ldots, v_n$ sono linearmente dipendenti, uno di loro è combinazione lineare dei rimanenti. Supponiamo per semplicità che sia v_n:

$$v_n = \lambda_1 v_1 + \cdots + \lambda_{n-1} v_{n-1}.$$

Allora

$$\begin{aligned} D(v_1, \ldots, v_n) &= D(v_1, \ldots, v_{n-1}, \lambda_1 v_1 + \cdots + \lambda_{n-1} v_{n-1}) = \\ &= \sum_{i=1}^{n-1} \lambda_i D(v_1, \ldots, v_{n-1}, v_i) = \\ &= \sum_{i=1}^{n-1} \lambda_i 0 = 0 \end{aligned}$$

perchè D è alternante. □

Esempi 8.4.4. Un esempio banale di funzione determinante è la funzione identicamente nulla: $D(v_1, \ldots, v_n) = 0$ per ogni scelta di $(v_1, \ldots, v_n) \in V^n$.

8.5 Formula di Leibniz per i determinanti

Fissiamo una funzione determinante su V e sia $\mathcal{B} = (v_1, \ldots, v_n)$ una base di V. Per n vettori $w_1, \ldots, w_n \in V$, vogliamo esprimere $D(w_1, \ldots, w_n)$ in termini delle loro coordinate rispetto a $\mathcal{B}$. Scriviamo dunque ogni w_i come combinazione lineare di $v_1, \ldots, v_n$ nella forma

$$w_i = \sum_{j_i=1}^{n} x_{ij_i} v_{j_i}.$$

Abbiamo bisogno di usare indici diversi per ogni vettore perchè ora li faremo variare tutti n contemporaneamente:

$$\begin{aligned}
D(w_1, \ldots, w_n) &= D(\sum_{j_1=1}^{n} x_{1j_1} v_{j_1}, \sum_{j_2=1}^{n} x_{2j_2} v_{j_2}, \ldots, \sum_{j_n=1}^{n} x_{nj_n} v_{j_n}) = \text{multilinearità I argomento} \\
&= \sum_{j_1=1}^{n} x_{1j_1} D(v_{j_1}, \sum_{j_2=1}^{n} x_{2j_2} v_{j_2}, \ldots, \sum_{j_n=1}^{n} x_{nj_n} v_{j_n}) = \text{multilin. per gli altri} \\
&= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} x_{1j_1} x_{2j_2} \cdots x_{nj_n} D(v_{j_1}, \ldots, v_{j_n}).
\end{aligned}$$

Ma $D(v_{j_1}, \ldots, v_{j_n})$ può essere $\neq 0$ soltanto se gli indici $j_1, \ldots, j_n$ sono tutti diversi, cioè se formano una permutazione di $1, \ldots, n$. Quindi possiamo riscrivere così:

$$\begin{aligned}
D(w_1, \ldots, w_n) &= \sum_{\sigma \in \mathcal{S}_n} x_{1\sigma(1)} \ldots x_{n\sigma(n)} D(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \\
&= \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) x_{1\sigma(1)} \ldots x_{n\sigma(n)} D(v_1, \ldots, v_n). \qquad (8.1)
\end{aligned}$$

L'espressione (8.1) è nota come **formula di Leibniz per il determinante**. È una somma con $n!$ addendi quindi è utilizzabile solo per n piccolo. Come conseguenza della formula di Leibniz, abbiamo il seguente corollario.

Corollario 8.5.1. *Sia D una funzione determinante non nulla su V. Allora $(v_1, \ldots, v_n)$ è una base di V se e solo se $D(v_1, \ldots, v_n) \neq 0$.*

Dimostrazione. Se $v_1, \ldots, v_n$ non formano una base, sono linearmente dipendenti, dunque $D(v_1, \ldots, v_n) = 0$ per la Proposizione 8.4.3. Viceversa: essendo $D \neq 0$ esistono n vettori $w_1, \ldots, w_n$ tali che $D(w_1, \ldots, w_n) \neq 0$. Allora se $v_1, \ldots, v_n$ formano una base, applicando la (8.1) otteniamo che dev'essere $D(v_1, \ldots, v_n) \neq 0$. □

Il seguente teorema, molto importante, ci dà l'unicità della funzione determinante a meno di costanti.

Teorema 8.5.2. *Sia V un K-spazio vettoriale di dimensione n, $(v_1, \ldots, v_n)$ una base di V. Allora esiste una e una sola funzione determinante su V, $D : V^n \to K$, tale che $D(v_1, \ldots, v_n) = 1$. Ogni altra funzione determinante è del tipo λD, con $\lambda \in K$.*

Dimostrazione. Come si fa usualmente, dimostriamo prima l'unicità e poi l'esistenza. Supponendo che esista una funzione determinante D tale che $D(v_1, \ldots, v_n) = 1$, la formula di Leibniz (8.1) ci dà che, presi n vettori $w_i = \sum_j x_{ij} v_j$, dev'essere

$$D(w_1, \ldots, w_n) = \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) x_{1\sigma(1)} \ldots x_{n\sigma(n)}, \tag{8.2}$$

e perciò abbiamo l'unicità. Dobbiamo dunque dimostrare che, prendendo la (8.2) come definizione di D, otteniamo una funzione multilineare alternante tale che $D(v_1, \ldots, v_n) = 1$. Per l'ultima affermazione osserviamo che $v_i = \sum_j \delta_{ij} v_j$ e quindi dalla (8.2) otteniamo $D(v_1, \ldots, v_n) = \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) \delta_{1\sigma(1)} \ldots \delta_{n\sigma(n)} = 1$, perchè rimane un unico addendo corrispondente alla permutazione identica.

Verifichiamo che D è multilineare:

$$\begin{aligned} D(w_1, \ldots, \lambda w_i + \mu w_i', \ldots, w_n) &= \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) x_{1\sigma(1)} \cdots (\lambda x_{i\sigma(i)} + \mu x'_{i\sigma(i)}) \ldots x_{n\sigma(n)} = \\ &= \lambda D(w_1, \ldots, w_i, \ldots w_n) + \mu D(w_1, \ldots, w_i', \ldots w_n). \end{aligned}$$

Verifichiamo infine che D è alternante; consideriamo vettori $w_1, \ldots, w_n$ con $w_i = w_j$ dove i, j sono due indici diversi. Ciò significa che $x_{ik} = x_{jk}$ per ogni indice k. Sia $\tau = (i\ j)$ la trasposizione che scambia proprio gli indici i e j. Allora $\mathcal{S}_n = \mathcal{A}_n \cup \tau \mathcal{A}_n$ (Osservazione 33). Quindi

$$\begin{aligned} D(w_1, \ldots, w_i, \ldots, w_j, \ldots, w_n) &= \sum_{\sigma \in \mathcal{S}_n} x_{1\sigma(1)} \cdots x_{n\sigma(n)} = \\ &= \sum_{\sigma \in \mathcal{A}_n} x_{1\sigma(1)} \cdots x_{n\sigma(n)} - \sum_{\sigma \in \mathcal{A}_n} x_{1\sigma(\tau(1))} \cdots x_{n\sigma(\tau(n))}. \end{aligned}$$

Ma

$$\begin{aligned} x_{1\sigma(\tau(1))} \cdots x_{i\sigma(\tau(i))} \cdots x_{j\sigma(\tau(j))} \cdots x_{n\sigma(\tau(n))} &= \quad \text{per definizione di } \tau \\ = x_{1\sigma(1)} \cdots x_{i\sigma(j)} \cdots x_{j\sigma(i)} \cdots x_{n\sigma(n)} &= \quad \text{per l'ipotesi che } w_i = w_j \\ = x_{1\sigma(1)} \cdots x_{j\sigma(j)} \cdots x_{i\sigma(i)} \cdots x_{n\sigma(n)}, & \end{aligned}$$

quindi la somma è nulla. □

8.6 Determinante di una matrice

Consideriamo ora il caso in cui $V = K^n$ in cui è fissata la base canonica $\mathcal{C} = (e_1, \ldots, e_n)$. Il determinante standard, denotato det, è l'unica funzione determinante

$$\det : K^n \times \cdots \times K^n = (K^n)^n \to K$$

tale che $\det(e_1, \ldots, e_n) = 1$. In questo caso il dominio $(K^n)^n$ si interpreta come spazio delle matrici quadrate di ordine n, ossia se $(w_1, \ldots, w_n) \in K^n \times \cdots \times K^n$ le n-uple $w_1, \ldots, w_n$ sono pensate come le n **righe** di una matrice $n \times n$. Quindi $\det(E_n) = \det(e_1, \ldots, e_n) = 1$: è il determinante della matrice identica E_n.

Se $A = (a_{ij})_{i,j=1,\ldots,n}$ è una matrice $n \times n$, useremo la notazione

$$\det(A) = \det(\underbrace{a_1, \ldots, a_n}_{\text{righe di } A}) = \mid A \mid .$$

Esempi 8.6.1 (Determinante delle matrici di ordine 2 e 3)**.**
Se A è una matrice di ordine 2, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, per la regola di Leibniz si ha:

$$\det(A) = \det((a_{11}, a_{12}), (a_{21}, a_{22})) = \underbrace{a_{11}a_{22}}_{\sigma=\mathrm{id}_{I_2}} - \underbrace{a_{12}a_{21}}_{\sigma=(1\ 2)}.$$

Se A è una matrice di ordine 3, $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, lo sviluppo del suo determinante secondo la regola di Leibniz contiene 6 addendi, corrispondenti alle 6 permutazioni di $\mathcal{S}_3$ (Esempio 8.3.7):

$$\det(A) = a_{11}a_{22}a_{33} - \underbrace{a_{11}a_{23}a_{32}}_{\sigma=(2\ 3)} + \underbrace{a_{12}a_{23}a_{31}}_{\sigma=(1\ 2\ 3)} - \underbrace{a_{12}a_{21}a_{33}}_{\sigma=(1\ 2)} + \underbrace{a_{13}a_{21}a_{32}}_{\sigma=(1\ 3\ 2)} - \underbrace{a_{13}a_{22}a_{31}}_{\sigma=(1\ 3)}.$$

La **regola di Sarrus** è un trucco per scrivere velocemente il determinante di una matrice 3×3 (*solo matrici* 3×3). Riscriviamo la matrice A, ripetendo poi le prime due colonne ed evidenziando tre diagonali:

$$\left(\begin{array}{ccc|cc} \color{red}{a_{11}} & \color{green}{a_{12}} & \color{blue}{a_{13}} & a_{11} & a_{12} \\ a_{21} & \color{red}{a_{22}} & \color{green}{a_{23}} & \color{blue}{a_{21}} & a_{22} \\ a_{31} & a_{32} & \color{red}{a_{33}} & \color{green}{a_{31}} & \color{blue}{a_{32}} \end{array}\right)$$

I primi tre addendi di $\det(A)$ sono i tre prodotti degli elementi sulle tre diagonali evidenziate: $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$. Ora dobbiamo invece sottrarre i tre prodotti degli elementi contenuti nelle tre diagonali secondarie evidenziate di seguito:

$$\begin{pmatrix} a_{11} & a_{12} & \color{red}{a_{13}} & | & \color{green}{a_{11}} & \color{blue}{a_{12}} \\ a_{21} & \color{red}{a_{22}} & \color{green}{a_{23}} & | & \color{blue}{a_{21}} & a_{22} \\ \color{red}{a_{31}} & \color{green}{a_{32}} & \color{blue}{a_{33}} & | & a_{31} & a_{32} \end{pmatrix}$$

Si ottiene: $-a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$. Confrontando con l'Esempio 8.6.1, si vede che quello che si trova è proprio il determinante della matrice considerata.

Proposizione 8.6.2. *Sia A una matrice quadrata $n \times n$, e sia ${}^tA$ la sua trasposta. Allora $\det(A) = \det({}^tA)$.*

Dimostrazione. Sia $A = (a_{ij})_{i,j=1,\dots,n}$, ${}^tA = (a'_{ij})_{i,j=1,\dots,n}$ con $a'_{ij} = a_{ji}$ per ogni coppia di indici. Si ha:

$$\begin{aligned} \det({}^tA) &= \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn}\sigma) a'_{1\sigma(1)} \cdots a'_{n\sigma(n)} = \\ &= \sum_{\sigma} (\operatorname{sgn}\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \text{cambiando eventualmente l'ordine} \\ &= \sum_{\sigma} (\operatorname{sgn}\sigma) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} = \\ &= \det(A) \ \text{ perchè } \sigma^{-1} \text{ descrive tutto } \mathcal{S}_n. \end{aligned}$$

□

Osservazione 34. Come conseguenza osserviamo che il determinante è una funzione multilineare alternante anche sulle colonne di A.

8.7 Comportamento di un determinante per operazioni elementari

Vediamo ora come cambia il determinante di una matrice A se si opera con operazioni elementari sulle righe. Scriviamo la matrice mettendo in evidenza le sue righe.

(OE1) : due righe della matrice vengono scambiate, perciò con **(OE1) il determinante cambia segno**.

(OE2) $A = \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix} \stackrel{(OE2)}{\rightarrow} \begin{pmatrix} A_{(1)} \\ \vdots \\ \lambda A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix}$; per la multilinearità $\det \begin{pmatrix} A_{(1)} \\ \vdots \\ \lambda A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix} = \lambda \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix}$:

il determinante viene moltiplicato per λ;

(OE3) $A = \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix} \stackrel{(OE3)}{\rightarrow} \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} + \lambda A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix}$; ancora per la multilinearità si ha

$$\det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} + \lambda A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix} = \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix} + \lambda \det \begin{pmatrix} A_{(1)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix}.$$

ma la seconda matrice contiene due righe uguali, e perciò il determinante è nullo; quindi **(OE3) lascia il determinante invariato**.

Come conseguenza abbiamo il seguente Corollario.

Corollario 8.7.1. *Se A è una matrice quadrata di ordine n e λ è uno scalare, allora det$(\lambda A) = \lambda^n$det(A). In particolare det$(-A) = (-1)^n$det(A).*

Quanto abbiamo appena osservato riguardo alle operazioni elementari potrà essere applicato al calcolo di determinanti, alla luce della seguente proposizione.

Proposizione 8.7.2. *Sia A una matrice triangolare superiore, $A = (a_{ij})_{i,j=1,\dots,n}$ con $a_{ij} = 0$ per ogni coppia di indici tali che $i > j$: $\begin{pmatrix} a_{11} & * & * & \cdots & * \\ 0 & a_{22} & * & \cdots & * \\ \vdots & & \ddots & & * \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$. Allora det$(A) = a_{11}a_{22}\cdots a_{nn}$: è il prodotto degli elementi che stanno sulla diagonale principale.*

Dimostrazione. $\det(A) = \sum_\sigma (\operatorname{sgn}\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}$; osserviamo che se $\sigma \neq \mathrm{id}_{I_n}$ c'è almeno un indice i tale che $i > \sigma(i)$.

Per l'ipotesi che A sia triangolare superiore, questo implica che nel prodotto $a_{1\sigma(1)} \dots a_{n\sigma(n)}$ c'è almeno un fattore nullo, e quindi in $\det(A)$ rimane solo l'addendo corrispondente a id_{In}. □

La stessa proprietà vale anche per le matrici triangolari inferiori (con lo stesso ragionamento).

Allora per calcolare il determinante di una matrice A, possiamo trasformarla in una matrice a scala con operazioni elementari; ricordiamo che una matrice a scala è in particolare triangolare superiore. Infine, teniamo conto del fatto che la OE1 cambia il segno, OE2 fa moltiplicare il determinante per uno scalare e OE3 lascia il determinante invariato.

Esempi 8.7.3. Presentiamo un esempio di una matrice 4×4; lo sviluppo del suo determinante secondo la regola di Leibniz contiene $4! = 24$ addendi; con l'algoritmo appena esposto il calcolo è molto più rapido.

$$\begin{vmatrix} 0 & 3 & 1 & 2 \\ 5 & 1 & 3 & 3 \\ 0 & 2 & 0 & 0 \\ 1 & 3 & 1 & 1 \end{vmatrix} = \text{OE1 scambio la prima con la quarta riga} = -\begin{vmatrix} \boxed{1} & 3 & 1 & 1 \\ 5 & 1 & 3 & 3 \\ 0 & 2 & 0 & 0 \\ 0 & 3 & 1 & 2 \end{vmatrix} = \text{OE3} =$$

$$-\begin{vmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & -14 & -2 & -2 \\ 0 & 2 & 0 & 0 \\ 0 & 3 & 1 & 2 \end{vmatrix} = \text{OE1 e OE2} = 2\begin{vmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & -14 & -2 & -2 \\ 0 & 3 & 1 & 2 \end{vmatrix} = \text{OE3} = 2\begin{vmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 1 & 2 \end{vmatrix} =$$

$$\text{OE2} =$$

$$= -4\begin{vmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 1 & 2 \end{vmatrix} = \text{OE3} = -4\begin{vmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & \boxed{1} \end{vmatrix} = -4.$$

Abbiamo visto nel Corollario 8.5.1 che una funzione determinante non nulla, definita su V, non si annulla su vettori $v_1, \dots, v_n$ se e solo se questi formano una base. Ne segue:

Proposizione 8.7.4. *Sia A una matrice quadrata. Allora* $\det(A) = \begin{vmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ A_{(n)} \end{vmatrix} \neq 0$ *se e solo se le righe di A formano una base di $\mathbb{K}^n$.*

Corollario 8.7.5. *Sia A una matrice quadrata. Allora det$(A) \neq 0$ se e solo se* $\mathrm{rg}A = n$.

Proposizione 8.7.6. *Sia* $A = \begin{pmatrix} B & | & * \\ --- & | & --- \\ 0 & | & C \end{pmatrix}$ *una matrice a blocchi, con B, C matrici quadrate. Allora*

$$\mathit{det}(A) = \mathit{det}(B)\mathit{det}(C).$$

Dimostrazione. Si usano operazioni elementari, trasformando B e C in matrici triangolari superiori senza moltiplicare nessuna riga per uno scalare, ma operando solo con OE1 e OE3. Prima si passa da B a B' triangolare con operazioni elementari che comprendono k scambi di righe, lavorando solo sulle righe di B; poi si passa da C a C' triangolare, con operazioni elementari sulle righe di C, con l scambi di righe: $A \to \begin{pmatrix} B' & | & * \\ --- & | & --- \\ 0 & | & C' \end{pmatrix} = A'$. Allora det$(A')$ = triangolare = det(B')det(C') = $(-1)^k$det$(B)(-1)^l$det(C) = $(-1)^{k+l}$det(A). Si conclude che det(A) = det(B)det(C). $\square$

8.8 Teorema di Binet

La dimostrazione di questo importante teorema, di enunciato molto semplice, si basa sulla teoria dei determinanti che abbiamo costruito.

Teorema 8.8.1. *Siano A, B matrici quadrate $n \times n$ su K. Allora det(AB) = det(A)det(B). In particolare l'applicazione det : GL$(n, \mathbb{K}) \to \mathbb{K} \setminus \{0\}$ è un omomorfismo di gruppi (moltiplicativi).*

Dimostrazione. Consideriamo dapprima il caso particolare in cui det$(B) = 0$: allora le righe di B sono linearmente dipendenti, quindi il sistema lineare omogeneo $BX = 0$ ha soluzioni non nulle: esiste $v \neq 0$, $v \in K^n$, tale che $Bv = 0$. Allora anche $(AB)v = A(Bv) = A0 = 0$: ciò significa che anche il sistema lineare omogeneo $(AB)X = 0$ ha una soluzione non nulla, dunque AB ha rango $< n$, e quindi det$(AB) = 0$.

Sia ora det$(B) \neq 0$. Consideriamo l'applicazione

$$f_B : M(n \times n, \mathbb{K}) \to \mathbb{K}$$

definita da

$$f_B(A) = \frac{\det(AB)}{\det(B)}.$$

Vogliamo dimostrare che $f_B(A) = \det(A)$. Per il Teorema 8.5.2 basta dimostrare che:

1. f_B è multilineare e alternante nelle righe di A;
2. $f_B(\mathbb{I}_n) = 1$.

La 2. è immediata: $f_B(\mathbb{I}_n) = \frac{\det(\mathbb{I}_n B)}{\det(B)} = 1$. Per dimostrare la 1. consideriamo una matrice A con $a_j = \lambda a'_j + \mu a''_j$; allora al posto di indici jk di AB abbiamo

$$a_j b^k = (\lambda a'_j + \mu a''_j) b^k = \lambda(a'_j b^k) + \mu(a''_j b^k);$$

perciò la riga j-esima di AB è una combinazione lineare di due righe. Ma il determinante è lineare nella j-esima variabile, perciò $\det(AB) = \lambda \det(A'B) + \mu(A''B)$, dove A' è ottenuta da A sostituendo la riga a_j con a'_j, e analogamente per A''. Quindi

$$f_B(A) = \lambda f_B(A') + \mu f_B(A'').$$

Infine: se A ha due righe uguali, anche AB ha due righe uguali, e quindi $\det(AB) = 0$, che implica $f_B(AB) = 0$.

Concludiamo così che $\det(A) = f_B(A)$. □

Corollario 8.8.2. *Sia A una matrice quadrata invertibile. Allora det$(A^{-1}) = ($det$A)^{-1}$.*

Dimostrazione. $AA^{-1} = \mathbb{I}_n$; allora per il Teorema di Binet 8.8.1, $\det(AA^{-1}) = 1 = \det(A)\det(A^{-1})$. □

Corollario 8.8.3. *Se A, B sono matrici quadrate simili, allora det(A) = det(B).*

Dimostrazione. Da $B = S^{-1}AS$ segue $\det(B) = \det(S^{-1}AS) = \det(S^{-1})\det(A)\det(S) =$ per il Corollario precedente $= \det(S)^{-1}\det(A)\det(S) = \det(A)$. □

Il Corollario 8.8.3 permette di dare la definizione di determinante di un endomorfismo.

Definizione 8.8.4. Sia $f : V \to V$ un endomorfismo di un $\mathbb{K}$-spazio vettoriale V di dimensione finita. Si definisce il **determinante di** f ponendo $\det(f) = \det M^{\mathcal{B}}_{\mathcal{B}}(f)$, qualunque sia la base $\mathcal{B}$ di V considerata.

La definizione è ben posta perchè se si prende un'altra base $\mathcal{B}'$ di V, le matrici $M^{\mathcal{B}}_{\mathcal{B}}(f)$ e $M^{\mathcal{B}'}_{\mathcal{B}'}$ di f sono simili, e hanno perciò lo stesso determinante. Con questa definizione, abbiamo che f è un isomorfismo di V in sè se e solo se $\det(f) \neq 0$.

8.9 Matrice aggiunta

Sia $A = (a_{ij})_{i,j=1,\ldots,n}$ una matrice $n \times n$. Per ogni i, j denotiamo A'_{ij} la seguente matrice $n \times n$:

$$\begin{array}{cc} & \begin{array}{ccccccccc} & & & & \overset{j}{\downarrow} & & & & \end{array} \\ i \rightarrow & \begin{pmatrix} a_{11} & \ldots & a_{1,j-1} & | & 0 & | & a_{1,j+1} & \ldots & a_{1n} \\ \vdots & & & | & \vdots & | & & & \\ a_{i-1,1} & \ldots & a_{i-1,j-1} & | & 0 & | & a_{i-1,j+1} & \ldots & a_{i-1,n} \\ 0 & & 0 & & 1 & & 0 & \ldots & 0 \\ a_{i+1,1} & \ldots & a_{i+1,j-1} & | & 0 & | & a_{i+1,j+1} & \ldots & a_{i+1,n} \\ \vdots & & & | & \vdots & | & & & \\ a_{n1} & \ldots & a_{n,j-1} & | & 0 & | & a_{n,j+1} & \ldots & a_{nn} \end{pmatrix} \end{array} \tag{8.3}$$

Abbiamo sostituito all'i-esima riga e_j e alla j-esima colonna e_i. Osserviamo che, se procediamo con trasformazioni elementari del III tipo su A'_{ij} come segue: alla I riga sommiamo la riga i-esima moltiplicata per a_{1j}, alla II riga sommiamo la riga i-esima moltiplicata per a_{2j}, e così via, lasciando invariata soltanto la riga i-esima, otteniamo la matrice

$$\begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ e_j \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}.$$

Con queste trasformazioni elementari il determinante non è cambiato.

Definizione 8.9.1. La matrice $\widetilde{A} = (\widetilde{a}_{ij})_{ij=1,\ldots,n}$, dove $\widetilde{a}_{ij} := | A'_{ji} |$, è detta **matrice aggiunta di** A.

Notiamo che gli indici sono scambiati, dunque $\widetilde{A} = {}^t(| A'_{ij} |)_{i,j=1,\ldots,n}$.

Teorema 8.9.2. *Sia A una matrice quadrata. Si ha $\widetilde{A}A = A\widetilde{A} = det(A)E_n$.*

Dimostrazione. Denotiamo con b_{ik} l'elemento di $A\widetilde{A}$ di indici ik. Si ha:

$$b_{ik} = \sum_{j=1}^{n} a_{ij}\widetilde{a}_{jk} = \sum_{j=1}^{n} a_{ij} \mid A'_{kj} \mid = \sum_{j=1}^{n} a_{ij}\text{det}\begin{pmatrix} a_1 \\ \vdots \\ a_{k-1} \\ e_j \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix} = \text{det}\begin{pmatrix} a_1 \\ \vdots \\ a_{k-1} \\ \sum_{j=1}^{n} a_{ij}e_j \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix}.$$

L'ultima uguaglianza segue dalla linearità del determinante rispetto al k-esimo argomento. Ma $\sum_j a_{ij}e_j = a_i$: è la riga i-esima di A che ora va nella posizione k-esima. Quindi si trova:

$$b_{ik} = \text{det}\begin{pmatrix} a_1 \\ \vdots \\ a_{k-1} \\ a_i \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix} = \begin{cases} 0 & \text{se } i \neq k \\ \text{det}(A) & \text{se } i = k \end{cases} = \text{det}(A)\delta_{ik}.$$

Perciò $A\widetilde{A} = \text{det}(A)E_n$. Per calcolare $\widetilde{A}A$ si lavora in maniera analoga sulle colonne, ricordando che il determinante è multilineare alternante anche sulle colonne della matrice (Osservazione 34). □

Corollario 8.9.3 (Formula per l'inversa di una matrice)**.** *Se A è una matrice quadrata invertibile, si ha*

$$A^{-1} = \frac{1}{det(A)}\widetilde{A}.$$

Notiamo che si tratta di una formula che dà direttamente l'espressione dell'inversa di una matrice, da distinguersi dall'algoritmo della Sezione **??**.

Esempi 8.9.4. Sia A una matrice 2×2. Se

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \text{ allora}$$

$$\widetilde{a}_{11} = \begin{vmatrix} 1 & 0 \\ 0 & a_{22} \end{vmatrix} = a_{22},\ \widetilde{a}_{22} = \begin{vmatrix} a_{11} & 0 \\ 0 & 1 \end{vmatrix} = a_{11},\ \widetilde{a}_{21} = \begin{vmatrix} 0 & 1 \\ a_{21} & 0 \end{vmatrix} = -a_{21},\ \widetilde{a}_{12} = -a_{12}.$$

Dunque

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

8.10 Sviluppo di Laplace di un determinante

Data una matrice quadrata A $n \times n$, denotiamo con A_{ij} la sua sottomatrice $(n-1) \times (n-1)$ ottenuta cancellando la riga i-esima e la colonna j-esima di A. A_{ij} è detta sottomatrice complementare dell'elemento a_{ij} di A e il suo determinante **minore complementare di** a_{ij}.

Confrontiamo il determinante della matrice A_{ij} con quello della matrice A'_{ij} definita nella Sezione 8.9.

Proposizione 8.10.1. *Si ha det*$(A'_{ij}) = \widetilde{a}_{ji} = (-1)^{i+j}$*det*$(A_{ij}$.

Dimostrazione. Infatti da A'_{ij} con $i-1$ scambi di righe e $j-1$ scambi di colonne passiamo alla matrice a blocchi

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_{ij} & \\ 0 & & & \end{pmatrix}$$

che ha lo stesso determinante di A_{ij}. Otteniamo quindi l'espressione voluta perchè $(i-1)+(j-1)$ ha la stessa parità di $i+j$. $\square$

Definizione 8.10.2. L'elemento $(-1)^{i+j}\det(A_{ij})$ è detto **complemento algebrico o cofattore dell'elemento** a_{ij} di A. La matrice aggiunta è detta anche **matrice cofattore** di A.

Teorema 8.10.3 (I Teorema di Laplace per lo sviluppo di un determinante)**.**

Sia $A = (a_{ij})_{i,j=1,\dots,n}$ *una matrice* $n \times n$.

1. Sia i *un indice di riga fissato. Allora*

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij}),$$

ogni addendo è prodotto dell'elemento a_{ij} *per il suo complemento algebrico. Questo è detto sviluppo di det*(A) *seconda la riga* i*-esima.*

2. Sia j *un indice di colonna fissato. Allora*

$$det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij}),$$

anche qui ogni addendo è prodotto dell'elemento a_{ij} per il suo complemento algebrico. Questo è detto sviluppo di det(A) seconda la colonna j-esima.

Dimostrazione. Sappiamo che $A\widetilde{A} = (b_{ij})_{i,j} = \det(A)E_n$. Allora per ogni i

$$b_{ii} = \det(A) = \sum_{j=1}^{n} a_{ij}\widetilde{a}_{ji} = \sum_{j} a_{ij} \mid A'_{ij} \mid = \sum_{j} (-1)^{i+j} a_{ij} \mid A_{ij} \mid .$$

Se si considera $\widetilde{A}A$ si ottiene l'analoga espressione per le colonne. □

Esempi 8.10.4. 1. Consideriamo il seguente determinante:

$$\begin{vmatrix} 0 & 1 & 2 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \text{sviluppiamo secondo la terza riga e otteniamo} =$$

$$= (-1)^{3+1} \cdot 1 \cdot \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} + (-1)^{3+2} \cdot 1 \cdot \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} + 0 = -3 + 6 = 3.$$

2. $\begin{vmatrix} 4 & 0 & 5 & 6 \\ 7 & 0 & 8 & 2 \\ 1 & 3 & 2 & 1 \\ 0 & 0 & -1 & 0 \end{vmatrix}$: lo sviluppiamo secondo la seconda colonna $= -3 \begin{vmatrix} 4 & 5 & 6 \\ 7 & 8 & 2 \\ 0 & -1 & 0 \end{vmatrix} =$ sviluppiamo second

$-3 \begin{vmatrix} 4 & 6 \\ 7 & 2 \end{vmatrix} = -3(8 - 42) = 102.$

3. Risoluzione dei sistemi lineari omogenei di due equazioni in tre incognite.

Consideriamo il sistema lineare

$$\begin{cases} ax + by + cz & = 0 \\ a'x + b'y + c'z & = 0 \end{cases} \text{avente come matrice } \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}.$$

Tale sistema ha la seguente soluzione $\left(\begin{vmatrix} b & c \\ b' & c' \end{vmatrix}, - \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}, \begin{vmatrix} a & b \\ a' & b' \end{vmatrix} \right)$.

Per vederlo, considerare lo sviluppo di Laplace secondo la prima riga dei seguenti determinanti con due righe uguali:

$$\begin{vmatrix} a & b & c \\ a & b & c \\ a' & b' & c' \end{vmatrix} = 0, \quad \begin{vmatrix} a' & b' & c' \\ a & b & c \\ a' & b' & c' \end{vmatrix} = 0.$$

Una proprietà analoga vale per i sistemi lineari di $n - 1$ equazioni in n incognite.

Osservazione 35. Come conseguenza del Teorema di Laplace 8.10.3, si ha che, se si fissano due indici di riga diversi $i \neq k$, allora

$$\sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{kj}) = 0, \tag{8.4}$$

ossia se si sommano i prodotti degli elementi di una riga per i complementi algebrici degli elementi di un'altra riga, si ottiene 0. In effetti la (8.4) si può pensare come lo sviluppo di Laplace del determinante di una matrice con due righe uguali, cioè la matrice ottenuta sostituendo in A alla riga k-esima la riga i-esima. Questo è talvolta citato come **II Teorema di Laplace.**

La proprietà analoga vale anche per le colonne.

8.11 Teorema di Cramer

Consideriamo un sistema lineare di n equazioni in n incognite la cui matrice dei coefficienti A ha rango massimo n. Equivalentemente A è invertibile, o ha determinante non nullo; si dice anche che A è non singolare o non degenere. In tal caso il sistema ha una e una sola soluzione qualunque sia la colonna dei termini noti (Sezione **??**). Il Teorema di Cramer fornisce una formula esplicita per l'unica soluzione.

Teorema 8.11.1. *Sia A una matrice $n \times n$ invertibile a coefficienti in K, siano $a^1, \ldots, a^n$ le sue colonne e sia $Ax = b$ un sistema lineare con $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in K^n$. Allora l'unica soluzione del sistema è la n-upla $(y_1, \ldots, y_n)$ di coordinate*

$$y_j = \frac{det(a^1, \cdots, a^{j-1}, b,\ a^{j+1}, \cdots, a^n)}{det(A)}$$

per ogni $j = 1, \ldots, n$: è un quoziente di due determinanti, a denominatore quello di A, a numeratore quello della matrice ottenuta sostituendo alla colonna j-esima di A la colonna dei termini noti.

Dimostrazione. Sia $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ la soluzione. Allora $Ay = b$; ciò equivale a

$$y_1a^1 + y_2a^2 + \cdots + y_na^n = b.$$

Perciò $\det(a^1, \cdots, a^{j-1}, b,\ a^{j+1}, \cdots, a^n) = \det(a^1, \cdots, a^{j-1},\ y_1a^1+y_2a^2+\cdots+y_na^n,\ a^{j+1}, \cdots, a^n) =$ per la multilinearità $= \sum_{i=1}^n y_i \det(a^1, \cdots, a^{j-1}, a^i,\ a^{j+1}, \cdots, a^n) = y_j \det(A)$. Essendo $\det(A) \neq 0$, possiamo dividere per $\det(A)$ e otteniamo la tesi. □

Osserviamo che l'unica soluzione è anche esprimibile come $y = A^{-1}b$.

8.12 Rango e minori di una matrice

Definizione 8.12.1. I minori di una matrice sono i determinanti delle sue sottomatrici quadrate. Si dice che un minore ha ordine s se è il determinante di una sottomatrice $s \times s$.

Vale la seguente caratterizzazione del rango di una matrice.

Proposizione 8.12.2. *Sia $A \in M(m\times n, K)$ una matrice $m\times n$. Allora A ha rango r se e solo se A ha almeno un minore non nullo di ordine r e ogni minore di A di ordine $s > r$ è nullo. Equivalentemente, il rango di A è il massimo ordine di un minore non nullo di A.*

Dimostrazione. Se A ha rango r, A contiene r righe linearmente indipendenti; possiamo supporre che siano $a_1, \cdots, a_r$. Allora la sottomatrice $A' = \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}$ formata dalle prime r righe di A è una matrice $r \times n$ di rango r. Quindi A' ha r colonne linearmente indipendenti: questo ci dà una sottomatrice $r \times r$ invertibile di A. Dato che, per ipotesi, se $s > r$, s righe o colonne di A comunque prese sono linearmente dipendenti, ogni sottomatrice quadrata di ordine s ha righe e colonne linearmente dipendenti.

Viceversa, supponiamo che la sottomatrice formata dalla prime r righe e colonne di A sia invertibile; allora le prime r righe e colonne di A sono linearmente indipendenti, quindi $\operatorname{rg} A \geq r$. Ma $\operatorname{rg} A$ non può essere $> r$ altrimenti per la precedente implicazione potrei trovare un minore non nullo di ordine $> r$. □

Si può dimostrare il seguente teorema di Kronecker, o degli orlati, che permette di semplificare il procedimento espresso dal Teorema 8.12.2. In una matrice A, considerata una sua sottomatrice quadrata B di ordine p con determinante diverso da zero, si definiscono **orlati di** B tutte le sottomatrici quadrate di ordine $p + 1$, ottenute aggiungendo a B (parte di) una riga e una colonna di A. Il Teorema afferma che, se tutti gli orlati di B hanno determinante nullo, allora $\operatorname{rg} A = p$.

8.13 Il determinante come area o volume

Siano $v_1, \dots, v_n$ vettori non nulli di $\mathbb{R}^n$.

Definizione 8.13.1. Il parallelepipedo generato da $v_1, \dots, v_n$ è il sottinsieme di $\mathbb{R}^n$:

$$P(v_1, \dots, v_n) := \{v = \lambda_1 v_1 + \dots \lambda_n v_n \mid 0 \leq \lambda_i \leq 1,\ i = 1, \dots, n\}.$$

Per esempio, per $n = 1$ troviamo $P(v) = \{\lambda v \mid 0 \leq \lambda \leq 1\} \subset \mathbb{R}$: è il segmento corrispondente al vettore v. Per $n = 2$, troviamo il parallelogramma $P(v_1, v_2)$ avente i due vettori v_1, v_2 come lati. Se v_1, v_2 sono linearmente dipendenti, il parallelogramma è detto degenere, è il caso di due vettori paralleli.

Proposizione 8.13.2. *Siano $v_1, v_2 \in \mathbb{R}^2$ vettori linearmente indipendenti. L'area di $P(v_1, v_2)$ è il valore assoluto del determinante dei due vettori: $\mid det \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \mid$.*

Figura 8.1: Parallelogramma di lati v_1, v_2 e rettangolo di lati v_1', v_2'

Cenno di dimostrazione. Consideriamo il parallelogramma di lati v_1, v_2 come in figura. L'area è il prodotto "base $\times$ altezza". Modifichiamo ora v_1 senza cambiare tale area, sostituiamo cioè a v_1 un vettore $v_1' = v_1 + \lambda v_2$, dove λ è scelto in modo che $v_1 + \lambda v_2$ sia parallelo a e_1, primo vettore della base canonica, quindi $v_1' = ae_1 = (a, 0)$. Osserviamo che passando da $P(v_1, v_2)$ a $P(v_1', v_2)$ l'area non cambia perchè non cambia l'altezza perpendicolare a v_2.

Ripetiamo ora sostituendo a v_2 un vettore $v_2' = \mu v_1' + v_2$ con μ scelto in modo che v_2' risulti parallelo a e_2, quindi $v_2' = be_2 = (0, b)$. Questa volta non cambia l'altezza perpendicolare a v_1', dunque l'area è invariata.

Abbiamo ottenuto $v_1' = ae_1 = (a, 0)$, $v_2' = be_2 = (0, b)$, con $a, b \in \mathbb{R}$; sono i lati di un rettangolo di area $\mid a \mid\mid b \mid$. D'altra parte

$$\det \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \det \begin{pmatrix} v_1 + \lambda v_2 \\ v_2 \end{pmatrix} = \det \begin{pmatrix} v_1' \\ v_2 + \mu v_1' \end{pmatrix} = \det \begin{pmatrix} v_1' \\ v_2' \end{pmatrix} = \det \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = ab.$$

$\square$

Questo ragionamento si estende a dimensione n qualunque: si definisce per induzione il volume del parallelepipedo $P(v_1, \dots, v_n)$ come valore assoluto del determinante $\det(v_1, \dots, v_n)$: è il prodotto del volume n-dimensionale di una base per la relativa altezza.

Se $f : \mathbb{R}^2 \to \mathbb{R}^2$ è un endomorfismo del tipo $f = L(A)$, dove A una matrice 2×2, e $v_1, v_2 \in \mathbb{R}^2$ sono due vettori, si dimostra con un conto diretto che $\det(f(v_1), f(v_2)) = \det(A)\det(v_1, v_2)$. Dunque il determinante di A, che coincide con il determinante di f, misura il rapporto fra le aree dei due parallelogrammi, di lati v_1, v_2 e $f(v_1), f(v_2)$.