

Determinante

1 Definizione, esistenza e unicità, calcolo e principali proprietà

Sia $\mathbf{M}(n, n, K)$ lo spazio delle matrici quadrate $n \times n$ a coefficienti in un campo K , vogliamo provare il seguente Teorema:

Theorem 1.1 *Esiste un'unica mappa F dallo spazio delle matrici $n \times n$ a valori in K che abbia le seguenti proprietà :*

- 1) F è multilineare sulle colonne
- 2) Se la matrice A ha due colonne uguali allora $F(A) = 0$,
- 3) $F(I) = 1$ dove I è la matrice identica $n \times n$.

Questa mappa F la chiameremo determinante e la indicheremo con \det .

Qualche volta il determinante viene anche indicato scrivendo la matrice con due linee verticali simili al valore assoluto, ad esempio $\det \begin{pmatrix} 1, 2 \\ 4, 5 \end{pmatrix}$ si può anche indicare con $\begin{vmatrix} 1, 2 \\ 4, 5 \end{vmatrix}$.

Dire che F è multilineare sulle colonne vuol dire che è lineare su ogni colonna, dire che il determinante è lineare sulla i esima colonna vuol dire che se X ed Y sono vettori colonna di n componenti e α e β sono scalari in K , e se $A^1, \dots, A^{i-1}, A^{i+1}, \dots, A^n$ sono altri vettori colonna di n componenti, allora vale la formula

$$\begin{aligned} & F(A^1, A^2, \dots, A^{i-1}, \alpha X + \beta Y, A^{i+1}, \dots, A^n) \\ &= \alpha F(A^1, A^2, \dots, A^{i-1}, X, A^{i+1}, \dots, A^n) + \beta F(A^1, A^2, \dots, A^{i-1}, Y, A^{i+1}, \dots, A^n). \end{aligned}$$

Lemma 1.2 *Una mappa che soddisfi le proprietà 1) 2) è alternante sulle colonne cioè se in una matrice scambiamo due colonne la F cambia segno.*

Proof.

Fissiamo due indici i e j con $1 \leq i < j \leq n$ ed una matrice A $n \times n$. Consideriamo la matrice $A' = (A^1, \dots, A^{i-1}, A^i + A^j, A^{i+1}, \dots, A^{j-1}, A^i + A^j, A^{j+1}, \dots, A^n)$. Dalla proprietà 2) sappiamo che $F(A') = 0$ quindi per la linearità sulla i esima e sulla j esima colonna abbiamo

$$\begin{aligned}
0 &= F((A^1, \dots, A^{i-1}, A^i, A^{i+1}, \dots, A^{j-1}, A^i + A^j, A^{j+1} \dots A^n)) \\
&+ F(A^1, \dots, A^{i-1}, A^j, A^{i+1}, \dots, A^{j-1}, A^i + A^j, A^{j+1} \dots A^n) \\
&= F((A^1, \dots, A^{i-1}, A^i, \dots, A^{j-1}, A^i, A^{j+1} \dots A^n)) \\
&\quad F((A^1, \dots, A^{i-1}, A^i, \dots, A^{j-1}, A^j, A^{j+1} \dots A^n)) \\
&\quad F((A^1, \dots, A^{i-1}, A^j, \dots, A^{j-1}, A^i, A^{j+1} \dots A^n)) \\
&+ F(A^1, \dots, A^{i-1}, A^j, A^{i+1}, \dots, A^{j-1}, A^j, A^{j+1} \dots A^n).
\end{aligned}$$

Ancora della proprieta' 2) sappiamo che il primo e l'ultimo termine di questa somma sono nulli, si trova quindi

$$\begin{aligned}
0 &= F((A^1, \dots, A^{i-1}, A^i, A^{i+1}, \dots, A^{j-1}, A^j, A^{j+1} \dots A^n)) \\
&+ F(A^1, \dots, A^{i-1}, A^j, A^{i+1}, \dots, A^{j-1}, A^i, A^{j+1} \dots A^n) \\
\text{cioe' } &F((A^1, \dots, A^{i-1}, A^i, A^{i+1}, \dots, A^{j-1}, A^j, A^{j+1} \dots A^n)) \\
&= -F(A^1, \dots, A^{i-1}, A^j, A^{i+1}, \dots, A^{j-1}, A^i, A^{j+1} \dots A^n).
\end{aligned}$$

■

Lemma 1.3 *Se il campo K ha caratteristica diversa da 2 ed una mappa F e' alter-nante sulle colonne allora soddisfa la 2).*

Proof.

Se una matrice A ha due colonne uguali, se scambiamo le due colonne la matrice non cambia quindi non cambia il valore di F su quella matrice, d'altra parte per l'alternanza sulle colonne il valore di F su quella matrice cambia segno, quindi se il campo ha caratteristica diversa da due il valore di F su quella matrice e' zero. ■

Lemma 1.4 *Se una mappa F soddisfa la 1) e vale zero su ogni matrice che abbia due colonne uguali una accanto all'altra, allora F soddisfa anche la 2).*

Proof.

Ripetendo la dimostrazione del lemma 1.2 nel caso di una coppia i, j con $j = i + 1$ si vede che scambiando due colonne vicine la F cambia segno. Ora prendiamo una matrice A che abbia la colonna i esima e la colonna j esima uguali con $i < j$ ma j non necessariamente uguale ad $i + 1$. Se a partire dalla matrice A scambiamo successivamente la colonna i con la colonna $i + 1$, di questa nuova matrice scambiamo la colonna $i + 1$ con la $i + 2$ e cosi' via fino a che la colonna i esima di A non si trovi accanto alla sua colonna j esima. Applicando la F a questa sequenza di matrici a partire dalla matrice A otteniamo tutti numeri o uguali o di segno opposto, ma alla fine troveremo il valore 0. Percio' $F(A) = 0$. ■

Vogliamo provare ora l'esistenza della funzione determinante, definiamo un candidato determinante sulle matrici $n \times n$ per induzione su n .

Per $n = 1$ poniamo $\det(a) = a$, questa funzione soddisfa tutte e tre le proprietà. Supponiamo di aver definito una funzione che soddisfi tutte e tre le proprietà per tutte le matrici $(n - 1) \times (n - 1)$, vogliamo definire un candidato determinante per le matrici $n \times n$.

Sia A e' una matrice $n \times n$ e siano dati due indici i e j tra 1 ed n , indichiamo con $A(i, j)$ la matrice $(n - 1) \times (n - 1)$ ottenuta da A rimuovendo la riga i esima e la colonna j esima. Supponiamo di aver definito una F candidato determinante che soddisfi tutte e tre le proprietà per le matrici $(n - 1) \times (n - 1)$ e indichiamo con $D_{i,j}$ il valore di questa funzione applicata alla matrice $A(i, j)$ Abbiamo allora il seguente

Theorem 1.5 *Sviluppo di Laplace per righe del determinante. Fissato un indice i compreso tra 1 ed n Se prendiamo come F la funzione*

$$F(A) = (-1)^{i+1}a_{i,1}D_{i,1} + (-1)^{i+2}a_{i,2}D_{i,2} + \dots (-1)^{i+n}a_{i,n}D_{i,n}$$

questa F soddisfa tutte e tre le proprietà del teorema 1.1.

Come si vede la riga i esima della matrice resta fissa e cambiano le varie colonne, per questo si chiama sviluppo di Laplace rispetto alla i esima riga.

Example 1.6 *Nel caso di matrici 2×2 prendendo $i = 1$ troviamo $D_{1,1} = a_{2,2}$, $D_{1,2} = a_{2,1}$ e otteniamo $\det(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$.*

Proviamo ora il teorema 1.5.

Partiamo dalla proprietà 3) e consideriamo la matrice identità. Dato che l'identità ha tutti zeri fuori dalla diagonale nella somma dello sviluppo di Laplace rispetto alla riga i esima sopravvive solo il termine $(-1)^{i+i}a_{i,i}D_{i,i}$. D'altra parte se togliamo dalla matrice identità ($n \times n$) la riga i esima e la colonna i esima resta la matrice identità $(n - 1) \times (n - 1)$. Per l'ipotesi di induzione il nostro candidato determinante per le matrici $(n - 1) \times (n - 1)$ deve valere 1 sulla matrice identità, Cioè $D_{i,i} = 1$. Dato che gli elementi sulla diagonale della matrice identità sono tutti 1 otteniamo che per la matrice identità $n \times n$ si ha che $(-1)^{i+i}a_{i,i}D_{i,i} = 1$.

Per quanto riguarda la proprietà 1) fissiamo un indice k compreso tra 1 ed n . Dato un indice j tra 1 ed n diverso da k , se nella matrice A modifichiamo la k esima colonna sostituendola con una combinazione lineare di due vettori colonna e lasciamo invariate tutte le altre colonne i coefficienti $a_{i,j}$ restano invariati. D'altra parte per ipotesi induttiva i termini $D_{i,j}$ cambiano linearmente. il termine $(-1)^{i+j}a_{i,j}D_{i,j}$ cambia allora linearmente. Consideriamo ora il termine $(-1)^{i+k}a_{i,k}D_{i,k}$. Se si modifica solo la k esima colonna sostituendola con una combinazione lineare di vettori colonna, la matrice $A(i, k)$ resta invariata, quindi resta invariato anche $D_{i,k}$ mentre il coefficiente $a_{i,k}$ cambia linearmente. Quindi il termine $(-1)^{i+k}a_{i,k}D_{i,k}$ cambia linearmente. La somma di termini che cambiano linearmente cambia linearmente, questo dimostra la proprietà 1).

Usando il Lemma 1.4 vediamo che per provare la proprietà 2) sarà sufficiente provare che il nostro candidato determinante vale zero su una matrice che abbia due colonne vicine uguali. Supponiamo quindi che la matrice A sia tale che $A^k = A^{k+1}$.

Se j e' diverso da k la matrice $A(i, j)$ ha allora due colonne vicine uguali e per ipotesi induttiva $D_{i,j} = 0$. Il nostro candidato determinante risulta allora uguale a $(-1)^{i+k} a_{i,k} D_{i,k} + (-1)^{i+k+1} a_{i,k+1} D_{i,k+1}$. D'altra parte dato che $A^k = A^{k+1}$ se ne deduce che $a_{i,k} = a_{i,k+1}$. Inoltre per $1 \leq s \leq k-1$ la s esima colonne di $A(i, k)$ si ottiene togliendo il termine $a_{i,s}$ dalla colonne A^s di A , la colonna k esima della matrice $A(i, k)$ e' ottenuta togliendo il termine $a_{i,k+1}$ dalla colonna A^{k+1} , e le colonne $k+s$ esima di $A(i, k)$ si ottiene togliendo il termine $a_{i,s+1}$ dalla colonna A^{s+1} . Similmente per la matrice $A(i, k+1)$. Dato che $A^k = A^{k+1}$ risulta che $A(i, k) = A(i, k+1)$, quindi $D_{i,k} = D_{i,k+1}$. In conclusione

$$(-1)^{i+k} a_{i,k} D_{i,k} + (-1)^{i+k+1} a_{i,k+1} D_{i,k+1} = (-1)^{i+k} a_{i,k} D_{i,k} (1 - 1) = 0.$$

Abbiamo quindi verificato tutte le tre proprieta', provando cosi' l'esistenza di un candidato funzione determinante.

Vogliamo ora provare l'unicita' della funzione determinante.

Dato un intero positivo n , sia S_n il gruppo delle permutazioni cioe' delle applicazioni biunivoche dell' insieme $\{1, 2, \dots, n\}$ in se' stesso. Questo e' un gruppo rispetto all' operazione di composizione di applicazioni.

Dati due indici i e j con $1 \leq i < j \leq n$, Chiamiamo trasposizione che scambia i e j l'elemento $\tau_{i,j} \in S_n$ cosi' definito :

$\tau_{i,j}(i) = j, \tau_{i,j}(j) = i, \tau_{i,j}(k) = k$ se $k \neq i, k \neq j$. Dai corsi di algebra sappiamo che valgono i due seguenti teoremi:

Theorem 1.7 *Ogni elemento di S_n e' composizione di trasposizioni.*

Theorem 1.8 *Esiste una applicazione segno : $S_n \rightarrow \{-1, 1\}$ con le seguenti proprieta':*

$$\begin{aligned} segno(\sigma_1 \circ \sigma_2) &= segno(\sigma_1) segno(\sigma_2). \\ \text{Per ogni trasposizione } \tau, \text{ si ha } segno(\tau) &= -1 \end{aligned}$$

In particolare dal teorema 1.8 segue che $segno(Id) = 1$, e $segno(\sigma^{-1}) = \frac{1}{segno(\sigma)} = segno(\sigma)$.

Inoltre dai teoremi sopra si puo' dedurre che, benché il modo di decomporre una permutazione σ come prodotto di trasposizioni non sia unico, se una permutazione si puo' scrivere come prodotto di un numero pari, rispettivamente dispari, di trasposizioni, allora ogni altra decomposizione di σ come prodotto di trasposizioni avra' un numero pari, rispettivamente dispari di elementi.

Vediamo ora l'unicita' della funzione determinante provando la validita' di una formula per il calcolo del determinante che benché poco utilizzabile in pratica, ha alcune interessanti implicazioni

Theorem 1.9 *Sia F una funzione che soddisfi le proprieta' 1), 2) 3) del teorema 1.1 allora vale la formula :*

$$F(A) = \sum_{\sigma \in S_n} segno(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} a_{\sigma(3),3} \dots a_{\sigma(n),n}$$

In particolare questa formula mostra l'unicita' della funzione determinante.

Proof. Sia e_1, e_2, \dots, e_n la base canonica in k^n .

Dalla linearit' di F sulla prima colonna abbiamo

$$F(A) = F(A^1, A^2, \dots, A^n) = F\left(\sum_{j_1=1}^n a_{j_1,1} e_{j_1}, A^2, \dots, A^n\right) = \sum_{j_1=1}^n a_{j_1,1} F(e_{j_1}, A^2, \dots, A^n).$$

Dalla linearita' di F sulla seconda colonna abbiamo

$$F(e_{j_1}, A^2, \dots, A^n) = F\left(e_{j_1}, \sum_{j_2=1}^n a_{j_2,2} e_{j_2}, A^3, \dots, A^n\right) = \sum_{j_2=1}^n a_{j_2,2} F(e_{j_1}, e_{j_2}, A^3, \dots, A^n).$$

Procedendo allo stesso modo su le altre colonne si trova alla fine la formula:

$$F(A) = \sum_{j_1, j_2, \dots, j_n=1}^n a_{j_1,1} a_{j_2,2} \dots a_{j_n,n} F(e_{j_1}, e_{j_2}, \dots, e_{j_n}).$$

Dalla proprieta' 2) sappiamo che se nella sequenza (j_1, j_2, \dots, j_n) ci sono due indici uguali abbiamo $F(e_{j_1}, e_{j_2}, \dots, e_{j_n}) = 0$. Ora nella sequenza (j_1, j_2, \dots, j_n) non ci sono indici uguali se e solo se esiste $\sigma \in S_n$ tale che $(j_1, j_2, \dots, j_n) = (\sigma(1), \sigma(2), \dots, \sigma(n))$. Ora se σ e' una permutazione e τ una trasposizione in S_n segue dalla alternanza della F (vedi Lemma 1.2) che $F(e_{\tau(\sigma(1))}, e_{\tau(\sigma(2))}, \dots, e_{\tau(\sigma(n))}) = -F(e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)})$. D'altra parte se σ si scrive come prodotto di trasposizioni nella forma $\sigma = \tau_1 \circ \tau_2 \circ \tau_3 \dots \circ \tau_k$, allora $\tau_1 \circ \sigma = \tau_2 \circ \tau_3 \dots \circ \tau_k$. Quindi procedendo per induzione sul numero k di trasposizioni che decompongono σ si trova che $F(e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}) = \text{segno}(\sigma) F(e_1, \dots, e_n)$. In fine per la proprieta' 3) $F(e_1, \dots, e_n) = 1$. ■

Theorem 1.10 *Vale la formula*

$$\sum_{\sigma \in S_n} \text{segno}(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} a_{\sigma(3),3} \dots a_{\sigma(n),n} = \sum_{\sigma \in S_n} \text{segno}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} a_{3,\sigma(3)} \dots a_{n,\sigma(n)}.$$

Se ne deduce che $\det(A) = \det(A^t)$.

Proof. Dato che la trasformazione $\sigma \rightarrow \sigma^{-1}$ da S_n in se' e' iniettiva e suriettiva, abbiamo:

$$\sum_{\sigma \in S_n} \text{segno}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} a_{3,\sigma(3)} \dots a_{n,\sigma(n)} = \sum_{\sigma \in S_n} \text{segno}(\sigma^{-1}) a_{1,\sigma^{-1}(1)} a_{2,\sigma^{-1}(2)} a_{3,\sigma^{-1}(3)} \dots a_{n,\sigma^{-1}(n)}.$$

Se $\sigma^{-1}(h) = k$ allora $a_{h,\sigma^{-1}(h)} = a_{\sigma(k),k}$ da cui segue che per ogni $\sigma \in S_n$ abbiamo

$$a_{1,\sigma^{-1}(1)} a_{2,\sigma^{-1}(2)} a_{3,\sigma^{-1}(3)} \dots a_{n,\sigma^{-1}(n)} = a_{\sigma(1),1} a_{\sigma(2),2} \dots a_{\sigma(n),n}.$$

Ricordando che $\text{segno}(\sigma) = \text{segno}(\sigma^{-1})$ si conclude la dimostrazione del teorema. ■

Sfruttando il teorema 1.10 e con dimostrazione analoga a quella del lemma 1.2 si ottiene subito il seguente corollario:

Corollary 1.11 *La funzione determinante e' multilineare sulle righe, alternante sulle righe, vale zero sulle matrici che abbiano due righe uguali e vale 1 sulla matrice identita'.*

Con una dimostrazione analoga a quella del teorema 1.9 si vede che la funzione determinante e' l'unica che soddisfa le propriet' del corollario 1.11.

Inoltre dal teorema 1.10 e dal teorema 1.12 si ottiene subito il seguente:

Theorem 1.12 *Sviluppo di Laplace per colonne del determinante. Fissato un indice i compreso tra 1 ed n Si ha*

$$\det(A) = (-1)^{1+i}a_{1,i}D_{1,i} + (-1)^{2+i}a_{2,i}D_{2,i} + \dots (-1)^{n+i}a_{n,i}D_{n,i}.$$

Esempio

Se A e' una matrice $n \times n$ e λ e' un elemento di K , allora $\det(\lambda A) = \lambda^n \det(A)$. Infatti le colonne di λA sono $\lambda A^1, \lambda A^2, \dots, \lambda A^n$.

Per linearita' sulla prima colonna abbiamo

$$\det(\lambda A^1, \lambda A^2, \dots, \lambda A^n) = \lambda \det(A^1, \lambda A^2, \dots, \lambda A^n).$$

Per linearita' sulla seconda colonna abbiamo

$$\det(\lambda A^1, \lambda A^2, \dots, \lambda A^n) = \lambda^2 \det(A^1, A^2, \lambda A^3, \dots, \lambda A^n).$$

Così procedendo alla fine otteniamo $\det(\lambda A) = \lambda^n \det(A)$.

Una matrice quadrata A si dice triangolare superiore se tutti i suoi elementi al di sotto della diagonale sono nulli, cio' se $a_{i,j} = 0$ quando $i > j$.

Proposition 1.13 *Se A e' una matrice $n \times n$ triangolare superiore allora $\det(A) = a_{1,1}a_{2,2}a_{3,3} \dots a_{n,n}$.*

Proof. Procediamo per induzione su n , per $n = 1$ il risultato e' immediato, Ammettiamo il risultato vero per matrici triangolari superiori $(n-1) \times (n-1)$ e sia A una matrice triangolare superiore $n \times n$. Facendo lo sviluppo di Laplace del determinante di A secondo la prima riga otteniamo $\det(A) = (-1)^{1+1}a_{1,1}D_{1,1} + (-1)^{1+2}a_{1,2}D_{1,2} + \dots (-1)^{1+n}a_{1,n}D_{1,n}$. D'altra parte ogni matrice $A(1, j)$ per $j > 1$ ha la prima colonna nulla, quindi per linearita' del determinante sulla prima colonna $\det(A(1, j)) = D_{1,j} = 0$ per $j > 1$. Percio' $\det(A) = a_{1,1}D_{1,1}$. D'altra parte $D_{1,1} = \det(A(1, 1))$ e' la matrice $A(1, 1)$ e' una matrice $(n-1) \times (n-1)$ triangolare superiore con elementi sulla diagonale $a_{2,2}, a_{3,3} \dots a_{n,n}$. Allora per ipotesi induttiva otteniamo $D_{1,1} = a_{2,2}a_{3,3} \dots a_{n,n}$ e in conclusione $\det(A) = a_{1,1}a_{2,2}a_{3,3} \dots a_{n,n}$. ■

Example 1.14 *Sia S una matrice $n \times n$ a scala, se l'ultima riga di S e' nulla $\det(S) = 0$, altrimenti la matrice S ha n pivot (tutti non nulli), e' una matrice triangolare superiore e i suoi pivot si trovano sulla diagonale. Quindi $\det(S) = p_1 p_2 \dots p_n \neq 0$.*

Infatti se S ha l'ultima riga nulla per linearita' del determinante sulle righe (corollario 1.11), si ha che $\det(S) = 0$. Se l'ultima riga di S non e' nulla la matrice ha rango n ed e' una matrice triangolare superiore con i pivot sulla diagonale, la formula $\det(S) = p_1 p_2 \dots p_n$ segue allora dalla proposizione 1.13.

Proposition 1.15 *Sia A una matrice $n \times n$ e sia S una sua ridotta a scala attraverso il procedimento di eliminazione di Gauss. Se l'ultima riga di S e' nulla $\det(A) = 0$, altrimenti $\det(A) = (-1)^m p_1 p_2 \dots p_n \neq 0$. Dove p_1, p_2, \dots, p_n sono i pivot di S ed m e' il numero di scambi di righe che e' stato necessario per passare dalla matrice A alla matrice S .*

Proof. Fissati due indici i e j con $1 \leq i < j \leq n$ abbiamo chiamato operazione elementare sulle righe il passaggio da una matrice A con righe A_1, A_2, \dots, A_n ad una matrice A' le cui righe siano $A_1, A_2, \dots, A_{i-1}, A_i, \dots, A_{j-1}, \lambda A_i + A_j, A_{j+1}, \dots, A_n$ con $\lambda \in K$. Ora per multilinearita' del determinante sulle righe,

$$\det(A') = \lambda \det(A_1, A_2, \dots, A_{i-1}, A_i, A_{i+1}, \dots, A_{j-1}, A_i, A_{j+1}, \dots, A_n) + \det(A) = \det(A)$$

perche' una matrice con due righe uguali ha determinante nullo. D'altra parte se scambiamo due righe in una matrice lasciando le altre invariate il determinante cambia segno. Ora il passaggio dalla matrice A alla matrice a scala S attraverso l'eliminazione di Gauss si ottiene con una sequenza di operazioni elementari e di scambi di righe, quindi $\det(S) = (-1)^m \det(A)$. La proposizione segue allora dall'esempio 1.14. ■

Proposition 1.16 *Una matrice A $n \times n$ ha rango n se e solo se $\det(A) \neq 0$.*

Proof. Sia S una ridotta a scala di Gauss di A , allora A ha rango n se e solo se S ha n pivot, se e solo se l'ultima riga di S non e' nulla, la proposizione segue allora dalla proposizione 1.15.

■

Theorem 1.17 *Siano A e B matrici $n \times n$, allora $\det(AB) = \det(A)\det(B)$.*

Proof.

Supponiamo che $\det(A) = 0$, allora $\det(A)\det(B) = 0$. Dobbiamo verificare che $\det(AB) = 0$. Ora se $\det(A) = 0$ il rango di A e' minore di n , Ricordando che il rango per righe coincide con il rango per colonne, troviamo che le righe di A sono linearmente dipendenti, cioe' esiste una combinazione lineare $\lambda_1 A_1 + \lambda_2 A_2, \dots, + \lambda_n A_n = 0$ e $\lambda_i \neq 0$ per qualche i tra 1 ed n . Le righe di AB sono $A_1 B, A_2 B, \dots, A_n B$ e vale $\lambda_1 A_1 B + \lambda_2 A_2 B, \dots, + \lambda_n A_n B = 0$ con $\lambda_i \neq 0$, quindi il rango di AB e' minore di n e $\det(AB) = 0$. Se $\det(A) \neq 0$ definiamo

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

Allora si ha che $F(I) = 1$ dove I e' la matrice identita'. Inoltre se X ed Y sono vettori colonna di n componenti e α e β sono scalari in K , e se $B^1, \dots, B^{i-1}, B^{i+1}, \dots, B^n$ sono altri vettori colonna di n componenti, allora vale la formula

$$\begin{aligned} & F\left(B^1, B^2, \dots, B^{i-1}, \alpha X + \beta Y, B^{i+1}, \dots, B^n\right) \\ &= \frac{\det\left(B^1, B^2, \dots, B^{i-1}, \alpha X + \beta Y, B^{i+1}, \dots, B^n\right)}{\det(A)} \end{aligned}$$

$$\begin{aligned}
&= \frac{\alpha \det(B^1, B^2, \dots, B^{i-1}, X, B^{i+1}, \dots, B^n)}{\det(A)} + \frac{\beta \det(A^1, A^2, \dots, A^{i-1}, Y, A^{i+1}, \dots, A^n)}{\det(A)} \\
&= \alpha F(B^1, B^2, \dots, B^{i-1}, X, B^{i+1}, \dots, B^n) + \beta F(B^1, B^2, \dots, B^{i-1}, Y, B^{i+1}, \dots, B^n).
\end{aligned}$$

cioè F è multilineare sulle colonne. In fine se B^1, B^2, \dots, B^n sono le colonne di B allora AB^1, AB^2, \dots, AB^n sono le colonne di AB . Quindi se $B^i = B^j$ per $1 \leq i < j \leq n$, anche $AB^i = AB^j$, quindi $\det(AB) = 0$, da cui $F(B) = 0$. Dall'unicità della funzione che soddisfa le proprietà del teorema 1.1 si trova che $\frac{\det(AB)}{\det(A)} = \det(B)$. ■

2 Uso del determinante

Ricordiamo che una matrice A $n \times n$ si dice invertibile se e solo se esiste una matrice B $n \times n$ tale che $AB = BA = I$. In tal caso la matrice B si dice inversa di A .

Proposition 2.1 *Se l'inversa di una matrice A esiste essa è unica.*

Proof.

Se $AB_1 = B_1A = I$ e $AB_2 = B_2A = I$, allora $B_2AB_1 = B_2(AB_1) = B_2I = B_2$. Mentre $B_2AB_1 = (B_2A)B_1 = IB_1 = B_1$. ■

L'inversa di A quando esiste si indica con A^{-1} .

Theorem 2.2 *Sia A una matrice $n \times n$, allora i fatti seguenti sono equivalenti:*

- 1) Il rango di A è n
- 2) Le righe di A sono linearmente indipendenti
- 3) Le colonne di A sono linearmente indipendenti
- 4) L'unica soluzione del sistema lineare omogeneo $AX = 0$ è il vettore $X = 0$
- 5) Per ogni vettore colonna b di n componenti, esiste una ed una sola soluzione del sistema lineare non omogeneo $AX = b$
- 6) La matrice A è invertibile
- 7) $\det(A) \neq 0$.

Proof. Il rango per colonne di A è uguale al rango per righe, quindi 1) equivale a 2) equivale a 3).

Sappiamo dalla proposizione 1.16 che 1) equivale a 7).

Ricordiamo che se X è un vettore colonna con componenti x_1, x_2, \dots, x_n , allora $AX = x_1A^1 + x_2A^2 + \dots + x_nA^n$. Quindi 4) equivale a 3).

Ovviamente 5) implica 4) implica 3). La 3) implica che gli n vettori colonna di A sono una base di K^n e, grazie all'osservazione precedente questo implica 5). Quindi 1), 2), 3), 4), 5) e 7) sono equivalenti.

Supponiamo sia vera la 6) e sia X un vettore colonna tale che $AX = 0$, allora $X = IX = A^{-1}(AX) = 0$ vale quindi la 4). Se vale la 4) vale la 5). Siano e_1, e_2, \dots, e_n i vettori colonna della base canonica di K^n . Siano b^1, b^2, \dots i vettori soluzione dei sistemi $Ab^i = e^i$, (che esistono e sono unici grazie alla 5), e sia B la matrice che ha per colonne i vettori b^1, b^2, \dots, b^n . Abbiamo allora la relazione $AB = I$. D'altra parte

se vale la 5) per la matrice A vale la 7) quindi $\det(A) \neq 0$, e dalla proposizione 1.10 $\det(A^t) \neq 0$ e deve valere la 5) per A^t . Dalle considerazioni sopra sappiamo che esiste una matrice C $n \times n$ tale che $A^t C = I$ e facendo la trasposta troviamo $C^t A = I$. Poniamo $B_1 = B$, $B_2 = C^t$, abbiamo allora: Quindi $B_2 A B_1 = (B_2 A) B_1 = I B_1 = B_1$, mentre $B_2 A B_1 = B_2 (A B_1) = B_2 I = B_2$. Quindi vale la 6). ■

Grazie al teorema precedente sappiamo che se $\det(A) \neq 0$ ogni sistema lineare non omogeneo $AX = b$ deve avere una ed una sola soluzione. Vogliamo ora determinare questa soluzione usando il determinante.

Theorem 2.3 *Teorema di Cramer* sia A una matrice $n \times n$ con $\det(A) \neq 0$, sia b un vettore colonna in n componenti, e sia X l'unica soluzione (grazie al teorema 2.2) del sistema lineare non omogeneo $AX = b$. Siano x_1, x_2, \dots, x_n le componenti di X . Vale allora per ogni i compreso tra 1 ed n , la formula

$$x_i = \frac{\det(A^1, A^2, \dots, A^{i-1}, b, A^{i+1}, \dots, A^n)}{\det(A)}.$$

Proof.

Come abbiamo precedentemente osservato possiamo scrivere

$$b = AX = x_1 A^1 + x_2 A^2 \dots + x_n A^n$$

quindi per la multilinearità del determinante sulle colonne abbiamo:

$$\begin{aligned} & \det(A^1, A^2, \dots, A^{i-1}, b, A^{i+1}, \dots, A^n) \\ &= \det(A^1, A^2, \dots, A^{i-1}, x_1 A^1 + x_2 A^2 \dots + x_n A^n, A^{i+1}, \dots, A^n) \\ &= x_1 \det(A^1, A^2, \dots, A^{i-1}, A^1, A^{i+1}, \dots, A^n) \\ &+ x_2 \det(A^1, A^2, \dots, A^{i-1}, A^2, A^{i+1}, \dots, A^n) + \dots \\ &+ x_n \det(A^1, A^2, \dots, A^{i-1}, A^n, A^{i+1}, \dots, A^n). \end{aligned}$$

D'altra parte se k è diverso da i la matrice $(A^1, A^2, \dots, A^{i-1}, A_k, A^{i+1}, \dots, A^n)$ ha due colonne uguali, quindi $x_k \det(A^1, A^2, \dots, A^{i-1}, b, A^{i+1}, \dots, A^n) = 0$. Se ne ricava quindi la relazione

$$\det(A^1, A^2, \dots, A^{i-1}, b, A^{i+1}, \dots, A^n) = x_i \det(A).$$

■

Il teorema di Cramer ci permette anche di ricavare una formula per il calcolo della matrice inversa di una matrice A $n \times n$ con $\det(A) \neq 0$. in termini di determinanti. Tale matrice è infatti invertibile grazie al teorema 2.2.

Ricordiamo che data una matrice A $n \times n$ abbiamo indicato con $A(i, j)$ la matrice $(n-1) \times (n-1)$ ottenuta togliendo da A la riga i esima e la colonna j esima ed abbiamo indicato con $D_{i,j}$ il determinante di $A(i, j)$. Ora indichiamo con P la matrice che come componente i, j abbia il numero $(-1)^{i+j} D_{i,j}$. ed indichiamo con $Agg(A)$ (matrice aggiunta di A), la matrice P^t .

Theorem 2.4 *Se A e' una matrice $n \times n$ con determinante non nullo (quindi invertibile), vale la formula*

$$A^{-1} = \frac{\text{Agg}(A)}{\det(A)} = \frac{P^t}{\det(A)}.$$

Proof.

Siano e_1, e_2, \dots, e_n i vettori della base canonica di K^n . Sia $B = A^{-1}$ e siano b^1, b^2, \dots, b^n le colonne di B . Per ogni i tra 1 ed n e' allora soddisfatto il sistema lineare $Ab^i = e_i$. Osserviamo ora che il termine $b_{j,i}$ della matrice B e' la j esima componente del vettore colonna b^i . Per il teorema di Cramer si ha

$$b_{j,i} = \frac{\det(A^1, A^2, \dots, A^{j-1}, e_i, A^{j+1}, \dots, A^n)}{\det(A)} = \frac{\det(C)}{\det(A)}$$

dove $C = (A^1, A^2, \dots, A^{i-1}, e_i, A^{i+1}, \dots, A^n)$. Calcoliamo ora il determinante di C usando lo sviluppo di Laplace rispetto alla i esima riga. Tutte le matrici $C(i, k)$ con k diverso da j hanno una colonna nulla, quindi hanno determinante 0. Percio' vale la formula

$$\det(C) = (-1)^{i+j} c_{i,j} \det(C(i, j)). \text{ Ora } c_{i,j} = 1, \text{ e } C(i, j) = A(i, j).$$

$$\text{In altre parole } B^t = \frac{P}{\det(A)}. \blacksquare$$

Esempio

Consideriamo una matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

tale che $\det(A) = ad - bc \neq 0$ allora

$$\text{Agg}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

cioe' $\text{Agg}(A)$ si ottiene scambiando gli elementi sulla diagonale di A e lasciando gli altri al loro posto ma con il segno cambiato. Quindi

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Sia A una matrice $m \times n$ (quindi non necessariamente quadrata) siano dati due interi p e q con $1 \leq p \leq m$, $1 \leq q \leq n$. Scegliamo p indici i_1, \dots, i_p e q indici j_1, \dots, j_q tali che $1 \leq i_1 < i_2 < i_3 \dots i_p \leq m$ e $1 \leq j_1 < j_2 < j_3 < \dots < j_q \leq n$. L'insieme degli elementi della matrice A che si trovano simultaneamente sulle righe i_1, i_2, \dots, i_p e sulle colonne j_1, j_2, \dots, j_q formano una matrice B $p \times q$ "contenuta" nella matrice A . Ogni matrice di questo tipo si dice "sottomatrice" di A .

Lemma 2.5 *Sia A una matrice $m \times n$ e sia k un intero tale che $1 \leq k \leq m$, e $1 \leq k \leq n$. Sia B una sottomatrice $k \times k$ di A . Siano $1 \leq i_1 < i_2 < i_3 < i_k \leq m$ gli indici delle righe di A che contengono qualche elemento di B , e siano $1 \leq j_1 < j_2 < j_k \leq n$ gli indici delle colonne che contengono qualche elemento di B . Se $\det(B) \neq 0$ allora le righe $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ sono linearmente indipendenti, e similmente le colonne $A^{j_1}, A^{j_2}, \dots, A^{j_k}$ sono linearmente indipendenti.*

Proof.

Se del vettore riga A_{i_1} consideriamo solo gli elementi che si trovano sulle colonne di indice j_1, j_2, \dots, j_k otteniamo il vettore riga B_1 della matrice B , se del vettore riga A_{i_2} consideriamo solo gli elementi che si trovano sulle colonne di indice j_1, j_2, \dots, j_k otteniamo il vettore riga B_2 della matrice B , e così' via. Supponiamo allora di avere una combinazione lineare

$$\lambda_1 A_{i_1} + \lambda_2 A_{i_2} + \lambda_3 A_{i_3} + \lambda_k A_{i_k} = 0.$$

Se dei vari vettori riga consideriamo solo gli elementi che si trovano sulle colonne di indice j_1, j_2, \dots, j_k otteniamo la combinazione lineare

$$\lambda_1 B_1 + \lambda_2 B_2 + \lambda_3 B_3 + \dots \lambda_k B_k = 0.$$

Dato che $\det(B) \neq 0$, per il teorema 2.2 i vettori riga di B sono linearmente indipendenti, quindi $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$. Percio' i vettori A_{i_1}, \dots, A_{i_k} sono linearmente indipendenti. Similmente se del vettore colonna A^{j_1} consideriamo solo gli elementi che si trovano sulle righe di indice i_1, i_2, \dots, i_k otteniamo il vettore colonna B^1 , se del vettore colonna A^{j_2} consideriamo solo gli elementi che si trovano sulle righe di indice i_1, i_2, \dots, i_k otteniamo il vettore colonna B^2 , e così' via. Supponiamo allora di avere una combinazione lineare

$$\lambda_1 A^{j_1} + \lambda_2 A^{j_2} + \lambda_3 A^{j_3} + \lambda_k A^{j_k} = 0.$$

Se dei vari vettori colonna consideriamo solo gli elementi che si trovano sulle righe di indice i_1, i_2, \dots, i_k otteniamo la combinazione lineare

$$\lambda_1 B^1 + \lambda_2 B^2 + \lambda_3 B^3 + \dots \lambda_k B^k = 0.$$

Dato che $\det(B) \neq 0$, per il teorema 2.2 i vettori colonna di B sono linearmente indipendenti, quindi $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$. Percio' i vettori A^{j_1}, \dots, A^{j_k} sono linearmente indipendenti.

■

Ad esempio se A e' una matrice $m \times n$ non nulla ogni suo elemento non nullo si puo' vedere come una sottomatrice 1×1 di A con determinante non nullo.

Il precedente Lemma ci permette di dimostrare il:

Theorem 2.6 *Teorema degli orlati*

Sia A una matrice $m \times n$ di rango r , allora $r = 0$ se e solo se A e' la matrice nulla. Altrimenti sia B una sottomatrice $h \times h$ di A con $\det(B) \neq 0$. Allora $h \leq r$, inoltre se $h < r$ esiste una sottomatrice B' $(h+1) \times (h+1)$ di A con $\det(B') \neq 0$ tale che B sia una sottomatrice di B' . Iterando il procedimento si ottiene alla fine una sottomatrice \tilde{B} $r \times r$ di A con $\det(\tilde{B}) \neq 0$ e tale che B sia sottomatrice di \tilde{B} . In particolare

$$r = \max\{h : \text{esiste una sottomatrice } B \text{ } h \times h \text{ di } A \text{ tale che } \det(B) \neq 0\}.$$

Piu' precisamente sia B una sottomatrice $h \times h$ di A con $\det(B) \neq 0$. Allora

$$r = \max\{l : \text{esiste una sottomatrice } B' \text{ } l \times l \text{ di } A \text{ tale che } \det(B') \neq 0$$

e tale che B e' una sottomatrice di B' \}.

Proof. Chiaramente $r = 0$ se e solo se $A = 0$. Supponiamo allora $A \neq 0$. Siano $A_{i_1}, A_{i_2}, \dots, A_{i_h}$ le righe di A che contengono qualche elemento di B , Per il lemma 2.5 le righe $A_{i_1}, A_{i_2}, \dots, A_{i_h}$ sono linearmente indipendenti, quindi h e' minore o uguale alla dimensione dello sottospazio vettoriale U di K^n generato dalle righe di A , cioe' dato che r e' la dimensione di U otteniamo $h \leq r$. Se $h < r$ esiste una riga A_{i_0} di A con i_0 diverso da i_1, i_2, \dots, i_h tale che le righe $A_{i_0}, A_{i_1}, \dots, A_{i_h}$ siano linearmente indipendenti. Sia C la sottomatrice $(h+1) \times n$ di A le cui righe siano $A_{i_0}, A_{i_1}, A_{i_2}, \dots, A_{i_h}$. Dato che le righe di C sono linearmente indipendenti il rango di C e' $h+1$. Inoltre B e' una sottomatrice di C con determinante non nullo. Siano $C^{j_1}, C^{j_2}, \dots, C^{j_h}$ le colonne di C che contengono qualche elemento di B , sempre per il lemma 2.5 le colonne $C^{j_1}, C^{j_2}, \dots, C^{j_h}$ sono linearmente indipendenti. Dato che il rango di C e' $h+1$ per lo stesso ragionamento fatto in precedenza, esiste una colonna C^{j_0} di C con j_0 diverso da j_1, \dots, j_h tale che le colonne $C^{j_0}, C^{j_1}, \dots, C^{j_h}$ siano linearmente indipendenti. Sia D la sottomatrice $(h+1) \times (h+1)$ di C le cui colonne sono $C^{j_0}, C^{j_1}, C^{j_2}, \dots, C^{j_h}$. Dato che le colonne di D sono linearmente indipendenti, per il teorema 2.2 il determinante di D e' non nullo. Inoltre D e' una sottomatrice di A e B e' una sottomatrice di D . Ponendo allora $B' = D$ concludiamo la dimostrazione. ■

Remark 2.7 *Sia A una matrice $m \times n$ di rango $r > 0$. Il teorema degli orlati permette quindi di calcolare r con un procedimento iterativo. Si parte da una sottomatrice B $h \times h$ di A con $\det(B)$ diverso da zero. Siano A_{i_1}, \dots, A_{i_h} le righe di A che contengono qualche elemento di B , e siano A^{j_1}, \dots, A^{j_h} le colonne di A che contengono qualche elemento di B . Sia i compreso tra 1 ed m diverso da i_1, \dots, i_h , e sia j compreso tra 1 ed n diverso da j_1, \dots, j_h . Esiste un'unica sottomatrice B' di A , $(h+1) \times (h+1)$ tale che B sia sottomatrice di B' e tale che B' abbia elementi sulla i esima riga e sulla j esima colonna di A . La matrice B' viene detta sottomatrice orlata di B . Di ogni tale sottomatrice orlata di B si calcola il determinante, se si ottiene sempre il risultato 0 il rango di A e' h , altrimenti una di tali sottomatrici $(h+1) \times (h+1)$, chiamiamola B^* ha determinante non nullo. Possiamo ora ripetere il procedimento e considerare tutte le sottomatrici $(h+2) \times (h+2)$ di A orlate di B^* e cos' via finche' non arriviamo alla grandezza massima possibile delle sottomatrici a determinante non zero che siano ottenute a partire da B con una successione di orlature. Tale grandezza massima e' il rango di A .*

3 Volume

Sia V uno spazio vettoriale reale di dimensione n con un prodotto scalare definito positivo. Sia $\beta = \{v_1, v_2, \dots, v_n\}$ una base di V , chiamiamo parallelepipedo generato dalla base β il sottoinsieme P_β cosi' definito:

$$P_\beta = \{X \in \mathbf{R}^n : X = t_1 v_1 + \dots + t_n v_n \text{ con la condizione } 0 \leq t_j \leq 1 \text{ per } 1 \leq j \leq n\}.$$

Vogliamo definire la nozione di volume n dimensionale di un parallelepipedo.

Per $n = 1$ per $\beta = \{v_1\}$, P_β e' un segmento ed e' ragionevole definire il volume come la lunghezza del segmento cioe' $vol(P_\beta) = \|v_1\|$.

Per $n = 2$, $\beta = \{v_1, v_2\}$, e' ragionevole definire volume bidimensionale di P_β come l'area di P_β . Sia u la proiezione ortogonale di v_2 sulla retta generata da v_1 e sia $h = v_2 - u$ la relativa altezza. La geometria elementare ci dice che $\text{vol}(P_\beta) = \|v_1\| \|h\|$. In dimensione 3 il volume e' l'usuale volume tridimensionale.

Ispirandoci a questi casi consideriamo la base $\beta = \{v_1, v_2, \dots, v_n\}$, e il sottoinsieme $\beta' = \{v_1, v_2, \dots, v_{n-1}\}$, se W e' il sottospazio vettoriale di V generato da β' , allora β' e' una base di W . Definiamo allora induttivamente

$\text{vol}(P_\beta) = \|h\| \text{vol}(P_{\beta'})$, dove $h = v_n - u$ ed u e' la proiezione ortogonale di v_n sul sottospazio W .

Ora ricordiamo il procedimento di ortogonalizzazione e ortonormalizzazione di Gram Schmidt della base $\beta = \{v_1, v_2, \dots, v_n\}$.

Per costruire la ortogonalizzata w_1, \dots, w_n di β si procede induttivamente nel modo seguente:

$w_1 = v_1$, $w_{k+1} = v_{k+1} - u_{k+1}$. Dove u_{k+1} e' la proiezione ortogonale di v_{k+1} sul sottospazio vettoriale generato di w_1, \dots, w_k . In fine per trovare l'ortonormalizzata si pone $z_j = \frac{w_j}{\|w_j\|}$ per $1 \leq j \leq n$.

Proposition 3.1 *Sia β una base di V e sia w_1, \dots, w_n l'ortogonalizzata di Gram Schmidt di β , allora $\text{vol}(P_\beta) = \|w_1\| \|w_2\| \|w_3\| \dots \|w_n\|$.*

Proof.

L'affermazione e' ovviamente vera se la dimensione n di V e' 1, procediamo allora per induzione su n . Se l'asserto e' vero per tutti gli spazi vettoriali di dimensione $n - 1$ con prodotto scalare definito positivo e se $\beta' = v_1, \dots, v_{n-1}$, allora per ipotesi induttiva e per la costruzione dell'ortogonalizzato di Gram-Schmidt $\text{vol}(P_{\beta'}) = \|w_1\| \|w_2\| \dots \|w_{n-1}\|$. D'altra parte se come base del parallelepipedo P_β scegliamo $P_{\beta'}$ allora la relativa altezza e' proprio il vettore w_n . ■

Theorem 3.2 *Sia β una base di V e sia P_β il parallelogramma associato a β , sia k_1, \dots, k_n una base ortonormale di V , e sia S la matrice di componenti $s_{i,j}$ dove $v_i = \sum_{j=1}^n s_{j,i} k_j$. Allora il volume di P_β e' il valore assoluto del determinante della matrice S , cioe' $\text{vol}(P_\beta) = |\det(S)|$.*

Abbiamo bisogno di dimostrare preliminarmente alcuni lemmi

Lemma 3.3 *Siano $\beta_1 = \{u_1, u_2, \dots, u_n\}$ e $\beta_2 = \{h_1, h_2, \dots, h_n\}$ due basi ortonormali di V , sia S la matrice di componenti $s_{i,j}$ dove $h_i = \sum_{j=1}^n s_{j,i} u_j$ allora vale la formula $S^t S = I$ in particolare $\det(S) = \pm 1$.*

Proof.

Dalla formula $h_i = \sum_{j=1}^n s_{j,i} u_j$, facendo il prodotto scalare si ottiene

$$\langle h_k, h_m \rangle = \left\langle \sum_{j=1}^n s_{j,k} u_j, \sum_{l=1}^n s_{l,m} u_l \right\rangle = \sum_{j=1}^n s_{j,k} s_{j,m}$$

dove abbiamo usato la relazione $\langle u_k, u_k \rangle = 1$ e $\langle u_k, u_l \rangle = 0$ se $l \neq k$. D'altra parte $\sum_{j=1}^n s_{j,k} s_{j,m} = (S^t S)_{m,k}$. D'altra parte abbiamo anche $\langle h_k, h_k \rangle = 1$ e $\langle h_k, h_m \rangle = 0$ se $h \neq k$, quindi si ottiene $S^t S = I$. Dato che $\det(S^t) = \det(S)$ e $\det(I) = 1$ troviamo $1 = \det(S^t S) = \det(S^t) \det(S) = (\det(S))^2$. ■

Lemma 3.4 *Siano $\beta_1 = \{v_1, v_2, \dots, v_n\}$, $\beta_2 = \{w_1, w_2, \dots, w_n\}$ e $\beta_3 = \{r_1, r_2, \dots, r_n\}$ basi di V , sia S la matrice di componenti $s_{i,j}$ dove $v_i = \sum_{j=1}^n s_{j,i} w_j$, sia P la matrice di componenti $p_{i,j}$ dove $w_i = \sum_{j=1}^n p_{j,i} r_j$, e sia Q la matrice di componenti $q_{i,j}$ dove $v_i = \sum_{j=1}^n q_{j,i} r_j$, allora si ha $Q = SP$.*

Proof. La matrice S e' la matrice associata all'applicazione identita' da V in V rispetto alle basi β_1 in partenza e β_2 in arrivo, mentre la matrice P e' la matrice associata all'applicazione identita' da V in V rispetto alle basi β_2 in partenza e β_3 in arrivo. La matrice Q e' invece la matrice associata all'applicazione identita' da V in V rispetto alle basi β_1 in partenza e β_3 in arrivo, la formula enunciata nel lemma segue dalla corrispondenza tra composizione di applicazioni lineari e prodotti di matrici. ■

Dimostriamo ora il teorema 3.2. Sia $\{w_1, \dots, w_n\}$ la base ortogonalizzata con Gram-Schmidt della base $\beta = \{v_1, v_2, \dots, v_n\}$. Abbiamo allora per costruzione $w_i = \sum_{j=1}^n t_{j,i} v_j$. Per costruzione la matrice T di componenti $t_{i,j}$ e' triangolare superiore ed ha tutti 1 sulla diagonale. Sia ora $z_i = \frac{w_i}{\|w_i\|}$ la corrispondente base ortonormalizzata. Abbiamo allora $z_i = \sum_{j=1}^n \lambda_{j,i} w_j$ dove la matrice Λ di componenti $\lambda_{i,j}$ e' una matrice diagonale tale che $\lambda_{i,i} = \frac{1}{\|w_i\|}$. Sia M la matrice di componenti $m_{i,j}$ tale che $z_i = \sum_{j=1}^n m_{j,i} v_j$, e sia D la matrice di componenti $d_{i,j}$ tale che $z_i = \sum_{j=1}^n d_{j,i} k_j$. Per il Lemma 3.4 $MS = D$. Sempre per il Lemma 3.4 sappiamo che $M = \Lambda T$. Quindi $\Lambda TS = D$. calcolandone ora il determinante troviamo $\det(\Lambda)\det(T)\det(S) = \det(D)$. Per la proposizione 1.13 sappiamo che $\det(T) = 1$ e che

$$\det(\Lambda) = \frac{1}{\|w_1\| \|w_2\| \|w_3\| \dots \|w_n\|}.$$

Per il lemma 3.3 sappiamo che $\det(D) = \pm 1$. In conclusione troviamo $\det(S) = \pm(\|w_1\| \|w_2\| \|w_3\| \dots \|w_n\|)$. La dimostrazione si conclude allora usando la proposizione 3.1.

Example 3.5 *Consideriamo lo spazio vettoriale \mathbf{R}^n con il prodotto scalare canonico, Se $\beta = v_1, v_2, \dots, v_n$ e' una base di \mathbf{R}^n e se S e' la matrice che ha per colonne i vettori v_1, v_2, \dots, v_n , allora il volume di P_β e' il valore assoluto del determinante di S . Infatti basta applicare il teorema 3.2 scegliendo come base ortonormale k_1, \dots, k_n la base canonica di \mathbf{R}^n .*

Corollary 3.6 *Sia β una base di V e sia P_β il parallelogramma associato a β , sia G la matrice di componenti $g_{i,j} = \langle v_i, v_j \rangle$ allora $\det(G) > 0$ e $\text{vol}(P_\beta) = \sqrt{\det(G)}$.*

.

Proof. Scegliamo una base ortonormale k_1, \dots, k_n di V . Sia $v_\alpha = \sum_{j=1}^n s_{\alpha,j} k_j$. Allora

$$g_{\alpha,\beta} = \langle v_\alpha, v_\beta \rangle = \left\langle \sum_{j=1}^n s_{\alpha,j} k_j, \sum_{l=1}^n s_{\beta,l} k_l \right\rangle = \sum_{j,l=1}^n s_{\alpha,j} s_{\beta,l} \langle k_j, k_l \rangle.$$

Nel passaggio precedente abbiamo usato la bilinearita' del prodotto scalare. Dato che la base k_1, \dots, k_n e' ortonormale abbiamo $\langle k_j, k_l \rangle = 0$ se j e' diverso da l e $\langle k_j, k_j \rangle = 1$. Quindi si ottiene

$$g_{\alpha, \beta} = \sum_{j=1}^n s_{\alpha, j} s_{\beta, j} = (SS^t)_{\alpha, \beta}.$$

Cioe' $G = SS^t$. Dato che S e' la matrice di passaggio tra due basi $\det(S) \neq 0$ e $\det(G) = \det(SS^t) = \det(S)\det(S^t) = \det(S)^2 > 0$. Inoltre per il teorema 3.2 $\text{vol}(P_\beta) = |\det(S)| = \sqrt{\det(S)^2} = \sqrt{\det(G)}$.

■

4 Prodotto vettoriale in \mathbf{R}^3

Sia V uno spazio vettoriale su \mathbf{R} con prodotto scalare $\langle \rangle$ definito positivo, iniziamo con un paio di lemmi:

Lemma 4.1 *Siano X ed Y vettori di V , allora $\langle X, v \rangle = \langle Y, v \rangle$ per ogni $v \in V$ se e solo se $X = Y$.*

Proof.

Chiaramente se $X = Y$, allora $\langle X, v \rangle = \langle Y, v \rangle$ per ogni $v \in V$. Viceversa se $\langle X, v \rangle = \langle Y, v \rangle$ per ogni $v \in V$, avremmo $\langle X - Y, v \rangle = 0$ per ogni $v \in V$, se fosse $X - Y \neq 0$ scegliendo $v = X - Y$ arriveremmo ad una contraddizione dato che il prodotto scalare si suppone essere definito positivo. ■

Lemma 4.2 *Sia u un vettore dello spazio V , allora l'applicazione $l_u : V \rightarrow \mathbf{R}$ definita da $l(X) = \langle X, u \rangle$ e' lineare. Viceversa, per ogni $l : V \rightarrow \mathbf{R}$ esiste uno ed un solo vettore $u \in V$ tale che $l(X) = \langle X, u \rangle$ per ogni $X \in V$.*

Proof. La linearita' della mappa l_u segue dalla linearita' del prodotto scalare. Viceversa data $l : V \rightarrow \mathbf{R}$ lineare, l'unicita' del vettore u che soddisfi la condizione sopra segue dal lemma 4.1. Resta da mostrare l'esistenza. Sia u_1, u_2, \dots, u_n una base ortonormale di V , poniamo $u = l(u_1)u_1 + l(u_2)u_2 + \dots + l(u_n)u_n$. Vogliamo provare che $l = l_u$. Dato che u_1, u_2, \dots, u_n e' una base sara' sufficiente provare che $l(u_j) = l_u(u_j)$ per ogni j tra 1 ed n . Ma

$$l_u(u_j) = \langle u_j, u \rangle = \langle u_j, \sum_{k=1}^n l(u_k)u_k \rangle = \sum_{k=1}^n l(u_k) \langle u_j, u_k \rangle = l(u_j)$$

dato che la base e' ortonormale. ■

Consideriamo ora lo spazio vettoriale $V = \mathbf{R}^3$ con il prodotto scalare canonico, fissati due vettori $v, w \in \mathbf{R}^3$, per la multilinearita' della funzione determinante, l'applicazione $l(X) = \det(v, w, X)$ e' lineare. Esiste quindi per il lemma 4.2 un unico vettore $u \in \mathbf{R}^3$ (che dipende da v e w), tale che $\det(v, w, X) = \langle X, u \rangle$ per ogni $X \in \mathbf{R}^3$. Il vettore u cosi' definito si chiama prodotto vettoriale di v e w e si indica con $v \wedge w$. Abbiamo quindi provato la seguente

Proposition 4.3 Per ogni coppia di vettori v e w in \mathbf{R}^3 esiste uno ed un solo vettore, che indicheremo con $v \wedge w \in \mathbf{R}^3$, che soddisfi la proprieta' seguente:

$$\langle X, v \wedge w \rangle = \det(v, w, X) \text{ per ogni } X \in \mathbf{R}^3$$

Theorem 4.4 Il prodotto vettoriale ha le seguenti proprieta':

- 1) *Linearita' sul primo fattore cioe',*
 $(\alpha v_1 + \beta v_2) \wedge w = \alpha(v_1 \wedge w) + \beta(v_2 \wedge w)$
 per ogni $\alpha, \beta \in \mathbf{R}$ e per ogni $v_1, v_2, w \in V$
- 2) *Linearita' sul secondo fattore cioe',*
 $v \wedge (\alpha w_1 + \beta w_2) = \alpha(v \wedge w_1) + \beta(v \wedge w_2)$
 per ogni $\alpha, \beta \in \mathbf{R}$ e per ogni $v, w_1, w_2 \in V$.
- 3) *Antisimmetria* $v \wedge w = -w \wedge v$
- 4) $v \wedge w = 0$ se e solo se v e w sono linearmente dipendenti.
- 5) Se v e w sono linearmente indipendenti, il vettore (non nullo) $v \wedge w$ e' perpendicolare al sottospazio vettoriale generato da v e w .
- 6) Se v e w sono linearmente indipendenti, $\det(v, w, v \wedge w) > 0$
- 7) Il modulo del vettore $v \wedge w$ coincide con l'area del parallelogramma individuato da v e w

Proof.

Per la definizione del prodotto vettoriale, per la linearita' del determinante sulla prima colonna, e per la linearita' del prodotto scalare sui due fattori, fissato X in \mathbf{R}^3 abbiamo:

$$\langle (\alpha v_1 + \beta v_2) \wedge w, X \rangle = \det((\alpha v_1 + \beta v_2), w, X) = \alpha \det(v_1, w, X) + \beta \det(v_2, w, X) =$$

$$\alpha \langle v_1 \wedge w, X \rangle + \beta \langle v_2 \wedge w, X \rangle = \langle \alpha(v_1 \wedge w) + \beta(v_2 \wedge w), X \rangle$$

Il lemma 4.1 prova allora la proprieta' 1), la dimostrazione della proprieta' 2) e' analoga. Per quel che riguarda la proprieta' 3) abbiamo:

$$\langle w \wedge v, X \rangle = \det(w, v, X) = -\det(v, w, X) = -\langle v \wedge w, X \rangle = \langle -v \wedge w, X \rangle .$$

Concludiamo quindi la dimostrazione della proprieta' 3) ancora grazie al lemma 4.1. Se v e w sono linearmente dipendenti lo sono anche i vettori v, w, X quindi per la proposizione 1.16 $\langle 0, X \rangle = 0 = \det(v, w, X) = \langle v \wedge w, X \rangle$. e invociamo ancora il lemma 4.1. D'altra parte se v e w sono linearmente indipendenti, esiste $u \in \mathbf{R}^3$ tale che v, w e u sono anche essi linearmente indipendenti. Quindi $\langle v \wedge w, u \rangle = \det(v, w, u) \neq 0$, percio' $v \wedge w \neq 0$. Questo conclude la dimostrazione della proprieta' 4). Siano ora v e w vettori linearmente indipendenti, abbiamo $\det(v, w, v) = 0 = \det(v, w, w)$. Quindi $\langle v \wedge w, v \rangle = \langle v \wedge w, w \rangle = 0$. Percio' se α e β sono numeri reali $\langle v \wedge w, \alpha v + \beta w \rangle = \alpha \langle v \wedge w, v \rangle + \beta \langle v \wedge w, w \rangle = 0$, ovvero il vettore non nullo $v \wedge w$ e' ortogonale al piano generato da v e da w e si ottiene la 5). Ora se v e w sono linearmente indipendenti $\det(v, w, v \wedge w) = \langle v \wedge w, v \wedge w \rangle > 0$ dato che il

vettore $v \wedge w$ e' non nullo, e il prodotto scalare canonico di \mathbf{R}^3 e' definito positivo, ovvero la 6). In fine per la 7) se v e w sono linearmente indipendenti, poniamo $n = \frac{v \wedge w}{\|v \wedge w\|}$. Abbiamo

$$\det(v, w, n) = \frac{1}{\|v \wedge w\|} \det(v, w, v \wedge w) = \frac{\langle v \wedge w, v \wedge w \rangle}{\|v \wedge w\|^2} = \frac{\|v \wedge w\|^2}{\|v \wedge w\|^2} = \|v \wedge w\| > 0.$$

Quindi per il teorema 3.2 il modulo di $v \wedge w$ e' il volume del parallelogramma individuato dai vettori v, w ed n . Ma n e' ortogonale al piano generato da v e w , quindi n e' il vettore altezza di tale parallelogramma. Inoltre n ha norma 1, quindi il volume del parallelogramma individuato da v, w ed n , coincide con l'area del parallelogramma individuato da v e w . ■

Osserviamo che le proprieta' 4), 5) 6), 7) caratterizzano il prodotto vettoriale. Infatti la proprieta' 4) determinano quando tale prodotto e' zero, inoltre se non e' zero la proprieta' 5) ne determina la direzione, la proprieta' 6) ne determina il verso e la proprieta' 7) ne determina il modulo.

Ricordiamo inoltre che se θ e' l'angolo compreso tra due vettori non nulli v e w in \mathbf{R}^n , con $0 \leq \theta \leq \pi$, vale la formula $\langle v, w \rangle = \|v\| \|w\| \cos(\theta)$. D'altra parte il vettore altezza h del parallelogramma individuato da v e w e' tale che $\|h\| = \|v\| \sin(\theta)$. (infatti $0 \leq \theta \leq \pi$ quindi $\sin(\theta) \geq 0$). Percio' l'area del parallelogramma individuato da v e da w e' $\|v\| \|w\| \sin(\theta)$. In particolare se v e w sono in \mathbf{R}^3 , vale $\|v \wedge w\| = \|v\| \|w\| \sin(\theta)$.

Vogliamo ora calcolare le coordinate del prodotto vettoriale. Sia $v \in \mathbf{R}^3$ un vettore colonna di coordinate a, b, c , sia $w \in \mathbf{R}^3$ un altro vettore colonna di coordinate d, e, f e siano p, q, r le coordinate del prodotto vettoriale $v \wedge w$. Abbiamo quindi $v \wedge w = pe_1 + qe_2 + re_3$, dove e_1, e_2, e_3 e' la base canonica di \mathbf{R}^3 . Allora dato che la base canonica e' ortonormale si ricava:

$$\begin{aligned} \langle v \wedge w, e_1 \rangle &= p \langle e_1, e_1 \rangle + q \langle e_2, e_1 \rangle + r \langle e_3, e_1 \rangle = p \\ \langle v \wedge w, e_2 \rangle &= p \langle e_1, e_2 \rangle + q \langle e_2, e_2 \rangle + r \langle e_3, e_2 \rangle = q \\ \langle v \wedge w, e_3 \rangle &= p \langle e_1, e_3 \rangle + q \langle e_2, e_3 \rangle + r \langle e_3, e_3 \rangle = r. \end{aligned}$$

Quindi dalla definizione del prodotto vettoriale si ottiene

$$\det(v, w, e_1) = p, \det(v, w, e_2) = q, \det(v, w, e_3) = r.$$

Scambiando la seconda e la terza colonna, e poi la prima e la seconda colonna, il determinante cambia segno due volte, rimane quindi invariato. Il determinante resta inoltre invariato se sostituiamo una matrice con la sua trasposta, possiamo quindi anche scrivere

$$\det((e_1, v, w)^t) = p, \det((e_2, v, w)^t) = q, \det((e_3, v, w)^t) = r.$$

Sviluppando con il metodo di Laplace questi ultimi determinanti rispetto alla prima riga ricaviamo:

$$p = (bf - ce), q = (cd - af), r = (ae - bd).$$

Per memorizzare queste formule si usa spesso un trucco formale che non ha realmente senso ma e' comodo per fare i conti. Si scrive una "matrice" che abbia sulla prima riga i tre vettori della base canonica, sulla seconda riga le tre componenti a, b, c del vettore v e sulla terza riga le tre componenti d, e, f del vettore w . Sviluppando formalmente il determinante di questa matrice rispetto alla prima riga si ottiene una combinazione lineare dei vettori della base canonica i cui coefficienti sono le coordinate p, q, r del prodotto vettoriale $v \wedge w$.

Corollary 4.5 *Siano v e w vettori colonna linearmente indipendenti in \mathbf{R}^3 , che individuano un parallelepipedo P nel piano generato da v e w . Supponiamo che v abbia componenti a, b, c e che w abbia componenti d, e, f , allora*

$$(\text{area}(P))^2 = \|v \wedge w\|^2 = (bf - ce)^2 + (cd - af)^2 + (ae - bd)^2 = \langle v, v \rangle \langle w, w \rangle - \langle v, w \rangle^2 .$$

Proof.

Basta usare l'espressione in coordinate del vettore $v \wedge w$ e il corollario 3.6. ■