Analisi Matematica II, Anno Accademico 2017-2018.
Ingegneria Edile e Architettura
Vincenzo M. Tortorelli
FOGLIO DI TEORIA n. 11
FUNZIONI MULTILINEARI, DERIVATE DI PRODOTTI DI CAMMINI, DETERMINANTE

Funzioni multilineari.

Funzioni multilineari. - Siano $V^1, \dots V^k$, W spazi vettoriali: una funzione $P : V^1 \times \cdots \times V^k \to W$ si dice *multilineare* se fissate $k-1$ tra le variabili vettoriali, si ottiene una funzione lineare nella variabile vettoriale rimanente. Si dice *forma multilineare* se $W = \mathbf{R}$ o $W = \mathbf{C}$.
- In altri termini è *distributiva rispetto* alle somme degli spazi vettorilai e *"commutativa" rispetto* al prodotto per numeri reali. È quindi suggestiva la notazione che è quella del prodotto di numeri: $P(v^1, \dots, v^k) = v^1 \cdot_P \cdots \cdot_P v^k$ per mettere in risalto tali proprietà.
Se $\mathbf{b_j} = \{b_j^i\}_{i\in I}$ è una base di V^j, $1 \le j \le k$, il prodotto P è determinato dai valori sulle $k-ple$ di elementi delle basi: $P(b_1^{i_1}, \dots, b_k^{i_k})$.

Funzioni multilineari alternanti: Per $V^1 = \cdots = V^k = V$ una funzione multilineare su V ($V \times \cdots \times V$ k volte) si dice *alternante* se si annullla quando *due fattori sono uguali.* Ovvero è *antisimmetrica* cioè cambia segno se due fattori si scambiano di posto.
Osservazione: se $k > \dim V$ la sola funzione multilineare alternante su V è quella nulla.

Esempio: - il tipico esempio di forma multilineare è il determinante di una matrice A quadrata $m \times m$, come funzione delle colonne (o delle righe). In questo caso $k = m$, $V^1 = \cdots = V^k = \mathbf{R}^m$, $W = \mathbf{R}$. Essa risulta *alternante* cioè *anticommutativa* rispetto allo scambio delle variabili vettoriali. Si ricorda il seguente teorema:
Teorema: vi è un'unica funzione m-lineare da $\mathbf{R}^m \times \cdots \times \mathbf{R}^m$ad $\mathbf{R}$, alternante, che vale uno sulla base canonica di $\mathbf{R}^m$: il determinante della matrice che ha gli m vettori come colonne.

Dimostrazione: siano dati: $\mathbf{e}_1, \dots, \mathbf{e}_m$ la base canonica $\mathbf{R}^m$, e, per $1 \le j \le m$, i vettori $v^j = (v_1^j, \dots, v_m^j) = v_1^j \mathbf{e}_1 + \cdots + v_m^j \mathbf{e}_m$. Se Λ è multilineare su $\mathbf{R}^m$ per cui $\Lambda(\mathbf{e}_1, \dots, \mathbf{e}_m) = 1$, si ha: $\Lambda(v^1, v^2, \dots, v^m) = \Lambda(v_1^1\mathbf{e}_1 + \cdots + v_m^1\mathbf{e}_m, v^2, \dots, v^m) =$

$$= v_1^1\Lambda(\mathbf{e}_1, v^2, \dots, v^m) + \cdots + v_m^1\Lambda(\mathbf{e}_m, v^2, \dots, v^m) = \sum_{i_1=1}^m v_{i_1}^1 \Lambda(\mathbf{e}_{i_1}, v^2, \dots, v^m) =$$

$$= \sum_{i_1=1}^m \sum_{i_2=1}^m v_{i_1}^1 v_{i_2}^2 \Lambda(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, v^m) = \cdots = \sum_{i_1,\dots,i_m=1}^m v_{i_1}^1 \cdot \cdots \cdot v_{i_m}^m \Lambda(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_m}) =$$

$$= \sum_{i_p \ne i_q \text{ se } p\ne q,\ 1}^m v_{i_1}^1 \cdot \cdots \cdot v_{i_m}^m \Lambda(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_m}) =$$

$$= \sum_{i_p \ne i_q \text{ se } p\ne q,\ 1}^m \text{segno}(i_1, \dots, i_m) v_{i_1}^1 \cdot \cdots \cdot v_{i_m}^m \Lambda(\mathbf{e}_1, \dots, \mathbf{e}_m) =$$

$$= \sum_{i_p \ne i_q \text{ se } p\ne q,\ 1}^m \text{segno}(i_1, \dots, i_m) v_{i_1}^1 \cdots\cdots v_{i_m}^m = \sum_{j_p \ne j_q \text{ se } p\ne q,\ 1}^m \text{segno}(j_1, \dots, j_m) v_1^{j_1} \cdots\cdots v_m^{j_m} =: \det(v_i^j)$$

Notazione: - Se A è una matrice $h \times k$ con A_i^j si indica la sua componente di i-esima riga e j-esima colonna: considerando gli elementi di $\mathbf{R}^h$ e di $\mathbf{R}^k$ come colonne $A_i^j = {}^t\mathbf{e}_i^{\mathbf{R}^h} A \mathbf{e}_j^{\mathbf{R}^k}$.
- Con A_i si indica la i-esima riga, con A^j la j-esima colonna: $A_i = {}^t\mathbf{e}_i^{\mathbf{R}^h} A$, $A^j = A\mathbf{e}_j^{\mathbf{R}^k}$.
- Siano dati: A matrice $h \times k$, due *multi-indici crescenti* di numeri naturali non nulli: $I = (i_1, \dots, i_r)$, $r \le h$, $1 \le i_1 < \cdots < i_r \le h$ e, rispettivamente, $J = (j_1, \dots, j_s)$, $s \le k$, $1 \le j_1 < \cdots < j_s \le k$.
- - Con A_I o $A_{i_1,\dots,i_r}$ rispettivamente A^J o $A^{j_1,\dots,j_s}$, si indicano le matrici $r \times k$ ed $h \times s$ ottenute selezionando le righe $i_1 \dots i_r$ rispettivamente le colonne $j_1 \dots j_s$;

- - con A_I^J o $A_{i_1,\dots,i_r}^{j_1,\dots,j_s}$ la matrice $r\times s$ ottenuta selezionando sia le r righe che le s colonne;

- - posto quindi $E^{\mathbf{R}^n}=(\mathbf{e}_1^{\mathbf{R}^n}|\dots|\mathbf{e}_n^{\mathbf{R}^n})=Id_{n\times n}=\begin{pmatrix} {}^t\mathbf{e}_1^{\mathbf{R}^n}\\ \vdots \\ {}^t\mathbf{e}_n^{\mathbf{R}^n}\end{pmatrix}$, si ha ($A$ è $h\times k$):

$$A_I=E_I^{\mathbf{R^h}}A,\quad A^J=AE^{\mathbf{R^k}J},\ A_I^J=E_I^{\mathbf{R^h}}AE^{\mathbf{R^k}J}=(A_I)^J=(A^J)_I,\ \ {}^t(E^J)=E_J;$$

- - $E_I=E_{i_1\dots i_r}$ ha nulle le colonne di indice non in I, e, essendo $i_h<i_{h+1}$, $m\le r$ colonne *adiacenti consecutive non nulle* formano sottomatrici $Id_{m\times m}$; analogamente per E^J;
- - per le matrici identità se I ha qualche indice che non compare in J, o viceversa, la $E_IE^J=E_IId_{n\times n}E^J=E_I^J$, matrice $r\times s$, ha la riga di quell'indice, rispettivamente di quella colonna, nulla: in particolare $E_IE^I=E_I^I$ è la matrice identica della dimensione di I;
- - non ha senso il prodotto E^JE_I se gli indici hanno dimensione diversa, $r\neq s$; se $r=s$ la matrice $n\times n$, E^JE_I ha nulle le colonne di indice non in I e le righe non in J.

In particolare E^IE_I è la matrice diagonale con elementi nulli per gli indici non in I;
- - con $A_{\not I}$ o $A_{\not i_1,\dots,\not i_r}$, $\quad A^{\not J}$ o $A^{\not j_1,\dots,\not j_s}$, $\quad A_{\not I}^{\not J}$ o $A_{\not i_1,\dots,\not i_r}^{\not j_1,\dots,\not j_s}$ le matrici $(h-r)\times k$, $h\times(k-s)$, $(h-r)\times(k-s)$ ottenute nei vari casi scartando le righe o le colonne degli indici barrati.
- - Se con ${}^cI$, ${}^cJ$ si indicano i rimanenti indici si ha: $A_{\,{}^cI}=A_{\not I}$ e analoghi per $A^{\,{}^cJ}$, $A_{\,{}^cI}^{\,{}^cJ}$.

- Quindi per una matrice quadrata A invertibile si ha $\left[A^{-1}\right]_i^j = (-1)^{i+j}\dfrac{\det A_{\not j}^{\not i}}{\det A}$.

<u>Osservazione:</u> si considerano $E=Id_{n\times n}$, $H=(h_1,\dots,h_m)$, $1\le h_1<\dots<h_m\le n$, $m\le n$.
- Allora la matrice $m\times n$, E_H è la matrice associata (nelle basi canoniche) alla funzione lineare di *proiezione* di $\mathbf{R}^n$ su $\mathbf{R}^m$, con le componenti specificate da H.
- La matrice $n\times m$, E^H è associata (nelle basi canoniche) alla funzione lineare *immersione* di $\mathbf{R}^m$ in $\mathbf{R}^n$, nel sottospazio coordinato con componenti non specificate da H nulle.
- Come osservato $E_HE^H=E_H^H=Id_{m\times m}$; E^HE_H, $n\times n$ (immersione della proiezione) è associata alla proiezione ortogonale da $\mathbf{R}^n$ in sè sul sottospazio coordinato di $\mathbf{R}^n$ generato dalla colonne di E^H, quello appunto in cui le coomponenti non specificate da H sono nulle.
- Se J ed I hanno stessa lunghezza r la $E^JE_I+E^{\not J}E_{\not I}$, indicati con $h_1<\dots<h_{n-r}$ gli indici non in I, $k_1<\dots<k_{m-r}$ quelli non in J, dà la *permutazione di coordinate crescente a blocchi* (shuffle) che porta x_{i_h} al posto j_h, x_{h_a} al posto k_a.

<u>Notazione:</u> - Per la sostituzione con una colonna v, della colonna j^a di una matrice A, $h\times k$ si usa la notazione $A\,[v/A^j]$, o $(\dots|A^{j-1}|V|A^{j+1}\dots)$. Analogamente per le righe.
- Quindi la regola di Cramer per le soluzioni u del sistema lineare $Au=v$ $h\times h$, ovvero $u_i=\sum_{j=1}^{h}(-1)^{i+j}v_j\det A_{\not j}^{\not i}(\det A)^{-1}$ diventa, essendo la sommatoria lo sviluppo del determinante per la i^a colonna della matrice $A[v/A^i]$:

$u_i = \dfrac{\det A[v/A^i]}{\det \text{ A}}$. Per un sistema matriciale $AU=V$, $U=A^{-1}V$: $\quad U_i^j = \dfrac{\det A[V^j/A^i]}{\det A}$.

Teorema: se $(\mathbf{e}_1,\dots,\mathbf{e}_m)$ è la base canonica di $\mathbf{R}^m$, $k\le m$ ogni funzione k lineare alternante Λ su $\mathbf{R}^m$ è del tipo: $\quad\Lambda(A^1,\dots,A^k)=\sum_{1\le i_1<\dots<i_k\le m}\Lambda(\mathbf{e}_{i_1},\dots,\mathbf{e}_{i_k})\det A_{i_1\dots i_k}^{1\dots k}$.

La dimostrazione elementare si basa sul ragionamento che si usa per provare l'unicità del determinante, ovvero lo sviluppo di Leibniz, usando la multilinearità (distributività ed omogeneità) e l'alternanza per raggruppare i fattori che differiscono solo per l'ordine.

Corollario: due funzioni k lineari su uno spazio vettoriale V di dimensione m, coincidono se coincido sulle k-ple, $\mathbf{b}_{i_1},\dots,\mathbf{b}_{i_k}$ di elementi di una base $\mathbf{b}_1,\dots,\mathbf{b}_k$.

- Si definiscono le funzioni k-lineari $dx_{i_1}\wedge\dots\wedge dx_{i_k}(A^1,\dots,A^k)=\det A_{i_1\dots i_k}^{1\dots k}$, $A^j\in\mathbf{R}^m$.

- Tenendo presente il significato geometrico del determinante ("volume con segno" di un "parellelepipedo orientato", cfr.paragrafi seguenti) queste funzioni intuitivamente calcolano i "volumi k-dimensionali con segno" dei "parallelepipedi k-dimensionali" in $\mathbf{R}^k$ ottenuti proiettando sui $\binom{m}{k}$ piani coordinati di $\mathbf{R}^m$ di dimensione k, il "parallelpipedo k-dimensionale" in $\mathbf{R}^m$ con spigoli dati dalle somme di $0_{\mathbf{R}^m}, A^1, \dots, A^k$.

- Quindi le funzioni k lineari alternanti su $\mathbf{R}^m$ sono uno spazio vettoriale di dimensione $\binom{m}{k}$, con base data da $dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ al variare di $i_i < \cdots < i_k \leq m$:

Esempio: cfr. FT10 - il prodotto vettoriale $\cdot \times \cdot$ in $\mathbf{R}^3$, è una funzione bilineare alternante a valori in $\mathbf{R}^3$: $\cdot \times \cdot : \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}^3$, $V^1 = V^2 = W = \mathbf{R}^3$:

$$\begin{aligned} u \times v &= dx_2 \wedge dx_3(u,v)\mathbf{e}_1 - dx_1 \wedge dx_3(u,v)\mathbf{e}_2 + dx_1 \wedge dx_2(u,v)\mathbf{e}_3 = \\ &= (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1) \end{aligned} .$$

- Fissato un vettore $u = (u_1, u_2, u_3)$ ad esso si associa la matrice *antisimmetrica* $\tilde{u}$ associata alla funzione lineare che da il prodotto vettore sinistro con u:

$$u \times x = \tilde{u}x = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Esempio: Indicato con $\mathcal{M}_{a\times b}$ lo spazio vettoriale delle matrici reali con a righe e b colonne, il prodotto righe per colonne $P : \mathcal{M}_{h\times m} \times \mathcal{M}_{m\times k} \to \mathcal{M}_{h\times k}$: $(P(A,B))_i^j = \sum_r A_i^r B_r^j$, è bilineare.

Prodotti scalari tra matrici Identificando lo spazio $\mathcal{M}_{h\times m}$ delle matrici $h \times m$ con $\mathbf{R}^{hm}$, *e.g.* allineando verticalmente le colonne per ottenere vettori colonna $A = (A_i^j) \mapsto {}^t(A_1^1, \dots, A_h^1, A_1^2, \dots, A_1^m, \dots, A_h^m)$, e orizzontalmente le righe per ottenere vettori riga $A \mapsto (A_1^1, \dots, A_1^m, A_2^1, \dots, A_h^1, \dots, A_h^m)$, il prodotto scalare in $\mathbf{R}^{hm}$ diventa un prodotto scalare $\cdot_{\mathcal{M}}$ tra matrici $h \times m$

$$\langle A \cdot B \rangle_{\mathbf{R}^{hm}} = \sum_i \sum_r A_r^i B_r^i = tr\,{}^t\!AB =: A \cdot_{\mathcal{M}} B$$

corrispondente al prodotto righe per colonne

$$\begin{aligned} tr\,{}^t\!AB &= (({}^tA)_1^1, \dots, ({}^tA)_1^h, ({}^tA)_2^1, \dots, ({}^tA)_m^1, \dots, ({}^tA)_m^h)\ {}^t(B_1^1, \dots, B_h^1, B_1^2, \dots, B_1^m, \dots, B_h^m) = \\ &= (A_1^1, \dots, A_h^1, A_1^2, \dots, A_1^m, \dots, A_h^m)\ {}^t(B_1^1, \dots, B_h^1, B_1^2, \dots, B_1^m, \dots, B_h^m) \end{aligned} .$$

Funzioni bilineari e sesquilinari. - Una forma bilineare reale definita su coppie di elementi di uno spazio vettoriale reale, che sia simmetrica (commutativa nelle due variabili vettoriali), non negativa, non degenere è un prodotto scalare (cfr.FT2, proprietà s1, s2, s3, s4).
- Una forma definita sulle coppie di elementi di uno spazio vettoriale *complesso*, a valori complessi, $\langle x \cdot y \rangle \in \mathbf{C}$, si dice *sesquilineare* se soddisfa:
h2.1) *additività:* $\langle (x+u) \cdot (y+v) \rangle = \langle x \cdot y \rangle + \langle u \cdot y \rangle + \langle x \cdot v \rangle + \langle u \cdot v \rangle$,
h2.2) *omogeneità (linearità) nella prima variabile:* $\langle \lambda x \cdot y \rangle = \lambda \langle x \cdot y \rangle$ per ogni $\lambda \in \mathbf{C}$,
h2.3) *antiomogeneità (antilinearità) nella prima variabile:* $\langle x \cdot \mu y \rangle = \overline{\mu} \langle x \cdot y \rangle$ per ogni $\mu \in \mathbf{C}$;
- una tale funzione si dice *hermitiana* se soddisfa anche la condizione:
h3)*simmetria:* $\langle x \cdot y \rangle = \overline{\langle y \cdot x \rangle}$, da cui segue anche: $\langle x \cdot x \rangle \in \mathbf{R}$,
- si dice *prodotto scalare hermitiano* o prodotto scalare complesso se inoltre soddisfa:
h1) *non negatività:* $\langle x \cdot x \rangle \geq 0$,
h4) *non degenere:* $\langle x \cdot x \rangle = 0 \iff x = \underline{0}$.
- Come per le forme su $\mathbf{R}^m$, ogni forma sesquilineare su $\mathbf{C}^m$ individua una matrice B, $m \times m$, a coefficienti in $\mathbf{C}$, e viceversa: si ottiene con i valori della forma sulle coppie di elementi della base canonica $B_i^j = \langle \mathbf{e}_i \cdot \mathbf{e}_j \rangle$, per cui $\langle x \cdot y \rangle = {}^t x B \overline{y}$ (considerando le m-ple come colonne). La matrice B associata ad una funzione hermitiana risulta pertanto *autoaggiunta* cioè $B = {}^t\overline{B}$, e con autovalori reali.
Il prodotto scalare complesso "canonico" su $\mathbf{C}^m$ è quello associato alla matrice identica: $\langle x \cdot y \rangle = x_1 \overline{y_1} + \cdots + x_m \overline{y_m}$.
- Se una forma sesquilineare è non negativa è hermitiana e si pone $\|x\| = \sqrt{\langle x \cdot x \rangle}$.

Diseguaglianza di Cauchy-Schwarz: Come per i prodotti scalari reali, per i prodotti scalari hermitiani vale la diseguaglianza di Cauchy-Schwarz ($|x|$ è il modulo di un numero complesso): $$|\langle A\cdot B\rangle|\le ||A||\,||B||,$$
avendosi l'eguaglianza se e solo se A e B sono linearmente dipendenti.
La dimostrazione esposta in FT2, nel caso di prodotti scalari reali, porta solo a $|Re(\langle A\cdot B\rangle)|\le ||A||||B||$. Si applica quindi tale diseguaglianza con $\dfrac{\langle A\cdot B\rangle}{|\langle A\cdot B\rangle|}B$ al posto di B.
- Pertanto per prodotti scalari complessi $\|x\|$ è una norma e $d(x,y)=\|x-y\|$ una distanza.

Le eguaglianze Cauchy-Binet e Cauchy-Pitagora

Cauchy-Binet in $\mathbf{R}^3$: - come mostrato in FT 10, le uniche funzioni reali bilineari antisimmetriche $B:\mathbf{R}^3\times\mathbf{R}^3\to\mathbf{R}$, sono:

$$B\,(u,v)=\ \det\begin{pmatrix} a & u_1 & v_1\\ b & u_2 & v_2\\ c & u_3 & v_3\end{pmatrix}=a(u_2v_3-u_3v_2)+b(v_1u_3-u_1v_3)+c(u_1v_2-u_2v_1)=$$
$$=\left\langle\begin{pmatrix}a\\b\\c\end{pmatrix}\cdot\left[\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}\times\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}\right]\right\rangle=(u_1,u_2,u_3)\begin{pmatrix}0 & c & -b\\ -c & 0 & a\\ b & -a & 0\end{pmatrix}\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}.$$

- Equivalentemente $A=\ {}^t(a_1,a_2,a_3),\ B=\ {}^t(b_1,b_2,b_3),U=\ {}^t(u_1,u_2,u_3),\ V=\ {}^t(v_1,v_2,v_3)$
$$\langle A\times B\cdot U\times V\rangle\ =\ \left\langle\left[\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}\times\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}\right]\cdot\left[\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}\times\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}\right]\right\rangle\ =$$
$$=\det(A\ |B)_{23}\det(U\ |V)_{23}+\det(A\ |B)_{13}\det(U\ |V)_{13}+\det(A\ |B)_{12}\det(U\ |V)_{12}=$$
$$=\ (a_2b_3-a_3b_2)(u_2v_3-u_3v_2)+(b_1a_3-a_1b_3)(v_1u_3-u_1v_3)+(a_1b_2-a_2b_1)(u_1v_2-u_2v_1).$$

D'altronde $$\det\begin{pmatrix}{}^tA\\ {}^tB\end{pmatrix}(U\mid V)=\ \det\begin{pmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{pmatrix}\begin{pmatrix}u_1 & v_1\\ u_2 & v_2\\ u_3 & v_3\end{pmatrix}=\det\begin{pmatrix}\langle A\cdot U\rangle & \langle A\cdot V\rangle\\ \langle B\cdot U\rangle & \langle B\cdot V\rangle\end{pmatrix}=$$
$$=(a_1u_1+a_2u_2+a_3u_3)(b_1v_1+b_2v_2+b_3v_3)-(a_1v_1+a_2v_2+a_3v_3)(b_1u_1+b_2u_2+b_3u_3).$$
Sviluppando i calcoli si ottiene alla fine:
$$\det\begin{pmatrix}{}^tA\\ {}^tB\end{pmatrix}(U\,|\,V)=\langle A\times B\cdot U\times V\rangle=$$
$$=\det(A\ |B)_{23}\det(U\ |V)_{23}+\det(A\ |B)_{13}\det(U\ |V)_{13}+\det(A\ |B)_{12}\det(U\ |V)_{12}.$$

- Alternativamente: osservando che fissati A e B le funzioni $(U,V)\mapsto\det{}^t(A|B)(U|V)$ e $(U,V)\mapsto\langle A\times B\cdot U\times V\rangle$, sono bilineari alternanti su $\mathbf{R}^3$, per ottenere l'identità basta verificarla sulle coppie di elementi della base canonica $\mathbf{e}_1=\ {}^t(1,0,0),\ \mathbf{e}_2=\ {}^t(0,1,0),\ \mathbf{e}_3=\ {}^t(0,0,1)$,
$$\det{}^t(A|B)(\mathbf{e}_i|\mathbf{e}_j)\overset{?}{=}\langle A\times B\cdot\mathbf{e}_i\times\mathbf{e}_j\rangle,\ i<j$$

ancora osservando che fissati $\mathbf{e}_i,\ \mathbf{e}_j$ $(A,B)\mapsto\det{}^t(A|B)(\mathbf{e}_i|\mathbf{e}_j)$ e $(A,B)\mapsto\langle A\times B\cdot\mathbf{e}_i\times\mathbf{e}_j\rangle$ sono bilineari alternanti basta verificare: $\det{}^t(\mathbf{e}_h|\mathbf{e}_k)(\mathbf{e}_i|\mathbf{e}_j)\overset{?}{=}\langle\mathbf{e}_h\times\mathbf{e}_k\cdot\mathbf{e}_i\times\mathbf{e}_j\rangle,\ i<j,\ \ h<k$:

infatti se $\{h,k\}\neq\{i,j\}$ la matrice prodotto del primo membro ha una colonna e una riga nulla e quindi il determinante è nullo. Il secondo membro sarà invece nullo essendo i prodotti vettori di due elementi della base canonica eguali, a meno del segno, a quello mancante: se le coppie sono diverse i prodotti vettori sono, segno a parte, diversi vettori della base canonica, quindi ortogonali. Quando invece $(i,j)=(h,k)$ entrambi i termini sono uguali ad 1, avendo a priori imposto $i<j$ e $h<k$.

Osservazione: - il determinante di matrici 2×2, $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ intuitivamente rappresenta il rapporto tra l'area con segno del parallelogramma di vertici $O=(0,0)$, ${}^tA=(a_1,a_2)$, ${}^tB=(b_1,b_2)$, $S=(a_1+b_1,a_2+b_2)$, e quella "unitaria" del quadrato generato dai versori e_1, e_2 del sistema di coordinate. Se questo parallelogramma è *orientato* come (e_1,e_2), cioè il "primo " lato orientato $\vec{OA}$, si allinea sul "secondo" lato orientato $\vec{OB}$, spazzando la zona dell'angolo convesso tra i due, in senso antiorario, (ovvero i vertici nell'ordine (O,A,S,B) sono in senso antiorario), si assegna segno positivo negativo altrimenti. Per esempio con il teorema del seno:
$[\det\ (A|B)]^2=\det\ {}^t(A|B)(A|B)=$

$$=\det\begin{pmatrix} \langle A\cdot A\rangle_{\mathbf{R}^2} & \langle A\cdot B\rangle_{\mathbf{R}^2} \\ \langle B\cdot A\rangle_{\mathbf{R}^2} & \langle B\cdot B\rangle_{\mathbf{R}^2}\end{pmatrix}=|A|^2_{\mathbf{R}^2}|B|^2_{\mathbf{R}^2}(1-\cos^2\widehat{AOB})=|A|^2|B|^2\sin^2\widehat{AOB}.$$

In un qualsiasi sistema di riferimento cartesiano con stessa orientazione (ortonormale dato da una rotazione R) il determinante della matrice delle nuove coordinate, ${}^tRA$, ${}^tRB$, non cambia.

Interpretazione geometrica di $\langle A\times B\cdot U\times V\rangle$: *ortogonalizzando* (A,B) senza normalizzare, (sostituendo $C=B-\langle B\cdot\widehat{A}\rangle\widehat{A}$ a B) la quantità $\det\begin{pmatrix} {}^tA \\ {}^tB\end{pmatrix}(U\,|\,V)$ non cambia, così come l' area del parallegramma generato da A, B è uguale ad $|A||C|$, quella del rettangolo generato da A, C. Nel caso $\langle A\cdot U\rangle$, $\langle C\cdot U\rangle$, e $\langle A\cdot V\rangle$, $\langle C\cdot V\rangle$ sono le coordinate nel sistema di riferimento *ortogonale (non ortonormale)* $(A|A|^{-2},C|C|^{-2})$, delle proiezioni di U e di V sul piano Π generato da A, B. Quindi, un'interpretazione geometrica di $\langle A\times B\cdot U\times V\rangle_{\mathbf{R}^3}$, è quella di rapporto tra le "aree con segno" (il segno del determinante tiene conto dell'orientazione relativa): quella del parallelogramma proiezione su Π di quello associato ad (U,V), e quella $\frac{1}{|A||C|}$ del rettangolo della base $(A|A|^{-2},C|C|^{-2})$. Brevemente $\langle A\times B\cdot U\times V\rangle_{\mathbf{R}^3}$ è l'"area con segno" *rispetto alle unità di misura* $|A|^{-1}_{\mathbf{R}^3}$, $|C|^{-1}_{\mathbf{R}^3}$ *di un sistema di riferimento ortogonale* che dia un rettangolo con stessa area e orientazione del parallelogramma generato da A, B.

Estensione del teorema di Pitagora: - d'altronde grazie a tale interpretazione intuitiva si arriva direttamente a convincersi dell'identità di Cauchy-Binet nel caso $A=U$, $B=V$.
Infatti da una parte: $\det\ {}^t(A|B)(A|B)=\det\begin{pmatrix} \langle A\cdot A\rangle_{\mathbf{R}^3} & \langle A\cdot B\rangle_{\mathbf{R}^3} \\ \langle B\cdot A\rangle_{\mathbf{R}^3} & \langle B\cdot B\rangle_{\mathbf{R}^3}\end{pmatrix}=$

$$=|A|^2_{\mathbf{R}^3}|B|^2_{\mathbf{R}^3}(1-\cos^2\widehat{AOB})=|A|^2|B|^2\sin^2\widehat{AOB}$$

Lemma : se R è una rotazione, ${}^tR=R^{-1}$, $\det R=1$, si ha $R(A\times B)=RA\times RB$.

$$\langle C\cdot(A\times B)\rangle=\det\ (C|A|B)=\det R\ \det\ (C|A|B)=\det(RC|RA|RB)=$$
$$=\langle RC\cdot(RA\times RB)\rangle=\langle C\cdot\ {}^tR(RA\times RB)\rangle$$

per cui $R(A\times B)=RA\times RB$, essendo C arbitrario.

- D'altra parte, grazie al lemma, considerando la rotazione che porta il piano generato da A, B sul piano coordinato $z=0$ (rotazione attorno alla retta intersezione dei due sottospazi vettoriali in questione) ci si riconduce al caso piano e si ha:

$$|A\times B|^2_{\mathbf{R}^3}=|RA\times RB|^2_{\mathbf{R}^3}=(\det(RA|RB))^2=|RA|^2_{\mathbf{R}^2}|RB|^2_{\mathbf{R}^2}|\sin(\widehat{RA)O(}RB)|^2=$$
$$=|A|^2_{\mathbf{R}^3}|B|^2_{\mathbf{R}^3}|\sin\widehat{AOB}|^2,$$

che si interpreta come area del parallelogramma di vertici $O,\ A,\ B,\ A+B$ nello spazio .

- In definitiva si ottiene l'analogo del *Teorema di Pitagora*:

$$\mathrm{Area}^2(O,\ A,\ B,\ A+B)=\det\ {}^t(A|B)(AB)=|A\times B|^2_{\mathbf{R}^3}=$$
$$=\text{somma quadrati aree proiezioni ortogonali sui piani coordinati.}$$

La formula di Cauchy-Binet in $\mathbf{R}^m$: se S e P sono matrici $m\times k$, con $m>k$ si ha:

$$\det\ {}^tSP\ =\sum_{1\le t_1<\cdots<t_k\le m}\det\ S_{t_1\ldots t_k}\det P_{t_1\ldots t_k}=\sum_{1\le t_1<\cdots<t_k\le m}\det\ S_{t_1\ldots t_k}P_{t_1\ldots t_k}.$$

Dimostrazione: - si segue la traccia della seconda dimostrazione fatta nel caso $m=3$ e $k=2$.

Preliminarmente si osserva $\qquad\qquad \det A_T \det B_U = \det\, A_T B_U = \det\, E_T A E_U B.$

i- Fissata S, essendo ${}^tSP = ({}^tSP^1|\dots|{}^tSP^k)$, la $(P^1,\dots,P^k) \sim P \mapsto \det{}^tSP$ è k-lineare alternante. Più direttamente ognuno degli addendi del secondo membro definisce una funzione k-lineare alternante: $(P^1,\dots,P^k) \sim P \mapsto \det\, S_{t_1\dots t_k}\det P_{t_1\dots t_k}$.

ii- Quindi, denotando con $(\mathbf{e}_1,\dots,\mathbf{e}_m)$ la base canoniica di $\mathbf{R}^m$, basta considerare
$P = (\mathbf{e}_{i_1}|\dots|\mathbf{e}_{i_k}) =: E^I$, $1 \le i_1 < \dots < i_k \le m$. Ripetendo il ragionamento su S vista per righe, fissati $1 \le i_1 < \dots < i_k \le m$, basta considerare
$S = (\mathbf{e}_{h_1}|\dots|\mathbf{e}_{h_k}) =: E^H$, $1 \le h_1 < \dots < h_k \le m$.
Ci si riduce quindi a verificare dati $1 \le i_1 < \dots < i_k \le m$, $1 \le h_1 < \dots < h_k \le m$

$$\det{}^t(\mathbf{e}_{h_1}|\dots|\mathbf{e}_{h_k})(\mathbf{e}_{i_1}|\dots|\mathbf{e}_{i_k}) \overset{?}{=} \sum_{1\le t_1<\dots<t_k\le m} \det(\mathbf{e}_{h_1}|\dots|\mathbf{e}_{h_k})_{t_1\dots t_k}\det(\mathbf{e}_{i_1}|\dots|\mathbf{e}_{i_k})_{t_1\dots t_k},$$

indicando i k-multiindici crescenti con le maiuscole delle lettere rispettivamente indicizzate

$$\det{}^t(E^H)E^I \overset{?}{=} \sum_{T\uparrow}\det(E^H)_T\det(E^I)_T.$$

iii- Ora basta osservare:
- per il primo membro che ${}^t(E^H) = E_H$, quindi ${}^t(E^H)E^I = E_H E^I = E^H_I$;
- per il secondo membro che: $(E^H)_T = E_T E^H = E^H_T$, $(E^I)_T = E_T E^I = E^I_T$.

Se $H{\ne}I$: - per il primo membro $\det E^H_I = 0$ avendo la matrice una colonna e una riga nulle;
- per il secondo membro: per ogni addendo se $T \ne H$ la prima matrice avrebbe una colonna e una riga nulle, per $T \ne I$ ciò accadrebbe per la seconda matrice: in entrambi i casi i determinanti sarebbero nulli. Pertanto se $H \ne I$ sarebbero tutti addendi nulli.

Se $H{=}I$: - per il primo membro $\det E^I_I = \det Id_{k\times k} = 1$;
- per il secondo membro: per quanto sopra analizzato, sarebbe non nullo solo l'addendo con $T = H = I$, e si avrebbe $\det E_T E^H = \det E_T E^I = \det Id_{k\times k} = 1$.

<u>Osservazione:</u> $\det S\,{}^tP$ è sempre nullo. Già le matrici E^H hanno $m-k$ righe nulle e le E_I, di tipo trasposto, $m-k$ colonne nulle. Il prodotto $E^H E_I$, $m\times m$, avrà determinante nullo.

<u>Osservazione:</u> L'interpretazione geometrica del determinante si estende al determinante di matrici quadrate Q $m\times m$ considerandolo come *"m-volume con segno"* del parallepipedo m-dimensionale di vertici $O, Q^1,\dots,Q^m$, e le loro somme, *orientato considerando nell'ordine* $(Q^1,\dots Q^m)$. L'argomento è la versione piuttosto semplificata del seguente:

Interpretazione geometrica: - tale interpretazione si estende appunto anche ai determinanti del tipo $\det{}^tPP \ge 0$, con P matrice, "lunga", $m\times k$, $m>k$.
Si interpreterà tale determinante, sempre non negativo, come "*quadrato del k-volume*" del parallepipedo k-dimensionale in $\mathbf{R}^m$, di vertici $O, P^1,\dots,P^k$, e le loro somme.

- Se le colonne di P sono dipendenti ("k-parellelpipedo degenere")anche le colonne di ${}^tPP = ({}^tPP^1|\dots|{}^tPP^k)$ sono dipendenti e il determinante è nullo.

- Un modo di convincersi di tale interpretazione, quando le colonne siano indipendenti, è fare, iterativamente nell'ordine di $P^1,\dots,P^k$, partendo da $R^1 = P^1$, la proiezione R^i ortogonale della colonna P^i sullo spazio vettoriale *ortogonale a quello generato dalle precedenti*: è il *procedimento di ortogonalizzazione di Gram-Schmidt senza normalizzare* i vettori ottenuti, indicando con $\widehat{v}$ il versore di un vettore $v \ne \vec{0}$:

$$R^h = P^h - \sum_{i=1}^{h-1}\langle P^h\cdot\widehat{R^i}\rangle_m\widehat{R^i},\ h>1,\quad R^1 = P^1.$$

- - A livello algebrico si ottiene una matrice $R = (R^1|\dots|R^k)$,$m\times k$, con colonne ortogonali, per cui ${}^tRR$ è la matrice *diagonale quadrata* $k\times k$, con elementi diagonali $|R^1|^2_{\mathbf{R}^m},\dots,|R^k|^2_{\mathbf{R}^m}$, per cui $\det{}^tRR = |R^1|^2_{\mathbf{R}^m}\cdot\dots\cdot|R^k|^2_{\mathbf{R}^m}$.
- - A livello geometrico si ottiene un *rettangolo* di dimensione k.

- - - Algebricamente si ha: ${}^tRR = {}^tR(R^1|\dots|R^k) = ({}^tRR^1|\dots|{}^tRR^k)$, ogni colonna ${}^tRR^h$ è ${}^tRP^h$+*combinazione lineare delle precedenti*: $\quad \det {}^tRR = \det {}^tRP = \det {}^tPR = \det {}^tPP$.
- - - Geometricamente: l'h-volume, *qual che sia la sua defìnzione*, $V_k(S)$ di un parallelepipedo h-dimensionale S, si vuole coincidente con il *prodotto* del volume $(h-1)$-dimensionale di una *"base "* (che è $(h-1)$-parellelepipedo generato dai primi $h-1$ spigoli dall'origine,le prime $h-1$ colonne delle matrici in giuoco), *moltiplicato per "l'altezza"* (la lunghezza della proiezione ortogonale dell'h-simo spigolo dall'origine sull'ortogonale del sottospazio generato dai precedenti). Quindi *qual che sia la definizone di volume*, iterativamente si deve ottenere:

$$V_k(P) = V_k(R).$$

Nel caso del k-rettangolo, generato da $R^1,\dots,R^k$ tale volume $V(R)$ si vuole che *sia anche il prodotto delle lunghezze degli spigoli* generatori. Quindi concludendo:

$$\det {}^tPP = \det {}^tRR = |R^1|^2_{\mathbf{R}^m}\cdot\dots\cdot|R^k|^2_{\mathbf{R}^m} = V(R)^2 = V(P)^2.$$

Osservazione: da tale interpretazione e dalla formula di Cauchy-Binet:
Pitagora per k-parallelepipedi: se P è un k-parellelepipedo in $\mathbf{R}^m$, $k<m$:

$$V_k(P) = \sqrt{\text{somma quadrati dei } (k-1)\text{-volumi delle proiezioni ortogonali sui } (k-1)\text{-piani coordinati.}}$$

Derivate di prodotti di cammini

Derivate di prodotti: se $P : \mathbf{R}^{M_1}\times\dots\times\mathbf{R}^{M_k} \to \mathbf{R}^m$ è multilineare, e $f^i : I \to \mathbf{R}^{M_i}$, $1\le i\le k$, sono cammini derivabili in t, allora $g(t) = P(f^1(t),\dots,f^k(t))$ è derivabile in t e:

$$\frac{d\left(f^1\cdot_P\dots\cdot_P f^k\right)}{dt}(t) = \sum_{i=1}^{k}\dots f^{i-1}(t)\cdot_P\frac{df^i}{dt}(t)\cdot_P f^{i+1}(t)\dots.$$

La dimostrazione è di tipo induttivo sul numero di fattori k. Si può supporre, passando alle componenti, che $m=1$. Il passo induttivo ricalca la dimostrazione per $k=2$ (caso di funzioni bilineari). A sua volta questa ricalca quella della derivata del prodotto di funzioni a valori reali, oppure si fa per componenti. Si esemplifica quindi solo la dimostrazione nel caso di funzioni bilineari e si denota con A la matrice $M_1\times M_2$ associata a P:

$$\frac{f^1(t+h)\cdot_P f^2(t+h) - f^1(t)\cdot_P f^2(t)}{h} - \left(f^{1\prime}(t)\cdot_P f^2(t) + f^1(t)\cdot_P f^2(t)'\right) =$$
$$= \frac{f^1(t+h)\cdot_P f^2(t+h) - f^1(t+h)\cdot_P f^2(t) + f^1(t+h)\cdot_P f^2(t) - f^1(t)\cdot_P f^2(t)}{h} +$$
$$-\left(f^{1\prime}(t)\cdot_P f^2(t) + f^1(t)\cdot_P f^2(t)'\right) = \text{per bilinearità}$$
$$= f^1(t+h)\cdot_P\frac{f^2(t+h)-f^2(t)}{h} + \frac{f^1(t+h)-f^1(t)}{h}\cdot_P f^2(t) - \left(f^{1\prime}(t)\cdot_P f^2(t) + f^1(t)\cdot_P f^2(t)'\right) =$$
$$= f^1(t)\cdot_P\left(\frac{f^2(t+h)-f^2(t)}{h} - f^2(t)'\right) +$$
$$+(f^1(t+h)-f^1(t))\cdot_P\frac{f^2(t+h)-f^2(t)}{h} + \left(\frac{f^1(t+h)-f^1(t)}{h} - f^{1\prime}(t)\right)\cdot_P f^2(t),$$

per Cauchy-Schwarz $|v^1\cdot_P v^2| \le C(A)|v^1|_{M_1}|v^2|_{M_2}$: tutti e tre gli addendi sono infinitesimi.
Derivata del determinante. - Data una matrice A, quadrata $m\times m$, si definiscono: la matrice $cofA$ dei *cofattori* $(cofA)^k_h = (-1)^{h+k}detA^{\not k}_{\not h}$ e la sua trasposta, la matrice *aggiunta* $adjA$, $(adjA)^k_h = (-1)^{h+k}detA^{\not h}_{\not k}$. Se A è invertibile $(detA)A^{-1} = adjA$.
- A titolo di esempio se $A : I \to \mathcal{M}_{m\times m}$ è un cammino di matrici derivabili il determinante (essendo un polinomio nelle funzione coefficienti $A^j_i(t)$) è derivabile e si ha:

$$(detA)' = cofA\cdot_{\mathcal{M}} A' = tr(adjA\ A')$$

nel caso in cui A sia invertibile si ha quindi $\qquad (detA)' = (detA)\ tr(A^{-1}A')$

Dimostrazione: applicando la formula per funzioni m lineari e lo sviluppo per la colonna j^a del determinante: $(detA)' = \sum_{j=1}^m det\left(\dots A^{j-1}A^{j\prime}A^{j+1}\dots\right) = \sum_{j=1}^m\sum_{i=1}^m (A^j_i)'(-1)^{i+j}det\left(A^{\not j}_{\not i}\right) =$
$= A'\cdot_{\mathcal{M}}(cofA)$.

Moti piani di rotazione. - un moto di rotazione attorno ad un asse per l'origine nello spazio, su piani ad esso ortogonali, è dato da $P(t) = A(t)P_0$, $V(t) = P'(t)$: P_0 posizione iniziale, $A : \mathbf{R} \to \mathcal{M}_{3\times 3}$ cammino di matrici ortonormali ${}^tAA = Id$, $A(0) = Id$, con stesso autovettore v di autovalore 1: Av=v, che identifica il comune *asse di rotazione costante*.
- Siano Ω il *vettore di velocità angolare* $\Omega = \pm\omega v$, e (x_1, x_2, x_3) le coordinate cartesiane di $P(t)$: $\quad \Omega \times P = \tilde{\Omega} = \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $\; P'(t) = V(t) = \Omega(t) \times P(t)$.
- Derivando le relazione di ortogonalità si ottiene ${}^tA'A + {}^tAA' = 0_{\mathcal{M}_{3\times 3}}$, la matrice nulla. In altri termini $A'\,{}^tA$ è *antisimmetrica* ${}^tB = -B$.
- Essendo fisso l'asse di rotazione si ottiene anche $A'v = 0_{\mathbf{R}^3}$
- Si ha $P'(t) = A'(t)P_0$ da una parte, dall'altra $P'(t) = \Omega(t) \times P(t) = \tilde{\Omega}A(t)P_0$ e questo per ogni condizione iniziale P_0. Pertanto

$$A'(t) = \tilde{\Omega}A(t) \quad , \quad \text{per ortonormalità di } A(t) \text{ quindi } \tilde{\Omega} = A'(t)\,{}^tA(t).$$

[B] per V.Barutello et al. Analisi mat. vol. 2;

[F] per N.Fusco et al. An.Mat. due;

[FS] per N.Fusco et al. Elem. di An. Mat. due, versione semplificata.

[B] cap. III, cap.IV, in particolare per le forme quadratiche IV.5, 6 pagg.197-210.
[F] cap.2.16, 17, 19, in particolare per le forme quadratiche 3.36, 40.

DETERMINANTI E PRODOTTO VETTORIALE
[B] pagg. 112-126;
[F] pagg. 334-338.