MATHESIS DI PADOVA

PSEUDO-VOLUME, PSEUDO-INVERSA E DECOMPOSIZIONI IN VALORI SINGOLARI DI MATRICI REALI

Luigi Salce

Padova, 19 Novembre 2010

Ho letto recentemente un vecchio articolo divulgativo di G.B. Price, apparso sull'American Mathematical Mothly del 1947, dal titolo: "Some identities in the theory of determinants", in cui sono fornite varie dimostrazioni di alcuni tra i più importanti risultati sui determinanti.

Ricordo che i determinanti erano noti nell'antica Cina fin dal 1300 d.C.. L'espansione di un determinante secondo righe o colonne era nota al matematico giapponese Seki Shinsuke Kowa, detto Takakazu, nella seconda metà del 1600, mentre in Europa i determinanti furono introdotti da Cramer a Ginevra nel 1750 e da Vandermonde nel 1771.

I risultati ricordati nell'articolo di Price testimoniano il grande interesse ed il fervore delle ricerche sui determinanti in tutto il XIX secolo, interesse che si affievolì però con il passare dei decenni. A testimonianza di ciò, Price mette a confronto due frasi di Felix Klein (1849-1925) e di C.C. MacDuffee, autore del noto libro: "Vectors and Matrices" del 1943, scritte a distanza di mezzo secolo.

Scrive Klein: "Cayley mi disse un giorno, conversando, che se avesse dovuto fare quindici conferenze su tutta la matematica, ne avrebbe riservata una ai determinanti".

Scrive MacDuffee: "L'importanza della nozione di determinante è stata, ed è a tutt'oggi, molto sovra-stimata. I sistemi di equazioni si possono risolvere in modo semplice e pulito senza i determinanti, come illustrato in questo libro. In realtà, forse il 90% della teoria delle matrici si può sviluppare senza neppure menzionare i determinanti. Tuttavia questa nozione è necessaria in qualche occasione, ed è molto utile in molte altre, perciò non è il caso di insistere troppo con questo punto di vista".

Price nel suo articolo conclude dicendo: " I matematici al giorno d'oggi (ricordo che siamo nel 1947), con giudizio avvalorato dall'esperienza, probabilmente sarebbero d'accordo nel sostenere che la teoria dei determinanti appartiene alla categoria delle "teorie strumentali" e non è oggetto di grande interesse di per sè".

Con l'esperienza di altri 60 anni, si può dire che leggendo le riviste specializzate di Algebra Lineare o di Teoria delel Matrici, sono sempre molti gli articoli, più o meno brillanti, più o meno profondi, che trattano dei determinanti. In questa conferenza mostrerò un interessante sviluppo recente (del 1992) di questa nozione.

PREMESSA. Una matrice reale 3x3 è invertibile esattamente quando il suo determinante è non nullo. In tal caso il determinante rappresenta il volume orientato del parallelepipedo generato dai vettori colonna della matrice.
Se la matrice non è invertibile, il parallelepipedo è degenere ed il suo volume è nullo, come pure il determinante della matrice.

DOMANDA. Per le matrici quadrate nxn a determinante nullo si può trovare una qualche nozione che generalizza la invertibilità ed una qualche nozione che generalizza il volume?

La risposta è positiva e si estende a matrici non solo quadrate, ma qualunque; essa si basa sulla **decomposizione in valori singolari** delle matrici e si ottengono le nozioni di **matrice pseudo-inversa** e di **pseudo-volume.**

1. MATRICI E APPLICAZIONI LINEARI

Ogni matrice reale nxn A si può interpretare come una applicazione lineare

$$f_A : R^n \to R^n \quad , \quad \mathbf{v} \mapsto A\mathbf{v} .$$

Mettiamo in evidenza le colonne di A:

$$A = [\mathbf{a}_1\ \mathbf{a}_2\ ....\ \mathbf{a}_n].$$

La j-esima colonna $\mathbf{a}_j$ di A non è altro che $A\mathbf{e}_j$, dove $\mathbf{e}_j = [0\ ...\ 010\ ...\ 0]^T$ è il j-esimo versore di R.

Quindi il parallelepipedo generato dalle colonne $\mathbf{a}_1, \mathbf{a}_2, ... , \mathbf{a}_n$ di A (che risulta non degenere se e solo se tali colonne sono linearmente indipendenti, ovvero, se $Det(A) \neq 0$) non è altro che il trasformato tramite f_A dell'ipercubo unitario generato dagli n versori $\mathbf{e}_1, \mathbf{e}_2, ... , \mathbf{e}_n$ di R^n; il suo volume è $|\ Det(A)\ |$.

Se invece dell'ipercubo unitario generato dai versori consideriamo un altro ipercubo unitario di R^n generato da altri n vettori di norma 1 a due a due ortogonali $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$, il trasformato tramite f_A di tale ipercubo ha lo stesso volume (non orientato) del precedente:

infatti il suo volume è il modulo di

$$\mathrm{Det}([A\mathbf{q}_1\ A\mathbf{q}_2\ \ldots.\ A\mathbf{q}_n])$$

ed è facile vedere che

$$[A\mathbf{q}_1\ A\mathbf{q}_2\ \ldots.\ A\mathbf{q}_n] = A \cdot Q$$

dove $Q = [\mathbf{q}_1\ \mathbf{q}_2\ \ldots.\ \mathbf{q}_n]$. Essendo Q una matrice a colonne ortonormali, essa soddisfa a $Q^T \cdot Q = I_n$. Ciò comporta che $1 = \mathrm{Det}(I_n) = \mathrm{Det}(Q^TQ) = \mathrm{Det}(Q)^2$.

Perciò si ha:

$$\mathrm{Det}([A\mathbf{q}_1\ A\mathbf{q}_2\ \ldots.\ A\mathbf{q}_n]) = \mathrm{Det}(A \cdot Q) = \mathrm{Det}(A)\mathrm{Det}(Q) = \pm\ \mathrm{Det}(A).$$

Possiamo concludere che:

il modulo del determinante di una matrice reale nxn A coincide con il volume dell'iperparallelepipedo che si ottiene trasformando per mezzo della applicazione lineare f_A un qualunque ipercubo generato da n vettori ortonormali di R^n

Se invece Det(A) = 0 (equivalentemente, se A non è invertibile), l'iper-parallelepipedo formato dalle colonne linearmente dipendenti di A diventa degenere, ovvero, passa dalla dimensione n ad una inferiore (uguale al rango della matrice).

Ma trattare alla stessa stregua ogni matrice A con Det(A) = 0 è un modo di procedere un pò "grezzo".

2. I 4 SOTTOSPAZI FONDAMENTALI DI UNA MATRICE

Una generica matrice reale mxn A si può interpretare, come già visto nel caso quadrato, come una applicazione lineare

$$f_A : R^n \to R^m \quad , \quad \mathbf{v} \mapsto A\mathbf{v} .$$

Il dominio R^n si decompone nella somma diretta di due sottospazi, uno il **complemento ortogonale** dell'altro (il prodotto interno tra due vettori **u** e **v** è dato al solito da $\mathbf{u}^T\mathbf{v}$):

$$R^n = C(A^T) \oplus N(A)$$

dove

- $C(A^T)$ è lo **spazio "delle righe"** di A, cioè il sottospazio di R^n generato dalle righe di A viste come vettori colonna (quindi dalle colonne della trasposta A^T)

- N(A) è lo **spazio nullo** di A, formato dai vettori **v** annullati da A, tali cioè che $A\mathbf{v} = \mathbf{0}$ (quindi dalle soluzioni del sistema omogeneo associato ad A).

Il codominio R^m si decompone a sua volta nella somma diretta di due sottospazi, uno il **complemento ortogonale** dell'altro:

$$R^m = C(A) \oplus N(A^T)$$

dove

- $C(A)$ è lo **spazio delle colonne** di A, che è il sottospazio di R^m generato dalle colonne di A

- $N(A^T)$ è lo **spazio nullo sinistro** di A, formato dai vettori **u** annullati da A^T, tali cioè che $A^T\mathbf{u} = \mathbf{0}$.

Si ha dunque

$$R^n = \begin{array}{c} C(A^T) \\ \oplus \\ N(A) \end{array} \xrightarrow{f_A} \begin{array}{c} C(A) \\ \oplus \\ N(A^T) \end{array} = R^m$$

dove $f_A(N(A)) = 0$, mentre f_A induce un **isomorfismo** tra $C(A^T)$ e $C(A)$

($C(A^T) \cap N(A) = 0$ dà l'iniettività e $rk(A^T) = rk(A)$ la suriettività).

Quindi f_A trasforma k vettori ortogonali di norma 1 di $C(A^T)$ (cioè un ipercubo k-dimensionale unitario dello spazio $C(A^T)$ che ha dimensione k) in un iperparallelepipedo k-dimesionale dello spazio C(A) che ha pure dimensione k

Nel caso di una matrice quadrata invertibile A , accade che $N(A) = 0 = N(A^T)$, f_A è un isomorfismo, e risulta:

$$C(A^T) = R^m = C(A) .$$

Quindi lo spazio delle righe (come pure lo spazio delle colonne) contiene l'ipercubo generato dai versori $\mathbf{e}_1$, $\mathbf{e}_2$, ... , $\mathbf{e}_n$ di R^n, cosa che non accade in generale per il sottospazio $C(A^T)$ di R^n per una matrice qualunque A.

ESEMPIO. Consideriamo l'esempio di una matrice 2x4 di rango 1:

$$A = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix}, \qquad f_A : R^4 \to R^2$$

E' immediato verificare che risulta

$$R^4 = \underset{C(A^T)}{< \begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix} >} \oplus \underset{N(A)}{< \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 0 \\ 2 \end{bmatrix} >}$$

Si noti che i vettori nella base di $N(A)$, ortogonali a quello che genera $C(A^T)$, non sono ortogonali tra loro; si possono ortonormalizzare per poter entrare come colonne in una matrice V a colonne ortonormali (il che equivale a dire che $V^{-1} = V^T$, ovvero che V è ortogonale), producendo:

$$V = \begin{bmatrix} 2/\sqrt{14} & 0 & 0 & 5/\sqrt{35} \\ 0 & 1 & 0 & 0 \\ 1/\sqrt{14} & 0 & 3/\sqrt{10} & -1/\sqrt{35} \\ 3/\sqrt{14} & 0 & -1/\sqrt{10} & -3/\sqrt{35} \end{bmatrix} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \mathbf{v}_4]$$

Per il codominio si ha :

$$R^2 = \underbrace{< \begin{bmatrix} 2 \\ 1 \end{bmatrix} >}_{C(A)} \oplus \underbrace{< \begin{bmatrix} 1 \\ -2 \end{bmatrix} >}_{N(A^T)}$$

ed i due vettori normalizzati diventano colonne della matrice ortogonale

$$U = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} = [\mathbf{u}_1 \ \mathbf{u}_2]$$

Un calcolo diretto mostra che $f_A(\mathbf{v}_1) = \sqrt{70}\ \mathbf{u}_1$.

Ci si chiede: c'è un modo di "invertire" l'azione di f_A tramite una applicazione $f_{A^+}: R^2 \to R^4$ indotta da una opportuna matrice 4x2 A^+ in modo che la restrizione di f_A da $C(A^T)$ a $C(A)$, sia invertita dalla restrizione di f_{A^+} a $C(A)$?

3. PSEUDO-INVERSA E PSEUDO-VOLUME DI UNA MATRICE

L'idea che guida il tentativo di estendere ad una matrice rettangolare la nozione di matrice inversa e quella di volume, è di considerare ciò che accade nello spazio delle righe $C(A^T)$, in cui f_A si comporta come un isomorfismo.

A tal fine, risulta determinante la **decomposizione in valori singolari** di A, valida per una generica matrice complessa. Noi esamineremo il caso reale.

Una matrice reale mxn A di rango k si decompone nel prodotto di tre matrici:

$$\text{(SVD)} \qquad A = U \cdot S \cdot V^T$$

- U è ortogonale mxm (cioè a colonne ortonormali, quindi $U^T = U^{-1}$)
- V è ortogonale nxn (cioè a colonne ortonormali, quindi $V^T = V^{-1}$)
- S è mxn con k coefficienti positivi sulla pseudo-diagonale: $s_1 \geq s_2 \geq \ldots \geq s_k > 0$

e tutti gli altri coefficienti nulli.

I numeri reali positivi $s_1 \geq s_2 \geq \ldots \geq s_k$ si chiamano **valori singolari** di A e sono da A univocamente individuati mentre U e V possono opportunamente variare.

In quanto matrici ortogonali, U e V hanno le colonne che costituiscono basi ortonormali rispettivamente di R^m e di R^n. Ma si può dire ben di più:

- le prime k colonne di U formano una base ortonormale di C(A)
- le ultime m-k colonne di U formano una base ortonormale di $N(A^T)$

$$U = [\underbrace{\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_{k-1}\ \mathbf{u}_k}_{\text{base o. n. di } C(A)}\quad \underbrace{\mathbf{u}_{k+1} \ldots\ \mathbf{u}_{m-1}\ \mathbf{u}_m}_{\text{base o. n. di } N(A^T)}] = [U_1\ \ U_0]$$

- le prime k colonne di V formano una base ortonormale di $C(A^T)$
- le ultime n-k colonne di V formano una base ortonormale di N(A)

$$V = [\underbrace{\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_{k-1}\ \mathbf{v}_k}_{\text{base o.n. di } C(A^T)}\quad \underbrace{\mathbf{v}_{k+1} \ldots\ \mathbf{v}_{n-1}\ \mathbf{v}_n}_{\text{base o.n. di } N(A)}] = [V_1\ \ V_0]$$

Ricordando che $V^T = V^{-1}$, la SVD è equivalente a

$$A \cdot V = U \cdot S$$

ovvero a:

$$(*) \qquad A\,\mathbf{v}_j = s_j\,\mathbf{u}_j \qquad \text{per } j \le k$$

$$A\,\mathbf{v}_j = \mathbf{0} \qquad \text{per } k < j \le n\,.$$

Mentre per $j \le k$ i vettori $\mathbf{v}_j$ e $\mathbf{u}_j$ sono legati dalle relazioni (*), le ultime m-k colonne di U e le ultime n-k colonne di V possono essere scelte in modo qualunque, purchè formino basi ortonormali di $N(A^T)$ e di $N(A)$, rispettivamente.

Quindi la applicazione lineare f_A manda l'ipercubo k-dimensionale contenuto in $C(A^T)$ generato dalle prime k colonne di V nell'iperparallelepipedo rettangolo k-dimensionale di $C(A)$ generato dai vettori $s_j\mathbf{u}_j$, ognuno di lunghezza s_j. Tale iperparallelepipedo ha ovviamente volume $\prod_j s_j$.

DEFINIZIONE. Chiamasi **pseudo-volume** di una matrice reale mxn A il prodotto dei valori singolari di A:

$$\mathbf{vol^+(A) = \prod_j s_j} .$$

DEFINIZIONE. Chiamasi **pseudo-inversa** di una matrice mxn A la matrice

$$\mathbf{A^+ = V \cdot S^+ \cdot U^T}$$

dove S^+ è la matrice nxm con k coefficienti positivi $s_1^{-1} \leq s_2^{-1} \leq \ldots \leq s_k^{-1}$ sulla pseudo-diagonale e tutti gli altri coefficienti nulli.

Quindi S^+ si ottiene trasponendo S e invertendone i coefficienti "ove possibile".

Naturalmente le nuove definizioni sono buone se, nel caso di una matrice nxn invertibile A, si riottengono le usuali nozioni di volume e di matrice inversa.

Se A è invertibile (quindi $k=n$), la pseudo-inversa A^+ viene a coincidere con A^{-1} giacchè in tal caso $S^+ = S^{-1}$ (e $U^T = U^{-1}$, $V^T = V^{-1}$).

Per quanto riguarda lo pseudo-volume, bisogna vedere che $\prod_j s_j = |\ \mathrm{Det}(A)\ |$.

Ciò dipende dal seguente fatto:

i valori singolari s_j sono le radici quadrate degli autovalori λ_j (>0) di AA^T

come si vede calcolando $AA^T = USV^TVS^TU^T = USS^TU^T$.

Quindi si ha:

$$\mathrm{vol}^+(A)^2 = \prod_j s_j^2 = \prod_j \lambda_j = \mathrm{Det}(AA^T) = \mathrm{Det}(A)^2.$$

Si osservi che per una generica matrice A mxn si ha:

$$\mathrm{vol}^+(A^+) = \prod_j s_j^{-1} = 1 / \prod_j s_j = 1 / \mathrm{vol}^+(A).$$

NOTIZIE STORICHE

1) *La decomposizione in valori singolari venne provata per matrici quadrate reali da Sylvester nel 1889, ma la prima dimostrazione per una generica matrice complessa mxn venne data solo nel 1939 da Eckart e Young.*

2) *La nozione di matrice pseudo-inversa (detta "di Moore-Penrose") risale agli anni '20 del secolo scorso; fu introdotta dal E.H. Moore, che ne provò l'unicità e l'esistenza senza usare la decomposizione in valori singolari. La sua scoperta non fu adeguatamente divulgata. La pseudo-inversa divenne invece una solida e usatissima acquisizione quando R. Penrose la riscoprì nel 1955.*

3) *La nozione di pseudo-volume è molto più recente, essendo stata introdotta da Ben Israel in un articolo del 1992, anche se è intrinsecamente legata alla nozione di "compound matrix", nota a Sylvester già a metà '800.*

ESEMPI

I) Riconsideriamo l'esempio della matrice 2x4 di rango 1:

$$A = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix}$$

Si ha la SVD della matrice $A = U\, S\, V^T$:

$$A = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{70} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2/\sqrt{14} & 0 & 1/\sqrt{14} & 3/\sqrt{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3/\sqrt{10} & -1/\sqrt{10} \\ 5/\sqrt{35} & 0 & -1/\sqrt{35} & -3/\sqrt{35} \end{bmatrix}$$

L'unico valore singolare di A è $s_1 = \sqrt{70}$.

Infatti

$$A\,A^T = \begin{bmatrix} 56 & 28 \\ 28 & 14 \end{bmatrix}$$

ha come polinomio caratteristico $p_{AA^T}(X) = X^2 - 70\,X$, quindi AA^T ha come autovalori 70 e 0, conseguentemente

$$vol^+(A) = \sqrt{70}.$$

La pseudo-inversa di A risulta essere:

$$A^+ = \begin{bmatrix} 2/\sqrt{14} & 0 & 0 & 5/\sqrt{35} \\ 0 & 1 & 0 & 0 \\ 1/\sqrt{14} & 0 & 3/\sqrt{10} & -1/\sqrt{35} \\ 3/\sqrt{14} & 0 & -1/\sqrt{10} & -3/\sqrt{35} \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{70} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}$$

che, fatti i conti, coincide con $70^{-1}\,A^T$.

II) Consideriamo la matrice 1x3:

$$A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \qquad f_A : R^3 \to R$$

$$A\,A^T = \begin{bmatrix} 3 \end{bmatrix} \qquad p_{AA^T}(X) = X - 3$$

SVD: $$A = \begin{bmatrix} 1 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{3} & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -2/\sqrt{3} \end{bmatrix}$$

$$vol^+(A) = \sqrt{3} = (\, \Sigma_{ij}\, a_i^2 \,)^{1/2} = \; \|\, A \,\|_2$$

$$A^+ = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & 0 & -2/\sqrt{3} \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{3} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = 1/(\|A\|_2)^2\, A^T$$

III) Consideriamo la matrice 2x3:

$$B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad f_B : R^3 \rightarrow R^2$$

$$B\,B^T = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \qquad p_{BB^T}(X) = X^2 - 6X = X(X-6)$$

SVD:

$$B = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -2/\sqrt{3} \end{bmatrix}$$

$$vol^+(B) = \sqrt{6} = \left(\Sigma_{ij}\, b_{ij}^2 \right)^{1/2} = \|\, B \,\|_F$$

$$B^+ = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & 0 & -2/\sqrt{3} \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{6} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = 1/6 \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = 1/(\|B\|_F)^2\, B^T$$

IV) Consideriamo la matrice 3x3:

$$C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad f_C : R^3 \to R^3$$

$$C\,C^T = \begin{bmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{bmatrix} \qquad p_{CC^T}(X) = X^3 - 9X^2 = X^2(X-9)$$

$$\text{SVD:} \quad C = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & 0 & -2/\sqrt{3} \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -2/\sqrt{3} \end{bmatrix}$$

$$\mathrm{vol}^+(C) = \sqrt{9} = \left(\textstyle\sum_{ij} c_{ij}^2 \right)^{1/2} = \| C \|_F$$

$$C^+ = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & 0 & -2/\sqrt{3} \end{bmatrix} \cdot \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & -1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -2/\sqrt{3} \end{bmatrix} = 1/9\, C^T$$

Negli esempi precedenti le matrici hanno sempre rango 1. E' questo il motivo per cui lo pseudo-volume coincide con la norma di Frobenius. Però in generale non è vero che $vol^+(A) = \| A \|_F$, come mostra l'esempio seguente.

V) Consideriamo la matrice 3x2 di rango 2:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} \qquad AA^T = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

valori singolari: $\sqrt{2}$, 2

quindi: $vol^+(A) = 2\sqrt{2} \neq \sqrt{6} = \| A \|_F$

Lo pseudo-volume $vol^+(A)$ si può sì vedere come norma di Frobenius, ma non della matrice A, bensì di un'altra matrice "composta" a partire da A.

4. PROPRIETA' DELLA MATRICE PSEUDO-INVERSA

(1) *La matrice pseudo-inversa A^+ di una matrice reale mxn A è l'unica matrice nxm tale che*

$$A\,A^+A = A \quad , \quad A^+A\,A^+ = A^+$$

$$A\,A^+ = (A\,A^+)^T \quad , \quad A^+A = (A^+A)^T$$

Si deduce che AA^+ e A^+A sono matrici di proiezione, cioè sono idempotenti e simmetriche; AA^+ proietta R^m su $C(A)$, AA^+ proietta R^n su $C(A^T)$.

Si conferma che $f_A : R^n \rightarrow R^m$ e $f_{A^+} : R^m \rightarrow R^n$ sono una l'inversa dell'altra, se ristrette rispettivamente a $C(A^T)$ e $C(A)$; infatti

$$f_{A^+}f_A(A^T\underline{v}) = A^+AA^T\underline{v} = A^T\underline{v} \qquad \text{giacchè } A^+AA^T = A^T$$

$$f_Af_{A^+}(A\underline{v}) = AA^+A\underline{v} = A\underline{v} \qquad \text{giacchè } AA^+A = A.$$

(2) *Dato un sistema di equazioni lineari* $A\,\underline{x} = \underline{b}$, *la soluzione ai minimi quadrati di norma minima è il vettore* $A^{+}\underline{b}$.

Ricordiamo che le soluzioni ai minimi quadrati di $A\underline{x} = \underline{b}$ sono i vettori $\underline{z}$ di $C(A)$ (quindi del tipo $\underline{z} = A\underline{v}$) che minimizzano la distanza da $\underline{b}$, cioè $\| \underline{b} - \underline{z} \|_2$.

5. PROPRIETA' DELLO PSEUDO-VOLUME

Sia A una generica matrice mxn di rango k.

(1) $$vol^{+}(A) = \left(\Sigma_{I,J}\, Det^{2}(A_{IJ}) \right)^{1/2}$$

dove:

- *I è un insieme di k indici di riga, compresi tra 1 ed m*
- *J è un insieme di k indici di colonna, compresi tra 1 ed n*
- *A_{IJ} è la sottomatrice kxk di A ottenuta intersecando le righe indiciate da I con le colonne indiciate da J.*

Si capisce che entra in gioco la matrice *k-composta di A:*

$$A^{(k)} = [\, Det(A_{IJ}) \,]_{(IJ)}$$

in cui, avendo ordinato lessicograficamente le k-ple di indici I e J, il coefficiente di posto (I,J) è il determinante $Det(A_{IJ})$, $A^{(k)}$ ha dimensioni $\binom{m}{k} \times \binom{n}{k}$.

Ad esempio, se m=3, n=4 e k=2, la matrice 2-composta di A è la seguente matrice 3x6 $A^{(2)}$:

$$\begin{bmatrix} \mathrm{Det}\ A_{(12,12)} & \mathrm{Det}\ A_{(12,13)} & \mathrm{Det}\ A_{(12,14)} & \mathrm{Det}\ A_{(12,23)} & \mathrm{Det}\ A_{(12,24)} & \mathrm{Det}\ A_{(12,34)} \\ \mathrm{Det}\ A_{(13,12)} & \mathrm{Det}\ A_{(13,13)} & \mathrm{Det}\ A_{(13,14)} & \mathrm{Det}\ A_{(13,23)} & \mathrm{Det}\ A_{(13,24)} & \mathrm{Det}\ A_{(13,34)} \\ \mathrm{Det}\ A_{(23,12)} & \mathrm{Det}\ A_{(23,13)} & \mathrm{Det}\ A_{(23,14)} & \mathrm{Det}\ A_{(23,23)} & \mathrm{Det}\ A_{(23,24)} & \mathrm{Det}\ A_{(23,34)} \end{bmatrix}$$

La formula (1) dice che lo pseudo-volume è la norma di una matrice:

$$\mathrm{vol}^{+}(A) = \left(\Sigma_{I,J}\ \mathrm{Det}^{2}(A_{IJ}) \right)^{1/2} = \|\ A^{(k)}\ \|_{F}\ .$$

(2) *Date una matrice A mxn ed una matrice B nxp, risulta*

$$vol^{+}(AB) \leq vol^{+}(A)\, vol^{+}(B)$$

e vale l'uguaglianza se $N(A) = N(B^T)$ *e* $N(A^T) = N(B)$.

La disuguaglianza segue da una nota formula di R. Horn sui valori singolari del prodotto di due matrici:

$$\prod_i s_i(AB) \leq \prod_i s_i(A) \prod_i s_i(B)$$

dove l' indice i fino a $k = rk(AB)$ ($\leq \min \{rk(A)\, ,\, rk(B)\}$).

Si noti che per matrici quadrate A e B invertibili le condizioni $N(A) = N(B^T)$ e $N(A^T) = N(B)$ sono certamente soddisfatte, riottenendo il fatto che il volume del prodotto di due matrici invertibili è il prodotto dei due volumi.

LO PSEUDO-VOLUME COME UN DETERMINANTE

(3) *se* $A \neq O$ *è mxn, allora:* $\qquad vol^+(A) = |\det(B(A))|$

dove B è una matrice quadrata invertibile di ordine m+n-k, detta <u>*matrice bordata associata*</u> *ad A:*

$$B(A) \quad = \begin{bmatrix} A & U_o \\ V_o^T & O \end{bmatrix}$$

con U_o *è* V_o *formate dalle ultime m-k e n-k colonne di U e V, rispettivamente.*

(4) *se* $A \neq O$ *è nxn, allora:* $\qquad vol^+(A) = |\det(C(A))|$

dove C è una matrice nxn invertibile, detta <u>*matrice complementare*</u> *di A:*

$$C(A) \quad = \quad A + U_o V_o^T .$$

ESEMPIO. Consideriamo sempre l'Esempio precedente:

$$A = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix}.$$

La matrice bordata associata di A risulta essere

$$B(A) = \left[\begin{array}{cccc|c} 4 & 0 & 2 & 6 & 1/\sqrt{50} \\ 2 & 0 & 1 & 3 & -2/\sqrt{5} \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3/\sqrt{10} & -1/\sqrt{10} & 0 \\ 5/\sqrt{35} & 0 & -1/\sqrt{35} & -3/\sqrt{35} & 0 \end{array}\right]$$

che, come si può verificare, ha come modulo del determinante $\sqrt{70}$.

ESEMPIO. Consideriamo la semplice matrice:

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$$

Una decomposizione in valori singolari di A è:

$$A = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = U \cdot S \cdot V^T$$

quindi $\mathrm{vol}^+(A) = \sqrt{5}$. Calcoliamo $B(A)$ e $C(A)$:

$$B(A) = \left[\begin{array}{cc|c} 1 & 0 & -2/\sqrt{5} \\ 2 & 0 & 1/\sqrt{5} \\ \hline 0 & 1 & 0 \end{array}\right]$$

che si verifica avere determinante $-\sqrt{5}$, mentre

$$C(A) = A + U_0 V_0^T = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2/\sqrt{5} \\ 2 & 1/\sqrt{5} \end{bmatrix}$$

che si verifica avere determinante $\sqrt{5}$.

La matrice bordata associata ad una generica matrice A

$$B(A) = \begin{bmatrix} A & U_o \\ V_o^T & O \end{bmatrix}$$

è invertibile.

Dim. $\begin{bmatrix} A & U_o \\ V_o^T & O \end{bmatrix} \cdot \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} = \begin{bmatrix} \underline{0} \\ \underline{0} \end{bmatrix} \Leftrightarrow \begin{cases} A\,\underline{x} + U_o\,\underline{y} = \underline{0} \\ V_o^T\underline{x} = \underline{0} \end{cases}$ $\Longrightarrow U_o^TA\underline{x} + U_o^TU_o\underline{y} = \underline{0}$

Ma $U_o^TA = O$ perchè $A^TU_o = O$ (le colonne di U_o sono base di $N(A^T)$) e

$U_o^TU_o = I_k$, quindi $\underline{y} = \underline{0}$.

- $A\,\underline{x} = \underline{0}$ comporta che $\underline{x} \in N(A)$
- $V_o^T\underline{x}=\underline{0}$ comporta che $\underline{x}$ è ortogonale alle colonne di V_o che generano $N(A)$

quindi anche $\underline{x} = \underline{0}$.

La matrice complementare associata ad una generica matrice A

$$C(A) = A + U_o V_o^T$$

è invertibile.

Dim.

$$(A + U_o V_o^T)\,\underline{x} = \underline{0} \quad \Leftrightarrow \quad \begin{cases} A\,\underline{x} + U_o\,\underline{y} = \underline{0} \quad \Rightarrow \quad \underline{y} = \underline{0} \\ V_o^T \underline{x} = \underline{y} \end{cases}$$

e si procede come sopra.

ESEMPIO. Matrice 3x3 di rango 2

$$A = \begin{bmatrix} -\sqrt{2} & \sqrt{2} & 1 \\ \sqrt{2} & -\sqrt{2} & 1 \\ 2\sqrt{2} & -2\sqrt{2} & 0 \end{bmatrix}$$

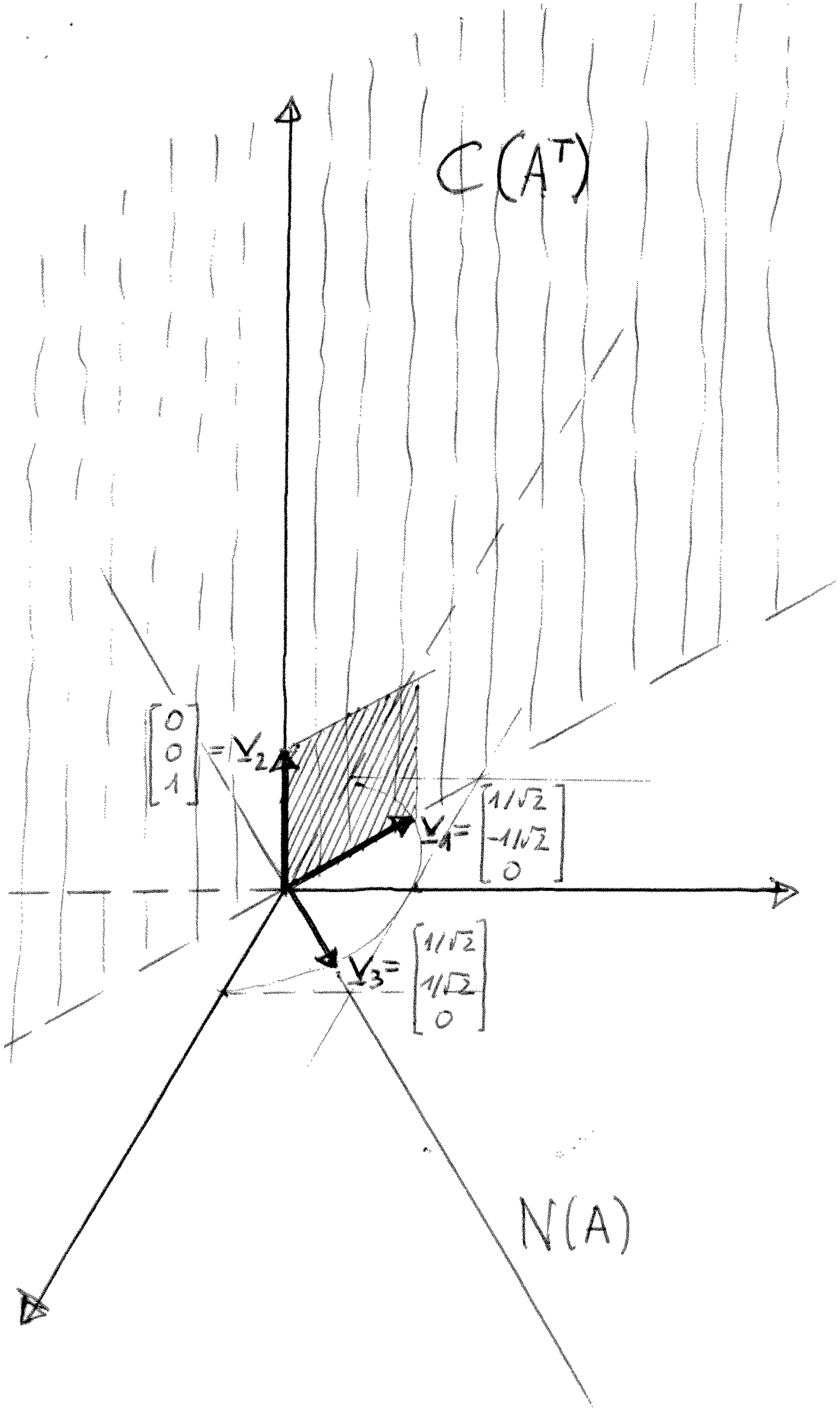
$$\underbrace{\begin{bmatrix} -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \end{bmatrix}}_{U} \cdot \underbrace{\begin{bmatrix} 2\sqrt{6} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{S} \cdot \underbrace{\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 1 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}}_{V^T}$$

$$\mathrm{vol}^{+}(A) = 2\sqrt{6}\sqrt{2} = 4\sqrt{3}$$

$$A^+ = VS^+U^T = \begin{bmatrix} -1/\sqrt{6} & 1/12\sqrt{2} & 1/6\sqrt{2} \\ 1/\sqrt{6} & -1/12\sqrt{2} & -1/6\sqrt{2} \\ 1/\sqrt{6} & 1/2 & 0 \end{bmatrix}$$

$$C(A^T) = \left\langle \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle \longleftarrow f_A \longrightarrow \left\langle \begin{bmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \right\rangle = C(A)$$

figura



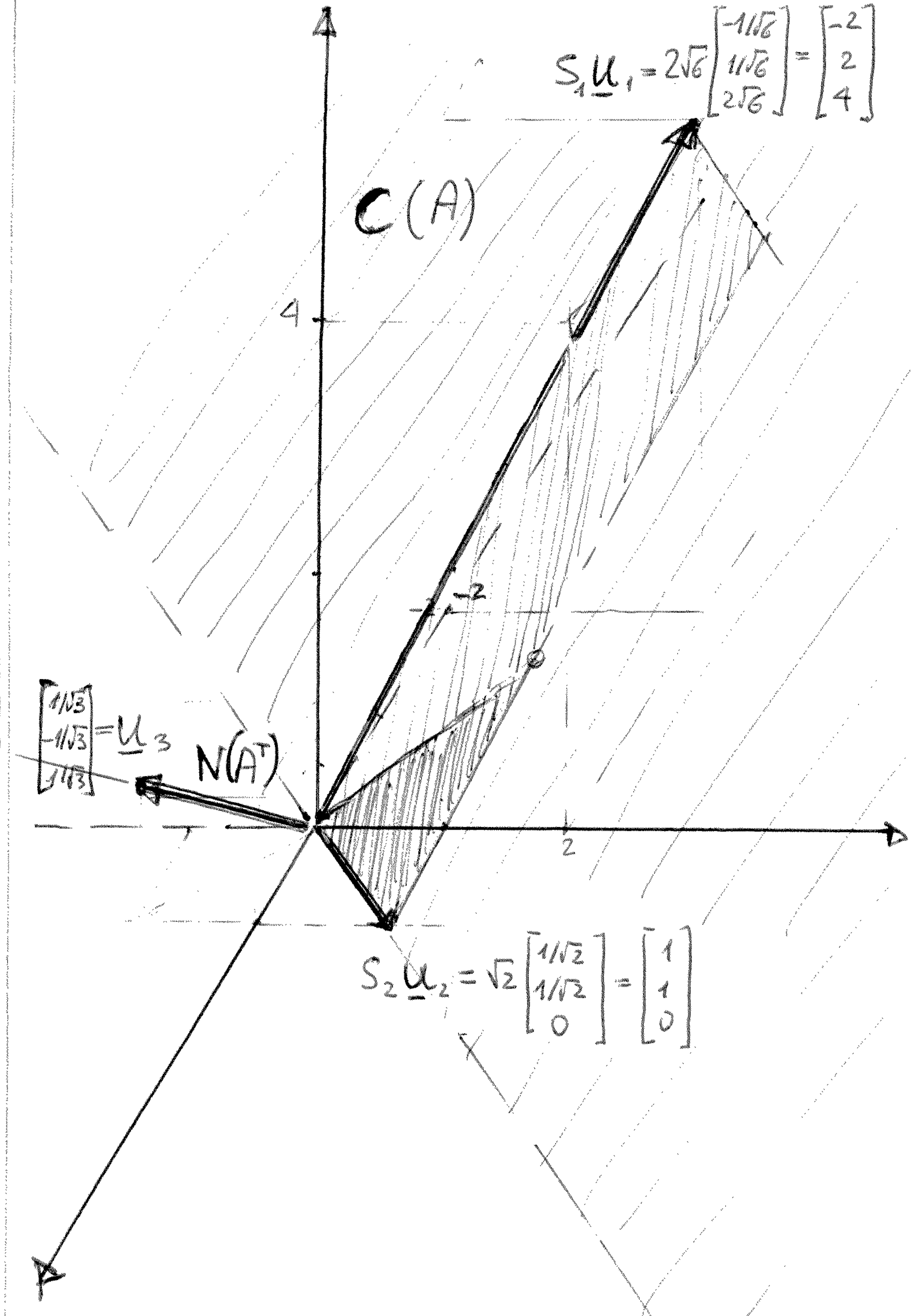
$$\xrightarrow{f_A}$$

$$A = \begin{bmatrix} -\sqrt{2} & \sqrt{2} & 1 \\ \sqrt{2} & -\sqrt{2} & 1 \\ 2\sqrt{2} & -2\sqrt{2} & 0 \end{bmatrix}$$

$$A\underline{v}_1 = s_1\underline{u}_1$$

$$A\underline{v}_2 = s_2\underline{u}_2$$

$$A\underline{v}_3 = \underline{0}$$



Calcoliamo $vol^+(A) = (\sum_{I,J} \det^2(A_{IJ}))^{1/2}$

Poichè m=3, n=3 e k=2, si hanno le seguenti coppie (I,J):

$$\begin{matrix} (12,12) & (12,13) & (12,23) \\ (13,12) & (13,13) & (13,23) \\ (23,12) & (23,13) & (23,23) \end{matrix}$$

$$A^{(2)} = \begin{bmatrix} \begin{vmatrix} -\sqrt{2} & \sqrt{2} \\ \sqrt{2} & -\sqrt{2} \end{vmatrix} & \begin{vmatrix} -\sqrt{2} & 1 \\ \sqrt{2} & 1 \end{vmatrix} & \begin{vmatrix} \sqrt{2} & 1 \\ -\sqrt{2} & 1 \end{vmatrix} \\ \begin{vmatrix} -\sqrt{2} & \sqrt{2} \\ 2\sqrt{2} & -2\sqrt{2} \end{vmatrix} & \begin{vmatrix} -\sqrt{2} & 1 \\ 2\sqrt{2} & 0 \end{vmatrix} & \begin{vmatrix} \sqrt{2} & 1 \\ -2\sqrt{2} & 0 \end{vmatrix} \\ \begin{vmatrix} \sqrt{2} & -\sqrt{2} \\ 2\sqrt{2} & -2\sqrt{2} \end{vmatrix} & \begin{vmatrix} \sqrt{2} & 1 \\ 2\sqrt{2} & 0 \end{vmatrix} & \begin{vmatrix} -\sqrt{2} & 1 \\ -2\sqrt{2} & 0 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 0 & -2\sqrt{2} & 2\sqrt{2} \\ 0 & -2\sqrt{2} & 2\sqrt{2} \\ 0 & -2\sqrt{2} & 2\sqrt{2} \end{bmatrix}$$

$$(vol^+(A))^2 = \sum_{I,J} \det^2(A_{IJ}) = 6(2\sqrt{2})^2 = 48 = (4\sqrt{3})^2.$$