Product Cosines of Angles between Subspaces

Jianming Miao and Adi Ben-Israel *

July 1, 1993

Dedicated to Professor C.R. Rao on his 75th birthday

Abstract

Let

$$\cos\{L, M\} \; := \; \prod_{i=1} \cos\, \theta_i \; ,$$

denote the product of the cosines of the **principal angles** $\{\theta_i\}$ between the subspaces L and M. The **direction cosines** of an r-dimensional subspace L are the $\binom{n}{r}$ numbers $\{\cos\{L, \mathbb{R}^n_J\} : J \in Q_{r,n}\}$ where

$Q_{r,n}$:= the set of increasing sequences of r elements from $\{1, \ldots, n\}$, and
$\mathbb{R}^n_J \; := \; \{\mathbf{x} = (x_k) \in \mathbb{R}^n : \; x_k = 0 \text{ for } k \notin J\}$.

The **basic decomposition** of a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$, with $\operatorname{rank} A = r > 0$, is

$$A \;=\; \sum_{I \in \mathcal{I}(A)} \; \sum_{J \in \mathcal{J}(A)} \cos^2\{R(A), \mathbb{R}^m_I\} \, \cos^2\{R(A^T), \mathbb{R}^n_J\} \, B_{IJ} \; ,$$

a convex combination of nonsingular linear operators $B_{IJ} : \mathbb{R}^n_J \to \mathbb{R}^n_I$. Here
$\mathcal{I}(A) \;:=\; \{I \in Q_{r,m} : \; \operatorname{rank} A_{I*} \;=\; r\}$, and $\mathcal{J}(A) \;:=\; \{J \in Q_{r,n} : \; \operatorname{rank} A_{*J} \;=\; r\}$.

The product cosines are related to the matrix **volume**, defined as the product of its nonzero singular values. The Moore-Penrose inverse $A^\dagger$ is characterized as having the minimal volume among all $\{1,2\}$-inverses of A. Indeed, if G is a $\{1,2\}$-inverse of A, with range $R(G) = T$ and null-space $N(G) = S$ then

$$\operatorname{vol} G \;=\; \frac{\operatorname{vol} A^\dagger}{\cos\{T, R(A^T)\} \, \cos\{S, N(A^T)\}} \; .$$

Key words: Principal angles. Singular values. Volume. Generalized Inverses.

*RUTCOR-Rutgers Center for Operations Research, Rutgers University, P.O. Box 5062, New Brunswick, NJ 08903-5062

1 Introduction

1.1 Notation

We use the notation and terminology of [2] and [5]. In particular:

1.1.A The **volume** of $A \in \mathbb{R}_r^{m \times n}$ is

$$\operatorname{vol} A \ := \ \begin{cases} 0 & , \text{ if } r = 0 \ , \\ \prod_{i=1}^{r} \sigma_i & , \text{ if } r > 0 \ , \end{cases} \tag{1.1}$$

where σ_i are the nonzero **singular values** of A .

1.1.B Let L, M be subspaces in $\mathbb{R}^n$, and $\dim\, L = l \ \leq \dim\, M \ = \ m$. Then the **principal angles** between L and M,

$$0 \ \leq \ \theta_1 \ \leq \ \theta_2 \ \leq \ \cdots \ \leq \ \theta_l \ \leq \ \frac{\pi}{2} \tag{1.2}$$

are defined by

$$\cos\theta_i \ := \ \frac{<\mathbf{x}_i, \ \mathbf{y}_i>}{\|\mathbf{x}_i\| \, \|\mathbf{y}_i\|} \ = \ \max \left\{ \frac{<\mathbf{x}, \ \mathbf{y}>}{\|\mathbf{x}\| \, \|\mathbf{y}\|} \ : \ \begin{array}{l} \mathbf{x} \in L \ , \ \mathbf{x} \perp \mathbf{x}_k, \\ \mathbf{y} \in M \ , \ \mathbf{y} \perp \mathbf{y}_k, \end{array} \ , \ k = 1, \ldots, i-1 \right\}, \tag{1.3}$$

where

$$(\mathbf{x}_i, \ \mathbf{y}_i) \ \in \ L \times M, \ i = 1, \ldots, l \ , \tag{1.4}$$

are the corresponding l pairs of **principal vectors**.

1.1.C The product of principal sines, and the product of principal cosines, are denoted by

$$\sin\{L, M\} \ := \ \sin\theta_1 \cdots \sin\theta_l \ , \tag{1.5}$$

$$\cos\{L, M\} \ := \ \cos\theta_1 \cdots \cos\theta_l \ . \tag{1.6}$$

Note that (1.5) and (1.6) are just notation, and not ordinary trigonometrical functions. In particular, $\sin^2\{L, M\} + \cos^2\{L, M\} \ \leq \ 1$.

1.1.D Let $Q_{r,n}$ denote the set of increasing sequences of r elements from $\{1, \ldots, n\}$. For $A \in \mathbb{R}_r^{m \times n}$ we denote by

$$\mathcal{I}(A) \ := \ \{I \in Q_{r,m} : \ \operatorname{rank} A_{I*} = r\} \ , \tag{1.7}$$

$$\mathcal{J}(A) \ := \ \{J \in Q_{r,n} : \ \operatorname{rank} A_{J*} = r\} \ . \tag{1.8}$$

1.1.E The **basic subspaces** of dimension r of $\mathbb{R}^n$ are the $\binom{n}{r}$ subspaces

$$\mathbb{R}_J^n \ := \ \{\mathbf{x} = (x_k) \in \mathbb{R}^n : x_k = 0 \text{ if } k \notin J\} \ , \quad J \in Q_{r,n} \ , \tag{1.9}$$

which, for $r = 1$, reduce to the n coordinate lines

$$\mathbb{R}_{\{j\}}^n \ := \ \{\mathbf{x} = (x_k) \in \mathbb{R}^n : x_k = 0 \text{ if } k \neq j\} \ , \quad j = 1, \ldots, n \ . \tag{1.10}$$

1.2 Results

This paper studies relations between the volume function, principal angles and generalized inverses. In § 2 we study the direction cosines $\cos\{L, \mathbb{R}_J^n\}$ of a subspace L. The basic decomposition of linear operators $\mathbb{R}^n \to \mathbb{R}^m$ is given in § 3. In § 4 we prove related extremal properties of the Moore–Penrose. The Moore–Penrose inverse $A^\dagger$ is of minimal volume among the $\{1, 2\}$-inverses of A.

2 Direction cosines

Let L be a line in $\mathbb{R}^n$ passing through the origin, spanned by the vector $\boldsymbol{\ell} = (\ell_j)$. The **direction cosines** of L are the n cosines

$$\{\cos\{L, \mathbb{R}^n_{\{j\}}\} : \ j = 1, \ldots, n\} \tag{2.1}$$

of the non-obtuse angles between L and the n coordinate axes. The direction cosines are the moduli of the cosines of the angles between $\boldsymbol{\ell}$ and the unit vectors $\{\mathbf{e}_j : \ j = 1, \ldots, n\}$

$$\cos\{L, \mathbb{R}^n_{\{j\}}\} \ = \ |\cos \angle\{\boldsymbol{\ell}, \mathbf{e}_j\}| \ , \tag{2.2}$$

and satisfy

$$\sum_{j=1}^{n} \cos^2\{L, \mathbb{R}^n_{\{j\}}\} \ = \ 1 \ . \tag{2.3}$$

For any line M through the origin, spanned by the vector $\boldsymbol{m} = (m_j)$,

$$\begin{aligned} \cos\{L, M\} \ = \ |\cos \angle\{\boldsymbol{\ell}, \boldsymbol{m}\}| \ &= \ \left| \sum_{j=1}^{n} \cos \angle\{\boldsymbol{\ell}, \mathbf{e}_j\} \cos \angle\{\boldsymbol{m}, \mathbf{e}_j\} \right| \\ &\leq \ \sum_{j=1}^{n} |\cos \angle\{\boldsymbol{\ell}, \mathbf{e}_j\}| \ |\cos \angle\{\boldsymbol{m}, \mathbf{e}_j\}| \\ &= \ \sum_{j=1}^{n} \cos\{L, \mathbb{R}^n_{\{j\}}\} \ \cos\{M, \mathbb{R}^n_{\{j\}}\} \ , \end{aligned} \tag{2.4}$$

with equality in (2.4) if and only if $\cos \angle\{\boldsymbol{\ell}, \mathbf{e}_j\}$ and $\cos \angle\{\boldsymbol{m}, \mathbf{e}_j\}$ have the same signs for all j, or equivalently,

$$\operatorname{sign}(\ell_j) \ = \ \operatorname{sign}(m_j) \ , \ j = 1, \ldots, n \ . \tag{2.5}$$

The analogous results for general subspaces of $\mathbb{R}^n$ are given below. First the analog of the identity (2.3).

Theorem 1 Let L be a subspace of $\mathbb{R}^n$, $\dim L \ = \ \ell > 0$ and let $r \in \{1, \ldots, n\}$. Then

$$\sum_{J \in Q_{r,n}} \cos^2\{L, \mathbb{R}^n_J\} \ = \ \begin{cases} \dfrac{\binom{r}{\ell}\binom{n}{r}}{\binom{n}{\ell}} & , \ \text{if } r \ \geq \ \ell \\[2ex] \dfrac{\binom{n-r}{n-\ell}\binom{n}{r}}{\binom{n}{\ell}} & , \ \text{if } r \ < \ \ell \end{cases} \tag{2.6}$$

<u>Proof</u>: Let $r \ \geq \ \ell$. For any $J \ = \ \{j_1, \cdots, j_r\} \in Q_{r,n}$, let $P \ := \ (\mathbf{e}_{j_1}, \cdots, \mathbf{e}_{j_r}) \in \mathbb{R}^{n \times r}$ denote the matrix with columns $\mathbf{e}_j$, $j \in J$. Let the columns of $Q \in \mathbb{R}^{n \times \ell}$ form an orthonormal basis for L. Then

$$\begin{aligned} \cos^2\{L, \mathbb{R}^n_J\} \ &= \ \prod_{i=1}^{\ell} \sigma_i^2(P^T Q), \quad \text{by [5, Lemma 1]} \\ &= \ \det(Q_{J*}^T Q_{J*}) \\ &= \ \sum_{\substack{K \in Q_{\ell,n} \\ K \subseteq J}} \det(Q_{K*}^T Q_{K*}), \end{aligned}$$

where $\sigma_i(P^TQ)$ are singular values of P^TQ. Therefore

$$\begin{aligned}
\sum_{J\in Q_{r,n}} \cos^2\{L,\mathbb{R}^n_J\} &= \sum_{J\in Q_{r,n}} \sum_{\substack{K\in Q_{\ell,n}\\ K\subseteq J}} \det(Q^T_{K*}Q_{K*}),\\
&= \frac{\binom{r}{\ell}\binom{n}{r}}{\binom{n}{\ell}} \sum_{K\in Q_{\ell,n}} \det(Q^T_{K*}Q_{K*}),\\
&= \frac{\binom{r}{\ell}\binom{n}{r}}{\binom{n}{\ell}} \det(Q^TQ),\\
&= \frac{\binom{r}{\ell}\binom{n}{r}}{\binom{n}{\ell}}, \qquad (2.7)
\end{aligned}$$

where the second equality follows that for each term $\det(Q^T_{K*}Q_{K*}),\ K\in Q_{\ell,n}$, it appears in the summation exactly $\frac{\binom{r}{\ell}\binom{n}{r}}{\binom{n}{\ell}}$ times. The result for $r < \ell$ is obtained from (2.7) using the fact that the nonzero principal angles between L and M are the same as the nonzero principal angles between $L^\perp$ and $M^\perp$, [5, Theorem 3]. Therefore if $r < \ell$,

$$\sum_{J\in Q_{r,n}} \cos^2\{L,\mathbb{R}^n_J\} = \sum_{J\in Q_{n-r,n}} \cos^2\{L^\perp,\mathbb{R}^n_J\} = \frac{\binom{n-r}{n-\ell}\binom{n}{n-r}}{\binom{n}{n-\ell}}.$$

□

We see from (2.6) that

$$\sum_{J\in Q_{r,n}} \cos^2\{L,\mathbb{R}^n_J\} = 1 \qquad (2.8)$$

only if $r = \dim L$ or $r = n$. The special case $r = \dim L = 1$ gives the identity (2.3).
The following theorem gives the analog of inequality (2.4) for equi-dimensional subspaces.

Theorem 2 If L and M are subspaces of $\mathbb{R}^n$ of dimension r, then

$$\cos\{L,M\} \le \sum_{J\in Q_{r,n}} \cos\{L,\mathbb{R}^n_J\}\ \cos\{M,\mathbb{R}^n_J\}. \qquad (2.9)$$

Proof: Let the columns of E and F be orthonormal bases for L and M respectively. Then

$$\begin{aligned}
\cos\{L,M\} &= |\det(E^TF)|, \quad \text{by [5, Theorem 5]},\\
&= |\sum_{J\in Q_{r,n}} \det(E^T_{J*}F_{J*})|,\\
&\le \sum_{J\in Q_{r,n}} |\det(E_{J*})|\,|\det(F_{J*})|,\\
&= \sum_{J\in Q_{r,n}} \cos\{L,\mathbb{R}^n_J\}\ \cos\{M,\mathbb{R}^n_J\}|, \quad \text{by [5, Corollary 2]}. \qquad (2.10)
\end{aligned}$$

□

The proof shows that equality holds in (2.9) if and only if

$$\text{sign}\,\det\,(E_{J*})\ =\ \text{sign}\,\det\,(F_{J*})\ ,\ \forall J\in Q_{r,n}\ , \tag{2.11}$$

or equivalently, corresponding Plücker coordinates of L and M have the same signs.

3 The basic decomposition of linear operators

A linear operator $A:\mathbb{R}^n\to\mathbb{R}^m$ of $\text{rank}\,A=r>0$ can be written as a convex combination

$$A\ =\ \sum_{I\in\mathcal{I}(A)}\ \sum_{J\in\mathcal{J}(A)}\frac{\det^2 A_{IJ}}{\text{vol}^2 A}\,\widehat{(A^\dagger)_{JI}^{-1}} \tag{3.1}$$

where $\widehat{(A^\dagger)_{JI}^{-1}}$ is an $m\times n$ matrix with the inverse of the (J,I)th submatrix of $A^\dagger$ in position (I,J) and zeros elsewhere, see [2, Theorem 6.2]. Each $\widehat{(A^\dagger)_{JI}^{-1}}$ is a one-to-one mapping of $\mathbb{R}^n_J$ onto $\mathbb{R}^n_I$. The operator A of rank r is therefore a convex combination of nonsingular operators between basic subspaces of dimension r. The representation (3.1) is called a **basic decomposition** of A. We interpret the convex weights $\det^2 A_{IJ}/\text{vol}^2 A$ of (3.1) in terms of direction cosines as follows.

Theorem 3 If A is a linear operator $:\mathbb{R}^n\to\mathbb{R}^m$ of $\text{rank}\,A=r>0$, then there exist linear operators $\{B_{IJ}:I\in\mathcal{I}(A),J\in\mathcal{J}(A)\}$ such that

$$\begin{aligned} B_{IJ}:\mathbb{R}^n_J\ &\to\ \mathbb{R}^m_I \text{ is one-to-one and onto}\ ,\\ N(B_{IJ})\ &=\ (\mathbb{R}^n_J)^\perp\ , \end{aligned}$$

and

$$A\ =\ \sum_{I\in\mathcal{I}(A)}\ \sum_{J\in\mathcal{J}(A)}\cos^2\{R(A),\mathbb{R}^m_I\}\,\cos^2\{R(A^T),\mathbb{R}^n_J\}\,B_{IJ}\ . \tag{3.2}$$

Proof: Let

$$A\ =\ CR \tag{3.3}$$

be a **full rank factorization** of A, and apply (3.1) to C and R separately, to get

$$A\ =\ \sum_{I\in\mathcal{I}(C)}\frac{\det^2 C_{I*}}{\text{vol}^2 C}\,\widehat{(C^\dagger)_{*I}^{-1}}\ \ \sum_{J\in\mathcal{J}(R)}\frac{\det^2 R_{J*}}{\text{vol}^2 R}\,\widehat{(R^\dagger)_{J*}^{-1}} \tag{3.4}$$

We recall

$$\cos\{R(A),\ \mathbb{R}^m_I\}\ =\ \frac{|\det A_{IJ}|}{\text{vol}\,A_{*J}}\ , \tag{3.5}$$

for any $J\in\mathcal{J}(A)$, see [5, Corollary 2]. Then (3.2) follows from (3.4) and (3.5) since $\mathcal{I}(A)=\mathcal{I}(C)$, $R(A)=R(C)$, $\mathcal{J}(A)=\mathcal{J}(R)$, $R(A^T)=R(R^T)$ and $A^\dagger=R^\dagger C^\dagger$. □

It follows from [2] that the basic decomposition of $A^\dagger$ has the same convex weights as (3.2),

$$A^\dagger\ =\ \sum_{I\in\mathcal{I}(A)}\ \sum_{J\in\mathcal{J}(A)}\cos^2\{R(A),\mathbb{R}^m_I\}\,\cos^2\{R(A^T),\mathbb{R}^n_J\}\,\widehat{A_{IJ}^{-1}}\ , \tag{3.6}$$

here A_{IJ} is the $(I,J)^{th}$ submatrix of A, and $\widehat{\ }$ denote padding with zeros.

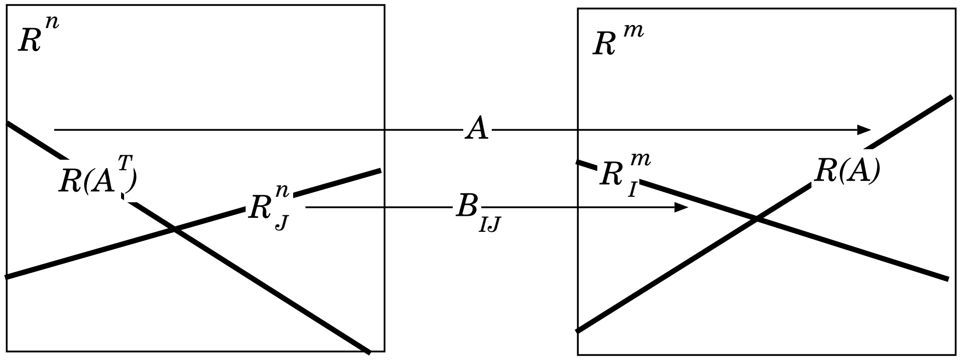

Figure 1: A linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ and one of the basic operators B_{IJ}

The following example shows that the basic decomposition (3.1) of A may not be unique even if we fix the convex weights.

Example 1 Let $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$. The basic decomposition of A is given by (3.2) as

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

with basic operators

$$B_{\{1,2\},\{1,2\}} \;=\; \begin{pmatrix} 5 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix} \;,\; B_{\{1,2\},\{1,3\}} \;=\; \begin{pmatrix} 1 & 0 & 2 \\ -2 & 0 & 2 \end{pmatrix} \;,\; B_{\{1,2\},\{2,3\}} \;=\; \begin{pmatrix} 0 & -\frac{1}{2} & \frac{5}{2} \\ 0 & 1 & 1 \end{pmatrix} .$$

However A can also be expressed as

$$A \;=\; \frac{1}{6}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \frac{2}{3}\begin{pmatrix} 0 & 0 & 3 \\ 0 & 1 & 1 \end{pmatrix}$$

with the same convex weights, but different basic operators. Note that the above two expressions also have the same corresponding minors.

$$\det\begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix} \;=\; \det\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \;=\; 6\,.$$

$$\det\begin{pmatrix} 1 & 2 \\ -2 & 2 \end{pmatrix} \;=\; \det\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \;=\; 6\,.$$

$$\det\begin{pmatrix} -\frac{1}{2} & \frac{5}{2} \\ 1 & 1 \end{pmatrix} \;=\; \det\begin{pmatrix} 0 & 3 \\ 1 & 1 \end{pmatrix} \;=\; -3\,.$$

4 Extremal volumes of $\{1,2\}$-inverses

Let A be a linear operator : $\mathbb{R}^n \to \mathbb{R}^m$ of rank r, range $R(A)$ and null space $N(A)$. The $\{1,2\}$-inverses of A are the operators $G : \mathbb{R}^m \to \mathbb{R}^n$ satisfying

$$AGA \;=\; A \quad \text{and} \quad GAG \;=\; G\,. \tag{4.1}$$

The set of all $\{1,2\}$-inverses of A is denoted by $A\{1,2\}$. For any two subspaces S and T such that

$$\mathbb{R}^n \;=\; N(A) \oplus T\;, \quad \mathbb{R}^m \;=\; R(A) \oplus S \tag{4.2}$$

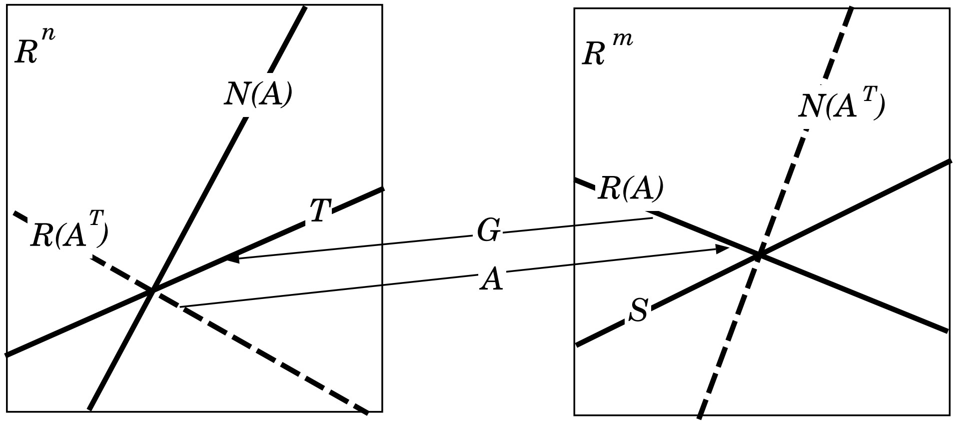

Figure 2: A $\{1,2\}$-inverse G of A with range T and null-space S

there is a unique $\{1,2\}$-inverse G of A, with

$$R(G) \;=\; T \;, \quad N(G) = S \;, \tag{4.3}$$

see Figure 2. In particular, if $S = N(A^T)$ and $T = R(A^T)$ then G is the Moore-Penrose inverse $A^\dagger$. The volume of $A^\dagger$ is

$$\text{vol}\,(A^\dagger) = \frac{1}{\text{vol}\,(A)} \;. \tag{4.4}$$

Theorem 4 Let G be a $\{1,2\}$-inverse of A with range $R(G) = T$ and null space $N(G) = S$. Then

$$\text{vol}\,G \;=\; \frac{\text{vol}\,A^\dagger}{\cos\{T, R(A^T)\}\,\cos\{S, N(A^T)\}} \;. \tag{4.5}$$

Proof: The rank of G is r , since $G \in A\{1,2\}$. Let

$$G \;=\; EPF \;, \qquad E\,,\; F^T \in \mathbb{R}_r^{n\times r} \;, \qquad P \in \mathbb{R}_r^{r\times r} \;, \tag{4.6}$$

be a full rank factorization of G , where E , F^T have orthonormal columns. Then

$$R(G) \;=\; R(E) \;=\; T \;, \quad N(G) \;=\; N(F) \;=\; S \;. \tag{4.7}$$

It follows from (4.1) that

$$PFCRE \;=\; I_r \;. \tag{4.8}$$

Therefore

$$P \;=\; (RE)^{-1}\,(FC)^{-1} \;, \tag{4.9}$$

and

$$\begin{aligned} \text{vol}\,G &= \text{vol}\,(E)\,\text{vol}\,(P)\,\text{vol}\,(F)\;, \\ &= \frac{1}{|\det(RE)|\,|\det(FC)|}\;, \quad \text{by (4.9)} \\ &= \frac{1}{\text{vol}\,(R)\cos\{R(E), R(R^T)\}\,\text{vol}\,(C)\cos\{R(F^T), R(C)\}}\;, \quad \text{by [5, Theorem 5]} \\ &= \frac{\text{vol}\,A^\dagger}{\cos\{T, R(A^T)\}\,\cos\{S^\perp, R(A)\}}\;, \quad \text{by (4.4)} \\ &= \frac{\text{vol}\,A^\dagger}{\cos\{T, R(A^T)\}\,\cos\{S, N(A^T)\}}, \quad \text{by [5, Theorem 3]}\;. \end{aligned} \tag{4.10}$$

□

Corollary 1 The Moore-Penrose inverse $A^\dagger$ is of minimal volume among all $\{1,2\}$-inverses of A .

Proof: If $T \neq R(A^T)$ or $S \neq N(A^T)$ then the denominator in (4.5) is < 1 . □

Remark. $\mathrm{vol}\, G$ is unbounded in $A\{1,2\}$, although

$$\begin{aligned} \cos\{S, N(A^T)\} &= 0 \text{ violates } AGA = A \,, \\ \cos\{T, R(A^T)\} &= 0 \text{ violates } GAG = G \,. \end{aligned}$$

References

[1] S.N. Afriat, "Orthogonal and oblique projectors and the characteristics of pairs of vector spaces", *Proc. Cambridge Phil. Soc.* **53** (1957), 800-816

[2] A. Ben-Israel, "A volume associated with $m \times n$ matrices", *Lin. Algeb. and its Appl.* **167**(1992), 87-111

[3] A. Ben-Israel and T.N.E. Greville, *Generalized Inverses: Theory and Applications* , Wiley-Interscience, 1974

[4] H. Hotelling, "Relations between two sets of variates", *Biometrika* **28**(1935), 321-377

[5] J. Miao and A. Ben-Israel, "On principal angles between subspaces in $\mathbb{R}^n$ ", *Lin. Algeb. and its Appl.* **171**(1992), 81-98