Lin. & Multilin. Algeb. **40**(1995), 153–161

Singular Values and Maximum Rank Minors of Generalized Inverses

R.B. Bapat * Adi Ben-Israel †‡

February 2, 1995
Revised June 15, 1995

Abstract

Singular values and maximum rank minors of generalized inverses are studied. Proportionality of maximum rank minors is explained in terms of space equivalence. The Moore–Penrose inverse $A^\dagger$ is characterized as the $\{1\}$–inverse of A with minimal volume.

Key words: Singular values. Volume. Generalized Inverses. The Moore–Penrose Inverse. Compound Matrices. Space Equivalent Matrices.

*Indian Statistical Institute, New Delhi, 110016, India

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

‡Research done at the Indian Statistical Institute, New Delhi, and supported by a Fulbright Grant and the US Educational Foundation in India.

1 Introduction

Throughout this paper A is an $m \times n$ real matrix of rank r, a fact denoted by $A \in \mathbb{R}_r^{m\times n}$. The singular values of A are denoted $\{\sigma_i(A) : i = 1, \cdots, r\}$. The vector in $\mathbb{R}^{mn}$ obtained by reading the columns of A one by one is denoted $\operatorname{vec} A$.

For $k = 1, \cdots, r$, the k-th compound of A, denoted $C_k(A)$, is the $\binom{m}{k} \times \binom{n}{k}$ matrix whose elements are the $k \times k$ minors of A, i.e. the determinants of its $k \times k$ submatrices ordered lexicographically. The $r \times r$ minors of A (i.e. the elements of $C_r(A)$) are called its **maximum rank minors**.

We denote by $Q_{k,n}$ the set of increasing sequences of k elements from $\{1, 2, \cdots, n\}$. Given index sets $I \subset \{1, \cdots, m\}$ and $J \subset \{1, \cdots, n\}$ we denote by A_{IJ} the corresponding submatrix of A. The submatrix of columns in J is denoted A_{*J}.

Definition 1 For $k = 1, \ldots, r$, the k–**volume** of A is defined as the Frobenius norm of the k-th compound matrix $C_k(A)$,

$$\operatorname{vol}_k A \ := \ \sqrt{\sum_{I\in Q_{k,m}\ ,\ J\in Q_{k,n}} |\det A_{IJ}|^2} \tag{1.1a}$$

or equivalently,

$$\operatorname{vol}_k A \ = \ \sqrt{\sum_{I\in Q_{k,r}} \left(\prod_{i\in I} \sigma_i^2(A)\right)} \tag{1.1b}$$

the square root of the k-th symmetric function[1] of $\{\sigma_1^2(A), \cdots, \sigma_r^2(A)\}$.

We use the convention

$$\operatorname{vol}_k A \ := \ 0 \ , \quad \text{for } k = 0 \text{ or } k > \operatorname{rank} A \ . \tag{1.2}$$

It helps to think of the k–volume of A as the (ordinary) Euclidean norm of $\operatorname{vec} C_k(A)$. In particular, for $k = 1$, the 1–**volume** of $A = (a_{ij})$ is its Frobenius norm

$$\operatorname{vol}_1(A) \ = \ \sqrt{\sum_{i,j} |a_{ij}|^2} \ = \ \sqrt{\operatorname{tr} A^T A} \tag{1.3}$$

and for $r = \operatorname{rank} A$, the r–**volume** of A is

$$\operatorname{vol}_r A \ := \ \sqrt{\sum_{I\in Q_{r,m}\ ,\ J\in Q_{r,n}} |\det A_{IJ}|^2} \tag{1.4a}$$

$$= \ \prod_{i=1}^{r} \sigma_i(A) \ . \tag{1.4b}$$

The r–volume $\operatorname{vol}_r A$ is sometimes called just the **volume** of A, as in [3], and denoted by $\operatorname{vol} A$.

It should be noted that the k–volume of A is not the volume of its k-th compound. Indeed, for $k = 1, \cdots, r = \operatorname{rank} A$, the rank of $C_k(A)$ is $\binom{r}{k}$. Its volume (i.e. its $\binom{r}{k}$–volume) is given in terms of the r–volume of A as

$$\operatorname{vol}_{\binom{r}{k}} C_k(A) \ = \ (\operatorname{vol}_r A)^{\binom{r-1}{k-1}} \ , \quad k = 1, \cdots, r \ . \tag{1.5}$$

[1] The k–volume was defined in [6] as the product of the k largest singular values of A. Definition (1.1) is more natural.

The left side is a product of the singular values of A, each appearing exactly $\binom{r-1}{k-1}$ times, and the result follows from (1.4b).

The study of generalized inverses reveals instances where corresponding maximum rank minors of two matrices A, B are proportional, i.e.

$$\det A_{IJ} \;=\; \alpha \det B_{IJ} \tag{1.6}$$

for some $\alpha \neq 0$. For example, the corresponding maximum rank minors of $A^{\dagger}$ and A^T satisfy

$$\det \left(A^{\dagger}\right)_{IJ} \;=\; \frac{1}{\mathrm{vol}^2 A} \det \left(A^T\right)_{IJ} \tag{1.7}$$

see [3, Lemma 3.2]. Proportionality of maximum rank minors is an essential feature in the study of generalized inverses for matrices over integral domains, see [1]. We explain this proportionality in § 2, through the conecpt of state equivalence. Singular values of generalized inverses are studied in § 3. The Moore–Penrose inverse is characterized as the $\{1\}$–inverse of minimal volume in § 4.

In § 2 we have occasion to use Plücker coordinates, a concept from multilinear algebra, see e.g. [8], [9]. The **Plücker coordinates** of an r–dimensional subspace $L \subset \mathbb{R}^n$ are the components of the exterior product $\mathbf{x}_1 \wedge \mathbf{x}_2 \wedge \cdots \wedge \mathbf{x}_r$ where $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_r\}$ is <u>any</u> basis of L . The Plücker coordinates of L are determined up to a scalar multiple, i.e. they span a line in $\mathbb{R}^{\binom{n}{r}}$. Thus there is a one–to–one correspondence between r–dimensional subspaces L in $\mathbb{R}^n$ and 1–dimensional subspaces in $\mathbb{R}^{\binom{n}{r}}$, see e.g. [10, Theorem 4.9].

For example, given $A \in \mathbb{R}_r^{m\times n}$, the Plücker coordinates of $R(A)$, the **range** of A, are the components of $\mathrm{vec}\, C_r(A)$, i.e. the maximum rank minors of A.

2 Space equivalent matrices

The following definition describes matrices representing linear transformations between the same subspaces.

Definition 2 Two $m \times n$ matrices A , B are called **space equivalent** if

$$R(A) \;=\; R(B)\;, \tag{2.1a}$$
$$\text{and} \quad R(A^T) \;=\; R(B^T)\;. \tag{2.1b}$$

Let L, M be subspaces of $\mathbb{R}^n$, with dimensions ℓ, m respectively, and let $\ell \leq m$. We denote by $\cos\{L, M\}$ the product of the cosines of the ℓ principal angles between L and M, see e.g. [6]. In particular, $\cos\{L, M\} = 1$ if and only if $L \subset M$. The following version of the Cauchy-Schwarz inequality was proved in [6, Theorem 5], for full column–rank matrices A, $B \in \mathbb{R}_r^{m\times r}$,

$$\mathrm{vol}\left(B^T A\right) \;=\; \mathrm{vol}\, A\, \mathrm{vol}\, B\, \cos\{R(A), R(B)\} \tag{2.2}$$

We extend this result to matrices of arbitrary rank in Theorem 1 below. First we need

Lemma 1 Let $S \in \mathbb{R}^{m\times m}$, $A \in \mathbb{R}_m^{m\times n}$. Then

$$\mathrm{vol}_m\,(SA) \;=\; |\det S|\, \mathrm{vol}\, A\;. \tag{2.3}$$

Proof: If S is singular, then both sides of (2.3) are zero. Let S be nonsingular. Then $\operatorname{rank}(SA) = m$, and

$$\begin{aligned} \operatorname{vol}_m(SA) = \operatorname{vol}(SA) &= \sqrt{\sum_{J \in Q_{m,n}} \det^2 (SA)_{*J}} \\ &= \sqrt{\sum_{J \in Q_{m,n}} \det^2 S \det^2 A_{*J}} \\ &= |\det S| \operatorname{vol} A . \end{aligned}$$

□

Theorem 1 Let $A, B \in \mathbb{R}_r^{m \times n}$. Then

$$\operatorname{vol}_r \left(B^T A\right) = \operatorname{vol}_r A \operatorname{vol}_r B \cos\{R(A), R(B)\} \tag{2.4a}$$

$$\operatorname{vol}_r \left(A B^T\right) = \operatorname{vol}_r A \operatorname{vol}_r B \cos\{R(A^T), R(B^T)\} . \tag{2.4b}$$

Proof of (2.4a): If $\operatorname{rank} B^T A < r$ then there is an $x \in \mathbb{R}^n$ such that $Ax \neq 0$ and $B^T Ax = 0$. Therefore one of the principal angles between $R(A)$ and $R(B)$ is $\frac{\pi}{2}$, and (2.4a) gives $0 = 0$.

Assume $\operatorname{rank} B^T A = r$, and let all volumes below be r–volumes. Let $A = C_A R_A$ and $B = C_B R_B$ be full rank factorizations of A and B. Then

$$\begin{aligned} B^T A &= (C_B R_B)^T (C_A R_A) \\ &= R_B^T \left(C_B^T C_A R_A\right) \end{aligned}$$

is a full rank factorization if $\operatorname{rank} B^T A = r$. Its volume is

$$\begin{aligned} \operatorname{vol}\left(B^T A\right) &= \operatorname{vol} R_B \operatorname{vol}\left(C_B^T C_A R_A\right) , \quad \text{by [3, Lemma 2.2]} , \\ &= \operatorname{vol} R_B \left|\det\left(C_B^T C_A\right)\right| \operatorname{vol} R_A , \quad \text{by Lemma 1} \\ &= \operatorname{vol} R_B \operatorname{vol} R_A \left(\operatorname{vol} C_B \operatorname{vol} C_A \cos\{R(C_A), R(C_B)\}\right) , \quad \text{by [6, Theorem 5]} \\ &= (\operatorname{vol} C_A \operatorname{vol} R_A)(\operatorname{vol} C_B \operatorname{vol} R_B) \cos\{R(A), R(B)\} , \quad \text{since } R(C_A) = R(A),\ R(C_B) = R(B) \\ &= \operatorname{vol} A \operatorname{vol} B \cos\{R(A), R(B)\} . \end{aligned}$$

The proof of (2.4b) is similar. □

Example 1 If P is idempotent then its eigenvalues are 1, 0 and its nonzero singular values are all ≥ 1. Thus $\operatorname{vol} P \geq 1$. More precisely,

$$\operatorname{vol} P = \frac{1}{\cos\{R(P), R(I-P)^{\perp}\}} ,$$

where $R(P)$ is the range of P, and $R(I-P)$ is its null–space. This follows from (2.4a) with $A = P$, $B = P^T$ so that $B^T A = P^2 = P$.

Therefore $\operatorname{vol} P = 1$ if and only if $P = P^T$, i.e. P is an orthogonal projector. □

The vectors $\mathrm{vec}\, C_r(A)$ and $\mathrm{vec}\, C_r(B)$ give the Plücker coordinates of the subspaces $R(A)$ and $R(B)$ respectively. The (ordinary) angle between these vectors, in the space $\mathbb{R}^{\binom{m}{r}\binom{n}{r}}$, has cosine equal to $\cos\{R(A), R(B)\}$. Statements (2.4a) and (2.4b) are Cauchy–Schwarz inequalities for the vectors $\mathrm{vec}\, C_r(A)$ and $\mathrm{vec}\, C_r(B)$. As expected, equality holds if their components (i.e. the maximum rank minors of A, B) are proportional, see (2.6) below.

Theorem 2 Let $A,\, B \in \mathbb{R}_r^{m\times n}$. Then the following are equivalent:
(a) A and B are space equivalent.
(b) There are matrices $X,\, Y \in \mathbb{R}^{n\times m}$ such that

$$\begin{aligned} A &= BXB && (2.5a) \\ B &= AYA && (2.5b) \end{aligned}$$

(c) $\mathrm{vol}_r\left(B^T A\right) = \mathrm{vol}_r\left(A\, B^T\right) = \mathrm{vol} A\, \mathrm{vol} B$.
(d) The r–compounds of A, B satisfy

$$C_r(A) = \alpha\, C_r(B)\ , \quad \text{for some } \alpha \neq 0\ . \tag{2.6}$$

Proof: (b) $\Longrightarrow$ (a) is obvious. To prove (a) $\Longrightarrow$ (b), we use $R(A) = R(B) \Longrightarrow A = BB^\dagger A$ and $R(A^T) = R(B^T) \Longrightarrow A = AB^\dagger B$ to show that $A = BB^\dagger A = BB^\dagger AB^\dagger B$, proving (2.5a) for $X = B^\dagger AB^\dagger$. (2.5b) is similarly proved.
(a) $\Longrightarrow$ (c) from (2.4a) and (2.4b), and (c) $\Longrightarrow$ (d) by the Cauchy–Schwarz inequality for $\mathrm{vec}\, C_r(A)$ and $\mathrm{vec}\, C_r(B)$. To prove (d) $\Longrightarrow$ (a) we note that the matrix $C_r(A)$ is of rank 1, and of the form xy^T where x and y are the Plücker coordinates of the subspaces $R(A)$ and $R(A^T)$, respectively. From (d) it follows that $C_r(B) = \alpha\, xy^T$, proving that $R(A)$ and $R(B)$ have the same Plücker coordinates and therefore $R(A) = R(B)$. Similarly $R(A^T) = R(B^T)$. □

Example 2 The matrices $A^\dagger$ and A^T are space equivalent. Therefore

$$\det\left(A^\dagger\right)_{IJ} = \alpha \det\left(A^T\right)_{IJ}$$

for all indices IJ of $r \times r$ submatrices. Adding the squares of these expressions we get

$$\mathrm{vol}^2 A^\dagger = \alpha^2\, \mathrm{vol}^2 A^T$$

and

$$\alpha = \frac{1}{\mathrm{vol}^2 A}\ , \quad \text{since } \mathrm{vol}\, A^T = \mathrm{vol}\, A \text{ and } \mathrm{vol}\, A^\dagger = \frac{1}{\mathrm{vol}\, A}\ ,$$

proving (1.7).

3 Singular values of generalized inverses

Let $A \in \mathbb{R}_r^{m\times n}$ have the singular value decomposition (SVD)

$$A = U \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} V^T \tag{3.1}$$

where U, V are orthogonal, and Σ is a diagonal matrix, with the singular values of A

$$\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_r(A) \ . \tag{3.2}$$

The general $\{1\}$–inverse of A is

$$G \ = \ V \begin{pmatrix} \Sigma^{-1} & X \\ Y & Z \end{pmatrix} U^T \tag{3.3}$$

where X, Y, Z are arbitrary submatrices of appropriate sizes. In particular,

$Z \ = \ Y \Sigma X$ gives the general $\{1, 2\}$ inverses, i.e. the solutions of $AXA = A$, $XAX = X$,

$X = O$ gives the general $\{1, 3\}$–inverses (the solutions of $AXA = A$, $(AX)^T = AX$),

$Y = O$ gives the general $\{1, 4\}$–inverses (the solutions of $AXA = A$, $(XA)^T = XA$),

finally, the Moore–Penrose inverse is (3.3) with $X = O$, $Y = O$ and $Z = O$.

We show next that each singular value of the Moore–Penrose inverse $A^\dagger$ is dominated by a corresponding singular value of any $\{1\}$–inverse of A.

Theorem 3 Let G be a $\{1\}$–inverse of A with singular values

$$\sigma_1(G) \geq \sigma_2(G) \geq \cdots \geq \sigma_s(G) \tag{3.4}$$

where $s \ = \ \operatorname{rank} G$ ($\geq \ \operatorname{rank} A$). Then

$$\sigma_i(G) \ \geq \ \sigma_i(A^\dagger) \ , \quad i = 1, \ldots, r \ . \tag{3.5}$$

Proof: Dropping U, V we write

$$\begin{aligned} G\,G^T \ &= \ \begin{pmatrix} \Sigma^{-1} & X \\ Y & Z \end{pmatrix} \begin{pmatrix} \Sigma^{-1} & Y^T \\ X^T & Z^T \end{pmatrix} \\ &= \ \begin{pmatrix} \Sigma^{-2} + XX^T & ? \\ ? & ? \end{pmatrix} , \end{aligned}$$

where ? denotes a submatrix not needed in this proof. Then for $i = 1, \ldots, r$,

$$\begin{aligned} \sigma_i^2(G) \ &:= \ \lambda_i \left(G\,G^T \right) \\ &\geq \ \lambda_i \left(\Sigma^{-2} + XX^T \right) \ , \quad \text{(e.g. [5, Chapter 11, Theorem 11])} \\ &\geq \ \lambda_i \left(\Sigma^{-2} \right) \ , \quad \text{(e.g. [5, Chapter 11, Theorem 9])} \\ &= \ \sigma_i^2 \left(A^\dagger \right) \ , \end{aligned}$$

proving the theorem. □

Corollary 1 If G is a $\{2\}$–inverse of A of rank q ($\leq \ \operatorname{rank} A$), then

$$\sigma_i(A) \ \geq \ \sigma_i(G^\dagger) \ , \quad i = 1, \ldots, q \ . \tag{3.6}$$

Proof: The statement that G is a $\{2\}$–inverse of A is equivalent to the statement that A is a $\{1\}$–inverse of G. Then (3.6) follows from (3.5) by reversing the roles of A and G. □

Note: For a $\{1, 2\}$–inverse the inequalities (3.6) are equivalent to (3.5), and give no further information.

If G is a $\{1, 3\}$–inverse of A, the inequalities (3.5) can be reversed in the following sense.

Theorem 4 Let $A \in \mathbb{R}_r^{m\times n}$ and let G be a $\{1,3\}$–inverse of A, with singular values

$$\sigma_1(G) \geq \sigma_2(G) \geq \cdots \geq \sigma_s(A) \ , \quad \text{where } s = \min\{m,n\} \ .$$

Then

$$\sigma_i(G) \ \geq \ \sigma_i(A^\dagger) \ \geq \ \sigma_{n-r+i}(G) \ , \quad i = 1,\cdots,r \ . \tag{3.7}$$

In particular, if $m = n$ and $r = n-1$, then

$$\sigma_i(G) \ \geq \ \sigma_i(A^\dagger) \ \geq \ \sigma_{i+1}(G) \ , \quad i = 1,\cdots,r \ . \tag{3.8}$$

Proof: With $X = O$ in (3.3), the matrix GG^T becomes

$$GG^T \ = \ \begin{pmatrix} \Sigma^{-2} & ? \\ ? & ? \end{pmatrix}$$

and the results follow from Poincare's Separation Theorem, see [5, Chapter 11, Theorem 12]. □

4 Minimal volume characterization of the Moore–Penrose inverse

It was shown in [7] that the Moore–Penrose inverse $A^\dagger$ is of minimal r–volume among all $\{1,2\}$–inverses of A, and it is the unique minimizer, i.e. this property characterizes the Moore–Penrose inverse. The Moore–Penrose inverse was also shown in [4] to be the unique minimizer among all $\{1,3\}$–inverses of a class of functions including the unitarily invariant matrix norms.

From Theorem 3 we conclude that for each $k = 1,\cdots,r$, the Moore–Penrose inverse $A^\dagger$ is of minimal k–volume among all $\{1\}$–inverses G of A,

$$\text{vol}_k G \ \geq \ \text{vol}_k A^\dagger \ , \quad k = 1,\ldots,r \ . \tag{4.1}$$

Moreover, this property is a characterization of $A^\dagger$, as indicated in the following results.

Theorem 5 Let $A \in \mathbb{R}_r^{m\times n}$, and let k be any integer in $\{1,\cdots,r\}$. Then the Moore–Penrose inverse $A^\dagger$ is the unique $\{1\}$–inverse of A with minimal k–volume.

Proof: We prove this result directly, by solving the k–volume minimization problem, showing it to have the Moore–Penrose inverse as the unique solution.

The easiest case is $k = 1$. The claim is that $A^\dagger$ is the unique solution $X = (x_{ij})$ of the minimization problem

(P.1) $$\text{minimize } \tfrac{1}{2}\,\text{vol}_1^2 X \quad \text{such that } AXA \ = \ A \ ,$$

where by (1.3)

$$\text{vol}_1^2\,(x_{ij}) \ = \ \sum_{ij} |x_{ij}|^2 \ = \ \text{tr}\, X^T X \ .$$

We use the Lagrangian function

$$L(X,\Lambda) \ := \ \frac{1}{2}\,\text{tr}\, X^T X \ - \ \text{tr}\,\Lambda^T(AXA - A) \tag{4.2}$$

where $\Lambda = (\lambda_{ij})$ is a matrix Lagrange multiplier. The Lagrangian can be written, using the "vec" notation, as

$$L(X,\Lambda) \ = \ \frac{1}{2}\,(\text{vec}\, X)^T(\text{vec}\, X) \ - \ (\text{vec}\,\Lambda)^T \left(A^T \otimes A\right) \text{vec}\, X$$

and its derivative with respect to $\mathrm{vec}\, X$ is

$$(\nabla_X L(X,\Lambda))^T \;=\; (\mathrm{vec}\, X)^T - (\mathrm{vec}\,\Lambda)^T \left(A^T \otimes A\right)$$

see e.g. [5]. The necessary condition for optimality is that the derivative vanishes,

$$\begin{aligned} (\mathrm{vec}\, X)^T - (\mathrm{vec}\,\Lambda)^T \left(A^T \otimes A\right) &= \mathrm{vec}\, O \\ \text{or equivalently,} \quad X &= A^T \Lambda A^T \; . \end{aligned} \tag{4.3}$$

This condition is also sufficient, since (P.1) is a problem of minimizing a convex function subject to linear constraints. Indeed, the Moore–Penrose inverse $A^\dagger$ is the unique $\{1\}$–inverse of A satisfying (4.3) for some Λ (see e.g. [2]). Therefore $A^\dagger$ is the unique solution of (P.1).

An alternative (simpler) way to show this is by writing (3.3) as

$$G \;=\; U \begin{pmatrix} \Sigma^{-1} & X \\ Y & Z \end{pmatrix} V^T \;=\; U \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} V^T \;+\; U \begin{pmatrix} O & X \\ Y & Z \end{pmatrix} V^T \;=\; A^\dagger + \left(G - A^\dagger\right) \; . \tag{4.4}$$

We conclude that

$$\mathrm{vol}_1^2\, G \;=\; \mathrm{vol}_1^2\, A^\dagger \;+\; \mathrm{vol}_1^2\, (G - A^\dagger) \;, \quad \text{whenever } AGA = A \tag{4.5}$$

proving that $A^\dagger$ is the unique minimal norm $\{1\}$–inverse of A.

For any $1 \le k \le r$ the problem analogous to (P.1) is

(P.k) $\qquad$ minimize $\frac{1}{2}\,\mathrm{vol}_k^2 X$ such that $AXA \;=\; A$.

We note that $AXA = A$ implies

$$C_k(A) C_k(X) C_k(A) \;=\; C_k(A) \; . \tag{4.6}$$

Taking (4.6) as the constraint in (P.k), we get the Lagrangian

$$L(X,\Lambda) \;:=\; \frac{1}{2} \sum_{I \in Q_{k,n}\;,\; J \in Q_{k,m}} |\det X_{IJ}|^2 \;-\; \mathrm{tr}\, C_k(\Lambda)^T (C_k(A) C_k(X) C_k(A) - C_k(A)) \; .$$

It follows, in analogy with the case $k = 1$, that a necessary and sufficient condition for optimality of X is

$$C_k(X) \;=\; C_k(A^T) C_k(\Lambda) C_k(A^T) \; . \tag{4.7}$$

Moreover, $A^\dagger$ is the unique $\{1\}$–inverse satisfying (4.7), and is therefore the unique solution of (P.k). $\square$

Note: The rank s of a $\{1\}$–inverse G may be greater than r, in which case the volumes

$$\mathrm{vol}_{r+1}(G), \mathrm{vol}_{r+2}(G), \cdots, \mathrm{vol}_s(G)$$

are positive. However, the corresponding volumes of $A^\dagger$ are zero, by Definition (1.2), so the inequalities (4.1) still hold.

The optimality characterization (4.1) has an interesting geometric interpretation. Consider first the case $k = 1$. Simplifying the identity (4.5) we get an equivalent condition

$$\mathrm{tr}\, (A^\dagger)^T (G - A^\dagger) \;=\; 0 \;, \quad \text{whenever } AGA = A \;, \tag{4.8}$$

i.e. $A^\dagger$ is orthogonal to all matrices $G - A^\dagger$, where G ranges over $\{1\}$–inverses of A, and the inner product $< X, Y > := \mathrm{tr}\, X^T Y$ is used. This makes sense since:

the set $A\{1\} \;=\; \{X : AXA = A\}$ of $\{1\}$–inverses of A is an affine set in $\mathbb{R}^{n \times m}$, the set $A\{1\} - A^\dagger \;=\; \{X : AXA = O\}$ is a subspace in $\mathbb{R}^{n\times m}$, and

$A^\dagger$ is the minimal norm element of $A\{1\}$,

therefore $A^\dagger$ is orthogonal to the subspace $A\{1\} - A^\dagger$.

For $k \geq 1$, the result analogous to (4.5) is

$$\begin{aligned} \mathrm{vol}_k^2\, G \;&:=\; \mathrm{vol}_1^2\, C_k(G) \\ &=\; \mathrm{vol}_1^2 C_k(A^\dagger) \;+\; \mathrm{vol}_1^2\,(G - A^\dagger)\;, \quad \text{from (4.4)} \\ &=\; \mathrm{vol}_k^2\, A^\dagger \;+\; \mathrm{vol}_1^2\,(G - A^\dagger) \end{aligned} \tag{4.9}$$

and the equivalent orthogonality condition (analogous to (4.8)) is

$$(\mathrm{vec}\, C_k(A^\dagger))^T \left(\mathrm{vec}\, C_k(G) - \mathrm{vec}\, C_k(A^\dagger)\right) \;=\; 0\;, \tag{4.10}$$

for all $k = 1, \ldots, r$ and $\{1\}$–inverses G of A. The geometric interpretation is again that the set $C_k(A)\{1\}$ of $\{1\}$–inverses of $C_k(A)$ is an affine set in $\mathbb{R}^{\binom{n}{k}\times\binom{m}{k}}$, and the vector $\mathrm{vec}\, C_k(A^\dagger)$ is orthogonal to the subspace $C_k(A)\{1\} - C_k(A^\dagger)$.

References

[1] R.B. Bapat, "Generalized inverses with proportional minors", *Lin. Algeb. and its Appl.* **211**(1994), 27–33

[2] A. Ben-Israel, "On decompositions of matrix spaces with applications to matrix equations", *Atti Accad Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* (8) **45** (1968), 54-60

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

[4] D.R. Jensen, "Minimal properties of Moore–Penrose inverses", *Lin. Algeb. and its Appl.* **196**(1994), 175–182

[5] J.R. Magnus and H. Neudecker, *Matrix Differential Calculus*, J. Wiley, 1988

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

[7] J. Miao and A. Ben-Israel, "Product cosines of angles between subspaces", *Lin. Algeb. and its Appl.* (to appear)

[8] G.D. Mostow and J.H. Sampson, *Linear Algebra*, McGraw–Hill, New York, 1969

[9] G.D. Mostow, J.H. Sampson and J.-P. Meyer, *Fundamental Structures of Algebra*, McGraw–Hill, New York, 1963

[10] G.C. Shephard, *Vector Spaces of Finite Dimension*, University Mathematical Texts, Oliver & Boyd, Edinburgh, 1966