Benítez López, J.; Liu, X. (2013). A short proof of a matrix decomposition with applications. Linear Algebra and its Applications. 438(3):1398-1414. doi:10.1016/j.laa.2012.10.002.

A short proof of a matrix decomposition with applications

Julio Benítez[*] Xiaoji Liu[†]

We give a very short proof of the main result of J. Benítez, *A new decomposition for square matrices*, Electron. J. Linear Algebra, 20, 207-225 (2010). Also, we present some consequences of this result.
Keywords: CS decomposition, canonical angles, generalized inverses
AMS-Subject Classification: 15A09, 15A23.

1 A short proof of a known decomposition

Let $\mathbb{C}_{m,n}$ be the set of $m \times n$ complex matrices. Let A^*, $\mathcal{R}(A)$, $\mathcal{N}(A)$, and $\mathrm{rank}(A)$ denote the conjugate transpose, column space, null space, and rank, respectively, of $A \in \mathbb{C}_{m,n}$. Furthermore, let $A^\dagger$ stand for the Moore-Penrose inverse of A, i.e., the unique matrix satisfying the equations

$$AA^\dagger A = A, \qquad A^\dagger AA^\dagger = A^\dagger, \qquad AA^\dagger = (AA^\dagger)^*, \qquad A^\dagger A = (A^\dagger A)^*.$$

We shall denote the zero matrix in $\mathbb{C}_{n,m}$ by $0_{n,m}$, and when there is no danger of confusion, we will simply write 0. In addition, $\mathbf{1}_n$ and $\mathbf{0}_n$ will denote the $n \times 1$ column vectors all of whose components are 1 and 0, respectively. If $\mathcal{S}$ is a subspace of $\mathbb{C}^n$, then $P_\mathcal{S}$ stands for the orthogonal projector onto the subspace $\mathcal{S}$.

We shall use the concept of canonical angles (also called principal angles) which will be defined in the next paragraph [13]:

Definition 1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two nontrivial subspaces of $\mathbb{C}^n$ and $r = \min\{\dim \mathcal{X}, \dim \mathcal{Y}\}$. We define the* canonical angles $\theta_1, \ldots, \theta_r \in [0, \pi/2]$ *between $\mathcal{X}$ and $\mathcal{Y}$ by*

$$\cos\theta_i = \sigma_i(P_\mathcal{X} P_\mathcal{Y}), \qquad i = 1, \ldots, r, \tag{1}$$

where the nonnegative real numbers $\sigma_1(P_\mathcal{X} P_\mathcal{Y}), \ldots, \sigma_r(P_\mathcal{X} P_\mathcal{Y})$ are the singular values of $P_\mathcal{X} P_\mathcal{Y}$. We will have in mind the possibility that one canonical angle is repeated.

[*]Departamento de Matemática Aplicada, Instituto de Matemática Multidisciplinar, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain (jbenitez@mat.upv.es).
[†]College of Mathematics and Computer Science, Guangxi University for Nationalities, Nanning 530006, China (xiaojiliu72@yahoo.com.cn).

In [2] it was given the following theorem:

Theorem 1. *Let $A \in \mathbb{C}_{n,n}$, $r = \operatorname{rank}(A)$, and let $\theta_1, \ldots, \theta_p$ be the canonical angles between $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$ belonging to $]0, \pi/2[$. Denote by x and y the multiplicities of the angles 0 and $\pi/2$ as a canonical angle between $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$, respectively. There exists a unitary matrix $V \in \mathbb{C}_{n,n}$ such that*

$$A = V \begin{bmatrix} MC & MS \\ 0 & 0 \end{bmatrix} V^*, \tag{2}$$

where $M \in \mathbb{C}_{r,r}$ is nonsingular,

$$C = \operatorname{diag}(\mathbf{0}_y, \cos\theta_1, \ldots, \cos\theta_p, \mathbf{1}_x), \tag{3}$$

$$S = \begin{bmatrix} \operatorname{diag}(\mathbf{1}_y, \sin\theta_1, \ldots, \sin\theta_p) & 0_{p+y,n-(r+p+y)} \\ 0_{x,p+y} & 0_{x,n-(r+p+y)} \end{bmatrix},$$

and $r = y + p + x$. Furthermore, x and $y + n - r$ are the multiplicities of the singular values 1 and 0 in $P_{\mathcal{R}(A)} P_{\mathcal{R}(A^*)}$, respectively.

The usefulness of this result was proved in [2] by studying several important classes of matrices, partial orderings in $\mathbb{C}_{n,n}$, the dimensions of $\mathcal{R}(A) \cap \mathcal{R}(A^*)$ and $\mathcal{R}(A) \cap \mathcal{R}(A^*)^{\perp}$, and the norm of $AA^\dagger - A^\dagger A$. We shall use the CS decomposition which is now established (see e.g. [3, 10, 11] and for a survey of this decomposition, [12]).

Lemma 1 (CS decomposition)**.** *Let $P_1, P_2 \in \mathbb{C}_{n,n}$ be two orthogonal projectors. Then there exists a unitary matrix $U \in \mathbb{C}_{n,n}$ such that*

$$P_1 = U \begin{bmatrix} I & & & & & \\ & 0 & & & & \\ & & I & & & \\ & & & I & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix} U^*, \qquad P_2 = U \begin{bmatrix} \widehat{C}^2 & \widehat{C}\widehat{S} & & & & \\ \widehat{C}\widehat{S} & \widehat{S}^2 & & & & \\ & & I & & & \\ & & & 0 & & \\ & & & & I & \\ & & & & & 0 \end{bmatrix} U^*,$$

where $\widehat{C}, \widehat{S}$ are positive diagonal real matrices such that $\widehat{C}^2 + \widehat{S}^2 = I$, the symbol I denotes identity matrices of various sizes, and the corresponding blocks in the two projection matrices are of the same size.

Proof. (*of Theorem 1*). Since $AA^\dagger$ and $A^\dagger A$ are orthogonal projectors, by Lemma 1, there exist a unitary matrix $U \in \mathbb{C}_{n,n}$ and $p, x, y, z \in \{0\} \cup \mathbb{N}$ such that

$$AA^\dagger = U(T_1 \oplus R_1)U^*, \qquad A^\dagger A = U(T_2 \oplus R_2)U^*, \tag{4}$$

where

$$T_1 = \begin{bmatrix} I_p & \\ & 0 \end{bmatrix} \in \mathbb{C}_{2p,2p}, \quad T_2 = \begin{bmatrix} \widehat{C}^2 & \widehat{C}\widehat{S} \\ \widehat{C}\widehat{S} & \widehat{S}^2 \end{bmatrix} \in \mathbb{C}_{2p,2p}, \tag{5}$$

$$R_1 = I_x \oplus I_y \oplus 0 \oplus 0 \in \mathbb{C}_{n-2p,n-2p}, \quad R_2 = I_x \oplus 0 \oplus I_z \oplus 0 \in \mathbb{C}_{n-2p,n-2p}, \tag{6}$$

and in addition $\widehat{C}, \widehat{S} \in \mathbb{C}_{p,p}$ have the same meaning as in Lemma 1. Let us denote $t = (n-2p)-(x+y+z)$ in order to the last summands in (6) have order t. If $p=0$, then blocks T_1 and T_2 do not appear in (4). Moreover, some blocks in the representation of R_1 and R_2 in (6) can also be absent.

From representations (4), (5), and (6) we get $\operatorname{rank}(AA^\dagger) = p+x+y$ and $\operatorname{rank}(A^\dagger A) = p+x+z$, because $\operatorname{rank}(T_1) = \operatorname{rank}(T_2) = p$. Since $AA^\dagger$ and $A^\dagger A$ are the orthogonal projectors onto $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$, respectively, we have $\operatorname{rank}(AA^\dagger) = \operatorname{rank}(A)$ and $\operatorname{rank}(A^\dagger A) = \operatorname{rank}(A^*)$. Since $\operatorname{rank}(A) = \operatorname{rank}(A^*)$, we deduce $y = z$. Since $r = \operatorname{rank}(A)$ we have $r = p + x + y$.

By a suitable permutation matrix, there is a unitary matrix $V \in \mathbb{C}_{n,n}$ such that

$$AA^\dagger = V \begin{bmatrix} I_y & 0 & 0 & 0 & 0 & 0 \\ 0 & I_p & 0 & 0 & 0 & 0 \\ 0 & 0 & I_x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0_{y,y} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0_{p,p} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0_{t,t} \end{bmatrix} V^*$$

and

$$A^\dagger A = V \begin{bmatrix} 0_{y,y} & 0 & 0 & 0 & 0 & 0 \\ 0 & \widehat{C}^2 & 0 & 0 & \widehat{C}\widehat{S} & 0 \\ 0 & 0 & I_x & 0 & 0 & 0 \\ 0 & 0 & 0 & I_y & 0 & 0 \\ 0 & \widehat{C}\widehat{S} & 0 & 0 & \widehat{S}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0_{t,t} \end{bmatrix} V^*.$$

If we define

$$C = \begin{bmatrix} 0_{y,y} & 0 & 0 \\ 0 & \widehat{C} & 0 \\ 0 & 0 & I_x \end{bmatrix} \in \mathbb{C}_{r,r} \qquad \text{and} \qquad S = \begin{bmatrix} I_y & 0 & 0 \\ 0 & \widehat{S} & 0 \\ 0 & 0 & 0_{x,t} \end{bmatrix} \in \mathbb{C}_{r,n-r},$$

we have

$$AA^\dagger = V \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r,n-r} \end{bmatrix} V^*, \qquad A^\dagger A = V \begin{bmatrix} C^2 & CS \\ S^*C & S^*S \end{bmatrix} V^* \tag{7}$$

and

$$C^2 + SS^* = I_r, \qquad (\widehat{C} \oplus I_{n-r-p})^2 + S^*S = I_{n-r}, \qquad CS = S(\widehat{C} \oplus I_{n-r-p}). \tag{8}$$

Now, let us represent

$$A = V \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} V^*, \qquad X_1 \in \mathbb{C}_{r,r},\ X_4 \in \mathbb{C}_{n-r,n-r}.$$

From $A = (AA^\dagger)A$ and the first identity of (7) we have

$$\begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix}.$$

Hence $X_3 = 0$ and $X_4 = 0$. From $A = A(A^\dagger A)$ and the second identity of (7) we obtain

$$\begin{bmatrix} X_1 & X_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} X_1 & X_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} C^2 & CS \\ S^*C & S^*S \end{bmatrix}.$$

Therefore $X_1 = X_1C^2 + X_2S^*C$ and $X_2 = X_1CS + X_2S^*S$. If we define $M = X_1C + X_2S^*$, we get $X_1 = MC$ and $X_2 = MS$. Now, we shall prove the nonsingularity of M. From the first identity of (8) we easily get

$$AA^* = V \begin{bmatrix} MM^* & 0 \\ 0 & 0 \end{bmatrix} V^*,$$

and thus, $r = \mathrm{rank}(A) = \mathrm{rank}(AA^*) = \mathrm{rank}(MM^*) = \mathrm{rank}(M)$. Since $M \in \mathbb{C}_{r,r}$ and $r = \mathrm{rank}(M)$ we get that M is nonsingular.

Since $\widehat{C}^2 + \widehat{S}^2 = I_p$ and $\widehat{C}, \widehat{S}$ are positive, real, and diagonal matrices, there exist $\theta_1, \ldots, \theta_p \in]0, \pi/2[$ such that $\widehat{C} = \mathrm{diag}(\cos\theta_1, \ldots, \cos\theta_p)$ and $\widehat{S} = \mathrm{diag}(\sin\theta_1, \ldots, \sin\theta_p)$. It remains to prove that $\theta_1, \ldots, \theta_p$ are the canonical angles between $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$ belonging to $]0, \pi/2[$, and x and y are the multiplicities of the singular values 0 and 1 in $P_{\mathcal{R}(A)}P_{\mathcal{R}(A^*)}$, respectively. To this end, we will use (1). From (7), we obtain

$$P_{\mathcal{R}(A)}P_{\mathcal{R}(A^*)} = V \begin{bmatrix} C^2 & CS \\ 0 & 0 \end{bmatrix} V^*.$$

Next, we are going to find the sigular value decomposition of $P_{\mathcal{R}(A)}P_{\mathcal{R}(A^*)}$. Let us remark that from (8) we get that the matrix

$$T = \begin{bmatrix} C & S \\ -S^* & \widehat{C} \oplus I_{n-r-p} \end{bmatrix} \in \mathbb{C}_{n,n} \tag{9}$$

is unitary. Hence, the singular value decomposition of $P_{\mathcal{R}(A)}P_{\mathcal{R}(A^*)}$ is

$$P_{\mathcal{R}(A)}P_{\mathcal{R}(A^*)} = V(C \oplus 0_{n-r,n-r})(TV^*)$$

since V and TV^* are unitary and $C \oplus 0_{n-r,n-r}$ is a diagonal matrix with real and nonnegative numbers on its diagonal. Therefore, these numbers are the singular values of $P_{\mathcal{R}(A)}P_{\mathcal{R}(A^*)}$. □

Note: From now on the symbols A, M, C, and S will denote the matrices appearing in Theorem 1.

It is straightforward by checking the four conditions of the Moore-Penrose inverse that if A is written as in (2), then

$$A^\dagger = V \begin{bmatrix} CM^{-1} & 0 \\ S^*M^{-1} & 0 \end{bmatrix} V^*. \tag{10}$$

2 Some consequences

We will show that the decomposition given in Theorem 1 permits give a unified approach to many different results in matrix algebra. We present a few (some of them are known). In the forthcoming, we shall denote by $\|\cdot\|$ the Euclidean norm and we shall use the so called C^*-identity: for any matrix X, one has $\|X\|^2 = \|XX^*\| = \|X^*X\|$. Another useful fact is that the Euclidean norm is unitary invariant, i.e., $\|W_1XW_2\| = \|X\|$ for any $X \in \mathbb{C}_{m,n}$ and unitary matrices $W_1 \in \mathbb{C}_{m,m}$, $W_2 \in \mathbb{C}_{n,n}$. Also, we will need that under the notation of Theorem 1 and if $X \in \mathbb{C}_{r,r}$, then

$$\left\| \begin{bmatrix} XC & XS \\ 0 & 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} CX & 0 \\ S^*X & 0 \end{bmatrix} \right\| = \|X\|. \tag{11}$$

The proof of this last affirmation is quite easy: let us define $D = \left[\begin{smallmatrix} XC & XS \\ 0 & 0 \end{smallmatrix}\right]$ and $E = D^*$. Observe that $\|D\|^2 = \|DD^*\| = \|XX^*\| = \|X\|^2$ and $\|E\| = \|D\|$. In particular from (2) and (10) we have $\|A\| = \|M\|$ and $\|A^\dagger\| = \|M^{-1}\|$.

2.1 On the Drazin inverse

We review some elementary facts about the Drazin inverse and the index of a matrix (see [1, Chapter 4] for more information). For $A \in \mathbb{C}_{n,n}$ the *index* of A is the smallest integer $k \geq 0$ such that $\mathrm{rank}(A^{k+1}) = \mathrm{rank}(A^k)$. Such integer always exists. It can be proved that there is a unique matrix, denoted by A^D, such that

$$A^{k+1}A^D = A^k, \qquad AA^D = A^DA, \qquad A^DAA^D = A^D, \qquad (k \text{ is the index of } A).$$

This matrix A^D is *the Drazin inverse of A*. The Drazin inverse of a matrix of index 1 is called *the group inverse* and is customary written $A^\#$. We shall see in next result how to represent the Drazin inverse of a matrix A when A is written as in (2).

Theorem 2. *Let A be represented as in (2). Then*

$$A^D = V \begin{bmatrix} (MC)^D & [(MC)^D]^2MS \\ 0 & 0 \end{bmatrix} V^*.$$

Proof. Let us denote $X = V \left[\begin{smallmatrix} (MC)^D & [(MC)^D]^2MS \\ 0 & 0 \end{smallmatrix}\right] V^*$. The equalities $XA = AX$ and $XAX = X$ are easy to check in view of the defintion of the Drazin inverse of MC. Let m be the index of MC. Thus, $(MC)^{m+1}(MC)^D = (MC)^m$ holds.

On the other hand, by induction, one easily has

$$A^{p+1} = V \begin{bmatrix} (MC)^{p+1} & (MC)^pMS \\ 0 & 0 \end{bmatrix} V^*, \qquad \forall\, p \in \mathbb{N} \cup \{0\}.$$

Thus,

$$\begin{aligned} A^{m+2}X &= V \begin{bmatrix} (MC)^{m+2} & (MC)^{m+1}MS \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (MC)^D & [(MC)^D]^2MS \\ 0 & 0 \end{bmatrix} V^* \\ &= V \begin{bmatrix} (MC)^{m+2}(MC)^D & (MC)^{m+2}[(MC)^D]^2MS \\ 0 & 0 \end{bmatrix} V^*. \end{aligned}$$

Since

$$(MC)^{m+2}(MC)^D = MC(MC)^{m+1}(MC)^D = MC(MC)^m = (MC)^{m+1}$$

and

$$\begin{aligned}&(MC)^{m+2}[(MC)^D]^2MS \\ &\quad = (MC)^{m+1}\left[MC[(MC)^D]^2\right]MS = (MC)^{m+1}(MC)^DMS = (MC)^mMS,\end{aligned}$$

then $A^{m+2}X = A^{m+1}$, which proves $A^D = X$. □

Observe that the proof of the above result distills that if C is nonsingular (which is equivalent to the invertibility of MC, or in other words, 0 is the index of MC), then A is group invertible and

$$A^{\#} = V \begin{bmatrix} C^{-1}M^{-1} & C^{-1}M^{-1}C^{-1}S \\ 0 & 0 \end{bmatrix} V^*. \tag{12}$$

In fact, in [2, Theorem 3.7], it was proved that for a matrix A represented as in (2), then A has group inverse if and only if C is nonsingular. In the next subsection, we shall give another proof of this fact based on the full-rank factorization of the matrix A.

2.2 On the full-rank factorization of a square matrix

If $A \in \mathbb{C}_{n,n}$ is represented as in (2), then we can write explicitly one full-rank factorization of A.

Theorem 3. *Let $A \in \mathbb{C}_{n,n}$ be represented as in (2) and $r = \operatorname{rank}(A)$. Then a full-rank factorization of A is $A = FG$, where*

$$F = V \begin{bmatrix} M \\ 0 \end{bmatrix} \in \mathbb{C}_{n,r}, \qquad G = \begin{bmatrix} C & S \end{bmatrix} V^* \in \mathbb{C}_{r,n}. \tag{13}$$

Furthermore, one has $GG^* = I_r$ and $\|A\| = \|F\|$.

Proof. The proof of $A = FG$ is trivial. The equality $GG^* = I_r$ follows from the first identity of (8). Since multiplying by a nonsingular matrix (and V is nonsingular) does not change the rank, then the rank of F is the rank of $\left[\begin{smallmatrix} M \\ 0 \end{smallmatrix}\right]$, which is r since M is nonsingular and $M \in \mathbb{C}_{r,r}$. Now, $\operatorname{rank}(G) = \operatorname{rank}(GG^*) = \operatorname{rank}(I_r) = r$. Finally, taking into account that the Euclidean norm is unitary invariant and V is unitary, then $\|F\| = \| \left[\begin{smallmatrix} M \\ 0 \end{smallmatrix}\right] \| = \|M\|$. □

The full-rank factorization turns out to be a powerful tool in the study of generalized inverses (see e.g., [1]). We give an easy example in Theorem 4 in where the following result of Cline [7] is used: *Let a square matrix A have the full-rank factorization $A = FG$. Then A has a group inverse if and only if GF is nonsingular, in which case $A^{\#} = F(GF)^{-2}G$.*

Theorem 4. *Let $A \in \mathbb{C}_{n,n}$ be represented as in (2). Then A has group inverse if and only if C is invertible, in which case, (12) holds.*

Proof. We use the aforementioned result of Cline and the nonsingularity of M: Since $GF = CM$, we have $\exists A^{\#} \iff \exists (GF)^{-1} \iff \exists (CM)^{-1} \iff \exists C^{-1}$. Finally, (12) follows from $A^{\#} = F(GF)^{-2}G$ and (13). $\square$

2.3 A result of Djoković

Let $P \in \mathbb{C}_{n,n}$ be a projector (i.e., $P^2 = P$). If we apply to P the decomposition given in Theorem 1 we get $MCMC = MC$ and $MCMS = MS$. The invertibility of M leads to $CM[C\ S] = [C\ S]$. Postmultiplying by $\left[\begin{smallmatrix} C \\ S^* \end{smallmatrix}\right]$ and using (8), we get $CM = I_r$, which yields that C is nonsingular and $M = C^{-1}$ (let us recall that always C is square). Hence we can decompose

$$P = V \begin{bmatrix} I_r & C^{-1}S \\ 0 & 0 \end{bmatrix} V^*. \tag{14}$$

Since C is nonsingular and by recalling the definition of matrix C given in (3), there is no canonical angle between $\mathcal{R}(P)$ and $\mathcal{R}(P^*)$ equal to $\pi/2$. Let $\theta_1, \ldots, \theta_p$ be the canonical angles between $\mathcal{R}(P)$ and $\mathcal{R}(P^*)$ belonging to $]0, \pi/2[$. The angle 0 is a canonical angle between $\mathcal{R}(P)$ and $\mathcal{R}(P^*)$ with multiplicity $x = r - p$ (it may happen that $x = 0$). Now,

$$\begin{aligned} C^{-1}S &= \begin{bmatrix} \operatorname{diag}(\cos\theta_1, \ldots, \cos\theta_p) & 0 \\ 0 & I_x \end{bmatrix}^{-1} \begin{bmatrix} \operatorname{diag}(\sin\theta_1, \ldots, \sin\theta_p) & 0 \\ 0 & 0_{x,n-(r+p)} \end{bmatrix} \\ &= \begin{bmatrix} \operatorname{diag}(\tan\theta_1, \ldots, \tan\theta_p) & 0 \\ 0 & 0_{x,n-(r+p)} \end{bmatrix}. \end{aligned}$$

This is a result given by Djoković. In fact, the original statement of Djoković is the following:

Theorem 5 ([8]). *Let $P \in \mathbb{C}_{n,n}$ be a projector whose rank is r. Then there is a unitary similarity that reduces P to the block diagonal form*

$$\operatorname{diag}\left(\begin{bmatrix} 1 & \sigma_1 \\ 0 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 1 & \sigma_p \\ 0 & 0 \end{bmatrix}, I_x, 0_{n-r,n-r} \right),$$

where $\sigma_1, \ldots, \sigma_p > 0$ and x are uniquely defined by the projector P.

Let us remark that we give a geometrical vision of the numbers $\sigma_1, \ldots, \sigma_p$. See [9] for another geometrical explanation of these numbers.

The expression (14) permits also give a geometric explanation of the unitary similarity reduction stated in the original theorem of Djoković. Let $\mathbf{v} \in \mathcal{R}(P)$ and let us denote

$$\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = V^*\mathbf{v}, \qquad \mathbf{v}_1 \in \mathbb{C}_{r,1},\ \mathbf{v}_2 \in \mathbb{C}_{n-r,1}.$$

Now,

$$V \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \mathbf{v} = P\mathbf{v} = V \begin{bmatrix} I_r & C^{-1}S \\ 0 & 0 \end{bmatrix} V^*V \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = V \begin{bmatrix} \mathbf{v}_1 + C^{-1}S\mathbf{v}_2 \\ \mathbf{0} \end{bmatrix}.$$

Thus $\mathcal{R}(P) \subset \{V \left[\begin{smallmatrix} \mathbf{w}_1 \\ \mathbf{0} \end{smallmatrix}\right] : \mathbf{w}_1 \in \mathbb{C}_{r,1}\}$, being the opposite inclusion obvious. Hence we have proved that in the equality (14), the r first columns of V form an orthonormal basis of $\mathcal{R}(P)$, while the $n - r$ last rows of V form an orthonormal basis of $\mathcal{R}(P)^{\perp}$.

2.4 More on group inverses

Now, let us consider an arbitrary $A \in \mathbb{C}_{n,n}$ being represented as in (2). If A has group inverse, then C is nonsingular, and we can construct the matrix appearing in (14). We shall show that this matrix has a specific meaning: it is known [1, Exercise 30, Chapter 6] that $AA^{\#}$ is the projector onto $\mathcal{R}(A)$ along $\mathcal{N}(A)$. Employing (2) and (12) we have

$$AA^{\#} = V \begin{bmatrix} MC & MS \\ 0 & 0 \end{bmatrix} \begin{bmatrix} C^{-1}M^{-1} & C^{-1}M^{-1}C^{-1}S \\ 0 & 0 \end{bmatrix} V^* = V \begin{bmatrix} I_r & C^{-1}S \\ 0 & 0 \end{bmatrix} V^*, \quad (15)$$

which gives another geometrical vision of the canonical angles between $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$ when A has group inverse.

We can further extract more information from (12) and (15).

Corollary 1. *Let $A \in \mathbb{C}_{n,n}$ have rank r and let $\theta_1 \le \cdots \le \theta_r < \pi/2$ be the canonical angles between $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$. Then*

(i) *The nonzero singular values of $AA^{\#}$ are $1/\cos\theta_1 \le \cdots \le 1/\cos\theta_r$. In particular we have $\|AA^{\#}\| = 1/\cos\theta_r$.*

(ii) $\|A^{\#}\| = \|C^{-1}M^{-1}C^{-1}\|$. *In particular we have $\|A^{\#}\| \le \|A^{\dagger}\|/\cos^2\theta_r$.*

Proof. (i) Observe that (15) can be written as

$$AA^{\#} = V \begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} TV^*, \quad (16)$$

where the matrix T is defined in (9). Since T and V are unitary, and C is real and diagonal, then (16) is the singular value decomposition of $AA^{\#}$, which proves (i).

(ii) In view of (12), we can write

$$A^{\#} = V \begin{bmatrix} (C^{-1}M^{-1}C^{-1})C & (C^{-1}M^{-1}C^{-1})S \\ 0 & 0 \end{bmatrix} V^*.$$

It is enough to apply (11) to get (ii). In particular, we have $\|C^{-1}M^{-1}C^{-1}\| \le \|C^{-1}\|^2\|M^{-1}\|$. Recalling $\|A^{\dagger}\| = \|M^{-1}\|$ finishes the proof. □

2.5 Some expressions involving limits and generalized inverses

If $A \in \mathbb{C}_{m,n}$, then one has

$$\lim_{t\to 0+} A^*(AA^* + tI_m)^{-1} = A^{\dagger}. \quad (17)$$

See [1, Section 3.3, Section 4.4] for three proofs and [5] for the original statement. Also, it is known that

$$t > 0 \qquad \Rightarrow \qquad \|A^*(AA^* + tI_m)^{-1} - A^{\dagger}\| < t\|A^{\dagger}\|^3, \quad (18)$$

which was proved by Boyarintsev in [4, Theorem 1.2.3]. Evidently, (18) implies (17). We shall prove (18) by using Theorem 1.

Theorem 6. *Let $A \in \mathbb{C}_{m,n}$. Then (18) holds.*

Proof. First, assume that A is square. If we represent A as in (2), then by using the first identity of (8) we get

$$\begin{aligned}(AA^* + tI_n)^{-1} &= V\left(\begin{bmatrix} MC & MS \\ 0 & 0 \end{bmatrix}\begin{bmatrix} CM^* & 0 \\ S^*M^* & 0 \end{bmatrix} + \begin{bmatrix} tI_r & 0 \\ 0 & tI_{n-r} \end{bmatrix}\right)^{-1} V^* \\ &= V\begin{bmatrix} (MM^* + tI_r)^{-1} & 0 \\ 0 & t^{-1}I_{n-r} \end{bmatrix} V^*. \end{aligned} \tag{19}$$

And now, by employing (10), we get

$$\begin{aligned}&A^\dagger - A^*(AA^* + tI_n)^{-1} \\ &\quad = V\left\{\begin{bmatrix} CM^{-1} & 0 \\ S^*M^{-1} & 0 \end{bmatrix} - \begin{bmatrix} CM^* & 0 \\ S^*M^* & 0 \end{bmatrix}\begin{bmatrix} (MM^* + tI_r)^{-1} & 0 \\ 0 & t^{-1}I_{n-r} \end{bmatrix}\right\} V^* \\ &\quad = V\begin{bmatrix} C[M^{-1} - M^*(MM^* + tI_r)^{-1}] & 0 \\ S^*[M^{-1} - M^*(MM^* + tI_r)^{-1}] & 0 \end{bmatrix} V^*. \end{aligned} \tag{20}$$

By using (11) one gets

$$\|A^\dagger - A^*(AA^* + tI_n)^{-1}\| = \|M^{-1} - M^*(MM^* + tI_r)^{-1}\|. \tag{21}$$

Let us take N in such a way that $N^{-1} = M^*(MM^* + tI_r)^{-1}$ (observe that this is feasible since both M^* and $MM^* + tI_r$ are nonsingular). From the definition of N we get $N = (MM^* + tI_r)(M^*)^{-1} = M + t(M^*)^{-1}$. Thus

$$\begin{aligned}&M^{-1} - M^*(MM^* + tI_r)^{-1} = M^{-1} - N^{-1} = M^{-1}(N - M)N^{-1} \\ &\quad = M^{-1}(t(M^*)^{-1})M^*(MM^* + tI_r)^{-1} = tM^{-1}(MM^* + tI_r)^{-1}. \end{aligned} \tag{22}$$

By applying the spectral theorem to the definite positive matrix MM^*, there exist $\lambda_1 \geq \cdots \geq \lambda_r > 0$ and a unitary matrix $U \in \mathbb{C}_{r,r}$ such that $MM^* = U \operatorname{diag}(\lambda_1, \ldots, \lambda_r)U^*$. Hence

$$\begin{aligned}&\|(MM^* + tI_r)^{-1}\| = \left\| U \operatorname{diag}\left(\frac{1}{\lambda_1 + t}, \ldots, \frac{1}{\lambda_r + t}\right) U^* \right\| \\ &\quad = \frac{1}{\lambda_r + t} < \frac{1}{\lambda_r} = \|(MM^*)^{-1}\| = \|(M^{-1})^* M^{-1}\| = \|M^{-1}\|^2 = \|A^\dagger\|^2. \end{aligned} \tag{23}$$

Now, (21), (22), and (23) finish the proof when A is square. If $A \in \mathbb{C}_{m,n}$ is not square, then $n < m$ or $n > m$. If $n < m$, then define $\widetilde{A} = [A \ 0_{m-n,n}] \in \mathbb{C}_{m,m}$ and apply the result for square matrices to get (18). If $n > m$, then similarly, add zero rows to A to prove the theorem. ☐

In general for $A \in \mathbb{C}_{m,n}$, one has that $\exists \lim_{t \to 0+}(AA^* + tI_m)^{-1}$ if and only if A is nonsingular. However, as we saw, $\lim_{t \to 0+} A^*(AA^* + tI_n)^{-1}$ always exists. In the following result we investigate the matrices B such that $\lim_{t \to 0+} B(AA^* + tI_n)^{-1}$ exists.

Theorem 7. *Let $A \in \mathbb{C}_{m,n}$ and $B \in \mathbb{C}_{q,m}$. Then $\lim_{t\to 0+} B(AA^* + tI_m)^{-1}$ exists if and only if $\mathcal{R}(B^*) \subset \mathcal{R}(A)$. In such case, for any $t > 0$ one has*

$$\|B(AA^* + tI_m)^{-1} - B(AA^*)^\dagger\| < t\|B\|\|A^\dagger\|^4. \tag{24}$$

In particular, $\lim_{t\to 0+} B(AA^* + tI_m)^{-1} = B(AA^*)^\dagger$.

Proof. In the first part of the proof, we will assume that $m = n$. Let us prove the equivalence. If A is nonsingular, then $\mathcal{R}(B^*) \subset \mathcal{R}(A)$ is evident. If A is singular, let $r = \text{rank}(A) < n$, and we can represent A as in (2). Furthermore, we write B as follows:

$$B = \begin{bmatrix} B_1 & B_2 \end{bmatrix} V^*, \qquad B_1 \in \mathbb{C}_{q,r},\ B_2 \in \mathbb{C}_{q,n-r}.$$

Now, (19) leads to

$$B(AA^* + tI_n)^{-1} = \begin{bmatrix} B_1(MM^* + tI_r)^{-1} & t^{-1}B_2 \end{bmatrix} V^*,$$

which shows that $\lim_{t\to 0+} B(AA^* + tI_n)^{-1}$ exists if and only if $B_2 = 0$ since M is nonsingular. On the other hand, by (2), (8), and (10) we have

$$BAA^\dagger = \begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} MC & MS \\ 0 & 0 \end{bmatrix} \begin{bmatrix} CM^{-1} & 0 \\ S^*M^{-1} & 0 \end{bmatrix} V^* = \begin{bmatrix} B_1 & 0 \end{bmatrix} V^*.$$

Hence, $B_2 = 0$ if and only if $BAA^\dagger = B$. It is simple to prove that $BAA^\dagger = B$ is equivalent to $\mathcal{R}(B^*) \subset \mathcal{R}(A)$.

Assume that $\mathcal{R}(B^*) \subset \mathcal{R}(A)$. It is clear that

$$\begin{aligned} B(AA^* + tI_n)^{-1} - B(AA^*)^\dagger &= \left(\begin{bmatrix} B_1(MM^* + tI_r)^{-1} & 0 \end{bmatrix} - \begin{bmatrix} B_1(MM^*)^{-1} & 0 \end{bmatrix} \right) V^* \\ &= \begin{bmatrix} B_1[(MM^* + tI_r)^{-1} - (MM^*)^{-1}] & 0 \end{bmatrix} V^* \\ &= \begin{bmatrix} B_1 & 0 \end{bmatrix} \begin{bmatrix} (MM^* + tI_r)^{-1} - (MM^*)^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^*. \end{aligned}$$

Therefore,

$$\|B(AA^* + tI_n)^{-1} - B(AA^*)^\dagger\| \le \|B\|\ \|(MM^* + tI_r)^{-1} - (MM^*)^{-1}\|.$$

The last norm in the above equation can be bounded by a standard way: by using the equality $P^{-1} - Q^{-1} = P^{-1}(Q - P)Q^{-1}$ valid for any two invertible matrices P and Q we have

$$\|B(AA^* + tI_n)^{-1} - B(AA^*)^\dagger\| \le t\|B\|\ \|(MM^* + tI_r)^{-1}\|\ \|(MM^*)^{-1}\|.$$

By the proof of Theorem 6 (see the inequality (23)) one has $\|(MM^* + tI_r)^{-1}\| < \|A^\dagger\|^2$ and $\|(MM^*)^{-1}\| = \|A^\dagger\|^2$. This proves this theorem when A is a square matrix.

If $m < n$, let us define $\widetilde{A} = \left[\begin{smallmatrix} A \\ 0_{n-m,m} \end{smallmatrix}\right] \in \mathbb{C}_{n,n}$ and $\widetilde{B} = [B\ 0_{p,n-m}] \in \mathbb{C}_{p,n}$. By taking into account the following elementary facts and applying the theorem for the square matrix $\widetilde{A}$ and $\widetilde{B}$, the theorem can be proved when $m < n$.

- $\exists \lim_{t\to 0+} B(AA^* + tI_m)^{-1} \iff \exists \lim_{t\to 0+} \widetilde{B}(\widetilde{A}\widetilde{A}^* + tI_n)^{-1}$.

- $\mathcal{R}(B^*) \subset \mathcal{R}(A) \iff \mathcal{R}(\widetilde{B}^*) \subset \mathcal{R}(\widetilde{A})$.
- $\|B(AA^* + tI_m)^{-1} - B(A^*A)^\dagger\| = \|\widetilde{B}(\widetilde{A}\widetilde{A}^* + tI_n)^{-1} - \widetilde{B}(\widetilde{A}^*\widetilde{A})^\dagger\|$, $\|B\| = \|\widetilde{B}\|$, and $\|A^\dagger\| = \|\widetilde{A}^\dagger\|$.

If $m > n$, let us define $\widetilde{A} = [A\ 0_{m,m-n}] \in \mathbb{C}_{m,m}$. Now we can easily check $B(\widetilde{A}\widetilde{A}^* + tI_m)^{-1} = B(AA^* + tI_m)^{-1}$, $\mathcal{R}(\widetilde{A}) = \mathcal{R}(A)$, $B(\widetilde{A}^*\widetilde{A})^\dagger = B(A^*A)^\dagger$, and $\|\widetilde{A}^\dagger\| = \|A^\dagger\|$, which finishes the proof. □

Observe that (under the notation of Theorem 7), a natural choice for $\mathcal{R}(B^*) \subset \mathcal{R}(A)$ holds is $B = A^*$. If we apply Theorems 6 and 7 to $B = A^*$, one gets $A^\dagger = A^*(AA^*)^\dagger$, a well known identity of the Moore-Penrose inverses. Furthermore, observe that the inequality (24) when $B = A^*$ is weaker than the inequality (18) since $1 = \|AA^\dagger\| \leq \|A\|\|A^\dagger\|$.

By a similar technique than in Theorem 7, one gets the following result.

Theorem 8. *Let $A \in \mathbb{C}_{m,n}$ and $B \in \mathbb{C}_{n,m}$. Then $\lim_{t\to 0^+} B(AA^\dagger + tI_m)^{-1}$ exists if and only if $\mathcal{R}(B^*) \subset \mathcal{R}(A)$. In such case, that for any $t > 0$ one has $B(AA^\dagger + tI_m)^{-1} = (1+t)^{-1}B$. In particular, $\lim_{t\to 0^+} B(AA^\dagger + tI_m)^{-1} = B$.*

It is known (see for example [1, Section 4.4]) that for a given $A \in \mathbb{C}_{n,n}$, then there exists $A^\#$ if and only if $\lim_{\lambda\to 0}(\lambda I_n + A)^{-1}A$ exists, in which case, $\lim_{\lambda\to 0}(\lambda I_n + A)^{-1}A = AA^\#$. Here and in the following, $\lambda \to 0$ means $\lambda \to 0$ through any neighborhood $\mathcal{U}$ of 0 in $\mathbb{C}$ such that $A + \lambda I_n$ is nonsingular for $\lambda \in \mathcal{U} \setminus \{0\}$ (observe that such net of neighborhoods exists since the cardinal of $\{z \in \mathbb{C} : \det(A + zI_n) = 0\}$ is finite). Also it is known [1, Section 4.4] that for $A \in \mathbb{C}_{n,n}$, then

$$AA^\dagger = A^\dagger A \qquad \iff \qquad \lim_{\lambda\to 0}(\lambda I_n + A)^{-1}P_{\mathcal{R}(A)} = A^\dagger. \tag{25}$$

Recall that $P_{\mathcal{R}(A)}$ denotes the orthogonal projection onto $\mathcal{R}(A)$.

Observe that for $A \in \mathbb{C}_{n,n}$ as in (2), then (in view of the nonsingularity of M) $AA^\#A^\dagger = A^\dagger$ implies $S = 0$. Since $C^2 + SS^* = I_r$ and C is a diagonal matrix all of whose components of its main diagonal are nonnegative, we have $C = I_r$. Hence we can decompose $A = V(M \oplus 0)V^*$, where $V \in \mathbb{C}_{n,n}$ is unitary and $M \in \mathbb{C}_{r,r}$ nonsingular, being this last statement equivalent to $AA^\dagger = A^\dagger A$ (see [6, Theorem 4.3.1]). Reciprocally, if $A \in \mathbb{C}_{n,n}$ have rank r and can be decomposed as $A = V(K \oplus 0)V^*$, being $V \in \mathbb{C}_{n,n}$ unitary and $K \in \mathbb{C}_{r,r}$ nonsingular, then $A^\dagger = A^\# = V(K^{-1} \oplus 0)V^*$, which yields $AA^\#A^\dagger = A^\dagger$. Therefore, for $A \in \mathbb{C}_{n,n}$, one has $AA^\#A^\dagger = A^\dagger \iff AA^\dagger = A^\dagger A$, which explains the link between (25) and the second item of the next theorem.

Parenthetically, a matrix A such that $AA^\dagger = A^\dagger A$ is said *EP* (this name comes from Equal Projection) or *Range-Hermitian.*

We shall see how the Theorem 1 works in these situations.

Theorem 9. *Let $A \in \mathbb{C}_{n,n}$ be singular. Then*

(i) *There exists $\lim_{\lambda\to 0}(A + \lambda I_n)^{-1}A$ if and only if A is group invertible, in which case, $\lim_{\lambda\to 0}(A + \lambda I_n)^{-1}A = AA^\#$.*

(ii) *There exists* $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}AA^\dagger$ *if and only if* A *is group invertible, in which case,* $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}AA^\dagger = AA^\# A^\dagger$.

(iii) *There exists* $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^*$ *if and only if* A *is Range-Hermitian, in which case,* $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^* = A^\dagger A^*$.

(iv) *There exists* $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^\dagger$ *if and only if* A *is Range-Hermitian, in which case,* $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^* = (A^\dagger)^2$.

(v) *If* A *is group invertible, then* $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^\# = (A^\#)^2$.

Proof. Let us represent A as in (2). Pick $\mathcal{U}$ any neighborhood of 0 in $\mathbb{C}$ such that $A+\lambda I_n$ is nonsingular for $\lambda\in\mathcal{U}\setminus\{0\}$ and take a fixed $\lambda\in\mathcal{U}\setminus\{0\}$. Since

$$A+\lambda I_n = V\begin{bmatrix} MC+\lambda I_r & MS \\ 0 & \lambda I_{n-r}\end{bmatrix} V^*,$$

we get that $MC+\lambda I_r$ is nonsingular and

$$(A+\lambda I_n)^{-1} = V\begin{bmatrix} (MC+\lambda I_r)^{-1} & -\lambda^{-1}(MC+\lambda I_r)^{-1}MS \\ 0 & \lambda^{-1}I_{n-r}\end{bmatrix} V^*. \tag{26}$$

(i) We get

$$(A+\lambda I_n)^{-1}A = V\begin{bmatrix} (MC+\lambda I_r)^{-1}MC & (MC+\lambda I_r)^{-1}MS \\ 0 & 0\end{bmatrix} V^*. \tag{27}$$

If we assume the existence of $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A$, then (27) implies that $(MC+\lambda I_n)^{-1}MC$ and $(MC+\lambda I_n)^{-1}MS$ have limit when $\lambda\to 0$. Thus, $\exists\lim_{\lambda\to 0}(MC+\lambda I_n)^{-1}MC^2$ and $\exists\lim_{\lambda\to 0}(MC+\lambda I_n)^{-1}MSS^*$, and therefore exists (observe that we use $C^2+SS^*=I_r$)

$$\lim_{\lambda\to 0}(MC+\lambda I_n)^{-1}MC^2 + (MC+\lambda I_n)^{-1}MSS^* = \lim_{\lambda\to 0}(MC+\lambda I_n)^{-1}M.$$

Since M is nonsingular, there exists $\lim_{\lambda\to 0}(MC+\lambda I_n)^{-1}$, which is equivalent to the nonsingularity of MC, which, as we have seen, is equivalent to the group invertibility of A. Furthermore, (15) and (27) imply

$$(A+\lambda I_n)^{-1}A \xrightarrow{\lambda\to 0} V\begin{bmatrix} I_r & (MC)^{-1}MS \\ 0 & 0\end{bmatrix} V^* = V\begin{bmatrix} I_r & C^{-1}S \\ 0 & 0\end{bmatrix} V^* = AA^\#.$$

If MC is nonsingular, it is clear from (27) that $\exists\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A$.

(ii) Since $AA^\dagger A = A$ and (i) have just been proved, it is clear that (ii) holds.

(iii) We have

$$\begin{aligned}(A+\lambda I_n)^{-1}A^* &= V\begin{bmatrix} (MC+\lambda I_r)^{-1} & -\lambda^{-1}(MC+\lambda I_r)^{-1}MS \\ 0 & \lambda^{-1}I_{n-r}\end{bmatrix}\begin{bmatrix} CM^* & 0 \\ S^*M^* & 0\end{bmatrix} V^* \\ &= V\begin{bmatrix} (MC+\lambda I_r)^{-1}CM^* - \lambda^{-1}(MC+\lambda I_r)^{-1}MSS^*M^* & 0 \\ \lambda^{-1}S^*M^* & 0\end{bmatrix} V^*. \end{aligned} \tag{28}$$

If $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^*$ exists, then $\lim_{\lambda\to 0}\lambda^{-1}S^*M^*$ exists, which in view of the nonsingularity of M, leads to $S=0$. Since $C^2+SS^*=I_r$ and C is a diagonal matrix being nonnegative its entries, we obtain $C=I_r$. Hence, A can be decomposed as $A=V(M\oplus 0)V^*$, being $M\in\mathbb{C}_{r,r}$ nonsingular, $V\in\mathbb{C}_{n,n}$ unitary, and $r=\operatorname{rank}(A)$. Therefore, A is Range-Hermitian, and (28) together with $A=V(M\oplus 0)V^*$, $C=I_r$ and $S=0$ shows that

$$\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^*=\lim_{\lambda\to 0}V\begin{bmatrix}(M+\lambda I_r)^{-1}M^* & 0\\ 0 & 0\end{bmatrix}V^*=V\begin{bmatrix}M^{-1}M^* & 0\\ 0 & 0\end{bmatrix}V^*=A^\dagger A^*.$$

If A is Range-Hermitian, by the decomposition $A=V(M\oplus 0)V^*$, being $M\in\mathbb{C}_{r,r}$ nonsingular, $V\in\mathbb{C}_{n,n}$ unitary, and $r=\operatorname{rank}(A)$, then $(A+\lambda I_n)^{-1}A^*=V((M+\lambda I_r)^{-1}M^*\oplus 0)V^*$, which proves the existence of $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^*$.

(iv) We have

$$\begin{aligned}(A+\lambda I_n)^{-1}A^\dagger &= V\begin{bmatrix}(MC+\lambda I_r)^{-1} & -\lambda^{-1}(MC+\lambda I_r)^{-1}MS\\ 0 & \lambda^{-1}I_{n-r}\end{bmatrix}\begin{bmatrix}CM^{-1} & 0\\ S^*M^{-1} & 0\end{bmatrix}V^*\\ &= V\begin{bmatrix}(MC+\lambda I_r)^{-1}CM^{-1}-\lambda^{-1}(MC+\lambda I_r)^{-1}MSS^*M^{-1} & 0\\ \lambda^{-1}S^*M^{-1} & 0\end{bmatrix}V^*. \quad (29)\end{aligned}$$

If $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}A^\dagger$ exists, then (29) shows that $\lim_{\lambda\to 0}\lambda^{-1}S^*M^{-1}$ exists, which leads to $S=0$. As in the previous proof of (iii), we get $C=I_r$ and A is Range-Hermitian. Furthermore, (29), $C=I_r$, and $S=0$ imply

$$(A+\lambda I_n)^{-1}A^\dagger=V\begin{bmatrix}(M+\lambda I_r)^{-1}M^{-1} & 0\\ 0 & 0\end{bmatrix}V^*\xrightarrow{\lambda\to 0}V\begin{bmatrix}M^{-2} & 0\\ 0 & 0\end{bmatrix}V^*=(A^\dagger)^2.$$

If A is Range-Hermitian, as in the proof of the previous item (iii), by means the decomposition $A=V(M\oplus 0)V^*$, where $V\in\mathbb{C}_{n,n}$ is unitary and M is nonsingular, we can easily get that $\lim_{\lambda\to 0}(A+\lambda I_n)A^\dagger$ exists.

(v) We have

$$\begin{aligned}&(A+\lambda I_n)^{-1}A^\#\\ &\quad=V\begin{bmatrix}(MC+\lambda I_r)^{-1} & -\lambda^{-1}(MC+\lambda I_r)^{-1}MS\\ 0 & \lambda^{-1}I_{n-r}\end{bmatrix}\begin{bmatrix}C^{-1}M^{-1} & C^{-1}M^{-1}C^{-1}S\\ 0 & 0\end{bmatrix}V^*\\ &\quad=V\begin{bmatrix}(MC+\lambda I_r)^{-1}C^{-1}M^{-1} & (MC+\lambda I_r)^{-1}C^{-1}M^{-1}C^{-1}S\\ 0 & 0\end{bmatrix}V^*\\ &\quad\xrightarrow{\lambda\to 0}V\begin{bmatrix}(MC)^{-2} & (MC)^{-2}C^{-1}S\\ 0 & 0\end{bmatrix}V^*=(A^\#)^2.\end{aligned}$$

□

In the next theorem 10, we shall investigate expressions of the form $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}B$ for matrices $A\in\mathbb{C}_{n,n}$ and $B\in\mathbb{C}_{n,m}$. Evidently, if A is singular, then $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}$ does not exist, but it may happen that $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}B$ exists for concrete matrices B. Also, it is worthy to note that if A is nonsingular, then $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}B$ exists for any matrix $B\in\mathbb{C}_{n,m}$ (and this limit is $A^{-1}B$). But before, we will prove a simple lemma.

Lemma 2. *Let $X \in \mathbb{C}_{r,m}$. Then X have rank r if and only if then there exists $Y \in \mathbb{C}_{m,r}$ such that $XY = I_r$.*

Proof. If $m = r$, the lemma is evident. Thus, we can assume in the following that $m \neq r$.

Assume that $\operatorname{rank}(X) = r$. Let $\mathbf{x}_1, \ldots, \mathbf{x}_r \in \mathbb{C}_{1,m}$ be the rows of X. These r vectors are linearly independent since $r = \operatorname{rank}(X)$, and thus, $r \leq m$. Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ be a basis of $\mathbb{C}_{1,m}$ and let us define the matrices $X_1 = \begin{bmatrix} \mathbf{x}_{r+1} \\ \cdots \\ \mathbf{x}_m \end{bmatrix}$ and $T = \begin{bmatrix} X \\ X_1 \end{bmatrix}$. Observe that T is nonsingular since $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ is a basis of $\mathbb{C}_{1,m}$ and let us partition $T^{-1} = \begin{bmatrix} Y & Y_1 \end{bmatrix}$, where $Y \in \mathbb{C}_{m,r}$ and $Y_1 \in \mathbb{C}_{m,r-m}$. Now, $\begin{bmatrix} I_r & 0 \\ 0 & I_{m-r} \end{bmatrix} = I_m = TT^{-1} = \begin{bmatrix} X \\ X_1 \end{bmatrix} \begin{bmatrix} Y & Y_1 \end{bmatrix}$ gives $I_r = XY$.

If there exists $Y \in \mathbb{C}_{m,r}$ such that $XY = I_r$, then by using $\operatorname{rank}(X) \leq r$ since $X \in \mathbb{C}_{r,m}$, we have $r = \operatorname{rank}(I_r) = \operatorname{rank}(XY) \leq \operatorname{rank}(X) \leq r$. $\square$

Theorem 10. *Let $A \in \mathbb{C}_{n,n}$ have rank $r < n$ and represented as in (2) and $B \in \mathbb{C}_{n,m}$. Then*

(i) *If $\lim_{\lambda \to 0}(A + \lambda I_n)^{-1}B = E$, then $B = AE$ and $\mathcal{R}(E) \subset \mathcal{R}(A)$.*

(ii) *If $\lim_{\lambda \to 0}(A + \lambda I_n)^{-1}B = E$, and $\operatorname{rank}(E) = r$, then $B = AE$, $\mathcal{R}(E) \subset \mathcal{R}(A)$, and MC is group invertible.*

(iii) *If $\lim_{\lambda \to 0}(A + \lambda I_n)^{-1}B$ exists and A is group invertible, then $\lim_{\lambda \to 0}(A + \lambda I_n)^{-1}B = A^{\#}B$.*

(iv) *If there exists $F \in \mathbb{C}_{n,m}$ such that $B = AF$ and $\mathcal{R}(F) \subset \mathcal{R}(A)$ and MC is group invertible, then there exists $\lim_{\lambda \to 0}(A + \lambda I_n)^{-1}B$.*

(v) *If there exists $F \in \mathbb{C}_{n,m}$ such that $B = AF$ and $\mathcal{R}(F) \subset \mathcal{R}(A)$ and A is group invertible, then $\lim_{\lambda \to 0}(A + \lambda I_n)^{-1}B = F$ and*

$$|\lambda| < \frac{1}{\|(MC)^{-1}\|} \quad \Rightarrow \quad \|(A + \lambda I_n)^{-1}B - F\| \leq |\lambda| \frac{\|F\| \|(MC)^{-1}\|}{1 - |\lambda| \|(MC)^{-1}\|}. \tag{30}$$

Proof. Let us represent A as in (2). Pick $\mathcal{U}$ any neighborhood of 0 in $\mathbb{C}$ such that $A + \lambda I_n$ is nonsingular for $\lambda \in \mathcal{U}$ and take a fixed $\lambda \in \mathcal{U}$. As in Theorem 9, we will use (26). Pick now any matrix $B \in \mathbb{C}_{n,m}$ represented by

$$B = V \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \qquad B_1 \in \mathbb{C}_{r,m}, \; B_2 \in \mathbb{C}_{n-r,m} \tag{31}$$

Thus,

$$(A + \lambda I)^{-1}B = V \begin{bmatrix} (MC + \lambda I_r)^{-1}B_1 - \lambda^{-1}(MC + \lambda I_r)^{-1}MSB_2 \\ \lambda^{-1}B_2 \end{bmatrix}. \tag{32}$$

(i): Assume that there exists $\lim_{\lambda \to 0}(A + \lambda I)^{-1}B$. By (32), one can deduce $B_2 = 0$. Using $B_2 = 0$ and (32) we obtain that $\lim_{\lambda \to 0}(MC + \lambda I_r)^{-1}B_1$ exists. Let us denote

$$X = \lim_{\lambda \to 0}(MC + \lambda I_r)^{-1}B_1. \tag{33}$$

Therefore, $B_2 = 0$ and (32) lead to

$$E = \lim_{\lambda\to 0}(A+\lambda I)^{-1}B = \lim_{\lambda\to 0} V \begin{bmatrix} (MC+\lambda I_r)^{-1}B_1 \\ 0 \end{bmatrix} = V \begin{bmatrix} X \\ 0 \end{bmatrix}. \tag{34}$$

Now, $B_1 = (MC+\lambda I_r)(MC+\lambda I_r)^{-1}B_1 \xrightarrow{\lambda\to 0} MCX$. Hence $B_1 = MCX$. By (2), (31), (34), and $B_2 = 0$ one gets

$$B = V \begin{bmatrix} MCX \\ 0 \end{bmatrix} = V \begin{bmatrix} MC & MS \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X \\ 0 \end{bmatrix} = AE.$$

Furthermore, it is easy to check $AA^\dagger E = E$, which is equivalent to $\mathcal{R}(E) \subset \mathcal{R}(A)$.

(ii): We can use the proof of the previous item (i). It remains to prove that MC is group invertible. By hypothesis, the rank of E is r, hence (34) leads to $\text{rank}(X) = r$. By Lemma 2, there exists $Y \in \mathbb{C}_{m,r}$ such that $XY = I_r$. Now,

$$(MC+\lambda I_r)^{-1}B_1Y = (MC+\lambda I_r)^{-1}MCXY = (MC+\lambda I_r)^{-1}MC.$$

Since $\lim_{\lambda\to 0}(MC+\lambda I_r)^{-1}B_1Y$ exists (because from (33), $\lim_{\lambda\to 0}(MC+\lambda I_r)^{-1}B_1$ exists) we obtain that there exists $\lim_{\lambda\to 0}(MC+\lambda I_r)^{-1}MC$. Now, item (i) of Theorem 9 finishes the proof of this implication.

(iii): Also, we shall use the proof of item (i). Since A is group invertible, Theorem 4 together the nonsingularity of M lead to the existence of $(MC)^{-1}$. From (33) we get $X = (MC)^{-1}B_1$. Thus, from (34), we obtain

$$E = V \begin{bmatrix} X \\ 0 \end{bmatrix} = V \begin{bmatrix} (MC)^{-1}B_1 \\ 0 \end{bmatrix} = V \begin{bmatrix} (MC)^{-1} & (MC)^{-1}C^{-1}S \\ 0 & 0 \end{bmatrix} V^*V \begin{bmatrix} B_1 \\ 0 \end{bmatrix}.$$

By observing that in the proof of (i) we obtained $B_2 = 0$ and taking into account the representations (12), (31) for $A^\#$, B, respectively, we get $E = A^\# B$.

(iv): Let us represent $F = V \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$, where $F_1 \in \mathbb{C}_{r,m}$ and $F_2 \in \mathbb{C}_{n-r,m}$. One has $\mathcal{R}(F) \subset \mathcal{R}(A) \iff AA^\dagger F = F \iff F_2 = 0$. Therefore,

$$B = AF = V \begin{bmatrix} MC & MS \\ 0 & 0 \end{bmatrix} VV^* \begin{bmatrix} F_1 \\ 0 \end{bmatrix} = V \begin{bmatrix} MCF_1 \\ 0 \end{bmatrix}.$$

Now we use (31) and (32) to get

$$(A+\lambda I_n)^{-1}B = V \begin{bmatrix} (MC+\lambda I_r)^{-1}MCF_1 \\ 0 \end{bmatrix}. \tag{35}$$

By hypothesis, MC is group invertible, and having in mind item (i) of Theorem 9, there exists $\lim_{\lambda\to 0}(MC+\lambda I_r)^{-1}MC$, which in conjunction with (35), shows that exists $\lim_{\lambda\to 0}(A+\lambda I_n)^{-1}B$.

(v): We use the proof of the previous item (iii). Since A is group invertible, Theorem 4 and the nonsingularity of M lead to the nonsingularity of MC. Therefore, the representation of F in the proof of the previous item (iii), $F_2 = 0$, and (35) yield to

$$(A+\lambda I_n)^{-1}B \xrightarrow{\lambda\to 0} V \begin{bmatrix} (MC)^{-1}MCF_1 \\ 0 \end{bmatrix} = V \begin{bmatrix} F_1 \\ 0 \end{bmatrix} = F. \tag{36}$$

We shall prove (30). From (35) and (36) one has

$$\|(A+\lambda I_n)^{-1}B - F\| = \|(MC+\lambda I_r)^{-1}MCF_1 - F_1\| \le \|(MC+\lambda I_r)^{-1}MC - I_r\|\|F_1\|.$$

Observe that (36) implies that $\|F_1\| = \|F\|$. Also we have

$$\begin{aligned}(MC+\lambda I_r)^{-1}MC - I_r &= [MC(I_r+\lambda(MC)^{-1}]^{-1}MC - I_r \\ &= [I_r+\lambda(MC)^{-1}]^{-1}(MC)^{-1}MC - I_r = [I_r+\lambda(MC)^{-1}]^{-1} - I_r.\end{aligned}$$

Denote $H = -\lambda(MC)^{-1}$. If $|\lambda| < 1/\|(MC)^{-1}\|$, then $\|H\| < 1$ and $I_r + H + H^2 + \cdots = (I_r - H)^{-1}$. Therefore

$$\begin{aligned}&\|[I_r+\lambda(MC)^{-1}]^{-1} - I_r\| \\ &\quad = \|(I_r-H)^{-1} - I_r\| \le \sum_{n=1}^{\infty}\|H\|^n = \frac{\|H\|}{1-\|H\|} = \frac{|\lambda|\|(MC)^{-1}\|}{1-|\lambda|\|(MC)^{-1}\|}.\end{aligned}$$

The previous computations prove (30). □

Remark: Let $A \in \mathbb{C}_{n,n}$ have group inverse and represented as in (2). Let $\theta_1 \le \cdots \le \theta_r < \pi/2$ be the canonical angles between $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$. Since $\|(MC)^{-1}\| \le \|C^{-1}\|\|M^{-1}\| = \|A^\dagger\|/\cos\theta_r$, then the implication (30) can be changed by the following weaker but somehow simpler.

$$|\lambda| < \frac{\cos\theta_r}{\|A^\dagger\|} \quad \Rightarrow \quad \|(A+\lambda I_n)^{-1}B - F\| \le |\lambda|\frac{\|F\|\|A^\dagger\|}{\cos\theta_r - |\lambda|\|A^\dagger\|}.$$

Acknowledgments

The first author is supported by Work funded by Vicerrectorado de Investigación U.P.V. PAID 06-2010-2285. The second author is supported by the National Natural Science Foundation of China (11061005), the Ministry of Education Science and grants (HCIC201103) of Guangxi Key Laborarory of Hybrid Computational and IC Design Analysis Open Fund.

References

[1] A. Ben-Israel, T.N.E Greville, *Generalized Inverses: Theory and Applications*, Wiley-Interscience New York, 1974, second ed. Springer, New York, 2002.

[2] J. Benítez, *A new decomposition for square matrices*, Electron. J. Linear Algebra, 20, 207-225 (2010).

[3] Å. Björck, *Numerical Methods for Least Squares Problems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1996.

[4] Yu. E. Boyarintsev, Methods of Solving Singular Systems of Ordinary Differential Equations, Wiley, Chichester, 1992, (translation of the 1988 Russian original).

[5] C. G. den Broeder Jr., A. Charnes, Contributions to the theory of generalized inverses for matrices, Technical Report, Purdue University, Department of Mathematics, Lafayette, IN, 1957, (reprinted as ONR Res. Memo. No. 39, Northwestern University, Evanston, IL, 1962).

[6] S.L. Campbell, C.D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.

[7] R.E. Cline, Inverses of rank invariant powers of a matrix, SIAM J. Numer. Anal., 5, 182-197 (1968).

[8] D. Djoković, Unitary similarity of projectors, Aequationes Mathematicae, 42, 220-224 (1991).

[9] A. Galántai, Subspaces, angles and pairs of orthogonal projections, Linear Multilinear Algebra, 56, 227-260 (2008).

[10] G. H. Golub, C. F. Van Loan, *Matrix Computations*, third ed. Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, 1996.

[11] M. Hegland, J. Garcke, V. Challis, The combination technique and some generalisations, Linear Algebra Appl., 420, 249-275 (2007).

[12] C. C. Paige, M. Wei, History and generality of the CS decomposition, Linear Algebra Appl., 208-209, 303-326 (1994).

[13] H.K. Wimmer, Canonical angles of unitary spaces and perturbations of direct complements, Linear Algebra Appl., 287, 373-379 (1999).