A generalization of the Moore-Penrose inverse related to matrix subspaces of $\mathbb{C}^{n \times m}$☆

Antonio Suárez, Luis González*

*Department of Mathematics, University of Las Palmas de Gran Canaria,
35017 Las Palmas de Gran Canaria, Spain*

Abstract

A natural generalization of the classical Moore-Penrose inverse is presented. The so-called S-Moore-Penrose inverse of a $m \times n$ complex matrix A, denoted by $A_S^\dagger$, is defined for any linear subspace S of the matrix vector space $\mathbb{C}^{n \times m}$. The S-Moore-Penrose inverse $A_S^\dagger$ is characterized using either the singular value decomposition or (for the nonsingular square case) the orthogonal complements with respect to the Frobenius inner product. These results are applied to the preconditioning of linear systems based on Frobenius norm minimization and to the linearly constrained linear least squares problem.

Keywords: Moore-Penrose inverse; Frobenius norm; S-Moore-Penrose inverse; Approximate inverse preconditioning; Constrained least squares problem

1. Introduction

Let $\mathbb{C}^{m \times n}$ ($\mathbb{R}^{m \times n}$) denote the set of all $m \times n$ complex (real) matrices. Throughout this paper, the notations A^T, A^*, A^{-1}, $r(A)$ and $tr(A)$ stand for the transpose, conjugate transpose, inverse, rank and trace of matrix A, respectively. The *general reciprocal* was described by E. H. Moore in 1920 [1] and independently rediscovered by R. Penrose in 1955 [2], and it is nowadays called the Moore-Penrose inverse. The Moore-Penrose inverse of a matrix $A \in \mathbb{C}^{m \times n}$ (sometimes referred to as the pseudoinverse of A) is the unique

---

matrix $A^\dagger \in \mathbb{C}^{n \times m}$ satisfying the four Penrose equations

$$AA^\dagger A = A, \qquad A^\dagger A A^\dagger = A^\dagger, \qquad \left(AA^\dagger\right)^* = AA^\dagger, \qquad \left(A^\dagger A\right)^* = A^\dagger A. \quad (1.1)$$

For more details on the Moore-Penrose inverse, see [3]. The pseudoinverse has been widely studied during the last decades from both theoretical and computational points of view. Some of the most recent works can be found, e.g., in [4, 5, 6, 7, 8, 9, 10] and in the references contained therein.

As is well-known, the Moore-Penrose inverse $A^\dagger$ can be alternatively defined as the unique matrix that gives the minimum Frobenius norm among all solutions to any of the matrix minimization problems

$$\min_{M \in \mathbb{C}^{n \times m}} \|MA - I_n\|_F, \quad (1.2)$$

$$\min_{M \in \mathbb{C}^{n \times m}} \|AM - I_m\|_F, \quad (1.3)$$

where I_n denotes the identity matrix of order n and $\|\cdot\|_F$ stands for the matrix Frobenius norm.

Also, the pseudoinverse of A can be defined via a limiting process as [11]

$$A^\dagger = \lim_{\delta \to 0} A^* \left(AA^* + \delta I_m\right)^{-1} = \lim_{\delta \to 0} \left(A^*A + \delta I_n\right)^{-1} A^*$$

and by the explicit algebraic expression [3]

$$A^\dagger = C^* \left(CC^*\right)^{-1} \left(B^*B\right)^{-1} B^*$$

based on the full rank factorization

$$A = BC, \ B \in \mathbb{C}^{m \times r}, \ C \in \mathbb{C}^{r \times n}, \ r(A) = r(B) = r(C) = r.$$

In particular, if $A \in \mathbb{C}^{m \times n}$ has full row rank (full column rank, respectively) then $A^\dagger$ is simply the unique solution to problem (1.2) (to problem (1.3), respectively), given by the respective explicit expressions [3]

$$A^\dagger = \left\{ \begin{array}{ll} (i) \ A^* \left(AA^*\right)^{-1} & \text{iff} \ r(A) = m, \\ (ii) \ \left(A^*A\right)^{-1} A^* & \text{iff} \ r(A) = n, \end{array} \right. \quad (1.4)$$

so that $A^\dagger$ is a right inverse (left inverse, respectively) of A if $r(A) = m$ ($r(A) = n$, respectively).

A natural generalization of the Moore-Penrose inverse, the so-called left or right S-Moore-Penrose inverse, arises simply by taking in Equations (1.2) or (1.3), respectively, the minimum over an arbitrary matrix subspace S of $\mathbb{C}^{n \times m}$, instead of over the whole matrix space $\mathbb{C}^{n \times m}$. This is the starting point of this paper, that has been organized as follows.

In Section 2, we define the S-Moore-Penrose inverse and we provide an explicit expression based on the singular value decomposition (SVD) of matrix A, as well as an alternative expression for the nonsingular case, in terms of the orthogonal complements with respect to the Frobenius inner product. Next, in Section 3 we apply these results to the preconditioning of linear systems based on Frobenius norm minimization and to the linear least squares problem subject to some linear restrictions. Finally, in Section 4 we present our conclusions.

2. The S-Moore-Penrose inverse

Equations (1.2) and (1.3), regarding the least squares characterization of the Moore-Penrose inverse, can be generalized as follows.

Definition 2.1. Let $A \in \mathbb{C}^{m \times n}$ and let S be a linear subspace of $\mathbb{C}^{n \times m}$. Then
(i) The left Moore-Penrose inverse of A with respect to S or, for short, the left S-Moore-Penrose inverse of A, denoted by $A_{S,l}^{\dagger}$, is the minimum Frobenius norm solution to the matrix minimization problem

$$\min_{M \in S} \|MA - I_n\|_F. \tag{2.1}$$

(ii) The right Moore-Penrose inverse of A with respect to S or, for short, the right S-Moore-Penrose inverse of A, denoted by $A_{S,r}^{\dagger}$, is the minimum Frobenius norm solution to the matrix minimization problem

$$\min_{M \in S} \|AM - I_m\|_F. \tag{2.2}$$

Remark 2.1. Note that for $S = \mathbb{C}^{n \times m}$, the left and right S-Moore-Penrose inverses become the Moore-Penrose inverse. That is, the left and right S-Moore-Penrose inverses generalizes the definition, via the matrix minimization problems (1.2) and (1.3), respectively, of the classical pseudoinverse.

3

Moreover, when $A \in \mathbb{C}^{n \times n}$ is nonsingular and $A^{-1} \in S$ then the left and right S-Moore-Penrose inverses are just the inverse. Briefly:

$$A^{\dagger}_{\mathbb{C}^{n \times m}, l} = A^{\dagger} = A^{\dagger}_{\mathbb{C}^{n \times m}, r} \quad \text{and} \quad A^{\dagger}_{S,l} = A^{-1} = A^{\dagger}_{S,r} \text{ for all } S \text{ s.t. } A^{-1} \in S.$$

Remark 2.2. In particular, if matrix A has full row rank, (full column rank, respectively) then it has right inverses, (left inverses, respectively). Thus, due to the uniqueness of the orthogonal projection of the identity matrix I_n (I_m, respectively) onto the matrix subspace $SA \subseteq \mathbb{C}^{n \times n}$ ($AS \subseteq \mathbb{C}^{m \times m}$, respectively), we conclude that, for these special cases, Definition 2.1 simply affirms that $A^{\dagger}_{S,l}$ ($A^{\dagger}_{S,r}$, respectively) is just the unique solution to problem (2.1) (to problem (2.2), respectively). The following simple example illustrates this fact.

Example 2.1. For $n = m = 2$ and

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}, \quad S = span \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix} \right\}_{\alpha \in \mathbb{C}},$$

on one hand, we have

$$\left\| A^{\dagger}_{S,l} A - I_2 \right\|_F^2 = \min_{M \in S} \| M A - I_2 \|_F^2 = \min_{\alpha \in \mathbb{C}} |\alpha - 1|^2 + |\alpha|^2 + 1,$$

so that

$$A^{\dagger}_{S,l} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}, \quad \left\| A^{\dagger}_{S,l} A - I_2 \right\|_F = \frac{3}{2}$$

and on the other hand, we have

$$\left\| A A^{\dagger}_{S,r} - I_2 \right\|_F^2 = \min_{M \in S} \| A M - I_2 \|_F^2 = \min_{\alpha \in \mathbb{C}} |\alpha - 1|^2 + |2\alpha|^2 + 1,$$

so that

$$A^{\dagger}_{S,r} = \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & 0 \end{pmatrix}, \quad \left\| A A^{\dagger}_{S,r} - I_2 \right\|_F = \frac{9}{5}.$$

Remark 2.3. Example 2.1 illustrates three basic differences between the Moore-Penrose and the S-Moore-Penrose inverses. First, we have that, in general, $A^{\dagger}_{S,l} \neq A^{\dagger}_{S,r}$. Second, neither $A^{\dagger}_{S,l}$, nor $A^{\dagger}_{S,r}$ satisfies, in general, none of the Penrose equations (1.1). Finally, although matrix A has full rank, neither $A^{\dagger}_{S,l}$ is a left inverse, nor $A^{\dagger}_{S,r}$ is a right inverse of matrix A.

Throughout this paper we address only the case of the right S-Moore-Penrose inverse $A_{S,r}^{\dagger}$, but analogous results can be obtained for the left S-Moore-Penrose inverse $A_{S,l}^{\dagger}$.

Next theorem provides us with an explicit expression for $A_{S,r}^{\dagger}$ from an SVD of matrix A.

Theorem 2.1. *Let $A \in \mathbb{C}^{m \times n}$ and let S be a linear subspace of $\mathbb{C}^{n \times m}$. Let $A = U\Sigma V^*$ be an SVD of matrix A. Then*

$$A_{S,r}^{\dagger} = V\Sigma_{V^*SU,r}^{\dagger}U^* \quad \text{and} \quad \left\|AA_{S,r}^{\dagger} - I_m\right\|_F = \left\|\Sigma\Sigma_{V^*SU,r}^{\dagger} - I_m\right\|_F. \quad (2.3)$$

Moreover,

$$\left\|A^{\dagger} - A_{S,r}^{\dagger}\right\|_F = \left\|\Sigma^{\dagger} - \Sigma_{V^*SU,r}^{\dagger}\right\|_F. \quad (2.4)$$

Proof. For all $M \in S$, we have $N = V^*MU \in V^*SU$ and $M = VNU^*$. Using the invariance property of the Frobenius norm under multiplication by unitary matrices, we get

$$\begin{aligned}
\left\|A\widetilde{M} - I_m\right\|_F &= \min_{M \in S} \|AM - I_m\|_F = \min_{N \in V^*SU} \|U\Sigma V^*VNU^* - I_m\|_F \\
&= \min_{N \in V^*SU} \|U(\Sigma N - I_m)U^*\|_F = \min_{N \in V^*SU} \|\Sigma N - I_m\|_F \\
&= \left\|\Sigma\widetilde{N} - I_m\right\|_F,
\end{aligned} \quad (2.5)$$

that is, $\widetilde{N} \in V^*SU$ is a solution to the matrix minimization problem

$$\min_{N \in V^*SU} \|\Sigma N - I_m\|_F \quad (2.6)$$

if and only if $\widetilde{M} = V\widetilde{N}U^* \in S$ is a solution to problem (2.2).

Now, let $\widetilde{N}$ be any solution to problem (2.6). Since $\Sigma_{V^*SU,r}^{\dagger}$ is the minimum Frobenius norm solution to problem (2.6), we have

$$\left\|\Sigma_{V^*SU,r}^{\dagger}\right\|_F \le \left\|\widetilde{N}\right\|_F$$

and then, using again the fact that the Frobenius norm is unitarily invariant, we have

$$\left\|V\Sigma_{V^*SU,r}^{\dagger}U^*\right\|_F \le \left\|V\widetilde{N}U^*\right\|_F \quad \text{for any solution } \widetilde{N} \text{ to problem (2.6), i.e.,}$$

$$\left\| V\Sigma^\dagger_{V^*SU,r}U^* \right\|_F \leq \left\| \widetilde{M} \right\|_F \quad \text{for any solution } \widetilde{M} \text{ to problem (2.2).}$$

This proves that the minimum Frobenius norm solution $A^\dagger_{S,r}$ to problem (2.2) is $V\Sigma^\dagger_{V^*SU,r}U^*$, and the right-hand equality of Equation (2.3) immediately follows from Equation (2.5).

Finally, Equation (2.4) immediately follows using the well-known expression $A^\dagger = V\Sigma^\dagger U^*$ and the left-hand equality of Equation (2.3), that is,

$$\left\| A^\dagger - A^\dagger_{S,r} \right\|_F = \left\| V\Sigma^\dagger U^* - V\Sigma^\dagger_{V^*SU,r}U^* \right\|_F = \left\| \Sigma^\dagger - \Sigma^\dagger_{V^*SU,r} \right\|_F. \qquad \square$$

Remark 2.4. Formula (2.3) for the right S-Moore-Penrose inverse generalizes the usual expression for the Moore-Penrose inverse in terms of an SVD of A, since (see Remark 2.1)

$$S = \mathbb{C}^{n\times m} \Rightarrow V^*SU = \mathbb{C}^{n\times m} \Rightarrow A^\dagger = A^\dagger_{S,r} = V\Sigma^\dagger_{V^*SU,r}U^* = V\Sigma^\dagger U^*.$$

For nonsingular square matrices, the next theorem provides an expression for $A^\dagger_{S,r}$ involving the orthogonal complement of subspace S (denoted, as usual, by $S^\perp$) with respect to the Frobenius inner product $\langle \cdot, \cdot \rangle_F$. This expression generalizes Equation (1.4)-(ii).

Theorem 2.2. *Let $A \in \mathbb{C}^{n\times n}$ with $r(A) = n$ and let S be a linear subspace of $\mathbb{C}^{n\times n}$. Then*

$$A^\dagger_{S,r} = (A^*A)^{-1}(A^* + Q), \quad \text{for some} \quad Q \in S^\perp. \qquad (2.7)$$

Proof. Since $AA^\dagger_{S,r}$ is the orthogonal projection of the identity matrix onto the subspace AS, we have

$$AA^\dagger_{S,r} - I_n \in (AS)^\perp. \qquad (2.8)$$

Now, for all $M \in S$ and for all $N \in S^\perp$ we have

$$\left\langle AM, (A^{-1})^* N \right\rangle_F = tr\left(AMN^*A^{-1} \right) = tr\left(A^{-1}AMN^* \right) = \langle M, N \rangle_F = 0$$

and this proves the inclusion $(A^{-1})^* S^\perp \subseteq (AS)^\perp$. Moreover, since A and $(A^{-1})^*$ are nonsingular, we have

$$\dim\left((AS)^\perp \right) = n^2 - \dim(AS) = n^2 - \dim(S) = \dim\left(S^\perp \right) = \dim\left((A^{-1})^* S^\perp \right)$$

and this proves the set equality

$$\left(AS\right)^{\perp} = \left(A^{-1}\right)^{*} S^{\perp}. \tag{2.9}$$

Finally, using Equations (2.8) and (2.9), we get

$$AA_{S,r}^{\dagger} - I_n = \left(A^{-1}\right)^{*} Q, \quad \text{for some} \quad Q \in S^{\perp}, \quad \text{i.e.,}$$

$$A_{S,r}^{\dagger} = \left(A^{*}A\right)^{-1}\left(A^{*} + Q\right), \quad \text{for some} \quad Q \in S^{\perp}. \quad \square$$

An example where Theorem 2.2 can be applied is given in the next corollary.

Corollary 2.1. *Let $A \in \mathbb{C}^{n \times n}$ with $r\left(A\right) = n$. Let $s \in \mathbb{C}^n - \{0\}$ and let $[0/s]$ be the annihilator subspace of s in $\mathbb{C}^{n \times n}$, i.e.,*

$$[0/s] = \left\{M \in \mathbb{C}^{n \times n} \mid Ms = 0\right\}.$$

Then

$$A_{[0/s],r}^{\dagger} = A^{-1}\left(I_n - \|s\|_2^{-2}\, ss^{*}\right), \tag{2.10}$$

where $\|\cdot\|_2$ stands for the usual vector Euclidean norm.

Proof. Using Equation (2.7) for $S = [0/s]$, we have

$$A_{[0/s],r}^{\dagger} = \left(A^{*}A\right)^{-1}\left(A^{*} + Q\right), \quad \text{for some} \quad Q \in [0/s]^{\perp}$$

and since the orthogonal complement of the annihilator subspace $[0/s]$ is given by [12]

$$[0/s]^{\perp} = \left\{vs^{*} \mid v \in \mathbb{C}^n\right\},$$

we get

$$A_{[0/s],r}^{\dagger} = \left(A^{*}A\right)^{-1}\left(A^{*} + vs^{*}\right), \quad \text{for some} \quad v \in \mathbb{C}^n. \tag{2.11}$$

Multiplying both sides of Equation (2.11) by vector s, we get

$$A_{[0/s],r}^{\dagger}s = \left(A^{*}A\right)^{-1}\left(A^{*} + vs^{*}\right)s$$

and, since $A_{[0/s],r}^{\dagger} \in [0/s]$, we get

$$0 = \left(A^{*} + vs^{*}\right)s \quad \Rightarrow \quad v = -\frac{1}{\|s\|_2^2}A^{*}s$$

and substituting v by its above expression in Equation (2.11), we get

$$A_{[0/s],r}^{\dagger} = \left(A^{*}A\right)^{-1}\left(A^{*} - \frac{1}{\|s\|_2^2}A^{*}ss^{*}\right) = A^{-1}\left(I_n - \|s\|_2^{-2}\, ss^{*}\right). \quad \square$$

Remark 2.5. Writing Equation (2.10) as

$$A^\dagger_{[0/s],r} = A^\dagger \left(I_n - \frac{ss^*}{s^*s} \right),$$

we have a representation of the left S-Moore-Penrose inverse of A (S being the annihilator subspace $[0/s]$ of s) as the product of the Moore-Penrose inverse of A and the elementary annihilator $I_n - \frac{ss^*}{s^*s}$ of vector s.

Regarding the full column rank case ($A \in \mathbb{C}^{m \times n}, r(A) = n$), let us mention that, in addition to formulas (2.3) and (2.7), other explicit expressions for $A^\dagger_{S,r}$ -more appropriate for computational purposes- can be given using an arbitrary basis of subspace S. The basic idea simply consists of expressing the orthogonal projection $AA^\dagger_{S,r}$ of I_m onto the subspace AS by its expansion with respect to an orthonormal basis of AS (after using the Gram-Schmidt orthonormalization procedure, if necessary). These expressions have been developed in [13] for the special case of real $n \times n$ nonsingular matrices, and they have been applied to the preconditioning of large linear systems and illustrated with some numerical experiments corresponding to real-world cases.

The next two lemmas state some spectral properties of the matrix product $AA^\dagger_{S,r}$ that will be used in the next section.

Lemma 2.1. *Let $A \in \mathbb{C}^{m \times n}$ and let S be a linear subspace of $\mathbb{C}^{n \times m}$. Then*

$$\left\| AA^\dagger_{S,r} - I_m \right\|_F^2 = m - tr\left(AA^\dagger_{S,r} \right), \tag{2.12}$$

$$\left\| AA^\dagger_{S,r} \right\|_F^2 = tr\left(AA^\dagger_{S,r} \right). \tag{2.13}$$

Proof. Using the fact that $AA^\dagger_{S,r}$ is the orthogonal projection of the identity onto the matrix subspace AS, we have

$$\left\langle AA^\dagger_{S,r} - I_m, AM \right\rangle_F = 0, \text{ for all } M \in S$$

and, in particular, for $M = A^\dagger_{S,r}$ we have

$$\left\langle AA^\dagger_{S,r} - I_m, AA^\dagger_{S,r} \right\rangle_F = 0$$

8

and then we get, on one hand,

$$
\begin{aligned}
\left\| AA_{S,r}^\dagger - I_m \right\|_F^2 &= \left\langle AA_{S,r}^\dagger - I_m, AA_{S,r}^\dagger \right\rangle_F + \left\langle I_m - AA_{S,r}^\dagger, I_m \right\rangle_F \\
&= m - tr\left(AA_{S,r}^\dagger \right)
\end{aligned}
$$

and, on the other hand,

$$
\left\| AA_{S,r}^\dagger \right\|_F^2 = \left\langle AA_{S,r}^\dagger, AA_{S,r}^\dagger \right\rangle_F = \left\langle AA_{S,r}^\dagger, I_m \right\rangle_F = tr\left(AA_{S,r}^\dagger \right). \qquad \square
$$

Lemma 2.2. *Let $A \in \mathbb{C}^{m \times n}$ and let S be a linear subspace of $\mathbb{C}^{n \times m}$. Let $\{\lambda_i\}_{i=1}^m$ and $\{\sigma_i\}_{i=1}^m$ be the sets of eigenvalues and singular values, respectively, of matrix $AA_{S,r}^\dagger$ arranged, as usual, in nonincreasing order of their modules. Then*

$$
\sum_{i=1}^m \sigma_i^2 = \sum_{i=1}^m \lambda_i \leq m. \tag{2.14}
$$

Proof. Taking into account that

$$
\left\| AA_{S,r}^\dagger \right\|_F^2 = tr\left(AA_{S,r}^\dagger \left(AA_{S,r}^\dagger \right)^* \right) = \sum_{i=1}^m \sigma_i^2, \qquad tr\left(AA_{S,r}^\dagger \right) = \sum_{i=1}^m \lambda_i
$$

and using Equations (2.12) and (2.13), the proof is concluded. $\square$

3. Applications

3.1. Applications to preconditioning

In numerical linear algebra, the convergence of the iterative Krylov methods [14, 15, 16] used for solving the linear system

$$
Ax = b, \quad A \in \mathbb{R}^{n \times n}, \quad A \text{ nonsingular}, \quad x, b \in \mathbb{R}^{n \times 1} \tag{3.1}
$$

can be improved by transforming it into the right preconditioned system

$$
AMy = b, \quad x = My. \tag{3.2}
$$

The approximate inverse preconditioning (3.2) of system (3.1) is performed in order to get a preconditioned matrix AM as close as possible to

the identity in some sense, taking into account the computational cost required for this purpose. The closeness of AM to I_n may be measured by using a suitable matrix norm like, for instance, the Frobenius norm. In this way, the optimal (in the sense of the Frobenius norm) right preconditioner of system (3.1), among all matrices M belonging to a given subspace S of $\mathbb{R}^{n \times n}$, is precisely the right S-Moore-Penrose inverse $A_{S,r}^{\dagger}$, i.e., the unique solution to the matrix minimization problem (Remark 2.2)

$$\min_{M \in S} \|AM - I_n\|_F, \quad A \in \mathbb{R}^{n \times n}, \quad r(A) = n, \quad S \subseteq \mathbb{R}^{n \times n}. \quad (3.3)$$

Alternatively, we can transform system (3.1) into the left preconditioned system $MAx = Mb$, which is now related to the left S-Moore-Penrose inverse $A_{S,l}^{\dagger}$.

In the next theorem, we apply the spectral property (2.14) to the special case of the optimal preconditioner $A_{S,r}^{\dagger}$, i.e., the solution to problem (3.3). We must assume that matrix $A_{S,r}^{\dagger}$ is nonsingular, in order to get a nonsingular coefficient matrix $AA_{S,r}^{\dagger}$ of the preconditioned system (3.2). Hence, all singular values of matrix $AA_{S,r}^{\dagger}$ will be greater than zero, and they will be denoted as

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0.$$

Theorem 3.1. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and let S be a vector subspace of $\mathbb{R}^{n \times n}$ such that the solution $A_{S,r}^{\dagger}$ to problem (3.3) is nonsingular. Then the smallest singular value σ_n of the preconditioned matrix $AA_{S,r}^{\dagger}$ lies in the interval $(0, 1]$. Moreover, we have the following lower and upper bounds on the distance $d(I_n, AS)$*

$$(1 - \sigma_n)^2 \leq \left\| AA_{S,r}^{\dagger} - I_n \right\|_F^2 \leq n \left(1 - \sigma_n^2 \right). \quad (3.4)$$

Proof. Using Lemma 2.2 for $m = n$, we get

$$n\sigma_n^2 \leq \sum_{i=1}^{n} \sigma_i^2 \leq n \implies \sigma_n^2 \leq 1 \implies \sigma_n \leq 1.$$

To prove the left-hand inequality of Equation (3.4), we use the well-known fact that singular values continuously depend on their arguments [17], i.e., for all $A, B \in \mathbb{R}^{n \times n}$ and for all $i = 1, 2, \ldots, n$

$$|\sigma_i(A) - \sigma_i(B)| \leq \|A - B\|_F \implies |\sigma_n - 1| \leq \left\| AA_{S,r}^{\dagger} - I_n \right\|_F.$$

Finally, to prove the right-hand inequality of Equation (3.4), we use Equations (2.12) and (2.13) for $m = n$

$$\left\| AA_{S,r}^{\dagger} - I_n \right\|_F^2 = n - tr\left(AA_{S,r}^{\dagger} \right) = n - \left\| AA_{S,r}^{\dagger} \right\|_F^2 = n - \sum_{i=1}^{n} \sigma_i^2 \leq n\left(1 - \sigma_n^2 \right). \square$$

Remark 3.1. Equation (3.4) states that $\left\| AA_{S,r}^{\dagger} - I \right\|_F$ decreases to 0 at the same time as the smallest singular value σ_n of the preconditioned matrix $AA_{S,r}^{\dagger}$ increases to 1. That is, the closeness of $AA_{S,r}^{\dagger}$ to the identity is determined by the closeness of σ_n to the unity. In the extremal case, $\sigma_n = 1$ iff $A_{S,r}^{\dagger} = A^{-1}$ iff $A^{-1} \in S$. A more complete version of Theorem 3.1, involving not only the smallest singular value, but also the smallest eigenvalue's modulus, of the preconditioned matrix, can be found in [18]. The proof presented here is completely different, more direct and simpler than the one given in [18].

Remark 3.2. In numerical linear algebra, the theoretical effectiveness analysis of the preconditioners M to be used, is based on the spectral properties of the preconditioned matrix AM (e.g., condition number, clustering of eigenvalues and singular values, departure from normality, etc.). For this reason, the spectral properties of matrix $AA_{S,r}^{\dagger}$, stated in Lemmas 2.1 and 2.2 and in Theorem 3.1, are especially useful when the left S-Moore-Penrose inverse is the optimal preconditioner $A_{S,r}^{\dagger}$ defined by problem (3.3). We refer the reader to [18] for more details about the theoretical effectiveness analysis of these preconditioners.

3.2. Applications to the constrained least squares problem

The S-Moore-Penrose inverse can be applied to the linear least squares problem subject to a set of homogeneous linear constraints. This is in accordance with the well-known fact that the Moore-Penrose inverse can be applied to the linear least squares problem without restrictions [19], the major application of the Moore-Penrose inverse [20]. More precisely, let us recall that, regarding the linear system $Ax = b$, $A \in \mathbb{R}^{m \times n}$, the unique solution (if A has full column rank) or the minimum norm solution (otherwise) of the least squares problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$$

is given by

$$\bar{x} = A^{\dagger}b.$$

11

Here we consider as well the linear system $Ax = b$, $A \in \mathbb{R}^{m \times n}$, but now subject to the condition $x \in T$, where T is a given vector subspace of $\mathbb{R}^n$. Denoting by $Bx = 0$ the system of Cartesian equations of subspace T, this problem is usually formulated as

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2 \ \text{ subject to } Bx = 0 \tag{3.5}$$

and we refer the reader to [11] and to the references contained therein, for more details about the linearly constrained linear least squares problem.

Next theorem provides us with a new, different approach -in terms of the S-Moore-Penrose inverse- to problem (3.5). Our theorem covers all possible situations that one can imagine for both the full rank and the rank deficient cases, e.g., $Ax = b$ consistent but $Ax = b, x \in T$ inconsistent; $Ax = b$ consistent and underdetermined while $Ax = b, x \in T$ consistent and uniquely determined; both $Ax = b$ and $Ax = b, x \in T$ inconsistent, etc.

Theorem 3.2. *Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Let T be a linear subspace of $\mathbb{R}^n$ of dimension p. Then the minimum norm solution to the constrained least squares problem*

$$\min_{x \in T} \|Ax - b\|_2 \tag{3.6}$$

is given by

$$\bar{x}_T = A^\dagger_{S_P, r} b, \tag{3.7}$$

where P is the $n \times p$ matrix whose columns are the vectors of an arbitrary basis of T and $S_P = P\mathbb{R}^{p \times m} \subseteq \mathbb{R}^{n \times m}$. In particular, if A has full column rank then $\bar{x}_T$ is the unique least squares solution to problem (3.6).

Proof. Let $B = \{v_1, \ldots, v_p\}$ be a given orthonormal basis of T and let P be the $n \times p$ matrix whose columns are $v_1, \ldots, v_p$. For a clearer exposition, we subdivide the proof into the following three parts.
(i) Writing any vector $x \in T$ as $x = P\alpha$ ($\alpha \in \mathbb{R}^p$), we get

$$\|A\widetilde{x} - b\|_2 = \min_{x \in T} \|Ax - b\|_2 = \min_{\alpha \in \mathbb{R}^p} \|(AP)\alpha - b\|_2 = \|(AP)\widetilde{\alpha} - b\|_2,$$

that is, $\widetilde{\alpha} \in \mathbb{R}^p$ is a solution to the unconstrained least squares problem

$$\min_{\alpha \in \mathbb{R}^p} \|(AP)\alpha - b\|_2 \tag{3.8}$$

12

if and only if $\widetilde{x} = P\widetilde{\alpha} \in T$ is a solution to the constrained least squares problem (3.6).

Now, let $\widetilde{\alpha}$ be any solution to problem (3.8). Since $\bar{\alpha} = (AP)^{\dagger} b$ is the minimum norm solution to problem (3.8), we have

$$\left\| (AP)^{\dagger} b \right\|_{2} \leq \|\widetilde{\alpha}\|_{2}$$

and then, using the fact that P has orthonormal columns $\left(P^{T} P = I_{p}\right)$, we have

$$\|P\alpha\|_{2}^{2} = (P\alpha)^{T} (P\alpha) = \alpha^{T}\alpha = \|\alpha\|_{2}^{2} \quad \text{for all} \ \alpha \in \mathbb{R}^{p}$$

and then

$$\left\| P (AP)^{\dagger} b \right\|_{2} \leq \|P\widetilde{\alpha}\|_{2} \ \text{for any solution} \ \widetilde{\alpha} \ \text{to problem (3.8), i.e.,}$$

$$\left\| P (AP)^{\dagger} b \right\|_{2} \leq \|\widetilde{x}\|_{2} \ \text{for any solution} \ \widetilde{x} \ \text{to problem (3.6).}$$

This proves that the minimum norm solution $\bar{x}_{T}$ to problem (3.6) can be expressed as

$$\bar{x}_{T} = P (AP)^{\dagger} b \ \text{ for any orthonormal basis} \ B \ \text{of} \ T. \tag{3.9}$$

(ii) Writing any matrix $M \in S_{P} = P\mathbb{R}^{p \times m}$ as $M = PN \ (N \in \mathbb{R}^{p \times m})$, we get

$$\begin{aligned}
\left\| A\widetilde{M} - I_{m} \right\|_{F} &= \min_{M \in S_{P}} \|AM - I_{m}\|_{F} \\
&= \min_{N \in \mathbb{R}^{p \times m}} \|(AP) N - I_{m}\|_{F} = \left\| (AP) \widetilde{N} - I_{m} \right\|_{F},
\end{aligned}$$

that is, $\widetilde{N} \in \mathbb{R}^{p \times m}$ is a solution to the unconstrained matrix minimization problem

$$\min_{N \in \mathbb{R}^{p \times m}} \|(AP) N - I_{m}\|_{F} \tag{3.10}$$

if and only if $\widetilde{M} = P\widetilde{N} \in S_{P}$ is a solution to the constrained matrix minimization problem

$$\min_{M \in S_{P} = P\mathbb{R}^{p \times m}} \|AM - I_{m}\|_{F} \tag{3.11}$$

Now, let $\widetilde{N}$ be any solution to problem (3.10). Since $(AP)^{\dagger}$ is the minimum Frobenius norm solution to problem (3.10), we have

$$\left\| (AP)^\dagger \right\|_F \leq \left\| \widetilde{N} \right\|_F$$

and then, using again the fact that P has orthonormal columns $\left(P^T P = I_p \right)$, we have

$$\| PN \|_F^2 = tr \left((PN)^T (PN) \right) = tr \left(N^T N \right) = \| N \|_F^2 \quad \text{for all} \ N \in \mathbb{R}^{p \times m}$$

and then

$$\left\| P (AP)^\dagger \right\|_F \leq \left\| P \widetilde{N} \right\|_F \quad \text{for any solution} \ \widetilde{N} \ \text{to problem (3.10), i.e.,}$$

$$\left\| P (AP)^\dagger \right\|_F \leq \left\| \widetilde{M} \right\|_F \quad \text{for any solution} \ \widetilde{M} \ \text{to problem (3.11).}$$

This proves that the minimum Frobenius norm solution $A_{S_P,r}^\dagger$ to problem (3.11) can be expressed as

$$A_{S_P,r}^\dagger = P (AP)^\dagger \quad \text{for any orthonormal basis} \ B \ \text{of} \ T. \tag{3.12}$$

(iii) Using Equations (3.9) and (3.12), we get

$$\bar{x}_T = A_{S_P,r}^\dagger b \quad \text{for any orthonormal basis} \ B \ \text{of} \ T. \tag{3.13}$$

To conclude the derivation of formula (3.7), we must proof that, although the assumption that B is an orthonormal basis of T is (in general) essential for Equations (3.9) and (3.12) to be hold, formula (3.13) for $\bar{x}_T$ is valid not only for orthonormal bases, but in fact for any basis (not necessarily orthonormal) of subspace T. In other words, the vector $A_{S_P,r}^\dagger b$ is independent of the chosen basis. Indeed, let B, B' be two arbitrary bases of T and let P, P' be the $n \times p$ matrices whose columns are the vectors of B and B', respectively. Writing matrix P' as $P' = PC$, C being a $p \times p$ nonsingular matrix, we have

$$S_{P'} = P' \mathbb{R}^{p \times m} = P \left(C \mathbb{R}^{p \times m} \right) = P \mathbb{R}^{p \times m} = S_P \ \Rightarrow \ A_{S_P,r}^\dagger b = A_{S_{P'},r}^\dagger b.$$

Finally, if in particular A has full column rank then A is left invertible and, due to the uniqueness of the orthogonal projection $A \bar{x}_T$ of vector b onto the subspace $AT \subseteq \mathbb{R}^m$, we derive that the minimum norm solution $\bar{x}_T = A_{S_P,r}^\dagger b$ to problem (3.6) is indeed the unique solution to this problem.□

Remark 3.3. In particular, for the special full rank case $r(A) = n$, the expression (3.9) for the unique least squares solution $\bar{x}_T$ of problem (3.6), is valid for any basis (not necessarily orthonormal) of T. Indeed, let B, B' be two arbitrary bases of T and let P, P' be the $n \times p$ matrices whose columns are the vectors of B and B', respectively. Writing matrix P' as $P' = PC$, C being a $p \times p$ nonsingular matrix, we have

$$P'(AP')^\dagger = PC((AP)C)^\dagger = PCC^\dagger(AP)^\dagger = PCC^{-1}(AP)^\dagger = P(AP)^\dagger,$$

where the reverse order law for the product $(AP)C$ holds because AP has full column rank, i.e.,

$$p = r(A) + r(P) - n \le r(AP) \le min(r(A), r(P)) = p$$

and C has full row rank; see, e.g., [3]. In conclusion, the minimum norm solution to problem (3.6) is given by

$$\bar{x}_T = P(AP)^\dagger b \begin{cases} \text{for any orthonormal basis of } T \text{ if } r(A) < n, \\ \text{for any basis of } T \text{ if } r(A) = n \end{cases}$$

and also, for both cases $(r(A) \le n)$, by $\bar{x}_T = A_{S_P,r}^\dagger b$ for any basis of T.

The following simple example illustrates Theorem 3.2.

Example 3.1. Let $A = (1\,1\,1) \in \mathbb{R}^{1\times3}$ and $b = (2) \in \mathbb{R}^{1\times1}$. Consider the subspace $T \equiv x_3 = 0$ of $\mathbb{R}^3$. Then $m = 1, n = 3, p = 2$ and taking the (nonorthonormal) basis of T: $B = \{(1,0,0),(1,1,0)\}$, we get

$$P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \Rightarrow S_P = P\mathbb{R}^{2\times1} = \left\{ \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \right\}_{a,b\in\mathbb{R}} \subset \mathbb{R}^{3\times1} \Rightarrow A_{S_P,r}^\dagger = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}$$

and then the minimum norm solution to problem (3.6) is given by

$$\bar{x}_T = A_{S_P,r}^\dagger b = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

4. Concluding remarks

The left and right S-Moore-Penrose inverses, $A_{S,l}^\dagger$ and $A_{S,r}^\dagger$, extend the idea of the standard inverse from nonsingular square matrices not only to rectangular and singular square matrices (as the Moore-Penrose inverse does), but also to invertible matrices, themselves, by considering a subspace S of $\mathbb{C}^{n \times m}$ not containing A^{-1}. Restricted to the right case, we have expressed $A_{S,r}^\dagger$ using either an SVD of A or (for the nonsingular case) the orthogonal complements with respect to the Frobenius inner product. $A_{S,r}^\dagger$ can be seen as an useful alternative to $A^\dagger$: It provides us with a computationally feasible approximate inverse preconditioner for large linear systems (in contrast to the unfeasible computation of $A^\dagger = A^{-1}$), and it also provides the minimum norm solution $A_{S_P,r}^\dagger b$ to the least squares problem subject to homogeneous linear constraints (in contrast to the minimum norm solution $A^\dagger b$ to the unconstrained least squares problem). The behavior of the S-Moore-Penrose inverse with respect to both theoretical results and other applications, already known for the Moore-Penrose inverse, can be considered for future researches.

References

[1] E. H. Moore, On the reciprocal of the general algebraic matrix. *Bull. Amer. Math. Soc.* 26:394-395 (1920).

[2] R. Penrose, A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.* 51:406-413 (1955).

[3] A. Ben-Israel, T. N. E. Greville, *Generalized Inverses: Theory and Applications*, 2nd. Ed. Springer-Verlag, NewYork, 2003.

[4] J. K. Baksalary, O. M. Baksalary, Particular formulae for the Moore-Penrose inverse of a columnwise partitioned matrix. *Linear Algebra Appl.* 421:16-23 (2007).

[5] T. Britz, D. D. Olesky, P. van den Driessche, The Moore-Penrose inverse of matrices with an acyclic bipartite graph. *Linear Algebra Appl.* 390:47-60 (2004).

[6] M. A. Rakha, On the Moore-Penrose generalized inverse matrix. *Appl. Math. Comput.* 158:185-200 (2004).

[7] Y. Tian, Using rank formulas to characterize equalities for Moore-Penrose inverses of matrix products. *Appl. Math. Comput.* 147:581-600 (2004).

[8] Y. Tian, Approximation and continuity of Moore-Penrose inverses of orthogonal row block matrices. *Appl. Math. Comput.* (2008), doi: 10.1016/j.amc.2008.10.027, to appear.

[9] F. Toutounian, A. Ataei, A new method for computing Moore-Penrose inverse matrices. *J. Comput. Appl. Math.* (2008), doi: 10.1016/j.cam.2008.10.008, to appear.

[10] X. Zhang, J. Cai, Y. Wei, Interval iterative methods for computing Moore-Penrose inverse. *Appl. Math. Comput.* 183:522-532 (2006).

[11] G. H. Golub, C. F. Van Loan, *Matrix Computations*, 3rd Ed. Johns Hopkins Press, Baltimore, USA, 1996.

[12] A. Griewank, A short proof of the Dennis-Schnabel theorem. *BIT* 22:252-256 (1982).

[13] G. Montero, L. González, E. Flórez, M. D. García, A. Suárez, Approximate inverse computation using Frobenius inner product. *Numer. Linear Algebra Appl.* 9:239-247 (2002).

[14] O. Axelsson, *Iterative Solution Methods*. Cambridge University Press, Cambridge, MA, 1994.

[15] A. Greenbaum, *Iterative Methods for Solving Linear Systems*. Frontiers Appl. Math., SIAM, Vol. 17, Philadelphia, PA, 1997.

[16] Y. Saad, *Iterative Methods for Sparse Linear Systems*. PWS Publishing Co., Boston, MA, 1996.

[17] R. A. Horn, C.R. Johnson, *Matrix Analysis*. Cambridge University Press, Cambridge, MA, 1985.

[18] L. González, Orthogonal projections of the identity: Spectral analysis and applications to approximate inverse preconditioning. *SIAM Rev.* 48:66-75 (2006).

[19] R. Penrose, On best approximate solutions of linear matrix equations. *Proc. Cambridge Philos. Soc.* 52:17-19 (1956).

[20] A. Ben-Israel, The Moore of the Moore-Penrose inverse. *Electron. J. Linear Algebra* 9:150-157 (2002).