Improving approximate inverses based on Frobenius norm minimization*

Luis González*, Antonio Suárez

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

Abstract

Approximate inverses, based on Frobenius norm minimization, of real nonsingular matrices are analyzed from a purely theoretical point of view. In this context, this paper provides several sufficient conditions, that assure us the possibility of improving (in the sense of the Frobenius norm) some given approximate inverses. Moreover, the optimal approximate inverses of matrix $A \in \mathbb{R}^{n \times n}$, among all matrices belonging to certain subspaces of $\mathbb{R}^{n \times n}$, are obtained. Particularly, a natural generalization of the classical normal equations of the system $Ax = b$ is given, when searching for approximate inverses $N \neq A^T$ such that AN is symmetric and $\|AN - I\|_F < \|AA^T - I\|_F$.

Keywords: Approximate inverses; Frobenius norm minimization; Trace; Spectrum; Singular values; Normal equations

1. Introduction

Let $\mathbb{R}^{n \times n}$ be the set of all $n \times n$ real matrices. In the following, $A \in \mathbb{R}^{n \times n}$ is assumed to be nonsingular, the symbols A^T, A^{-1} and $tr\,(A)$ stand for the transpose, the inverse, and the trace of matrix A, respectively, and I denotes the $n \times n$ identity matrix.

Roughly speaking, by a left (right, respectively) approximate inverse of A, we mean a matrix $N \in \mathbb{R}^{n \times n}$ such that the matrix product NA (AN, respectively) is "close to the identity" in a certain sense. This closeness may

---

*Short running title: Improving approximate inverses
*Corresponding author.
Email address: luisglez@dma.ulpgc.es (Luis González)

be measured by using an adequate matrix norm, and throughout this paper we use the Frobenius norm $\|\cdot\|_F$. Moreover, we address only the case of the right approximate inverses and, for simplicity, they will be referred to as approximate inverses (but analogous results can be obtained for the left ones). More precisely, we begin with the following definition.

Definition 1.1. *Let $A, N, N' \in \mathbb{R}^{n \times n}$. Assume that A is nonsingular. Then we say that N is better approximate inverse of A than N', or that N improves N' as approximate inverse of A if and only if*

$$\|AN - I\|_F < \|AN' - I\|_F .$$

In this context, given a linear subspace S of $\mathbb{R}^{n \times n}$, we consider the problem of obtaining the optimal approximate inverse N of matrix A in the subspace S. In accordance with Definition 1.1, throughout this paper the terms "the optimal" or "the best", mean that matrix $N \in S$ minimizes the Frobenius norm on the residual matrix $AM - I$. More precisely, "the optimal" approximate inverse N of A in S is the solution to the minimization problem

$$\min_{M \in S} \|AM - I\|_F = \|AN - I\|_F , \tag{1.1}$$

but the approximate inverse N is not necessarily optimal in any other sense of the word.

The optimization problem (1.1) is tightly connected to the numerical analysis of linear systems

$$Ax = b, \qquad A \in \mathbb{R}^{n \times n}, \quad x, b \in \mathbb{R}^n, \tag{1.2}$$

where A is a large, sparse and nonsingular matrix.

Indeed, in numerical linear algebra, the resolution of these systems is usually performed by iterative methods based on Krylov subspaces [1, 2]. In general, the convergence of such Krylov methods is not assured or may be too slow. To improve their behavior, a preconditioning matrix N is used to transform the system (1.2) into the following equivalent system,

$$ANy = b, \quad x = Ny, \tag{1.3}$$

the so-called right preconditioned system, which is performed in order to get a preconditioned matrix AN, as close as possible to the identity [3], and N is called a (right) approximate inverse preconditioner of system (1.2).

Often, the preconditioners are parametrized by prescribed sparsity patterns [4, 5], but we consider here a more general case of linear parametrization where the approximate inverse N, defined by Eq. (1.1), belongs to an arbitrary matrix subspace S of $\mathbb{R}^{n \times n}$. In [6], the authors introduce the idea to use Frobenius norm minimization for preconditioning problems. Some of the methods for constructing sparse approximate inverse preconditioners that are best approximations in the Frobenius norm, can be found, for instance, in [7, 8, 9, 10, 11, 12, 13, 14, 15] and in the references therein.

It is important to highlight here that the purpose of this paper is to provide purely theoretical results about the optimal approximate inverses $N \in S \subseteq \mathbb{R}^{n \times n}$ given by Eq. (1.1), in the theoretical context of Definition 1.1. However, our purpose is not to propose new algorithms for the numerical problem of preconditioning linear systems. Only, in a few cases, our results are related to computational strategies using special approximate inverse preconditioners based on Frobenius minimization.

The main goal of this paper is to apply several spectral properties of the matrix product AN (N being the solution to problem (1.1)) to obtain sufficient conditions for the existence of approximate inverses improving (in the sense of Definition 1.1) some given approximate inverses. By the way, we obtain the optimal approximate inverses of matrix A, among all the matrices belonging to certain linear subspaces of $\mathbb{R}^{n \times n}$.

For this purpose, in Section 2, we recall some useful expressions for matrix N, and several spectral properties of matrix AN. Next, Section 3 is devoted to establish our new results: the above mentioned sufficient conditions for improving approximate inverses, as well as the optimal approximate inverses of matrix A in certain matrix subspaces of $\mathbb{R}^{n \times n}$. Finally, conclusions are presented in Section 4.

2. Some preliminaries

Now, we present some preliminary results required to make this paper self-contained. For more details about these previous results and for their proofs, we refer the reader to [16, 17, 18].

2.1. Expressions for matrix N

Taking advantage of the well-known fact that the matrix Frobenius norm derives from an inner product, the solution N to problem (1.1) can be directly

3

obtained via orthogonal projections. Here, and in the following, orthogonality is with respect to the Frobenius inner product $\langle \cdot, \cdot \rangle_F$. More precisely, using the orthogonal projection theorem, the matrix product AN is the orthogonal projection of the identity onto the subspace AS, as stated by the following Lemma [16].

Lemma 2.1. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n \times n}$. Then, the solution to problem (1.1) is characterized by*

$$tr\left(A^t A N M^t\right) = tr\left(AM\right), \quad \forall M \in S, \tag{2.1}$$

and the minimum Frobenius norm is

$$\|AN - I\|_F^2 = n - tr(AN).$$

Moreover,

$$\|AN\|_F^2 = tr(AN).$$

Equation (2.1) implicitly gives the solution N to problem (1.1). For the purposes of this paper, it suffices to recall here the following simple explicit formula. The basic idea simply consists of expressing the orthogonal projection AN of the identity matrix onto the subspace AS by its expansion with respect to an orthonormal basis of AS [16].

Lemma 2.2. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular. Let S be a linear subspace of $\mathbb{R}^{n \times n}$ of dimension d, and $\{M_1, ..., M_d\}$ a basis of S such that $\{AM_1, ..., AM_d\}$ is an orthogonal basis of AS. Then, the solution to problem (1.1) is*

$$N = \sum_{i=1}^{d} \frac{tr\left(AM_i\right)}{\|AM_i\|_F^2} M_i, \tag{2.2}$$

and the minimum Frobenius norm is

$$\|AN - I\|_F^2 = n - \sum_{i=1}^{d} \frac{[tr\left(AM_i\right)]^2}{\|AM_i\|_F^2}. \tag{2.3}$$

Remark 2.1. In addition to formula (2.2), other explicit expressions for the optimal preconditioners N can be given using an arbitrary basis of subspace S. The idea consists of using the Gram-Schmidt orthonormalization procedure to obtain an orthonormal basis of AS, and then to apply Eq. (2.2). Such expressions have been developed for computational purposes in [16], and they have been applied to the preconditioning of large linear systems arising from real-world cases.

2.2. Spectral properties of matrix AN

Let us denote by $\{\lambda_i (AN)\}_{i=1}^n$ and $\{\sigma_i (AN)\}_{i=1}^n$ the sets of eigenvalues and singular values, respectively, of the matrix product AN arranged, as usual, in nonincreasing order of their modules, i.e.,

$$
\begin{array}{ccccccccc}
|\lambda_1 (AN)| & \geq & |\lambda_2 (AN)| & \geq & \cdots & \geq & |\lambda_n (AN)| & \geq & 0, \\
\sigma_1 (AN) & \geq & \sigma_2 (AN) & \geq & \cdots & \geq & \sigma_n (AN) & \geq & 0.
\end{array}
$$

The following proposition states some spectral properties of matrix AN. We refer the reader to [17, 18] for different proofs of these properties.

Proposition 2.1. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and let S be a linear subspace of $\mathbb{R}^{n \times n}$. Then*

$$\|AN\|_F^2 \leq n, \tag{2.4}$$

$$0 \leq tr\,(AN) \leq n, \tag{2.5}$$

$$\sigma_n (AN) \leq |\lambda_n (AN)| \leq 1 \tag{2.6}$$

and

$$
\begin{aligned}
(1 - |\lambda_n (AN)|)^2 &\leq (1 - \sigma_n (AN))^2 \leq \|AN - I\|_F^2 \\
&\leq n \left(1 - |\lambda_n (AN)|^2\right) \leq n \left(1 - \sigma_n^2 (AN)\right).
\end{aligned}
$$

Moreover, $\|AN - I\|_F$ decreases to 0 at the same time as the trace of AN increases to n, and also, at the same time as the smallest singular value $\sigma_n (AN)$ (or the smallest eigenvalue's modulus $|\lambda_n (AN)|$) of matrix AN increases to 1.

3. Improving approximate inverses of matrix A

To avoid confusion, from now on $|\lambda_n (A)|$ and $|\lambda_n (AN)|$ denote the smallest eigenvalue's modulus of matrices A and AN, respectively. Similarly, $\sigma_n (A)$ and $\sigma_n (AN)$ stand for the smallest singular value of matrices A and AN, respectively.

Now, using Eqs. (2.4), (2.5) and (2.6), we prove the following corollaries of Proposition 2.1, which provide different sufficient conditions for improving some given approximate inverses, i.e., for reducing the Frobenius norm on the residual matrix $AM - I$.

It is important to recall here that the sentences "is better approximate inverse of A than..." or "the best or the optimal approximate inverse of

A among...", used throughout this paper, have always the precise meaning stated by Definition 1.1 (and not any other meaning).

The first corollary provides us the best approximate inverse of matrix A, among all the $n \times n$ scalar matrices.

Corollary 3.1. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular.*
(i) If $tr(A) \notin [0, n]$ or if $|\lambda_n(A)| > 1$, then there exists a scalar matrix αI which is better approximate inverse of A than the identity matrix, i.e.,

$$\exists\, \alpha \in \mathbb{R},\ \alpha \ne 1 \ \text{ s.t. } \ \|A(\alpha I) - I\|_F = \|\alpha A - I\|_F < \|A - I\|_F.$$

(ii) The scalar $\overline{\alpha}$ that minimizes $\|\alpha A - I\|_F$ is

$$\overline{\alpha} = \frac{tr(A)}{\|A\|_F^2}. \tag{3.1}$$

Proof. To prove (i) consider the matrix subspace $S = span\{I\}$ of $\mathbb{R}^{n \times n}$. Suppose that $N = I$ is the solution to problem (1.1) for subspace S. Then, on one hand, using Eq. (2.5), we have

$$N = I \ \Rightarrow \ tr(AN) = tr(A) \in [0, n],$$

which contradicts the hypothesis on $tr(A)$. On the other hand, using Eq. (2.6), we have

$$N = I \ \Rightarrow \ |\lambda_n(AN)| = |\lambda_n(A)| \le 1,$$

which contradicts the hypothesis on $\lambda_n(A)$.

To prove (ii), just consider that the solution N to problem (1.1) for the subspace $S = span\{I\}$ can be obtained from Eq. (2.2), using the basis $\{I\}$ of S. That is,

$$N = \frac{tr(A)}{\|A\|_F^2} I \ \Rightarrow \ \overline{\alpha} = \frac{tr(A)}{\|A\|_F^2}. \ \square$$

The following corollary provides the optimal approximate inverse of matrix A, among all scalar multiples of A.

Corollary 3.2. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular.*
(i) If $tr(A^2) \notin [0, n]$ or if $|\lambda_n(A)| > 1$, then there exists a scalar multiple αA of A which is better approximate inverse of A than A itself, i.e.,

$$\exists\, \alpha \in \mathbb{R},\ \alpha \ne 1 \ \text{ s.t. } \ \|A(\alpha A) - I\|_F = \|\alpha A^2 - I\|_F < \|A^2 - I\|_F.$$

6

(ii) The scalar $\overline{\alpha}$ that minimizes $\left\|\alpha A^2 - I\right\|_F$ is

$$\overline{\alpha} = \frac{tr\left(A^2\right)}{\left\|A^2\right\|_F^2}.$$

We omit the proof of Corollary 3.2, since it is similar to that of Corollary 3.1 if we use matrix A instead of I.

Next corollary deals with the search of symmetric approximate inverses for symmetric matrices. Throughout this paper, the set of all $n \times n$ real symmetric matrices is denoted by $S_n(\mathbb{R})$.

Corollary 3.3. *Let $A \in S_n(\mathbb{R})$ be nonsingular.*
(i) If $\|A\|_F^2 > n$ or if $|\lambda_n(A)| > 1$, then there exists a symmetric matrix M which is better approximate inverse of A than A itself, i.e.,

$$\exists\, M \in S_n(\mathbb{R}), \quad s.t. \quad \left\|AM - I\right\|_F < \left\|A^2 - I\right\|_F.$$

(ii) The symmetric matrix N that minimizes $\left\|AM - I\right\|_F$ is given by

$$tr\left(A^2 N M\right) = tr\left(AM\right), \quad \forall M \in S_n(\mathbb{R}).$$

Proof. To prove (i) consider the matrix subspace $S = S_n(\mathbb{R})$ of all symmetric matrices in $\mathbb{R}^{n \times n}$. Suppose that $N = A$ is the solution to problem (1.1) for subspace S. Then, on one hand, using Eq. (2.5), we have

$$N = A \;\Rightarrow\; tr\left(AN\right) = tr\left(A^2\right) = tr\left(AA^T\right) = \|A\|_F^2 \in [0, n]\,,$$

which contradicts the hypothesis on $\|A\|_F^2$. On the other hand, using Eq. (2.6), we have

$$N = A \;\Rightarrow\; |\lambda_n(AN)| = \left|\lambda_n\left(A^2\right)\right| = |\lambda_n(A)|^2 \leq 1,$$

which contradicts the hypothesis on $\lambda_n(A)$.

To prove (ii), it suffices to apply Eq. (2.1) for subspace $S = S_n(\mathbb{R})$. $\square$

The following three corollaries are close related to the Frobenius norm based preconditioning of linear systems. For this reason, from now on we must assume that the solution N to problem (1.1) is nonsingular, in order to get a nonsingular preconditioned matrix AN.

First we need to set some notations. In the following, $M_{i,j}$ denotes the $n \times n$ matrix whose only nonzero term is $m_{ij} = 1$, e_i denotes the i-th canonical vector (i.e., Ae_i is the i-th column of A), and the symbols $\|\cdot\|_2$ and $\langle\cdot, \cdot\rangle_2$ stand for the Euclidean vector norm and inner product, respectively.

Remark 3.1. Since the only non-null column of matrix $AM_{i,j}$ is its j-th one, which coincides with the i-th column Ae_i of matrix A, then we have

$$tr\left(AM_{i,j}\right) = a_{ji}, \qquad \|AM_{i,j}\|_F^2 = \|Ae_i\|_2^2. \tag{3.2}$$

Moreover

$$\langle AM_{i,j}, AM_{i',j}\rangle_F = \langle Ae_i, Ae_{i'}\rangle_2 \quad \text{for all } i \neq i'$$

and

$$\langle AM_{i,j}, AM_{i',j'}\rangle_F = 0 \quad \text{for all } j \neq j',$$

so that any system of matrices $\{AM_{i,j}\}_{j=1}^n$ is orthogonal with respect to the Frobenius inner product.

Often, the diagonal matrix $diag\left(a_{11}^{-1}, a_{22}^{-1}, \ldots, a_{nn}^{-1}\right)$ is used as an approximate inverse preconditioner of system (1.2). However, this is not always the optimal choice among the diagonal approximate inverses, as the following corollary states.

Corollary 3.4. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular, such that $a_{ii} \neq 0$, for all $i = 1, 2, \ldots, n$.*
(i) If matrix A is not diagonal, then there exists an $n \times n$ diagonal matrix D which is better approximate inverse of A than $diag\left(a_{11}^{-1}, a_{22}^{-1}, \ldots, a_{nn}^{-1}\right)$, i.e., there exists $D \in \mathbb{R}^{n \times n}$ (D diagonal) such that

$$\|AD - I\|_F < \left\|A \cdot diag\left(a_{11}^{-1}, a_{22}^{-1}, \ldots, a_{nn}^{-1}\right) - I\right\|_F.$$

(ii) The diagonal matrix $\overline{D}$ that minimizes $\|AD - I\|_F$ is

$$\overline{D} = diag\left(\frac{a_{11}}{\|Ae_1\|_2^2}, \cdots, \frac{a_{nn}}{\|Ae_n\|_2^2}\right). \tag{3.3}$$

Moreover, the corresponding minimum residual Frobenius norm is given by

$$\left\|A\overline{D} - I\right\|_F^2 = n - \sum_{i=1}^n \frac{a_{ii}^2}{\|Ae_i\|_2^2}. \tag{3.4}$$

Proof. To prove (i) consider the matrix subspace S of all diagonal matrices in $\mathbb{R}^{n \times n}$. Suppose that $N = diag\left(a_{11}^{-1}, a_{22}^{-1}, \ldots, a_{nn}^{-1}\right)$ is the solution to problem (1.1) for subspace S. Then,

$$(AN)_{ij} = \begin{cases} 1, & \text{if } i = j, \\ a_{ij} \cdot a_{jj}^{-1}, & \text{if } i \neq j \end{cases}$$

8

and thus, we have

$$\|AN\|_F^2 = \sum_{i=1}^{n} (AN)_{ii}^2 + \sum_{i\neq j} (AN)_{ij}^2 = n + \sum_{i\neq j} \left(a_{ij} \cdot a_{jj}^{-1}\right)^2 > n,$$

since there exists $i \neq j$ such that $a_{ij} \neq 0$, because, by assumption, A is not diagonal. But the inequality $\|AN\|_F^2 > n$ contradicts Eq. (2.4).

To prove (ii), consider the canonical basis $\{M_{i,i}\}_{i=1}^{n}$ of subspace S. Since the basis $\{AM_{i,i}\}_{i=1}^{n}$ of AS is orthogonal (Remark 3.1), then the solution N to problem (1.1) for the subspace S can be obtained using Eqs. (2.2) and (3.2)

$$\overline{D} = \sum_{i=1}^{n} \frac{tr\,(AM_{i,i})}{\|AM_{i,i}\|_F^2} M_{i,i} = \sum_{i=1}^{n} \frac{a_{ii}}{\|Ae_i\|_2^2} M_{i,i} = diag\left(\frac{a_{11}}{\|Ae_1\|_2^2}, \cdots, \frac{a_{nn}}{\|Ae_n\|_2^2}\right).$$

Finally, to prove (3.4), it suffices to use Eqs. (2.3) and (3.2)

$$\left\|A\overline{D} - I\right\|_F^2 = n - \sum_{i=1}^{n} \frac{[tr\,(AM_{i,i})]^2}{\|AM_{i,i}\|_F^2} = n - \sum_{i=1}^{n} \frac{a_{ii}^2}{\|Ae_i\|_2^2}. \quad \square$$

Remark 3.2. The optimal diagonal (left) preconditioner $\widehat{D}$ that minimizes $\|DA - I\|_F$ was already presented in [15], and it is given by the following expression

$$\widehat{D} = diag\left(\frac{a_{11}}{\|e_1^T A\|_2^2}, \cdots, \frac{a_{nn}}{\|e_n^T A\|_2^2}\right).$$

Note the similarity/duality between the above expression for $\widehat{D}$ and Eq. (3.3) for $\overline{D}$: The only difference between both expressions is that $\widehat{D}$ involves the Euclidean norms of the rows $e_i^T A$ of matrix A, while $\overline{D}$ involves the Euclidean norms of its columns Ae_i. Anyway, our proof for $\overline{D}$ is different from that given in [15] for $\widehat{D}$.

Obviously, the optimal diagonal approximate inverse (3.3) of matrix A, has exactly one nonzero element per column. This suggests the following question: Which is the best approximate inverse of matrix A, among all the $n \times n$ matrices that have exactly one nonzero element per column? The answer is given in the next corollary, which generalizes Corollary 3.4.

Corollary 3.5. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular. The best approximate inverse of matrix A, among all $n \times n$ matrices that have exactly one nonzero element per column is*

$$N = \sum_{j=1}^{n} \frac{a_{ji_j}}{\left\| Ae_{i_j} \right\|_2^2} M_{i_j,j},$$

where, for each fixed index $j = 1, 2, \ldots, n$, its corresponding index i_j is given by

$$\frac{\left| a_{ji_j} \right|}{\left\| Ae_{i_j} \right\|_2} = \max \left\{ \frac{\left| a_{j1} \right|}{\left\| Ae_1 \right\|_2}, \frac{\left| a_{j2} \right|}{\left\| Ae_2 \right\|_2}, \ldots, \frac{\left| a_{jn} \right|}{\left\| Ae_n \right\|_2} \right\}. \tag{3.5}$$

Moreover, the corresponding minimum residual Frobenius norm is given by

$$\| AN - I \|_F^2 = n - \sum_{j=1}^{n} \frac{a_{ji_j}^2}{\left\| Ae_{i_j} \right\|_2^2}. \tag{3.6}$$

Proof. Let $j \in \{1, 2, \ldots, n\}$ be arbitrary, but fixed. The optimal approximate inverse $N_{i,j}$ of matrix A, among all the $n \times n$ matrices whose only nonzero term is placed at the i-th row, j-th column, can be obtained as the solution to problem (1.1) for subspace $S = span\{M_{i,j}\}$. That is, using Eqs. (2.2) and (3.2), we get

$$N_{i,j} = \frac{tr\left(AM_{i,j}\right)}{\left\| AM_{i,j} \right\|_F^2} M_{i,j} = \frac{a_{ji}}{\left\| Ae_i \right\|_2^2} M_{i,j},$$

for which, using Eqs. (2.3) and (3.2), we have

$$\| AN_{i,j} - I \|_F^2 = n - \frac{\left[tr\left(AM_{i,j}\right) \right]^2}{\left\| AM_{i,j} \right\|_F^2} = n - \frac{a_{ji}^2}{\left\| Ae_i \right\|_2^2}.$$

Consequently, the index $i \in \{1, 2, \ldots, n\}$ that minimizes $\| AN_{i,j} - I \|_F^2$ is the one that maximizes the quotient $a_{ji}^2 / \left\| Ae_i \right\|_2^2$, that is, the index i_j defined by Eq. (3.5).

Now, consider the subspace S of all $n \times n$ matrices that have exactly one nonzero term per column, placed at the optimal position (i_j, j), for each column $j = 1, 2, \ldots, n$. Obviously, $S = span \left\{ M_{i_j,j} \right\}_{j=1}^{n}$.

Since the basis $\left\{ AM_{i_j,j} \right\}_{j=1}^{n}$ of AS is orthogonal (Remark 3.1), then the solution N to problem (1.1) for the subspace S can be obtained using Eqs.

(2.2) and (3.2)

$$N = \sum_{j=1}^{n} \frac{tr\left(AM_{i_j,j}\right)}{\left\|AM_{i_j,j}\right\|_F^2} M_{i_j,j} = \sum_{j=1}^{n} \frac{a_{ji_j}}{\left\|Ae_{i_j}\right\|_2^2} M_{i_j,j}.$$

Finally, to prove (3.6), it suffices to use Eqs. (2.3) and (3.2)

$$\|AN - I\|_F^2 = n - \sum_{j=1}^{n} \frac{\left[tr\left(AM_{i_j,j}\right)\right]^2}{\left\|AM_{i_j,j}\right\|_F^2} = n - \sum_{j=1}^{n} \frac{a_{ji_j}^2}{\left\|Ae_{i_j}\right\|_2^2}. \quad \square$$

Let us consider the dual question to the one solved by Corollary 3.5 namely, to obtain the best approximate inverse N of matrix A, among all the $n \times n$ matrices that have exactly one nonzero element per row. For this problem, we can proceed in a similar way to the proof of Corollary 3.5. Then, we consider the subspace S' of all $n \times n$ matrices that have exactly one nonzero term per row, placed at the optimal position (i, j_i), for each row $i = 1, 2, \ldots, n$. However, in the case that $j_i = j_{i'}$, we have (see Remark 3.1)

$$\left\langle AM_{i,j_i}, AM_{i',j_{i'}} \right\rangle_F = \left\langle Ae_i, Ae_{i'} \right\rangle_2 \quad \text{for all } i \neq i',$$

so that the basis $\{AM_{i,j_i}\}_{i=1}^{n}$ of subspace AS' is not necessarily orthogonal. Thus, Lemma 2.2 can not be applied, and we need other explicit expressions for the optimal approximate inverse N (as explained in Remark 2.1).

Our last corollary, is related to the problem of transforming system (1.2) into a symmetric system. As is well-known, using the transpose of the coefficients matrix A as preconditioner is not, in general, a good strategy. For instance, the spectral condition number of the preconditioned matrix AA^T satisfies

$$\kappa_2\left(AA^T\right) = \kappa_2^2\left(A\right).$$

In this context, next corollary provides an approximate inverse of A that improves A^T (in the sense of Definition 1.1), preserving the symmetry of the preconditioned matrix.

Corollary 3.6. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular. Let S' be a matrix subspace of $\mathbb{R}^{n \times n}$ such that*

$$A^T \in S' \subseteq \left\{ M \in \mathbb{R}^{n \times n} \mid (AM)^T = AM \right\}. \tag{3.7}$$

11

(i) If $\|A\|_F^2 > n$ or if $\sigma_n(A) > 1$, then there exists a matrix N which is better approximate inverse of A than A^T, and such that AN becomes symmetric. More precisely,

$$\exists\, N \in S', \ \ N \neq A^T, \ \ \text{s.t.} \ \ \|AN - I\|_F < \|AA^T - I\|_F. \tag{3.8}$$

(ii) The matrix N that minimizes $\|AN - I\|_F$ over subspace S' is given by

$$tr\,(ANAM) = tr\,(AM)\,, \ \ \ \forall M \in S'. \tag{3.9}$$

Proof. To prove (i) consider the matrix subspace $S = S'$. Suppose that $N = A^T$ is the solution to problem (1.1) for subspace S. Then, on one hand, using Eq. (2.5), we have

$$N = A^T \ \Rightarrow \ tr\,(AN) = tr\,\left(AA^T\right) = \|A\|_F^2 \in [0, n]\,,$$

which contradicts the hypothesis on $\|A\|_F^2$. On the other hand, using Eq. (2.6), we have

$$N = A^T \ \Rightarrow \ |\lambda_n\,(AN)| = \left|\lambda_n\,\left(AA^T\right)\right| = \sigma_n^2\,(A) \leq 1,$$

which contradicts the hypothesis on $\sigma_n(A)$.

To prove (ii), it suffices to apply Eq. (2.1) for subspace $S = S'$ using the fact that AM is symmetric for all $M \in S'$. $\square$

From a theoretical point of view, Corollary 3.6 leads to a natural generalization of the normal equations of system (1.2).

For this purpose, by defining $M = PA^T$, where P is a symmetric matrix, AM is symmetric and the right preconditioned system

$$AMy = b, \ \ \ x = My$$

can be rewritten as

$$APA^Ty = b, \ \ \ x = PA^Ty \ \ \left(P^T = P\right), \tag{3.10}$$

a natural generalization of the classical normal equations $(P = I)$ related to system (1.2)

$$AA^Ty = b, \ \ \ x = A^Ty. \tag{3.11}$$

Corollary 3.6 has stated that whenever $\|A\|_F^2 > n$ (or whenever $\sigma_n(A) > 1$), we can assure that there exists a matrix M such that AM is symmetric

12

and, at the same time, M improves A^T as approximate inverse of A. But, due to the definition of M and to the fact that AM is symmetric if and only if P is symmetric, this is equivalent to saying that there exists a symmetric matrix P for which the generalized normal equations (3.10) improve (in the sense of the Frobenius norm minimization) the classical normal equations (3.11).

Moreover, from a practical point of view, note that Eq. (3.7) can be rewritten as

$$A^T \in S' \subseteq \left\{ PA^T \mid P^T = P \right\} \tag{3.12}$$

and this enables one to obtain, for several adequate subspaces S', an explicit expression for the optimal preconditioners N satisfying Eq. (3.8) –note that Eq. (3.9) only defines N implicitly. Two examples that illustrate this procedure are given below.

Example 3.1. Let $\{P_1, \ldots, P_d\}$ be a linearly independent set of $n \times n$ real symmetric matrices, containing the identity. The subspace

$$S' = span \left\{ P_1 A^T, \ldots, P_d A^T \right\}$$

clearly satisfies Eq. (3.12) (or, equivalently, Eq. (3.7)).

Hence, we can explicitly obtain the solution N to problem (1.2) for subspace S' from its basis $\left\{ P_1 A^T, \ldots, P_d A^T \right\}$. Two cases are possible:
(i) If $\left\{ A P_1 A^T, \ldots, A P_d A^T \right\}$ is an orthogonal basis of subspace AS' then we just use Lemma 2.2 for obtaining N.
(ii) If $\left\{ A P_1 A^T, \ldots, A P_d A^T \right\}$ is not an orthogonal basis of subspace AS' then we use the formulas developed in [16] for obtaining N (see Remark 2.1).

Example 3.2. As a special case of Example 3.1, consider the canonical basis of the subspace $S_n(\mathbb{R})$ of $n \times n$ real symmetric matrices

$$\{M_{i,i}\}_{i=1}^{n} \cup \{M_{i,j} + M_{j,i}\}_{i<j},$$

where recall that $M_{i,j}$ denotes the $n \times n$ matrix whose only nonzero term is $m_{ij} = 1$. Then we have the following basis of the subspace of all matrices PA^T such that P is symmetric

$$\left\{ PA^T \mid P^T = P \right\} = span \left(\left\{ M_{i,i} A^T \right\}_{i=1}^{n} \cup \left\{ (M_{i,j} + M_{j,i}) A^T \right\}_{i<j} \right) \tag{3.13}$$

and then one can consider the subspace $S' = span \left\{ M_{i,i} A^T \right\}_{i=1}^{n}$, and iteratively augmenting its bases in order to obtain different preconditioners N satisfying Eq. (3.8).

Another practical interest of Corollary 3.6 stands in the possibility of using the conjugate gradient method for solving the right preconditioned system $ANy = b$, $x = Ny$ when matrix AN is, besides symmetric, positive definite.

Remark 3.3. Due to the well-known inequality $\sigma_n(A) \leq |\lambda_n(A)|$ for every nonsingular matrix A [19], Corollaries 3.1-3.3 are also valid if we replace the assumption $|\lambda_n(A)| > 1$ by the more restrictive one $\sigma_n(A) > 1$.

Remark 3.4. Let $A \in \mathbb{R}^{n \times n}$ such that $\sigma_n(A) > 1$. Then, Corollaries 3.1-3.3 and 3.6 are also valid for all matrices obtained by any permutations of the rows and/or columns of matrix A. Indeed, let QAP be any matrix obtained by any permutations of the rows and/or columns of A. Since the permutation matrices are orthogonal, and since the spectral matrix norm $\|\cdot\|_2$ is unitarily invariant [20], then we have

$$\sigma_n(QAP) = \left\|(QAP)^{-1}\right\|_2^{-1} = \left\|P^T A^{-1} Q^T\right\|_2^{-1} = \left\|A^{-1}\right\|_2^{-1} = \sigma_n(A)$$

and thus

$$\sigma_n(A) > 1 \quad \Rightarrow \quad \sigma_n(QAP) > 1 \quad \Rightarrow \quad |\lambda_n(QAP)| > 1,$$

so that matrix QAP also satisfies the spectral hypothesis of those corollaries.

4. Concluding remarks

As starting point, we have considered the problem of obtaining the optimal (in the sense of the Frobenius norm) approximate inverse N of a non-singular matrix $A \in \mathbb{R}^{n \times n}$ over an arbitrary matrix subspace $S \subseteq \mathbb{R}^{n \times n}$, i.e., the solution to the minimization problem (1.1). Using the spectral properties of the matrix product AN (e.g., $tr(AN) \in [0, n]$, $|\lambda_n(AN)| \leq 1$, etc.), we have established different sufficient conditions that assure us the possibility of improving some given approximate inverses (i.e., reducing the Frobenius norm on matrix $AM - I$). Moreover, we have explicitly obtained the optimal approximate inverse N, for different matrix subspaces S of $\mathbb{R}^{n \times n}$. The most relevant cases in this sense correspond to the optimal diagonal approximate inverse and to an approximate inverse $N \neq A^T$ such that AN is symmetric and $\|AN - I\|_F < \|AA^T - I\|_F$. The latter led us to generalize the classical normal equations of system $Ax = b$.

For future research, one can consider the problem of obtaining the optimal approximate inverse of matrix A for different matrix subspaces $S \subset \mathbb{R}^{n \times n}$, not analyzed in this paper. This can be performed, not only in the same way, but also in a different way to the one used here like, for instance, using the orthogonal complement of S with respect to the Frobenius inner product. An additional line for future research, from a numerical/computational point of view, could consist of applying the results presented in this theoretical paper to the preconditioning problem of large linear systems.

Acknowledgments

References

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

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

[3] M. Benzi, Preconditioning techniques for large linear systems: a survey, J. Comput. Phys. 182 (2002) 418-477.

[4] M. Benzi, M. Tůma, A comparative study of sparse approximate inverse preconditioners, Appl. Numer. Math. 30 (1999) 305-340.

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

[6] M.W. Benson, P.O. Frederickson, Iterative solution of large sparse linear systems arising in certain multidimensional approximation problems, Util. Math. 22 (1982) 127-140.

[7] B. Carpentieri, I.S. Duff, L. Giraud, Sparse pattern selection strategies for robust Frobenius-norm minimization preconditioners in electromagnetism, Numer. Linear Algebra Appl. 7 (2000) 667-685.

[8] E. Chow, Y. Saad, Approximate inverse preconditioners via sparse-sparse iterations, SIAM J. Sci. Comput. 19 (1998) 995-1023.

[9] N.I.M. Gould, J.A. Scott, Sparse approximate-inverse preconditioners using norm-minimization techniques, SIAM J. Sci. Comput. 19 (1998) 605-625.

[10] M. Grote, T. Huckle, Parallel preconditioning with sparse approximate inverses, SIAM J. Sci. Comput. 18 (1997) 838-853.

[11] T. Huckle, Approximate sparsity patterns for the inverse of a matrix and preconditioning, Appl. Numer. Math. 30 (1999) 291303.

[12] T. Huckle, A. Kallischko, Frobenius norm minimization and probing for preconditioning, Int. J. Comput. Math. 84 (2007) 12251248.

[13] T. Huckle, A. Kallischko, A. Roy, M. Sedlacek, T. Weinzierl, An efficient parallel implementation of the MSPAI preconditioner, Parallel Comput. 36 (2010) 273-284.

[14] H.L. Ong, Fast approximate solution of large-scale sparse linear systems, J. Comput. Appl. Math. 10 (1984) 45-54.

[15] P. Tarazaga, D. Cuellar, Preconditioners generated by minimizing norms, Comput. Math. Appl. 57 (2009) 1305-1312.

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

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

[18] A. Suárez, L. González, A generalization of the Moore-Penrose inverse related to matrix subspaces of $\mathbb{C}^{n \times m}$, Appl. Math. Comput. 216 (2010) 514-522.

[19] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, MA, 1991.

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