

# Geometrical and Spectral Properties of the Orthogonal Projections of the Identity

Luis González, Antonio Suárez, and Dolores García

*Research Institute SIANI, Department of Mathematics,  
University of Las Palmas de Gran Canaria, Campus de Tafira,  
Las Palmas de Gran Canaria 35017, Canary Islands, Spain*

Correspondence should be addressed to Luis González, luisglez@dma.ulpgc.es

In this paper, we analyze the best approximation  $AN$  (in the Frobenius sense) to the identity matrix in an arbitrary matrix subspace  $AS$  ( $A \in \mathbb{R}^{n \times n}$  nonsingular,  $S$  being any fixed subspace of  $\mathbb{R}^{n \times n}$ ). Some new geometrical and spectral properties of the orthogonal projection  $AN$  are derived. In particular, new inequalities for the trace and for the eigenvalues of matrix  $AN$  are presented for the special case that  $AN$  is symmetric and positive definite.

## 1 Introduction

The set of all  $n \times n$  real matrices is denoted by  $\mathbb{R}^{n \times n}$ , and  $I$  denotes the identity matrix of order  $n$ . In the following,  $A^T$  and  $tr(A)$  denote, as usual, the transpose and the trace of matrix  $A \in \mathbb{R}^{n \times n}$ . The notations  $\langle \cdot, \cdot \rangle_F$  and  $\|\cdot\|_F$  stand for the Frobenius inner product and matrix norm, defined on the matrix space  $\mathbb{R}^{n \times n}$ . Throughout this paper, the terms orthogonality, angle and cosine will be used in the sense of the Frobenius inner product.

Our starting point is the linear system

$$Ax = b, \quad A \in \mathbb{R}^{n \times n}, \quad x, b \in \mathbb{R}^n, \quad (1.1)$$

where  $A$  is a large, nonsingular and sparse matrix. The resolution of this system is usually performed by iterative methods based on Krylov subspaces (see, e.g., [2, 15]). The coefficient matrix  $A$  of system (1.1) is often extremely ill-conditioned and highly indefinite, so that in this case, Krylov subspace methods are not competitive without a good preconditioner (see, e.g., [15, 19]). Then to improve the convergence of these Krylov methods, system (1.1) can be preconditioned with an adequate nonsingular preconditioning matrix  $N$ , transforming it into any of the equivalent systems

$$NAx = Nb, \quad (1.2)$$

$$ANy = b, \quad x = Ny, \quad (1.3)$$

the so-called left and right preconditioned systems, respectively. In this paper, we address only the case of the right-hand side preconditioned matrices  $AN$ , but analogous results can be obtained for the left-hand side preconditioned matrices  $NA$ .

The preconditioning of system (1.1) is often performed in order to get a preconditioned matrix  $AN$  as close as possible to the identity in some sense, and the preconditioner  $N$  is called an approximate inverse of  $A$ . The closeness of  $AN$  to  $I$  may be measured by using a suitable matrix norm like, for instance, the Frobenius norm [3]. In this way, the problem of obtaining the best preconditioner  $N$  (with respect to the Frobenius norm) of system (1.1) in an arbitrary subspace  $S$  of  $\mathbb{R}^{n \times n}$  is equivalent to the minimization problem; see, e.g., [14]

$$\min_{M \in S} \|AM - I\|_F = \|AN - I\|_F. \quad (1.4)$$

The solution  $N$  to problem (1.4) will be referred to as the “optimal” or the “best” approximate inverse of matrix  $A$  in the subspace  $S$ . Since matrix  $AN$  is the best approximation to the identity in subspace  $AS$ , it will be also referred to as the orthogonal projection of the identity matrix onto the subspace  $AS$ . Although many of the results presented in this paper are also valid for the case that matrix  $N$  is singular, from now on, we assume that the optimal approximate inverse  $N$  (and thus also the orthogonal projection  $AN$ ) is a nonsingular matrix. The solution  $N$  to problem (1.4) has been studied as a natural generalization of the classical Moore-Penrose inverse in [16], where it has been referred to as the  $S$ -Moore-Penrose inverse of matrix  $A$ .

The main goal of this paper is to derive new geometrical and spectral properties of the best approximations  $AN$  (in the sense of formula (1.4)) to the identity matrix. Such properties could be used to analyze the quality and theoretical effectiveness of the optimal approximate inverse  $N$  as preconditioner of system (1.1). However, it is important to highlight that the purpose of this paper is purely theoretical and we are not looking for immediate numerical or computational approaches (although our theoretical results could be potentially applied to the preconditioning problem). In particular, the term “optimal (or best) approximate inverse” is used in the sense of formula (1.4), and not in any other sense of this expression.

Among the many different works dealing with practical algorithms that can be used to compute approximate inverses, we refer the reader to e.g., [3, 6, 7, 10] and to the references therein. In [3], the author presents an exhaustive survey of preconditioning techniques and, in particular, describes several algorithms for computing sparse approximate inverses based on Frobenius norm minimization like, for instance, the well-known SPAI and FSAI algorithms. Also focused on approximate inverses based on minimizing  $\|AM - I\|_F$ , in [6], an iterative descent-type method is used to approximate each column of the inverse, and the iteration is done with “sparse-matrix by sparse-vector” operations. When the system matrix is expressed in block-partitioned form, some preconditioning options are explored in [7]. In [10], the idea of “target” matrix is introduced, in the context of sparse approximate inverse preconditioners, and the generalized Frobenius norms  $\|B\|_{F,H}^2 = \text{tr}(BHB^T)$  ( $H$  symmetric positive definite) are used, for minimization purposes, as an alternative to the classical Frobenius norm.

The last results of our work are devoted to the special case that matrix  $AN$  is symmetric and positive definite. In this sense, let us recall that the cone of symmetric and positive definite matrices has a rich geometrical structure and, in this context, the angle that any symmetric and positive definite matrix forms with the identity plays a very important role [1]. In that paper, the authors extend this geometrical point of view, and analyze the geometrical structure of the subspace of symmetric matrices of order  $n$ , including the location of all orthogonal matrices, not only the identity matrix.

This paper has been organized as follows. In Section 2, we present some preliminary results required to make the paper self-contained. Sections 3 and 4 are devoted to obtain new geometrical and spectral relations, respectively, for the orthogonal projections  $AN$  of the identity matrix. Finally, Section 5 closes the paper with its main conclusions.

## 2 Some Preliminaries

Now, we present some preliminary results concerning the orthogonal projection  $AN$  of the identity onto the matrix subspace  $AS \subset \mathbb{R}^{n \times n}$ . For more details about these results and for their proofs, we refer the reader to [14, 16, 8].

Taking advantage of the prehilbertian character of the matrix Frobenius norm, the solution  $N$  to problem (1.4) can be obtained using the orthogonal projection theorem. More precisely, the matrix product  $AN$  is the orthogonal projection of the identity onto the subspace  $AS$ , and it satisfies the conditions stated by the following lemma; see [14, 8].

**Lemma 2.1** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then*

$$0 \leq \|AN\|_F^2 = \text{tr}(AN) \leq n. \quad (2.1)$$

$$0 \leq \|AN - I\|_F^2 = n - \text{tr}(AN) \leq n. \quad (2.2)$$

An explicit formula for matrix  $N$  can be obtained by expressing the orthogonal projection  $AN$  of the identity matrix onto the subspace  $AS$  by its expansion with respect to an orthonormal basis of  $AS$  [14]. This is the idea of the following lemma.

**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  $\{M_1, \dots, M_d\}$  a basis of  $S$  such that  $\{AM_1, \dots, AM_d\}$  is an orthogonal basis of  $AS$ . Then, the solution  $N$  to problem (1.4) is*

$$N = \sum_{i=1}^d \frac{\text{tr}(AM_i)}{\|AM_i\|_F^2} M_i, \quad (2.3)$$

and the minimum (residual) Frobenius norm is

$$\|AN - I\|_F^2 = n - \sum_{i=1}^d \frac{[\text{tr}(AM_i)]^2}{\|AM_i\|_F^2}. \quad (2.4)$$

Let us mention two possible options, both taken from [14], for choosing in practice the subspace  $S$  and its corresponding basis  $\{M_i\}_{i=1}^d$ . The first example consists of considering the subspace  $S$  of  $n \times n$  matrices with a prescribed sparsity pattern, i.e.,

$$S = \{M \in \mathbb{R}^{n \times n} : m_{ij} = 0 \text{ for all } (i, j) \notin K\}, \quad K \subset \{1, 2, \dots, n\} \times \{1, 2, \dots, n\}.$$

Then, denoting by  $M_{i,j}$  the  $n \times n$  matrix whose only nonzero entry is  $m_{ij} = 1$ , a basis of subspace  $S$  is clearly  $\{M_{i,j} : (i, j) \in K\}$ , and then  $\{AM_{i,j} : (i, j) \in K\}$  will be a basis of subspace  $AS$  (since we have assumed that matrix  $A$  is nonsingular). In general, this basis of  $AS$  is not orthogonal, so that we only need to use the Gram-Schmidt procedure to obtain an orthogonal basis of  $AS$ , in order to apply the orthogonal expansion (2.3).

For the second example, consider a linearly independent set of  $n \times n$  real symmetric matrices  $\{P_1, \dots, P_d\}$  and the corresponding subspace

$$S' = \text{span} \{P_1 A^T, \dots, P_d A^T\},$$

which clearly satisfies

$$S' \subseteq \{M = P A^T : P \in \mathbb{R}^{n \times n} \quad P^T = P\} = \{M \in \mathbb{R}^{n \times n} : (AM)^T = AM\}.$$

Hence, we can explicitly obtain the solution  $N$  to problem (1.4) for subspace  $S'$ , from its basis  $\{P_1A^T, \dots, P_dA^T\}$ , as follows. If  $\{AP_1A^T, \dots, AP_dA^T\}$  is an orthogonal basis of subspace  $AS'$  then we just use the orthogonal expansion (2.3) for obtaining  $N$ . Otherwise, we use again the Gram-Schmidt procedure to obtain an orthogonal basis of subspace  $AS'$ , and then we apply formula (2.3). The interest of this second example stands in the possibility of using the conjugate gradient method for solving the preconditioned linear system, when the symmetric matrix  $AN$  is positive definite. For a more detailed exposition of the computational aspects related to these two examples, we refer the reader to [14].

Now, we present some spectral properties of the orthogonal projection  $AN$ . From now on, we denote by  $\{\lambda_i\}_{i=1}^n$  and  $\{\sigma_i\}_{i=1}^n$  the sets of eigenvalues and singular values, respectively, of matrix  $AN$  arranged, as usual, in nonincreasing order, i.e.,

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| > 0, \quad (2.5)$$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0. \quad (2.6)$$

The following lemma [8] provides some inequalities involving the eigenvalues and singular values of the preconditioned matrix  $AN$ .

**Lemma 2.3** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Then*

$$\sum_{i=1}^n \lambda_i^2 \leq \sum_{i=1}^n |\lambda_i|^2 \leq \sum_{i=1}^n \sigma_i^2 = \|AN\|_F^2 = \text{tr}(AN) = \sum_{i=1}^n \lambda_i \leq \sum_{i=1}^n |\lambda_i| \leq \sum_{i=1}^n \sigma_i. \quad (2.7)$$

The following fact [8] is a direct consequence of Lemma 2.3.

**Lemma 2.4** *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 smallest singular value and the smallest eigenvalue's modulus of the orthogonal projection  $AN$  of the identity onto the subspace  $AS$  are never greater than 1. That is,*

$$0 < \sigma_n \leq |\lambda_n| \leq 1. \quad (2.8)$$

The following theorem [8] establishes a tight connection between the closeness of matrix  $AN$  to the identity matrix and the closeness of  $\sigma_n$  ( $|\lambda_n|$ ) to the unity.

**Theorem 2.1** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then*

$$(1 - |\lambda_n|)^2 \leq (1 - \sigma_n)^2 \leq \|AN - I\|_F^2 \leq n \left(1 - |\lambda_n|^2\right) \leq n \left(1 - \sigma_n^2\right). \quad (2.9)$$

**Remark 2.1** Theorem 2.1 states that the closer the smallest singular value  $\sigma_n$  of matrix  $AN$  is to the unity, the closer matrix  $AN$  will be to the identity, i.e., the smaller  $\|AN - I\|_F$  will be; and conversely. The same happens with the smallest eigenvalue's modulus  $|\lambda_n|$  of matrix  $AN$ . In other words, we get a good approximate inverse  $N$  of  $A$  when  $\sigma_n$  ( $|\lambda_n|$ ) is sufficiently close to 1.

To finish this section, let us mention that, recently, lower and upper bounds on the normalized Frobenius condition number of the orthogonal projection  $AN$  of the identity onto the subspace  $AS$ , have been derived in [17]. In addition, that work proposes a natural generalization (related to an arbitrary matrix subspace  $S$  of  $\mathbb{R}^{n \times n}$ ) of the normalized Frobenius condition number of the nonsingular matrix  $A$ .

### 3 Geometrical Properties

In this section, we present some new geometrical properties for matrix  $AN$ ,  $N$  being the optimal approximate inverse of matrix  $A$ , defined by (1.4). Our first lemma states some basic properties involving the cosine of the angle between matrix  $AN$  and the identity, i.e.,

$$\cos(AN, I) = \frac{\langle AN, I \rangle_F}{\|AN\|_F \|I\|_F} = \frac{\text{tr}(AN)}{\|AN\|_F \sqrt{n}}. \quad (3.1)$$

**Lemma 3.1** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then*

$$\cos(AN, I) = \frac{\text{tr}(AN)}{\|AN\|_F \sqrt{n}} = \frac{\|AN\|_F}{\sqrt{n}} = \frac{\sqrt{\text{tr}(AN)}}{\sqrt{n}}. \quad (3.2)$$

$$0 \leq \cos(AN, I) \leq 1. \quad (3.3)$$

$$\|AN - I\|_F^2 = n(1 - \cos^2(AN, I)). \quad (3.4)$$

*Proof.* First, using (3.1) and (2.1) we immediately obtain (3.2). As a direct consequence of (3.2) we derive that  $\cos(AN, I)$  is always nonnegative. Finally, using (2.2) and (3.2), we get

$$\|AN - I\|_F^2 = n - \text{tr}(AN) = n(1 - \cos^2(AN, I)) \quad (3.5)$$

and the proof is concluded.  $\square$

**Remark 3.1** In [5], the authors consider an arbitrary approximate inverse  $Q$  of matrix  $A$  and derive the following equality

$$\|AQ - I\|_F^2 = (\|AQ\|_F - \|I\|_F)^2 + 2(1 - \cos(AQ, I))\|AQ\|_F \|I\|_F, \quad (3.6)$$

that is, the typical decomposition (valid in any inner product space) of the strong convergence into the convergence of the norms  $(\|AQ\|_F - \|I\|_F)^2$  and the weak convergence  $(1 - \cos(AQ, I))\|AQ\|_F \|I\|_F$ . Note that for the special case that  $Q$  is the optimal approximate inverse  $N$  defined by (1.4), formula (3.4) has stated that the strong convergence is reduced just to the weak convergence and, indeed, just to the cosine  $\cos(AN, I)$ .

**Remark 3.2** More precisely, formula (3.4) states that the closer  $\cos(AN, I)$  is to the unity (i.e., the smaller the angle  $\angle(AN, I)$  is), the smaller  $\|AN - I\|_F$  will be; and conversely. This gives us a new measure of the quality (in the Frobenius sense) of the approximate inverse  $N$  of matrix  $A$ , by comparing the minimum residual norm  $\|AN - I\|_F$  with the cosine of the angle between  $AN$  and the identity, instead of with  $\text{tr}(AN)$ ,  $\|AN\|_F$  (Lemma 2.1) or with  $\sigma_n, |\lambda_n|$  (Theorem 2.1). So, for a fixed nonsingular matrix  $A \in \mathbb{R}^{n \times n}$  and for different subspaces  $S \subset \mathbb{R}^{n \times n}$ , we have

$$\begin{aligned} \text{tr}(AN) \nearrow n &\Leftrightarrow \|AN\|_F \nearrow \sqrt{n} \Leftrightarrow \sigma_n \nearrow 1 \Leftrightarrow |\lambda_n| \nearrow 1 \\ &\Leftrightarrow \cos(AN, I) \nearrow 1 \Leftrightarrow \|AN - I\|_F \searrow 0. \end{aligned} \quad (3.7)$$

Obviously, the optimal theoretical situation corresponds to the case

$$\begin{aligned} \text{tr}(AN) = n &\Leftrightarrow \|AN\|_F = \sqrt{n} \Leftrightarrow \sigma_n = 1 \Leftrightarrow \lambda_n = 1 \\ &\Leftrightarrow \cos(AN, I) = 1 \Leftrightarrow \|AN - I\|_F = 0 \\ &\Leftrightarrow N = A^{-1} \Leftrightarrow A^{-1} \in S. \end{aligned} \quad (3.8)$$

**Remark 3.3** Note that the ratio between  $\cos(AN, I)$  and  $\cos(A, I)$  is independent of the order  $n$  of matrix  $A$ . Indeed, assuming that  $\text{tr}(A) \neq 0$  and using (3.2), we immediately obtain

$$\frac{\cos(AN, I)}{\cos(A, I)} = \frac{\|AN\|_F : \sqrt{n}}{\text{tr}(A) : \|A\|_F \sqrt{n}} = \|AN\|_F \frac{\|A\|_F}{\text{tr}(A)}. \quad (3.9)$$

The following lemma compares the trace and the Frobenius norm of the orthogonal projection  $AN$ .

**Lemma 3.2** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then*

$$\text{tr}(AN) \leq \|AN\|_F \Leftrightarrow \|AN\|_F \leq 1 \Leftrightarrow \text{tr}(AN) \leq 1 \Leftrightarrow \cos(AN, I) \leq \frac{1}{\sqrt{n}}, \quad (3.10)$$

$$\text{tr}(AN) \geq \|AN\|_F \Leftrightarrow \|AN\|_F \geq 1 \Leftrightarrow \text{tr}(AN) \geq 1 \Leftrightarrow \cos(AN, I) \geq \frac{1}{\sqrt{n}}. \quad (3.11)$$

*Proof.* Using (2.1), we immediately obtain the four left-most equivalences. Using (3.2), we immediately obtain the two right-most equivalences.  $\square$

The next lemma provides us with a relationship between the Frobenius norms of the inverses of matrices  $A$  and its best approximate inverse  $N$  in subspace  $S$ .

**Lemma 3.3** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then*

$$\|A^{-1}\|_F \|N^{-1}\|_F \geq 1. \quad (3.12)$$

*Proof.* Using (2.1), we get

$$\|AN\|_F \leq \sqrt{n} = \|I\|_F = \|(AN)(AN)^{-1}\|_F \leq \|AN\|_F \|(AN)^{-1}\|_F \Rightarrow \|(AN)^{-1}\|_F \geq 1 \quad (3.13)$$

and hence

$$\|A^{-1}\|_F \|N^{-1}\|_F \geq \|N^{-1}A^{-1}\|_F = \|(AN)^{-1}\|_F \geq 1 \quad (3.14)$$

and the proof is concluded.  $\square$

The following lemma compares the minimum residual norm  $\|AN - I\|_F$  with the distance (with respect to the Frobenius norm)  $\|A^{-1} - N\|_F$  between the inverse of  $A$  and the optimal approximate inverse  $N$  of  $A$  in any subspace  $S \subset \mathbb{R}^{n \times n}$ . First, note that for any two matrices  $A, B \in \mathbb{R}^{n \times n}$  ( $A$  nonsingular), from the submultiplicative property of the Frobenius norm, we immediately get

$$\|AB - I\|_F^2 = \|A(B - A^{-1})\|_F^2 \leq \|A\|_F^2 \|B - A^{-1}\|_F^2. \quad (3.15)$$

However, for the special case that  $B = N$  (the solution to problem (1.4)), we also get the following inequality.

**Lemma 3.4** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then*

$$\|AN - I\|_F^2 \leq \|A\|_F^2 \|N - A^{-1}\|_F^2. \quad (3.16)$$

*Proof.* Using the Cauchy-Schwarz inequality and (2.2), we get

$$\begin{aligned}
|\langle A^{-1} - N, A^T \rangle_F| &\leq \|A^{-1} - N\|_F \|A^T\|_F \Rightarrow \\
|\operatorname{tr}((A^{-1} - N)A)| &\leq \|A^{-1} - N\|_F \|A^T\|_F \Rightarrow \\
|\operatorname{tr}(A(A^{-1} - N))| &\leq \|N - A^{-1}\|_F \|A\|_F \Rightarrow \\
|\operatorname{tr}(I - AN)| &\leq \|N - A^{-1}\|_F \|A\|_F \Rightarrow \\
n - \operatorname{tr}(AN) &\leq \|N - A^{-1}\|_F \|A\|_F \Rightarrow \\
\|AN - I\|_F^2 &\leq \|A\|_F \|N - A^{-1}\|_F. \quad \square
\end{aligned} \tag{3.17}$$

The following extension of the Cauchy-Schwarz inequality, in a real or complex inner product space  $(H, \langle \cdot, \cdot \rangle)$ , was obtained by Buzano [4]. For all  $a, x, b \in H$ , we have

$$|\langle a, x \rangle \cdot \langle x, b \rangle| \leq \frac{1}{2} (\|a\| \|b\| + |\langle a, b \rangle|) \|x\|^2. \tag{3.18}$$

The next lemma provides us with lower and upper bounds on the inner product  $\langle AN, B \rangle_F$ , for any  $n \times n$  real matrix  $B$ .

**Lemma 3.5** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then for every  $B \in \mathbb{R}^{n \times n}$ , we have*

$$|\langle AN, B \rangle_F| \leq \frac{1}{2} (\sqrt{n} \|B\|_F + |\operatorname{tr}(B)|). \tag{3.19}$$

*Proof.* Using (3.18) for  $a = I$ ,  $x = AN$ ,  $b = B$ , and (2.1), we get

$$\begin{aligned}
|\langle I, AN \rangle_F \cdot \langle AN, B \rangle_F| &\leq \frac{1}{2} (\|I\|_F \|B\|_F + |\langle I, B \rangle_F|) \|AN\|_F^2 \Rightarrow \\
|\operatorname{tr}(AN) \cdot \langle AN, B \rangle_F| &\leq \frac{1}{2} (\sqrt{n} \|B\|_F + |\operatorname{tr}(B)|) \|AN\|_F^2 \Rightarrow \\
|\langle AN, B \rangle_F| &\leq \frac{1}{2} (\sqrt{n} \|B\|_F + |\operatorname{tr}(B)|). \quad \square
\end{aligned} \tag{3.20}$$

The next lemma provides an upper bound on the arithmetic mean of the squares of the  $n^2$  terms in the orthogonal projection  $AN$ . By the way, it also provides us with an upper bound on the arithmetic mean of the  $n$  diagonal terms in the orthogonal projection  $AN$ . These upper bounds (valid for any matrix subspace  $S$ ) are independent of the optimal approximate inverse  $N$  and thus they are independent of the subspace  $S$ , and only depend on matrix  $A$ .

**Lemma 3.6** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular with  $\operatorname{tr}(A) \neq 0$  and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then*

$$\frac{\|AN\|_F^2}{n^2} \leq \frac{\|A\|_F^2}{[\operatorname{tr}(A)]^2}, \quad \frac{\operatorname{tr}(AN)}{n} \leq \frac{n \|A\|_F^2}{[\operatorname{tr}(A)]^2}. \tag{3.21}$$

*Proof.* Using (3.18) for  $a = A^T$ ,  $x = I$ ,  $b = AN$ , the Cauchy-Schwarz inequality for  $\langle A^T, AN \rangle_F$ , and (2.1), we get

$$\begin{aligned}
|\langle A^T, I \rangle_F \cdot \langle I, AN \rangle_F| &\leq \frac{1}{2} (\|A^T\|_F \|AN\|_F + |\langle A^T, AN \rangle_F|) \|I\|_F^2 \Rightarrow \\
|tr(A) \cdot tr(AN)| &\leq \frac{n}{2} (\|A\|_F \|AN\|_F + |\langle A^T, AN \rangle_F|) \\
&\leq \frac{n}{2} (\|A\|_F \|AN\|_F + \|A^T\|_F \|AN\|_F) = n \|A\|_F \|AN\|_F \Rightarrow \\
|tr(A)| \|AN\|_F^2 &\leq n \|A\|_F \|AN\|_F \Rightarrow \frac{\|AN\|_F}{n} \leq \frac{\|A\|_F}{|tr(A)|} \Rightarrow \\
\frac{\|AN\|_F^2}{n^2} &\leq \frac{\|A\|_F^2}{[tr(A)]^2} \Rightarrow \frac{tr(AN)}{n} \leq \frac{n \|A\|_F^2}{[tr(A)]^2}. \quad \square
\end{aligned} \tag{3.22}$$

**Remark 3.4** Lemma 3.6 has the following interpretation in terms of the quality of the quality of the optimal approximate inverse  $N$  of matrix  $A$  in subspace  $S$ . The closer the ratio  $n \|A\|_F / |tr(A)|$  is to zero, the smaller  $tr(AN)$  will be, and thus, due to (2.2), the larger  $\|AN - I\|_F$  will be, and this happens for any matrix subspace  $S$ .

**Remark 3.5** By the way, from Lemma 3.6 we obtain the following inequality for any nonsingular matrix  $A \in \mathbb{R}^{n \times n}$ . Consider any matrix subspace  $S$  s.t.  $A^{-1} \in S$ . Then  $N = A^{-1}$  and using Lemma 3.6, we get

$$\frac{\|AN\|_F^2}{n^2} = \frac{\|I\|_F^2}{n^2} = \frac{1}{n} \leq \frac{\|A\|_F^2}{[tr(A)]^2} \Rightarrow |tr(A)| \leq \sqrt{n} \|A\|_F. \tag{3.23}$$

## 4 Spectral Properties

In this section, we present some new spectral properties for matrix  $AN$ ,  $N$  being the optimal approximate inverse of matrix  $A$ , defined by (1.4). Mainly, we focus on the case that matrix  $AN$  is symmetric and positive definite. This has been motivated by the following reason. When solving a large non-symmetric linear system (1.1) by using Krylov methods, a possible strategy consists of searching for an adequate optimal preconditioner  $N$  such that the preconditioned matrix  $AN$  is symmetric positive definite [14]. This enables one to use the conjugate gradient method (CG-method), which is, in general, a computationally efficient method for solving the new preconditioned system [15, 9].

Our starting point is Lemma 2.3, which has established that the sets of eigenvalues and singular values of any orthogonal projection  $AN$  satisfy

$$\sum_{i=1}^n \lambda_i^2 \leq \sum_{i=1}^n |\lambda_i|^2 \leq \sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n \lambda_i \leq \sum_{i=1}^n |\lambda_i| \leq \sum_{i=1}^n \sigma_i. \tag{4.1}$$

Let us particularize (4.1) for some special cases.

First, note that if  $AN$  is normal (i.e., for all  $1 \leq i \leq n$ :  $\sigma_i = |\lambda_i|$  [12]) then (4.1) becomes

$$\sum_{i=1}^n \lambda_i^2 \leq \sum_{i=1}^n |\lambda_i|^2 = \sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n \lambda_i \leq \sum_{i=1}^n |\lambda_i| = \sum_{i=1}^n \sigma_i. \tag{4.2}$$

In particular, if  $AN$  is symmetric ( $\sigma_i = |\lambda_i| = \pm \lambda_i \in \mathbb{R}$ ) then (4.1) becomes

$$\sum_{i=1}^n \lambda_i^2 = \sum_{i=1}^n |\lambda_i|^2 = \sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n \lambda_i \leq \sum_{i=1}^n |\lambda_i| = \sum_{i=1}^n \sigma_i. \tag{4.3}$$

In particular, if  $AN$  is symmetric and positive definite ( $\sigma_i = |\lambda_i| = \lambda_i \in \mathbb{R}^+$ ), then the equality holds in all inequalities (4.1), i.e.,

$$\sum_{i=1}^n \lambda_i^2 = \sum_{i=1}^n |\lambda_i|^2 = \sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n \lambda_i = \sum_{i=1}^n |\lambda_i| = \sum_{i=1}^n \sigma_i. \quad (4.4)$$

The next lemma compares the traces of matrices  $AN$  and  $(AN)^2$ .

**Lemma 4.1** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Then*

(i) *For any orthogonal projection  $AN$*

$$\text{tr} \left( (AN)^2 \right) \leq \|AN\|_F^2 = \text{tr} (AN). \quad (4.5)$$

(ii) *For any symmetric orthogonal projection  $AN$*

$$\left\| (AN)^2 \right\|_F \leq \text{tr} \left( (AN)^2 \right) = \|AN\|_F^2 = \text{tr} (AN). \quad (4.6)$$

(iii) *For any symmetric positive definite orthogonal projection  $AN$*

$$\left\| (AN)^2 \right\|_F \leq \text{tr} \left( (AN)^2 \right) = \|AN\|_F^2 = \text{tr} (AN) \leq [\text{tr} (AN)]^2. \quad (4.7)$$

*Proof.*

(i) Using (4.1), we get

$$\sum_{i=1}^n \lambda_i^2 \leq \sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n \lambda_i. \quad (4.8)$$

(ii) It suffices to use the obvious fact that  $\left\| (AN)^2 \right\|_F \leq \|AN\|_F^2$  and the following equalities taken from (4.3)

$$\sum_{i=1}^n \lambda_i^2 = \sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n \lambda_i. \quad (4.9)$$

(iii) It suffices to use (4.6) and the fact that (see, e.g., [13, 18]) if  $P$  and  $Q$  are symmetric positive definite matrices then  $\text{tr} (PQ) \leq \text{tr} (P) \text{tr} (Q)$  for  $P = Q = AN$ .  $\square$

The rest of the paper is devoted to obtain new properties about the eigenvalues of the orthogonal projection  $AN$  for the special case that this matrix is symmetric positive definite.

First, let us recall that the smallest singular value and the smallest eigenvalue's modulus of the orthogonal projection  $AN$  are never greater than 1 (see Lemma 2.4). The following theorem establishes the dual result for the largest eigenvalue of matrix  $AN$  (symmetric positive definite).

**Theorem 4.1** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Suppose that matrix  $AN$  is symmetric and positive definite. Then the largest eigenvalue of the orthogonal projection  $AN$  of the identity onto the subspace  $AS$  is never less than 1. That is,*

$$\sigma_1 = \lambda_1 \geq 1. \quad (4.10)$$

*Proof.* Using (4.4), we get

$$\sum_{i=1}^n \lambda_i^2 = \sum_{i=1}^n \lambda_i \Rightarrow \lambda_n^2 - \lambda_n = \sum_{i=1}^{n-1} (\lambda_i - \lambda_i^2). \quad (4.11)$$

Now, since  $\lambda_n \leq 1$  (Lemma 2.4), then  $\lambda_n^2 - \lambda_n \leq 0$ . This implies that at least one summand in the right-most sum in (4.11) must be less than or equal to zero. Suppose that such summand is the  $k$ -th one ( $1 \leq k \leq n-1$ ). Since  $AN$  is positive definite, then  $\lambda_k > 0$  and thus

$$\lambda_k - \lambda_k^2 \leq 0 \Rightarrow \lambda_k \leq \lambda_k^2 \Rightarrow \lambda_k \geq 1 \Rightarrow \lambda_1 \geq 1 \quad (4.12)$$

and the proof is concluded.  $\square$

In Theorem 4.1, the assumption that matrix  $AN$  is positive definite is essential for assuring that  $|\lambda_1| \geq 1$ , as the following simple counter-example shows. Moreover, from Lemma 2.4 and Theorem 4.1, respectively, we have that the smallest and largest eigenvalue of  $AN$  (symmetric positive definite) satisfy  $\lambda_n \leq 1$  and  $\lambda_1 \geq 1$ , respectively. Nothing can be asserted about the remaining eigenvalues of the symmetric positive definite matrix  $AN$ , which can be greater than, equal to, or less than the unity, as the same counter-example also shows.

**Example 4.1** For  $n = 3$ , let

$$A_k = \begin{pmatrix} 3 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad k \in \mathbb{R}, \quad (4.13)$$

let  $I_3$  be identity matrix of order 3, and let  $S$  be the subspace of all  $3 \times 3$  scalar matrices, i.e.,  $S = \text{span}\{I_3\}$ . Then the solution  $N_k$  to problem (1.4) for subspace  $S$  can be immediately obtained by using formula (2.3) as follows

$$N_k = \frac{\text{tr}(A_k)}{\|A_k\|_F^2} I_3 = \frac{k+4}{k^2+10} I_3 \quad (4.14)$$

and then we get

$$A_k N_k = \frac{k+4}{k^2+10} A_k = \frac{k+4}{k^2+10} \begin{pmatrix} 3 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.15)$$

Let us arrange the eigenvalues and singular values of matrix  $A_k N_k$ , as usual, in nonincreasing order (as shown in (2.5) and (2.6)).

On one hand, for  $k = -2$  we have

$$A_{-2} N_{-2} = \frac{1}{7} \begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.16)$$

and then

$$\sigma_1 = |\lambda_1| = \frac{3}{7} \geq \sigma_2 = |\lambda_2| = \frac{2}{7} \geq \sigma_3 = |\lambda_3| = \frac{1}{7}. \quad (4.17)$$

Hence,  $A_{-2} N_{-2}$  is indefinite and  $\sigma_1 = |\lambda_1| = 3/7 < 1$ .

On the other hand, for  $1 < k < 3$ , we have (see matrix (4.15))

$$\sigma_1 = \lambda_1 = 3 \frac{k+4}{k^2+10} \geq \sigma_2 = \lambda_2 = k \frac{k+4}{k^2+10} \geq \sigma_3 = \lambda_3 = \frac{k+4}{k^2+10} \quad (4.18)$$

and then

$$\begin{cases} k = 2 : & \sigma_1 = \lambda_1 = \frac{9}{7} > 1, & \sigma_2 = \lambda_2 = \frac{6}{7} < 1, & \sigma_3 = \lambda_3 = \frac{3}{7} < 1. \\ k = \frac{5}{2} : & \sigma_1 = \lambda_1 = \frac{6}{5} > 1, & \sigma_2 = \lambda_2 = 1, & \sigma_3 = \lambda_3 = \frac{2}{5} < 1. \\ k = \frac{8}{3} : & \sigma_1 = \lambda_1 = \frac{90}{77} > 1, & \sigma_2 = \lambda_2 = \frac{80}{77} > 1, & \sigma_3 = \lambda_3 = \frac{30}{77} < 1. \end{cases} \quad (4.19)$$

Hence, for  $A_k N_k$  positive definite, we have (depending on  $k$ ):  $\lambda_2 < 1$ ,  $\lambda_2 = 1$  or  $\lambda_2 > 1$ .

The following corollary improves the lower bound zero, on both  $\text{tr}(AN)$ , given in (2.1), and  $\cos(AN, I)$ , given in (3.3).

**Corollary 4.1** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Suppose that matrix  $AN$  is symmetric and positive definite. Then*

$$1 \leq \|AN\|_F \leq \text{tr}(AN) = \|AN\|_F^2 \leq n. \quad (4.20)$$

$$\cos(AN, I) \geq \frac{1}{\sqrt{n}}. \quad (4.21)$$

*Proof.* Denote by  $\|\cdot\|_2$  the spectral norm. Using the well-known inequality  $\|\cdot\|_2 \leq \|\cdot\|_F$  [11], Theorem 4.1 and (2.1), we get

$$\|AN\|_F \geq \|AN\|_2 = \sigma_1 = \lambda_1 \geq 1 \Rightarrow 1 \leq \|AN\|_F \leq \text{tr}(AN) = \|AN\|_F^2 \leq n. \quad (4.22)$$

Finally, (4.21) follows immediately from (4.20) and (3.11).  $\square$

Let us mention that an upper bound on all the eigenvalues moduli and on all singular values of any orthogonal projection  $AN$ , can be immediately obtained from (4.1) and (2.1) as follows

$$\sum_{i=1}^n |\lambda_i|^2 \leq \sum_{i=1}^n \sigma_i^2 = \|AN\|_F^2 \leq n \Rightarrow |\lambda_i|, \sigma_i \leq \sqrt{n} \text{ for all } i = 1, 2, \dots, n. \quad (4.23)$$

Our last theorem improves the upper bound given in (4.23) for the special case that the orthogonal projection  $AN$  is symmetric positive definite.

**Theorem 4.2** *Let  $A \in \mathbb{R}^{n \times n}$  be nonsingular and let  $S$  be a linear subspace of  $\mathbb{R}^{n \times n}$ . Let  $N$  be the solution to problem (1.4). Suppose that matrix  $AN$  is symmetric and positive definite. Then all the eigenvalues of matrix  $AN$  satisfy*

$$\sigma_i = \lambda_i \leq \frac{1 + \sqrt{n}}{2} \text{ for all } i = 1, 2, \dots, n. \quad (4.24)$$

*Proof.* First, note that the assertion is obvious for the smallest singular value since  $|\lambda_n| \leq 1$  for any orthogonal projection  $AN$  (Lemma 2.4). For any eigenvalue of  $AN$ , we use the fact that  $x - x^2 \leq 1/4$  for all  $x > 0$ . Then from (4.4), we get

$$\begin{aligned} \sum_{i=1}^n \lambda_i^2 = \sum_{i=1}^n \lambda_i &\Rightarrow \lambda_1^2 - \lambda_1 = \sum_{i=2}^n (\lambda_i - \lambda_i^2) \leq \frac{n-1}{4} \Rightarrow \\ \lambda_1 &\leq \frac{1 + \sqrt{n}}{2} \Rightarrow \lambda_i \leq \frac{1 + \sqrt{n}}{2} \text{ for all } i = 1, 2, \dots, n. \quad \square \end{aligned} \quad (4.25)$$

## 5 Conclusion

In this paper, we have considered the orthogonal projection  $AN$  (in the Frobenius sense) of the identity matrix onto an arbitrary matrix subspace  $AS$  ( $A \in \mathbb{R}^{n \times n}$  nonsingular,  $S \subset \mathbb{R}^{n \times n}$ ). Among other geometrical properties of matrix  $AN$ , we have established a strong relation between the quality of the approximation  $AN \approx I$  and the cosine of the angle  $\angle(AN, I)$ . Also, the distance between  $AN$  and the identity has been related to the ratio  $n \|A\|_F / |\text{tr}(A)|$  (which is independent of the subspace  $S$ ). The spectral analysis has provided lower and upper bounds on the largest eigenvalue of the symmetric positive definite orthogonal projections of the identity.

## Acknowledgments

The authors are grateful to the anonymous referee for valuable comments and suggestions, which have improved the earlier version of this paper. This work was partially supported by the “Ministerio de Economía y Competitividad” (Spanish Government), and FEDER, through Grant contract: CGL2011-29396-C03-01.

## References

- [1] R. Andreani, M. Raydan, and P. Tarazaga, “On the geometrical structure of symmetric matrices,” *Linear Algebra and its Applications*, vol. 438, no. 3, pp. 1201-1214, 2013.
- [2] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, MA, USA, 1994.
- [3] M. Benzi, “Preconditioning techniques for large linear systems: a survey,” *Journal of Computational Physics*, vol. 182, no. 2, pp. 418-477, 2002.
- [4] M. L. Buzano, “Generalizzazione della diseuguaglianza di Cauchy-Schwarz,” *Rendiconti del Seminario Matematico Università e Politecnico di Torino*, vol. 31 (1971/73), pp. 405-409, 1974 (Italian).
- [5] J.-P. Chehab and M. Raydan, “Geometrical properties of the Frobenius condition number for positive definite matrices,” *Linear Algebra and its Applications*, vol. 429, no. 8-9, pp. 2089-2097, 2008.
- [6] E. Chow and Y. Saad, “Approximate inverse preconditioners via sparse-sparse iterations,” *SIAM Journal on Scientific Computing*, vol. 19, no. 3, pp. 995-1023, 1998.
- [7] E. Chow and Y. Saad, “Approximate inverse techniques for block-partitioned matrices,” *SIAM Journal on Scientific Computing*, vol. 18, no. 6, pp. 1657-1675, 1997.
- [8] L. González, “Orthogonal projections of the identity: spectral analysis and applications to approximate inverse preconditioning,” *SIAM Review*, vol. 48, no. 1, pp. 66-75, 2006.
- [9] M. R. Hestenes and E. Stiefel, “Methods of conjugate gradients for solving linear systems,” *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409-436, 1952.
- [10] R. M. Holland, A. J. Wathen, and G. J. Shaw, “Sparse approximate inverses and target matrices,” *SIAM Journal on Scientific Computing*, vol. 26, no. 3, pp. 1000-1011, 2005.

- [11] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [12] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
- [13] E. H. Lieb and W. Thirring, “Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities,” in *Studies in Mathematical Physics: Essays in Honor of Valentine Bargmann*, E. Lieb, B. Simon, and A. Wightman, Eds., pp. 269-303, Princeton University Press, Princeton, NJ, USA, 1976.
- [14] G. Montero, L. González, E. Flórez, M. D. García, and A. Suárez, “Approximate inverse computation using Frobenius inner product,” *Numerical Linear Algebra with Applications*, vol. 9, no. 3, pp. 239-247, 2002.
- [15] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing Co., Boston, MA, USA, 1996.
- [16] A. Suárez and L. González, “A generalization of the Moore-Penrose inverse related to matrix subspaces of  $\mathbb{C}^{n \times m}$ ,” *Applied Mathematics and Computation*, vol. 216, no. 2, pp. 514-522, 2010.
- [17] A. Suárez and L. González, “Normalized Frobenius condition number of the orthogonal projections of the identity,” *Journal of Mathematical Analysis and Applications*, vol. 400, no. 2, pp. 510-516, 2013.
- [18] Z. Ulukök and R. Türkmen, “On some matrix trace inequalities,” *Journal of Inequalities and Applications*, vol. 2010, Article ID 201486, 2010.
- [19] S.-L. Wu, F. Chen, and X.-Q. Niu, “A note on the eigenvalue analysis of the SIMPLE preconditioning for incompressible flow,” *Journal of Applied Mathematics*, vol. 2012, Article ID 564132, 2012.