Research Article

Geometrical and Spectral Properties of the Orthogonal Projections of the Identity

Luis González, Antonio Suárez, and Dolores García

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

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

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 $\text{tr}(A)$ 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

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

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., [1, 2]). The coefficient matrix A of the system (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., [2, 3]). Then, to improve the convergence of these Krylov methods, the system (1) can be preconditioned with an adequate nonsingular preconditioning matrix N, transforming it into any of the equivalent systems

$$NAx = Nb,$$
$$ANy = b, \qquad x = Ny, \tag{2}$$

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 the system (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 [4]. In this way, the problem of obtaining the best preconditioner N (with respect to the Frobenius norm) of the system (1) in an arbitrary subspace S of $\mathbb{R}^{n \times n}$ is equivalent to the minimization problem; see, for example, [5]

$$\min_{M \in S} \|AM - I\|_F = \|AN - I\|_F. \tag{3}$$

The solution N to the problem (3) 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 the problem (3) has been studied as a natural generalization of the classical Moore-Penrose inverse in [6], 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 (3)) 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 the system (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 (3) 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 for example, [4, 7–9] and to the references therein. In [4], 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. A different approach (which is also focused on approximate inverses based on minimizing $\|AM - I\|_F$) can be found in [7], where 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 [8]. In [9], the idea of "target" matrix is introduced, in the context of sparse approximate inverse preconditioners, and the generalized Frobenius norms $\|B\|_{F,H}^2 = \operatorname{tr}(BHB^T)$ (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 [10]. In this 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 [5, 6, 11].

Taking advantage of the prehilbertian character of the matrix Frobenius norm, the solution N to the problem (3) 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 lemmas; see [5, 11].

Lemma 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 the problem* (3). *Then,*

$$0 \le \|AN\|_F^2 = \operatorname{tr}(AN) \le n, \tag{4}$$

$$0 \le \|AN - I\|_F^2 = n - \operatorname{tr}(AN) \le n. \tag{5}$$

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 [5]. This is the idea of the following lemma.

Lemma 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 the problem* (3) *is*

$$N = \sum_{i=1}^{d} \frac{\operatorname{tr}(AM_i)}{\|AM_i\|_F^2} M_i, \tag{6}$$

and the minimum (residual) Frobenius norm is

$$\|AN - I\|_F^2 = n - \sum_{i=1}^{d} \frac{[\operatorname{tr}(AM_i)]^2}{\|AM_i\|_F^2}. \tag{7}$$

Let us mention two possible options, both taken from [5], 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, that is,

$$S = \left\{ M \in \mathbb{R}^{n \times n} : m_{ij} = 0 \ \forall (i, j) \notin K \right\}, \tag{8}$$

$$K \subset \{1, 2, \dots, n\} \times \{1, 2, \dots, n\}.$$

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 (6).

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

$$S' = \operatorname{span} \left\{ P_1 A^T, \dots, P_d A^T \right\}, \tag{9}$$

which clearly satisfies

$$S' \subseteq \left\{ M = PA^T : P \in \mathbb{R}^{n \times n} \; P^T = P \right\}$$
$$= \left\{ M \in \mathbb{R}^{n \times n} : (AM)^T = AM \right\}. \tag{10}$$

Hence, we can explicitly obtain the solution N to the problem (3) for subspace S', from its basis $\{P_1 A^T, \ldots, P_d A^T\}$, as follows. If $\{AP_1 A^T, \ldots, AP_d A^T\}$ is an orthogonal basis of subspace AS', then we just use the orthogonal expansion (6) for obtaining N. Otherwise, we use again the Gram-Schmidt procedure to obtain an orthogonal basis of subspace AS', and then we apply formula (6). 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 [5].

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, that is,

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n| > 0,$$
$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0. \tag{11}$$

The following lemma [11] provides some inequalities involving the eigenvalues and singular values of the preconditioned matrix AN.

Lemma 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,*

$$\sum_{i=1}^n \lambda_i^2 \leq \sum_{i=1}^n |\lambda_i|^2 \leq \sum_{i=1}^n \sigma_i^2 = \|AN\|_F^2$$
$$= \operatorname{tr}(AN) = \sum_{i=1}^n \lambda_i \leq \sum_{i=1}^n |\lambda_i| \leq \sum_{i=1}^n \sigma_i. \tag{12}$$

The following fact [11] is a direct consequence of Lemma 3.

Lemma 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,*

$$0 < \sigma_n \leq |\lambda_n| \leq 1. \tag{13}$$

The following theorem [11] 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 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 the problem (3). Then,*

$$(1 - |\lambda_n|)^2 \leq (1 - \sigma_n)^2 \leq \|AN - I\|_F^2$$
$$\leq n\left(1 - |\lambda_n|^2\right) \leq n\left(1 - \sigma_n^2\right). \tag{14}$$

Remark 6. Theorem 5 states that the closer the smallest singular value σ_n of matrix AN is to the unity, the closer matrix AN will be to the identity, that is, 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 σ_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 [12]. In addition, this 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 (3). Our first lemma states some basic properties involving the cosine of the angle between matrix AN and the identity, that is,

$$\cos(AN, I) = \frac{\langle AN, I \rangle_F}{\|AN\|_F \|I\|_F} = \frac{\operatorname{tr}(AN)}{\|AN\|_F \sqrt{n}}. \tag{15}$$

Lemma 7. *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 the problem (3). Then,*

$$\cos(AN, I) = \frac{\operatorname{tr}(AN)}{\|AN\|_F \sqrt{n}} = \frac{\|AN\|_F}{\sqrt{n}} = \frac{\sqrt{\operatorname{tr}(AN)}}{\sqrt{n}}, \tag{16}$$

$$0 \leq \cos(AN, I) \leq 1, \tag{17}$$

$$\|AN - I\|_F^2 = n\left(1 - \cos^2(AN, I)\right). \tag{18}$$

Proof. First, using (15) and (4) we immediately obtain (16). As a direct consequence of (16), we derive that $\cos(AN, I)$ is always nonnegative. Finally, using (5) and (16), we get

$$\|AN - I\|_F^2 = n - \operatorname{tr}(AN) = n\left(1 - \cos^2(AN, I)\right) \tag{19}$$

and the proof is concluded. $\square$

Remark 8. In [13], the authors consider an arbitrary approximate inverse Q of matrix A and derive the following equality:

$$\|AQ - I\|_F^2 = \left(\|AQ\|_F - \|I\|_F\right)^2$$
$$+ 2\left(1 - \cos(AQ, I)\right)\|AQ\|_F \|I\|_F, \tag{20}$$

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 (3), formula (18) has stated that the strong convergence is reduced just to the weak convergence and, indeed, just to the cosine $\cos(AN, I)$.

Remark 9. More precisely, formula (18) 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 $\mathrm{tr}(AN)$, $\|AN\|_F$ (Lemma 1), or σ_n, $|\lambda_n|$ (Theorem 5). 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

$$\mathrm{tr}(AN) \nearrow n \Longleftrightarrow \|AN\|_F \nearrow \sqrt{n} \Longleftrightarrow \sigma_n \nearrow 1 \Longleftrightarrow |\lambda_n| \nearrow 1$$
$$\Longleftrightarrow \cos(AN, I) \nearrow 1 \Longleftrightarrow \|AN - I\|_F \searrow 0. \tag{21}$$

Obviously, the optimal theoretical situation corresponds to the case

$$\mathrm{tr}(AN) = n \Longleftrightarrow \|AN\|_F = \sqrt{n} \Longleftrightarrow \sigma_n = 1 \Longleftrightarrow \lambda_n = 1$$
$$\Longleftrightarrow \cos(AN, I) = 1 \Longleftrightarrow \|AN - I\|_F = 0$$
$$\Longleftrightarrow N = A^{-1} \Longleftrightarrow A^{-1} \in S. \tag{22}$$

Remark 10. Note that the ratio between $\cos(AN, I)$ and $\cos(A, I)$ is independent of the order n of matrix A. Indeed, assuming that $\mathrm{tr}(A) \neq 0$ and using (16), we immediately obtain

$$\frac{\cos(AN, I)}{\cos(A, I)} = \frac{\|AN\|_F : \sqrt{n}}{\mathrm{tr}(A) : \|A\|_F \sqrt{n}} = \|AN\|_F \frac{\|A\|_F}{\mathrm{tr}(A)}. \tag{23}$$

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

Lemma 11. *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 the problem (3). Then,*

$$\mathrm{tr}(AN) \leq \|AN\|_F \Longleftrightarrow \|AN\|_F \leq 1 \Longleftrightarrow \mathrm{tr}(AN) \leq 1$$
$$\Longleftrightarrow \cos(AN, I) \leq \frac{1}{\sqrt{n}}, \tag{24}$$

$$\mathrm{tr}(AN) \geq \|AN\|_F \Longleftrightarrow \|AN\|_F \geq 1 \Longleftrightarrow \mathrm{tr}(AN) \geq 1$$
$$\Longleftrightarrow \cos(AN, I) \geq \frac{1}{\sqrt{n}}. \tag{25}$$

Proof. Using (4), we immediately obtain the four leftmost equivalences. Using (16), we immediately obtain the two rightmost equivalences. □

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 12. *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 the problem (3). Then,*

$$\|A^{-1}\|_F \|N^{-1}\|_F \geq 1. \tag{26}$$

Proof. Using (4), we get

$$\|AN\|_F \leq \sqrt{n} = \|I\|_F = \|(AN)(AN)^{-1}\|_F$$
$$\leq \|AN\|_F \|(AN)^{-1}\|_F \Longrightarrow \|(AN)^{-1}\|_F \geq 1, \tag{27}$$

and hence

$$\|A^{-1}\|_F \|N^{-1}\|_F \geq \|N^{-1}A^{-1}\|_F = \|(AN)^{-1}\|_F \geq 1, \tag{28}$$

and the proof is concluded. □

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

$$\|AB - I\|_F^2 = \|A(B - A^{-1})\|_F^2$$
$$\leq \|A\|_F^2 \|B - A^{-1}\|_F^2. \tag{29}$$

However, for the special case that $B = N$ (the solution to the problem (3)), we also get the following inequality.

Lemma 13. *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 the problem (3). Then,*

$$\|AN - I\|_F^2 \leq \|A\|_F \|N - A^{-1}\|_F. \tag{30}$$

Proof. Using the Cauchy-Schwarz inequality and (5), we get

$$\left|\left\langle A^{-1} - N, A^T \right\rangle_F\right| \leq \|A^{-1} - N\|_F \|A^T\|_F$$
$$\Longrightarrow \left|\mathrm{tr}\left((A^{-1} - N)A\right)\right| \leq \|A^{-1} - N\|_F \|A^T\|_F$$
$$\Longrightarrow \left|\mathrm{tr}\left(A(A^{-1} - N)\right)\right| \leq \|N - A^{-1}\|_F \|A\|_F$$
$$\Longrightarrow |\mathrm{tr}(I - AN)| \leq \|N - A^{-1}\|_F \|A\|_F$$
$$\Longrightarrow n - \mathrm{tr}(AN) \leq \|N - A^{-1}\|_F \|A\|_F$$
$$\Longrightarrow \|AN - I\|_F^2 \leq \|A\|_F \|N - A^{-1}\|_F. \tag{31}$$

□

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 [14]. For all $a, x, b \in H$, we have

$$|\langle a, x \rangle \cdot \langle x, b \rangle| \leq \frac{1}{2} \left(\|a\| \|b\| + |\langle a, b \rangle| \right) \|x\|^2. \quad (32)$$

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 14. *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 the problem* (3). *Then, for every $B \in \mathbb{R}^{n \times n}$, we have*

$$\left| \langle AN, B \rangle_F \right| \leq \frac{1}{2} \left(\sqrt{n} \|B\|_F + |\mathrm{tr}\,(B)| \right). \quad (33)$$

Proof. Using (32) for $a = I$, $x = AN$, $b = B$, and (4), we get

$$\left| \langle I, AN \rangle_F \cdot \langle AN, B \rangle_F \right| \leq \frac{1}{2} \left(\|I\|_F \|B\|_F + |\langle I, B \rangle_F| \right) \|AN\|_F^2$$

$$\implies \left| \mathrm{tr}\,(AN) \cdot \langle AN, B \rangle_F \right|$$

$$\leq \frac{1}{2} \left(\sqrt{n} \|B\|_F + |\mathrm{tr}\,(B)| \right) \|AN\|_F^2$$

$$\implies \left| \langle AN, B \rangle_F \right| \leq \frac{1}{2} \left(\sqrt{n} \|B\|_F + |\mathrm{tr}\,(B)| \right). \quad (34)$$

$\square$

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 15. *Let $A \in \mathbb{R}^{n \times n}$ be nonsingular with $\mathrm{tr}(A) \neq 0$ and let S be a linear subspace of $\mathbb{R}^{n \times n}$. Let N be the solution to the problem* (3). *Then,*

$$\frac{\|AN\|_F^2}{n^2} \leq \frac{\|A\|_F^2}{[\mathrm{tr}\,(A)]^2},$$

$$\frac{\mathrm{tr}\,(AN)}{n} \leq \frac{n\|A\|_F^2}{[\mathrm{tr}\,(A)]^2}. \quad (35)$$

Proof. Using (32) for $a = A^T$, $x = I$, and $b = AN$, and the Cauchy-Schwarz inequality for $\langle A^T, AN \rangle_F$ and (4), we get

$$\left| \langle A^T, I \rangle_F \cdot \langle I, AN \rangle_F \right|$$

$$\leq \frac{1}{2} \left(\|A^T\|_F \|AN\|_F + \left| \langle A^T, AN \rangle_F \right| \right) \|I\|_F^2$$

$$\implies |\mathrm{tr}\,(A) \cdot \mathrm{tr}\,(AN)|$$

$$\leq \frac{n}{2} \left(\|A\|_F \|AN\|_F + \left| \langle A^T, AN \rangle_F \right| \right)$$

$$\leq \frac{n}{2} \left(\|A\|_F \|AN\|_F + \left\| A^T \right\|_F \|AN\|_F \right)$$

$$= n \|A\|_F \|AN\|_F$$

$$\implies |\mathrm{tr}\,(A)| \, \|AN\|_F^2 \leq n \|A\|_F \|AN\|_F$$

$$\implies \frac{\|AN\|_F}{n} \leq \frac{\|A\|_F}{|\mathrm{tr}\,(A)|}$$

$$\implies \frac{\|AN\|_F^2}{n^2} \leq \frac{\|A\|_F^2}{[\mathrm{tr}\,(A)]^2} \implies \frac{\mathrm{tr}\,(AN)}{n} \leq \frac{n\|A\|_F^2}{[\mathrm{tr}\,(A)]^2}. \quad (36)$$

$\square$

Remark 16. Lemma 15 has the following interpretation in terms of the quality of the optimal approximate inverse N of matrix A in subspace S. The closer the ratio $n\|A\|_F/|\mathrm{tr}(A)|$ is to zero, the smaller $\mathrm{tr}(AN)$ will be, and thus, due to (5), the larger $\|AN - I\|_F$ will be, and this happens for any matrix subspace S.

Remark 17. By the way, from Lemma 15, 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 15, we get

$$\frac{\|AN\|_F^2}{n^2} = \frac{\|I\|_F^2}{n^2} = \frac{1}{n} \leq \frac{\|A\|_F^2}{[\mathrm{tr}\,(A)]^2} \quad (37)$$

$$\implies |\mathrm{tr}\,(A)| \leq \sqrt{n} \|A\|_F.$$

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 (3). 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 nonsymmetric linear system (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 [5]. 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 [2, 15].

Our starting point is Lemma 3, which has established that the sets of eigenvalues and singular values of any orthogonal projection AN satisfy

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

$$= \sum_{i=1}^{n} \lambda_i \leq \sum_{i=1}^{n} |\lambda_i| \leq \sum_{i=1}^{n} \sigma_i. \quad (38)$$

Let us particularize (38) for some special cases.

First, note that if AN is normal (i.e., for all $1 \leq i \leq n$: $\sigma_i = |\lambda_i|$ [16]), then (38) becomes

$$\sum_{i=1}^{n} \lambda_i^2 \leq \sum_{i=1}^{n} |\lambda_i|^2 = \sum_{i=1}^{n} \sigma_i^2$$
$$= \sum_{i=1}^{n} \lambda_i \leq \sum_{i=1}^{n} |\lambda_i| = \sum_{i=1}^{n} \sigma_i. \tag{39}$$

In particular, if AN is symmetric ($\sigma_i = |\lambda_i| = \pm\lambda_i \in \mathbb{R}$), then (38) becomes

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} |\lambda_i|^2 = \sum_{i=1}^{n} \sigma_i^2$$
$$= \sum_{i=1}^{n} \lambda_i \leq \sum_{i=1}^{n} |\lambda_i| = \sum_{i=1}^{n} \sigma_i. \tag{40}$$

In particular, if AN is symmetric and positive definite ($\sigma_i = |\lambda_i| = \lambda_i \in \mathbb{R}^+$), then the equality holds in all (38), that is,

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} |\lambda_i|^2 = \sum_{i=1}^{n} \sigma_i^2$$
$$= \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} |\lambda_i| = \sum_{i=1}^{n} \sigma_i. \tag{41}$$

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

Lemma 18. *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 the problem (3). Then,*

(i) *for any orthogonal projection AN*

$$\mathrm{tr}\left((AN)^2\right) \leq \|AN\|_F^2 = \mathrm{tr}(AN), \tag{42}$$

(ii) *for any symmetric orthogonal projection AN*

$$\left\|(AN)^2\right\|_F \leq \mathrm{tr}\left((AN)^2\right) = \|AN\|_F^2 = \mathrm{tr}(AN), \tag{43}$$

(iii) *for any symmetric positive definite orthogonal projection AN*

$$\left\|(AN)^2\right\|_F \leq \mathrm{tr}\left((AN)^2\right) = \|AN\|_F^2 = \mathrm{tr}(AN) \leq [\mathrm{tr}(AN)]^2. \tag{44}$$

Proof. (i) Using (38), we get

$$\sum_{i=1}^{n} \lambda_i^2 \leq \sum_{i=1}^{n} \sigma_i^2 = \sum_{i=1}^{n} \lambda_i. \tag{45}$$

(ii) It suffices to use the obvious fact that $\|(AN)^2\|_F \leq \|AN\|_F^2$ and the following equalities taken from (40):

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} \sigma_i^2 = \sum_{i=1}^{n} \lambda_i. \tag{46}$$

(iii) It suffices to use (43) and the fact that (see, e.g., [17, 18]) if P and Q are symmetric positive definite matrices then $\mathrm{tr}(PQ) \leq \mathrm{tr}(P)\,\mathrm{tr}(Q)$ for $P = Q = AN$. $\qquad \square$

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 4). The following theorem establishes the dual result for the largest eigenvalue of matrix AN (symmetric positive definite).

Theorem 19. *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 the problem (3). 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,*

$$\sigma_1 = \lambda_1 \geq 1. \tag{47}$$

Proof. Using (41), we get

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} \lambda_i \implies \lambda_n^2 - \lambda_n = \sum_{i=1}^{n-1} \left(\lambda_i - \lambda_i^2\right). \tag{48}$$

Now, since $\lambda_n \leq 1$ (Lemma 4), then $\lambda_n^2 - \lambda_n \leq 0$. This implies that at least one summand in the rightmost sum in (48) must be less than or equal to zero. Suppose that such summand is the kth one ($1 \leq k \leq n-1$). Since AN is positive definite, then $\lambda_k > 0$, and thus

$$\lambda_k - \lambda_k^2 \leq 0 \implies \lambda_k \leq \lambda_k^2 \implies \lambda_k \geq 1 \implies \lambda_1 \geq 1 \tag{49}$$

and the proof is concluded. $\qquad \square$

In Theorem 19, the assumption that matrix AN is positive definite is essential for assuring that $|\lambda_1| \geq 1$, as the following simple counterexample shows. Moreover, from Lemma 4 and Theorem 19, respectively, we have that the smallest and largest eigenvalues 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 counterexample also shows.

Example 20. For $n = 3$, let

$$A_k = \begin{pmatrix} 3 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad k \in \mathbb{R}, \tag{50}$$

let I_3 be identity matrix of order 3, and let S be the subspace of all 3×3 scalar matrices; that is, $S = \mathrm{span}\{I_3\}$. Then the solution N_k to the problem (3) for subspace S can be immediately obtained by using formula (6) as follows:

$$N_k = \frac{\mathrm{tr}(A_k)}{\|A_k\|_F^2} I_3 = \frac{k+4}{k^2+10} I_3 \tag{51}$$

and then we get

$$A_k N_k = \frac{k+4}{k^2+10} A_k = \frac{k+4}{k^2+10} \begin{pmatrix} 3 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{52}$$

Let us arrange the eigenvalues and singular values of matrix $A_k N_k$, as usual, in nonincreasing order (as shown in (11)).

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

$$A_{-2}N_{-2} = \frac{1}{7}\begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad (53)$$

and then

$$\sigma_1 = |\lambda_1| = \frac{3}{7}$$

$$\geq \sigma_2 = |\lambda_2| = \frac{2}{7} \qquad (54)$$

$$\geq \sigma_3 = |\lambda_3| = \frac{1}{7}.$$

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 (52))

$$\sigma_1 = \lambda_1 = 3\frac{k+4}{k^2+10}$$

$$\geq \sigma_2 = \lambda_2 = k\frac{k+4}{k^2+10} \qquad (55)$$

$$\geq \sigma_3 = \lambda_3 = \frac{k+4}{k^2+10},$$

and then

$$k = 2: \quad \sigma_1 = \lambda_1 = \frac{9}{7} > 1,$$

$$\sigma_2 = \lambda_2 = \frac{6}{7} < 1,$$

$$\sigma_3 = \lambda_3 = \frac{3}{7} < 1,$$

$$k = \frac{5}{2}: \quad \sigma_1 = \lambda_1 = \frac{6}{5} > 1,$$

$$\sigma_2 = \lambda_2 = 1, \qquad (56)$$

$$\sigma_3 = \lambda_3 = \frac{2}{5} < 1,$$

$$k = \frac{8}{3}: \quad \sigma_1 = \lambda_1 = \frac{90}{77} > 1,$$

$$\sigma_2 = \lambda_2 = \frac{80}{77} > 1,$$

$$\sigma_3 = \lambda_3 = \frac{30}{77} < 1.$$

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 (4), and $\cos(AN, I)$, given in (17).

Corollary 21. *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 the* problem (3). *Suppose that matrix AN is symmetric and positive definite. Then,*

$$1 \leq \|AN\|_F \leq \text{tr}(AN) = \|AN\|_F^2 \leq n, \qquad (57)$$

$$\cos(AN, I) \geq \frac{1}{\sqrt{n}}. \qquad (58)$$

Proof. Denote by $\|\cdot\|_2$ the spectral norm. Using the well-known inequality $\|\cdot\|_2 \leq \|\cdot\|_F$ [19], Theorem 19, and (4), we get

$$\|AN\|_F \geq \|AN\|_2 = \sigma_1 = \lambda_1 \geq 1$$
$$\Longrightarrow 1 \leq \|AN\|_F \leq \text{tr}(AN) = \|AN\|_F^2 \leq n. \qquad (59)$$

Finally, (58) follows immediately from (57) and (25). □

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 (38) and (4) as follows:

$$\sum_{i=1}^{n} |\lambda_i|^2 \leq \sum_{i=1}^{n} \sigma_i^2 = \|AN\|_F^2 \leq n$$
$$\Longrightarrow |\lambda_i|, \quad \sigma_i \leq \sqrt{n}, \ \forall i = 1, 2, \ldots, n. \qquad (60)$$

Our last theorem improves the upper bound given in (60) for the special case that the orthogonal projection AN is symmetric positive definite.

Theorem 22. *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 the problem (3). Suppose that matrix AN is symmetric and positive definite. Then, all the eigenvalues of matrix AN satisfy*

$$\sigma_i = \lambda_i \leq \frac{1 + \sqrt{n}}{2} \quad \forall i = 1, 2, \ldots, n. \qquad (61)$$

Proof. First, note that the assertion is obvious for the smallest singular value since $|\lambda_n| \leq 1$ for any orthogonal projection AN (Lemma 4). For any eigenvalue of AN, we use the fact that $x - x^2 \leq 1/4$ for all $x > 0$. Then from (41), we get

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} \lambda_i$$

$$\Longrightarrow \lambda_1^2 - \lambda_1 = \sum_{i=2}^{n} (\lambda_i - \lambda_i^2) \leq \frac{n-1}{4} \qquad (62)$$

$$\Longrightarrow \lambda_1 \leq \frac{1+\sqrt{n}}{2} \Longrightarrow \lambda_i \leq \frac{1+\sqrt{n}}{2}$$

$$\forall i = 1, 2, \ldots, n.$$

□

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 $n\|A\|_F / |\operatorname{tr}(A)|$ (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

References

[1] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, Mass, USA, 1994.

[2] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing, Boston, Mass, USA, 1996.

[3] 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, 7 pages, 2012.

[4] M. Benzi, "Preconditioning techniques for large linear systems: a survey," *Journal of Computational Physics*, vol. 182, no. 2, pp. 418–477, 2002.

[5] 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.

[6] A. Suárez and L. González, "A generalization of the Moore-Penrose inverse related to matrix subspaces of $C^{n \times m}$," *Applied Mathematics and Computation*, vol. 216, no. 2, pp. 514–522, 2010.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] M. L. Buzano, "Generalizzazione della diseguaglianza di Cauchy-Schwarz," *Rendiconti del Seminario Matematico Università e Politecnico di Torino*, vol. 31, pp. 405–409, 1974 (Italian).

[15] 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.

[16] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[17] 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.

[18] Z. Ulukök and R. Türkmen, "On some matrix trace inequalities," *Journal of Inequalities and Applications*, vol. 2010, Article ID 201486, 8 pages, 2010.

[19] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.