Advances in Linear Algebra & Matrix Theory, 2014, 4, 172-185
Published Online September 2014 in SciRes. http://www.scirp.org/journal/alamt
http://dx.doi.org/10.4236/alamt.2014.43015

Low-Rank Positive Approximants of Symmetric Matrices

Achiya Dax

Hydrological Service, Jerusalem, Israel
Email: dax20@water.gov.il

Received 25 July 2014; revised 25 August 2014; accepted 23 September 2014

Abstract

Given a symmetric matrix *X*, we consider the problem of finding a low-rank positive approximant of *X*. That is, a symmetric positive semidefinite matrix, *S*, whose rank is smaller than a given positive integer, ℓ, which is nearest to *X* in a certain matrix norm. The problem is first solved with regard to four common norms: The Frobenius norm, the Schatten *p*-norm, the trace norm, and the spectral norm. Then the solution is extended to any unitarily invariant matrix norm. The proof is based on a subtle combination of Ky Fan dominance theorem, a modified pinching principle, and Mirsky minimum-norm theorem.

Keywords

Low-Rank Positive Approximants, Unitarily Invariant Matrix Norms

1. Introduction

Let X be a given real symmetric $n \times n$ matrix. In this paper we consider the problem of finding a low-rank symmetric positive semidefinite matrix which is nearest to X with regard to a certain matrix norm. Let $\|\cdot\|$ be a given unitarily invariant matrix norm on $\mathbb{R}^{n\times n}$. (The basic features of such norms are explained in the next section.) Let ℓ be a given positive integer such that $1 \le \ell \le n-1$, and define

$$\mathbb{S}^{+}_{n,\ell} = \left\{ S \,\middle|\, S \in \mathbb{R}^{n\times n},\, S \ge 0, \text{ and } \operatorname{rank}(S) \le \ell \right\},$$

where the notation $S \ge 0$ means that S is symmetric and positive semidefinite. Then the problem to solve has the form

$$\begin{aligned} &\text{minimize} && F(S) = \|X - S\| \\ &\text{subject to} && S \in \mathbb{S}^{+}_{n,\ell}. \end{aligned} \tag{1.1}$$

How to cite this paper: Dax, A. (2014) Low-Rank Positive Approximants of Symmetric Matrices. *Advances in Linear Algebra & Matrix Theory*, **4**, 172-185. http://dx.doi.org/10.4236/alamt.2014.43015

The need for solving such problems arises in certain matrix completion methods that consider Euclidean distance matrices, see [1] or [2]. Since X is assumed to be a symmetric matrix, it has a spectral decomposition

$$X = Q\Lambda Q^{\mathrm{T}}, \tag{1.2}$$

where $Q \in \mathbb{R}^{n\times n}$, $Q^{\mathrm{T}}Q = I$, is an orthonormal matrix

$$\Lambda = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$$

is a diagonal matrix, and

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$$

are the eigenvalues of X in decreasing order. If $X \ge 0$, then $\lambda_n \ge 0$ and the spectral decomposition (1.2) coincides with the SVD of X. In this case the solution of (1.1) is given by the Eckart-Young-Mirsky theorem. See the next section.

The rest of the paper assumes, therefore, that $\lambda_n < 0$. In this case the solution of (1.1) is related to that of the problem

$$\begin{aligned} &\text{minimize} && F(S) = \|X - S\| \\ &\text{subject to} && S \in \mathbb{S}_n^+, \end{aligned} \tag{1.3}$$

where

$$\mathbb{S}_n^+ = \left\{ S \,\middle|\, S \in \mathbb{R}^{n\times n} \text{ and } S \ge 0 \right\}.$$

Let q be a non-negative integer that counts the number of positive eigenvalues. That is

$$\lambda_j > 0 \ \text{ for } \ j = 1, \cdots, q, \ \text{ and } \ \lambda_j \le 0 \ \text{ for } \ j = q+1, \cdots, n. \tag{1.4}$$

Let the diagonal matrix Λ_q denotes the positive part of Λ,

$$\Lambda_q = \mathrm{diag}\{\lambda_1, \cdots, \lambda_q, 0, \cdots, 0\}. \tag{1.5}$$

(If $\lambda_1 \le 0$ then $q = 0$, and Λ_0 is the null matrix.) Then, as we shall see in Section 3, the matrix

$$X_q = Q\Lambda_q Q^{\mathrm{T}} \in \mathbb{S}_n^+$$

solves (1.3) in any unitarily invariant norm.

If $q \le \ell$ then, clearly, X_q is also a solution of (1.1). Hence in the rest of the paper we assume that $1 \le \ell < q$. This assumption implies that the diagonal matrix

$$\Lambda_\ell = \mathrm{diag}\{\lambda_1, \cdots, \lambda_\ell, 0, \cdots, 0\} \tag{1.6}$$

belongs to $\mathbb{S}_{n,\ell}^+$. The aim of the paper is to show that the matrix

$$X_\ell = Q\Lambda_\ell Q^{\mathrm{T}}$$

solves (1.1) for any unitarily invariant norm.

Let $A \in \mathbb{R}^{n\times n}$ be a given real $n \times n$ matrix. Then another related problem is

$$\begin{aligned} &\text{minimize} && F(S) = \|A - S\| \\ &\text{subject to} && S \in \mathbb{S}_n^+, \end{aligned} \tag{1.7}$$

The relation between (1.7) and (1.3) is seen when using the Frobenius matrix norm. Let $S \in \mathbb{R}^{n\times n}$ be a symmetric matrix and let $T \in \mathbb{R}^{n\times n}$ be a skew-symmetric matrix. Then, clearly,

$$\|S + T\|_F^2 = \|S\|_F^2 + \|T\|_F^2. \tag{1.8}$$

Recall also that any matrix $A \in \mathbb{R}^{n\times n}$ has a unique presentation as the sum $A = X + Y$ where $X = \left(A^{\mathrm{T}} + A\right)/2$ is symmetric and $Y = \left(A^{\mathrm{T}} - A\right)/2$ is skew-symmetric. Consequently, for any symmetric matrix, S say,

$$\|A-S\|_F^2=\|X-S+Y\|_F^2=\|X-S\|_F^2+\|Y\|_F^2. \tag{1.9}$$

Therefore, when using the Frobenius norm, a solution of (1.3) provides a solution of (1.7). This observation is due to Higham [3]. A matrix that solves (1.7) or (1.3) is called "**positive approximant**". Similarly, the term "**low-rank positive approximant**" refers to a matrix that solves (1.1).

The current interest in positive approximants was initiated in Halmos' paper [4], which considers the solution of (1.7) in the spectral norm. Rogers and Ward [5] considered the solution of (1.7) in the Schatten-*p* norm, Ando [6] considered this problem in the trace norm, and Higham [3] considered the Frobenius norm. Halmos [4] has considered the positive approximant problem in a more general context of linear operators on a Hilbert space. Other positive approximants problems (in the operators context) are considered in [7]-[11]. The problems (1.1), (1.3) and (1.7) fall into the category of "matrix nearness problems". Further examples of matrix (or operator) nearness problems are discussed in [12]-[18]. A review of this topic is given in Higham [19].

The plan of the paper is as follows. In the next section we introduce notations and tools which are needed for the coming discussions. In Section 3 we show that X_q solves (1.3). Section 4 considers the solution of (1.1) in Frobenius norm. This involves the Eckart-Young theorem. In the next sections Mirsky theorem extends the results to Schatten-*p* norms, the trace norm, and the spectral norm. Then it is proved that X_ℓ solves (1.1) in any unitarily invariant norm. The proof of this claim requires a subtle combination of Ky Fan dominance theorem, a modified pinching principle, and Mirsky theorem.

2. Notations and Tools

In this section we introduce notations and facts which are needed for coming discussions. Here A denotes a real $m\times n$ matrix with $m\ge n$. Let

$$A=USV^{\mathrm{T}} \tag{2.1}$$

be an SVD of A, where $U=[\mathbf{u}_1,\cdots,\mathbf{u}_m]$ is an $m\times m$ orthogonal matrix, $V=[\mathbf{v}_1,\cdots,\mathbf{v}_n]$ is an $n\times n$ orthogonal matrix, and $S=\mathrm{diag}\{\sigma_1,\cdots,\sigma_n\}$ is an $m\times n$ diagonal matrix. The singular values of A are assumed to be nonnegative and sorted to satisfy

$$\sigma_1\ge\sigma_2\ge\cdots\ge\sigma_n\ge 0. \tag{2.2}$$

The columns of U and V are called left singular vectors and right singular vectors, respectively. These vectors are related by the equalities

$$A\mathbf{v}_j=\sigma_j\mathbf{u}_j \quad\text{and}\quad A^{\mathrm{T}}\mathbf{u}_j=\sigma_j\mathbf{v}_j,\quad j=1,\cdots,n. \tag{2.3}$$

A further consequence of (2.1) is the equality

$$A=\sum_{j=1}^{n}\sigma_j\mathbf{u}_j\mathbf{v}_j^{\mathrm{T}}. \tag{2.4}$$

Moreover, let r denotes the rank of A. Then, clearly,

$$\sigma_1\ge\cdots\ge\sigma_r>0 \quad\text{and}\quad \sigma_j=0 \quad\text{for}\quad j=r+1,\cdots,n. \tag{2.5}$$

So (2.4) can be rewritten as

$$A=\sum_{j=1}^{r}\sigma_j\mathbf{u}_j\mathbf{v}_j^{\mathrm{T}}. \tag{2.6}$$

Let the matrices

$$U_k=[\mathbf{u}_1,\cdots,\mathbf{u}_k]\in\mathbb{R}^{m\times k} \quad\text{and}\quad V_k=[\mathbf{v}_1,\cdots,\mathbf{v}_k]\in\mathbb{R}^{n\times k} \tag{2.7}$$

be constructed from the first k columns of U and V, respectively. Let $S_k=\mathrm{diag}\{\sigma_1,\cdots,\sigma_k\}$ be a $k\times k$ diagonal matrix. Then the matrix

$$T_k=U_kS_kV_k^{\mathrm{T}}=\sum_{j=1}^{k}\sigma_j\mathbf{u}_j\mathbf{v}_j^{T} \tag{2.8}$$

is called a rank-*k* truncated SVD of A. (If $\sigma_k = \sigma_{k+1}$ then this matrix is not unique.)

Let a_{ij}, u_{ij}, v_{ij} denote the (i, j) entries of the matrices A, U, V, respectively. Then (2.4) indicates that

$$a_{ii} = \sum_{j=1}^{n} \sigma_j u_{ij} v_{ij} \quad \text{for } i = 1, \cdots, n \tag{2.9}$$

and

$$\sum_{i=1}^{n} |a_{ii}| \le \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_j |u_{ij}| \cdot |v_{ij}| = \sum_{j=1}^{n} \sigma_j \sum_{i=1}^{n} |u_{ij}| \cdot |v_{ij}| \le \sum_{j=1}^{n} \sigma_j, \tag{2.10}$$

where the last inequality follows from the Cauchy-Schwarz inequality and the fact that the columns of U and V have unit length.

Another useful property regards the concepts of majorization and unitarily invariant norms. Recall that a matrix norm $\|\cdot\|$ on $\mathbb{R}^{m\times n}$ is called unitarily invariant if the equalities

$$\|A\| = \|X^{\mathrm{T}} A\| = \|AY\| = \|X^{\mathrm{T}} AY\| \tag{2.11}$$

are satisfied for any matrix $A \in \mathbb{R}^{m\times n}$, and any pair of unitary matrices $X \in \mathbb{R}^{m\times m}$ and $Y \in \mathbb{R}^{n\times n}$. Let B and C be a given pair of $m \times n$ matrices with singular values

$$\beta_1 \ge \beta_2 \ge \cdots \ge \beta_n \ge 0 \quad \text{and} \quad \gamma_1 \ge \gamma_2 \ge \cdots \ge \gamma_n \ge 0,$$

respectively. Let $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_n)^{\mathrm{T}}$ and $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_n)^{\mathrm{T}}$ denote the corresponding *n*-vectors of singular values. Then the weak majorization relation $\boldsymbol{\beta} <_\omega \boldsymbol{\gamma}$ means that these vectors satisfy the inequality

$$\sum_{j=1}^{k} \beta_j \le \sum_{j=1}^{k} \gamma_j \quad \text{for } k = 1, \cdots, n. \tag{2.12}$$

In this case we say that $\boldsymbol{\beta}$ is weakly majorized by $\boldsymbol{\gamma}$, or that the singular values of B are weakly majorized by those of C. The dominance theorem of Fan [20] [21] relates these two concepts. It says that if the singular values of B are weakly majorized by those of C then the inequality

$$\|B\| \le \|C\| \tag{2.13}$$

holds for any unitarily invariant norm. For detailed proof of this fact see, for example, [8], [20]-[23]. The most popular example of an unitarily invariant norm is, perhaps, the Frobenius matrix norm

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}, \tag{2.14}$$

which satisfies

$$\|A\|_F^2 = \operatorname{trace}(A^{\mathrm{T}} A) = \operatorname{trace}(AA^{\mathrm{T}}) = \sum_{j=1}^{n} \sigma_j^2. \tag{2.15}$$

Other examples are the Schatten *p*-norms,

$$\|A\|_p = \left(\sum_{j=1}^{n} \sigma_j^p \right)^{1/p}, \quad 1 \le p < \infty \tag{2.16}$$

and Ky Fan *k*-norms,

$$\|A\|_{(k)} = \sum_{j=1}^{k} \sigma_j, \quad k = 1, \cdots, n. \tag{2.17}$$

The trace norm,

$$\|A\|_{\mathrm{tr}} = \sum_{j=1}^{n} \sigma_j \tag{2.18}$$

is obtained for $k = n$ and $p = 1$, while the spectral norm

$$\|A\|_{\text{sp}} = \sigma_1 = \max_j \sigma_j \tag{2.19}$$

corresponds to $k=1$ and $p=\infty$. The use of Ky Fan k-norms enables us to state the dominance principle in the following way.

Theorem 1 (Ky Fan dominance theorem) *The Inequality* (2.13) *holds for any unitarily invariant norm if and only if*

$$\|B\|_{(k)} \le \|C\|_{(k)} \quad \text{for } k=1,\cdots,n. \tag{2.20}$$

Another useful tool is the following "rectangular" version of Cauchy interlacing theorem. For a proof of this result see ([24], p. 229) or ([25], p. 1250).

Theorem 2 (A rectangular Cauchy interlace theorem) *Let the* $m\times n$ *matrix* A *have the singular values* (2.2). *Let the* $\tilde{m}\times\tilde{n}$ *matrix* $\tilde{A}$ *be a submatrix of* A *which is obtained by deleting* m' *rows and* n' *columns of* A. *That is,* $\tilde{m}+m'=m$ *and* $\tilde{n}+n'=n$. *Define* $\tilde{k}=\min\{\tilde{m},\tilde{n}\}$ *and let*

$$\tilde{\sigma}_1 \ge \tilde{\sigma}_2 \ge \cdots \ge \tilde{\sigma}_{\tilde{k}} \ge 0$$

denote the singular values of $\tilde{A}$. *Then*

$$\sigma_j \ge \tilde{\sigma}_j \ge \sigma_{m'+n'+j} \quad \text{for } j=1,\cdots,\tilde{k}. \tag{2.21}$$

To ease the coming discussions we return to square matrices. In the next assertions $W=(w_{ij})\in\mathbb{R}^{n\times n}$ is an arbitrary real $n\times n$ matrix. Combining the interlace theorem with the dominance theorem leads to the following corollary.

Theorem 3 *Let the* $n\times n$ *matrix* B_k *be obtained from* W *by setting to zero all the entries in the last* $n-k$ *rows and columns of* W. *Then the inequality*

$$\|B_k\| \le \|W\| \tag{2.22}$$

holds for any unitarily invariant norm.

Theorem 4 *Let the* $n\times n$ *diagonal matrix*

$$\text{diag}(W) = \text{diag}\{w_{11},\ldots,w_{nn}\},$$

be obtained from the diagonal entries of W. *Then*

$$\|\text{diag}(W)\| \le \|W\| \tag{2.23}$$

in any unitarily invariant norm.

Proof. There is no loss of generality in assuming that the diagonal entries of W are ordered such that

$$|w_{11}| \ge |w_{22}| \ge \cdots \ge |w_{nn}|.$$

Let the matrix B_k be defined as in Theorem 3. Then from (2.10) and (2.22) we conclude that

$$\|\text{diag}(W)\|_{(k)} \le \|B_k\|_{(k)} \le \|W\|_{(k)} \quad \text{for } k=1,\cdots,n,$$

which proves (2.23). □

Corollary 5 *The diagonal matrix*

$$\text{diag}(B_k) = \text{diag}\{w_{11},\cdots,w_{kk},0,0,\cdots,0\}$$

satisfies

$$\|\text{diag}(B_k)\| \le \|\text{diag}(W)\| \le \|W\| \tag{2.24}$$

in any unitarily invariant norm.

Lemma 6 *Let* X *and* Y *be a pair of real symmetric* $n\times n$ *matrices that satisfy*

$$0 \le X \le Y.$$

Then

$$\|X\| \le \|Y\|$$

in any unitarily invariant norm.

Proof. Using the spectral decomposition of X it is possible to assume that X is a diagonal matrix:

$$X = \operatorname{diag}\{\lambda_1, \cdots, \lambda_n\}.$$

The matrix $Y - X$ is positive semidefinite and, therefore, has non-negative diagonal entries. This observation implies the inequalities

$$y_{jj} \ge \lambda_j \quad \text{for } j = 1, \cdots, n,$$

and

$$\|X\| = \|\operatorname{diag}\{\lambda_1, \cdots, \lambda_n\}\| \le \|\operatorname{diag}\{y_{11}, \cdots, y_{nn}\}\| = \|\operatorname{diag}(Y)\|,$$

while (2.23) gives

$$\|\operatorname{diag}(Y)\| \le \|Y\|. \qquad \square$$

Theorem 7 (The pinching principle) *Let the matrix* $W \in \mathbb{R}^{n\times n}$ *be partitioned in the form*

$$\left(\begin{array}{c|c} W_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array}\right) \tag{2.25}$$

where $W_{11} \in \mathbb{R}^{q\times q}$ *and* $W_{22} \in \mathbb{R}^{(n-q)\times(n-q)}$. *Let the* $n \times n$ *matrix*

$$\hat{W} = \left(\begin{array}{c|c} W_{11} & 0 \\ \hline 0 & W_{22} \end{array}\right) \tag{2.26}$$

denote the "pinched" version of W. *Then the inequality*

$$\|\hat{W}\| \le \|W\| \tag{2.27}$$

holds in any unitarily invariant norm.

Proof. Using the SVD of W_{11} we obtain an pair of $q \times q$ orthonormal matrices, U_{11} and V_{11}, such that $U_{11}^T W_{11} V_{11}$ is a diagonal matrix that contains the singular values of W_{11}. Similarly there exists a pair of $(n-q)\times(n-q)$ orthonormal matrices, U_{22} and V_{22}, such that $U_{22}^{\mathrm{T}} W_{22} V_{22}$ is a diagonal matrix that contains the singular values of W_{22}. The related $n \times n$ matrices

$$U = \left(\begin{array}{c|c} U_{11} & 0 \\ \hline 0 & U_{22} \end{array}\right) \quad \text{and} \quad V = \left(\begin{array}{c|c} V_{11} & 0 \\ \hline 0 & V_{22} \end{array}\right)$$

are orthonormal matrices, and

$$U^{\mathrm{T}} \hat{W} V = \left(\begin{array}{c|c} U_{11}^{\mathrm{T}} W_{11} V_{11} & \\ \hline & U_{22}^{\mathrm{T}} W_{22} V_{22} \end{array}\right) \tag{2.28}$$

is a diagonal matrix. Moreover, comparing $U^{\mathrm{T}} W V$ with $U^{\mathrm{T}} \hat{W} V$ shows that

$$U^{\mathrm{T}} \hat{W} V = \operatorname{diag}(U^{\mathrm{T}} W V). \tag{2.29}$$

Hence a further use of (2.23) gives

$$\|\hat{W}\| = \|U^{\mathrm{T}} \hat{W} V\| = \|\operatorname{diag}(U^{\mathrm{T}} W V)\| \le \|U^{\mathrm{T}} W V\| = \|W\|. \qquad \square$$

Equality (2.28) relates the singular values of $\hat{W}$ with those of the matrices W_{11} and W_{22}: Each singular value of W_{11} is a singular value of $\hat{W}$. Similarly, each singular value of W_{22} is a singular value of $\hat{W}$. Conversely, each singular value of $\hat{W}$ is a singular value of W_{11} or a singular value of W_{22}. The last observation enables us to sharpen the results in certain cases. This is illustrated in Lemmas 8-11 below, which seem to be new. We will use these lemmas in the proofs of Theorems 18-21.

Lemma 8 (Pinching in Schatten *p*-norms)

$$\|W\|_p^p \ge \|\hat{W}\|_p^p = \|W_{11}\|_p^p + \|W_{22}\|_p^p. \tag{2.30}$$

Lemma 9 (Pinching in the trace norm)

$$\|W\|_{\mathrm{tr}} \ge \|\hat{W}\|_{\mathrm{tr}} = \|W_{11}\|_{\mathrm{tr}} + \|W_{22}\|_{\mathrm{tr}}. \tag{2.31}$$

Lemma 10 (Pinching in the spectral norm)

$$\|W\|_{\mathrm{sp}} \ge \|\hat{W}\|_{\mathrm{sp}} = \max\left\{\|W_{11}\|_{\mathrm{sp}}, \|W_{22}\|_{\mathrm{sp}}\right\}. \tag{2.32}$$

Lemma 11 (Pinching in Ky Fan *k*-norms) *Let* k_1 *and* k_2 *be a pair of positive integers that satisfy*

$$1 \le k_1 \le q \quad \text{and} \quad 1 \le k_2 \le n-q.$$

Then

$$\|W\|_{(k_1+k_2)} \ge \|\hat{W}\|_{(k_1+k_2)} \ge \|W_{11}\|_{(k_1)} + \|W_{22}\|_{(k_2)}. \tag{2.33}$$

Proof. The sum $\|W_{11}\|_{(k_1)} + \|W_{22}\|_{(k_2)}$ is formed from k_1+k_2 singular values of $\hat{W}$, while the sum defined by $\|\hat{W}\|_{(k_1+k_2)}$ is composed from the k_1+k_2 largest singular values of $\hat{W}$. □

The next tools consider the problem of approximating one matrix by another matrix of lower rank. Let $A \in \mathbb{R}^{m\times n}$ by a given matrix with SVD that satisfies (2.1)-(2.8). Let $1 \le k < n$ be a given integer, and let

$$\mathbb{B}_k = \left\{B \middle| B \in \mathbb{R}^{m\times n} \text{ and } \operatorname{rank}(B) \le k\right\}$$

denote the related set of low-rank matrices. Then here we seek a matrix $B \in \mathbb{B}$ that is nearest to A in a certain matrix norm. The difficulty stems from the fact that $\mathbb{B}_k$ is not a convex set. Let T_k denote a rank-*k* truncated SVD of A as defined in (2.8). Then the Eckart-Young theorem [26] says that T_k solves this problem in the Frobenius norm. The extension of this result to any unitarily invariant norm is due to Mirsky [27]. (Recall that T_k is not always unique. In such cases the nearest matrix is not unique.) A detailed statement of these assertions is given below. For recent discussions and proofs see [25].

Theorem 12 (Eckart-Young) *The inequality*

$$\|A-B\|_F^2 \ge \sum_{j=k+1}^{n} \sigma_j^2$$

holds for any matrix $B \in \mathbb{B}_k$. *Moreover, the matrix* T_k *solves the problem*

$$\begin{array}{ll} \text{minimize} & F(B) = \|A-B\|_F^2 \\ \text{subject to} & B \in \mathbb{B}_k, \end{array}$$

giving the optimal value of

$$\|A-T_k\|_F^2 = \left\|\sum_{j=k+1}^{n} \sigma_j \mathbf{u}_j \mathbf{v}_j^{\mathrm{T}}\right\|_F^2 = \sum_{j=k+1}^{n} \sigma_j^2.$$

Theorem 13 (Mirsky) *Let* $\|\cdot\|$ *be any unitarily invariant norm on* $\mathbb{R}^{m\times n}$. *Then the inequality*

$$\|A-B\| \ge \|A-T_k\|$$

holds for any matrix $B \in \mathbb{B}_k$. *In other words, the matrix* T_k *solves the problem*

$$\begin{array}{ll} \text{minimize} & \mu(B) = \|A-B\| \\ \text{subject to} & B \in \mathbb{B}_k. \end{array}$$

3. Positive Approximants of Symmetric Matrices

In this section we consider the solution of problem (1.3). Since $\|\cdot\|$ is a unitarily invariant norm, the spectral decomposition (1.2) enables us to convert (1.3) into the simpler form

$$\begin{aligned} &\text{minimize} && F(S) = \|\Lambda - S\| \\ &\text{subject to} && S \in \mathbb{S}_n^+ , \end{aligned} \tag{3.1}$$

whose solution provides a solution of (1.3).

Theorem 14 *Let the matrix* Λ_q *be defined as in (1.5). Then* Λ_q *solves (3.1) in any unitarily invariant norm.*

Proof. Let the diagonal matrix D_q be defined by the equality

$$\Lambda + D_q = \Lambda_q.$$

That is

$$D_q = \operatorname{diag}\left\{0,\cdots,0,\left|\lambda_{q+1}\right|,\cdots,\left|\lambda_n\right|\right\}. \tag{3.2}$$

Let $S = (s_{ij})$ be some matrix in $\mathbb{S}_n^+$ and let the matrix $W = (w_{ij}) \in \mathbb{R}^{n\times n}$ be defined by the equality

$$\Lambda + W = S. \tag{3.3}$$

Then the proof is concluded by showing that

$$\|W\| \ge \|D_q\|. \tag{3.4}$$

Let the diagonal matrix

$$W_q = \operatorname{diag}\left\{0,\cdots,0, w_{q+1,q+1},\cdots, w_{nn}\right\} \tag{3.5}$$

be obtained from the last $n-q$ diagonal entries of W. Then Corollary 5 implies that

$$\|W\| \ge \|W_q\|. \tag{3.6}$$

On the other hand, since $S \ge 0$, the diagonal entries of S are non-negative, which implies the inequalities

$$w_{jj} \ge \left|\lambda_j\right| \quad \text{for } j = q+1,\cdots,n, \tag{3.7}$$

and

$$\|W_q\| \ge \|D_q\|. \tag{3.8}$$

Now combining (3.6) and (3.8) gives (3.4) □

Theorem 14 is not new, e.g. ([8], p. 277) and [9]. However, the current proof is simple and short. In the next sections we extend these arguments to derive low-rank approximants.

4. Low-Rank Positive Approximants in the Frobenius Norm

In this section we consider the solution of problem (1.1) in the Frobenius norm. As before, the spectral decomposition (1.2) can be used to "diagonalize" the problem and the actual problem to solve has the form

$$\begin{aligned} &\text{minimize} && F(S) = \|\Lambda - S\|_F^2 \\ &\text{subject to} && S \in \mathbb{S}_{n,\ell}^+ . \end{aligned} \tag{4.1}$$

Theorem 15 *Let the matrix* Λ_ℓ *be defined as in* (1.6). *Then this matrix solves* (4.1)

Proof. Let the diagonal matrix D_ℓ be defined by the equality

$$\Lambda + D_\ell = \Lambda_\ell.$$

That is,

$$D_\ell = \operatorname{diag}\left\{0,\cdots,0,-\lambda_{\ell+1},\cdots,-\lambda_q,\left|\lambda_{q+1}\right|,\cdots,\left|\lambda_n\right|\right\}, \tag{4.2}$$

and

$$\|D_\ell\|_F^2 = \sum_{j=\ell+1}^{q} \lambda_j^2 + \sum_{j=q+1}^{n} \lambda_j^2. \tag{4.3}$$

Let $S = (s_{ij})$ be some matrix in $\mathbb{S}_{n,\ell}^{+}$ and let the matrix $W = (w_{ij}) \in \mathbb{R}^{n\times n}$ be defined by the equality

$$\Lambda + W = S. \tag{4.4}$$

Then the proof is concluded by showing that

$$\|W\|_F^2 \ge \|D_\ell\|_F^2. \tag{4.5}$$

This aim is achieved by considering a partition of W and S in the form

$$W = \left(\begin{array}{c|c} W_{11} & W_{12} \\ \hline W_{21} & W_{22} \end{array}\right) \quad \text{and} \quad S = \left(\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array}\right) \tag{4.6}$$

where W_{11} and S_{11} are $q\times q$ matrices, while W_{22} and S_{22} are $(n-q)\times(n-q)$ matrices. Then, clearly,

$$\|W\|_F^2 \ge \|W_{11}\|_F^2 + \|W_{22}\|_F^2. \tag{4.7}$$

Also, as before, since S is a positive semidefinite matrix it has non-negative diagonal entries, which implies the inequalities

$$w_{jj} \ge |\lambda_j| \quad \text{for } j = q+1,\cdots,n \tag{4.8}$$

and

$$\|W_{22}\|_F^2 \ge \sum_{j=q+1}^{n} w_{jj}^2 \ge \sum_{j=q+1}^{n} \lambda_j^2. \tag{4.9}$$

It is left, therefore, to show that

$$\|W_{11}\|_F^2 \ge \sum_{j=\ell+1}^{q} \lambda_j^2. \tag{4.10}$$

Observe that the matrices W_{11} and S_{11} are related by the equality

$$\Lambda_{11} + W_{11} = S_{11} \tag{4.11}$$

where

$$\Lambda_{11} = \operatorname{diag}\{\lambda_1,\cdots,\lambda_q\} \in \mathbb{R}^{q\times q} \tag{4.12}$$

and

$$\lambda_1 \ge \cdots \ge \lambda_q > 0. \tag{4.13}$$

Moreover, since S_{11} is a principal submatrix of S,

$$\operatorname{rank}(S_{11}) \le \operatorname{rank}(S) \le \ell. \tag{4.14}$$

Hence from the Eckart-Young theorem we obtain that

$$\|W_{11}\|_F^2 = \|\Lambda_{11} - S_{11}\|_F^2 \ge \sum_{j=\ell+1}^{q} \lambda_j^2. \qquad \square \tag{4.15}$$

Corollary 16 *Let* X *be a given real symmetric* $n\times n$ *matrix with the spectral decomposition* (1.2). *Then the matrix*

$$X_\ell = Q\Lambda_\ell Q^{\mathrm{T}} \tag{4.16}$$

solves the problem

$$\begin{aligned} &\text{minimize} && F(S) = \|X - S\|_F^2 \\ &\text{subject to} && S \in \mathbb{S}_{n,\ell}^{+}. \end{aligned} \tag{4.17}$$

Corollary 17 *Let* $A \in \mathbb{R}^{n\times n}$ *be a given matrix, let the matrix* $X = \left(A + A^{\mathrm{T}}\right)/2$ *have the spectral decomposition* (1.2), *and let the matrix* X_ℓ *be defined in* (4.16). *Then* X_ℓ *solves the problem*

$$\begin{aligned} &\text{minimize} \quad G(S) = \|A - S\|_F^2 \\ &\text{subject to} \quad S \in \mathbb{S}_{n,\ell}^{+}. \end{aligned} \tag{4.18}$$

5. Low-Rank Positive Approximants in the Schatten *p*-Norm

Let the diagonal matrix Λ be obtained from the spectral decomposition (1.2). In this section we consider the problem

$$\begin{aligned} &\text{minimize} \quad F(S) = \|\Lambda - S\|_p^p \\ &\text{subject to} \quad S \in \mathbb{S}_{n,\ell}^{+}. \end{aligned} \tag{5.1}$$

Theorem 18 *Let the matrix* Λ_ℓ *be defined in* (1.6). *Then this matrix solves* (5.1)

Proof. Let the matrices D_ℓ, W, and S be defined as in the proof of Theorem 15. Then here it is necessary to prove that

$$\|W\|_p^p \geq \|D_\ell\|_p^p, \tag{5.2}$$

where

$$\|D_\ell\|_p^p = \sum_{j=\ell+1}^{q} \lambda_j^p + \sum_{j=q+1}^{n} |\lambda_j|^p. \tag{5.3}$$

Let W and S be partitioned as in (4.6). Then from Lemma 8 we have

$$\|W\|_p^p \geq \|W_{11}\|_p^p + \|W_{22}\|_p^p. \tag{5.4}$$

Now Theorem 4 and (4.8) imply

$$\|W_{22}\|_p^p \geq \|\operatorname{diag}(W_{22})\|_p^p = \sum_{j=q+1}^{n} |w_{jj}|^p \geq \sum_{j=q+1}^{n} |\lambda_j|^p, \tag{5.5}$$

while applying Mirsky theorem on (4.11)-(4.14) gives

$$\|W_{11}\|_p^p = \|\Lambda_{11} - S_{11}\|_p^p \geq \sum_{j=\ell+1}^{q} \lambda_j^p. \tag{5.6}$$

Finally substituting (5.5) and (5.6) into (5.4) gives (5.2). □

6. Low-Rank Positive Approximants in the Trace Norm

Using the former notations, here we consider the problem

$$\begin{aligned} &\text{minimize} \quad F(S) = \|\Lambda - S\|_{\mathrm{tr}} \\ &\text{subject to} \quad S \in \mathbb{S}_{n,\ell}^{+}. \end{aligned} \tag{6.1}$$

Theorem 19 *The matrix* Λ_ℓ *solves* (6.1).

Proof. It is needed to show that

$$\|W\|_{\mathrm{tr}} \geq \|D_\ell\|_{\mathrm{tr}} \tag{6.2}$$

where

$$\|D_\ell\|_{\mathrm{tr}} = \sum_{j=\ell+1}^{q} |\lambda_j| + \sum_{j=q+1}^{n} |\lambda_j|. \tag{6.3}$$

The use of Lemma 9 yields

$$\|W\|_{\mathrm{tr}} \geq \|W_{11}\|_{\mathrm{tr}} + \|W_{22}\|_{\mathrm{tr}}. \tag{6.4}$$

Here Theorem 4 and (4.8) imply the inequalities

$$\|W_{22}\|_{\mathrm{tr}} \ge \|\mathrm{diag}(W_{22})\|_{\mathrm{tr}} = \sum_{j=q+1}^{n} |w_{jj}| \ge \sum_{j=q+1}^{n} |\lambda_j|, \tag{6.5}$$

and Mirsky theorem gives

$$\|W_{11}\|_{\mathrm{tr}} = \|\Lambda_{11} - S_{11}\|_{\mathrm{tr}} \ge \sum_{j=\ell+1}^{q} |\lambda_n|, \tag{6.6}$$

which completes the proof. □

7. Low-Rank Positive Approximants in the Spectral Norm

In this case we consider the problem

$$\begin{aligned} &\text{minimize} \quad F(S) = \|\Lambda - S\|_{\mathrm{sp}} \\ &\text{subject to} \quad S \in \mathbb{S}^{+}_{n,\ell}. \end{aligned} \tag{7.1}$$

Theorem 20 *The matrix* Λ_ℓ *solves* (7.1).

Proof. Following the former notations and arguments, here it is needed to show that

$$\|W\|_{\mathrm{sp}} \ge \|D_\ell\|_{\mathrm{sp}}.$$

Define

$$\alpha = \max_{j=\ell+1,\cdots,q} |\lambda_j| \quad \text{and} \quad \beta = \max_{j=q+1,\cdots,n} |\lambda_j|.$$

Then, clearly,

$$\|D_\ell\|_{\mathrm{sp}} = \max_{j=\ell+1,\cdots,n} |\lambda_j| = \max\{\alpha, \beta\}.$$

Using Lemma 10 we see that

$$\|W\|_{\mathrm{sp}} \ge \max\left\{\|W_{11}\|_{\mathrm{sp}}, \|W_{22}\|_{\mathrm{sp}}\right\}.$$

Now Theorem 4 and (4.8) imply

$$\|W_{22}\|_{\mathrm{sp}} \ge \|\mathrm{diag}(W_{22})\|_{\mathrm{sp}} = \max_{j=q+1,\cdots,n} |w_{jj}| \ge \max_{j=q+1,\cdots,n} |\lambda_j| = \beta,$$

while Mirsky theorem gives

$$\|W_{11}\|_{\mathrm{sp}} = \|\Lambda - S_{11}\|_{\mathrm{sp}} \ge \max_{j=\ell+1,\cdots,q} |\lambda_j| = \alpha.$$ □

8. Unitarily Invariant Norms

Let the diagonal matrices Λ and Λ_ℓ be defined as in Section 1, and let $\|\cdot\|$ denote any unitarily invariant norm on $\mathbb{R}^{n\times n}$. Below we will show that Λ_ℓ solves the problem

$$\begin{aligned} &\text{minimize} \quad F(S) = \|\Lambda - S\| \\ &\text{subject to} \quad S \in \mathbb{S}^{+}_{n,\ell}. \end{aligned} \tag{8.1}$$

The derivation of this result is based on the following assertion, which considers Ky Fan k-norms.

Theorem 21 *The matrix* Λ_ℓ *solves the problem*

$$\begin{aligned} &\text{minimize} \quad F(S) = \|\Lambda - S\|_{(k)} \\ &\text{subject to} \quad S \in \mathbb{S}^{+}_{n,\ell} \end{aligned} \tag{8.2}$$

for $k = 1,\cdots,n$.

Proof. We have already proved that Λ_ℓ solves (8.2) for the spectral norm $(k=1)$ and the trace norm $(k=n)$. Hence it is left to consider the case when $2 \le k \le n-1$. As before, the diagonal matrix D_ℓ is defined in (4.2), and the matrices S and W satisfy (4.4) as well as the partition (4.6). With these notations at hand it is needed to show that

$$\|W\|_{(k)} \ge \|D_\ell\|_{(k)}, \tag{8.3}$$

Let D_ℓ be partitioned in a similar way:

$$D_\ell = \left(\begin{array}{c|c} D_{11} & 0 \\ \hline 0 & D_{22} \end{array}\right), \tag{8.4}$$

where

$$D_{11} = \operatorname{diag}\left\{0,\cdots,0,-\lambda_{\ell+1},\cdots,-\lambda_q\right\} \in \mathbb{R}^{q\times q} \tag{8.5}$$

and

$$D_{22} = \operatorname{diag}\left\{\left|\lambda_{q+1}\right|,\cdots,\left|\lambda_n\right|\right\} \in \mathbb{R}^{(n-q)\times(n-q)}. \tag{8.6}$$

Then there are three different cases to consider.

The first case occurs when

$$\|D_\ell\|_{(k)} = \|D_{11}\|_{(k)}. \tag{8.7}$$

Here Theorem 3 implies the inequalities

$$\|W\|_{(k)} \ge \|W_{11}\|_{(k)}, \tag{8.8}$$

while from (4.11)-(4.14) and Mirsky theorem we obtain

$$\|W_{11}\|_{(k)} \ge \|D_{11}\|_{(k)} = \|D_\ell\|_{(k)}, \tag{8.9}$$

which proves (8.3).

The second case occurs when

$$\|D_\ell\|_{(k)} = \|D_{22}\|_{(k)}. \tag{8.10}$$

Here Theorem 3 implies

$$\|W\|_{(k)} \ge \|W_{22}\|_{(k)}, \tag{8.11}$$

while Theorem 4 and the inequalities (4.8) give

$$\|W_{22}\|_{(k)} \ge \left\|\operatorname{diag}(W_{22})\right\|_{(k)} \ge \|D_{22}\|_{(k)} = \|D_\ell\|_{(k)}, \tag{8.12}$$

which proves (8.3).

The third case occurs when neither (8.7) nor (8.10) hold. In this case there exist two positive integers, k_1 and k_2, such that

$$k_1 + k_2 = k \tag{8.13}$$

and

$$\|D_\ell\|_{(k)} = \|D_{11}\|_{(k_1)} + \|D_{22}\|_{(k_2)}. \tag{8.14}$$

Now Lemma 11 shows that

$$\|W\|_{(k)} \ge \|W_{11}\|_{(k_1)} + \|W_{22}\|_{(k_2)}. \tag{8.15}$$

A further use of (4.11)-(4.14) and Mirsky theorem give

$$\|W_{11}\|_{(k_1)} \ge \|D_{11}\|_{(k_1)}, \tag{8.16}$$

and from Theorem 4 and (4.8) we obtain

$$\|W_{22}\|_{(k_2)} \geq \|\mathrm{diag}(W_{22})\|_{(k_2)} \geq \|D_{22}\|_{(k_2)}. \tag{8.17}$$

Hence by substituting (8.16) and (8.17) into (8.15) we get (8.3). □

The fact that (8.3) holds for $k = 1, \cdots, n$ means that the inequality

$$\|W\| \geq \|D_\ell\| \tag{8.18}$$

holds for any unitarily invariant norm. This observation is a direct consequence of Ky Fan dominance theorem. The last inequality proves our final results.

Theorem 22 *The matrix* Λ_ℓ *solves* (8.1) *in any unitarily invariant norm.*

Theorem 23 *Using the notations of Section* 1, *the matrix*

$$X_\ell = Q\Lambda_\ell Q^{\mathrm{T}}$$

solves (*1.1*) *in any unitarily invariant norm*.

9. Concluding Remarks

In view of Theorem 14 and Mirsky theorem, the observation that Λ_ℓ solves (8.1) is not surprising. However, as we have seen, the proof of this assertion is not straightforward. A key argument in the proof is the inequality (8.15), which is based on Lemma 11.

Once Theorem 22 is proved, it is possible to use this result to derive Theorems 15-18. Yet the direct proofs that we give clearly illustrate why these theorems work. In fact, the proof of Theorem 15 paves the way for the other proofs. Moreover, as Corollary 17 shows, when using the Frobenius norm we get stronger results: In this case we are able to compute a low-rank positive approximant of any matrix $A \in \mathbb{R}^{n\times n}$.

References

[1] Dax, A. (2014) Imputing Missing Entries of a Data Matrix: A Review. Techical Report, Hydrological Service of Israel.

[2] Trosset, M.W. (2000) Distance Matrix Completion by Numerical Optimization. *Computational Optimization and Applications*, **17**, 11-22. http://dx.doi.org/10.1023/A:1008722907820

[3] Higham, N.J. (1988) Computing a Nearest Symmetric Positive Semidefinite Matrix. *Linear Algebra and Its Applications*, **103**, 103-118. http://dx.doi.org/10.1016/0024-3795(88)90223-6

[4] Halmos, P.R. (1972) Positive Approximants of Operators. *Indiana University Mathematics Journal*, **21**, 951-960. http://dx.doi.org/10.1512/iumj.1972.21.21076

[5] Rogers, D.D. and Ward, J.D. (1981) C_p-Minimal Positive Approximants. *Acta Scientiarum Mathematicarum*, **43**, 109-115.

[6] Ando, T. (1985) Approximation in Trace Norm by Positive Semidefinite Matrices. *Linear Algebra and Its Applications*, **71**, 15-21. http://dx.doi.org/10.1016/0024-3795(85)90230-7

[7] Ando, T., Sekiguchi, T. and Suzuki, T. (1973) Approximation by Positive Operators. *Mathematische Zeitschrift*, **131**, 273-282. http://dx.doi.org/10.1007/BF01174903

[8] Bhatia, R. (1997) Matrix Analysis. Springer, New York. http://dx.doi.org/10.1007/978-1-4612-0653-8

[9] Bhatia, R. and Kittaneh, F. (1992) Approximation by Positive Operators. *Linear Algebra and Its Applications*, **161**, 1-9. http://dx.doi.org/10.1016/0024-3795(92)90001-Q

[10] Bouldin, R. (1973) Positive Approximants. *Transactions of the American Mathematical Society*, **177**, 391-403. http://dx.doi.org/10.1090/S0002-9947-1973-0317082-6

[11] Sekiguchi, T. (1976) Positive Approximants of Normal Operators. *Hokkaido Mathematical Journal*, **5**, 270-279. http://dx.doi.org/10.14492/hokmj/1381758677

[12] Bouldin, R. (1980) Best Approximation of a Normal Operator in the Schatten *p*-Norm. *Proceedings of the American Mathematical Society*, **80**, 277-282.

[13] Bouldin, R. (1987) Best Approximation of a Normal Operator in the Trace Norm. *Acta Scientiarum Mathematicarum*, **51**, 467-474.

[14] Bouldin, R., (1988) Best Approximation of a Nonnormal Operator in the Trace Norm. *Journal of Mathematical Analysis and Applications*, **132**, 102-113. http://dx.doi.org/10.1016/0022-247X(88)90046-7

[15] Laszkiewicz, B. and Zietak, K. (2008) Approximation by Matrices with Restricted Spectra. *Linear Algebra and Its Applications*, **428**, 1031-1040. http://dx.doi.org/10.1016/j.laa.2007.09.009

[16] Phillips, J. (1977) Nearest Normal Approximation for Certain Normal Operators. *Proceedings of the American Mathematical Society*, **67**, 236-240. http://dx.doi.org/10.1090/S0002-9939-1977-0458212-4

[17] Ruhe, A. (1987) Closest Normal Matrix Finally Found! *BIT Numerical Mathematics*, **27**, 585-598. http://dx.doi.org/10.1007/BF01937278

[18] Zietak, K. (1997) Strict Spectral Approximation of a Matrix and Some Related Problems. *Applicationes Mathematicae*, **24**, 267-280.

[19] Higham, N.J. (1989) Matrix Nearness Problems and Applications. In: Gover, M.J.C. and Barnett, S., Eds., *Applications of Matrix Theory*, Oxford University Press, Oxford, 1-27.

[20] Fan, K. (1951) Maximum Properties and Inequalities for the Eigenvalues of Completely Continuous Operators. *Proceedings of the National Academy of Sciences of the United States of America*, **37**, 760-766. http://dx.doi.org/10.1073/pnas.37.11.760

[21] Fan, K. and Hoffman, A.J. (1955) Some Metric Inequalities in the Space of Matrices. *Proceedings of the American Mathematical Society*, **6**, 111-116. http://dx.doi.org/10.1090/S0002-9939-1955-0067841-7

[22] Horn, R.A. and Johnson, C.R. (1991) Topics in Matrix Analysis. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9780511840371

[23] Marshall, A.W., Olkin, I. and Arnold, B.C. (2011) Inequalities: Theory of Majorization and Its Applications. Springer Series in Statistics, 2nd Edition, Springer, New York.

[24] Zhang, F. (1999) Matrix Theory: Basic Results and Techniques. Springer-Verlag, New York. http://dx.doi.org/10.1007/978-1-4757-5797-2

[25] Dax, A. (2010) On Extremum Properties of Orthogonal Quotient Matrices. *Linear Algebra and Its Applications*, **432**, 1234-1257. http://dx.doi.org/10.1016/j.laa.2009.10.034

[26] Eckart, C. and Young, G. (1936) The Approximation of One Matrix by Another of Lower Rank. *Psychometrika*, **1**, 211-218. http://dx.doi.org/10.1007/BF02288367

[27] Mirsky, L. (1960) Symmetric Gauge Functions and Unitarily Invariant Norms. *Quarterly Journal of Mathematics*, **11**, 50-59. http://dx.doi.org/10.1093/qmath/11.1.50