Numer. Math. Theor. Meth. Appl. Vol. **3**, No. 4, pp. 449-460
doi: 10.4208/nmtma.2010.m9013 November 2010

Some Properties of the Optimal Preconditioner and the Generalized Superoptimal Preconditioner

Hong-Kui Pang* and Xiao-Qing Jin

Department of Mathematics, University of Macau, Macao, China.

Received 6 August 2009; Accepted (in revised version) 31 August 2009

Available online 9 September 2010

Abstract. The optimal preconditioner and the superoptimal preconditioner were proposed in 1988 and 1992 respectively. They have been studied widely since then. Recently, Chen and Jin [6] extend the superoptimal preconditioner to a more general case by using the Moore-Penrose inverse. In this paper, we further study some useful properties of the optimal and the generalized superoptimal preconditioners. Several existing results are extended and new properties are developed.

AMS subject classifications: 65F10, 65F15, 65F35, 65F99

Key words: Optimal preconditioner, generalized superoptimal preconditioner, Moore-Penrose inverse, unitarily invariant norm, semi-stability, singular value.

1. Introduction

Given a unitary matrix $U \in \mathbb{C}^{n\times n}$, let

$$\mathscr{M}_U \equiv \{U^*\Lambda_n U \mid \Lambda_n \text{ is any } n\text{-by-}n \text{ diagonal matrix}\}. \tag{1.1}$$

The optimal preconditioner $c_U(A_n)$ is defined to be the minimizer of

$$\min_{W_n\in\mathscr{M}_U} \left\|A_n - W_n\right\|_F .$$

This preconditioner was first proposed in [5] and then extended in [3, 12]. Due to be very efficient for solving a large class of structured systems [2, 4, 13, 14], the optimal preconditioner $c_U(A_n)$ has been studied deeply and widely. Many useful properties of $c_U(A_n)$ have been found.

Besides using the minimizer of $\min_{W_n} \left\|A_n - W_n\right\|_F$ as a preconditioner, Tyrtyshnikov [17] proposed another preconditioner $t_U(A_n)$, called superoptimal preconditioner, which is defined to be the minimizer of

$$\min_{W_n} \|I - W_n^{-1}A_n\|_F,$$

*Corresponding author. *Email addresses:* `ya87402@umac.mo` (H.-K. Pang), `xqjin@umac.mo` (X.-Q. Jin)

where W_n runs over all nonsingular matrices in $\mathscr{M}_U$ defined as in (1.1), I denotes the identity matrix. Recent results demonstrate that the superoptimal preconditioner has good filtering capabilities when applied in signal/image processing [8,9].

Very recently, the definition of the superoptimal preconditioner is generalized by Chen and Jin [6] by using the Moore-Penrose inverse [1]. For any arbitrary matrix A_n, the generalized superoptimal preconditioner $t_U(A_n)$ is defined to be the minimizer of

$$\min_{W_n \in \mathscr{M}_U} \|I - W_n^{\dagger} A_n\|_F, \tag{1.2}$$

where $W_n^{\dagger}$ denotes the Moore-Penrose inverse of W_n. In [6], the authors give an explicit formula for this generalized superoptimal preconditioner and discuss its stability properties.

In this paper, we further study the optimal preconditioner and the generalized superoptimal preconditioner defined as in (1.2). The rest part of the paper is arranged as follows. In Section 2, we extend some existing results and develop some new properties of $c_U(A_n)$. In Section 3, the properties of $t_U(A_n)$ are discussed. The relation between the singular values of the optimal preconditioned matrix $c_U(A_n)^{\dagger} A_n$ and the superoptimal preconditioned matrix $t_U(A_n)^{\dagger} A_n$ is given in Section 4. Here, $c_U(A_n)^{\dagger} \equiv (c_U(A_n))^{\dagger}$ and $t_U(A_n)^{\dagger} \equiv (t_U(A_n))^{\dagger}$. Our results generalize some results presented in [15,16].

2. The optimal preconditioner $c_U(A_n)$

In this section, we discuss the properties of the optimal preconditioner $c_U(A_n)$. Let $\delta(E_n)$ denote the diagonal matrix whose diagonal is equal to the diagonal of the matrix E_n. We first introduce some lemmas and theorems which will be used later.

Lemma 2.1. (Lemma 3.5 in [14]; Theorem 1 in [16]) *Let $A_n \in \mathbb{C}^{n\times n}$ with $n \geq 1$ and U be any unitary matrix. Then*

(i) $c_U(A_n) \equiv U^* \delta(U A_n U^*) U$ *which is uniquely determined by* A_n.

(ii) $c_U(A_n^*) = c_U(A_n)^*$.

(iii) $c_U(B_n A_n) = B_n c_U(A_n)$, $c_U(A_n B_n) = c_U(A_n) B_n$, *for any* $B_n \in \mathscr{M}_U$.

Lemma 2.2. *Let $A_n \in \mathbb{C}^{n\times n}$ be partitioned as*

$$A_n = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \textit{and} \quad D_n = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}.$$

Then for any unitarily invariant norm $\|\cdot\|$, we have

$$\left\|D_n\right\| \leq \left\|A_n\right\|.$$

Proof. Let

$$Q_n = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}.$$

Then we have from $2D_n = A_n + Q_n A_n Q_n$,

$$\|D_n\| \le \|A_n\|.$$

This completes the proof of the lemma. □

Theorem 2.1. (Corollary 3.5.9 in [11]) *Let $A, B \in \mathbb{C}^{m\times n}$ with $m \ge n$. Then for any unitarily invariant norm $\|\cdot\|$ defined on $\mathbb{C}^{m\times n}$, we have*

$$\|A\| \le \|B\| \tag{2.1a}$$

if and only if

$$\|A\|_{(k)} \le \|B\|_{(k)}, \qquad k = 1, \cdots, n, \tag{2.1b}$$

where $\|A\|_{(k)} \equiv \sum_{j=k}^{n} \sigma_j(A)$ are called the Ky Fan k-norms, and $\sigma_1(A) \le \sigma_2(A) \le \cdots \le \sigma_n(A)$ are the singular values of A.

Note that the Ky Fan k-norms are also unitarily invariant. We always assume that the singular values are arranged in the non-decreasing order in this paper. The following two lemmas are useful to study the properties of the optimal preconditioner $c_U(A_n)$.

Lemma 2.3. *Let $A_n \in \mathbb{C}^{n\times n}$ with $n \ge 1$. For any unitarily invariant norm $\|\cdot\|$, we have*

$$\left\|\delta(A_n)\right\| \le \left\|A_n\right\|. \tag{2.2}$$

Proof. For any $A_n \in \mathbb{C}^{n\times n}$, there exists a permutation matrix P such that entries of the main diagonal of $B_n = (b_{ij})_{n\times n} = PA_nP^*$ satisfy

$$|b_{11}| \le |b_{22}| \le \cdots \le |b_{nn}|.$$

Moreover, for any unitarily invariant norm $\|\cdot\|$,

$$\left\|\delta(B_n)\right\| = \left\|\delta(PA_nP^*)\right\| = \left\|P\delta(A_n)P^*\right\| = \left\|\delta(A_n)\right\|, \quad \left\|B_n\right\| = \left\|PA_nP^*\right\| = \left\|A_n\right\|.$$

Therefore, without loss of generality, we assume that A_n is a matrix with entries of the main diagonal satisfying

$$|a_{11}| \le |a_{22}| \le \cdots \le |a_{nn}|.$$

Obviously, $|a_{11}|, |a_{22}|, \cdots, |a_{nn}|$ are the singular values of $\delta(A_n)$. We partition the matrix A_n as in the following form

$$A_n = \begin{pmatrix} A_{k,k} & A_{k,n-k} \\ A_{n-k,k} & A_{n-k,n-k} \end{pmatrix}, \qquad 1 \le k \le n,$$

where $A_{k,k}$ is the k-by-k leading principal submatrix of A_n.

Next we prove this lemma by induction. When $k = 2$, it is easy to see by Lemma 2.2 that for any unitarily invariant norm $\|\cdot\|$,

$$\|\delta(A_{2,2})\| \le \|A_{2,2}\|.$$

Assume that for any unitarily invariant norm $\|\cdot\|$, the inequality $\|\delta(A_{k,k})\| \le \|A_{k,k}\|$ holds for a constant k with $2 \le k < n$. We then have by Theorem 2.1,

$$\sum_{i=l}^{k} |a_{ii}| \le \sum_{i=l}^{k} \sigma_i(A_{k,k}), \qquad l = 1, \cdots, k. \tag{2.3}$$

For the case of $k+1$, we partition the leading principal submatrix $A_{k+1,k+1}$ as in the following form

$$A_{k+1,k+1} = \begin{pmatrix} A_{k,k} & \alpha_k \\ \beta_k & a_{k+1,k+1} \end{pmatrix},$$

where $\alpha_k, \beta_k^* \in \mathbb{C}^k$. Let

$$D_{k+1} = \begin{pmatrix} A_{k,k} & 0 \\ 0 & a_{k+1,k+1} \end{pmatrix}.$$

We have from Lemma 2.2,

$$\|D_{k+1}\| \le \|A_{k+1,k+1}\|,$$

for any unitarily invariant norm $\|\cdot\|$. Thus, the following inequalities hold by Theorem 2.1,

$$\sum_{i=l}^{k+1} \sigma_i(D_{k+1}) \le \sum_{i=l}^{k+1} \sigma_i(A_{k+1,k+1}), \qquad l = 1, \cdots, k+1. \tag{2.4}$$

Note that $\sigma_1(A_{k,k})$, $\sigma_2(A_{k,k})$, $\cdots$, $\sigma_k(A_{k,k})$ and $|a_{k+1,k+1}|$ are the singular values of D_{k+1}. By (2.3) and (2.4), we obtain

$$|a_{k+1,k+1}| \le \sigma_{k+1}(D_{k+1}) \le \sigma_{k+1}(A_{k+1,k+1}), \tag{2.5}$$

and

$$\begin{aligned} \sum_{i=l}^{k+1} |a_{ii}| &\le \sum_{i=l}^{k} \sigma_i(A_{k,k}) + |a_{k+1,k+1}| \\ &\le \sum_{i=l}^{k+1} \sigma_i(D_{k+1}) \le \sum_{i=l}^{k+1} \sigma_i(A_{k+1,k+1}), \quad l = 1, \cdots, k. \end{aligned} \tag{2.6}$$

By Theorem 2.1 again, we have $\|\delta(A_{k+1,k+1})\| \le \|A_{k+1,k+1}\|$ for any unitarily invariant norm $\|\cdot\|$. Finally, we have (2.2) by using induction. $\square$

Lemma 2.4. *Let $A_n \in \mathbb{C}^{n\times n}$ be a Hermitian positive definite matrix and $B_n \in \mathbb{C}^{n\times n}$. Then*

$$\delta(B_nA_n^{-1}B_n^*) \geq \delta(B_n)\delta(A_n)^{-1}\delta(B_n^*), \tag{2.7}$$

where $\delta(A_n)^{-1} \equiv (\delta(A_n))^{-1}$, "$\geq$" means that $\delta(B_nA_n^{-1}B_n^*) - \delta(B_n)\delta(A_n)^{-1}\delta(B_n^*)$ is a positive semi-definite matrix.

Proof. Since

$$\begin{pmatrix} I & 0 \\ -B_nA_n^{-1} & I \end{pmatrix}\begin{pmatrix} A_n & B_n^* \\ B_n & B_nA_n^{-1}B_n^* \end{pmatrix}\begin{pmatrix} I & -A_n^{-1}B_n^* \\ 0 & I \end{pmatrix} = \begin{pmatrix} A_n & 0 \\ 0 & 0 \end{pmatrix},$$

and A_n is Hermitian positive definite, we know that by *Sylvester's law of inertia*,

$$\begin{pmatrix} A_n & B_n^* \\ B_n & B_nA_n^{-1}B_n^* \end{pmatrix}$$

is positive semi-definite. Let $d_k(M) = m_{k,k}$ be the kth element of the main diagonal of a matrix M. Note that for each k with $1 \leq k \leq n$, the matrix

$$\begin{pmatrix} d_k(A_n) & d_k(B_n^*) \\ d_k(B_n) & d_k(B_nA_n^{-1}B_n^*) \end{pmatrix}$$

is a 2-by-2 principal submatrix of

$$\begin{pmatrix} A_n & B_n^* \\ B_n & B_nA_n^{-1}B_n^* \end{pmatrix}.$$

Then

$$\begin{pmatrix} d_k(A_n) & d_k(B_n^*) \\ d_k(B_n) & d_k(B_nA_n^{-1}B_n^*) \end{pmatrix} \tag{2.8}$$

is also positive semi-definite. Consequently, we have

$$d_k(A_n)d_k(B_nA_n^{-1}B_n^*) \geq d_k(B_n)d_k(B_n^*). \tag{2.9}$$

Note that $\delta(A_n) > 0$ because A_n is Hermitian positive definite. Hence,

$$d_k(B_nA_n^{-1}B_n^*) \geq d_k(B_n)d_k(A_n)^{-1}d_k(B_n^*), \qquad k = 1, \cdots, n,$$

which yields (2.7). □

Now, by using the above lemmas, we can prove the following theorems which generalize some results in [14, 16].

Theorem 2.2. *Let $A_n \in \mathbb{C}^{n\times n}$ with $n \geq 1$. Then*

(i) *We have*

$$\sum_{i=k}^{n}\sigma_i(c_U(A_n)) \le \sum_{i=k}^{n}\sigma_i(A_n), \qquad k=1,\cdots,n, \tag{2.10}$$

where $\sigma_i(c_U(A_n))$ and $\sigma_i(A_n)$ denote the singular values of matrices $c_U(A_n)$ and A_n respectively.

(ii) *If A_n is a Hermitian positive definite matrix and $B_n \in \mathbb{C}^{n\times n}$. Then*

$$c_U(B_nA_n^{-1}B_n^*) \ge c_U(B_n)c_U(A_n)^{-1}c_U(B_n^*). \tag{2.11}$$

In particular, we have

$$c_U(A_n^{-1}) \ge c_U(A_n)^{-1}. \tag{2.12}$$

Proof. For (i), by Lemma 2.1 (i),

$$c_U(A_n) = U^*\delta(UA_nU^*)U.$$

For any unitarily invariant norm $\|\cdot\|$, by Lemma 2.3, we obtain

$$\left\|c_U(A_n)\right\| = \left\|\delta(UA_nU^*)\right\| \le \left\|UA_nU^*\right\| = \left\|A_n\right\|. \tag{2.13}$$

Thus, (i) holds by using Theorem 2.1. For (ii), since

$$\begin{aligned} c_U(B_nA_n^{-1}B^*) &= U^*\delta(UB_nA_n^{-1}B_n^*U^*)U = U^*\delta((UB_nU^*)(UA_nU^*)^{-1}(UB_nU^*)^*)U,\\ c_U(B_n)c_U(A_n)^{-1}c_U(B_n^*) &= U^*\delta(UB_nU^*)\delta(UA_nU^*)^{-1}\delta((UB_nU^*)^*)U, \end{aligned}$$

we only need to prove that

$$\delta((UB_nU^*)(UA_nU^*)^{-1}(UB_nU^*)^*) \ge \delta(UB_nU^*)\delta(UA_nU^*)^{-1}\delta((UB_nU^*)^*). \tag{2.14}$$

By Lemma 2.4, this inequality holds for A_n being Hermitian positive definite. In particular, when $B_n = I$, we have (2.12). □

Theorem 2.3. *Let $A_n \in \mathbb{C}^{n\times n}$ and $B_n \in \mathscr{M}_U$. If* $\mathrm{rank}(c_U(A_n)) = \mathrm{rank}(c_U(B_nA_n))$, *then we have*

$$B_nc_U(B_nA_n)^{\dagger} = c_U(A_n)^{\dagger} = B_nc_U(A_nB_n)^{\dagger}. \tag{2.15}$$

Proof. Let $B_n = U^*\Lambda_nU \in \mathscr{M}_U$ where $\Lambda_n = \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$. By Lemma 2.1, we know that

$$\mathrm{rank}(c_U(A_n)) = \mathrm{rank}(\delta(UA_nU^*)), \tag{2.16a}$$

$$\mathrm{rank}(c_U(B_nA_n)) = \mathrm{rank}(B_nc_U(A_n)) = \mathrm{rank}(\Lambda_n\delta(UA_nU^*)). \tag{2.16b}$$

Therefore, the given condition of $\mathrm{rank}(c_U(A_n)) = \mathrm{rank}(c_U(B_nA_n))$ is equivalent to

$$\mathrm{rank}(\delta(UA_nU^*)) = \mathrm{rank}(\Lambda_n\delta(UA_nU^*)). \tag{2.17}$$

Let $\delta(UA_nU^*) = \text{diag}(a_1, a_2, \cdots, a_n)$. We have $\delta(UA_nU^*)^\dagger = \text{diag}(a_1^\dagger, a_2^\dagger, \cdots, a_n^\dagger)$ where

$$a_k^\dagger = \begin{cases} a_k^{-1}, & \text{if } a_k \neq 0, \\ 0, & \text{if } a_k = 0, \end{cases}$$

for $k = 1, \cdots, n$. The equality (2.17) implies that $\lambda_k \neq 0$ if $a_k \neq 0$. Thus,

$$\Lambda_n \Lambda_n^\dagger \delta(UA_nU^*)^\dagger = \delta(UA_nU^*)^\dagger.$$

Hence, we have by Lemma 2.1,

$$\begin{aligned} B_n c_U(B_nA_n)^\dagger &= B_n(B_n c_U(A_n))^\dagger = U^*\Lambda_n U(U^*\Lambda_n \delta(UA_nU^*)U)^\dagger \\ &= U^*\Lambda_n(\Lambda_n\delta(UA_nU^*))^\dagger U = U^*(\Lambda_n\Lambda_n^\dagger\delta(UA_nU^*)^\dagger)U \\ &= U^*\delta(UA_nU^*)^\dagger U = (U^*\delta(UA_nU^*)U)^\dagger = c_U(A_n)^\dagger. \end{aligned}$$

Similarly, we can prove $c_U(A_n)^\dagger = B_n c_U(A_nB_n)^\dagger$. □

3. The generalized superoptimal preconditioner $t_U(A_n)$

In this section, we give some results of the generalized superoptimal preconditioner $t_U(A_n)$ (see (1.2)) introduced by Chen and Jin [6]. The following definition is needed.

Definition 3.1. *A matrix is said to be stable if the real parts of all the eigenvalues are negative. A matrix is said to be semi-stable if the real parts of all the eigenvalues are not larger than zero.*

Under certain conditions, $t_U(A_n)$ satisfies the following explicit formula.

Theorem 3.1. (Corollary 1 in [6]) *Let $A_n \in \mathbb{C}^{n\times n}$ and U be any unitary matrix with $\mathbf{u}_k$ being the kth row of U. Then $\mathbf{u}_kA_n \neq 0$ for any k if and only if $t_U(A_n)$ is uniquely determined by A_n. In this case,*

$$t_U(A_n) = c_U(A_nA_n^*)c_U(A_n^*)^\dagger.$$

Moreover,

$$t_U(A_n)^\dagger = c_U(A_nA_n^*)^{-1}c_U(A_n^*).$$

By means of this explicit formula, it is easy to verify the following theorem.

Theorem 3.2. *Let $A_n \in \mathbb{C}^{n\times n}$ and U be any unitary matrix with $\mathbf{u}_k$ being the kth row of U. If $\mathbf{u}_kA_n \neq 0$ for $k = 1, \cdots, n$, we then have,*

(i) $t_U(\alpha A_n) = \alpha t_U(A_n)$ *for all* $\alpha \in \mathbb{C}\backslash\{0\}$.

(ii) $t_U(A_n^*) = t_U(A_n)^*$ *for any normal matrix* A_n.

(iii) $t_U(B_nA_n) = B_n t_U(A_n)$ *for any invertible matrix* $B_n \in \mathscr{M}_U$.

(iv) $t_U(A_n)$ *is stable (semi-stable) if and only if* $c_U(A_n)$ *is stable (semi-stable).*

Proof. When $\mathbf{u}_k A_n \neq 0$ for $k = 1, \cdots, n$, we have from Theorem 3.1,

$$t_U(A_n) = c_U(A_n A_n^*) c_U(A_n^*)^\dagger. \tag{3.1}$$

For (i), we obtain by (3.1),

$$\begin{aligned} t_U(\alpha A_n) &= c_U(\alpha A_n \cdot \overline{\alpha} A_n^*) c_U(\overline{\alpha} A_n^*)^\dagger = \alpha \cdot \overline{\alpha} \cdot \overline{\alpha}^{-1} c_U(A_n A_n^*) c_U(A_n^*)^\dagger \\ &= \alpha c_U(A_n A_n^*) c_U(A_n^*)^\dagger = \alpha t_U(A_n). \end{aligned}$$

For (ii), since A_n is normal, we have by (3.1) and Lemma 2.1 (ii),

$$\begin{aligned} t_U(A_n)^* &= (c_U(A_n A_n^*) c_U(A_n^*)^\dagger)^* = (c_U(A_n^*)^\dagger)^* (c_U(A_n A_n^*))^* \\ &= c_U(A_n)^\dagger c_U(A_n A_n^*) = c_U(A_n^* A_n) c_U(A_n)^\dagger. \end{aligned}$$

From $\mathbf{u}_k A_n \neq 0$, we obtain that $\mathbf{u}_k A_n A_n^* \mathbf{u}_k^* > 0$ for $k = 1, \cdots, n$. Since A_n is normal, we have $\mathbf{u}_k A_n^* A_n \mathbf{u}_k^* > 0$, which implies $\mathbf{u}_k A_n^* \neq 0$ for $k = 1, \cdots, n$. Thus, we have by Theorem 3.1,

$$t_U(A_n^*) = c_U(A_n^* A_n) c_U(A_n)^\dagger.$$

Hence,

$$t_U(A_n^*) = t_U(A_n)^*.$$

For (iii), let $B_n = U^* \Lambda_n U \in \mathscr{M}_U$ with $\Lambda_n = \operatorname{diag}(b_1, b_2, \cdots, b_n)$. Since B_n is invertible, $b_k \neq 0$ for $k = 1, \cdots, n$. When $\mathbf{u}_k A_n \neq 0$ for $k = 1, \cdots, n$, it follows that

$$\mathbf{u}_k B_n A_n = \mathbf{u}_k U^* \Lambda_n U A_n = b_k \mathbf{u}_k A_n \neq 0.$$

Thus we obtain by Theorem 3.1,

$$t_U(B_n A_n) = c_U(B_n A_n A_n^* B_n^*) c_U(A_n^* B_n^*)^\dagger,$$

and then by Lemma 2.1 (iii),

$$t_U(B_n A_n) = B_n c_U(A_n A_n^*) B_n^* c_U(A_n^* B_n^*)^\dagger. \tag{3.2}$$

From the invertibility of B_n and Lemma 2.1 again,

$$\operatorname{rank}(c_U(B_n^* A_n^*)) = \operatorname{rank}(B_n^* c_U(A_n^*)) = \operatorname{rank}(c_U(A_n^*)).$$

Hence, we have by Theorem 2.3 and (3.2),

$$t_U(B_n A_n) = B_n c_U(A_n A_n^*) c_U(A_n^*)^\dagger = B_n t_U(A_n).$$

For (iv), it follows from (3.1) and Lemma 2.1 (i) that

$$t_U(A_n) = c_U(A_n A_n^*) c_U(A_n^*)^\dagger = U^* \delta(U A_n A_n^* U^*) \delta(U A_n^* U^*)^\dagger U.$$

Thus, $t_U(A_n)$ is stable (semi-stable) if and only if the diagonal matrix

$$\Lambda_n = \delta(UA_nA_n^*U^*)\delta(UA_n^*U^*)^\dagger$$

is stable (semi-stable). Since $\mathbf{u}_kA_n \neq 0$ for $k = 1, \cdots, n$, we know that $\delta(UA_nA_n^*U^*)$ is a diagonal matrix with positive diagonal entries. Therefore, Λ_n is stable (semi-stable) if and only if $\delta(UA_n^*U^*)^\dagger$ is stable (semi-stable). Note that $\delta(UA_n^*U^*)^\dagger$ is stable (semi-stable) if and only if $\delta(UA_n^*U^*)$ is stable (semi-stable) and $\delta(UA_n^*U^*)$ is stable (semi-stable) if and only if $\delta(UA_nU^*)$ is stable (semi-stable). Moreover, $\delta(UA_nU^*)$ is stable (semi-stable) if and only if $c_U(A_n) = U^*\delta(UA_nU^*)U$ is stable (semi-stable). Hence, $t_U(A_n)$ is stable (semi-stable) if and only if $c_U(A_n)$ is stable (semi-stable). □

Remark 3.1. According to Theorem 3.2 (iv), we see that if $\mathbf{u}_kA_n \neq 0$ for $k = 1, \cdots, n$, where $\mathbf{u}_k$ is the kth row of U, then $t_U(A_n)$ and $c_U(A_n)$ have the same stability property.

4. Singular value relation between $t_U(A_n)^\dagger A_n$ and $c_U(A_n)^\dagger A_n$

Jin and Wei proved the following result in [15].

Theorem 4.1. (Theorem 2.2 in [15]) *Let $A_n \in \mathbb{C}^{n\times n}$ such that $c_U(A_n)$ and $t_U(A_n)$ are invertible. If the singular values are arranged in the non-decreasing order. Then*

$$\sigma_k(t_U(A_n)^{-1}A_n) \leq \sigma_k(c_U(A_n)^{-1}A_n), \quad \textit{for} \;\; k = 1, \cdots, n.$$

We extend this result to a more general case. The following lemma and theorem will be used later.

Lemma 4.1. (Lemma 3.7 in [14]) *For $A_n \in \mathbb{C}^{n\times n}$, we have*

$$\delta(UA_nA_n^*U^*) \geq \delta(UA_nU^*)\delta(UA_n^*U^*) \geq 0.$$

Theorem 4.2. (Corollary 4.5.11 in [10]) *Let A_n be Hermitian and $S_n \in \mathbb{C}^{n\times n}$. Let the eigenvalues of A_n and $S_nS_n^*$ be arranged in the non-decreasing order. For each $k = 1, \cdots, n$, there exists $\theta_k \geq 0$ such that $\lambda_1(S_nS_n^*) \leq \theta_k \leq \lambda_n(S_nS_n^*)$ and*

$$\lambda_k(S_nA_nS_n^*) = \theta_k\lambda_k(A_n).$$

In particular, the number of positive (negative) eigenvalues of $S_nA_nS_n^*$ is less than or equal to the number of positive (negative) eigenvalues of A_n.

We always assume that the eigenvalues are arranged in the non-decreasing order.

Theorem 4.3. *Let $A_n \in \mathbb{C}^{n\times n}$ and $\mathbf{u}_k$ be the kth row of a unitary matrix U such that $\mathbf{u}_kA_n \neq 0$ holds for $k = 1, \cdots, n$. Let $S_n = \delta(UA_nA_n^*U^*)^{-1}\delta(UA_n^*U^*)\delta(UA_nU^*)$. Then we have*

$$\sigma_k(t_U(A_n)^\dagger A_n) = \theta_k\sigma_k(c_U(A_n)^\dagger A_n), \qquad k = 1, \cdots, n,$$

where

$$0 < \lambda_{n-m+1}(S_n) \leq \theta_k \leq \lambda_n(S_n) \leq 1$$

with $\lambda_{n-m+1}(S_n)$ and $\lambda_n(S_n)$ denoting the minimum and maximum nonzero eigenvalues of S_n respectively.

Proof. We have by Lemma 2.1 (i),

$$c_U(A_n)^\dagger A_n = U^*\delta(UA_nU^*)^\dagger UA_n.$$

Thus,

$$\begin{aligned}&(c_U(A_n)^\dagger A_n)(c_U(A_n)^\dagger A_n)^*\\&= U^*\delta(UA_nU^*)^\dagger UA_nA_n^*U^*\delta(UA_n^*U^*)^\dagger U \sim N, \quad (\text{“}\sim\text{” similar to})\end{aligned}$$

where $N \equiv \delta(UA_nU^*)^\dagger UA_nA_n^*U^*\delta(UA_n^*U^*)^\dagger$ is a Hermitian positive semi-definite matrix. Since $\mathbf{u}_kA_n \neq 0$ for all k, we know that $t_U(A_n)$ is uniquely determined by A_n (see Theorem 3.1) and in this case,

$$t_U(A_n)^\dagger A_n = c_U(A_nA_n^*)^{-1}c_U(A_n^*)A_n = U^*\delta(UA_nA_n^*U^*)^{-1}\delta(UA_n^*U^*)UA_n.$$

Therefore,

$$\begin{aligned}&(t_U(A_n)^\dagger A_n)(t_U(A_n)^\dagger A_n)^*\\&= U^*\delta(UA_nA_n^*U^*)^{-1}\delta(UA_n^*U^*)UA_nA_n^*U^*\delta(UA_nU^*)\delta(UA_nA_n^*U^*)^{-1}U \sim M,\end{aligned}$$

where

$$M \equiv \delta(UA_nA_n^*U^*)^{-1}\delta(UA_n^*U^*)UA_nA_n^*U^*\delta(UA_nU^*)\delta(UA_nA_n^*U^*)^{-1}$$

is a Hermitian positive semi-definite matrix. Let $S_n = \delta(UA_nA_n^*U^*)^{-1}\delta(UA_n^*U^*)\delta(UA_nU^*)$. It is easy to see from Lemma 4.1 that the eigenvalues of S_n satisfy

$$0 \le \lambda_k(S_n) \le 1, \qquad k = 1, \cdots, n. \tag{4.1}$$

Note that

$$\begin{aligned}\delta(UA_n^*U^*) &= \delta(UA_n^*U^*)\delta(UA_nU^*)\delta(UA_nU^*)^\dagger,\\ \delta(UA_nU^*) &= \delta(UA_n^*U^*)^\dagger\delta(UA_n^*U^*)\delta(UA_nU^*).\end{aligned}$$

Then we have

$$M = S_nNS_n^*. \tag{4.2}$$

Let P be a permutation matrix such that

$$P^{-1}\delta(UA_nU^*)P = \begin{pmatrix} D_m & 0\\ 0 & 0\end{pmatrix}, \tag{4.3}$$

where D_m is an m-by-m diagonal matrix whose diagonal entries are all the nonzero diagonal entries of $\delta(UA_nU^*)$. Therefore,

$$\begin{aligned}P^{-1}S_nP &= (P^{-1}\delta(UA_nA_n^*U^*)^{-1}P)(P^{-1}\delta(UA_n^*U^*)P)(P^{-1}\delta(UA_nU^*)P)\\&= \begin{pmatrix} G_m & 0\\ 0 & 0\end{pmatrix},\end{aligned} \tag{4.4}$$

where G_m is also an m-by-m diagonal matrix with nonzero diagonal entries. Obviously, the diagonal entries of G_m contain all the nonzero eigenvalues of S_n. It follows from (4.2) that

$$\begin{aligned} P^{-1}MP &= P^{-1}S_nNS_n^*P = (P^{-1}S_nP)(P^{-1}NP)(P^{-1}S_nP)^* \\ &= \begin{pmatrix} G_m & 0 \\ 0 & 0 \end{pmatrix} (P^{-1}NP) \begin{pmatrix} G_m & 0 \\ 0 & 0 \end{pmatrix}^*. \end{aligned} \tag{4.5}$$

We partition $P^{-1}MP$ and $P^{-1}NP$ as in the following forms:

$$P^{-1}MP = \begin{pmatrix} M_1 & M_2^* \\ M_2 & M_3 \end{pmatrix} \qquad \text{and} \qquad P^{-1}NP = \begin{pmatrix} N_1 & N_2^* \\ N_2 & N_3 \end{pmatrix},$$

where M_1 and N_1 are m-by-m submatrices. It is easy to see from (4.5) that

$$P^{-1}MP = \begin{pmatrix} M_1 & 0 \\ 0 & 0 \end{pmatrix} \tag{4.6}$$

with $M_1 = G_mN_1G_m^*$. Moreover, by (4.3) and

$$N = \delta(UA_nU^*)^\dagger UA_nA_n^*U^*\delta(UA_n^*U^*)^\dagger,$$

we have

$$\begin{aligned} \begin{pmatrix} N_1 & N_2^* \\ N_2 & N_3 \end{pmatrix} &= P^{-1}NP = (P^{-1}\delta(UA_nU^*)^\dagger P)(P^{-1}UA_nA_n^*U^*P)(P^{-1}\delta(UA_n^*U^*)^\dagger P) \\ &= \begin{pmatrix} D_m^{-1} & 0 \\ 0 & 0 \end{pmatrix} (P^{-1}UA_nA_n^*U^*P) \begin{pmatrix} (D_m^*)^{-1} & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Thus,

$$P^{-1}NP = \begin{pmatrix} N_1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.7}$$

Note that $M_1 = G_mN_1G_m^*$. By Theorem 4.2, we know that there exists $d_k \geq 0$ such that $\lambda_1(G_mG_m^*) \leq d_k \leq \lambda_m(G_mG_m^*)$ and

$$\lambda_k(M_1) = d_k\lambda_k(N_1), \qquad k = 1, \cdots, m, \tag{4.8}$$

where $\lambda_1(G_mG_m^*)$ and $\lambda_m(G_mG_m^*)$ denote the minimum and maximum eigenvalues of $G_mG_m^*$ respectively. From (4.6), (4.7) and (4.8), we derive

$$\lambda_k(M) = d_k\lambda_k(N),$$

where $\lambda_1(G_mG_m^*) \leq d_k \leq \lambda_m(G_mG_m^*)$ for $k = 1, \cdots, n$. Thus,

$$\sigma_k(t_U(A_n)^\dagger A_n) = \sqrt{d_k}\,\sigma_k(c_U(A_n)^\dagger A_n), \qquad k = 1, \cdots, n.$$

Let $\theta_k = \sqrt{d_k}$, $k = 1, \cdots, n$. From (4.4) and the fact that G_m is a diagonal matrix with nonzero diagonal entries, we have by (4.1),

$$0 < \lambda_{n-m+1}(S_n) \leq \theta_k \leq \lambda_n(S_n) \leq 1,$$

where $\lambda_{n-m+1}(S_n)$ and $\lambda_n(S_n)$ denote the minimum and maximum nonzero eigenvalues of S_n respectively. $\square$

Acknowledgments The work was supported by the research grant UL020/08-Y2/MAT/JXQ01/FST from University of Macau.

References

[1] A. Ben-Israel and T. Greville, *Generalized Inverses: Theory and Applications*, 2nd edition, Springer-Verlag, New York, 2003.
[2] R. Chan and X. Jin, *An Introduction to Iterative Toeplitz Solvers*, SIAM, Philadelpha, 2007.
[3] R. Chan, X. Jin and M. Yeung, *The circulant operator in the Banach algebra of matrices*, Linear Algebra Appl., 149 (1991), pp. 41–53.
[4] R. Chan and M. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Review, 38 (1996), pp. 427–482.
[5] T. Chan, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.
[6] J. Chen and X. Jin, *The generalized superoptimal preconditioner*, Linear Algebra Appl., 432 (2010), pp. 203–217.
[7] C. Cheng, X. Jin and Y. Wei, *Stability properties of superoptimal preconditioner from numerical range*, Numer. Linear Algebra Appl., 13 (2006), pp. 513–521.
[8] C. Estatico, *A class of filtering superoptimal preconditioners for highly ill-conditioned linear systems*, BIT, 42 (2002), pp. 753–778.
[9] C. Estatico and S. Serra-Capizzano, *Superoptimal approximation for unbounded symbols*, Linear Algebra Appl., 428 (2008), pp. 564–585.
[10] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[11] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[12] T. Huckle, *Circulant and skew circulant matrices for solving Toeplitz matrix problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 767–777.
[13] X. Jin, *Developments and Applications of Block Toeplitz Iterative Solvers*, Kluwer Academic Publishers, Dordrecht, and Science Press, Beijing, 2002.
[14] X. Jin, *Preconditioning Technique for Toeplitz Systems*, Higher Education Press, Beijing, 2010.
[15] X. Jin and Y. Wei, *A short note on singular values of optimal and superoptimal preconditioned matrices*, Inter. J. Comput. Math., 84 (2007), pp. 1261–1263.
[16] X. Jin and Y. Wei, *A survey and some extensions of T. Chan's preconditioner*, Linear Algebra Appl., 428 (2008), pp. 403–412.
[17] E. Tyrtyshnikov, *Optimal and superoptimal circulant preconditioners*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 459–473.