A Study for Spectral Properties of Preconditioner of Symmetric Toeplitz Systems

Ran Baik*

Abstract

In [9], Tyrtshnikov proposed a preconditioned approach to derive a general solution from a Toeplitz linear system. Furthermore, the process of selecting a preconditioner matrix from symmetric Toeplitz matrix, which has been used in previous studies, is introduced. This research introduces a new method for finding the preconditioner in a Toeplitz system. Also, through analyzing these preconditioners, it is derived that eigenvalues of a symmetric Toeplitz $T$ are very close to eigenvalues of a new preconditioner for $T$. It is shown that if the spectrum of the preconditioned system $C_0^{-1}T$ is clustered around 1, then the convergence rate of the preconditioned system is superlinear. From these results, it is determined to get the superlinear at the convergence rate by our good preconditioner $C_0$. Moreover, an advantage is driven by increasing various applications i.e. image processing, signal processing, etc. in this study from the proposed preconditioners for Toeplitz matrices. Another characteristic, which this research holds, is that the preconditioner retains the properties of the Toeplitz matrix.

Keywords : Toeplitz, Circulant, Preconditioner, Hermitian, Preconditioned Conjugate Gradient (PCG)

1. Introduction

In this paper we investigate a new preconditioner for preconditioned Toeplitz System. The studies on the preconditioning symmetric positive definite (SPD) Toeplitz matrices with circulant matrices have been [1], [3], [4], [5]. Toeplitz systems arise in a variety of practical applications in engineering fields. For instance, in signal processing, solutions of Toeplitz systems are
required in order to obtain the filter coefficients in the design of recursive digital filters [Chui, Chan] and time-series analysis also involves solutions of Toeplitz systems for unknown parameters of stationary autoregressive models [3]. An iterative method for solving the SPD system \( Ax = b \) can be derived by minimizing the quadratic functional \( \frac{1}{2}x^TAx - b^Tx \) with the conjugate gradient (CG) method, and the unique minimum gives the desired solution.

The convergence rate of the CG method depends on the spectrum of \( A \). In general, the CG method converges faster if \( A \) has a small condition number of clustered eigenvalues.

To accelerate its convergence rate, a pre-conditioning step is often introduced at each CG iteration, which leads to the PCG method.

A good choice of preconditioners for \( A \) is a matrix \( P \) that approximates \( A \) well (in the sense that the spectrum of the preconditioned matrix \( P^{-1}A \) is clustered around 1 or has a small condition number), and for which the matrix-vector product \( P^{-1}v \) can be computed efficiently for a given vector \( v \). With such a preconditioner, one solves in principle the preconditioned system \( \tilde{A}\tilde{x} = \tilde{b} \), where \( \tilde{A} = P^{1/2}AP^{-1/2} \), \( \tilde{x} = P^{-1/2}x \) and \( \tilde{b} = P^{-1/2}b \), by the CG method [3]. The idea of preconditioning is a simple one but is now recognized as critical to the effectiveness of the PCG method.

A Toeplitz preconditioner has also been proposed by Strang, and analyzed by Chan and Strang [3]. Strang's Preconditioner \( S \) is obtained by preserving the central half diagonals of \( A \) and using them to form a circulant matrix. Since \( S \) is a circulant, the matrix-vector product \( S^{-1}v \) can be conveniently computed via fast Fourier transformation (FFT) with \( O(n \log n) \) operations. It has been shown [1] - [5] that an eigenvalue class of \( S^{-1}A \) is clustered around 1 except a finite number of outlier. The convergence rate of preconditioned iterative methods depends on the singular value or eigen-value distribution of the preconditioned matrices.

In our research, we propose a new type of preconditioners \( C_0 \) for Toeplitz matrices as the following. We discuss about spectral properties of a circulant preconditioner \( C_0 \) in Section 2 and how to develop a circulant preconditioner \( C_0 \) for Toeplitz \( T_n \) in Section 3. It describes the eigenvalues distributions of \( C_0^{-1}T_n \), \( C_0 \), \( S \) and \( T_n \) in section 4. The concluding remarks are given in Sections 5.

### 2. Spectral properties of a new circulant preconditioner

we denote \( T_n \) by the set of \( n \) by \( n \) matrices and Toeplitz matrices, respectively:

\[
T_n = \begin{pmatrix}
t_0 & t_1 & \cdots & t_{n-1} \\
t_{-1} & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots \\
t_{-(n-1)} & \cdots & t_{-1} & t_0 \\
\end{pmatrix} \subseteq M_n. 
\]

We denote by \( T_n^R \) the set of all real symmetric Toeplitz matrices:

\[
T_n^R = \begin{pmatrix}
t_0 & t_1 & \cdots & t_{n-1} \\
t_{-1} & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots \\
t_{-(n-1)} & \cdots & t_{-1} & t_0 \\
\end{pmatrix} \subseteq T_n^S \quad (1)
\]

and by \( T_n^H \) the set of all hermitian Toeplitz matrices:

\[
T_n^H = \begin{pmatrix}
t_0 & t_1 & \cdots & t_{n-1} \\
t_1 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots \\
t_{n-1} & \cdots & t_1 & t_0 \\
\end{pmatrix} \subseteq T_n^H \quad (2)
\]

A Toeplitz matrix \( C = [c_{ij}] \subseteq T_n \) is called a circulant if \( c_{ij} = c_{(n+j-i)\mod n} \leq 0 \leq i, j \leq N-1 \), i.e.

\[
C = \begin{pmatrix}
c_0 & c_1 & \cdots & c_{n-1} \\
c_{n-1} & c_0 & \cdots & \vdots \\
\vdots & \ddots & \ddots & c_1 \\
c_{1} & \cdots & c_{n-1} & c_0 \\
\end{pmatrix} \subseteq T_n.
\]
From the structure of the circulant matrix, we see that
\[ C = c_0 J + c_1 J + \cdots + c_{n-1} J^{n-1} \]
where
\[ J = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \]

Therefore, every circulant matrix is generated by a simple permutation matrix \( J \).

Thus, \( C = F(c_0 J + c_1 J + \cdots + c_{n-1} J^{n-1}) \) and
\[
F = \text{Fidiag} \left( \sum_{j=0}^{n-1} c_j w^j, \sum_{j=0}^{n-1} c_j w^{j(n-1)} \right) F^T
\]
where \( \sum_{j=0}^{n-1} c_j w^{jk(n-1)} \) is the \( k \)th eigen-value,
\[
\frac{1}{\sqrt{n}} (w^{k(n-1)}, \ldots, w^{k(n-1)}(a-1))^T \in \mathbb{C}^n
\]
and
\[
w = e^{2\pi i/n} = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}
\]
is the corresponding \( k \)th eigen-vector of \( C \) for \( k = 1, 2, \ldots, n \).

We denote by \( C_n \) the set of all circulant matrices. Because all circulant matrices are generated by \( J \), \( C_n \) forms a commutative ring over real or complex field. Also it should be noted that \( C_n \) is completely characterized by the diagonalizability under \( F \)-unitary similarity.

Thus, \( C_n \) is a very special commutative subclass of the normal Toeplitz matrices. We denote by \( C_n^{R} \) the set of real symmetric circulant matrices and \( T_n^{R} \) the set of real symmetric Toeplitz matrices:
\[
C_n^{R} = \left\{ T_n^{R} (c_0, c_1, \ldots, c_k, c_{k-1}, \ldots, c_1) \right\}
\]
where \( k = \left\lfloor \frac{n}{2} \right\rfloor \).

Let \( T \in T_n^{R} \). Choose \( C_i \in C_n^{R} \) such that
\[
\| T - C_i \|_F = \min \left\{ \| T - C \|_F \right\}
\]
It is known that if
\[
T = T_n^{R} (t_0, t_1, \ldots, t_{n-1}) \subseteq T_n^{R}
\]
then
\[
C_i = T_n^{R} (c_{0i}, c_{1i}, \ldots, c_{ki}, c_{k-1}, \ldots, c_1) \subseteq C_n^{R}
\]
such that
\[
c_j = \frac{jt(n-j)j}{n} \quad j = 0, \ldots, k. \quad (4)
\]

Thus, (4) gives the formula for the initial symmetric circulant matrix choice for this case.

Note that
\[
|\lambda_j - \alpha_j|^2 \leq \sum_{j=0}^{n} (|\lambda_j - \alpha_j|^2) \leq \|T - C_i\|_F
\]
where \( \lambda_j \) and \( \alpha_j \) are the eigenvalues of the matrices \( T \) and \( C_i \) respectively. Thus, if \( \|T - C_i\|_F \) is small then the eigenvalues of \( C_i \) are close to the corresponding eigenvalues of \( T \) and we have a good choice of the preconditioner.

Let \( C_n^{H} \) be the set of real symmetric circulant matrices in (3). We denote by \( C_n^{H} \) the set of hermitian circulant matrices:
\[
C_n^{H} = \left\{ T_n^{H} (c_0, c_1, \ldots, c_k, c_{k-1}, \ldots, c_1) \right\}
\]
where \( k = \left\lfloor \frac{n}{2} \right\rfloor \) and \( T_n^{H} \) in (2).

Note that \( C_n^{H} \) is the \( F \)-real diagonalizable class of matrices: \( C \subseteq C_n^{H} \) if and only if
\[
C = F \text{idiag}(\alpha_1, \ldots, \alpha_n) F^T
\]
where \( \alpha_i \in \mathbb{R} \) for \( i = 1, \ldots, n \). It can be verified easily that \( C_n^{H} \subseteq C_n^{R} \) has the following finer spectral characteristic: \( C \subseteq C_n^{R} \) if and only if
\[
C = F \text{idiag}(\alpha_1, \ldots, \alpha_n) F^T
\]
where \( \alpha_i \in \mathbb{R} \) for \( i = 1, \ldots, n \) such that the algebraic multiplicity of each \( \alpha_i \) must be greater than or equal to 2, that is, there is no simple eigen value for real symmetric circulant matrices. A Toeplitz matrix \( K = [k_{ij}] \in T_n^{R} \) is called a skew circulant matrix if
\[
k_{ij} = \begin{cases} c_{j-i} & \text{for } 0 \leq i \leq j \leq n-1 \\ -c_{(i-j)} & \text{for } 0 \leq j \leq i \leq n-1 \end{cases}
\]

Thus, \( K = k_{01} + k_{12} + k_{23} + \cdots + k_{n-1}L^{n-1} \) where
Note that $P^*LP = \theta J$ where
\[
P = \text{diag}(1, \theta, \theta^2, \ldots, \theta^{n-1}), \quad \theta = e^{(\pi i)/n}
\]
and hence $F^*P^*KPE$ is a diagonal matrix. It is easy to identify that $P^*LP = \theta^s J$ for any odd number $s = 1, 3, \ldots$ and $K$ is diagonalizable under $P^*F$-unitarity for $s = 1, 3, \ldots$.

Note that $J = FWF^*$ and $J^* = F^*WF$ where
\[
P^2 = W^s = \text{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1}), \quad \omega = e^{2\pi i/n},
\]
hence $F^*PF$ must be a square root of $J^*$. Note that we need only to consider for $s = 1, \ldots, 2n - 1$, because of $P^s = P$. We denote by $K_n \subseteq T_n$ the set of all skew unitary matrices, and by $K_n^R \subseteq K_n$ the set of all real symmetric skew unitary matrices.

**Lemma 1:** The following are equivalent.

(i) $K \subseteq K_n^R$ (ii) $K$ is $P^*F$-real diagonalizable for $s = 1, 3, 5, \ldots, 2n - 1$

i.e., $K = C^{(s)}$ of all real diagonal matrices, $D = \text{diag}(\alpha_1^s, \alpha_2^s, \ldots, \alpha_n^s)$, $\alpha_i \in R$

(iii) $P^*KP \in C_n^R$ for $s = 1, 3, 5, \ldots, 2n - 1$

**Proof** (i) $\Rightarrow$ (iii) Suppose $K \subseteq K_n^R$ is a given real symmetric skew unitary matrix. Consider the set
\[
G = \{P^*KP\} s = 1, 3, 5, \ldots, 2n - 1
\]
\[
= \{W^*P^*KWP\} i = 0, \ldots, n - 1
\]
\[
= \{W^* C^{(s)} W\} i = 0, \ldots, n - 1
\]
where $C^{(s)}$ is a hermitian circular matrix, $P = \text{diag}(1, \theta, \theta^2, \ldots, \theta^{n-1})$ and $W = \text{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1})$, $\theta = e^{(\pi i)/n}$.

Since $FWF^* = J$, $G = \{F(P^{(s)} F^*) FWP F^* \} i = 0, \ldots, n - 1$

\[
= \{F(J^* \text{diag}(\alpha_1^s, \ldots, \alpha_n^s) J^*) F\} i = 0, \ldots, n - 1
\]
\[
= \{C^{(i)}\} i = 0, \ldots, n - 1 \subseteq C_n^R,
\]
where $\alpha_1^s, \ldots, \alpha_n^s$ are the eigenvalues of $K$.

(iii) $\Rightarrow$ (ii)

$P^*KP \in C_n^R, F^*P^* KPE$ is a diagonal. Thus, $K$ is $P^*F$-real diagonalizable for $s = 1, 3, 5, \ldots, 2n - 1$.

(ii) $\Rightarrow$ (i)

Let $C_n^R = C_n(G)$ be the convex hull of $G$, i.e., the set of all convex combination of the matrices from $G$.

We present the following general result about hermitian matrices. We denote by $H_n$ the set of all $n \times n$ hermitian matrices.

Suppose
\[
A \in H_n, \quad V = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix}, \quad \lambda_i \in R
\]
and $V \in M_n$ is unitary.

By $D$, we denote the set of all real diagonal matrices, $D = \{\text{diag}(\alpha_1, \ldots, \alpha_n) : \alpha_i \in R\}$.

For $a, b \in R$, $a \geq b$, we let $D(a, b) = \{\text{diag}(\alpha_1, \ldots, \alpha_n) : a_i \leq a \text{ for all } i = 1, \ldots, n\} \subseteq D$.

A subset $D^s(a, b) = \{\text{diag}(\alpha_1, \ldots, \alpha_n) : \sum_{i=1}^n \alpha_i = s \text{ and } b \leq \alpha_i \leq a\}$.

**Theorem 2:** Suppose
\[
A = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix}, \quad V^* \in H_n.
\]

Let $s = \sum_{i=1}^n \lambda_i$. Then $\min_{D(a, b), \sum_{i=1}^n \lambda_i = s} \|A - \Sigma\|_2$

\[
= \min_{D(a, b), \sum_{i=1}^n \lambda_i = s} \|A - \Sigma\|_2
\]

As a simple consequence of Theorem 2, we have the following.
Corollary 3: Let $K_0 \in K_n^R$ be given. Then
\[ \min_{C \in C_n^R} \| K_0 - C \|_2 = \min_{C \in C_n} \| K_0 - C \|_2 \]
where $C_0(G)$ is the convex hull of
\[ G = \left\{ W^* P^T K_0 P W^T P \right\}_{i=0, \ldots, n-1} \]
\[ = F' \left( J^* \text{diag}(\alpha_1, \ldots, \alpha_n) J' \right) F'_{i=0, \ldots, n-1} \]
\[ = \left\{ C^{(i)} \in C_n \right\}_{i=1, \ldots, n}. \]

3. A generation of circulant approximations

Now suppose $T \in T_n^R$. We define $T^C$, the complement of $T$ by $T^C = T_n^R$.

Then notice that
\[ T = \frac{T + T^C}{2} + \frac{T - T^C}{2} \]

where $T + T^C \in C_n^R$ and $T - T^C \in K_n^R$.

Thus we have
\[ T = C_T + K_T \quad (6) \]

where $C_T \in C_n^R$ and $K_T \in K_n^R$, and it is easy to verify that the decomposition is unique. Note that
\[ \min_{C \in C_n^R} \| T - C \|_2 = \min_{C \in C_n} \| C_T + K_T - C \|_2 \]
\[ = \min_{C \in C_n} \| C_T + K_T - (C_T + C) \|_2 \]
\[ \leq \min_{C \in C_n} \| K_T - C \|_2 \]

where the last equality is from the Corollary 3.

Since the object of this section is to obtain an initial hermitian circulant matrix, it is clear from above equality that we need to only consider a hermitian circulant matrix from the convex hull of $\left\{ P^T K_T P, \ldots, P^T (W^{-1} K_T W^{-1}) P \right\}$. We use the following steps to choose the initial hermitian circulant matrix for $T \in T_n^R$.

Set $S(\alpha, \beta) = \alpha P^T K_T P + \beta P^T W T K_T WP$,
\[ \alpha, \beta > 0, \alpha + \beta = 1. \]

Compute $\gamma(\alpha, \beta) = \frac{\| K_T - S(\alpha, \beta) \|_2}{\| C_T + S(\alpha, \beta) \|_2}$ for
\[ \alpha = 0, 0.1, \ldots, 0.9, 1. \]

Then the initial circulant matrix $S$ of given $T$ is the hermitian circulant matrix $C_0 = C_T + S(\alpha, \beta)$ with the minimum values of $\gamma(\alpha, \beta)$ where $T = C_T + K_T$.

Some examples of the distribution of the eigenvalues of $T$ and $C_0 \in C_n^R$ are given in Section 4.

**Example:** Choice of the scalars $\alpha$ and $\beta$.

We choose $\alpha$ and $\beta$, $\alpha + \beta = 1$ such that
\[ \gamma(\alpha, \beta) = \frac{\| K_T - S(\alpha, \beta) \|_2}{\| C_T + S(\alpha, \beta) \|_2} \]

is the minimum. Suppose $\gamma(\alpha_0, \beta_0) = \min_{\alpha, \beta} \gamma(\alpha, \beta)$, $\alpha + \beta = 1$. Then $C_0 = C_T + S(\alpha_0, \beta_0)$.

We provide the following two examples.

(a) $T \in T_n^R(t_1, \ldots, t_n)$ such that $t_i = \frac{1}{(i+1)^2}$ for $i = 0, \ldots, 49$. $\alpha_0 = 0.7$, $\beta_0 = 0.3$.

(b) $T \in T_n^R(t_1, \ldots, t_n)$ such that $t_i = \frac{(-1)^i}{i+1}$ for $i = 0, \ldots, 49$. $\alpha_0 = 1.0$, $\beta_0 = 0.3$.

Find $\alpha_0$ and $\beta_0$ and $\beta_0$ such that $\gamma(\alpha, \beta)$ is minimized. Then $C_0 = C_T + \alpha_0 P^T K_T P + \beta_0 P^T W T K_T WP$.

4. Numerical experiments

In this section, we compare our hermitian circulant preconditioner $C_0$ with Strang's circulant preconditioner $S_n$.

\[ \begin{bmatrix} t_0 & t_1 & \cdots & t_2 & t_1 \\ t_1 & t_0 & t_{-M} & \cdots & t_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ t_2 & t_{-M} & \cdots & t_{-M} & t_{-1} \\ t_{-1} & t_{-2} & \cdots & t_{-2} & t_1 \end{bmatrix} \]

\[ S_n = \begin{bmatrix} t_0 & t_1 & \cdots & t_2 & t_1 \\ t_1 & t_0 & t_{-M} & \cdots & t_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ t_2 & t_{-M} & \cdots & t_{-M} & t_{-1} \\ t_{-1} & t_{-2} & \cdots & t_{-2} & t_1 \end{bmatrix} \]

According to Strang's proposal shown above in the given matrix, we construct preconditioner $S_n$ by preserving $n$ consecutive diagonal in $T$.
and bring them around to form a circulant matrix.

We compare with Strang preconditioners, \( S_n \) and our preconditioner \( C_0 \) from the distribution of eigenvalues based on the objective matrix \( T \) with examples (a), (b) (Figure 1, 2). We also have an experiment for eigenvalue distributions of \( C_0^{-1}T \) with example (a), (b) (Figure 3).

\[ \text{(Figure 1) Eigenvalue Distribution of } S_n \text{ and } T \text{ for given example (a), (b)} \]

\[ \text{(Figure 2) Eigenvalue Distribution of } C_0 \text{ and } T \text{ for given example (a), (b)} \]

\[ \text{(Figure 3) Eigenvalue Distribution of } C_0^{-1}T \text{ for given example (a), (b)} \]

5. Conclusions

From (Figure 1, 2) our new preconditioner \( C_0 \) is closer than Strang preconditioners \( S_n \) to Toeplitz \( T \). We can apply a new and better preconditioner \( C_0 \) instead of the given matrix \( T \) by the iterative method for Toeplitz systems. We expect all eigenvalues of the given Toeplitz matrices to be close to all eigenvalues of new circulant preconditioners (Figure 2).

Also the distributions of eigenvalues of \( C_0^{-1}T \) are between 0.6 and 1.2 and are between 0.5 and 1.3 (Figure 3) as shown in section 3. We have the property that all eigenvalues of \( C_0^{-1}T \) are very close to 1 excluding the extreme eigenvalues. It supports to reduce the iterations for the iterative method on the Toeplitz system.

The objective matrix for case (a) is a symmetric positive matrix and \( ||I-C_0^{-1}T||_F \approx 1.1591 \) and (b) is a symmetric matrix and \( ||I-C_0^{-1}T||_F \approx 8.3689 \).

Acknowledgment

This work was supported by the KOSEF grant funded by the Korea government (MEST) (No. 2009-0083609)

REFERENCES

HOPKINS. 3. 1995

백란
1968년: North Carolina State University 대학원(박사후)
1995년: Northern Illinois University 대학원
(박사후-알고리즘)
1997년~현재: 호남대학교 컴퓨터공학과 교수
관심분야: 병렬알고리즘, 수치해석, 이미지 패턴인식 등