CIRCULANT AND SKEW–CIRCULANT PRECONDITIONERS FOR SKEW–HERMITIAN TYPE TOEPLITZ SYSTEMS

Raymond H. Chan and Xiao–Qing Jin

Department of Mathematics

University of Hong Kong

Hong Kong

November 1990

Abstract. We study the solutions of Toeplitz systems $A_n x = b$ by the preconditioned conjugate gradient method. The n-by-n matrix A_n is of the form $a_0 I + H_n$ where a_0 is a real number, I is the identity matrix and H_n is a skew-Hermitian Toeplitz matrix. Such kind of matrices appear in solving discretized hyperbolic differential equations. The preconditioners we considered here are circulant matrix C_n and skew-circulant matrix S_n where $A_n = \frac{1}{2}(C_n + S_n)$. The convergence rate of the iterative method depends on the spectra of the normalized preconditioned matrices $(C_n^{-1}A_n)^*(C_n^{-1}A_n)$ and $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$. For Toeplitz matrices A_n with entries which are Fourier coefficients of functions in the Wiener class, we show the invertibility of C_n and S_n and prove that the spectra of the normalized preconditioned matrices are clustered around one for large n. Hence, if the conjugate gradient method is applied to solve the normalized preconditioned systems, we expect fast convergence.

Abbreviated Title. Circulant and Skew-circulant Preconditioners

Key words. Skew-Hermitian type Toeplitz matrix, circulant matrix, skew-circulant matrix, preconditioned conjugate gradient method

AMS(MOS) subject classifications. 65F10, 65F15

§1 Introduction.

Let A_n be a Toeplitz matrix of order n having the following form:

$$A_n = \begin{bmatrix} a_0 & a_1 & \cdot & a_{n-2} & a_{n-1} \\ -\overline{a}_1 & a_0 & a_1 & \cdot & a_{n-2} \\ -\overline{a}_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_1 \\ -\overline{a}_{n-1} & \cdot & -\overline{a}_2 & -\overline{a}_1 & a_0 \end{bmatrix}, \qquad (1)$$

where a_0 is a real number. Obviously, $A_n = a_0 I + H_n$, where I is the identity matrix and H_n is a skew-Hermitian Toeplitz matrix. We call A_n the skew-Hermitian type Toeplitz matrix. We are interested in solving the system $A_n x = b$. This kind of problems often appear in solving discretized hyperbolic differential equations.

The idea of using preconditioned conjugate gradient method with circulant preconditioner for solving symmetric Toeplitz systems was first proposed by Strang [8]. Chan and Strang [1] then proved that if A_n is a symmetric Toeplitz matrix with entries which are Fourier coefficients of a positive function in the Wiener class, then the eigenvalues of $\tilde{C}_n^{-1} A_n$ are clustered around 1. Here $\tilde{C}_n$ is the symmetric circulant matrix which copies the central diagonals of A_n. These results are generalized to Hermitian positive definite Toeplitz systems in Chan [2]. However, only positive definite matrices are discussed in these papers. In this paper, we use the preconditioned conjugate gradient method with circulant preconditioner C_n or skew-circulant preconditioner S_n for solving the skew-Hermitian type Toeplitz systems. We show in §2 that if the generating function $f(\theta)$ is of the form $a_0 + ig(\theta)$, where $g(\theta)$ is a real-valued function in the Wiener class, then the spectra of $(C_n^{-1}A_n)^*(C_n^{-1}A_n)$ and $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$ are clustered around 1. In §3, we establish the superlinear convergence rate of the conjugate gradient method when applied to these normalized preconditioned systems. Finally, numerical results and some applications of the method to discretized hyperbolic systems are given in §4.

§2 Spectrum of the Normalized Preconditioned System.

Let us begin by supposing that the entries $a_{jl} = a_{l-j}$ of A_n are Fourier coefficients of the complex generating function

$$f(\theta) = \sum_{-\infty}^{\infty} a_k e^{-ik\theta}$$

defined on $[0, 2\pi)$. Since A_n is of the form given by (1), we see that $a_{-k} = -\bar{a}_k$. It is clear that $f(\theta)$ also can be written as

$$f(\theta) = a_0 + ig(\theta),$$

where $g(\theta)$ is a real-valued function. In fact,

$$\begin{aligned} g(\theta) = \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} (-ia_k)e^{-ik\theta} &= \sum_{k=1}^{\infty}(-i)(a_k e^{-ik\theta} + a_{-k}e^{ik\theta}) \\ &= \sum_{k=1}^{\infty}(-i)(a_k e^{-ik\theta} - \overline{a_k e^{-ik\theta}}) = 2\sum_{k=1}^{\infty} Im(a_k e^{-ik\theta}) \ , \end{aligned}$$

where $Im(x)$ denotes the imaginary part of the complex number x. We assume that $g(\theta)$ is a function in the Wiener class, i.e., its Fourier coefficients $\{-ia_k\}$ is in ℓ_1. We then have

$$\sum_{-\infty}^{\infty} |a_k| = |a_0| + \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} |-ia_k| \equiv c_0 < \infty \ . \tag{2}$$

It is clear that the solution of the system $A_n x = b$ is the same as the solution of the normalized preconditioned system $(P_n^{-1}A_n)^*(P_n^{-1}A_n)x = (P_n^{-1}A_n)^* P_n^{-1} b$ for any preconditioner P_n. When the conjugate gradient method is applied to solve the normalized preconditioned system, the convergence rate of the method depends on the spectrum of $(P_n^{-1}A_n)^*(P_n^{-1}A_n)$. The more clustered the eigenvalues are, the faster the convergence.

For skew-Hermitian type Toeplitz matrix A_n given by (1), we note that it can always be partitioned as

$$A_n = \frac{1}{2}\, C_n + \frac{1}{2}\, S_n,$$

where

$$C_n = \begin{bmatrix} a_0 & a_1 - \overline{a}_{n-1} & \cdot & a_{n-2} - \overline{a}_2 & a_{n-1} - \overline{a}_1 \\ -\overline{a}_1 + a_{n-1} & a_0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -\overline{a}_{n-2} + a_2 & \cdot & \cdot & a_0 & a_1 - \overline{a}_{n-1} \\ -\overline{a}_{n-1} + a_1 & \cdot & \cdot & -\overline{a}_1 + a_{n-1} & a_0 \end{bmatrix},$$

and

$$S_n = \begin{bmatrix} a_0 & a_1 + \overline{a}_{n-1} & \cdot & a_{n-2} + \overline{a}_2 & a_{n-1} + \overline{a}_1 \\ -(a_{n-1} + \overline{a}_1) & a_0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -(a_2 + \overline{a}_{n-2}) & \cdot & \cdot & a_0 & a_1 + \overline{a}_{n-1} \\ -(a_1 + \overline{a}_{n-1}) & \cdot & \cdot & -(a_{n-1} + \overline{a}_1) & a_0 \end{bmatrix}.$$

Clearly C_n is circulant and S_n is skew-circulant. We will use C_n and S_n as preconditioners for A_n. We show in the following that if the generating function $f(\theta) = a_0 + ig(\theta)$, where $g(\theta)$ is a real-valued function in the Wiener class, then

I. C_n, S_n and their inverses are uniformly bounded in the ℓ_2–norm;

II. $C_n^{-1}A_n$ and $S_n^{-1}A_n$ can be partitioned into the sum of $I + L_n + M_n$, where L_n is a low rank matrix and M_n is a matrix with small ℓ_2–norm;

III. the spectra of $(C_n^{-1}A_n)^*(C_n^{-1}A_n)$ and $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$ are clustered around one.

For claim I, we have

Theorem 1. *Let $f(\theta) = a_0 + ig(\theta)$, where $g(\theta)$ is a real-valued function in the Wiener class and a_0 is a nonzero real number. Then for all n, the circulant matrix C_n, the skew-circulant matrix S_n and their inverses are uniformly bounded in the ℓ_2–norm.*

Proof. We first prove the theorem for the skew-circulant matrix S_n. We note that $S_n = a_0 I + \widehat{H}_n$, where $\widehat{H}_n$ is a skew-Hermitian Toeplitz matrix. Therefore, the eigenvalues of S_n can be expressed as $\lambda_j(S_n) = a_0 + b_j i$, where b_j are real. Hence

$$|\lambda_j(S_n)| \geq |a_0| > 0, \quad \text{for} \quad j = 0, \cdots, n-1 . \tag{3}$$

In particular, we see that S_n is invertible. We also notice that the eigenvalues $\lambda_j(S_n)$ of S_n also can be expressed as

$$\lambda_j(S_n) = a_0 + \sum_{p=1}^{n-1} (a_p + \overline{a}_{n-p}) e^{\frac{(2j+1)\pi i}{n} p}, \quad \text{for} \quad j = 0, \cdots, n-1 .$$

see Davis [3]. Hence we have by (2)

$$|\lambda_j(S_n)| \le |a_0| + \sum_{p=1}^{n-1} |a_p + \overline{a}_{n-p}| \le 2 \sum_{p=-\infty}^{\infty} |a_p| = 2c_0 \ , \quad \text{for} \quad j = 0, \cdots, n-1 \ .$$

Since S_n is a normal matrix, we have

$$\|S_n\|_2 = |\lambda_{\max}(S_n)| \le 2c_0, \qquad \|S_n^{-1}\|_2 = |\lambda_{\min}^{-1}(S_n)| \le \frac{1}{|a_0|}. \tag{4}$$

Here $\lambda_{\max}(S_n)$ and $\lambda_{\min}(S_n)$ denote respectively the largest and the smallest eigenvalues in absolute value for S_n.

For circulant matrix C_n, the proof is similar. We only have to note that the eigenvalues of C_n are given by

$$\lambda_j(C_n) = a_0 + \sum_{p=1}^{n-1} (a_p - \overline{a}_{n-p}) e^{\frac{2\pi i j}{n} p}, \quad \text{for} \quad j = 0, \cdots, n-1 \ .$$

Hence we have, similar to (3) and (4),

$$|\lambda_j(C_n)| \ge |a_0| > 0, \quad \text{for} \quad j = 0, \cdots, n-1 \ ,$$

and

$$\|C_n\|_2 = |\lambda_{\max}(C_n)| \le 2c_0, \qquad \|C_n^{-1}\|_2 = |\lambda_{\min}^{-1}(C_n)| \le \frac{1}{|a_0|}. \qquad \square$$

We remark that when $a_0 = 0$, the matrix A_n is skew-Hermitian and the generating function $f(\theta) = ig(\theta)$. Then if we assume that $g(\theta) \ge g_{\min} > 0$ for all $\theta \in [0, 2\pi)$, then instead of getting (3), we can get $|\lambda_j(C_n)|, |\lambda_j(S_n)| \ge g_{\min} > 0$, for $j = 0, \cdots, n-1$, see Chan and Strang [1]. Thus we have

Theorem 1'. *Let $f(\theta) = ig(\theta)$, where $g(\theta)$ is a real positive function in the Wiener class. Then for all n sufficiently large, the circulant matrix C_n, the skew-circulant matrix S_n and their inverses are uniformly bounded in the ℓ_2–norm.*

For claim II, we first show that the matrix $C_n - A_n$ can be expressed as the sum of a low rank matrix and a matrix with small ℓ_2–norm.

Lemma 1. *Let $f(\theta) = a_0 + ig(\theta)$, where $g(\theta)$ is a real-valued function in the Wiener class. Then for all $\varepsilon > 0$, there exists an $N > 0$, such that for all $n > N$, $C_n - A_n = W_n^{(N)} + U_n^{(N)}$, where $\|W_n^{(N)}\|_2 \leq \varepsilon$ and rank$(U_n^{(N)}) \leq 2N$.*

Proof. Define $B_n = C_n - A_n$. It can be easily checked that B_n is skew-Hermitian Toeplitz matrix with entries $b_{ij} = b_{j-i}$ given by

$$b_l = \begin{cases} 0 & l = 0 \ , \\ a_{l-n} & 0 < l < n \ , \\ -\bar{b}_{-l} & 0 < -l < n \ . \end{cases}$$

Since g is in the Wiener class, by (2) for all given $\varepsilon > 0$, there exists an $N > 0$, such that

$$\sum_{l=N+1}^{\infty} |a_l| < \varepsilon.$$

Let $U_n^{(N)}$ be the n-by-n matrix obtained from B_n by replacing the $(n-N)$-by-$(n-N)$ leading principal submatrix of B_n by the zero matrix. Then rank $(U_n^{(N)}) \leq 2N$. Let $W_n^{(N)} \equiv B_n - U_n^{(N)}$. The leading $(n-N)$-by-$(n-N)$ block of $W_n^{(N)}$ is the leading $(n-N)$-by-$(n-N)$ principal submatrix of B_n, hence this block is a Toeplitz matrix. Thus

$$\begin{aligned} \|W_n^{(N)}\|_1 &\leq 2 \sum_{l=1}^{n-N-1} |b_l| = 2 \sum_{l=1}^{n-N-1} |a_{l-n}| = 2 \sum_{l=N+1}^{n-1} |a_{-l}| \\ &= 2 \sum_{l=N+1}^{n-1} |a_l| \leq 2 \sum_{l=N+1}^{\infty} |a_l| < 2\varepsilon \ . \end{aligned} \tag{5}$$

Since $W_n^{(N)}$ is skew-Hermitian, we have $\|W_n^{(N)}\|_\infty = \|W_n^{(N)}\|_1$. Thus

$$\|W_n^{(N)}\|_2 \leq (\|W_n^{(N)}\|_1 \cdot \|W_n^{(N)}\|_\infty)^{\frac{1}{2}} < \varepsilon.$$

Hence the spectrum of $W_n^{(N)}$ lies in $(-\varepsilon, \varepsilon)$. □

Since $C_n - A_n = A_n - S_n$, we have by Lemma 1,

$$\begin{aligned} S_n^{-1} A_n &= I + S_n^{-1}(A_n - S_n) = I + S_n^{-1}(C_n - A_n) \\ &= I + S_n^{-1}(W_n^{(N)} + U_n^{(N)}) = I + M_n + L_n \ , \end{aligned} \tag{6}$$

where $M_n = S_n^{-1}W_n^{(N)}$ and $L_n = S_n^{-1}U_n^{(N)}$. Similarly, for the circulant matrix C_n, we have

$$C_n^{-1}A_n = I + C_n^{-1}(A_n - C_n) = I - C_n^{-1}(W_n^{(N)} + U_n^{(N)}).$$

By Theorem 1, we thus have our main result.

Theorem 2. *Let $f(\theta) = a_0 + ig(\theta)$, where $g(\theta)$ is a real-valued function in the Wiener class and a_0 is a nonzero real number. Then for all $\varepsilon > 0$, there exists an $N > 0$, such that for all $n > N$, $C_n^{-1}A_n$ and $S_n^{-1}A_n$ can be written in the form $I + M_n + L_n$, where $\|M_n\|_2 \leq \varepsilon$ and rank$(L_n) \leq 2N$.*

Claim III now is an immediate Corollary of Theorem 2.

Corollary 1. *Let $f(\theta) = a_0 + ig(\theta)$, where $g(\theta)$ is a real-valued function in the Wiener class and a_0 is a nonzero real number. Then for all n sufficiently large, the spectra of $(C_n^{-1}A_n)^*(C_n^{-1}A_n)$ and $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$ are clustered around one. More precisely, for all $\varepsilon > 0$, there exists an $N > 0$ such that for all $n > N$, at most $4N$ eigenvalues of $(C_n^{-1}A_n)^*(C_n^{-1}A_n) - I$ and also of $(S_n^{-1}A_n)^*(S_n^{-1}A_n) - I$ can have absolute values exceeding ε.*

Proof. We prove the Corollary for the case of skew-circulant matrix S_n. The proof for C_n is similar and we therefore omit it. By Theorem 2, we have for all $\varepsilon > 0$, there exists an $N > 0$, such that for all $n > N$, $S_n^{-1}A_n = I + M_n + L_n$ where $\|M_n\|_2 \leq \varepsilon$ and rank$(L_n) \leq 2N$. Hence

$$\begin{aligned}(S_n^{-1}A_n)^*(S_n^{-1}A_n) - I &= (I + M_n^* + L_n^*)(I + M_n + L_n) - I \\ &= M_n + L_n + M_n^* + M_n^*M_n + M_n^*L_n + L_n^* + L_n^*M_n + L_n^*L_n \\ &= \widehat{M}_n + \widehat{L}_n \ ,\end{aligned}$$

where $\widehat{M}_n = M_n + M_n^* + M_n^*M_n$ and $\widehat{L}_n = L_n^*(I + L_n + M_n) + (I + M_n^*)L_n$. Thus $\|\widehat{M}_n\|_2 \leq 2\varepsilon + \varepsilon^2$ and rank$(\widehat{L}_n) \leq 4N$. Since both $\widehat{M}_n$ and $\widehat{L}_n$ are Hermitian, by Cauchy

Interlace Theorem, see Wilkinson [9], at most $4N$ eigenvalues of $(S_n^{-1}A_n)^*(S_n^{-1}A_n) - I$ have absolute values exceeding $2\varepsilon + \varepsilon^2$. $\square$

Similar to Theorem 1', when $a_0 = 0$ but $g(\theta)$ is positive, we have

Theorem 2'. *Let $f(\theta) = ig(\theta)$, where $g(\theta)$ is a real positive function in the Wiener class. Then for all $\varepsilon > 0$, there exists an $N > 0$, such that for all $n > N$, $C_n^{-1}A_n$ and $S_n^{-1}A_n$ can be written in the form $I + M_n + L_n$, where $\|M_n\|_2 \le \varepsilon$ and rank$(L_n) \le 2N$.*

Corollary 1'. *Let $f(\theta) = ig(\theta)$, where $g(\theta)$ is a real positive function in the Wiener class. Then for all n sufficiently large, the spectra of the normalized preconditioned matrices $(C_n^{-1}A_n)^*(C_n^{-1}A_n)$ and $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$ are clustered around one.*

Hence if the conjugate gradient method is applied to solve the normalized preconditioned system

$$(C_n^{-1}A_n)^*(C_n^{-1}A_n)x = (C_n^{-1}A_n)^*C_n^{-1}b,$$

or

$$(S_n^{-1}A_n)^*(S_n^{-1}A_n)x = (S_n^{-1}A_n)^*S_n^{-1}b,$$

we can expect a fast convergence rate.

§3 Superlinear Convergence Rate.

It follows easily from Corollary 1 or Corollary 1' that the conjugate gradient method converges superlinearly when applied to solve the normalized preconditioned systems, see Chan and Strang [1] for details. We note that if extra smoothness conditions are imposed on g, we can obtain a more precise bound on the convergence rate.

Theorem 3. *Let $f(\theta) = a_0 + ig(\theta)$, where $g(\theta)$ is a $(\ell+1)$-times differentiable function in the Wiener class with its $(\ell+1)$-th derivative in $L^1[0, 2\pi)$, $\ell > 0$, and a_0 is a nonzero* *real number. Then for large n,*

$$\|e_{8q}\| \le \frac{c^q}{[(q-1)!]^{8\ell}} \|e_0\| \ , \tag{7}$$

for some constant c that depends on f and ℓ only. Here e_q denotes the error vector at the q-th iteration, when the conjugate gradient method is applied to solve the normalized preconditioned system

$$(C_n^{-1}A_n)^*(C_n^{-1}A_n)x = (C_n^{-1}A_n)^*C_n^{-1}b \ ,$$

or

$$(S_n^{-1}A_n)^*(S_n^{-1}A_n)x = (S_n^{-1}A_n)^*S_n^{-1}b \ ;$$

and $||x||^2 = x^*(C_n^{-1}A_n)^*(C_n^{-1}A_n)x$ or $||x||^2 = x^*(S_n^{-1}A_n)^*(S_n^{-1}A_n)x$ accordingly.

Proof. Again for simplicity, we only give the proof for the skew-circulant case. The proof for the circulant case is similar. We recall that from the standard error analysis of the conjugate gradient method, we have

$$\frac{\|e_q\|}{\|e_0\|} \le \min_{P_q} \max_{\lambda} |P_q(\lambda)|, \tag{8}$$

where the minimum is taken over polynomials of degree q with constant term 1 and the maximum is taken over the spectrum of $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$, see for instance, Golub and van Loan [4]. In the following, we try to estimate that minimum.

We first recall that the Fourier coefficients of g are given by $-ia_j$. Hence the assumptions on g imply that

$$|-ia_j| \le \frac{c_1}{|j|^{\ell+1}} \ , \quad \forall j > 0 \ ,$$

where $c_1 = \|g^{(\ell+1)}\|_{L^1}$, see for instance, Katznelson [6]. Therefore we have

$$|a_j| \le \frac{c_1}{|j|^{\ell+1}} \ , \quad \forall j > 0 \ .$$

Hence for $W_n^{(k)}$ and $U_n^{(k)}$ defined as in Lemma 1, we have rank $(U_n^{(k)}) \le 2k$ and also by (5),

$$\|W_n^{(k)}\|_2 \le (\|W_n^{(k)}\|_\infty \|W_n^{(k)}\|_1)^{\frac{1}{2}} = \|W_n^{(k)}\|_1 \le 2 \sum_{j=k+1}^{n-1} |a_j|$$
$$\le 2c_1 \sum_{j=k+1}^{n-1} \frac{1}{|j|^{\ell+1}} \le 2c_1 \int_k^\infty \frac{dx}{x^{\ell+1}} \le \frac{2c_1}{k^\ell} \ , \quad \forall k \ge 1 \ .$$

Let $S_n^{-1}A_n = I + M_n^{(k)} + L_n^{(k)}$, where $M_n^{(k)} = S_n^{-1}W_n^{(k)}$ and $L_n^{(k)} = S_n^{-1}U_n^{(k)}$. For all $k \ge 1$, we have rank$(L_n^{(k)}) \le 2k$ and

$$\|M_n^{(k)}\|_2 = \|S_n^{-1}W_n^{(k)}\|_2 \le \|S_n^{-1}\|_2 \|W_n^{(k)}\|_2 \le \frac{1}{|a_0|}\|W_n^{(k)}\|_2 \le \frac{c_2}{k^\ell} \ ,$$

where $c_2 = 2c_1/|a_0|$. For the Hermitian matrix $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$, we have

$$(S_n^{-1}A_n)^*(S_n^{-1}A_n) = I + \widehat{M}_n^{(k)} + \widehat{L}_n^{(k)} \ ,$$

where

$$\widehat{M}_n^{(k)} = M_n^{(k)} + M_n^{(k)*} + M_n^{(k)*}M_n^{(k)} \ , \tag{9}$$

and

$$\begin{aligned} \widehat{L}_n^{(k)} =& L_n^{(k)} + L_n^{(k)*} + L_n^{(k)*}M_n^{(k)} + M_n^{(k)*}L_n^{(k)} + L_n^{(k)*}L_n^{(k)} \\ =& L_n^{(k)*}(I + L_n^{(k)} + M_n^{(k)}) + (I + M_n^{(k)*})L_n^{(k)} \ . \end{aligned} \tag{10}$$

From (9), it is clear that,

$$\|\widehat{M}_n^{(k)}\|_2 \le \|M_n^{(k)}\|_2 + \|M_n^{(k)*}\|_2 + \|M_n^{(k)}\|_2\|M_n^{(k)*}\|_2 \le \frac{2c_2}{k^\ell} + \frac{c_2^2}{k^{2\ell}} \le \frac{c_3}{k^\ell} \ ,$$

where $c_3 = 2c_2 + c_2^2$, and from (10), we see that, rank $(\widehat{L}_n^{(k)}) \le 2 \cdot 2k = 4k$.

Let the eigenvalues of $(S_n^{-1}A_n)^*(S_n^{-1}A_n) - I$ be ordered as

$$\mu_0^- \le \mu_{\frac{1}{4}}^- \le \mu_{\frac{2}{4}}^- \le \cdots \le \mu_{\frac{p}{4}}^- \le \cdots \le \mu_{\frac{p}{4}}^+ \le \cdots \le \mu_{\frac{2}{4}}^+ \le \mu_{\frac{1}{4}}^+ \le \mu_0^+ \ ,$$

then for all $p \geq 4$, we have, by Cauchy interlace theorem,

$$|\mu^{\pm}_{\frac{p}{4}}| \leq \|\widehat{M}_n^{([\frac{p}{4}])}\|_2 \leq \frac{c_3}{[\frac{p}{4}]^{\ell}} \ , \tag{11}$$

where $[x]$ denotes the integral part of x. Let the eigenvalues of $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$ be ordered as

$$\lambda_0^- \leq \lambda_{\frac{1}{4}}^- \leq \lambda_{\frac{2}{4}}^- \leq \cdots \leq \lambda_{\frac{p}{4}}^- \leq \cdots \leq \lambda_{\frac{p}{4}}^+ \leq \cdots \leq \lambda_{\frac{2}{4}}^+ \leq \lambda_{\frac{1}{4}}^+ \leq \lambda_0^+ \ .$$

Then for all $p \geq 4$, we have by (11)

$$1 - \frac{c_3}{[\frac{p}{4}]^{\ell}} \leq \lambda_{\frac{p}{4}}^- \leq \lambda_{\frac{p}{4}}^+ \leq 1 + \frac{c_3}{[\frac{p}{4}]^{\ell}} \ , \tag{12}$$

and for all $p \geq 0$, we have by (4)

$$\frac{a_0^2}{4c_0^2} \leq \lambda_{\frac{p}{4}}^- \leq \lambda_{\frac{p}{4}}^+ \leq \frac{c_0^2}{a_0^2} \ . \tag{13}$$

Our idea is to choose P_{8q} that annihilates the $4q$ pairs of extreme eigenvalues. Let

$$p_k(x) = \prod_{p=4k}^{4(k+1)-1} (1 - \frac{x}{\lambda_{\frac{p}{4}}^+})(1 - \frac{x}{\lambda_{\frac{p}{4}}^-}) \ , \qquad k = 0, \cdots, q-1,$$

and

$$P_{8q} = p_0 p_1 \cdots p_{q-1} \ .$$

We note that between the roots $\lambda^{\pm}_{\frac{p}{4}}$, the maximum value of $\left|(1 - \frac{x}{\lambda_{\frac{p}{4}}^+})(1 - \frac{x}{\lambda_{\frac{p}{4}}^-})\right|$ is obtained at the point $\frac{1}{2}(\lambda_{\frac{p}{4}}^+ + \lambda_{\frac{p}{4}}^-)$. Hence for $k = 0$, we have by (13),

$$\begin{aligned}\max_{x \in [\lambda_{\frac{3}{4}}^-, \lambda_{\frac{3}{4}}^+]} |p_0(x)| &\leq \prod_{p=0}^{3} \max_{x \in [\lambda_{\frac{p}{4}}^-, \lambda_{\frac{p}{4}}^+]} \left|(1 - \frac{x}{\lambda_{\frac{p}{4}}^+})(1 - \frac{x}{\lambda_{\frac{p}{4}}^-})\right| \\ &= \prod_{p=0}^{3} \frac{(\lambda_{\frac{p}{4}}^+ - \lambda_{\frac{p}{4}}^-)^2}{4\lambda_{\frac{p}{4}}^+ \lambda_{\frac{p}{4}}^-} \leq \prod_{p=0}^{3} (\frac{2c_0^2}{a_0^2})^2 \frac{1}{4} (\frac{4c_0^2}{a_0^2})^2 = c_4 \ ,\end{aligned}$$

for some constant c_4 that depends only on f and ℓ. For $1 \leq k \leq q-1$, we have by (12) and (13),

$$\begin{aligned}\max_{x \in [\lambda_{\frac{4(k+1)-1}{4}}^-, \lambda_{\frac{4(k+1)-1}{4}}^+]} |p_k(x)| &\leq \prod_{p=4k}^{4(k+1)-1} \max_{x \in [\lambda_{\frac{p}{4}}^-, \lambda_{\frac{p}{4}}^+]} \left|(1 - \frac{x}{\lambda_{\frac{p}{4}}^+})(1 - \frac{x}{\lambda_{\frac{p}{4}}^-})\right| \\ &= \prod_{p=4k}^{4(k+1)-1} \frac{(\lambda_{\frac{p}{4}}^+ - \lambda_{\frac{p}{4}}^-)^2}{4\lambda_{\frac{p}{4}}^+ \lambda_{\frac{p}{4}}^-} \leq \prod_{p=4k}^{4(k+1)-1} (\frac{2c_3}{[\frac{p}{4}]^{\ell}})^2 \frac{1}{4} (\frac{4c_0^2}{a_0^2})^2 \\ &= \frac{c_5}{k^{8\ell}} \ ,\end{aligned}$$

for some constant c_5 that also depends only on f and ℓ. Thus

$$\max_{x\in[\lambda_{q-1}^-,\lambda_{q-1}^+]} |P_{8q}(x)| \le c_4 c_5^{q-1} \prod_{k=1}^{q-1} \frac{1}{k^{8\ell}} = \frac{c^q}{[(q-1)!]^{8\ell}} \ , \tag{14}$$

for some constant c that depends on f and ℓ only. We remark that except for the $4q$ pairs of extreme eigenvalues, all eigenvalues are in the interval $[\lambda_{q-1}^-,\lambda_{q-1}^+]$. Since P_{8q} annihilates the $4q$ pairs of extreme eigenvalues and satisfies (14) in the interval $[\lambda_{q-1}^-,\lambda_{q-1}^+]$, inequality (7) now follows directly from (8) and (14). □

When $a_0 = 0$, if we assume that $g(\theta)$ is a positive function as before, we then have

Theorem 3'. *Let $f(\theta) = ig(\theta)$, where $g(\theta)$ is a $(\ell+1)$-times differentiable positive function in the Wiener class with its $(\ell+1)$-th derivative in $L^1[0,2\pi)$, $\ell > 0$. Then for large n,*

$$\|e_{8q}\| \le \frac{c^q}{[(q-1)!]^{8\ell}} \|e_0\| \ ,$$

for some constant c that depends on g and ℓ only.

§4 Numerical Results.

We consider the following two problems in this section.

Problem 1.

Let A_n be a Toeplitz matrix with diagonals given by

$$a_k = \begin{cases} 1 & k = 0 \ , \\ (1+k)^{-1.1} & k > 0 \ , \\ -a_{-k} & k < 0 \ . \end{cases}$$

Obviously the generating function of A_n is in the Wiener class. Table 1 shows the number of iterations required to get $\frac{\|r_k\|_2}{\|r_0\|_2} < 10^{-7}$. Here r_k is the residual vector at the k-th iteration. We use the vector of all ones for the right hand side vector b and the zero vector as our initial guess. We note that as n increases, the number of iterations increases for the normalized system, while it remains roughly a constant for both normalized preconditioned systems.

n	$A_n^* A_n$	$(S_n^{-1}A_n)^*(S_n^{-1}A_n)$	$(C_n^{-1}A_n)^*(C_n^{-1}A_n)$
16	8	8	7
32	12	8	7
64	16	8	7
128	18	8	8

Table 1. Number of Iterations for Different Systems

Problem 2.

We consider a model linear hyperbolic equation discussed in Holmgren and Otto [5]. For simplicity, we begin with the one-dimensional case first. The equation is:

$$\frac{\partial u(x,t)}{\partial t} + v\frac{\partial u(x,t)}{\partial x} = g(x) \tag{15}$$

which is defined on the domain

$$0 < x \le 1, \qquad t > 0 \ ,$$

with boundary and initial conditions

$$\begin{aligned} u(0,t) =& f(-at) \ , \\ u(x,0) =& f(x) \ . \end{aligned}$$

The right hand side function is given by

$$g(x) = (v-a)f'(x) \ .$$

Here v and a are positive constants and f is a scalar function with derivative f'. The analytical solution of (15) is given by $u = f(x - at)$.

Let k, h denote the time step and spatial step respectively. A time-discretization of (15) by using the trapezoidal rule gives:

$$4u^{m+1} + 2kv\left(\frac{\partial u}{\partial x}\right)^{m+1} = 2k(g^{m+1}+g^m) + 4u^m - 2kv\left(\frac{\partial u}{\partial x}\right)^{m} \ . \tag{16}$$

The spatial grid is uniform in the computational domain with $n+1$ gridpoints. Let u_i denote the approximative solution at the point x_i, where

$$x_i = ih \ , \quad i = 0, \cdots, n \ .$$

Clearly u_0 is given directly by the boundary condition. This implies that we have to solve for n unknowns in each time step. We use centered difference

$$\frac{\partial u}{\partial x} \approx \frac{u_{i+1} - u_{i-1}}{2h} \ , \quad i = 1, 2, \cdots, n-1 \ ,$$

to approximate the spatial derivative in (16) in the interior of the domain. At the outflow boundary we use one-sided difference

$$\frac{\partial u}{\partial x} \approx \frac{u_n - u_{n-1}}{h} \ .$$

Let $\alpha = vk/h$ and $u^m = (u_1^m, u_2^m, \cdots, u_n^m)$. By introducing the space-discretization in (16) for the unknowns at time level $m+1$, we have the following system,

$$\begin{cases} 4u_i^{m+1} + \alpha(u_{i+1}^{m+1} - u_{i-1}^{m+1}) = b_i & i = 1, 2, \cdots, n-1 \ , \\ (4+2\alpha)u_n^{m+1} - 2\alpha u_{n-1}^{m+1} = b_n \ . \end{cases}$$

In matrix form, we have

$$A_n u^{m+1} = b \ ,$$

where $b = (b_1, b_2, \cdots, b_n)^T$ contains known quantities and

$$A_n = \begin{bmatrix} 4 & \alpha & & & & \mathbf{0} \\ -\alpha & 4 & \ddots & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\alpha & 4 & \alpha \\ \mathbf{0} & & & & -2\alpha & 4+2\alpha \end{bmatrix} . \tag{17}$$

We define our circulant and skew-circulant preconditioners as

$$C_n = \begin{bmatrix} 4 & \alpha & & & -\alpha \\ -\alpha & 4 & \ddots & & \mathbf{0} \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & \ddots & 4 & \alpha \\ \alpha & & & -\alpha & 4 \end{bmatrix} ,$$

and

$$S_n = \begin{bmatrix} 4 & \alpha & & & \alpha \\ -\alpha & 4 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 4 & \alpha \\ -\alpha & & & -\alpha & 4 \end{bmatrix} .$$

Notice that since

$$S_n^{-1}A_n = I + S_n^{-1}(A_n - S_n) = I + L_n \ ,$$

where $L_n = S_n^{-1}(A_n - S_n)$, we have

$$(S_n^{-1}A_n)^*(S_n^{-1}A_n) = (I + L_n)^*(I + L_n) = I + \widehat{L}_n \ .$$

Here $\widehat{L}_n = L_n^* + (I + L_n^*)L_n$. Since rank (L_n) = rank $(A_n - S_n) = 2$, rank $(\widehat{L}_n) \leq 4$. Thus if the conjugate gradient method is applied to normalized preconditioned matrix $(S_n^{-1}A_n)^*(S_n^{-1}A_n)$, the number of iterations is around 5. The same conclusion holds for the circulant preconditioner C_n. Table 2 gives the number of iterations obtained experimentally. Here $\alpha = 100$, b is the vector of all ones and also as before, the zero vector is our initial guess. Table 2 shows the number of iterations required to get $\frac{\|r_k\|_2}{\|r_0\|_2} < 10^{-7}$.

n	$A_n^*A_n$	$(S_n^{-1}A_n)^*(S_n^{-1}A_n)$	$(C_n^{-1}A_n)^*(C_n^{-1}A_n)$
16	17	5	6
32	33	5	5
64	67	5	6
128	101	5	5

Table 2. Number of Iterations for Different Systems

Next we consider the two-dimensional case. The equation is:

$$\frac{\partial u(x_1,x_2,t)}{\partial t} + v_1\frac{\partial u(x_1,x_2,t)}{\partial x_1} + v_2\frac{\partial u(x_1,x_2,t)}{\partial x_2} = g(x_1,x_2) \tag{18}$$

which is defined on the domain

$$0 < x_1 \leq 1, \ 0 < x_2 \leq 1, \ t > 0 \ ,$$

with boundary and initial conditions

$$u(x_1, 0, t) = f(x_1 - at) \ ,$$

$$u(0, x_2, t) = f(x_2 - at) \ ,$$

$$u(x_1, x_2, 0) = f(x_1 + x_2) \ .$$

The right hand side function is given by

$$g(x) = (v_1 + v_2 - a) f'(x) \ .$$

Here v_1, v_2 and a are positive constants, f is a scalar function with derivative f'. The analytical solution of (18) is given by $u = f(x_1 + x_2 - at)$.

For simplicity, we assume that $v_1 = v_2 = v$ and the two spatial steps are equal, i.e., $h_1 = h_2 = h$. Then as before we let $\alpha = vk/h$, where k denotes the time step. For each time step, the discretized system to solve is of the following form:

$$A_N u^{m+1} = b,$$

where $A_N = A_n \otimes I + I \otimes A_n$ with A_n given by

$$A_n = \begin{bmatrix} 2 & \alpha & & & & \mathbf{0} \\ -\alpha & 2 & \ddots & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\alpha & 2 & \alpha \\ \mathbf{0} & & & & -2\alpha & 2+2\alpha \end{bmatrix} .$$

We note that the A_n here is different from the A_n in (17) only at the main diagoanl.

For the 2-dimensional problem, we have tested the following preconditioners:

$$C_N = C_n \otimes I + I \otimes C_n,$$

and

$$S_N = S_n \otimes I + I \otimes S_n,$$

where C_n and S_n are given by

$$C_n = \begin{bmatrix} 2 & \alpha & & & -\alpha \\ -\alpha & 2 & \ddots & & \mathbf{0} \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & \ddots & 2 & \alpha \\ \alpha & & & -\alpha & 2 \end{bmatrix},$$

and

$$S_n = \begin{bmatrix} 2 & \alpha & & & \alpha \\ -\alpha & 2 & \ddots & & \mathbf{0} \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & \ddots & 2 & \alpha \\ -\alpha & & & -\alpha & 2 \end{bmatrix}.$$

Different from that in one-dimensional case, here we use the preconditioned conjugate gradient squared method, see Sonneveld [7], to solve the following preconditioned systems

$$C_N^{-1}A_N u^{m+1} = C_N^{-1}b \quad \text{and} \quad S_N^{-1}A_N u^{m+1} = S_N^{-1}b.$$

Tables 3, 4 and 5 show the number of iterations required to get $\frac{\|r_k\|_2}{\|r_0\|_2} < 10^{-7}$ for different α's. We notice that as n increases, for $\alpha = 10, 100$, the number of iterations increases for the original system, while it stays almost the same for both preconditioned systems.

n	$N = n^2$	A_N	$S_N^{-1}A_N$	$C_N^{-1}A_N$
16	256	10	6	6
32	1024	10	6	6
64	4096	10	6	6
128	16384	10	5	5

Table 3. Number of Iterations for $\alpha = 1$

n	$N = n^2$	A_N	$S_N^{-1}A_N$	$C_N^{-1}A_N$
16	256	47	12	11
32	1024	65	11	11
64	4096	95	11	10
128	16384	124	11	10

Table 4. Number of Iterations for $\alpha = 10$

n	$N = n^2$	A_N	$S_N^{-1}A_N$	$C_N^{-1}A_N$
16	256	218	21	21
32	1024	329	20	20
64	4096	371	21	18
128	16384	454	23	21

Table 5. Number of Iterations for $\alpha = 100$

Finally, we would like to mention that our numerical results are similar to that obtained by Holmgren and Otto [5] where they used different choices of circulant preconditioners.

References

[1] R. Chan and G. Strang, *Toeplitz Equations by Conjugate Gradients with Circulant Preconditioner*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 104–119.

[2] R. Chan, *Circulant Preconditioners for Hermitian Teoplitz Systems*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 542–550.

[3] P. Davis, *Circulant Matrices*, John Wiley & Sons, Inc., New York, 1979.

[4] H. Golub and C. van Loan, *Matrix Computations*, The Johns Hopkins University Press, Maryland, 1983.

[5] S. Holmgren and K. Otto, *Iterative Solution Methods and Preconditioners for Block-tridiagonal Systems of Equations*, To appear.

[6] Y. Katznelson, *An Introduction to Harmonic Analysis*, 2nd Ed., Dover Publications, New York, 1976.

[7] P. Sonneveld, *CGS, a Fast Lanczos-type Solver for Nonsymmetric Linear Systems*, SIAM J. Sci. Stat. Comput., 10 (1989), pp. 36–52.

[8] G. Strang, *A Proposal for Toeplitz Matrix Calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.

[9] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.