Journal of Computational Mathematics, Vol.19, No.2, 2001, 167–176.

SINE TRANSFORM MATRIX FOR SOLVING TOEPLITZ MATRIX PROBLEMS*[1)]

Li-zhi Cheng
(*Department of Mathematics & System Sciences, National University of Defence Technology, Changsha 410073, China*)

Abstract

In recent papers, some authors studied the solutions of symmetric positive definite(SPD) Toeplitz systems $T_n x = b$ by the conjugate gradient method(CG) with different sine transforms based preconditioners. In this paper, we first discuss the properties of eigenvalues for the main known circulant, skew circulant and sine transform based preconditioners. A counter example shows that E.Boman's preconditioner is only positive semi-definite for the banded Toeplitz matrix. To use preconditioner effectively, then we propose a modified Boman's preconditioner and a new Cesaro sum type sine transform based preconditioner. Finally, the results of numerical experimentation with these two preconditioners are presented.

Key words: Preconditioner, Toeplitz systems, The fast sine transform, Conjugate gradient algorithm.

1. Introduction

Strang[1] first studied the use of circulant matrices $\mathbf{C}$ for solving systems of linear equations $T_n x = b$ with

$$T_n := T(t_0, t_1, \cdots, t_{n-1}) = \begin{pmatrix} t_0 & t_1 & \cdots & \cdots & t_{n-1} \\ t_1 & t_0 & t_1 & & \vdots \\ \vdots & t_1 & \ddots & \ddots & \vdots \\ t_{n-2} & \cdots & \cdots & t_0 & t_1 \\ t_{n-1} & t_{n-2} & \cdots & t_1 & t_0 \end{pmatrix} \tag{1.1}$$

a symmetric positive definite Toeplitz matrix.Numerous authors such as T.Chan[2], R.Chan,etc.[3],[4],[5], Tyrtyshnikov[6], Huckle[7] and T.Ku and C.Kuo[8] proposed different families of circulant/skew-circulant preconditioners.

Appling the preconditioned conjugate gradient algorithm(PCGA) to solve the systems $T_n x = b$, we must find a preconditioner $\mathbf{P}$ such that $\mathbf{P} y = d$ can be solved very fast and the eigenvalues of $\mathbf{P}^{-1} T_n$ are clustered around the point one. For the circulant and skew circulant $\mathbf{P}$,$\mathbf{P} y = d$ can be solved in $O(n \log n)$ operations by the fast Fourier transform(FFT). To avoid complex arithmetic, R.Chan,Ng and Wong[10] and E.Boman and Koltrach[11] presented two kinds of sine transform based preconditioners $S(T_n)$ and P_n respectively. For Toeplitz matrix with the bandwidth $2\beta + 1$, sine transform based preconditioners can also keeps banded,only $O(\beta^2 n) + O(\beta n) = O(n)$ operations is required per each iterative step when β is a constant independent of n. Since circulant/skewcirculant PCGA costs $O(n \log n)$ operations by the FFT algorithm, this implies that the complexity for sine transform based PCGA is reduced by an

* Received March 16, 1998; Final revised September 2, 1998.
1) This work is supported by Chinese Natural Science Foundation(No:9601012).

order of the magnitude compared with complexity of circulant/skewcirculant based the PCGA. Furthermore, numerical results given in [10],[11] show that the convergence performance of these two sine transform based preconditioners is better in terms of the number of iterations than that of the optimal circulant preconditioner. However, the known sine transform based preconditioner sometimes fails, hence the purpose of this paper is to seek a more effective sine transform based preconditioner.

This paper is organized as follows. In section §2 we first study the relationship between the eigenvalues of the main known circulant, skew circulant, sine transform based preconditioners and its Fourier series sums of generating function $f(x)$. For the banded Toeplitz matrix,we prove Boman's preconditioner$\mathbf{P}_n$ is only positive semi-definite, a counter example shows that preconditioner $\mathbf{P}_n$ will fail when $f(\frac{l\pi}{n+1}) = 0$ for some $1 \leq l \leq n$. This fact illustrates the conclusion in [11] that preconditioner $\mathbf{P}_n$ is positive definite is wrong. Since $S(T_n)$ may be positive semi-definite(R.Chan,etc.[10] only proved for sufficiently large n,$S(T_n)$ is positive definite with generating function $f(x) \geq m > 0$). Hence, in section §3 we first present a modified sine transform based preconditioner $\tilde{P}_n$ of $\mathbf{P}_n$,then we develop a Fourier series partial sum type and Cesàro sum sine transform based preconditioners P_N and $S_N(T_n)$ respectively.In section §4 we study the clustering properties for various sine transform based preconditioners. Finally, we present some numerical experimentations confirming our theoretical results.

2. The Eigenvalues of Various Preconditioners

Let us begin to introduce a real function $f(x)$ related to infinite Toeplitz matrx T_∞,namely

$$f(x) = \sum_{k=-\infty}^{+\infty} t_k \exp\{ikx\}, i = \sqrt{-1}, x \in [0, 2\pi] \tag{2.1}$$

The partial sums and Cesàro sums are defined by

$$f_n(x) = \sum_{k=-n}^{n} t_k \exp\{ikx\}, \sigma_N(x) = \frac{1}{N+1} \sum_{n=0}^{N} f_n(x), x \in [0, 2\pi] \tag{2.2}$$

respectively. For symmetric Toeplitz matrix,$t_k = t_{-k}$.

Let S_n stands for the discrete sine transform matrix

$$S_n = \sqrt{\frac{2}{n+1}}(\sin(\frac{ij\pi}{n+1}))_{i,j=1}^{n} \tag{2.3}$$

define the set

$$B_{n\times n} = \{B \in R^{n\times n} : S_n B S_n \text{ is a diagonal matrix}\} \tag{2.4}$$

If $T_n = T(t_0, t_1, \cdots, t_{n-1})$ is a Toeplitz matrix defined as Eq.(1.1), we indicate $H(T_n)$ as the following $n \times n$ matrix

$$H(T_n) = \begin{pmatrix} t_2 & \cdots & t_{n-1} & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ & & & & & 0 \\ t_{n-1} & & & & & t_{n-1} \\ 0 & \ddots & & \ddots & & \vdots \\ \vdots & \ddots & & & & \\ 0 & \cdots & 0 & t_{n-1} & \cdots & t_2 \end{pmatrix} \tag{2.5}$$

Lemma 1. (see [10]) $B_{n\times n} = \{T_n - H(T_n), T_n = T(t_0, t_1, \cdots, t_{n-1})$ *is a Toeplitz matrix.*$\}$

Let e_i denote the i-th unit vector in R^n, $Q_i = T_n(e_i) - H(T_n(e_i)), i = 1, 2, \cdots, n$, then we have

Lemma 2. (see [11]) *Let* $n \in Z^+$ *be given, then* $\{Q_i\}_{i=1}^n$ *is a basis for* $B_{n\times n}$. *Moreover,if* $i \geq 2$ *then the spectrum of* Q_i *is*

$$\{2\cos(\frac{(i-1)m\pi}{n+1})\}_{m=1}^n$$

Boman and Koltrach[11] proposed using the following sine transform based preconditioner P_n

$$P_n = \sum_{i=1}^n t_{i-1} Q_i \tag{2.6}$$

If T_n is a banded Toeplitz matrix with the bandwidth $2b+1$, then P_n become a banded preconditioner P_b(see [11]). Let $l = \lfloor \frac{n}{2} \rfloor, \alpha_k = \sum_{j=\frac{k}{2}+1}^{l-1} t_{2j}$ for $k \geq 0, k$ even; $\alpha_k = \sum_{j=\frac{k+1}{2}}^{l-1} t_{2j+1}$ for $k \geq 1$, k odd; $z_i = t_{i-1} + \frac{2}{n+1}\alpha_{i-1}, i = 1, 2, \cdots, n$, we can see that preconditioner $S(T_n)$ defined in [10] can be expressed as

$$S(T_n) = \sum_{i=1}^n z_i Q_i \tag{2.7}$$

Now we discuss the properties of these two preconditioners. We first compute the eigenvalues of P_n and $S(T_n)$.Construct function

$$g_n(x) = \frac{2}{n+1}[\alpha_0 + (\sum_{k=1}^{n-1}(t_k + 2\alpha_k)\cos kx - t_n \cos nx)]$$

It is easy to obtain $\lim_{n\to\infty} g_n(x) = 0$ uniformly holds for $x \in [-\pi, \pi]$, provided that $f(x) \in L_2$.

Theorem 1. *The eigenvalues of the matrices* P_n *and* $S(T_n)$ *can be expressed by*

$$\lambda_k(P_n) = f_{n-1}(\frac{k\pi}{n+1}), \lambda_k(S(T_n)) = \sigma_n(\frac{k\pi}{n+1}) + g_n(\frac{k\pi}{n+1}), k = 1, 2, 3, \cdots, n, \tag{2.8}$$

respectively.

Proof. By Lemma 2 and $t_j = t_{-j}, j = 1, 2, \cdots, n-1$,then $f_{n-1}(x) = t_0 + 2\sum_{j=1}^{n-1} t_j \cos jx$ and

$$\begin{aligned}\lambda_k(P_n(n)) &= \sum_{j=1}^n t_{j-1}\lambda_k(Q_j)\\ &= t_0 + 2\sum_{j=2}^n t_{j-1}\cos\tfrac{(j-1)k\pi}{n+1}\\ &= t_0 + 2\sum_{j=1}^{n-1} t_j \cos\tfrac{jk\pi}{n+1}\\ &= f_{n-1}(\tfrac{k\pi}{n+1})\end{aligned}$$

since

$$\begin{aligned}\sigma_n(\tfrac{k\pi}{n+1}) &= \tfrac{1}{n+1}\sum_{j=0}^n\sum_{m=-j}^j t_m \exp\{\tfrac{imk\pi}{n+1}\} = \tfrac{1}{n+1}(\sum_{j=0}^n(t_0 + 2\sum_{m=1}^j t_m\cos\tfrac{mk\pi}{n+1}))\\ &= t_0 + \tfrac{2}{n+1}\sum_{j=1}^n\sum_{m=1}^j t_m\cos\tfrac{mk\pi}{n+1} = t_0 + \tfrac{2}{n+1}\sum_{m=1}^n(n+1-m)t_m\cos\tfrac{mk\pi}{n+1}\end{aligned}$$

On the other hand, with the help of Lemma 2 we get

$$\begin{aligned}\lambda_k(S(T_n)) &= z_1 + 2\sum_{j=1}^{n-1} z_{j+1}\cos\tfrac{jk\pi}{n+1}\\ &= t_0 + \tfrac{2}{n+1}\alpha_0 + 2\sum_{j=1}^{n-1}(\tfrac{n-j+2}{n+1}t_j + \tfrac{2}{n+1}\alpha_j)\cos\tfrac{jk\pi}{n+1}\end{aligned}$$

If one defines $\alpha_{-k}=\alpha_k, k=1,2,\cdots,n$,then

$$\begin{array}{rcl}\lambda_k(S(T_n))-\sigma_n(\frac{k\pi}{n+1}) & = & \frac{2}{n+1}\alpha_0+2\sum_{j=1}^{n-1}(\frac{n-j+2}{n+1}t_j\\ & + & \frac{2}{n+1}\alpha_j-\frac{n-j+1}{n+1}t_j)\cos\frac{jk\pi}{n+1}-\frac{2}{n+1}t_n\cos\frac{nk\pi}{n+1}\\ & = & g_n(\frac{k\pi}{n+1})\end{array}$$

hence the results follows.

To obtain the eigenvalues of various mainly preconditioners, in the following we recall the definition of these preconditioners.

Strang's circulant preconditioner C_s(see [1]):if $n=2k, C_s:=T(t_0,t_1,\cdots,t_{k-1},2t_k,t_{k-1},\cdots,t_1)$, if $n=2k+1, C_s:=T(t_0,t_1,\cdots,t_k,t_k,\cdots,t_1)$.
note:for simplicity, we take C_s differs from the Strang's preconditioner costruction only by the diagonal with $\lfloor\frac{n}{2}\rfloor$.

T.Chan's circulant preconditioner C_f(see [2]). $C_f:=T(t_0,\frac{(n-1)t_1+t_{n-1}}{n},\cdots,\frac{t_{n-1}+(n-1)t_1}{n})$;

T.Ku and C.Kuo's circulant preconditioner C_K(see [8]).$C_K:=T(t_0,t_1+t_{n+1},t_2+t_{n-2},\cdots,t_{n-1}+t_1)$ and skew circulant preconditioner:$S_k:=T(t_0,t_1-t_{n-1},\cdots,t_{n-1}-t_1)$;

T.Huckle's circulant preconditioner C_T(see [7]) is determined by its eigenvalues $\lambda_m(C_T)=t_0+2\sum_{j=1}^{p}(1-\frac{j}{p})t_j\cos\frac{2jm\pi}{p}$ and skew circulant preconditioner $S_R:=T(t_0,\frac{(n-1)t_1-t_{n-1}}{n},\cdots,\frac{t_{n-1}-(n-1)t_1}{n}), S_s:=T(t_0,t_1,t_2,\cdots,-t_2,-t_1)$.

Make the similar discussion to those of Theorem 1, we have

Theorem 2. *The eigenvalues of preconditioners C_s,C_f,C_K,S_K,S_R and S_s can be expressed by*

$$\lambda_m(C_s)=f_{[\frac{n}{2}]}(\frac{2m\pi}{n}),\lambda_m(C_f)=\sigma_{n-1}(\frac{2m\pi}{n}) \tag{2.9}$$

$$\lambda_m(C_K)=f_{n-1}(\frac{2m\pi}{n}),\lambda_m(S_K)=\sigma_{n-1}(\frac{(2m+1)\pi}{n}) \tag{2.10}$$

$$\lambda_m(S_R)=f_{n-1}(\frac{(2m+1)\pi}{n}),\lambda_m(S_s)=\sigma_p(\frac{2m\pi}{n}),m=0,1,2,\cdots,n-1. \tag{2.11}$$

Remark. Eq.(2.11) see also [9],[13] for example.

By Theorem 1 and Theorem 2, the previous preconditioners can be divided into two kinds:

- The partial sums type such as P_n,C_s,C_K,S_R;
- Cesàro sums type such as $S(T_n),C_f,C_T,S_K$.

Futhermore, make similar numerical experimentation to those in [9],[14] we conclude that, for the nonbanded matrix, if generating function $f(x)$ is sufficiently smooth, then no essential difference between the partial sums and the Cesàro sums type preconditioners. While for the banded matrix, the convergence rate of sine transform based preconditioners P_n and $S(T_n)$ (see [10],[11]) behave "better" than circulant preconditioner C_f. Moreover, P_n and $S(T_n)$ can also keep the banded. This fact may lead to the saving of arithmetic operations per preconditioning iterations, this implies that sine transform based preconditioners are more effective than other preconditioners. However,in the following Theorem 3 we will see that matrices P_n and $S(T_n)$ may be positive semi-definite and are not suitable for preconditioners at this time.

Theorem 3. *If T_∞ is a symmetric positive definite(SPD) Toeplitz matrix with the bandwidth $2b+1$, then for $2b+1<n$ P_n is only positive semi-definite.*

Proof. Since T_∞ is a banded matrix, then $P_n=P_b=\sum_{k=1}^{b+1}t_{k-1}Q_k$, for $n>b, f_n(x)=f_b(x)=f(x)=\sum_{k=-b}^{b}t_k\exp\{ikx\}$. Moreover, T_∞ is positive definite, $T_\infty=U(\infty)U(\infty)^T$, where $U(\infty)$ is an upper triangular Toeplitz matrix(see e.g.,[12,proof of Theorem 1.1] or [11])

$$U(\infty)=\begin{pmatrix} c_0 & c_1 & \cdots & c_b & 0 & \cdots & \cdots \\ 0 & c_0 & \cdots & \cdots & c_b & 0 & \cdots \\ \vdots & & \ddots & & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & & \end{pmatrix}$$

hence

$$T_n = UU^T \tag{2.12}$$

where

$$U=\begin{pmatrix} c_0 & c_1 & \cdots & c_b & 0 & \cdots & \cdots & 0 \\ 0 & c_0 & \cdots & & c_b & 0 & \cdots & \\ \vdots & & \ddots & \ddots & & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & c_0 & c_1 & \cdots & c_b \end{pmatrix} \in R^{n\times(n+b)} \tag{2.13}$$

Make function $h(x)=\sum_{k=0}^{b} c_k \exp\{ikx\}$, by (2.12) we get

$$t_k=\sum_{j=0}^{b-k} c_j c_{j+k}, k=0,1,2,\cdots,b,$$

hence

$$\begin{aligned} \mid h(x)\mid^2 = & \ (\textstyle\sum_{k=0}^{b} c_k \exp\{ikx\})(\sum_{l=0}^{b} c_l \exp\{-ilx\}) \\ = & \ \textstyle\sum_{k=0}^{b}\sum_{l=0}^{b} c_k c_l \exp\{i(k-l)x\} = t_0 + 2\sum_{k=1}^{b} t_k \cos kx \end{aligned}$$

On the other hand, $t_{-k}=t_k$, then $f(x)=f_b(x)=t_0+2\sum_{k=1}^{b} t_k\cos kx = \mid h(x)\mid^2$,by Theorem 1, the eigenvalues of P_n is $\lambda_k(P_n)=f_{n-1}(\frac{k\pi}{n+1})=\mid h(\frac{k\pi}{n+1})\mid^2$,this shows $P_n(n)$ is a nonnegative definite matrix. On the other hand, take $n=6, b=2, c_0=1, c_1=-2\cos\frac{\pi}{7}, c_2=1$, by Eq.(2.12),(2.13), T_6 is a positive definite matrix and related infinite Toeplitz matrix T_∞ is also positive definite. However, $\lambda_1(P_6)=\mid h(\frac{\pi}{7})\mid^2=\mid c_0+c_1\exp\{\frac{\pi i}{7}\}+c_2\exp\{\frac{2\pi i}{7}\}\mid^2=0$, hence P_n is only positive semi-definite.

Remark. Theorem 4.2 in [11] that P_n is a positive definite matrix is wrong.

3. The Design of Positive Definite Sine Transform Based Preconditioners

Since preconditioner for solving positive definite equations must be positive definite. For a banded Toeplitz matrix, by Theorem 3 we can see that P_n will fail if $f_{n-1}(\frac{k\pi}{n+1})=0$ for some $1\le k\le n$. Moreover, for nonnegative generating function $f(x)$, under certain conditions, R.Chan,etc.[10] proved that for sufficiently large n,$S(T_n)$ is a positive definite matrix. However, for fixed n, this fact cannot guarantee that $S(T_n)$ is also positive definite. Hence preconditioners P_n and $S(T_n)$ may fail. To ensure preconditioner positive definite, we need to look for new effective preconditioners.

3.1 Modified sine transform based preconditioner $\tilde{P}_N$

Take a fixed $\epsilon>0, N\le n$ such that $\lim_{n\to\infty} N=+\infty$, preconditioner $\tilde{P}_N$ is defined by

$$S_n\tilde{P}_N S_n = diag\{\tilde{\lambda}_1,\tilde{\lambda}_2,\cdots,\tilde{\lambda}_n\} \tag{3.1}$$

where

$$\tilde{\lambda}_k=\begin{cases} f_{N-1}(\frac{k\pi}{n+1}), & if\, f_{N-1}(\frac{k\pi}{n+1})>0, \\ \epsilon, & \text{otherwise}; k=1,2,\cdots,n. \end{cases}$$

it is obviously that the preconditioner $\tilde{P}_N$ is a positive definite matrix. If $N = n, \tilde{P}_N$ can be considered as the improvement of P_n.

3.2 Cesàro sum type sine transform based preconditioner $S_N(T_n)$

For $N \leq n$,take

$$S_N(T_n) = \sum_{k=1}^{N}(1 - \frac{k-1}{N})t_{k-1}Q_k \tag{3.2}$$

Theorem 4. *The eigenvalues $\lambda_k(S_N(T_n))$ of matrix $S_N(T_n)$ are:*

$$\sigma_{N-1}(\frac{k\pi}{n+1}), k = 1, 2, \cdots, n.$$

Proof. By Eq.(2.1)

$$\sigma_{N-1}(x) = \frac{1}{N}\sum_{l=0}^{N-1} f_l(x) = \frac{1}{N}\sum_{l=0}^{N-1}\sum_{m=-l}^{l} t_m \exp\{imx\}$$

where,$i = \sqrt{-1}$.Whence from $t_{-m} = t_m$ yields

$$\begin{aligned} \sigma_{N-1}(\tfrac{k\pi}{n+1}) &= \tfrac{1}{N}\textstyle\sum_{l=0}^{N-1}\sum_{m=-l}^{l} t_m \exp\{\tfrac{imk\pi}{n+1}\} \\ &= \tfrac{1}{N}[\textstyle\sum_{l=0}^{N-1}(t_0 + 2\sum_{m=1}^{l} t_m \cos\tfrac{mk\pi}{n+1})] \\ &= t_0 + \tfrac{2}{N}[\textstyle\sum_{l=1}^{N-1}\sum_{m=1}^{l} t_m \cos\tfrac{mk\pi}{n+1}] = t_0 + \tfrac{2}{N}\sum_{m=1}^{N-1}(N-m)t_m \cos\tfrac{mk\pi}{n+1} \end{aligned}$$

on the other hand, by Eq.(3.2) and Lemma 2, then

$$\begin{aligned} \lambda_k(S_N(T_n)) &= \textstyle\sum_{j=1}^{N}(1 - \tfrac{j-1}{N})t_{j-1}\lambda(Q_j) \\ &= t_0 + 2\textstyle\sum_{j=2}^{N}(1 - \tfrac{j-1}{N})t_{j-1}\cos\tfrac{(j-1)k\pi}{n+1} = t_0 + \tfrac{2}{N}\sum_{j=1}^{N-1}(N-j)t_j \cos\tfrac{jk\pi}{n+1}, \end{aligned}$$

hence Theorem holds.

Corollary. *Let T_∞ be same as that in Theorem 3, then $S_N(T_n)$ is a positive definite Toeplitz matrix with the bandwidth $2b+1$.*

Proof. By the definition of Q_i and Eq.(3.2),we get $S_N(T_n)$ to be a matrix with the bandwidth $2b+1$. To prove $S_N(T_n)$ is positive definite, from Theorem 4, we only prove $\sigma_{N-1}(\frac{k\pi}{n+1}) > 0$ for $k = 1, 2, \cdots, n$. By the theory of the Fourier series

$$\sigma_{N-1}(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{N}(\frac{\sin\frac{N(t-x)}{2}}{\sin\frac{t-x}{2}})^2 f(t)\,dt$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{N}(\frac{\sin\frac{N(t-x)}{2}}{\sin\frac{t-x}{2}})^2\,dt \equiv 1 \tag{3.3}$$

From Theorem 3 we get $f(x) = \mid h(x) \mid^2 = t_0 + 2\sum_{k=1}^{b}\cos kx$, then the number of the roots for $f(x) = 0, x \in [0, 2\pi]$ is no more than $2b+1$. By Eq.(3.3) to obtain $\sigma_{N-1}(x) > 0$ uniformly holds for $x \in [-\pi, \pi]$, this shows $S_N(T_n)$ is a positive definite matrix.

4. The Clustering Property of the Preconditioners

To study the clustering property of the eigenvalues for preconditioners,we introduce two notations.

Definition 1. *Let $\gamma_n(\epsilon)$ be the number of $\lambda_k^{(n)} \notin (\mu - \epsilon, \mu + \epsilon)$,then the point μ is called a cluster of the sequences $\{\lambda_k^{(n)}\}_{k=1}^{n}$ if*

$$\lim_{n\to\infty}\frac{\gamma_n(\epsilon)}{n} = 0. \tag{4.1}$$

A cluster is called proper if $\gamma_n(\epsilon) \leq c(\epsilon)$,where $c(\epsilon)$ does not depend on n.

By the theory of numerical algebra,if the point one is a cluster of the preconditioned matrix,then the convergence rate for the PCGA will become superlinear(see [9] for example).

Definition 2. *Let $f(x)$ be Lebesgue integrable function with period 2π, then $f(x)$ is called slightly vanishing if*

$$\lim_{\epsilon\to 0^+} \int_{-\pi}^{\pi} \psi_\epsilon(| f(x) |)\, dx = 0 \tag{4.2}$$

where

$$\psi_\epsilon(x) = \begin{cases} 1, & x \in [0, \epsilon], \\ 0, & \text{otherwise.} \end{cases}$$

Theorem 5. *If T_n is a symmetric positive definite(SPD) Toeplitz matrix with bandwidth $2b+1, b \leq N < n$ such that $\lim_{n\to\infty} N = +\infty$, then the eigenvalues of $\tilde{P}_N^{-1} T_n$ is clustered around 1. More precisely*

$$\tilde{P}_N^{-1} T_n = I + R \tag{4.3}$$

where rank$R \leq 2b + r - 2, r$ is the number of zero for $f_N(\frac{k\pi}{n+1}) = 0$ and $r \leq b$, that is, all but $2b + r - 2$ eigenvalues of $(\tilde{P}_N)^{-1} T_n$ are precisely equal to one.

Proof. Let $z = \exp\{ix\}$, for $N \geq b, f_{N-1}(x) =| h(z) |^2 =| \sum_{l=0}^{b} c_l z^l |^2$, hence there exists at most b roots for $f_{N-1}(x) = 0, x \in [0, 2\pi]$, hence from Eq.(3.1) we get $\tilde{P}_N = P_N + R_1$, where rank$R_1 = r \leq b$. Moreover, by Lemma 1 then

$$\tilde{P}_N - T_n = \begin{bmatrix} t_2 & \cdots & t_b & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & 0 \\ & & & & & & \vdots \\ t_b & & & & & & \\ 0 & & & 0 & & & 0 \\ 0 & & & & & & t_b \\ \vdots & & & & & & \vdots \\ 0 & \cdots & & \cdots & t_b & & t_2 \end{bmatrix} + R_1 \equiv H_1, \; rank\, H_1 \leq 2b - 2 + r.$$

take $R = -(\tilde{P}_N)^{-1} H_1$,then $(\tilde{P}_N)^{-1} T_n = I + R$,where rank$R \leq 2b + r - 2$, hence the results follows.

Theorem 6. *Let $f(x) \in L_2$ be the generating function of Eq.(1.1) with $f(x) \geq m > 0$, then the point one is a cluster of the eigenvalues of $(S_n(T_n))^{-1} T_n$.*

Proof. By $f(x) \geq m > 0$ and Eq.(3.3), we have $\sigma_{n-1}(x) \geq m$, hence $\lambda_k(S_n(T_n)) = \sigma_{n-1}(\frac{k\pi}{n+1}) \geq m > 0, k = 1, 2, \cdots, n$. and

$$|| S_n(T_n)^{-1} ||_2 \leq \frac{1}{m} \tag{4.4}$$

let $z_0 = t_0, z_k = \frac{n-k}{n} t_k, k = 1, 2, \cdots, n-1, x_0 = 0, x_k = \frac{k}{n} t_k, k = 1, 2, \cdots, n-1; y_k = \frac{k+2}{n} t_{k+2}, k = 0, 1, 2, \cdots, n-3, y_{n-2} = y_{n-1} = 0; J(T_n) = (y_0, y_1, \cdots, y_{n-1})$,
from Eq.(3.2) and after some manipulations, we have

$$T_n - S_n(T_n) = T(x_0, x_1, \cdots, x_{n-1}) + H(J(T_n)) \tag{4.5}$$

and

$$\frac{1}{n} || T_n - S_n(T_n) ||_F^2 = \frac{2}{n} (\sum_{k=1}^{n-1} (n-k)(\frac{k}{n})^2 t_k^2 + \sum_{k=2}^{n-1} (k-1)(\frac{k}{n})^2 t_k^2)$$

since $\sum_{k=0}^{+\infty} t_k^2 < +\infty$,for all $\epsilon > 0$,we can always find positive integers N_3 such that $\sum_{k=N_3}^{+\infty} t_k^2 < \frac{\epsilon}{16}$ and for fixed N_3 exists N_1 only if $n \geq N_1$ then

$$\frac{2}{n}\sum_{k=1}^{N_3}(n-k)(\frac{k}{n})^2 t_k^2 \leq \frac{2}{n^2}\sum_{k=1}^{N_3} k^2 t_k^2 < \frac{\epsilon}{8}$$

$$\frac{2}{n}\sum_{k=2}^{N_3}(k-1)(\frac{k}{n})^2 t_k^2 < \frac{\epsilon}{4}$$

Thus we have

$$\begin{array}{ll} \frac{2}{n}[& \sum_{k=1}^{n-1}(n-k)(\frac{k}{n})^2 t_k^2 + \sum_{k=2}^{n-1}(k-1)(\frac{k}{n})^2 t_k^2] \\ \leq & \frac{2}{n}\sum_{k=1}^{N_3}(n-k)(\frac{k}{n})^2 t_k^2 + 2\sum_{k=N_3+1}^{+\infty} t_k^2 \\ + & \frac{2}{n}\sum_{k=2}^{N_3}(n-1)(\frac{k}{n})^2 t_k^2 + 2\sum_{k=N_3+1}^{+\infty} t_k^2 \\ < & \epsilon \end{array}$$

namely

$$\| T_n - S_n(T_n) \|_F^2 = 0(n) \tag{4.6}$$

by Eq.(4.4) then

$$\begin{array}{ll} \| & (S_n(T_n))^{-1}T_n - I_n \|_F^2 = \| (S_n(T_n))^{-1}(T_n - S_n(T_n) \|_F^2 \\ \leq & \| S_n(T_n)^{-1} \|_2^2 \| T_n - S_n(T_n) \|_F^2 \\ \leq & \frac{1}{m^2} \| T_n - S_n(T_n) \|_F^2 = 0(n) \end{array}$$

From Lemma 2.1 in [13] we get the conclusion of Theorem 6 holds.

Make the similar analysis to those in [13] we can obtain the following theorems.

Theorem 7. *Let $f(x) \in L_2$ be generating function of Eq.(1.1) with nonnegative slightly vanishing, if P_N and $S(T_n)$ are all symmetric positive definte, then the point one is a cluster of the eigenvalues of matrices $P_N^{-1}T_n$,*
$S(T_n)^{-1}T_n, \tilde{P}_N^{-1}T_n$.

Theorem 8. *Let $f(x) \in L_2$ be complex generating function of Eq.(2.1) with slightly vanishing, then the point one is a cluster of the singular values of matrix $P_N^{-1}T_n$(or $S(T_N)^{-1}T_n$, $\tilde{P}_N^{-1}T_n$) when P_N(or $S(T_n), \tilde{P}_N$) is an invertible matrix.*

5. Numerical Results

To test the behavor of the preconditioning stratgies introduced so far, we have used NDP-FORTRAN language, implemented on an IBM/586 machine. Our main term of comparsion will be T.Chan's optimal circulant preconditioner C_f[2], Boman's sine transform based preconditioner P_b[11], modifed preconditioner $\tilde{P}_N$ and Cesàro sum type preconditioner $S_n(T_n)$. In all the examples, the right hand side of the system was $(1,1,\cdots,1)^T$, and the zero vector was used as an the initial guess, the stopping criterion is residul vector was reduced by a factor less than 10^{-7}.

If one defines $P_1 \equiv C_f, P_2 \equiv P_n, P_3 \equiv \tilde{P}_n, P_4 \equiv S_n(T_n)$.

The PCGA was used to solve $T_n x = b$ for each of the following examples.

Example 1. $T_n^{(1)} = [\frac{\cos(|i-j|)}{|i-j|+1}]_{i,j=1}^n, n = 63, 255, 511, 1023, 2047, 4095.$

Example 2. $T_n^{(2)} = [\frac{1}{(i-j)^2+1}]_{i,j=1}^n, n = 63, 255, 511, 1023, 2047, 4095.$

Example 3. $T_n^{(3)} = [\frac{1}{2^{|i-j|}}]_{i,j=1}^n, n = 63, 255, 511, 1023, 2047, 4095.$

TABLE 1

Number of iterations for convergence

	$T_n^{(1)}$				$T_n^{(2)}$				$T_n^{(3)}$			
n	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4
63	6	6	6	6	4	4	4	4	3	3	3	3
255	6	6	6	7	4	4	4	4	3	3	3	3
511	7	7	7	7	4	4	4	4	3	3	3	3
1023	7	7	7	7	4	4	4	4	3	3	3	4
2047	7	7	7	7	4	4	4	4	3	3	3	4
4095	7	7	7	7	4	4	4	4	3	3	3	4

Example 4. The generating function of the Toeplitz matrix $T_\infty^{(4)}$ is $f(x) = \mid h(x) \mid^2 = \Pi_{j=1}^{b} \mid 1 - \alpha_j \exp\{ijx\} \mid^2, i = \sqrt{-1}, b = 20; n = 63, 255, 1023, 2047, 4095;$ $\alpha_j = 0.745(0.1 \times j - 1), j = 1, 2, \cdots, 20. T_n^{(4)}$ is a finite section of $T_\infty^{(4)}$ with the bandwidth 41.

Example 5. Let $h(x) = 1 - 2\cos x \exp\{ix\} + \exp\{2ix\}$,the generating function of $T_\infty^{(5)}$ is $f_1(x) = \mid h_1(x)h(x) \mid^2, n = 63, 255, 511, 1023, 2047, 4095;$ other assumptions are same as example 4,$T_n^{(5)}$ is a Toeplitz matrix with the bandwidth 44.

TABLE 2

Number of iterations for convergence

	$T_n^{(4)}$				$T_n^{(5)}$			
n	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4
63	55	10	10	10	56	fails	55	12
255	48	9	9	10	49	fails	48	11
511	37	8	8	9	38	fails	37	10
1023	30	9	9	9	31	fails	38	10
2047	21	9	9	10	21	fails	20	10
4095	17	10	10	11	18	fails	17	11

For the nonbanded problems, by **TABLE 1** we see that the convergence rate of sine transform based preconditioners $P_n, \tilde{P}_n$ and $S_n(T_n)$ are competitive optimal circulant preconditioner C_f. While for the banded problems, if generating function $f(x) \geq m > 0$, then all sine transform preconditioners such as $P_N, \tilde{P}_N$ and $S_n(T_n)$ are superior to optimal circulant preconditioner C_f. On the other hand, from example 5 we know that if nonnegative generating function $f(x)$ with $f(\frac{k\pi}{n+1}) = 0$ for some $1 \leq k \leq n$, then P_n will fail, but $S_n(T_n)$ is still positive definite, this fact confirms the conclusion in Theorem 3. Futhermore,from **TABLE 2** we see if preconditioner P_n fails, then the convergence rate of the preconditioners C_f and modified preconditioner $\tilde{P}_N$ is approximately same and inferior to that of preconditioner $S_n(T_n)$, the reason may be that $S_n(T_n)$ keeps the banded matrix structure while $\tilde{P}_N$ is no longer banded.

References

[1] G. Strang, A proposal for Toeplitz matrix calculations, *Stud. Appl. Math.,* **74** (1986), 171-176.

[2] T. Chan, An optimal circulants preconditioner for Toeplitz systems, *SIAM. J. Sci. Stat. Comput.,* **9** (1988), 766-771.

[3] R. Chan, M. Yeung, Circulant preconditioners for Toeplitz matrices with positive continuous generating function, *Math. Comp.,* **58** (1992), 233-240.

[4] ——, Circulant preconditioners constructed from kernels, *SIAM. J. Numer. Anal.,* **29** (1992), 1092-1103.

[5] R. Chan, Circulant preconditioners for Hermitian Toeplitz systems, *SIAM. J. Matrix Anal. Appl.,* **10** (1989), 542-550.

[6] E. Tyrtyshnikov, Optimal and superoptimal circulantpreconditioners, *SIAM. J. Matrix Anal. Appl.,* **13** (1992), 459-473.

[7] T.Huckle, Circulant and skew circulant matrices for solving Toeplitz matrix problems, *SIAM. J. Matrix Appl.,* **13** (1992), 767-777.

[8] T.Ku, C. Kuo, Design and Analysis of Toeplitz Preconditioners, IEEE Trans. Signal Process., **40** (1992), 129-141.

[9] R. Chan, M. Ng, Conjugate gradient methods for Toeplitz systems, *SIAM Review,* **38** (1996), 427-482.

[10] R. Chan, M. Ng, C. Wong, Sine transform based preconditioners for symmetric Toeplitz systems, *Linear Algebra Appl.,* **232** (1996), 237-259.

[11] E.Boman, I.Koltrach, Fast transform based preconditioners for Toeplitz equations, *SIAM J. Matrix Anal. Appl.,* **16** (1995), 628-645.

[12] I. Gohberg, I. Feldman, Convolution Equations and Projection Methods for Their Solution, American Mathematical Soceity, RI. 1974.

[13] E. Tyrtshnikov, Circulant preconditioner with unbounded inverses, *Linear Algebra Appl.,* **216** (1995), 1-24.

[14] V. Strela, E. Tyrtyshnikov, Which circulant preconditioner is better? *Math. Comp.,* **213** (1996), 137-150.