Circulant Preconditioned Toeplitz Least Squares Iterations

Raymond H. Chan,* James G. Nagy† and Robert J. Plemmons‡

November 1991, Revised April 1992

Abstract

We consider the solution of least squares problems $\min ||b - Tx||_2$ by the preconditioned conjugate gradient method, for m-by-n complex Toeplitz matrices T of rank n. A circulant preconditioner C is derived using the T. Chan optimal preconditioner on n-by-n Toeplitz row blocks of T. For Toeplitz T that are generated by 2π-periodic continuous complex-valued functions without any zeros, we prove that the singular values of the preconditioned matrix TC^{-1} are clustered around 1, for sufficiently large n. We show that if the condition number of T is of $O(n^\alpha)$, $\alpha > 0$, then the least squares conjugate gradient method converges in at most $O(\alpha \log n + 1)$ steps. Since each iteration requires only $O(m \log n)$ operations using the FFT, it follows that the total complexity of the algorithm is then only $O(\alpha m \log^2 n + m \log n)$. Conditions for *superlinear convergence* are given and regularization techniques leading to superlinear convergence for least squares computations with ill-conditioned Toeplitz matrices arising from inverse problems are derived. Numerical examples are provided illustrating the effectiveness of our methods.

Abbreviated Title. Toeplitz Least Squares Iterations.

Key Words. Least squares, Toeplitz matrix, circulant matrix, preconditioned conjugate gradients, regularization.

AMS(MOS) Subject Classifications. 65F10, 65F15

*Department of Mathematics, University of Hong Kong, Hong Kong.

†Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455.

‡Department of Mathematics and Computer Science, Wake Forest University, P.O. Box 7388, Winston-Salem, NC 27109. Research supported by the US Air Force under grant no. AFOSR-91-0163.

1 Introduction

The conjugate gradient (CG) method is an iterative method for solving Hermitian positive definite systems $Ax = b$, see for instance Golub and van Loan [21]. When A is a rectangular m-by-n matrix of rank n, one can still use the CG algorithm to find the solution to the least squares problem

$$\min \|b - Ax\|_2. \tag{1}$$

This can be done by applying the algorithm to the normal equations in factored form,

$$A^*(b - Ax) = 0, \tag{2}$$

which can be solved by conjugate gradients without explicitly forming the matrix A^*A, see Bjorck [7].

The convergence of the conjugate gradient algorithm and its variations depends on the the singular values of the data matrix A, see Axelsson [5]. If the singular values cluster around a fixed point, convergence will be rapid. Thus, to make the algorithm a useful iterative method, one usually *preconditions* the system. The preconditioned conjugate gradient (PCG) algorithm then solves (1) by transforming the problem with a preconditioner M, applying the conjugate gradient method to the transformed problem, and then transforming back. More precisely, one can use the conjugate gradient method to solve

$$\min \|b - AM^{-1}y\|_2,$$

and then set $x = M^{-1}y$.

In this paper we consider the least squares problem (1), with the data matrix $A = T$, where T is a rectangular m-by-n Toeplitz matrix of rank n. The matrix $T = (t_{jk})$ is said to be *Toeplitz* if $t_{jk} = t_{j-k}$, *i.e.*, T is constant along its diagonals. An n-by-n matrix C is said to be *circulant* if it is Toeplitz and its diagonals c_j satisfy $c_{n-j} = c_{-j}$ for $0 < j \leq n-1$. Toeplitz least squares problems occur in a variety of applications, especially in signal and image processing, see for instance Andrews and Hunt [3], Jain [24] and Oppenheim and Schafer [28].

Recall that the solution to the least squares problem

$$\min \|b - Tx\|_2 \tag{3}$$

can be found by the preconditioned conjugate gradient method by applying the method to the normal equations (2) in factored form, that is, using T and T^* without forming T^*T. The preconditioner M considered in this paper is given by an n-by-n circulant matrix $M = C$, where C^*C is then a circulant matrix that approximates T^*T.

The version of the PCG algorithm we use is given in [7] and can be stated as follows.

Algorithm PCG for Least Squares. *Let $x^{(0)}$ be an initial approximation to $Tx = b$, and let C be a given preconditioner. This algorithm computes the least squares solution, x, to $Tx = b$.*

$$r^{(0)} = b - Tx^{(0)}$$
$$p^{(0)} = s^{(0)} = C^{-*}T^{*}r^{(0)}$$
$$\gamma_0 = \|s^{(0)}\|_2^2$$

for $k = 0, 1, 2, \ldots$

$$q^{(k)} = TC^{-1}p^{(k)}$$
$$\alpha_k = \gamma_k / \|q^{(k)}\|_2^2$$
$$x^{(k+1)} = x^{(k)} + \alpha_k C^{-1}p^{(k)}$$
$$r^{(k+1)} = r^{(k)} - \alpha_k q^{(k)}$$
$$s^{(k+1)} = C^{-*}T^{*}r^{(k+1)}$$
$$\gamma_{k+1} = \|s^{(k+1)}\|_2^2$$
$$\beta_k = \gamma_{k+1}/\gamma_k$$
$$p^{(k+1)} = s^{(k+1)} + \beta_k p^{(k)}$$

The idea of using the preconditioned conjugate gradient method with circulant preconditioners for solving square positive definite Toeplitz systems was first proposed by Strang [30], although the application of circulant approximations to Toeplitz matrices has been used for some time in image processing, *e.g.*, [6] . The convergence rate of the method was analyzed in R. Chan and Strang [9] for Toeplitz matrices that are generated by positive Wiener class functions. Since then, considerable research have been done in finding other good circulant preconditioners or extending the class of generating functions for which the method is effective, see T. Chan [17], R. Chan [10], Tyrtyshnikov [32], Tismenetsky [31], Huckle [23], Ku and Kuo [25], R. Chan and Yeung [13], T. Chan and Olkin [18], R. Chan and Jin [12] and R. Chan and Yeung [14].

Recently, the idea of using circulant preconditioners has been extended to non-Hermitian square Toeplitz systems by R. Chan and Yeung [15] and to Toeplitz least squares problems by Nagy [26] and Nagy and Plemmons [27]. The main aim of this paper is to formalize and establish convergence results, and to provide applications, in the case where T is a rectangular Toeplitz (block) matrix.

For the purpose of constructing the preconditioner, we will see that by extending the Toeplitz structure of the matrix T and, if necessary, padding zeros to the bottom left-hand side, we may assume without loss of generality that $m = kn$ for some positive integer k. This padding is only for convenience in constructing the preconditioner and does not alter the original least squares problem. In the material to follow, we consider the case where k is a constant independent of n. More precisely, we consider in this paper, kn-by-n matrices T of the form

$$T = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_k \end{bmatrix}, \tag{4}$$

where each square block T_j is a Toeplitz matrix. Notice that if T itself is a rectangular Toeplitz matrix, then each block T_j is necessarily Toeplitz.

Following [26, 27], for each block T_j, we construct a circulant approximation C_j. Then our preconditioner is defined as a square circulant matrix C, such that

$$C^*C = \sum_{j=1}^{k} C_j^* C_j.$$

Notice that each C_j is an n-by-n circulant matrix. Hence they can all be diagonalized by the Fourier matrix F, i.e.

$$C_j = F\Lambda_j F^*$$

where Λ_j is diagonal, see Davis [19]. Therefore the spectrum of C_j, $j = 1, \cdots k$, can be computed in $O(n \log n)$ operations by using the Fast Fourier Transform (FFT). Since

$$C^*C = F \sum_{j=1}^{k} (\Lambda_j^* \Lambda_j) F^*,$$

C^*C is also circulant and its spectrum can be computed in $O(kn \log n)$ operations. Here we choose, as in [26, 27],

$$C = F(\sum_{j=1}^{k} \Lambda_j^* \Lambda_j)^{\frac{1}{2}} F^*. \tag{5}$$

The number of operations per iteration in Algorithm PCG for Least Squares depends mainly on the work of computing the matrix-vector multiplications. In our case, this amounts to computing products:

$$Ty, \quad T^*z, \quad C^{-1}y, \quad C^{-*}y$$

for some n-vectors y and m-vectors z. Since

$$C^{-1}y = F(\sum_{j=1}^{k} \Lambda_j^* \Lambda_j)^{-\frac{1}{2}} F^* y,$$

the products $C^{-1}y$ and $C^{-*}y$ can be found efficiently by using the FFT in $O(n \log n)$ operations. For the products Ty and T^*z, with T in block form with k n-by-n blocks T_j, we have to compute n products of the form $T_j w$ where T_j is an n-by-n Toeplitz matrix and w is an n-vector. However the product $T_j w$ can be computed using the FFT by first embedding T_j into a $2n$-by-$2n$ circulant matrix. The multiplication thus requires $O(2n \log(2n))$ operations. It follows that the operations for computing Ty and T^*z are of the order $O(m \log n)$, where $m = nk$. Thus we conclude that the cost per iteration in the preconditioned conjugate gradient method is of the order $O(m \log n)$.

As already mentioned in the beginning, the convergence rate of the method depends on the distribution of the singular values of the matrix TC^{-1} which are the same as the square roots of the eigenvalues of the matrix $(C^*C)^{-1}(T^*T)$. We will show, then, that if the generating functions of the blocks T_j are 2π-periodic continuous functions and if one of these functions has no zeros, then the spectrum of $(C^*C)^{-1}(T^*T)$ will be clustered around 1, for sufficiently large n. We remark that the class of 2π-periodic continuous functions contains the Wiener class of functions which in turn contains the class of rational functions considered in Ku and Kuo [25].

By using a standard error analysis of the conjugate gradient method, we then show that if the condition number $\kappa(T)$ of T is of $O(n^\alpha)$, then the number of iterations required for convergence, for sufficiently large n, is at most $O(\alpha \log n + 1)$ where $\alpha > 0$. Since the number of operations per iteration in the conjugate gradient method is of $O(m \log n)$, the total complexity of the algorithm is therefore of $O(\alpha m \log^2 n + m \log n)$. In the case when $\alpha = 0$, i.e. T is well-conditioned, the method converges in $O(1)$ steps. Hence the complexity is reduced to just $O(m \log n)$ operations, for sufficiently large n. On the other hand, the superfast direct algorithms by Ammar and Gragg [2] require $O(n \log^2 n)$ operations for n-by-n Toeplitz linear systems. The stability of fast direct methods has been studied by Bunch [8].

The outline of the paper is as follows. In §2, we construct the circulant preconditioners C for the Toeplitz least squares problem and study some of the spectral properties of these preconditioners. In §3, we show that the iteration matrix TC^{-1} has singular values clustered around 1. In §4, we then establish the convergence rate of the preconditioned conjugate gradient method when applied to the preconditioned system, and indicate when it is superlinear. In §5, we discuss the technique of regularization when the given Toeplitz matrix T is ill-conditioned. Numerical results and concluding remarks are given in §6.

2 Properties of the Circulant Preconditioner

In this section, we consider circulant preconditioners for least square problems and study their spectral properties. We begin by recalling some results for square Toeplitz systems.

For simplicity, we denote by $\mathcal{C}_{2\pi}$ the Banach space of all 2π-periodic continuous complex-valued functions equipped with the supremum norm $||\cdot||_\infty$. As already mentioned in §1, this class of functions contains the Wiener class of functions. For all $f \in \mathcal{C}_{2\pi}$, let

$$a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-ik\theta} d\theta, \ \ k = 0, \pm 1, \pm 2, \cdots,$$

be the Fourier coefficients of f. Let A be the n-by-n complex Toeplitz matrix with the (j,k)th entry given by a_{j-k}. The function f is called the generating function of the matrix A.

For a given n-by-n matrix A, we let C be the n-by-n circulant approximation of A as defined in T. Chan [17], i.e. C is the minimizer of $F(X) = ||A - X||_F$ over all circulant matrices X. For the special case where A is Toeplitz, the (j,ℓ)th entry of C is given by the diagonal $c_{j-\ell}$ where

$$c_k = \begin{cases} \dfrac{(n-k)a_k + k a_{k-n}}{n} & 0 \le k < n, \\ c_{n+k} & 0 < -k < n. \end{cases} \tag{6}$$

The following three Lemmas are proved in R. Chan and Yeung [15]. The first two give the bounds of $||A||_2$ and $||C||_2$ and the last one shows that $A - C$ has clustered spectrum for certain Toeplitz matrices A.

Lemma 1 *Let $f \in \mathcal{C}_{2\pi}$. Then we have*

$$||A||_2 \le 2||f||_\infty < \infty, \quad n = 1, 2, \cdots. \tag{7}$$

If moreover f has no zeros, i.e.

$$\min_{\theta \in [-\pi, \pi]} |f(\theta)| > 0,$$

then there exists a constant $c > 0$ such that for all n sufficiently large, we have

$$||A||_2 > c. \tag{8}$$

Lemma 2 *Let $f \in \mathcal{C}_{2\pi}$. Then we have*

$$||C||_2 \le 2||f||_\infty < \infty, \quad n = 1, 2, \cdots. \tag{9}$$

If moreover f has no zeros, then for all sufficiently large n, we also have

$$||C^{-1}||_2 \le 2||\frac{1}{f}||_\infty < \infty. \tag{10}$$

Lemma 3 *Let $f \in \mathcal{C}_{2\pi}$. Then for all $\epsilon > 0$, there exist N and $M > 0$, such that for all $n > N$,*

$$A - C = U + V$$

where

$$\text{rank } U \le M$$

and

$$||V||_2 \le \epsilon.$$

Now let us consider the general least squares problem (3) where T is an m-by-n matrix with $m \ge n$. For the purpose of constructing the preconditioner, we assume that $m = kn$, without loss of generality, since otherwise the final block T_k can be extended to an $n \times n$ Toeplitz matrix by extending the diagonals and padding the lower left part with zeros. (This modification is only for constructing the preconditioner. The original least squares problem (3) is not changed.) Thus we can partition T as (4), without loss of generality. We note that the solution to the least square problem (3) can be obtained by solving the normal equations

$$T^* T x = T^* b,$$

in factored form, where

$$T^*T = \sum_{j=1}^{k} T_j^* T_j.$$

Of course one can avoid actually forming T^*T for implementing the conjugate gradient method for the normal equations [7].

We will assume in the following that k is a constant independent of n and that each square block T_j, $j = 1, \cdots, k$ is generated by a generating function f_j in $\mathcal{C}_{2\pi}$. Following Nagy [26], and Nagy and Plemmons [27], we define a preconditioner for T based upon preconditioners for the blocks T_j.

For each block T_j, let C_j be the corresponding T. Chan's circulant preconditioner as defined in (6). Then it is natural to consider the square circulant matrix

$$C^*C = \sum_{j=1}^{k} C_j^* C_j \tag{11}$$

as a circulant approximation to T^*T [27]. Note, however, that C is computed (or applied) using the equation (5). Clearly C is invertible if one of the C_j is. In fact, using Lemma 2, we have

Lemma 4 *Let $f_j \in \mathcal{C}_{2\pi}$ for $j = 1, 2, \cdots, k$. Then we have*

$$||C||_2^2 \leq 4 \sum_{j=1}^{k} ||f_j||_\infty^2 < \infty, \quad n = 1, 2, \cdots. \tag{12}$$

If moreover one of the f_j, say f_ℓ, has no zeros, then for all sufficiently large n, we also have

$$||(C^*C)^{-1}||_2 \leq 4||\frac{1}{f_\ell}||_\infty^2 < \infty. \tag{13}$$

Proof: Equation (12) clearly follows from (11) and (9). To prove (13), we just note that $C_j^* C_j$ are positive semidefinite matrices for all $j = 1, \cdots, k$, hence

$$\lambda_{\min}(C^*C) \geq \lambda_{\min}(C_\ell^* C_\ell),$$

where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue. Thus by (10), we then have

$$||(C^*C)^{-1}||_2 \leq ||(C_\ell^* C_\ell)^{-1}||_2 = ||C_\ell^{-1}||_2^2 \leq 4||\frac{1}{f_\ell}||_\infty^2. \quad \square$$

3 Spectrum of TC^{-1}

In this section, we show that the spectrum of the matrix

$$(C^*C)^{-1}(T^*T)$$

is clustered around 1. It will follow then, that the singular values of TC^{-1} are also clustered around 1, since $(C^*C)^{-1}(T^*T)$ is similar to $(TC^{-1})^*(TC^{-1})$. We begin by analyzing the spectrum of each block.

Lemma 5 *For $1 \le j \le k$, if $f_j \in \mathcal{C}_{2\pi}$, then for all $\epsilon > 0$, there exist N_j and $M_j > 0$, such that for all $n > N_j$,*

$$T_j^*T_j - C_j^*C_j = U_j + V_j$$

where U_j and V_j are Hermitian matrices with

$$\text{rank } U_j \le M_j$$

and

$$||V_j||_2 \le \epsilon.$$

Proof: We first note that by Lemma 3, we have for all $\epsilon > 0$, there exist positive integers N_j and M_j such that for all $n > N_j$,

$$T_j - C_j = \tilde{U}_j + \tilde{V}_j$$

where rank $\tilde{U}_j \le M_j$ and $||\tilde{V}_j||_2 \le \epsilon$. Therefore,

$$\begin{aligned}
& T_j^*T_j - C_j^*C_j \\
= \;& T_j^*(T_j - C_j) + (T_j - C_j)^*C_j \\
= \;& T_j^*(T_j - C_j) - (T_j - C_j)^*(T_j - C_j) + (T_j - C_j)^*T_j \\
= \;& T_j^*(\tilde{U}_j + \tilde{V}_j) - (\tilde{U}_j + \tilde{V}_j)^*(\tilde{U}_j + \tilde{V}_j) + (\tilde{U}_j + \tilde{V}_j)^*T_j \\
\equiv \;& U_j + V_j.
\end{aligned}$$

Here

$$\begin{aligned}
U_j &= T_j^*\tilde{U}_j + \tilde{U}_j^*T_j - \tilde{U}_j^*\tilde{U}_j - \tilde{U}_j^*\tilde{V}_j - \tilde{V}_j^*\tilde{U}_j \\
&= \tilde{U}_j^*(T_j - \tilde{U}_j - \tilde{V}_j) + (T_j - \tilde{V}_j)^*\tilde{U}_j
\end{aligned}$$

and

$$V_j = \tilde{V}_j^*T_j + T_j^*\tilde{V}_j - \tilde{V}_j^*\tilde{V}_j.$$

It is clear that both U_j and V_j are Hermitian matrices. Moreover we have rank $U_j \le 2M_j$ and

$$||V_j||_2 \le 2\epsilon||T_j||_2 + \epsilon^2.$$

By (7), we then have

$$||V_j||_2 \leq 4\epsilon||f_j||_\infty + 2\epsilon^2. \quad \Box$$

Using the facts that

$$T^*T - C^*C = \sum_{j=1}^{k}(T_j^*T_j - C_j^*C_j)$$

and that k is independent of n, we immediately have

Lemma 6 *Let $f_j \in \mathcal{C}_{2\pi}$ for $j = 1, \cdots k$. Then for all $\epsilon > 0$, there exist N and $M > 0$, such that for all $n > N$,*

$$T^*T - C^*C = \tilde{U} + \tilde{V}$$

where $\tilde{U}$ and $\tilde{V}$ are Hermitian matrices with

$$\text{rank } \tilde{U} \leq M \tag{14}$$

and

$$||\tilde{V}||_2 \leq \epsilon. \tag{15}$$

We now show that the spectrum of the preconditioned matrix

$$(C^*C)^{-1}(T^*T)$$

is clustered around 1. We note that this is equivalent to showing that the spectrum of $(C^*C)^{-1}(T^*T) - I$, where I is the n-by-n identity matrix, is clustered around zero.

Theorem 1 *Let $f_j \in \mathcal{C}_{2\pi}$ for all $j = 1, \cdots, k$. If one of the f_j, say f_ℓ, has no zeros, then for all $\epsilon > 0$, there exist N and $M > 0$, such that for all $n > N$, at most M eigenvalues of the matrix*

$$(C^*C)^{-1}(T^*T) - I$$

have absolute values larger than ϵ.

Proof: By Lemma 6, we have

$$(C^*C)^{-1}(T^*T) - I = (C^*C)^{-1}(T^*T - C^*C) = (C^*C)^{-1}(\tilde{U} + \tilde{V}).$$

Therefore the spectra of the matrices

$$(C^*C)^{-1}(T^*T) - I \quad \text{and} \quad (C^*C)^{-1/2}(\tilde{U} + \tilde{V})(C^*C)^{-1/2}$$

are the same. However, by (14), we have

$$\text{rank } \left\{(C^*C)^{-1/2}\tilde{U}(C^*C)^{-1/2}\right\} \leq M$$

and by (15) and (13), we have

$$||(C^*C)^{-1/2}\tilde{V}(C^*C)^{-1/2}||_2 \le ||\tilde{V}||_2||(C^*C)^{-1}||_2 \le 4\hat{\epsilon}||\frac{1}{f_\ell}||_\infty^2,$$

where $\hat{\epsilon}$ replaces the ϵ specified in (15). Thus by applying Cauchy's interlace theorem (see Wilkinson [33]) to the Hermitian matrix

$$(C^*C)^{-1/2}\tilde{U}(C^*C)^{-1/2} + (C^*C)^{-1/2}\tilde{V}(C^*C)^{-1/2},$$

we see that its spectrum is clustered around zero. Hence the spectrum of the matrix $(C^*C)^{-1}(T^*T)$ is clustered around 1. □

From Theorem 1, we have the desired clustering result; namely, if $f_j \in \mathcal{C}_{2\pi}$ for all $j = 1, \cdots, k$ and if one of the f_j has no zeroes, then the *singular values of the preconditioned matrix* TC^{-1} *are clustered around* 1.

4 Convergence Rate of the Method

In this section, we analyze the convergence rate of Algorithm PCG for Least Squares, for our circulant preconditioned Toeplitz matrix TC^{-1}. We show first that the method converges, for sufficiently large n, in at most $O(\alpha \log n + 1)$ steps where $O(n^\alpha)$ is the condition number of T . We begin by noting the following error estimate of the conjugate gradient method.

Lemma 7 *Let* G *be a positive definite matrix and* x *be the solution to* $Gx = b$. *Let* x_j *be the* j*th iterant of the ordinary conjugate gradient method applied to the equation* $Gx = b$. *If the eigenvalues* $\{\delta_k\}$ *of* G *are such that*

$$0 < \delta_1 \le ... \le \delta_p \le \gamma_1 \le \delta_{p+1} \le ... \le \delta_{n-q} \le \gamma_2 \le \delta_{n-q+1} \le ... \le \delta_n,$$

then

$$\frac{||x - x_j||_G}{||x - x_0||_G} \le 2\left(\frac{\gamma - 1}{\gamma + 1}\right)^{j-p-q} \cdot \max_{\delta \in [\gamma_1, \gamma_2]} \left\{ \prod_{k=1}^{p} \left(\frac{\delta - \delta_k}{\delta_k} \right) \right\}. \tag{16}$$

Here

$$\gamma \equiv \left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{1}{2}} \ge 1$$

and $||v||_G \equiv v^*Gv$.

Proof: It is well-known that an error estimate of the conjugate gradient method is given by the following minimax inequality:

$$\frac{||x - x_j||_G}{||x - x_0||_G} \le \min_{\mathcal{P}_j} \max_{k=1,\cdots,n} |\mathcal{P}_j(\delta_k)|,$$

where $\mathcal{P}_j$ is any jth degree polynomial with constant term 1, see Axelsson and Barker [4]. To obtain an upper bound, we first use linear polynomials of the form $(\delta - \delta_k)/\delta_k$ that pass through the outlying eigenvalues δ_k, $1 \le k \le p$ and $n-q+1 \le k \le n$ to minimize the maximum absolute value of $\mathcal{P}_j$ at these eigenvalues. Then we use a $(j-p-q)$th degree Chebyshev polynomial $\mathcal{T}_{j-p-q}$ to minimize the maximum absolute value of $\mathcal{P}_j$ in the interval $[\delta_{p+1}, \delta_{n-q}]$. Then we get

$$\frac{||x-x_j||_G}{||x-x_0||_G} \le \mathcal{T}_{j-p-q}\left[\frac{\gamma_2+\gamma_1}{\gamma_2-\gamma_1}\right]^{-1} \max_{\delta\in[\gamma_1,\gamma_2]} \left\{ \prod_{k=1}^{p} \left(\frac{\delta-\delta_k}{\delta_k}\right) \prod_{k=n-q+1}^{n} \left(\frac{\delta_k-\delta}{\delta_k}\right) \right\}.$$

Equation (16) now follows by noting that for $\delta \in [\gamma_1, \gamma_2]$, we always have

$$0 \le \frac{\delta_k - \delta}{\delta_k} \le 1, \quad n-q+1 \le k \le n.$$

and that

$$\mathcal{T}_{j-p-q}\left[\frac{\gamma_2+\gamma_1}{\gamma_2-\gamma_1}\right]^{-1} \le 2\left(\frac{\gamma-1}{\gamma+1}\right)^{j-p-q},$$

see Axelsson and Barker [4]. $\Box$

For the system

$$(C^*C)^{-1}(T^*T)x = (C^*C)^{-1}T^*b, \tag{17}$$

the iteration matrix G is given by

$$G = (C^*C)^{-1/2}(T^*T)(C^*C)^{-1/2}.$$

By Theorem 1, we can choose $\gamma_1 = 1-\epsilon$ and $\gamma_2 = 1+\epsilon$. Then p and q are constants that depend only on ϵ but not on n. By choosing $\epsilon < 1$, we have

$$\frac{\gamma-1}{\gamma+1} = \frac{1-\sqrt{1-\epsilon^2}}{\epsilon} < \epsilon.$$

In order to use (16), we need a lower bound for δ_k, $1 \le k \le p$. We first note that

$$||G^{-1}||_2 = ||(T^*T)^{-1}(C^*C)||_2 \le \frac{||C||_2^2}{||T||_2^2}\kappa(T^*T).$$

If one of the f_ℓ has no zeros, then by (8), we have for n sufficiently large

$$||T||_2^2 \ge ||T_\ell||_2^2 \ge c$$

for some $c > 0$ independent of n. Combining this with (12), we then see that for all n sufficiently large,

$$||G^{-1}||_2 \le \tilde{c}\cdot\kappa(T^*T) \le \tilde{c}n^{\alpha},$$

for some constant $\tilde{c}$ that does not depend on n. Therefore,

$$\delta_k \geq \min_{\ell} \delta_\ell = \frac{1}{||G^{-1}||_2} \geq cn^{-\alpha}, \quad 1 \leq k \leq n.$$

Thus for $1 \leq k \leq p$ and $\delta \in [1-\epsilon, 1+\epsilon]$, we have,

$$0 \leq \frac{\delta - \delta_k}{\delta_k} \leq cn^{\alpha}.$$

Hence (16) becomes

$$\frac{||x - x_j||_G}{||x - x_0||_G} < c^p n^{p\alpha} \epsilon^{j-p-q}.$$

Given arbitrary tolerance $\tau > 0$, an upper bound for the number of iterations required to make

$$\frac{||x - x_j||_G}{||x - x_0||_G} < \tau$$

is therefore given by

$$j_0 \equiv p + q - \frac{p \log c + \alpha p \log n - \log \tau}{\log \epsilon} = O(\alpha \log n + 1).$$

Since by using FFTs, the matrix-vector products in Algorithm PCG for Least Squares can be done in $O(m \log n)$ operations for any n-vector v, the cost per iteration of the conjugate gradient method is of $O(m \log n)$. Thus we conclude that the work of solving (17) to a given accuracy τ is $O(\alpha m \log^2 n + m \log n)$ when $\alpha > 0$, and for sufficiently large n.

The convergence analysis given above can be further strengthened. For T an m-by-n matrix of the form (4) with $m = kn$, let $\lambda_{\min}(T_j^* T_j) = O(n^{-\alpha_j})$ for $j = 1, \cdots k$. By Lemma 1, we already know that

$$\lambda_{\min}(T_j^* T_j) \leq \lambda_{\max}(T_j^* T_j) \leq 2||f||_\infty^2,$$

therefore $\alpha_j \geq 0$. By the Cauchy interlace theorem, we see that

$$\lambda_{\min}(T^* T) \geq \sum_{j=1}^{k} \lambda_{\min}(T_j^* T_j) \geq O(n^{-\alpha}),$$

where

$$\alpha = \min_j \alpha_j \geq 0.$$

Therefore

$$\kappa(T^* T) \leq \frac{\lambda_{\max}(T^* T)}{\lambda_{\min}(T^* T)} \leq O(n^{\alpha}).$$

In the case when one of the $\alpha_j = 0$, i.e. the block T_j is well-conditioned independent of n, we see that the least squares problem is also well-conditioned, so that $\kappa(T) = O(1)$.

When at least one $\alpha_j = 0$, *i.e.*, $\kappa(T) = O(1)$, the number of iterations required for convergence is of $O(1)$. Hence the complexity of the algorithm reduces to $O(m \log n)$, for sufficiently large n. We remark that in this case, one can show further that the *method converges superlinearly* for the preconditioned least squares problem due to the clustering of the singular values for sufficiently large n (See R. Chan and Strang [9] or R. Chan [11] for details). In contrast, the method converges just linearly for the non-preconditioned case. This contrast is illustrated very well in the section on numerical tests.

5 Preconditioned Regularized Least Squares

In this section we consider solving least squares problems (3), where the rectangular matrix T is ill-conditioned. Such systems arise in many applications, such as signal and image restoration, see [3, 24, 28]. Often, the ill-conditioned nature of T results from discretization of ill-posed problems in partial differential and integral equations. Here for example, the problem of estimating an original image from a blurred and noisy observed image is an important case of an *inverse problem*, and was first studied by Hadamard [22] in the inversion of certain integral equations. Because of the ill-conditioning of T, naively solving $Tx = b$ will lead to extreme instability with respect to perturbations in b. The method of *regularization* can be used to achieve stability for these problems [7]. Stability is attained by introducing a stabilizing operator (called a regularization operator) which restricts the set of admissible solutions. Since this causes the regularized solution to be biased, a scalar (called a regularization parameter) is introduced to control the degree of bias. More specifically, the regularized solution is computed as

$$\min \left\| \begin{bmatrix} b \\ 0 \end{bmatrix} - \begin{bmatrix} T \\ \mu L \end{bmatrix} x(\mu) \right\|_2, \tag{18}$$

where μ is the regularization parameter and the $p \times n$ matrix L is the regularization operator.

The standard least squares solution to (3), given by $x = T^\dagger b$, is useless for these problems because it is dominated by rapid oscillations due to the errors. Hence in (18), one adds a term $\min \|Lx\|^2$ to (3) in order to *smooth* the solution x. Choosing L as a kth difference operator matrix forces the solution to have a small kth derivative. The regularization parameter μ controls the degree of smoothness (*i.e.*, degree of bias) of the solution, and is usually small. Choosing μ is not a trivial problem. In some cases a priori information about the signal and the degree of perturbations in b can be used to choose μ [1], or generalized cross-validation techniques may be used, *e.g.,* [7]. If no *a priori* information is known, then it may be necessary to solve (18) for several values of μ [20]. Recent analytical methods for choosing an optimal parameter μ are discussed by Reeves and Mersereau [29].

Based on the discussion above, the regularization operator L is usually chosen to be the identity matrix or some discretization of a differentiation operator [6, 20]. Thus L is typically a Toeplitz matrix. Hence, if T has the Toeplitz block form (4), then the matrix

$$\tilde{T} = \begin{bmatrix} T \\ \mu L \end{bmatrix}$$

retains this structure, with the addition of one block (or two blocks if L is a difference operator with more rows than columns). Since $\tilde{T}$ has the block structure (4), we can form the circulant preconditioner C for $\tilde{T}$ and use the PCG algorithm for least squares problems to solve (18).

Notice that if L is chosen to be the identity matrix, then the circulant preconditioner for $\tilde{T}$ can be constructed by simply adding μ to each of the eigenvalues of the circulant preconditioner for T. In addition, the last block in $\tilde{T}$ (*i.e.*, μI) has singular values μ. Thus, due to the remarks at the end of §4, if each block in T is generated by a function in $\mathcal{C}_{2\pi}$, and if $\mu \neq 0$, then $\kappa(\tilde{T}) \leq O(\mu^{-1})$ for all n. It follows then, for these problems, that (18) can be solved in $O(m \log n)$ operations, for sufficiently large n.

6 Numerical Tests

In this section we report on some numerical experiments which use the preconditioner C given by equation (5) in §1 for the conjugate gradient algorithm PCG for solving Toeplitz and block Toeplitz least squares problems. Here the preconditioner C is based on the T. Chan optimal preconditioner C_i, for each block T_i of T, as in §2. The experiments are designed to illustrate the performance of the preconditioner on a variety of problems, including some in which one or more Toeplitz blocks are very ill-conditioned.

For all numerical tests given in this section we use the stopping criteria $\|s^{(j)}\|_2/\|s^{(0)}\|_2 < 10^{-7}$, where $s^{(j)}$ is the (normal equations) residual after j iterations, and the zero vector is our initial guess. (Observe that the value $\|s^{(j)}\|_2$ is computed as part of the conjugate gradient algorithm.) All experiments were performed using the Pro-Matlab software on our workstations. The machine epsilon for Pro-Matlab on this system is approximately 2.2×10^{-16}.

To describe most of the Toeplitz matrices used in the examples below, we use the following notation. Let the m-vector c be the first column of T, and the n-vector r^T be the first row of T. Then

$$T = \mathrm{Toep}(c, r).$$

The right hand side vector b is generally chosen to be the vector of all ones.

Example 1: In this example we construct $m \times n$ Toeplitz matrices generated by a positive function in the Wiener class, varying the number of rows and columns and fixing the number of blocks in the block form (4) to $k = 3$. This example is a rectangular generalization of test data used by Strang [30], and is defined as follows. Let

$$c(i) = 1/2^{i-1}, \quad i = 1, \ldots, m \quad \text{and} \quad r(j) = 1/2^{j-1}, \quad j = 1, \ldots, n.$$

The convergence results for this example is shown in Table 1, which shows the number of iterations required for Algorithm PCG to converge using T (i.e. no preconditioner) and our C. We can see from Table 1 that the use of our preconditioner does accelerate the convergence rate of the CG algorithm for this problem. Moreover, for this example the number of iterations remains essentially constant as m and n increase.

In Figure 1 we plot the singular values of T and TC^{-1}. The plot of the singular value distributions shows that the preconditioner clusters the singular values very well for this example.

Figure 1. Singular values for T and TC^{-1} in Example 1.

Example 2: In this example we use the following three generating functions in the Wiener class to construct a $3n \times n$ block Toeplitz matrix.

(i) Example (a) from R. Chan and Yeung [15],

$$c_1(j) = r_1(j) = (|j-1|+1)^{-1.1} + \sqrt{-1}(|j-1|+1)^{-1.1}, \quad j = 1, 2, \ldots, n.$$

(ii) Example (b) from R. Chan and Yeung [15],

$$c_2(i) = (|i-1|+1)^{-1.1}, \quad i = 1, 2, \ldots, n,$$
$$r_2(j) = \sqrt{-1}(|j-1|+1)^{-1.1}, \quad j = 1, 2, \ldots, n.$$

(iii) Example (f) from R. Chan and Yeung [15],

$$c_3(1) = r_3(1) = \tfrac{1}{5}\pi^4$$
$$c_3(j) = r_3(j) = 4(-1)^{(j-1)}\left(\tfrac{\pi^2}{(j-1)^2} - \tfrac{6}{(j-1)^4}\right), \quad j = 2, 3, \ldots, n.$$

The matrix T is defined as

$$T^T = [T_1^T, \; T_2^T, \; T_3^T],$$

where T_1 =Toep(c_1, r_1), T_2 =Toep(c_2, r_2) and T_3 =Toep(c_3, r_3). For $n \times n$ systems R. Chan and Yeung [15] show that $\kappa_2(T_3) = O(n^4)$, while T_1 and T_2 are well-conditioned. They also show that T. Chan's preconditioner works well for T_1 and T_2, but not well for T_3.

In Table 1 we show the convergence results for this example, using no preconditioner and C as a preconditioner, for several values of m and n. Figure 2 shows the singular values of T and TC^{-1} for $m = 210$ and $n = 70$. These results illustrate the good convergence properties using the preconditioner C for this example containing an ill-conditioned block. Moreover, our computations verify the fact that $\kappa_2(T)$ remains almost constant as n increases from 40 to 80.

Figure 2. Singular values for T and TC^{-1} in Example 2.

Example 3: In this example we form a $2n \times n$ block Toeplitz matrix using generating functions from R. Chan and Yeung [15] which construct ill-conditioned $n \times n$ Toeplitz matrices. Here $T_1 = T_2$ and thus both blocks of T are ill-conditioned. The generating function, which is in the Wiener class, is:

Example (c) from R. Chan and Yeung [15],

$c_1(1) = r_1(1) = 0$

$c_1(j) = r_1(j) = (|j-1|+1)^{-1.1} + \sqrt{-1}(|j-1|+1)^{-1.1}, \quad j = 2, \ldots, n.$

Using the above generating functions, we let

$$T^T = [T_1,\ T_2]^T,$$

where $T_1 = T_2 =$Toep(c_1, r_1).

In Table 1 we show the convergence results for this example, using no preconditioner and C as a preconditioner, for several values of m and n. Figure 3 shows the singular values of T and TC^{-1} for $m = 140$ and $n = 70$. These results illustrate the good convergence properties of C for this example even though it contains all ill-conditioned blocks.

Figure 3. Singular values for T and TC^{-1} in Example 3.

	Example 1 ($m = 3n$)		Example 2 ($m = 3n$)		Example 3 ($m = 2n$)	
n	no prec.	with prec.	no prec.	with prec.	no prec.	with prec.
40	33	7	96	14	29	11
50	36	7	126	14	33	15
60	41	7	155	13	44	13
70	41	7	167	13	52	12
80	44	7	186	13	65	14

Table 1. Numbers of iterations for convergence in Examples 1 - 3.

Example 4: Here we consider an application to 1–dimensional image or signal reconstruction computations. In this example we construct the 100×100 Toeplitz matrix T, whose i, j entry is given by

$$t_{ij} = \begin{cases} 0 & \text{if } |i-j| > 8, \\ \frac{4}{51} g(0.15, x_i - x_j) & \text{otherwise}, \end{cases} \tag{19}$$

where

$$x_i = \frac{4i}{51}, \quad i = 1, 2, \ldots, 100,$$

and

$$g(\sigma, \gamma) = \frac{1}{2\sqrt{\pi}\sigma} \exp(-\frac{\gamma^2}{4\sigma^2}).$$

Matrices of this form occur in many image restoration contexts as a "prototype problem" and are used to model certain degradations in the recorded image [20, 24]. Due to the bandedness of T its generating function is in the Wiener class. The condition number of T is approximately 2.4×10^6.

Because of the ill–conditioning of T, the system $Tx = b$ will be very sensitive to any perturbations in b, see §5. To achieve stability we regularize the problem using the identity matrix as the regularization operator. Eldén [20] uses this approach to solve a linear system by direct methods with the same data matrix T defined in (19). To test our preconditioner we will fix $\mu = 0.01$, where μ is chosen based on some tests made by Eldén.

Let

$$\hat{T} = \begin{bmatrix} T \\ \mu I \end{bmatrix} \quad \text{and} \quad \hat{b} = \begin{bmatrix} b \\ 0 \end{bmatrix}.$$

Then $\hat{T}$ is simply a block Toeplitz matrix. Thus we can apply our preconditioner C, and the PCG algorithm, to solve (18). The convergence results for solving $Tx = b$ and $\hat{T}x = \hat{b}$ with no preconditioner and $\hat{T}x = \hat{b}$ using C as a preconditioner are shown in Table 2. The singular values of T and $\hat{T}C^{-1}$, and the convergence history for solving $Tx = b$ and $\hat{T}x = \hat{b}$ using our preconditioner C are shown in Figure 4. These results indicate that the PCG algorithm with our preconditioner C may be an effective method for solving this regularized least squares problem.

n	$Tx = b$	$\hat{T}x = \hat{b}$	$\hat{T}C^{-1}x = \hat{b}$
100	> 100	54	14

Table 2.Numbers of iterations for convergence in Example 4.

In summary, we have shown how to construct circulant preconditioners for the efficient solution of a wide class of Toeplitz least squares problems. The numerical experiments given collaborate our convergence analysis. Examples 1 and 2 both illustrate superlinear convergence for the PCG algorithm preconditioned by C, even when in Example 2 the matrix T contains an ill-conditioned block. In addition, even though the matrix T in Example 3 contains *all* ill-conditioned blocks, the scheme works well for the computations we performed.

Example 4 illustrates the applicability of the circulant PCG method to regularized least squares problems. The example comes from 1–dimensional signal restoration. 2–dimensional signal or image restoration computations often lead to very large least squares problems where the coefficient matrix is block Toeplitz with Toeplitz blocks. Block circulant preconditioners for this case are considered elsewhere [16].

In this paper we have used the T. Chan [17] optimal preconditioner for the Toeplitz blocks. Other circulant preconditioners such as ones studied by R. Chan [11], Huckle [23], Ku and Kuo [25], Strang [30], Tismenetsky [31], or Tyrtyshnikov [32], can be used, but the class of generating functions may need to be restricted for the convergence analysis to hold.

Figure 4. Singular values and convergence history for T and $\hat{T}C^{-1}$.

References

[1] J. Abbiss and P. Earwicker, *Compact Operator Equations, Regularization and Super-resolution*, in Mathematics in Signal Processing, Clarendon Press, Oxford, 1987.

[2] G. Ammar and W. Gragg, *Superfast Solution of Real Positive Definite Toeplitz Systems*, SIAM J. Matrix Appl. V9 (1988), pp. 61–76.

[3] H. Andrews and B. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

[4] O. Axelsson and V. Barker, *Finite Element Solution of Boundary Value Problems, Theory and Computation*, Academic Press, London, 1984.

[5] O. Axelsson and G. Lindskog, *On the rate of convergence of the preconditioned conjugate gradient algorithm*, Numer. Math. V48 (1986), pp. 499–523.

[6] J. Biedmond, R. Lagendijk and R. Mesereau, *Iterative Methods for Image Deblurring,* Proc. of the IEEE, V78 (1990), pp. 856–883.

[7] A. Bjorck, *Least Squares Methods*, in Handbook of Numerical Methods, ed. P. Ciarlet and J. Lions, Elsevier/North Holland Vol. 1, 1989.

[8] J. Bunch, *Stability of Methods for Solving Toeplitz Systems of Equations*, SIAM J. Sci. Statist. Comput., V6 (1985), pp. 349–364.

[9] R. Chan and G. Strang, *Toeplitz Equations by Conjugate Gradients with Circulant Preconditioner*, SIAM J. Sci. Statist. Comput., V10 (1989), pp. 104–119.

[10] R. Chan, *The Spectrum of a Family of Circulant Preconditioned Toe-plitz Systems*, SIAM J. Numer. Anal., V26 (1989), pp. 503-506.

[11] R. Chan, *Circulant Preconditioners for Hermitian Toeplitz Systems*, SIAM J. Matrix Anal. Appl., V10 (1989), pp. 542–550.

[12] R. Chan, X. Jin, *A family of Block Preconditioners for Block Systems*, SIAM J. Sci. Statist. Comput., to appear.

[13] R. Chan and M. Yeung, *Circulant Preconditioners for Toeplitz Matrices with Positive Continuous Generating Functions*, Math. Comp., to appear.

[14] R. Chan and M. Yeung, *Circulant Preconditioners Constructed from Kernels*, SIAM J. Numer. Anal., to appear.

[15] R. Chan and M. Yeung, *Circulant Preconditioners for Complex Toeplitz Matrices*, Research Report 91-6, Dept. of Math., Univ. of Hong Kong.

[16] R. Chan, J. Nagy and R. Plemmons, *Block Circulant Preconditioners for* 2*–dimensional Deconvolution Problems*, in preparation.

[17] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Statist. Comput., V9 (1988), pp. 766–771.

[18] T. Chan and J. Olkin, *Preconditioners for Toeplitz–block Matrices*, preprint, (1991).

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

[20] L. Eldén, *An Algorithm for the Regularization of Ill-conditioned, Banded Least Squares Problems*, SIAM J. Sci. Statist. Computing, V5 (1984), pp. 237-254.

[21] G. Golub and C. van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 2nd Edition, 1989.

[22] J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press, New Haven, CT, 1923.

[23] T. Huckle, *Circulant and Skew-circulant Matrices for Solving Toeplitz Matrix Problems*, in Proc. Copper Mountain Conference on Iterative Methods, Copper Mountain, Colorado, 1990.

[24] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Engelwood Cliffs, NJ 1989.

[25] T. Ku and C. Kuo, *Design and Analysis of Toeplitz Preconditioners*, IEEE Trans. Acoust. Speech Signal Process., to appear.

[26] J. Nagy, *Toeplitz Least Squares Computations,* Ph.D. thesis, North Carolina State University, directed by R. J. Plemmons.

[27] J. Nagy and R. Plemmons, *Some Fast Toeplitz Least Squares Algorithms*, Proc. SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations II, V1566, San Diego, CA, July 1991.

[28] A. Oppenheim and R. Schafer, *Discrete-time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ 1989.

[29] S. Reaves and R. Mersereau, *Optimal Regularization Parameter Estimation for Image Reconstruction*, Proc. SPIE Conference on Image Processing Algorithms and Techniques II, V1452 (1991), pp.127–137.

[30] G. Strang, *A Proposal for Toeplitz Matrix Calculations*, Stud. Appl. Math., V74 (1986), pp. 171–176.

[31] M. Tismenetsky, *A Decomposition of Toeplitz Matrices and Optimal Circulant Preconditioning,* Lin. Alg. and Applic. Special Issue on Iterative Methods, V154–156 (1991), pp. 105–121.

[32] E. Tyrtyshnikov, *Optimal and Super-optimal Circulant Preconditioners*, SIAM J. Matrix Anal. Appl., to appear.

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