The Spectrum of a Family of Circulant Preconditioned Toeplitz Systems

Raymond H. Chan*
Department of Mathematics
University of Hong Kong
Hong Kong

July 1987
Revised February 88

Abstract. We study the solutions of symmetric positive definite Toeplitz systems $Ax = b$ by the preconditioned conjugate gradient method. The preconditioner is the circulant matrix C that minimizes the Frobenius norm $||C - A||_F$, see T. Chan [5]. The convergence rate of these iterative methods is known to depend on the distribution of the eigenvalues of $C^{-1}A$. For Toeplitz matrix A with entries which are Fourier coefficients of a positive function in the Wiener class, we establish the invertiblity of C, find the asymptotic behaviour of the eigenvalues of the preconditioned matrix $C^{-1}A$ as the dimension increases and prove that they are clustered around one.

Abbreviated Title. Circulant Preconditioned Toeplitz Systems

Key words. Toeplitz matrix, circulant matrix, preconditioned conjugate gradient method

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

*This report is supported in part by the NSF Grants DCR84-05506, DCR86-02563 and DOE Grant DE-FG03-87ER25037.

1. Introduction

In this paper we discuss the solutions to a class of symmetric positive definite Toeplitz systems $Ax = b$ by the preconditioned conjugate gradient method. Direct methods that are based on the Levinson recursion formula are in constant use; see for instance, Levinson [7] and Trench [9]. For an n by n Toeplitz matrix A_n, these methods require $O(n^2)$ operations. Faster algorithms that require $O(n \log^2 n)$ operations have been developed, see Bitmead-Anderson [1] and Brent-Gustavson-Yun [2]. The stability properties of these direct methods for symmetric positive definite matrices are discussed in Bunch [3].

Strang [8] proposed using preconditioned conjugate gradient method with circulant preconditioners for solving symmetric positive definite Toe-plitz systems. The number of operations per iteration will be of order $O(n \log n)$ as circulant systems can be solved efficiently by the Fast Fourier Transform. R. Chan-Strang [4] then considered using a circulant preconditioner S_n that is obtained by copying the central diagonals of A_n and bringing them around to complete the circulant. More precisely, if $n = 2m$, and the entries a_{ij} of A_n are given by $a_{|i-j|}$ for $0 \le i, j < n$, then the entries $s_{ij} = s_{|i-j|}$ of S_n are given by

$$s_k = \begin{cases} a_k & 0 \le k \le m, \\ a_{n-k} & m \le k < n. \end{cases} \tag{1}$$

We proved in that paper that if the underlying generating function f, the Fourier coefficients of which give the entries of A_n, is a positive function in the Wiener class, then for n sufficiently large, S_n and S_n^{-1} are uniformly bounded in the l_2 norm and the eigenvalues of the preconditioned matrix $S_n^{-1}A_n$ are clustered around 1. We remark that the assumptions on f also imply that A_n are positive definite.

T. Chan [5] recently proposed another circulant matrix C_n that is obtained by averaging the corresponding diagonals of A_n with the diagonals of A_n being extended to length n by a wrap-around. More precisely, the entries $c_{ij} = c_{|i-j|}$ of C_n are given by

$$c_k = \frac{k a_{n-k} + (n-k) a_k}{n}, \quad 0 \le k < n. \tag{2}$$

He proved that such C_n minimizes the Frobenius norm $||C - A||_F$ and his experiments numerically showed that the spectrum of the preconditioned matrix $C_n^{-1}A_n$ is also clustered around one with the condition number of $C_n^{-1}A_n$ being often smaller than that of $S_n^{-1}A_n$.

In this paper, we will prove that if the generating function f is a positive function in the Wiener class, then the spectra of the preconditioners C_n and S_n are equal asymptotically. In particular, we will show that for n sufficiently large, C_n and C_n^{-1} are uniformly bounded in the l_2 norm and the eigenvalues of the preconditioned matrix $C_n^{-1}A_n$ are clustered around one. Hence, if the conjugate gradient method is applied to solve this preconditioned system, we can expect the method to have fast convergence.

2. The Spectrum of the Preconditioned Matrix $C_n^{-1}A_n$

Let us first assume that the Toeplitz matrices A_n are finite sections of a fixed singly infinite positive definite matrix A_∞, see Chan-Strang [4]. Thus the (i,j)-th entries of A_n and A_∞ are $a_{|i-j|}$. We associate to A_∞ the generating function

$$f(\theta) = \sum_{-\infty}^{\infty} a_{|k|} e^{-ik\theta},$$

defined on $[0, 2\pi)$. We will assume that f is a positive function in the Wiener class, i.e. the sequence $\{a_k\}$ is in l_1. It follows easily that A_n are symmetric positive definite matrices for all n, see for instance, Grenander-Szego [6]. Moreover, if

$$0 < f_{\min} < f < f_{\max} < \infty, \tag{3}$$

then the spectrum $\sigma(A_n)$ of A_n will lie in $[f_{\min}, f_{\max}]$.

We now show that the spectra of C_n and S_n are asymptotically the same. More precisely, we have

Lemma 1. *Let the generating function f be a positive function in the Wiener class, then*

$$\lim_{n\to\infty} \rho\left(S_n - C_n\right) = 0,$$

where $\rho(\cdot)$ denotes the spectral radius.

Proof: By (1) and (2), it is clear that $B_n \equiv S_n - C_n$ is circulant with entries

$$b_k = \begin{cases} \dfrac{k}{n}(a_k - a_{n-k}) & 0 \le k \le m, \\ \dfrac{n-k}{n}(a_{n-k} - a_k) & m \le k < n. \end{cases}$$

Here for simplicity, we are still assuming $n = 2m$. Using the fact that the j-th eigenvalue $\lambda_j(B_n)$ of B_n is given by $\sum_{k=0}^{n-1} b_k e^{2\pi ijk/n}$, we have

$$\lambda_j(B_n) = 2\sum_{k=1}^{m-1} \frac{k}{n}(a_k - a_{n-k})\cos(2\pi jk/n).$$

This implies

$$\rho(B_n) \le 2\sum_{k=1}^{m-1} \frac{k}{n}|a_k| \; + \; 2\sum_{k=m+1}^{n-1} |a_k|.$$

Since f is in the Wiener class, hence for all $\epsilon > 0$, we can always find an $M_1 > 0$ and an $M_2 > M_1$, such that

$$\sum_{k=M_1+1}^{\infty} |a_k| < \epsilon/6 \quad \text{and} \quad \frac{1}{M_2}\sum_{k=1}^{M_1} k|a_k| < \epsilon/6.$$

Thus for all $m > M_2$,

$$\rho(B_n) < \frac{2}{M_2}\sum_{k=1}^{M_1} k|a_k| + 2\sum_{k=M_1+1}^{m-1} |a_k| + 2\sum_{k=m+1}^{\infty} |a_k| < \epsilon. \quad \square$$

We remark that if f is positive and is in the Wiener class, then for n sufficiently large, S_n and S_n^{-1} are uniformly bounded in the l_2 norm, see R. Chan-Strang [4, Theorem 1]. Moreover, if (3) holds, then the spectrum $\sigma(S_n)$ lies in $[f_{\min}, f_{\max}]$ too. Using Lemma 1, we thus have,

Theorem 1. *Let f be a positive function in the Wiener class, then for all n sufficiently large, the circulant matrices C_n and C_n^{-1} are uniformly bounded in the l_2 norm. Moreover, $\sigma(C_n)$ lies in $[f_{\min}, f_{\max}]$.*

To prove that the spectrum of $C_n^{-1}A_n$ is clustered around 1, we first recall that the spectrum of $A_n - S_n$ is clustered around zero:

Lemma 2 [4, Theorem 4]. *Let f be a positive function in the Wiener class, then for all $\epsilon > 0$, there exist $N, M > 0$, such that for all $n > N$, at most M eigenvalues of $A_n - S_n$ have absolute value larger than ϵ.*

Notice that since

$$C_n^{-1}A_n = I_n + C_n^{-1}(A_n - S_n) + C_n^{-1}(S_n - C_n),$$

we have

Theorem 2. *Let f be a positive function in the Wiener class, then for all $\epsilon > 0$, there exist $N, M > 0$, such that for all $n > N$, at most M eigenvalues of $C_n^{-1}A_n - I_n$ have absolute value larger than ϵ.*

Thus the spectrum of $C_n^{-1}A_n$ is clustered around 1 for sufficiently large n. This is consistent with the numerical results obtained in T. Chan [5]. We note that since the spectra of $C_n^{-1}A_n$ and $S_n^{-1}A_n$ are equal asymptotically, we expect the convergence rates of the conjugate gradient method applied to $S_n^{-1}A_n$ and $C_n^{-1}A_n$ to be roughly the same for n sufficiently large. In particular, both will converge superlinearly, at least in exact arithmetic, see R. Chan-Strang [4] and the numerical results below.

3. Numerical Results and Concluding Remarks

For f in the Wiener class, the numerical results in T. Chan [5] show that the spectrum of $S_n^{-1}A_n$ is more clustered than that of $C_n^{-1}A_n$. This phenomenon is more pronounced when a_k decreases more rapidly with k. However, he also observed that in these cases, $C_n^{-1}A_n$ has a smaller condition number than $S_n^{-1}A_n$.

To test the convergence rates of both preconditioners, we apply the preconditioned conjugate gradient method on $A_n x = b$ with $a_k = (1+k)^{-1.1}$. We note that the generating function of A_n is in the Wiener class. The spectra of A_n, $S_n^{-1}A_n$ and $C_n^{-1}A_n$ for $n = 32$ are given in Figure 1. Table 1 shows the number of iterations required to make the l_2 norm of the residual vector

$< 10^{-7}$. The right hand side b is the vector of all ones and the zero vector is our initial guess. We see that as n increases, the number of iterations increases for the original matrix A_n, while it stays almost the same for the preconditioned matrices. Moreover, both preconditioned systems converge at the same rate for large n.

n	A_n	$S_n^{-1}A_n$	$C_n^{-1}A_n$
8	4	4	4
16	8	5	4
32	11	5	5
64	14	5	5

Table 1. Number of Iterations for Different Systems

We finally emphasize that since C_n is defined in terms of averaging the diagonals of A_n, it can be used for general non-Toeplitz matrix A_n. Thus if A_n is nearly Toeplitz, say a low rank perturbation of a Toeplitz matrix, then we still expect C_n to be a good preconditioner for A_n.

Acknowledgement

The author acknowledges helpful discussions with Prof. Tony Chan of UCLA and Prof. Zhengfang Zhou of MIT, and the hospitality of Prof. Gene Golub during my visit at Stanford University in the summer of 1987.

References

1. Bitmead, R. and Anderson, B., *Asymptotically Fast Solution of Toeplitz and Related Systems of Equations*, Lin. Alg. Appl., V34 (1980), pp. 103-116.

2. Brent, R., Gustavson, F. and Yun, D., *Fast Solution of Toeplitz Systems of Equations and Computations of Pade Approximations*, J. Algorithms, V1 (1980), pp. 259-295.

3. Bunch, J., *Stability of Methods for Solving Toeplitz Systems of Equations*, SIAM J. Sci. Stat. Comp., V6 (1985), pp. 349-364.

4. Chan, R. and Strang, G., *The Asymptotic Toeplitz-Circulant Eigenvalue Problem*, MIT App. Math. Dept., Numer. Anal. Report 87-5, May 1987.

5. Chan, T., *An Optimal Circulant Preconditioner for Toeplitz Systems*, UCLA Math. Dept., CAM Report 87-06, June 1987.

6. Grenander, U., and Szego, G., *Toeplitz Forms and Their Applications*, 2nd Ed., Chelsea Pub. Co., New York, 1984.

7. Levinson, N., *The Wiener rms (root-mean-square) Error Criterion in Filter Design and Prediction*, J. Math. Phys., V25 (1947), pp. 261-278.

8. Strang, G., *A Proposal for Toeplitz Matrix Calculations*, Studies in App. Math., V74 (1986), pp. 171-176.

9. Trench, W., *An Algorithm for the Inversion of Finite Toeplitz Matrices*, SIAM J. Appl. Math., V12 (1964), pp. 515-522.