Best Conditioned Circulant Preconditioners

Raymond H. Chan * and C.K. Wong
Department of Mathematics
University of Hong Kong
Pokfulam Road
Hong Kong

November 1992

Abstract

In this paper, we discuss the solutions to a class of Hermitian positive definite system $Ax = b$ by the preconditioned conjugate gradient method with circulant preconditioner C. In general, the smaller the condition number $\kappa(C^{-1/2}AC^{-1/2})$ is, the faster the convergence of the method will be. The circulant matrix C_b that minimizes $\kappa(C^{-1/2}AC^{-1/2})$ is called the best conditioned circulant preconditioner for the matrix A. We prove that if FAF^* has Property A where F is the Fourier matrix, then C_b minimizes $||C - A||_F$ over all circulant matrices C. Here $|| \cdot ||_F$ denotes the Frobenius norm. We also show that there exists non-circulant Toeplitz matrix A such that FAF^* has Property A.

Key Words. Toeplitz matrix, Circulant matrix, best conditioned circulant preconditioner, preconditioned conjugate gradient method

AMS(MOS) Subject Classifications. 65F10, 65F15, 65F35

---

*Research supported in part by HKRGC grant no. 221600070, ONR contract no. N00014-90-J-1695 and DOE grant no. DE-FG03-87ER25037.

1 Introduction

In this paper, we discuss the solutions to a class of Hermitian positive definite system $Ax = b$ by the preconditioned conjugate gradient method (PCG). The rate of convergence of the conjugate gradient method (CG) depends on the condition number $\kappa(A)$, see [1]. In general, the smaller $\kappa(A)$ is, the faster the convergence will be. In case $\kappa(A)$ is not small, the method is always used with a Hermitian positive definite matrix M to speed up the convergence rate of the method. More precisely, instead of applying the CG method to the system $Ax = b$, we apply the method to the transformed system $\tilde{A}\tilde{x} = \tilde{b}$ where $\tilde{A} = M^{-1/2}AM^{-1/2}$, $\tilde{x} = M^{1/2}x$ and $\tilde{b} = M^{-1/2}b$. The matrix M is called a preconditioner for A. The preconditioner M is chosen so as to minimize $\kappa(M^{-1/2}AM^{-1/2})$ and to allow efficient computation of the product $M^{-1/2}v$ for any given vector v. The preconditioner M for A can also be viewed as an approximation to A that is easily invertible.

A matrix B is said to have Property A if there exists a permutation matrix P such that

$$PBP^T = \begin{pmatrix} D_1 & H \\ K & D_2 \end{pmatrix},$$

where D_1 and D_2 are square diagonal matrices and H and K are arbitrary matrices.

An n-by-n matrix $A = [a_{i,j}]$ is said to be Toeplitz if $a_{i,j}=a_{i-j}$, i.e., A 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 $0 < j \leq n-1$. In this paper, any circulant matrix $C = [c_{k-j}]_{j,k=1}^{n}$ is denoted by $circ(c_0, c_1, ..., c_{n-1})$. We remark that all circulant matrices C can be diagonalized as

$$C = F^* \Lambda F. \tag{1}$$

where $F = \frac{1}{\sqrt{n}}[e^{\frac{2\pi jki}{n}}]_{j,k=0}^{n-1}$ is the Fourier matrix. Hence, for any vector v, the matrix-vector multiplication $C^{-1/2}v = F^*\Lambda^{-1/2}Fv$ can be computed efficiently by the Fast Fourier Transform (FFT) in $O(n \log n)$ operations. Since circulant matrices are Toeplitz matrices themselves, it is natural to consider using circulant matrices as preconditioners for Toeplitz systems, see Strang [11] and Olkin [10].

A circulant matrix C_b is said to be the best conditioned circulant preconditioner for a matrix A if

$$\kappa(C_b^{-1/2}AC_b^{-1/2}) \leq \kappa(C^{-1/2}AC^{-1/2})$$

for any circulant matrix C. In this paper, we prove that if FAF^* has Property A, then C_b is the minimizer of $||C - A||_F$ over all circulant matrices C. Clearly, by (1) if A is circulant then FAF^* has Property A. We will show that there exists non-circulant Toeplitz matrix A such that FAF^* has Property A.

2 Background of Circulant Preconditioners

For any n-by-n Hermitian positive definite Toeplitz matrix $A_n = [a_{j-k}]_{j,k=1}^n$, there are many possible circulant matrices C that one can define to be the preconditioners for the system $A_n x = b$. The most natural choice is the Strang preconditioner which is defined to be the circulant matrix C_S that copies the central diagonals of A_n and then wraps around to form the circulant, see [11]. In particular, $C_S = circ(a_0, ..., a_m, a_{m-1}, ..., a_1)$ when $n = 2m$.

If the diagonals a_j of the Toeplitz matrix A_n are Fourier coefficients of a positive function in the Wiener class (i.e. $\sum_{i=1}^{\infty} |a_i| < \infty$), Chan [2] proved that the eigenvalues of the preconditioned matrix $C_S^{-1/2} A_n C_S^{-1/2}$ will be clustered around one. More precisely, for all $\epsilon > 0$, there exist N_1, N_2, such that, for all $n > N_1$, at most N_2 eigenvalues of $C_S^{-1/2} A_n C_S^{-1/2} - I$ have absolute value larger than ϵ. Hence, the PCG method has a superlinear convergence rate for large n. Specifically, we have for all $\epsilon > 0$, there exists a constant $c(\epsilon) > 0$ such that the error vector e_q of the preconditioned conjugate gradient method at the qth iteration satisfies $||e_q|| \leq c(\epsilon)\epsilon^q ||e_0||$ when n is sufficiently large. Here $||x||^2 = x^* C_S^{-1/2} A_n C_S^{-1/2} x$. Hence, the number of iterations required for convergence is independent of the size of the matrix A_n when n is large.

Since the convergence rate of the PCG method depends on how small the condition number $\kappa(C^{-1/2} A C^{-1/2})$ is, and it is not easy to find the circulant preconditioner C that minimizes such condition number, much attention has been focused on searching a circulant matrix C which is close to the matrix A in certain norms. For any n-by-n Toeplitz matrix $A_n = [a_{j-k}]_{j,k=1}^n$, T. Chan in [7] proposed a circulant preconditioner $C_F = circ(c_0, c_1, ..., c_{n-1})$ which is the minimizer of $||C - A_n||_F$ over all circulant matrices C. Here $||\cdot||_F$ denotes the Frobenius norm. He showed that the entries c_j of C_F are given by

$$c_j = \frac{j a_{-(n-j)} + (n-j) a_j}{n}, \qquad j = 0, 1, ..., n-1.$$

It was then shown in Chan [3] that the spectrum of $C_F^{-1/2} A_n C_F^{-1/2}$ is also clustered around one if the underlying generating function of A_n is a positive function in the Wiener class. Tyrtyshnikov in [12] extended the definition of C_F to any general n-by-n matrix A. Also, he proved that C_F is symmetric positive definite whenever A is. Note that forming C_F only needs $O(n)$ operations for Toeplitz matrix A of order n, and $O(n^2)$ operations for general n-by-n matrix A. We note that instead of minimizing in the Frobenius norm as the T. Chan preconditioner does, Strang's preconditioner C_S actually minimizes $||C - A_n||_1$ and $||C - A_n||_\infty$ over all Hermitian circulant matrices C, see Chan [2].

Since the matrices $C^{-1}A$ and $C^{-1/2} A C^{-1/2}$ are similar, the PCG method converges superlinearly when the eigenvalues of $C^{-1}A$ are clustered around one. Tyrtyshnikov in [12] therefore proposed using the circulant preconditioner C_T that minimizes $||I - C^{-1}A||_F$

over all nonsingular circulant matrices C. Fast algorithms for finding C_T take $O(n^2 \log n)$ operations for general n-by-n matrix A and only $O(n \log n)$ operations when A is Toeplitz. He proved that the circulant matrix C_T is symmetric positive definite whenever A is. The analysis of such preconditioner for Toeplitz matrix A with positive generating function in the Wiener class was given by Chan, Jin and Yeung in [4]. They proved that the eigenvalues of the preconditioned system are clustered around one. The circulant matrix C_T is called the superoptimal circulant preconditioner for A and C_F is called the optimal circulant preconditioner for A.

Since the iterative matrix used in the PCG method is $C^{-1/2}AC^{-1/2}$, Huckle in [9] proposed the circulant preconditioner C_H which is defined to be the minimizer of $||I - C^{-1/2}AC^{-1/2}||_F$ over all positive definite circulant matrices C. He found that for any Toeplitz matrix A of order n, forming C_H requires a solution of another n-by-n linear system which can be solved iteratively in $O(n \log n)$ operations. He then proved that if the underlying function of A is positive and in the Wiener class, then the eigenvalues of the preconditioned system are clustered around one.

3 Best Conditioned Preconditioner

However, the above preconditioners are not the "true" optimal approximation. It is because the convergence rate of the PCG method depends on the condition number of the matrix $C^{-1/2}AC^{-1/2}$. Therefore, one should minimizes $\kappa(C^{-1/2}AC^{-1/2})$ instead of minimizing the norms of matrices. The aim of this paper is to find the circulant matrix which minimizes such condition number. The minimizer is called the best conditioned circulant preconditioner for A. We find that for matrices A such that the product FAF^* has Property A, the best conditioned circulant matrix is just the optimal circulant preconditioner C_F. The following two Lemmas relate matrices with Property A to the best conditioned circulant matrix.

Lemma 1 (Forsythe and Straus [8]) *Let Q be a Hermitian positive definite matrix of the form*

$$Q = \begin{pmatrix} I_p & B^* \\ B & I_q \end{pmatrix}$$

where I_p and I_q are identity matrices of order p and q respectively. Then $\kappa(\Lambda Q \Lambda) \geq \kappa(Q)$ for any diagonal matrix Λ.

Lemma 2 (Chan, Jin and Yeung [5]) *Let A be an n-by-n matrix and C_F be the optimal circulant preconditioner for A. Then*

$$C_F = F^* \delta(FAF^*) F,$$

where F is the Fourier matrix and $\delta(B)$ is the diagonal matrix such that $\delta(B)_{i,i} = (B)_{i,i}, 1 \leq i \leq n$.

With the help of the Lemmas, we prove the main result.

Theorem 1 *Let A be an n-by-n Hermitian positive definite matrix. If the matrix FAF^* has Property A, then C_F minimizes $\kappa(C^{-1/2}AC^{-1/2})$ over all Hermitian positive definite circulant matrices C.*

Proof: For any circulant matrix C, by (1), we have $C = F^*\Lambda F$ where Λ is a diagonal matrix. Hence

$$\kappa(C^{-1/2}AC^{-1/2}) = \kappa(F^*\Lambda^{-1/2}FAF^*\Lambda^{-1/2}F) = \kappa(\Lambda^{-1/2}FAF^*\Lambda^{-1/2}).$$

Let $\Lambda_F = \delta(FAF^*)$, the diagonal matrix with diagonal entries the same as the diagonal entries of FAF^*. Since FAF^* has Property A, the matrix $\Lambda_F^{-1/2}FAF^*\Lambda_F^{-1/2}$ also has Property A and has the same structure as the matrix Q in Lemma 1. Therefore, by Lemma 1, the diagonal matrix Λ_F minimizes $\kappa(\Lambda^{-1/2}FAF^*\Lambda^{-1/2})$ over all positive definite diagonal matrices Λ. By Lemma 2, $C_F = F^*\Lambda_F F$. Hence, C_F minimizes $\kappa(C^{-1/2}AC^{-1/2})$ over all Hermitian positive definite circulant matrices C. $\square$

Obviously, FAF^* has Property A for any circulant matrix A. Let us show that there exist non-circulant Hermitian Toeplitz matrices which have Property A. Let

$$A_1 = \begin{pmatrix} \alpha & \beta & \gamma \\ \bar{\beta} & \alpha & \beta \\ \bar{\gamma} & \bar{\beta} & \alpha \end{pmatrix}$$

where α is real number, β and γ are complex numbers with $\beta + \bar{\beta} = \gamma + \bar{\gamma}$; and

$$A_2 = \begin{pmatrix} a & b & c & d \\ \bar{b} & a & b & c \\ \bar{c} & \bar{b} & a & b \\ \bar{d} & \bar{c} & \bar{b} & a \end{pmatrix}$$

where a and c are real numbers, b and d are complex numbers with $b - \bar{d} = \sqrt{-1}(\bar{b} - d)$. It is not difficult to show that FA_1F^* and FA_2F^* have Property A. Finally, we emphasis that C_F is not the best conditioned circulant matrix for general Hermitian positive definite matrix A. For example, let us consider

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

and

$$C = circ(48/25, -4/5, 0, 0, -4/5).$$

Then we have

$$\kappa(C_F^{-1/2} A C_F^{-1/2}) \approx 3.4114 > 3.3156 \approx \kappa(C^{-1/2} A C^{-1/2}).$$

4 Concluding Remarks

For an n-by-n Hermitian matrix A such that FAF^* has Property A, we have proved that the optimal circulant preconditioner C_F is the best conditioned circulant preconditioner for A. Chan and Yeung [6] have proved that if A is a Toeplitz matrix with positive 2π-periodic continuous generating function, then the PCG method with C_F as preconditioner has a superlinear convergence rate. We refer the readers there for numerical performances of such preconditioner.

Acknowledgment: The authors wish to thank Professor Olof Widlund for his help in this work.

References

[1] O. Axelsson and V. Barker, *Finite Element Solution of Boundary Value Problems: Theory and Computation,* Academic Press, Orlando, Florada (1983).

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

[3] R. Chan, *The Spectrum of a Family of Circulant Preconditioned Toeplitz Systems,* SIAM J. Numer. Anal., Vol. 26 (1989), pp. 503-506

[4] R. Chan, X. Jin and M. Yeung, *The Spectra of Super-optimal Circulant Preconditioned Toeplitz Systems,* SIAM J. Numer. Anal., Vol. 28 (1991), pp. 871-879.

[5] R. Chan, X. Jin and M. Yeung, *The Circulant Operator in the Banach Algebra of Matrices,* Linear Algebra Appl., Vol. 149 (1991), pp. 41-53.

[6] R. Chan and M. Yeung, *Circulant Preconditioners for Toeplitz Matrices with Positive Continuous Generating Functions,* Math. Comp., Vol. 58 (1992), pp. 233-240.

[7] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems,* SIAM J. Sci. Statist. Comput., Vol. 9 (1988), pp. 766-771.

[8] G. Forsythe and E. Straus, *On Best Conditioned Matrices,* Proc. Amer. Math. Soc. Vol. 6 (1955), pp. 340-345.

[9] T. Huckle, *Some Aspects of Circulant Preconditioners*, to appear.

[10] J. Olkin, *Linear and Nonlinear Deconvolution Problems,* Ph.D. thesis, Rice University, Houston, Texas (1986)

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

[12] E. Tyrtyshnikov, *Optimal and Super-optimal Circulant Preconditioners,* SIAM J. Matrix Anal. Appl., Vol. 13 (1992), pp. 459-473.