Series in Information and Computational Science — 37

Developments and Applications of Block Toeplitz Iterative Solvers

Xiao-qing Jin

(块特普利茨迭代方法)

SCIENCE PRESS
SCIENCE PRESS USA Inc.

0241.6
Y5

Developments and Applications of Block Toeplitz Iterative Solvers

(块特普利茨迭代方法)

By

Xiao-Qing JIN

Faculty of Science and Technology,
University of Macau,
Macau, China.

SCIENCE PRESS
SCIENCE PRESS USA Inc.

Responsible Editors: Liu Jiashan   Fan Qingkui

Copyright©2002 by Science Press. Second Printing 2006.
Published by Science Press
16 Donghuangchenggen North Street
Beijing 100717, China

Printed in Beijing

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the copyright owner.

ISBN 7-03-010719-5 (Beijing)

Preface to the Series in Information and Computational Science

Since the 1970s, Science Press has published more than thirty volumes in its series Monographs in Computational Methods. This series was established and led by the late academician, Feng Kang, the founding director of the Computing Center of the Chinese Academy of Sciences. The monograph series has provided timely information of the frontier directions and latest research results in computational mathematics. It has had great impact on young scientists and the entire research community, and has played a very important role in the development of computational mathematics in China.

To cope with these new scientific developments, the Ministry of Education of the People's Republic of China in 1998 combined several subjects, such as computational mathematics, numerical algorithms, information science, and operations research and optimal control, into a new discipline called Information and Computational Science. As a result, Science Press also reorganized the editorial board of the monograph series and changed its name to Series in Information and Computational Science. The first editorial board meeting was held in Beijing in September 2004, and it discussed the new objectives, and the directions and contents of the new monograph series.

The aim of the new series is to present the state of the art in Information and Computational Science to senior undergraduate and graduate students, as well as to scientists working in these fields. Hence, the series will provide concrete and systematic expositions of the advances in information and computational science, encompassing also related interdisciplinary developments.

I would like to thank the previous editorial board members and assistants, and all the mathematicians who have contributed significantly to the monograph series on Computational Methods. As a result of their contributions the monograph series achieved an outstanding reputation in the community. I sincerely wish that we will extend this support to the new Series in Information and Computational Science, so that the new series can equally enhance the scientific development in information and computational science in this century.

Shi Zhongci

2005.7

Preface

In this book we introduce current developments and applications in using iterative methods for solving block Toeplitz systems. The block Toeplitz systems arise in a variety of applications in mathematics, scientific computing and engineering, for instance, image restoration problems in image processing; numerical differential equations and integral equations; time series analysis and control theory. Krylov subspace methods and multigrid methods are proposed. One of the main results of these iterative methods is that the operation cost of solving a large class of $mn \times mn$ block Toeplitz systems is only required $O(mn \log mn)$ operations.

This book consists of twelve chapters. Various bibliographies are placed at the end of thc book. In Chapter 1, we survey some background knowledge of matrix analysis and point Toeplitz iterative solvers that will be used later to develop our block Toeplitz iterative solvers.

In Chapter 2 we study block circulant preconditioners for the solution of block system $T_{mn}u = b$ by the preconditioned conjugate gradient (PCG) method where T_{mn} is an $m \times m$ block Toeplitz matrix with $n \times n$ Toeplitz blocks. The preconditioners $c_F^{(1)}(T_{mn})$, $\tilde{c}_F^{(1)}(T_{mn})$ and $c_{F,F}^{(2)}(T_{mn})$ are the matrices that preserve the block structure of T_{mn}. Specifically, they are defined to bc the minimizers of $\|T_{mn} - C_{mn}\|_F$ with C_{mn} over some special classes of matrices. We prove that if T_{mn} is positive definite, then $c_F^{(1)}(T_{mn})$, $\tilde{c}_F^{(1)}(T_{mn})$ and $c_{F,F}^{(2)}(T_{mn})$ are positive definite too. We also show that they are good preconditioners for solving some special block Toeplitz systems. Finally, we briefly discuss two other preconditioners $s_{F,F}^{(2)}(T_{mn})$ and $r_{F,F}^{(2)}(T_{mn})$. The invertibility of the preconditioners $s_{F,F}^{(2)}(T_{mn})$ and $r_{F,F}^{(2)}(T_{mn})$ is studied.

In Chapter 3 block circulant preconditioners for block Toeplitz systems are studied from the viewpoint of kernels. We show that most of the well known block circulant preconditioners can be derived from convoluting the generating functions of systems with some famous kernels. The convergence analysis is also given.

In Chapter 4 we study the solutions of a block Toeplitz systems $Tu = b$ by the PCG method where $T = T_{(1)} \otimes T_{(2)} \otimes \cdots \otimes T_{(m)}$ with Toeplitz blocks

此为试读,需要完整PDF请访问: www.ertongbook.com

$T_{(i)} \in \mathbb{R}^{n\times n}$, $i = 1, 2, \cdots, m$. Two preconditioners C and P are proposed. The preconditioner C is a matrix that preserves the tensor structure of T and is close to T in Frobenius norm over a special class of matrices. The preconditioner P is defined for ill conditioned problems. With a fast algorithm, we show that both C and P are good preconditioners for solving block Toeplitz systems with tensor structure. Only $O(mn^m \log n)$ operations are required for the solutions of preconditioned systems. The inverse heat problem is also discussed.

In Chapter 5 we study the constrained and weighted least squares problem

$$\min_x \frac{1}{2}(b - Tx)^T W(b - Tx)$$

where $W = \text{diag}(\omega_1, \cdots, \omega_m)$ with $\omega_1 \geq \cdots \geq \omega_m \geq 0$ and $T^T = \left(T_{(1)}^T, \cdots, T_{(k)}^T\right)$ with Toeplitz blocks $T_{(l)} \in \mathbb{R}^{n\times n}$, $l = 1, \cdots, k$. It is well known that this problem can be solved by solving the following linear system

$$\begin{cases} M\lambda + Tx = b, \\ \\ T^T\lambda = 0, \end{cases}$$

where $M = W^{-1}$. We use the PCG method with circulant-like preconditioner for solving the system and we obtain a fast convergence rate.

In Chapter 6 we study the solutions of ill conditioned block Toeplitz systems $T_{mn}u = b$ where T_{mn} are generated by a function $f(x, y) \geq 0$. Two important theorems [77], which give the relations between the values of $f(x, y)$ and the eigenvalues of T_{mn}, are proposed. Usually, the convergence rate of the conjugate gradient method for solving ill conditioned block Toeplitz systems is slow. To deal with such kind of problem, a block $\{\omega\}$-circulant preconditioner is proposed. We show that the block $\{\omega\}$-circulant preconditioner can work efficiently for ill conditioned block Toeplitz systems. A numerical comparison between the block $\{\omega\}$-circulant preconditioner and the preconditioner $c_{F,F}^{(2)}(T_{mn})$ is also given.

In Chapter 7 we first study block band Toeplitz preconditioners for the solutions of ill conditioned block Toeplitz systems $T_{mn}u = b$ by the PCG method. Here T_{mn} are assumed to be generated by a function $f(x, y) \geq 0$.

The generating function $g(x, y)$ of the block band Toeplitz preconditioners is a trigonometric polynomial of fixed degree and is determined by minimizing $|||(f - g)/f|||_\infty$. Remez algorithm is proposed to construct the preconditioners. We prove that the condition number of the preconditioned system is $O(1)$. A priori bound on the number of iterations for convergence is obtained. Finally, we briefly discuss the preconditioners based on some well known fast transforms.

In Chapter 8 we study the solutions of ill conditioned block Toeplitz systems $T_{mn}u = b$ by multigrid methods (MGMs). For a class of block Toeplitz matrices, we show that the convergence factor of the two-grid method is uniformly bounded below 1 and independent of m and n, and the full MGM has a convergence factor depending only on the number of levels. The cost per iteration for the MGM is of $O(mn \log mn)$ operations. Numerical results are given to explain the convergence rate.

In Chapter 9 we first review some results related to numerical solutions of elliptic boundary value problems. We then consider linear systems arising from implicit time discretizations and finite difference space discretizations of second-order hyperbolic equations in 2-dimensional space. We propose and analyse the use of block circulant preconditioners for the solutions of linear systems by the PCG method. For second-order hyperbolic equations with given initial and Dirichlet boundary conditions, we prove that the condition number of the preconditioned system is of $O(\alpha)$ or $O(m)$, where α is the grid ratio between the time and space steps and m is the number of interior grid points in each direction. The results are extended to parabolic equations. Numerical experiments also indicate that the preconditioned systems exhibit favorable clustering of eigenvalues that leads to a fast convergence rate. Block preconditioners based on the fast sine transform are discussed for discretized systems of second-order partial differential equations in 3-dimensional space.

In Chapter 10 block preconditioners based on the fast sine transform are proposed for solving non-symmetric and non-diagonally dominant linear systems that arise from discretizations of first-order partial differential equations. We prove that if the generalized minimal residual (GMRES) method is applied to solving the preconditioned systems, the asymptotic convergence factor of the method is independent of the mesh size and depends only on the grid ratio between the time and space steps. We compare the convergence rate of our preconditioned system with the one that preconditioned by the semi-Toeplitz preconditioner. We show that our preconditioned systems have a smaller asymptotic convergence factor and numerical experiments indicate that our preconditioned systems have a much faster convergence rate.

In Chapter 11 the block circulant preconditioner $\tilde{s}_F^{(1)}(M)$ is proposed for solving linear systems arising from numerical methods for ordinary differential equations (ODEs). We use linear multistep methods to discretize ODEs. These implicit numerical methods for solving ODEs require the solutions of non-symmetric, large and sparse linear systems at each integration step. Hence, the GMRES method is used. We show that when some stable boundary value methods are used to discretize ODEs, the preconditioner $\tilde{s}_F^{(1)}(M)$ is invertible and the eigenvalues of the preconditioned system are clustered around 1. When the GMRES method is applied to solving these preconditioned systems, we have a fast convergence rate. Numerical results are given to illustrate the effectiveness of the method. An algorithm for solving differential algebraic equations is also given.

In Chapter 12 we briefly study image restoration problems in image processing. The image of an object can be modeled as

$$g(\xi,\delta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t(\xi,\delta;\alpha,\beta)f(\alpha,\beta)d\alpha d\beta + \eta(\xi,\delta)$$

where $g(\xi,\delta)$ is the degraded image, $f(\alpha,\beta)$ is the orginal image, $\eta(\xi,\delta)$ represents an additive noise. The image restoration problem is that given the observed image g, compute an approximation to the original image f. The regularized PCG least squares method with the preconditioners based on some fast transforms is proposed for solving linear systems arising from image restoration problems.

This book contains main parts of my research work in the past twelve years. Some research results are joint work with Professor Raymond H.F. Chan of the Department of Mathematics, Chinese University of Hong Kong; Professor Q.S. Chang of the Institute of Applied Mathematics, Chinese Academy of Sciences; Dr. Michael K.P. Ng of the Department of Mathematics, University of Hong Kong; Dr. H.W. Sun of the Department of Mathematics and Physics, Guangdong University of Technology; and my students Miss K.I. Kou and Mr. S.L. Lei of the Faculty of Science and Technology, University of Macau. I wish to express my sincere gratitude to my former Ph.D supervisor, Professor Raymond H.F. Chan, for leading me to this interesting area of fast iterative Toeplitz solvers and for his continual guidance, constant encouragement, long standing friendship, financial support and help. I am indebted to Professor Tony F. Chan of the Department of Mathematics, University of California, Los Angeles, for his enlightening suggestions and comments, from which I benefited a great deal during my Ph.D studies. I would like to thank my friends Professor Z.C. Shi, Dr. C.M. Cheng, Dr. C.K. Wong and Dr. M.C. Yeung for their many helpful discussions and suggestions. Thanks are also due to my parents for their encouraging and financial support. Finally, I would like to express my appreciation to my dear wife, Kathy, who eased many burdens and provided an environment and the encouragement essential to the completion of this book.

The publication of the book is supported in part by the research grants No. RG009/98-99S/JXQ/FST, No. RG010/99-00S/JXQ/FST and No. RG026/00-01S/JXQ/FST from University of Macau.

CONTENTS

此为试读，需要完整PDF请访问：www.ertongbook.com

Chapter 1

Introduction

In this chapter we first introduce some background knowledge of matrix analysis which will be used throughout the book. We then give a brief survey of current developments in using preconditioned conjugate gradient (PCG) methods for solving Toeplitz systems in the point case.

1.1 Background

In this section an overview of the relevant concepts in matrix analysis is given. The material contained here will be helpful in developing our theory in later chapters.

1.1.1 Symmetric matrix, norms and tensor

A matrix $A \in \mathbb{R}^{n\times n}$ is said to be a symmetric matrix if $A^T = A$ where 'T' denotes the transposition. Real symmetric matrices have many elegant and important properties, see [95, 100], and here we present only several classical results that will be used later.

Theorem 1.1 (Spectral Theorem) *Let $A \in \mathbb{R}^{n\times n}$ be given. Then A is symmetric if and only if there exist an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\Lambda \in \mathbb{R}^{n\times n}$ such that $A = Q\Lambda Q^T$.*

We recall that a matrix $M \in \mathbb{R}^{n\times n}$ is called orthogonal if $M^{-1} = M^T$.

Theorem 1.2 (Cauchy's Interlace Theorem) *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix with eigenvalues*

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$$

and let

$$\mu_1 \le \mu_2 \le \cdots \le \mu_{n-1}$$

be the eigenvalues of a principal submatrix of A of order $n-1$. Then

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \mu_2 \le \cdots \le \mu_{n-1} \le \lambda_n.$$

Theorem 1.3 (Weyl's Theorem) *Let A, $E \in \mathbb{R}^{n\times n}$ be symmetric matrices and let the eigenvalues $\lambda_i(A)$, $\lambda_i(E)$ and $\lambda_i(A+E)$ be arranged in increasing order. Then for each $k = 1, 2, \cdots, n$, we have*

$$\lambda_k(A) + \lambda_1(E) \le \lambda_k(A+E) \le \lambda_k(A) + \lambda_n(E).$$

Theorem 1.4 (Courant–Fischer's Minimax Theorem) *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix with eigenvalues*

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$$

and let k be a given integer with $1 \le k \le n$. Then

$$\lambda_k = \min_{\dim\mathcal{X}=k}\ \max_{0\ne x\in\mathcal{X}} \frac{x^TAx}{x^Tx} = \min_{\dim\mathcal{X}=n-k+1}\ \max_{0\ne x\in\mathcal{X}} \frac{x^TAx}{x^Tx}.$$

In particular, for the smallest and largest eigenvalues, we have

$$\lambda_1 = \min_{x\ne 0} \frac{x^TAx}{x^Tx} \quad \textit{and} \quad \lambda_n = \max_{x\ne 0} \frac{x^TAx}{x^Tx}.$$

The results of Spectral Theorem, Cauchy's Interlace Theorem, Weyl's Theorem and Courant–Fischer's Minimax Theorem can be extended to the case of Hermitian matrices. We remark that a matrix $A \in \mathbb{C}^{n\times n}$ is said to be a Hermitian matrix if $A^* = A$ where '$*$' denotes the conjugate transposition. For any arbitrary $A \in \mathbb{C}^{n\times n}$, it is possible to decompose A into an 'almost diagonal form' – the Jordan canonical form.

Theorem 1.5 (Jordan's Decomposition Theorem) *If $A \in \mathbb{C}^{n\times n}$, then there exists an invertible matrix $X \in \mathbb{C}^{n\times n}$ such that*

$$X^{-1}AX = J \equiv \operatorname{diag}(J_1, J_2, \cdots, J_k)$$

which is called the Jordan canonical form of A, where diag$(\cdot)$ denotes the diagonal matrix and

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{pmatrix} \in \mathbb{C}^{n_i\times n_i},$$

for $i = 1, 2, \cdots, k$, are called Jordan blocks with $n_1 + \cdots + n_k = n$. The Jordan canonical form of A is unique up to the permutation of diagonal Jordan blocks. The eigenvalues $\lambda_i, i = 1, 2, \cdots, k$, are not necessarily distinct. If $A \in \mathbb{R}^{n\times n}$ with only real eigenvalues, then the matrix X can be taken to be real.

Let

$$x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{C}^n.$$

A vector norm on $\mathbb{C}^n$ is a function that assigns to each $x \in \mathbb{C}^n$ a real number $\|x\|$, called the norm of x, such that the following three properties are satisfied for all $x, y \in \mathbb{C}^n$ and all $\alpha \in \mathbb{C}$:

(i) $\|x\| > 0$ if $x \neq 0$ and $\|0\| = 0$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$;

(iii) $\|x + y\| \leq \|x\| + \|y\|$.

A useful class of vector norms is the p-norm defined by

$$\|x\|_p \equiv \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}.$$

The following p-norms are the most commonly used norms in practice:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}, \quad \|x\|_\infty = \max_{1\leq i\leq n} |x_i|.$$

Cauchy–Schwarz's inequality concerning $\|\cdot\|_2$ is given as follows,

$$|x^*y| \leq \|x\|_2\|y\|_2$$

for x, $y \in \mathbb{C}^n$. A very important property of vector norms on $\mathbb{C}^n$ is that all vector norms on $\mathbb{C}^n$ are equivalent, i.e., if $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are two norms on $\mathbb{C}^n$, then there exist two positive constants c_1 and c_2 such that

$$c_1\|x\|_\alpha \leq \|x\|_\beta \leq c_2\|x\|_\alpha$$

for all $x \in \mathbb{C}^n$. For example, if $x \in \mathbb{C}^n$, then we have

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2,$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

and

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty.$$

Let

$$A = (a_{i,j})_{i,j=1}^n \in \mathbb{C}^{n\times n}.$$

We now turn our attention to matrix norms. A matrix norm is a function that assigns to each $A \in \mathbb{C}^{n\times n}$ a real number $\|A\|$, called the norm of A, such that the following four properties are satisfied for all $A, B \in \mathbb{C}^{n\times n}$ and all $\alpha \in \mathbb{C}$:

(i) $\|A\| > 0$ if $A \neq 0$ and $\|0\| = 0$;
(ii) $\|\alpha A\| = |\alpha|\|A\|$;
(iii) $\|A + B\| \leq \|A\| + \|B\|$;
(iv) $\|AB\| \leq \|A\|\|B\|$.

For every vector norm, we can define a matrix norm in a natural way. Given the vector norm $\|\cdot\|_v$, the matrix norm induced by $\|\cdot\|_v$ is defined by

$$\|A\|_v \equiv \max_{x\neq 0} \frac{\|Ax\|_v}{\|x\|_v}.$$

The most important matrix norms are the matrix p-norms induced by the vector p-norms for $p = 1, 2, \infty$:

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^n |a_{i,j}|, \quad \|A\|_2 = \sigma_{\max}(A), \quad \|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^n |a_{i,j}|,$$

where $\sigma_{\max}(A)$ denotes the largest singular value of A, see [38]. The Frobenius norm is defined by

$$\|A\|_F \equiv \left(\sum_{j=1}^{n} \sum_{i=1}^{n} |a_{i,j}|^2 \right)^{1/2}.$$

One of the most important properties of $\|\cdot\|_2$ and $\|\cdot\|_F$ is that for any unitary matrices Q and Z,

$$\|A\|_2 = \|QAZ\|_2$$

and

$$\|A\|_F = \|QAZ\|_F.$$

We recall that a matrix $M \in \mathbb{C}^{n\times n}$ is called unitary if $M^{-1} = M^*$.

Let $A = (a_{i,j}) \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$. The $mp \times nq$ matrix

$$A \otimes B \equiv \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}$$

is called the tensor product of A and B.

The basic properties of the tensor product are summarized in the following theorem.

Theorem 1.6 *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$. Then we have,*

(i) *rank*$(A \otimes B) =$ *rank*$(A) \cdot$ *rank*(B);

(ii) $(A \otimes B)^* = A^* \otimes B^*$;

(iii) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, *where* $C \in \mathbb{C}^{n\times k}$, $D \in \mathbb{C}^{q\times r}$;

(iv) *If both A and B are invertible, then $A \otimes B$ is also invertible and*

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

1.1.2 Condition number and error estimates

When we solve a linear system $Ax = b$, a good measurement, which can tell us how sensitive the computed solution is to the input perturbations, is needed. The condition number of matrix is then defined. It relates the perturbations in x to the perturbations in A and b.

Definition 1.1 *Let $\|\cdot\|$ be any p-norm of matrix and A be an invertible matrix. The condition number of A is defined as follows,*

$$\kappa(A) \equiv \|A\|\|A^{-1}\|. \tag{1.1}$$

Obviously, the condition number depends on the matrix norm used. Since

$$1 = \|I\| = \|A \cdot A^{-1}\| \leq \|A\| \cdot \|A^{-1}\|$$

where I is the identity matrix, it follows that $\kappa(A) \geq 1$. When $\kappa(A)$ is small, then A is said to be well conditioned, whereas if $\kappa(A)$ is large, then A is said to be ill conditioned.

Let $\hat{x}$ be an approximation of the exact solution x of $Ax = b$. The error vector is defined as follows,

$$e = x - \hat{x},$$

i.e.,

$$x = \hat{x} + e. \tag{1.2}$$

The absolute error is given by

$$\|e\| = \|x - \hat{x}\|$$

for any vector norm. If $x \neq 0$, then the relative error is defined by

$$\frac{\|e\|}{\|x\|} = \frac{\|x - \hat{x}\|}{\|x\|}.$$

We have by substituting (1.2) into $Ax = b$,

$$A(\hat{x} + e) = A\hat{x} + Ae = b.$$

Therefore,

$$A\hat{x} = b - Ae = \hat{b}.$$

此为试读,需要完整PDF请访问: www.ertongbook.com