Applied Mathematics and Computation 126 (2002) 187–197
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

A note on the fast algorithm for block Toeplitz systems with tensor structure ☆

Kit-Ian Kou, Vai-Kuong Sin, Xiao-Qing Jin *

Faculty of Science and Technology, University of Macau, Macau

Abstract

We study the solutions of block Toeplitz systems $T_{mn}x = b$ by using the preconditioned conjugate gradient (PCG) method. Here $T_{mn} = T_m \otimes T_n$ and $T_i,\ i = m, n$ are Toeplitz matrices. In [X. Jin, Appl. Math. Comput. 73 (1995) 115–124], Jin introduced a fast algorithm for these systems by applying the PCG method. This fast algorithm allows a tensor problem to be reduced to a one-dimensional problem. It was proved that if the *mn*-by-*mn* system is well conditioned, then the PCG method converges superlinearly and only $O(mn \log mn)$ operations are required in solving the preconditioned system. However, only well-conditioned systems were considered in Jin, 1995. In this paper, we apply this fast algorithm with the $\{\omega\}$-circulant preconditioners proposed in [D. Potts, G. Steidl, Preconditioners for Ill-Conditioned toeplitz matrices, BIT, to be appeared] to solve the ill-conditioned systems. Numerical results are included to illustrate the effectiveness of the algorithm for solving the preconditioned systems by using the PCG method. An application in image restoration is also given. © 2002 Elsevier Science Inc. All rights reserved.

Keywords: Block Toeplitz matrix; Preconditioner; PCG method; Image restoration

1. Introduction

We are concerned with the numerical solutions of linear equation $T_{mn}x = b$ where $T_{mn} = T_m \otimes T_n$ and T_i, $i = m, n$ are Toeplitz matrices. This kind of systems occurs in a variety of applications, such as the digital image processing and the two-dimensional inverse heat problems, see [9,10]. In this paper, the conjugate

☆ Research supported by the research grant No. RG018/97-98S/JXQ/FST from the University of Macau and HKRGC grant No. CUHK4207/97P from the Chinese University of Hong Kong.

* Corresponding author.

PII: S0096-3003(00)00150-8

gradient (CG) method has been used efficiently in solving these linear systems. We use a preconditioner P_{mn} to accelerate the convergence rate of the CG method, i.e., instead of solving the original system $T_{mn}x = b$, we solve the following preconditioned system

$$P_{mn}^{-1}T_{mn}x = P_{mn}^{-1}b.$$

The preconditioner P_{mn} should be a good approximation to T_{mn} and $P_{mn}y = d$ should be easily solved for any vectors y and d.

For solving these block Toeplitz systems with tensor structure, Jin [10] introduced a fast algorithm for these systems by using the PCG method. The fast algorithm allows a tensor problem to be reduced to a one-dimensional problem. For the mn-by-mn block Toeplitz matrices generated by the positive, 2π-periodic and continuous functions, only $O(mn \log mn)$ operations are required for solving the preconditioned systems and a superlinear convergence rate of the PCG method is obtained by using the algorithm. The preconditioners used in [10] are the generalization of T. Chan's preconditioners [7] which are called the optimal block-circulant–circulant-block (BCCB) preconditioners, see also [4]. These BCCB preconditioners preserve the tensor structure of T_{mn} and approximate to T_{mn} under the Frobenius norm in the circulant algebra.

However, if the block Toeplitz matrices are generated by the functions with zeros, the corresponding Toeplitz systems are ill conditioned and the BCCB preconditioners are not good enough for solving such ill-conditioned systems by using the PCG method. Moreover, the number of iterations required for convergence will increase with the size of the systems, see [1]. To handle this problem, we emphasize that Jin's fast algorithm still works for solving the ill-conditioned block Toeplitz systems. We therefore apply this fast algorithm with the $\{\omega\}$-circulant preconditioners given in [12] to manage the ill-conditioned problems. If we have full knowledge about the generating functions of the block Toeplitz systems, then these $\{\omega\}$-circulant preconditioners ensure the superlinearly convergence rate for solving the ill-conditioned block Toeplitz systems by using the PCG method.

The outline of the paper is as follows. In Section 2, we introduce the $\{\omega\}$-circulant preconditioners for solving the ill-conditioned block Toeplitz systems with tensor structure by using the PCG method. We will review the fast algorithm given by Jin [10] in Section 3. Since there are no numerical tests in [10], we will give some numerical tests for both well-conditioned and ill-conditioned block Toeplitz systems in Section 4. An application in image restoration is also given.

2. Construction of $\{\omega\}$-circulant preconditioner

Let $\mathscr{C}_{2\pi}$ be the space of all 2π-periodic continuous real-valued functions. The Fourier coefficients of a function f in $\mathscr{C}_{2\pi}$ are given by

$$a_r = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\mathrm{e}^{-\mathrm{i}rx}\,\mathrm{d}x, \quad r = 0, \pm 1, \pm 2, \ldots$$

Obviously, $a_r = \bar{a}_{-r}$ for all r. Let $T_n[f]$ be the n-by-n Hermitian Toeplitz matrix with the (i,j)th entry given by a_{i-j}, $i,j = 0,\ldots,n-1$. We will use $\mathscr{C}_{2\pi}^{+}$ to denote the space of all non-negative functions in $\mathscr{C}_{2\pi}$ which have only a finite number of zeros. We note that the Toeplitz matrices $T_n[f]$ generated by $f \in \mathscr{C}_{2\pi}^{+}$ are positive definite for all n, see [2]. In this paper, we consider the generating functions $f \in \mathscr{C}_{2\pi}^{+}$ and construct the following n-by-n preconditioners $P_n[f]$ which were proposed by Potts and Steidl [12].

We first choose an equispaced grid

$$x_k := w_n - \pi + \frac{2\pi k}{n}, \quad w_n \in [0, \frac{2\pi}{n})$$

such that

$$f(x_k) \neq 0 \tag{1}$$

for all $k = 0,\ldots,n-1$. Note that the choice of the grids requires some preparatory information about the zeros of f. We now consider the preconditioner $P_n[f]$ that is of the form

$$P_n[f] := \begin{pmatrix} \hat{a}_0 & \hat{a}_{n-1}\mathrm{e}^{\mathrm{i}nw_n} & \cdots & \hat{a}_1\mathrm{e}^{\mathrm{i}nw_n} \\ \hat{a}_1 & \widehat{a_0} & & \\ \vdots & & \ddots & \vdots \\ \hat{a}_{n-1} & \ldots & \ldots & \hat{a}_0 \end{pmatrix}$$

with

$$\hat{a}_r = \hat{a}_r(f) := \frac{1}{n}\sum_{k=0}^{n-1} f(x_k)\mathrm{e}^{-\mathrm{i}rw_n}\mathrm{e}^{-2\pi\,\mathrm{i}rk/n}, \quad r = 0,1,2,\ldots,n-1.$$

These matrices $P_n[f]$ can also be written as

$$P_n[f] = \Omega_n F_n \Lambda_n \bar{F}_n \bar{\Omega}_n, \tag{2}$$

with

$$[F_n]_{j,k} = \frac{1}{\sqrt{n}}\left(\mathrm{e}^{-2\pi\,\mathrm{i}jk/n}\right)_{k=0}^{n-1}, \quad 0 \leqslant j, \;\; k \leqslant n-1,$$

$$\Omega_n = \mathrm{diag}\left(1, \mathrm{e}^{-\mathrm{i}w_n}, \mathrm{e}^{-\mathrm{i}2w_n}, \ldots, \mathrm{e}^{-\mathrm{i}(n-1)w_n}\right)$$

and

$$\Lambda_n = \mathrm{diag}(f(x_0), f(x_1), f(x_2), \ldots, f(x_{n-1})),$$

where diag$(\cdot)$ is the diagonal matrix.

These n-by-n preconditioners $P_n[f]$ have the following properties:

1. By (1) and $f \in \mathscr{C}_{2\pi}^{+}$, the matrices $P_n[f]$ are Hermitian and positive definite.
2. These matrices $P_n[f]$ are $\{e^{inw_n}\}$-circulant matrices [8]. Notice that $\{e^{inw_n}\}$-circulant matrices are Toeplitz matrices with the first entry of each column obtained by multiplying the last entry of the preceding column by e^{inw_n}. Specially, we obtain the circulant matrices for $w_n = 0$ and the skew-circulant matrices for $w_n = \pi/n$.
3. Similar to that of circulant matrix, once the diagonal matrix Λ_n in (2) is obtained, the products of $P_n[f]y$ and $P_n[f]^{-1}y$ for any vector y can be computed by FFTs in $O(n \log n)$ operations.
4. In view of (2), the matrix $P_n[f]$ can be constructed in $O(n \log n)$ operations and requires only $O(n)$ storage.
5. If $f \in \mathscr{C}_{2\pi}^{+}$ with a zero of even order, then the eigenvalues of $P_n[f]^{-1}T_n[f]$ are clustered around 1 and the spectra of $P_n[f]^{-1}T_n[f]$ are bounded away from zero.
6. The CG method, when applied to solving these preconditioned systems with the preconditioners $P_n[f]$, will converge superlinearly. Therefore, the total complexity in solving the preconditioned systems is reduced to $O(n \log n)$.

Let $f = f(x, y) = f_1(x)f_2(y)$ with $f_i \in \mathscr{C}_{2\pi}^{+}$ $(i = 1, 2)$ and let $T_m[f_1]$ be an m-by-m Toeplitz matrix and $T_n[f_2]$ be an n-by-n Toeplitz matrix. We consider the ill-conditioned block Toeplitz system

$$T_{mn}[f]x = b \quad \text{with} \quad T_{mn}[f] = T_m[f_1] \otimes T_n[f_2]. \tag{3}$$

As a preconditioner for (3), we suggest that the preconditioner is of the form

$$P_{mn}[f] = P_m[f_1] \otimes P_n[f_2]. \tag{4}$$

We note that the block $\{e^{inw_n}\}$-circulant preconditioner $P_{mn}[f]$ is a matrix that conserves the tensor structure of $T_{mn}[f]$. For solving both well-conditioned and ill-conditioned block Toeplitz systems by using the PCG method, the performances of these block $\{e^{inw_n}\}$-circulant preconditioners are more effective than those of the BCCB preconditioners [4], see Section 4.1.

3. Fast algorithm

In this section, we review the fast algorithm proposed in [10] for solving the block Toeplitz system (3). By using the preconditioner $P_{mn}[f]$ given by (4), the preconditioned system is given by

$$\left(P_m[f_1]^{-1}T_m[f_1] \otimes P_n[f_2]^{-1}T_n[f_2]\right)x = \tilde{b}, \tag{5}$$

where $\tilde{b} = P_{mn}[f]^{-1}b$. An algorithm for solving (5) is given as follows.

The Fast Algorithm ([10]).

(i) Solve the equation $\left(I_m \otimes P_n[f_2]^{-1}T_n[f_2]\right)y = \tilde{b}$;

(ii) solve the equation $(P_m[f_1]^{-1}T_m[f_1] \otimes I_n)x = y$.

In step (i), it is clear that the number of distinct eigenvalues of the preconditioned matrix $I_m \otimes P_n[f_2]^{-1}T_n[f_2]$ is the same as the number of distinct eigenvalues of $P_n[f_2]^{-1}T_n[f_2]$. In view of the properties of the $\{\omega\}$-circulant preconditioner $P_n[f_2]$, if $f_2 \in \mathscr{C}^+_{2\pi}$ with a zero of even order, then the eigenvalues of the n-by-n preconditioned matrix $P_n[f_2]^{-1}T_n[f_2]$ are clustered around 1 and the smallest eigenvalues of the preconditioned matrix are bounded from below by a positive constant independent of n. When the PCG method is applied to step (i), the convergence rate will be superlinear and the complexity of step (i) is equal to $\mathrm{O}(mn\log n)$.

For step (ii), with a similar analysis as for step (i), we need $\mathrm{O}(mn\log m)$ operations to solve the preconditioned system in step (ii) by using the PCG method and the convergence rate will be superlinear. Therefore the total number of operations of this algorithm in solving the preconditioned system (5) is of $\mathrm{O}(mn\log mn)$.

4. Numerical experiments

In this section, we illustrate by numerical examples the effectiveness of the fast algorithm (Section 3) with the block algorithm [4] in solving the preconditioned systems (5) by using the PCG method. An application in image restoration is also given.

4.1. Numerical tests

We apply the PCG method with the fast algorithm (Section 3) for solving the block Toeplitz systems (3). The preconditioners used are the $\{e^{inw_n}\}$-circulant preconditioner given by (4) and the BCCB preconditioner given in [4]. For comparisons, we also test the block algorithm proposed in [4].

We test the following six systems with different generating functions defined on $[-\pi, \pi]$. They are separated into two classes:

- Positive continuous generating functions:
  1. $f_1(x) = x^6 + 1,\quad f_2(y) = |y|^3 + 0.01;$
  2. $f_1(x) = (\cos x)^2 + 0.1,\quad f_2(y) = |y|^5 + \pi;$
  3. $f_1(x) = x^2 + \pi/2,\quad f_2(y) = y^4 + 1.$
- Non-negative continuous generating functions with zeros:
  4. $f_1(x) = (x^2 - 1)^2,\quad f_2(y) = y^2;$
  5. $f_1(x) = |x|^3,\quad f_2(y) = |y|^3;$
  6. $f_1(x) = x^4,\quad f_2(y) = y^4 + (\sin y)^2.$

Tables 1–6 show the number of iterations required for convergence with different choices of preconditioners. In these tables, I denotes that no precondi-

Table 1
Number of iterations for $f_1(x) = x^6 + 1$, $f_2(y) = |y|^3 + 0.01$

$m = n$	mn	I			BCCB			P		
		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.	
			(i)	(ii)		(i)	(ii)		(i)	(ii)
16	256	711	10	8	108	8	8	17	6	6
32	1024	5535	28	21	180	15	13	17	7	7
64	4096	☆	59	53	146	13	15	17	7	7
128	16 384	☆	107	132	109	11	18	17	7	7
256	65 536	☆	168	274	78	9	15	17	7	7

Table 2
Number of iterations for $f_1(x) = (\cos x)^2 + 0.1$, $f_2(y) = |y|^5 + \pi$

$m = n$	mn	I			BCCB			P		
		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.	
			(i)	(ii)		(i)	(ii)		(i)	(ii)
16	256	57	4	8	16	4	8	9	2	5
32	1024	142	8	19	18	6	8	10	2	5
64	4096	191	16	33	17	5	7	10	2	5
128	16 384	220	22	52	16	5	6	9	2	5
256	65 536	223	21	64	14	4	6	9	2	5

Table 3
Number of iterations for $f_1(x) = x^2 + \pi/2$, $f_2(y) = y^4 + 1$

$m = n$	mn	I			BCCB			P		
		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.	
			(i)	(ii)		(i)	(ii)		(i)	(ii)
16	256	108	8	8	19	6	8	11	4	5
32	1024	163	16	20	17	5	7	11	4	5
64	4096	190	19	36	16	5	7	11	4	5
128	16 384	202	19	55	14	5	6	10	4	5
256	65 536	203	19	67	14	4	6	10	4	5

tioner is used, BCCB denotes the optimal BCCB preconditioner given in [4] and P denotes the $\{e^{inw_n}\}$-circulant preconditioner given in Section 2. Iteration numbers more than 10 000 are denoted by "☆". In all the tests, we set $w_n = \pi/n$ for the construction of the preconditioner P, i.e., the matrix P is a skew-circulant preconditioner. When the PCG method is applied to such kind of systems, the stopping criteria is

$$\frac{\|r_k\|_2}{\|r_0\|_2} < 10^{-7},$$

Table 4
Number of iterations for $f_1(x) = (x^2 - 1)^2$, $f_2(y) = y^2$

$m = n$	mn	I			BCCB			P		
		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.	
			(i)	(ii)		(i)	(ii)		(i)	(ii)
16	256	359	9	8	64	8	8	22	6	4
32	1024	2608	25	17	125	14	10	25	6	5
64	4096	☆	71	38	271	17	12	24	6	6
128	16 384	☆	191	82	559	22	14	32	8	6
256	65 536	☆	460	172	1260	27	17	28	8	6

Table 5
Number of iterations for $f_1(x) = |x|^3$, $f_2(y) = |y|^3$

$m = n$	mn	I			BCCB			P		
		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.	
			(i)	(ii)		(i)	(ii)		(i)	(ii)
16	256	271	8	8	56	8	8	14	6	6
32	1024	3433	22	22	206	13	13	19	7	7
64	4096	☆	60	60	1196	17	17	25	9	9
128	16 384	☆	176	176	5860	24	24	33	9	9
256	65 536	☆	562	562	☆	36	36	51	9	9

Table 6
Number of iterations for $f_1(x) = x^4$, $f_2(y) = y^4 + (\sin y)^2$

$m = n$	mn	I			BCCB			P		
		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.		Block Algo.	Fast Algo.	
			(i)	(ii)		(i)	(ii)		(i)	(ii)
16	256	1166	10	9	146	8	8	22	6	6
32	1024	☆	33	24	843	16	14	31	8	6
64	4096	☆	113	64	4688	25	17	43	9	7
128	16 384	☆	487	170	☆	44	22	51	10	7
256	65 536	☆	2399	432	☆	136	29	70	12	7

where r_k is the residual vector after kth iterations. The right-hand side vector b is chosen to be the vector of all ones and the zero vector is the initial guess. All computations are produced by Matlab program.

In all tests, when we compare the fast algorithm with the block algorithm proposed in [4], the performance of the fast algorithm is better. In each step of the fast algorithm, the preconditioned systems with the skew-circulant preconditioners converge at a rate that is independent of the sizes of the matrices, while the preconditioned systems with the BCCB preconditioners converge at a rate that increases slowly with the sizes of the matrices. For the case that non-negative functions $f \in \mathscr{C}_{2\pi}^{+}$ having zeros of odd orders, we can obtain the same numerical behavior, see Table 5. We note that the BCCB preconditioners are not good enough for the last three systems which have generating functions with zeros on $[-\pi, \pi]$. This is the reason why we apply the skew-circulant preconditioners to handle these ill-conditioned systems. In these tables, Fast Algo. means the fast algorithm and Block Algo. means the block algorithm.

4.2. Image restoration

In this section, we consider the applications of the fast algorithm proposed in Section 3 for solving the linear systems arising from image restoration. The mathematical model of the linear image restoration problem is given as follows, see [9],

$$g(\xi, \delta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(\xi - \alpha, \delta - \beta) f(\alpha, \beta)\, \mathrm{d}\alpha\, \mathrm{d}\beta + \eta(\xi, \delta), \tag{6}$$

where $g(\xi, \delta)$ is the recorded (or degraded) image, $f(\alpha, \beta)$ is the ideal (or original) image and the vector $\eta(\xi, \delta)$ represents the additive noise. The function t is called the point spread function (PSF) and represents the degradation of the image. Since the PSF here is a function t of $\xi - \alpha$ and $\delta - \beta$, the function t is said to be spatially invariant. The integral in (6) is a two-dimensional convolution.

In the digital implementation of (6), the integral is discretized by using some quadrature rule to obtain the discrete scalar model

$$g(i, j) = \sum_{k=1}^{n} \sum_{l=1}^{n} t(i - k, j - l) f(k, l) + \eta(i, j).$$

In matrix form, we have the following linear algebraic system of the image restoration problem,

$$g = Tf + \eta, \tag{7}$$

where g, η and f are n^2-vectors and T is an n^2-by-n^2 block-Toeplitz–Toeplitz-block (BTTB) matrix. This is the square image formulation. The PCG method is proposed as a main tool to solve (7), see [3–6,11].

The following example is constructed. We first generate the 256×256 image f shown in Fig. 1 (left), and consider the spatially invariant discretized PSF matrix T with the entries given by

$$t_{i-j,k-l} = \begin{cases} \exp\{-0.05(i-j)^2 - 0.05(k-l)^2\}, & -8 \leqslant i-j,\ k-l \leqslant 8, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

The matrix for this example is a $65\,536 \times 65\,536$ BTTB matrix with tensor structure. Note that the noise function η is modeled as a random process. For simplicity, we neglect the noise vector η and only consider the ill-posed problem $g = Tf$. Here g is the observed image, T is defined by (8) and f is a vector formed by row ordering the original image in Fig. 1 (left), see [6]. By unstacking the vector g, we obtain the blurred image in Fig. 1 (right).

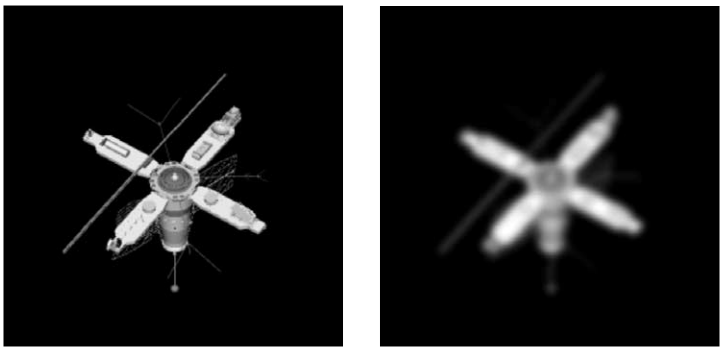

Fig. 1. Original image (left) and observed (blurred) image (right).

Table 7
Convergence results for the image restoration

Preconditioner used	Block Algo.	Fast Algo.	
		Step (i)	Step (ii)
I	909	71	135
BCCB	172	8	18

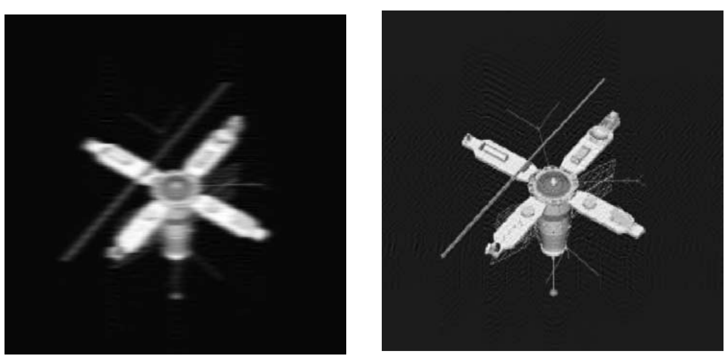

Fig. 2. Restored image with I: step (i) 71 (left) and step (ii) 135 iterations (right).

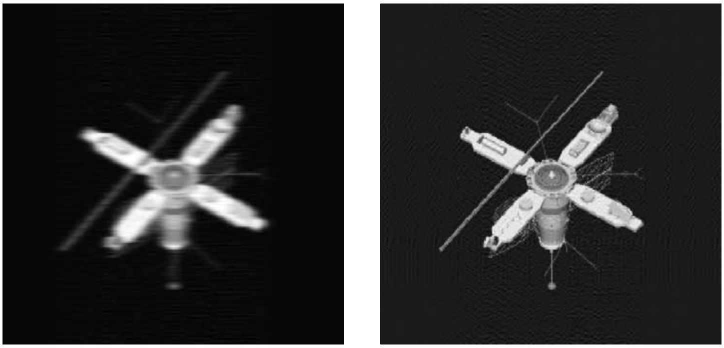

Fig. 3. Restored images with BCCB: step (i) 8 (left) and step (ii) 18 iterations (right).

Our goal is: given g and T, to recover an approximation to the original image f. Since T has tensor structure, we can apply the PCG method with the fast algorithm proposed in Section 3 to solve the BTTB system $Tf = g$. We compare the performance of the fast algorithm and the block Algorithm [4] with different preconditioners. The stopping criteria is $\|r_k\|_2/\|r_0\|_2 < 10^{-4}$. Table 7 shows the number of iterations required for convergence and the results for this image restoration problem are summarized in Figs. 1–3.

Acknowledgements

Part of the work was performed when the authors were visiting the Chinese University of Hong Kong. The hospitality of Prof. Raymond H. Chan is gratefully acknowledged. The authors would also like to thank Dr. H. Sun from the Guangdong University of Technology and Mr. W. Tang from the Chinese University of Hong Kong for their helpful suggestions and guidance in the programming of this paper.

References

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

[2] R. Chan, Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions, IMA J. Numer. Anal. 11 (1991) 333–345.

[3] R. Chan, M. Ng, Conjugate gradient methods for Toeplitz systems, SIAM Rev. 38 (1996) 427–482.

[4] R. Chan, X. Jin, A family of block preconditioners for block system, SIAM J. Sci. Statist. Comput. 13 (1992) 1218–1235.

[5] R. Chan, T. Chan, C. Wong, Cosine transform based preconditioners for total variation deblurring, IEEE Trans. Image Processing 8 (1999) 1472–1478.

[6] R. Chan, J. Nagy, R. Plemmons, FFT-based preconditioners for Toeplitz-block least squares problems, SIAM J. Sci. Numer. Anal. 30 (1993) 1740–1768.
[7] T. Chan, An optimal circulant preconditioner for Toeplitz systems, SIAM J. Sci. Statist. Comput. 9 (1988) 766–771.
[8] P. Davis, Circulant Matrices, Wiley, New York, 1979.
[9] R. Gonzalez, R. Woods, Digital Image Processing, Addison-Wesley, New York, 1992.
[10] X. Jin, A fast algorithm for block Toeplitz systems with tensor structure, Appl. Math. Comput. 73 (1995) 115–124.
[11] K. Kou, X. Jin, V. Sin, The sine transform operator in the Banach space of symmetric matrices and its application in image restoration, Numer. Math. J. Chinese Univ. 8 (1999) 231–242.
[12] D. Potts, G. Steidl, Preconditioners for ill-conditioned Toeplitz matrices, BIT (1999) 513–533.