Structured Matrix Computations and Applications

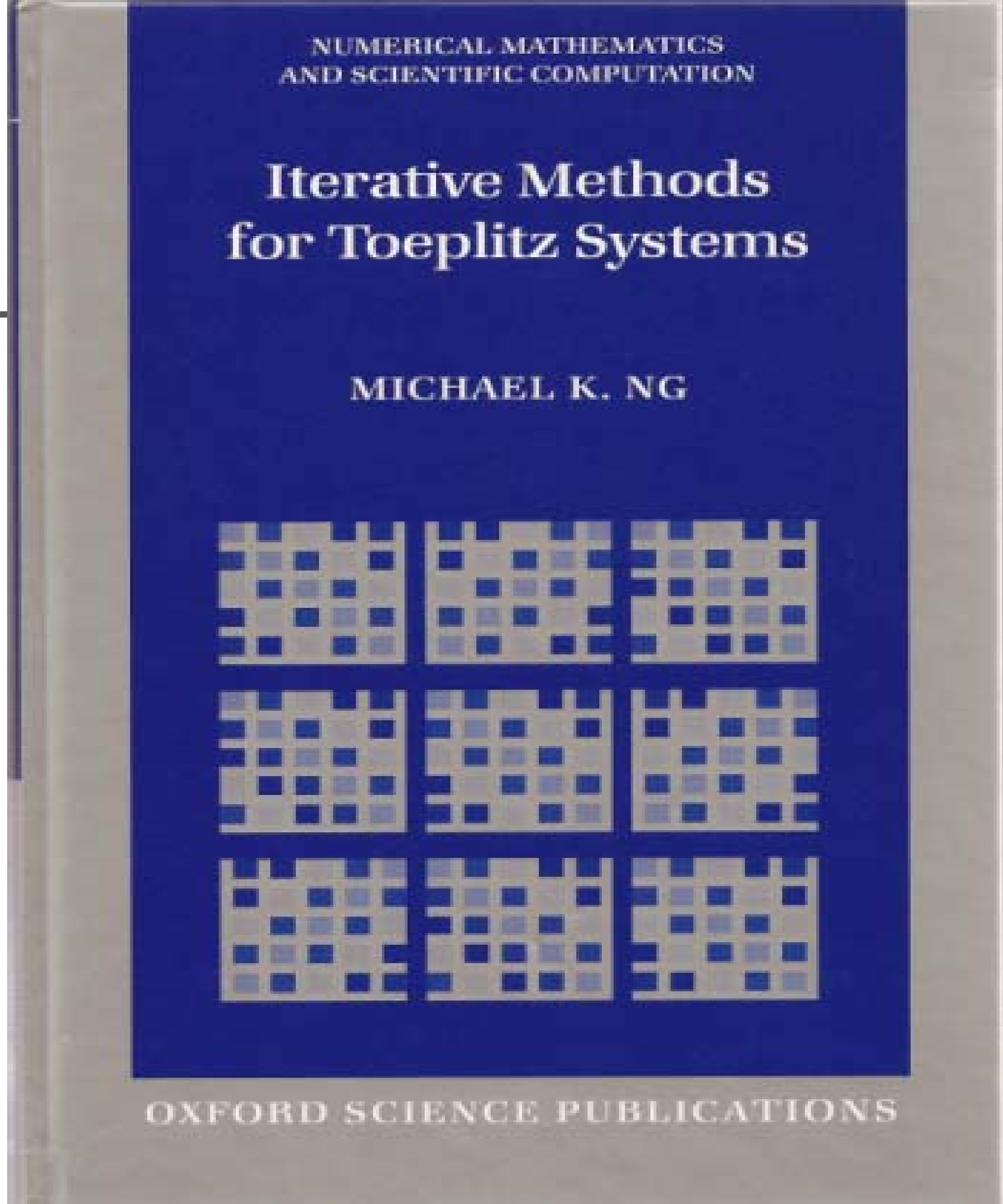
Michael K. Ng
Department of Mathematics
Hong Kong Baptist University
mng@math.hkbu.edu.hk

Outline

- Structured matrices have been around for a long time and are encountered in various fields of application.
- Toeplitz matrices, circulant matrices, Hankel matrices, semiseparable matrices, Kronecker product matrices, 2-by-2 block matrices …

Outline

- Toeplitz Matrices
- Overview
- Theory
- Applications
- Research Problems


NUMERICAL MATHEMATICS
AND SCIENTIFIC COMPUTATION
Iterative Methods
for Toeplitz Systems
MICHAEL K. NG
OXFORD SCIENCE PUBLICATIONS

Toeplitz Matrices

- A matrix is said to be a Toeplitz matrix if it is constant along its diagonals

$$\begin{bmatrix} t_0 & t_{-1} & \cdots & t_{2-n} & t_{1-n} \\ t_1 & t_0 & t_{-1} & & t_{2-n} \\ \vdots & t_1 & t_0 & \ddots & \vdots \\ t_{n-2} & & \ddots & \ddots & t_{-1} \\ t_{n-1} & t_{n-2} & \cdots & t_1 & t_0 \end{bmatrix}$$

Background

- The name Toeplitz originates from the work of Otto Toeplitz (1911) on blinear forms related to Laurent series
- Time series: Yule-Walker Equations (1927)
- Levinson's work (1947) in formulating the Wiener filtering problem

An example of Toeplitz system

- Linear prediction is a particularly important topic in digital signal processing
- The determination of the optimal linear filter for prediction requires the solution of a set of linear equations having a Toeplitz structure
- Stationary time series

$$\hat{x}(t) = -\sum_{k=1}^{n} h_n(k)x(t-k)$$

where $-h_n(k)$, $k = 1, \ldots, n$, represent the weights in the linear combination. These weights are called the prediction coefficients of the one-step forward predictor of order n.

The difference between the value $x(t)$ and the predicted value $\hat{x}(t)$ is called the forward prediction error, and is denoted by $e(t)$. Thus

$$e(t) = x(t) - \hat{x}(t) = x(t) + \sum_{k=1}^{n} h_n(k)x(t-k).$$

The mean square value of the forward prediction error is

$$\mathsf{E}(|e(t)|^2) = \mathsf{E}\left(\left|x(t) + \sum_{k=1}^{n} h_n(k)x(t-k)\right|^2\right),$$

where E is the expectation operator. Since this is a quadratic function of the predictor coefficients, the minimization of $\mathsf{E}(|e(t)|^2)$ yields the set of linear equations

$$\gamma_{xx}(j) = -\sum_{k=1}^{n} h_n(k)\gamma_{xx}(j-k), \quad j = 1, 2, \ldots, n.$$

Here

$$\gamma_{xx}(k) = \mathsf{E}(x(t)\overline{x(t-k)})$$

Direct Methods

- Schur algorithm (1917) – a test for determining the positive definiteness of a Toeplitz matrix
- Levinson (1947)
- Durbin (1960)
- Trench (1964)
- $O(n^2)$ algorithms
- Small → Large systems (Recursive)

Direct Methods

- The Gohberg-Semencul Formula

$$T_{m+1}^{-1} = \frac{1}{\delta_m}\left\{L_1^t L_1 - L_2^t L_2\right\}, \quad m = 0, 1, \ldots, n-1,$$

L_1 and L_2 are lower triangular Toeplitz matrices given by

$$L_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ r_{m-1} & 1 & 0 & & 0 \\ \vdots & r_{m-1} & 1 & \ddots & \vdots \\ r_2 & & \ddots & \ddots & 0 \\ r_1 & r_2 & \cdots & r_{m-1} & 1 \end{bmatrix} \quad \text{and} \quad L_2 = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ r_1 & 0 & 0 & & 0 \\ \vdots & r_1 & 0 & \ddots & \vdots \\ r_{m-2} & & \ddots & \ddots & 0 \\ r_{m-1} & r_{m-2} & \cdots & r_1 & 0 \end{bmatrix},$$

Superfast Direct Toeplitz Solvers

- Brent et al. (1980)
- Bitmead and Anderson (1980)
- Morf (1980)
- de Hong (1986)
- Ammar Gragg (1988)
- $O(n\log^2 n)$ algorithms
- Recursive from n→n/2→n/4→n/8→…

Look-ahead algorithms

- Singular or ill-conditioned principal submatrices
- Avoid breakdowns or near-breakdowns by skipping such submatrices
- Gueguen (1981), Delsarte et al (1985), Chan and Hansen (1992), Sweet (1993)
- Worst case: $O(n^3)$ algorithms

Stability

- The stability properties of symmetric positive definite Toepltiz matrices: Sweet (1984), Bunch (1985), Cybenko (1987), Bojanczyk et al (1995)
- Weakly stable (residual is small for well-conditioned matrices)
- Look-ahead methods are stable

Stability

- Toeplitz matrices → Cauchy matrices
- Partial pivoting → stable ? Gohberg et al (1995)
- Displacement representation → error growth
- Gu (1995), Chandrasekaran and Sayed (1996), Park and Elden (1996): QR-type algorithm on displacement representation → stable

Iterative Methods

- Rino (1970) and Ekstrom (1974): a decomposition of Toeplitz matrix into a circulant matrices and iterative methods
- Strang (1986), Olkin (1986): the use of preconditioned conjugate gradient method with circulant matrices as preconditioners for Toeplitz systems

Circulant Preconditioners

- Circulant matrices: Toeplitz matrices where each column is a circular shift of its preceding column

$$C = \begin{bmatrix} c_0 & c_{n-1} & \cdots & & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & & & c_2 \\ c_2 & c_1 & c_0 & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ c_{n-2} & & & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \cdots & c_2 & c_1 & c_0 \end{bmatrix}$$

Circulant Preconditioners

$$C_n = F_n^* \Lambda_n F_n$$

where the Fourier matrix F_n is the matrix with entries given by

$$[F_n]_{j,k} = \frac{1}{\sqrt{n}} e^{-2\pi i j k / n}, \quad 0 \le j, k \le n-1.$$

- Design of circulant matrices, Strang's preconditioners (1986), R. Chan's preconditioners (1988), Ku and Kuo's preconditioners (1992) …

Circulant Preconditioners

- T. Chan preconditioners (1988) → optimal preconditioners

  $\min \| C - T \|_F$

- Tyrtyshnikov preconditioners (1992) → superoptimal preconditioners

  $\min \| I - C^{-1}T \|_F$

Results

$$t_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{ik\theta}d\theta, \quad k = 0, \pm 1, \pm 2, \ldots.$$

Theorem 4.10. (Performance of T. Chan's preconditioners) *(Chan and Yeung, 1992b) Let f be a 2π-periodic continuous positive function. Then the spectra of $c(T_n)^{-1}T_n$ are clustered around 1 for large n.*

n	No	S_n	$c(T_n)$	R_n	$K_n^{(2)}$	P_n	Y_n	M_n
16	8	8	8	6	6	8	8	14
32	20	8	7	5	5	10	16	14
64	37	6	7	5	5	7	18	11
128	56	5	6	5	5	7	13	9
256	67	5	6	5	5	6	10	8
512	70	5	6	5	5	6	8	8

Table 4.1 *Number of iterations for different preconditioned systems.*

Results

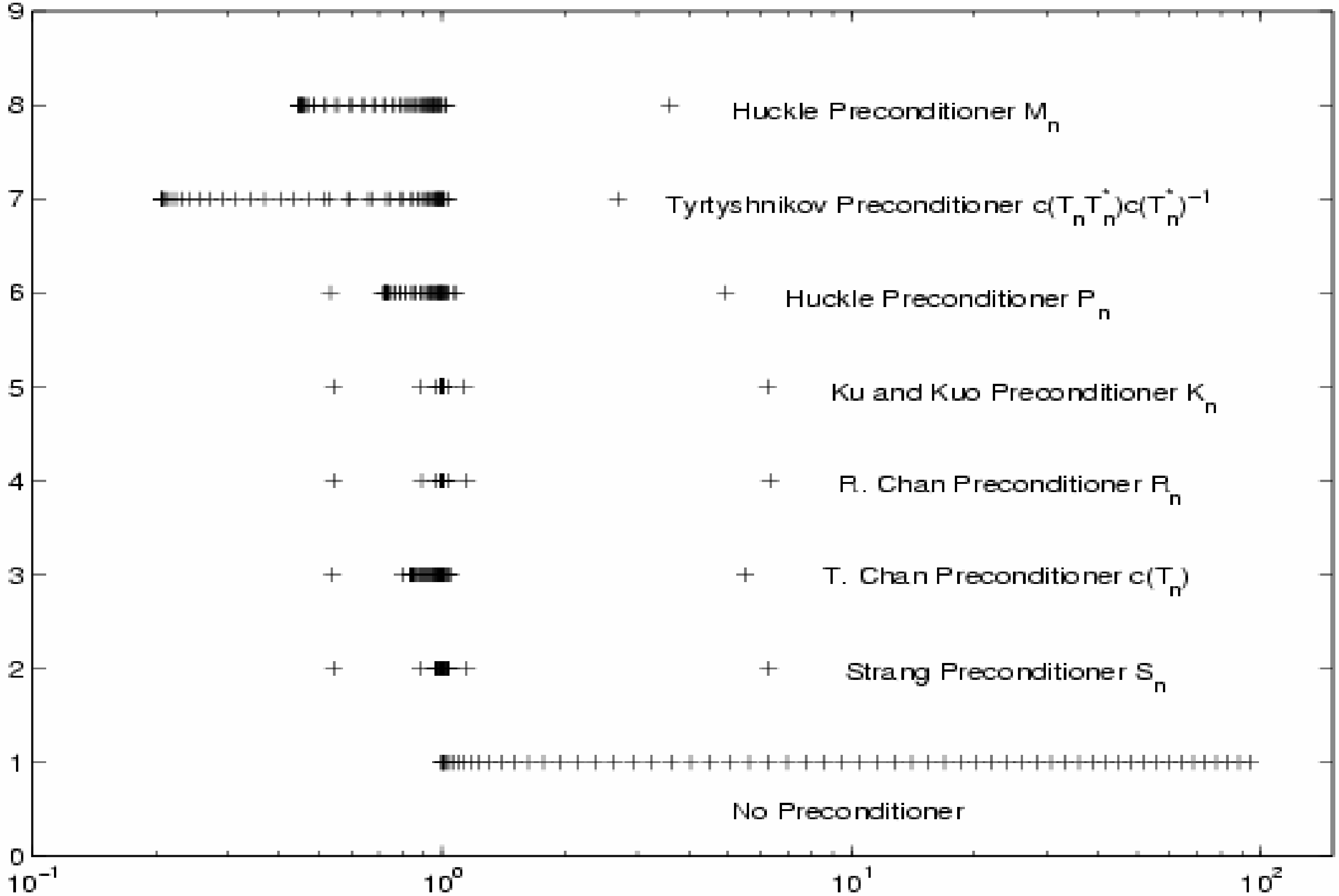


FIG. 4.1. Spectra of preconditioned matrices for $n = 64$.

Circulant Preconditioners

- The eigenvalues of Strang’s preconditioner is the values of the convolution product of the Dirichlet kernel
- The eigenvalues of T. Chan’s preconditioner is the values of the convolution product of the Fejer kernel
- Convergence results/different conditions

Transform-based Preconditioners

- Circulant matrices are precisely those matrices that can diagonalized by the discrete Fourier transform
- Sine transform
- Cosine transform
- Hartley transform
- A effective basis: shift matrices

Ill-conditioned systems

- Zeros of f

Theorem 3.8. (Spectra and Zeros) *(Kac et al., 1953; Kesten, 1962; Parter, 1962) Suppose that $f(\theta) - f_{\min}$ has a unique zero of order 2ν at $\theta = \theta_0$. Then $\lambda_{\min}(T_n)$ has the asymptotic expansion*

$$\lambda_{\min}(T_n) = f_{\min} + c\frac{f^{(2\nu)}(\theta_0)}{(2\nu)!}n^{-2\nu} + O(n^{-2\nu-1}), \quad n = 1, 2, \ldots$$

where c is a constant dependent on f and ν.

Ill-conditioned Systems

$$s(\theta) = \prod_{j=1}^{k} [2 - 2\cos(\theta - \theta_j)]^{\nu_j}$$

Theorem 6.3. (Band-Toeplitz Preconditioner II) *(Chan, 1993) Let f be a nonnegative periodic function defined in $[-\pi, \pi]$ with zeros attained at $\{\theta_j\}_{j=1}^{k}$ and their orders are $\{\nu_j\}_{j=1}^{k}$. Define $s(\theta)$ as in (6.6). Then there exist constants $c_1, c_2 > 0$, such that*

$$c_1 \le \frac{f(\theta)}{s(\theta)} \le c_2, \quad \forall \theta \in [-\pi, \pi].$$

In particular, $\kappa(T_n[s]^{-1} T_n[f]) \le c_2/c_1$ for all $n > 0$.

Ill-conditioned Systems

- The generalized Jackson kernel forms an approximate convolution identity → match the zeros automatically

	θ^2						θ^4					
n	32	64	128	256	512	1024	32	64	128	256	512	1024
I	36	79	170	362	753	1544	63	209	790	2149	†	†
S	–	–	–	–	–	–	–	–	–	–	–	–
C	12	16	19	23	29	39	26	42	71	161	167	247
$K_{N,4}$	8	9	10	9	9	9	15	17	20	24	26	26
$K_{N,6}$	10	10	10	10	9	9	15	16	18	18	17	18
$K_{N,8}$	9	10	10	10	10	10	16	17	19	19	19	20

Table 6.1 *Numbers of iterations required for difference preconditioners.*

Multigrid Methods

- Use projection/restriction operators to generate a sequence of sub-systems
- The zeros can be matched (zeros of f)

Theorem 6.25. (Level-Independent Convergence) *Let $f(\theta)$ be such that*

$$c_2(1 \pm \cos(l\theta)) \geq f(\theta) \geq c_1(1 \pm \cos(l\theta)),$$

for some integer l and positive constants c_1 and c_2. Then for any $1 \leq m \leq q$,

$$\|G^m T^m\|_1 \leq \sqrt{1 - \frac{c_1}{2c_2}}.$$

Recursive Preconditioners

- Use the principal submatrices as preconditioners
- Match the zeros automatically
- Solve the subsystems recursively
- Idea of direct methods
- Use the Gohberg-Semencul formula to represent the inverses of submatrices

Block-Toeplitz-Toeplitz-block Systems

$$T_{mn} = \begin{bmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,m} \\ T_{2,1} & T_{2,2} & \cdots & T_{2,m} \\ \vdots & \ddots & \ddots & \vdots \\ T_{m,1} & T_{m,2} & \cdots & T_{m,m} \end{bmatrix}$$

$$= \begin{bmatrix} T_{(0)} & T_{(1)} & \cdots & T_{(m-1)} \\ T_{(1)} & T_{(0)} & \cdots & T_{(m-2)} \\ \vdots & \ddots & \ddots & \vdots \\ T_{(m-1)} & T_{(m-2)} & \cdots & T_{(0)} \end{bmatrix}.$$

Results

Block-circulant-circulant-block preconditioners

can be defined based on the block structure

Theorem 7.14. (Clustered Spectra) *Let T_{mn} be given by (7.38) with an absolutely summable generating sequence. Then for all $\varepsilon > 0$, there exists an $N > 0$ such that for all $n > N$ and all $m > 0$, at most $O(m)$ eigenvalues of $c_F^{(1)}(T_{mn}) - T_{mn}$ have absolute values exceeding ϵ. Therefore if T_{mn} are positive definite for all m and n and $\lambda_{\min}(T_{mn}) \geq \delta > 0$, then for all $\epsilon > 0$ there exists an $N > 0$ such that for all $n > N$ and all $m > 0$, at most $O(m)$ eigenvalues of $[c_F^{(1)}(T_{mn})]^{-1}T_{mn} - I$ have absolute value larger than ϵ.*

Toeplitz Least Squares Problems

- Min || T x – b ||$^2_2$

$$c(k) = e^{-0.1*k^2}, \quad k = 1, 2, \ldots, m$$
$$r(k) = e^{-0.1*k^2}, \quad k = 1, 2, \ldots, n.$$

		Example (ii)		
n	m	no prec	disp prec	part prec
64	128	24	15	12
64	256	46	15	11
64	512	79	13	10
64	1024	132	11	9
64	2048	177	10	9

Applications to PDEs

- An elliptic problem on the unit-square with Dirichlet boundary conditions
- Circulant preconditioners are not optimal → condition number O(n)
- Sine transform based preconditioners are optimal → condition number O(1)
- Boundary conditions are matched

Example

$$\frac{\partial}{\partial x}\left[\left(1+\varepsilon e^{x+y}\right)\frac{\partial u}{\partial x}\right]+\frac{\partial}{\partial y}\left[\left(1+\frac{\varepsilon}{2}\sin(2\pi(x+y))\right)\frac{\partial u}{\partial y}\right]=g(x,y)$$

	$\varepsilon = 0.01$					$\varepsilon = 1.0$				
$1/h$	No	C_B	P	MILU	MINV	No	C_B	P	MILU	MINV
4	12	9	3	7	4	15	10	5	6	3
8	25	12	3	9	5	29	13	7	9	4
16	47	15	3	13	7	54	18	9	14	6
32	90	20	3	20	11	107	25	11	20	10
64	186	25	3	28	16	209	35	12	28	15
128	363	33	3	41	24	419	50	13	41	22

Table 8.1 *Number of iterations for the unit square.*

	$\varepsilon = 0.01$				$\varepsilon = 1.0$			
$1/h$	No	P	MILU	MINV	No	P	MILU	MINV
8	22	3	9	4	24	7	9	4
16	40	3	12	6	45	9	13	6
32	80	4	17	9	86	10	18	8
64	155	4	25	14	169	12	26	12
128	311	4	36	21	338	14	37	19

Table 8.2 *Number of iterations for the L-shaped domain.*

Sinc-Galerkin Methods for BVPs

- Toeplitz-plus-diagonal systems
- Toeplitz ←→ known f / banded prec.

M	No	E_s	E_u	B	E_s	E_u
4	79	4.1×10^{-3}	8.9×10^{-2}	9	4.1×10^{-3}	8.9×10^{-2}
8	522	3.7×10^{-4}	1.8×10^{-2}	9	3.7×10^{-4}	1.8×10^{-2}
16	>1000	***	***	9	1.1×10^{-5}	1.4×10^{-3}
32	>1000	***	***	9	7.6×10^{-8}	2.2×10^{-5}
64	>1000	***	***	9	1.1×10^{-9}	7.4×10^{-8}

Table 8.3 *Number of iterations required for convergence and the errors E_s and E_u between the numerical approximation and the true solution.*

M	No	E_s	E_u	B	E_s	E_u
4	108	6.6×10^{-4}	8.5×10^{-3}	17	6.6×10^{-4}	8.5×10^{-3}
8	961	1.1×10^{-4}	1.3×10^{-4}	21	1.1×10^{-4}	1.3×10^{-4}
16	>1000	***	***	25	5.8×10^{-6}	7.2×10^{-5}
32	>1000	***	***	30	6.0×10^{-8}	8.9×10^{-7}
64	>1000	***	***	31	5.3×10^{-9}	6.9×10^{-8}

Table 8.4 *Number of iterations required for convergence and the errors E_s and E_u between the numerical approximation and the true solution.*

PDEs

- Hyperbolic and parabolic equations

$$\frac{\partial u}{\partial t} + a(x,y)\frac{\partial u}{\partial x} + b(x,y)\frac{\partial u}{\partial y} = g(x,y)$$

- Block-circulant preconditioners by Holmgren and Otto (1992), Jin and Chan (1992), Hemmingsson (1996)

Applications to ODEs

$$\begin{cases} \frac{\mathrm{d}\mathbf{y}(t)}{\mathrm{d}t} = J\mathbf{y}(t) + \mathbf{g}(t),\ t \in (t_0, T]; \\ \mathbf{y}(t_0) = \mathbf{z}, \end{cases}$$

$$M_s = A_s \otimes I_m - hB_s \otimes J_m$$

Boundary value methods

Results

m	s	$\kappa(A_s)$	$\kappa(\hat{A}_s)$	$\kappa(B_s)$	$\kappa(\hat{B}_s)$	$\kappa(M_s)$	$\kappa(\hat{M}_s)$
10	8	10.6	10.7	3.2	9.3	44.6	49.0
10	16	21.0	21.0	3.7	9.5	44.5	48.0
10	32	41.3	41.0	3.9	9.5	44.1	48.0
20	8	10.6	10.7	3.2	9.3	164.2	181.0
20	16	21.0	21.0	3.7	9.5	163.8	179.0
20	32	41.3	41.0	3.9	9.5	163.6	178.0

Table 8.5 *Heat equation. Condition numbers for different sizes of the underlying Toeplitz matrices and their small rank perturbations counterparts. The formula (8.52) with* $k = 4$ *is used here.*

		CGN			BiCGStab		GMRES	
m	s	I	P	S	I	P	I	P
10	8	128	24	26	51	7	63	9
10	16	171	23	23	63	7	79	8
10	32	186	20	20	70	8	89	8
20	8	439	28	29	130	7	130	8
20	16	613	31	37	162	6	173	8
20	32	679	29	31	184	6	207	8

Table 8.6 *Number of iterations for the heat equation problem. The formula (8.52) with* $k = 4$ *is used here.*

Applications to Integral Equations

- Displacement kernel k(s,t)=k(s-t)
- Circulant integral operator
- Discretization schemes (modified prec.)

Accuracy	Rectangular			Trapezoidal			Simpson's		
	B	S	I	B	S	I	B	S	I
10^0	65.1	61.2	492.5	1.49	1.69	5.50	0.50	5.15	5.76
10^{-1}	**	**	**	5.62	6.09	51.28	0.84	9.27	9.89
10^{-2}	**	**	**	16.94	18.35	157.51	1.18	13.44	14.48
10^{-3}	**	**	**	89.48	98.62	930.19	2.68	27.04	29.21
10^{-4}	**	**	**	318.02	332.33	**	5.61	57.49	61.88
10^{-5}	**	**	**	**	**	**	11.64	109.98	116.66
10^{-6}	**	**	**	**	**	**	28.18	270.56	285.09

Table 12.2 *Number of megaflops for different quadrature rules and preconditioners.*

Boundary Integral Equations

$$g(x) = -\frac{1}{2\pi}\int_{\partial\Omega} \log|x-y|\sigma(y)\,dS_y + \eta, \qquad x \in \partial\Omega.$$

$$(\mathcal{B}u)(\phi) = \int_0^{2\pi} b(\theta-\phi)u(\theta)\,d\theta, \qquad 0 \le \phi \le 2\pi$$

with 2π-periodic kernel function $b(\phi)$. The *optimal circulant integral operator* for $\mathcal{A}$ is the unique circulant integral operator $\mathcal{C}$ that minimizes the Hilbert–Schmidt norm $|||\mathcal{B}-\mathcal{A}|||$ over all circulant integral operators $\mathcal{B}$, where

$$|||\mathcal{B}-\mathcal{A}|||^2 \equiv \int_0^{2\pi}\int_0^{2\pi} |a(\theta,\phi) - b(\theta-\phi)|^2\,d\theta\,d\phi,$$

Theorem 12.12. (Spectra of the Preconditioned Operators) *There exist positive constants $\gamma_2 \ge \gamma_1 > 0$ such that the spectrum of $\mathcal{C}^{-1}\mathcal{A}$ lies in $[\gamma_1, \gamma_2]$.*

Applications to Queueing Networks

- The previous talk given by W. K. Ching

Applications to Signal Processing

- Linear prediction filter
- Circulant preconditioners can be applied
- Probabilistic convergence result

Theorem 10.1. (Clustered Spectra of Preconditioned Matrices) *(Ng and Chan, 1994) Let the discrete-time process satisfy the above assumptions. Then for any given* $\epsilon > 0$ *and* $0 < \eta < 1$*, there exist positive integers* ρ_1 *and* ρ_2 *such that for* $n > \rho_1$*, the probability that at most* ρ_2 *eigenvalues of the matrix* $I - c(\tilde{A})^{-1}(A^*A)$ *have absolute value greater than* ϵ*, is greater than* $1-\eta$*, provided that* $m = O(n^{3+\nu})$ *with* $\nu > 0$*.*

Applications to Image Processing

- Deconvolution problem
- Point spread functions ←→ Toeplitz matrices subject to boundary conditions

$$
\begin{bmatrix}
h_m & \cdots & h_0 & \cdots & h_{-m} & & & & \\
 & h_m & & h_0 & & h_{-m} & & 0 & \\
 & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\
 & & & \ddots & \ddots & \ddots & \ddots & \ddots & \\
 & 0 & & & h_m & & h_0 & & h_{-m} \\
 & & & & & h_m & \cdots & h_0 & \cdots & h_{-m}
\end{bmatrix}
\begin{bmatrix}
f_{-m+1} \\ f_{-m+2} \\ \vdots \\ f_0 \\ f_1 \\ \vdots \\ f_n \\ f_{n+1} \\ \vdots \\ f_{n+m-1} \\ f_{n+m}
\end{bmatrix}
=
\begin{bmatrix}
g_1 \\ \vdots \\ g_n
\end{bmatrix}
$$

Deconvolution Problems

- Regularization
- Very ill-conditioned Toeplitz matrices
- Direct inversion → noises amplification
- Many possible solutions
- Regularization restricts the set of admissible solutions
- Tikhonov regularization: L_2 or H_1 norm

Deconvolution Problems

- Periodic boundary condition
- Zero boundary condition
- Reflective boundary condition

$$\min_{\mathbf{f}(\mu)} \left\{ \mu \| D\mathbf{f}(\mu) \|_2^2 + \| \mathbf{g} - A\mathbf{f}(\mu) \|_2^2 \right\}$$

$$(\mu D^t D + A^t A)\mathbf{f}(\mu) = A^t \mathbf{g}.$$

Example

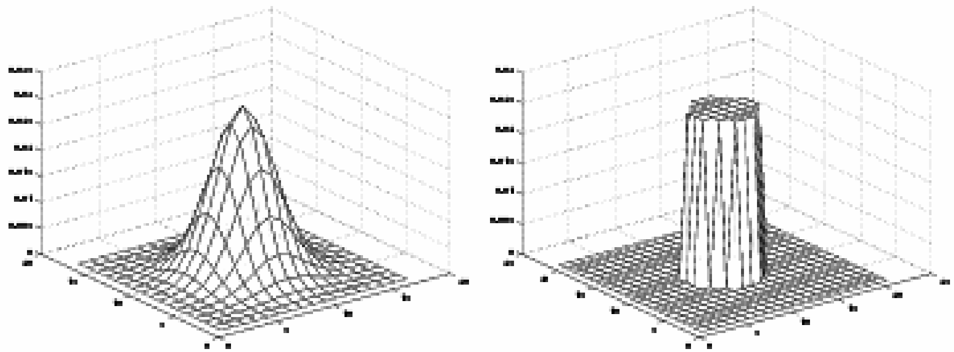


FIG. 11.2. Gaussian (atmospheric turbulence) blur (left) and out-of-focus blur (right).

FIG. 11.3. Noisy and blurred image by Gaussian (left) and out-of-focus blur (right).

Example

rel. error $= 1.24 \times 10^{-1}$ rel. error $= 1.15 \times 10^{-1}$ rel. error $= 6.59 \times 10^{-2}$

FIG. 11.4. Restoring Gaussian blur with zero boundary (left), periodic boundary (middle) and Neumann boundary (right) conditions.

Example

rel. error $= 1.20 \times 10^{-1}$ rel. error $= 1.09 \times 10^{-1}$ rel. error $= 4.00 \times 10^{-2}$

FIG. 11.5. Restoring out-of-focus blur with zero boundary (left), periodic boundary (middle) and Neumann boundary (right) conditions.

Image Restoration Problems

- Other deblurring matrices: spatial variant matrices
- Other measures in the fitting term: L1 norm (non-Gaussian noises)
- Other regularization methods: TV norm, edge-preserving methods (convex, nonconvex), Lipschitz regularization methods
- Other constraints: nonnegativity

Data-fitting term

- Data-fitting term is L1 norm
- $|| A f - g ||_1$ + regularization
- Non-Guassian noises
- Nonlinear problems
- Nonsmooth
- nonnegativity

Spatial-variant Matrices

- Example: Superresolution imaging
- Several low-resolution images
- Downsampling, missing pixels, motions, zooming, etc
- Transformed based preconditioners are not effective

TV-norm

$$\min_f F(f) = \min_u \frac{1}{2}\|\mathcal{H}f - g\|^2_{L^2(\Omega)} + \alpha \int_\Omega |\nabla f| \; dx \; dy$$

$$G(f) \equiv \mathcal{H}^*(\mathcal{H}f - g) - \alpha\nabla \cdot \left(\frac{\nabla f}{|\nabla f|}\right) = 0, \quad (x, y) \in \Omega,$$

$$\frac{\partial f}{\partial n} = 0, \quad (x, y) \in \partial\Omega$$

$$\kappa_\beta(f) = \frac{1}{\sqrt{|\nabla f|^2 + \beta}}, \quad \mathcal{L}_f v = -\nabla \cdot (\kappa_\beta(f)\nabla v)$$

$$\mathcal{A}_f v \equiv (\mathcal{H}^*\mathcal{H} + \alpha\mathcal{L}_f)v,$$

(a)

(b)

(c)

(d)

Figure 4: Multi-frame with a single blur. 4(a)-4(c) Observed image frames with blurring $\sigma = 0.8$. 4(d) Restored image with $\lambda = 0.0004$ reconstructed in 42.92 seconds, relative error = 0.0874 and PSNR = 24.06 dB.

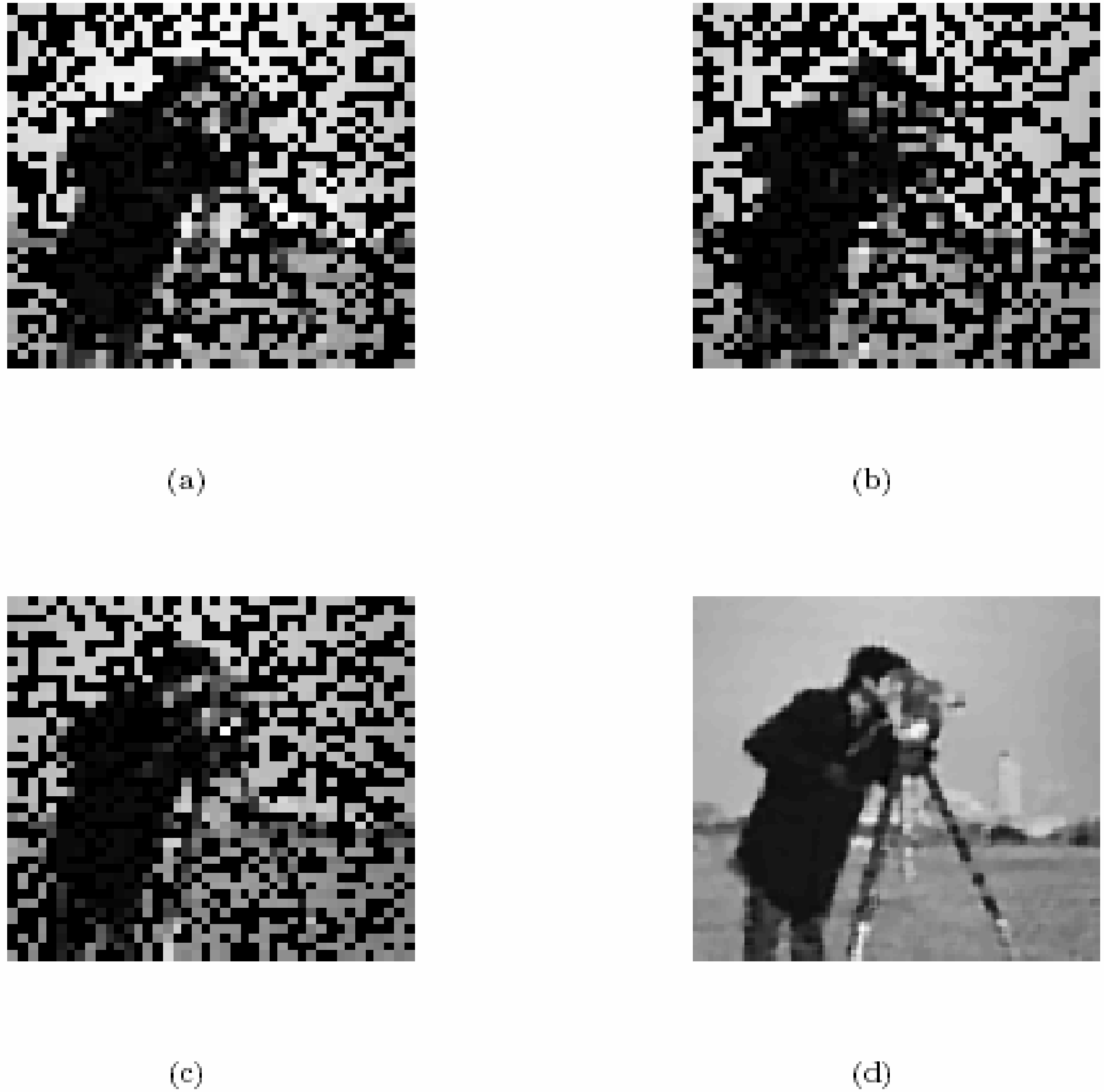


Figure 7: Random missing pixels without blur. 7(a)-7(c) Observed image frames. 7(d) Restored image with $\lambda = 0.0052$ reconstructed in 31.22 seconds, relative error = 0.110 and PSNR = 22.91 dB.

Results

Method	Our method (s)	Artificial time marching scheme (s)
Multi-frame, no blur	21.67	696.64
Multi-frame, single blur	42.92	950.55
Multi-frame, multi-blur	95.98	1870.00
Single-frame, multi-blur	110.47	1826.40

Table 3: Comparison with the artificial time marching scheme.

Exact TV Computation

- Chambolle (JMIV 2004) studied the discrete version of the exact total variation denoising model
- Chambolle has shown that the solution can be given by using the orthogonal projection of the noisy image onto the convex set with the magnitude of the discrete divergence being less than or equal to 1.
- Constrained minimization problems → projected gradient method (slow convergence)
- Semismooth Newton method (fast convergence)

Exact TV with Blur

$$\min_u \frac{1}{2\lambda} \|g - Au\|_X^2 + J(u)$$

is now changed in

$$\min_{u,w} F(u,w) = \min_{u,w} \{H(u,w) + J(u)\}$$

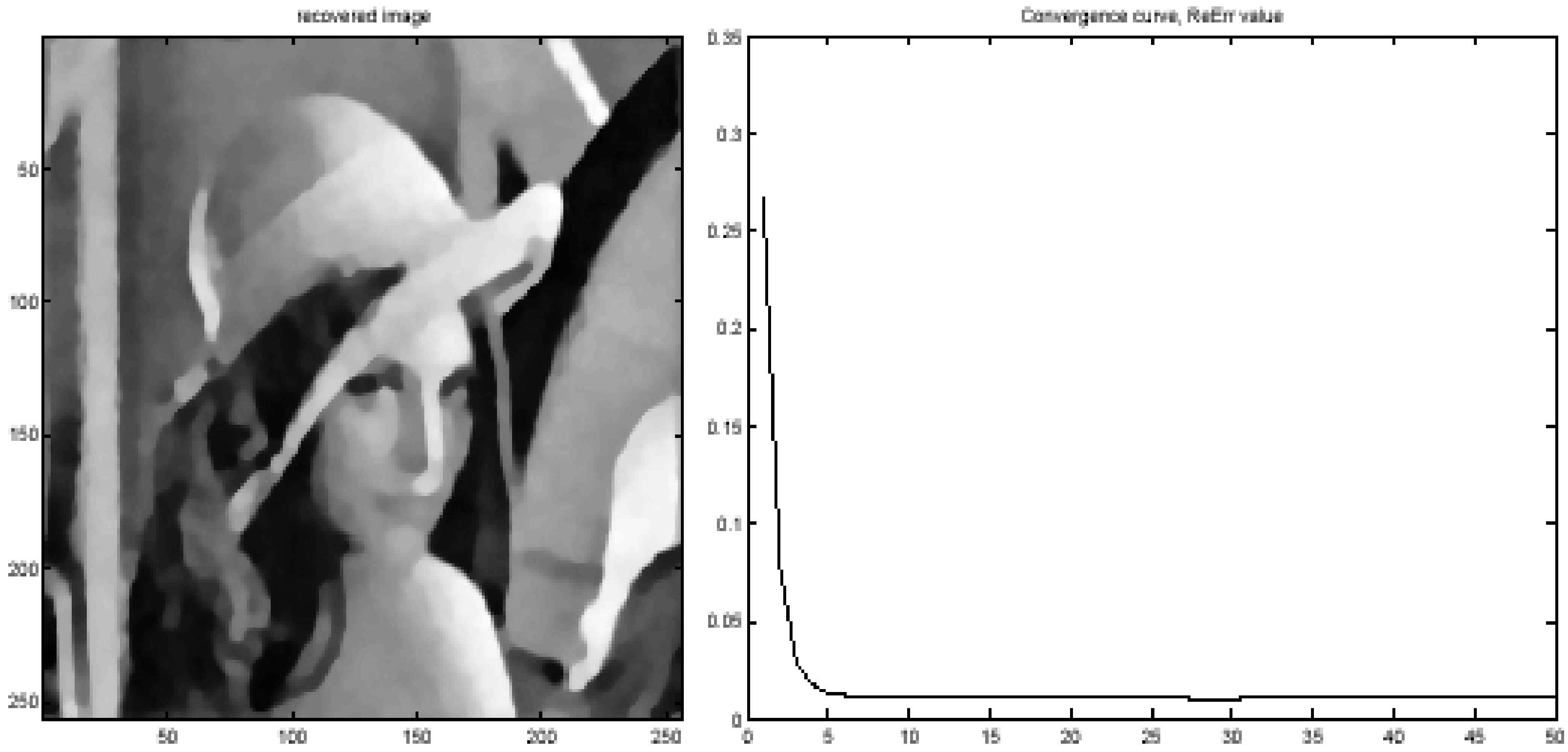
$$H(u,w) = \frac{1}{2\lambda\mu} \left(\|u - w\|^2 + \langle Cw, w \rangle \right) + \frac{1}{2\lambda} \left(\|g\|^2 - 2\langle Au, g \rangle \right)$$

Exact TV with Blur

- Alternating minimization algorithm
- Preconditioning techniques

$$w_n = (I - \mu A^* A)(u_n)$$
$$u_{n+1} = (I - \Pi_{\lambda\mu K_J})(w_n + \mu A^* g)$$

Example



Other Regularization Methods

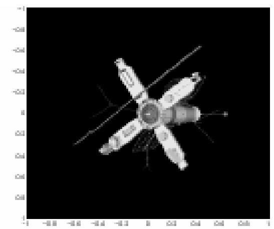


Figure 1: Original satellite image.

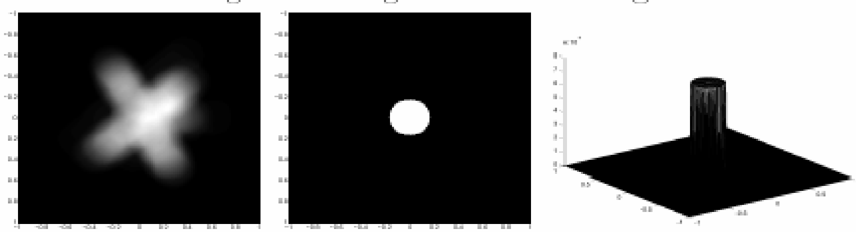


Figure 2: Degraded image(left), image relative error=0.6622, out of focus PSF(middle) and 3D graph of PSF(right).

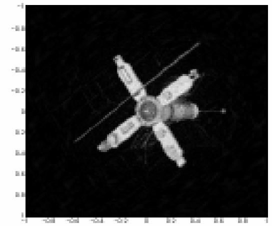


Figure 3: Recovered image using Lipschitz regularization method, $\alpha = 0.417$, $C_t = 0.99$, image relative error=0.2064, ISNR=10.1248

Blind Deconvolution Example

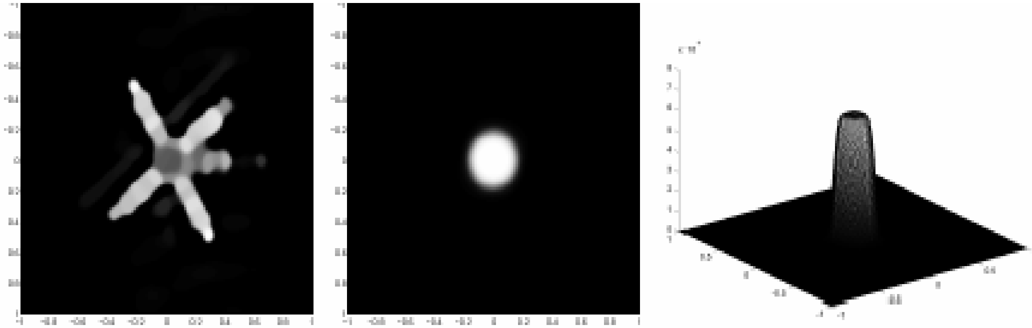


Figure 8: Recovered PSF(left), recovered image(middle) and 3D graph of recovered PSF, when number of AIM iteration=2, $\alpha_1 = 10^{-5}$ and $\alpha_2 = 5.10 \times 10^{-3}$,image relative error=0.3601, ISNR=5.2903, PSF relative error=0.2841, TV norm is used in both image and PSF.

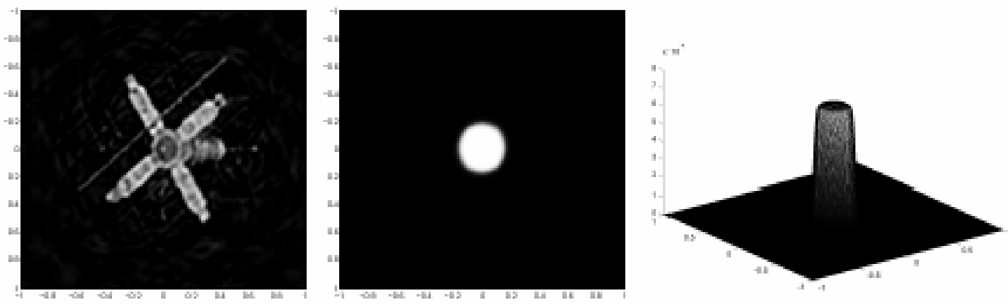


Figure 4: Recovered image(left), recovered PSF (middle) and 3D graph of recovered PSF, when number of AIM iteration=2, $\alpha_1 = 1.2 \times 10^{-6}$ and $\alpha_2 = 6.40 \times 10^{-3}$, image relative error=0.3545, ISNR=5.4265, PSF relative error=0.2055, Lipschitz regularization is used in image and TV norm is used in PSF.

Research Problems

For symmetric positive definite Toeplitz matrices, the spectra of the circulant preconditioned Toeplitz matrices are shown to be clustered. It is clear how this affects the convergence rate of the preconditioned conjugate gradient method. However, for the indefinite or non-Hermitian case, it is not clear how the clustered eigenvalues affect the convergence rate of the Krylov space methods.

Besides trigonometric transforms, wavelet transforms recently have many applications in signal and image processing. Lin *et al.* (2003) have studied wavelet transforms for Toeplitz matrices. The advantage is that the decay of the transformed entries is faster than that of the original entries. The wavelet-transformed Toeplitz matrix is no longer a Toeplitz matrix, but it still retains some displacement structure. A good question is how to design more effective preconditioners for Toeplitz matrices based on wavelet transforms.

Research Problems

- In the literature, the order of the zeros of generating functions is usually assumed to be even. What is the performance of the best circulant preconditioners when the order of the zeros is odd, or not an integer. If the preconditioners work, can we show that the spectra of these circulant preconditioned matrices are clustered around 1 or that the condition numbers of these circulant preconditioned matrices are uniformly bounded? Otherwise, can we develop other iterative methods or preconditioners for Toeplitz matrices that are generated by functions with zeros being odd or not integral? In R. Chan *et al.* (1998) some initial results show that multigrid methods may be a promising approach for these Toeplitz matrices.
- What happens if generation functions have poles instead of zeros? Can we apply the established results to this case? It is interesting to note that there are some applications in time series modelling where generating functions have poles (Lu, 2003).
- Little attention has been given to block-Toeplitz matrices that are generated by matrix-valued functions. Some results can be found in Serra (1999c). He has studied the preconditioners based on matrix-valued functions.

Research Problems

- Circulant preconditioners can be applied effectively and efficiently to solving Toeplitz least squares problems if Toeplitz matrices have full rank. How can we handle the case when Toeplitz matrices do not have full rank? One possibility is to consider the generalized inverses of Toeplitz matrices. In the literature, computing the inverses and the generalized inverses of structured matrices are important practical computational problems; see for instance Pan and Rami (2001), Bini *et al.* (2003), and Wei *et al.* (2004).
- When Toeplitz matrices have full rank, Toeplitz least squares problems $\min \|T\mathbf{x} - \mathbf{b}\|_2^2$ are equivalent to solving the normal equation matrices $T^*T\mathbf{x} = \mathbf{T}^*\mathbf{b}$. However, when we consider 1-norms instead of 2-norms in the computation we do not have this equivalence and cannot apply circulant preconditioners straightforwardly to the problem. It is still an open question how to solve $\min \|T\mathbf{x} - \mathbf{b}\|_1$ efficiently.
- It is interesting to find good preconditioners for Toeplitz-related systems with large displacement rank. Good examples are Toeplitz-plus-band systems studied in Section 7.2. Direct Toeplitz-like solvers cannot be employed because of the large displacement rank. However, iterative methods

Research Problems

are attractive since coefficient matrix–vector products can be computed efficiently at each iteration. For instance, for the Toeplitz-plus-band matrix, its matrix–vector product can be computed in $O(n \log n)$ operations. The main concern is how to design good preconditioners for such Toeplitz-related systems with large displacement rank. Recently, Lin *et al.* (2003) proposed and developed factorized banded inverse preconditioners for matrices with Toeplitz structure. Also Lin *et al.* (2004) studied incomplete factorization-based preconditioners for Toeplitz-like systems with large displacement ranks in image processing.

- Other interesting areas are to design efficient algorithms based on preconditioning techniques for finding eigenvalues and singular values of Toeplitz-like matrices. Ng (2000) has employed preconditioned Lanczos methods for the minimum eigenvalue of a symmetric positive definite Toeplitz matrix.

Thank you very much !

Q/A