# Preconditioning Block Toeplitz Matrices with small blocksize

Thomas Huckle, TUM

Cortona, September 2004

# Outline:

1. Iterative solvers for spd Toeplitz Matrices
2. Block Toeplitz Matrices and Generating Functions
3. Theoretical porperties of spd BT matrices
4. Special cases
5. Some numerical results
6. Problems of iterative methods for BT
7. Multigrid for BT matrices

# 1. Iterative Solvers for Toeplitz Matrices

$$T_n = \begin{pmatrix} t_0 & t_{-1} & . & . & t_{1-n} \\ t_1 & t_0 & t_{-1} & . & . \\ . & . & . & . & . \\ . & . & t_1 & t_0 & t_{-1} \\ t_{n-1} & . & . & t_1 & t_0 \end{pmatrix} = \left(t_{i-j}\right)_{i,j=1}^{n}$$

Assumption: $T_n$ , n=1,2,... , family of matrices

entries $t_j$ independent of n, decreasing.

$$f(x) = \cdots + t_{-2}e^{-2ix} + t_{-1}e^{-ix} + t_0 + t_1 e^{ix} + t_2 e^{2ix} + \cdots$$

$$t_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) e^{-ikx} dx$$

$$T_n(f) = toeplitz(t_{1-n}, \ldots, t_0, \ldots, t_{n-1})$$

# Example:

Toeplitz(0,...,0,-1,2,-1,0,...,0) = tridiag(-1,2,-1)

Generating function:

$$f(x) = -e^{ix} + 2 - e^{-ix} =$$
$$= 2 - 2 \cdot \cos(x) = 2\big(1 - \cos(x)\big)$$

**Properties of $T_n$ ←→ properties of f(x)** *(Szegö)*

**Properties of inv($P_n$)$T_n$ ←→ properties of f(x)/p(x)** *(Serra)*

$$range(T_n(f)) \subseteq range(f(x))$$

$$range(T_n^{-1}(g) \cdot T_n(f)) \subseteq range(f / g)$$

For banded Toeplitz matrix P = T(p) with trigonometric polynomial p(x)) and f / p continuous:

$$T_n^{-1}(p) \cdot T_n(f) = T_n(f / p) + low\_rank$$

Clustering for circulant preconditioner:

$$C_n^{-1}(f) \cdot T_n(f) = I + low\_rank + small\_norm$$

# 2. Block Toeplitz Matrix

$$A = \begin{pmatrix} T_0 & T_{-1} & \cdots & T_{1-n} \\ T_1 & T_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & T_{-1} \\ T_{n-1} & \cdots & T_1 & T_0 \end{pmatrix}$$

$T_j$ general m x m –Matrices. *(We will only consider m=2)*

Generating Matrix Function:

$$F(x) = \cdots + T_{-2}e^{-2ix} + T_{-1}e^{-ix} + T_0 + T_1e^{ix} + T_2e^{2ix} + \cdots$$

# Example

$$A = \begin{pmatrix}
2 & -1 & 0 & 0 & & & & \\
-1 & 3 & 1 & 0 & & & & \\
0 & 1 & 2 & -1 & & & & \\
0 & 0 & -1 & 3 & 1 & & & \\
 & & & 1 & \ddots & \ddots & & \\
 & & & & \ddots & \ddots & 1 & \\
 & & & & & 1 & 2 & -1 \\
 & & & & & & -1 & 3
\end{pmatrix},$$

$$T_0 = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}, \quad T_{-1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad T_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Generating Function: $F(x) = \begin{pmatrix} 2 & -1+e^{ix} \\ -1+e^{-ix} & 3 \end{pmatrix};$

$$A_n \Leftrightarrow \quad \tilde{A}_n = \left(\begin{array}{cccc|cccc} 2 & & & & -1 & & & \\ & 2 & & & 1 & -1 & & \\ & & \ddots & & & \ddots & \ddots & \\ & & & 2 & & & 1 & -1 \\ \hline -1 & & & & 3 & & & \\ 1 & -1 & & & & 3 & & \\ & \ddots & \ddots & & & & \ddots & \\ & & 1 & -1 & & & & 3 \end{array}\right)$$

Ill-posed Example:

$$F(x) = \begin{pmatrix} 1 & |x| \\ |x| & x^2 \end{pmatrix} \qquad T_n(F) = \begin{pmatrix} I & T_n(|x|) \\ T_n(|x|) & T_n(x^2) \end{pmatrix}$$

Eigenvalues of F(x):

0 and $1+x^2$

Eigenvalues of $T_n(F)$ in

$$\left[0, \quad 1+x^2\right]$$

with cluster at 0

# 3. Theoretical Properties

$$[\lambda_1, \lambda_n] = range(T_n(F)) \subseteq range(F(x)) = [\min \lambda(F), \max \lambda(F)]$$

$$range(T_n^{-1}(G) \cdot T_n(F)) \subseteq range(G^{-1} \cdot F)$$

Serra, Tilli,..

$$T_n^{-1}(P) \cdot T_n(F) = T_n(P^{-1} \cdot F) + low\_rank$$ for P banded

Circulant preconditioner:

$$\begin{pmatrix} C_1 & C_2 \\ C_2^H & C_3 \end{pmatrix} \approx \begin{pmatrix} T_1 & T_2 \\ T_2^H & T_3 \end{pmatrix}$$ with $C_i \approx T_i$

# Theorem 1: $T_n(f)^2 \leftrightarrow T_n(f^2)$ (f scalar and real)

1. $T_n(f^2) \geq T_n(f)^2$ :

Consider $\begin{pmatrix} T_n(f^2) & T_n(f) \\ T_n(f) & I_n \end{pmatrix} \leftrightarrow \begin{pmatrix} f^2 & f \\ f & 1 \end{pmatrix} \geq 0$

Therefore $T_n(f^2) - T_n^2(f) \geq 0$

# Clustering of $T_n(f)^{-2} \cdot T_n(f^2)$ :

$C_n(f)$ circulant preconditioner for $T_n(f)$ :

Clustering of eigenvalues and singular values by appropriate $C_n(f)$:

$$\mathrm{inv}(C_n(f))\, T_n(f) = I + R + E$$

$$\mathrm{inv}(C_n(f))^2\, T_n(f^2) = I + R + E$$

Therefore

$$T_n^{-2}(f) \cdot T_n(f^2) \Leftrightarrow \left(C \cdot T^{-2}(f) \cdot C\right) \cdot \left(C^{-1} \cdot T(f^2) \cdot C^{-1}\right) =$$
$$(I + R_1 + E_1) \cdot (I + R_2 + E_2) = I + R + E$$

# Regularized Least Squares:

Clustering of $T_n(f^2+\rho)^{-1} \cdot (T_n(f)^2+\rho \cdot I)$ :

$$T_n(f^2+\rho) \geq T_n^2(f) + \rho I$$ like before.

$$T(f^2+\rho)^{-1} \cdot \left(T^2(f)+\rho I\right) =$$
$$I + \left(T(f^2)+\rho I\right)^{-1} \cdot \left(T^2(f) - T(f^2)\right) =$$
$$I + \left(I + \rho T^{-1}(f^2)\right)^{-1} \cdot T^{-1}(f^2) \cdot \left(T^2(f) - T(f^2)\right) =$$
$$I - \underbrace{\left(I + \rho T^{-1}(f^2)\right)^{-1}}_{\leq 1} \cdot \underbrace{\left(I - T^{-1}(f^2) \cdot T^2(f)\right)}_{\text{clustered around zero}}$$

$\leq 1$ , clustered around zero

# Schur Complement for spd problems:

$$0 \le T_n \begin{pmatrix} |b|^2/g & b \\ \bar{b} & g \end{pmatrix} = \begin{pmatrix} T_n(|b|^2/g) & T_n(b) \\ T_n^H(b) & T_n(g) \end{pmatrix} \Leftrightarrow$$

$$\begin{pmatrix} |b|^2/g & b \\ \bar{b} & g \end{pmatrix} \in \left[0, g + |b|^2/g\right]$$

Therefore, it holds:

$$T_n(|b|^2/g) - T_n^H(b)T_n^{-1}(g)T_n(b) \ge 0,$$

$$T_n(g) - T_n(b)T_n^{-1}(|b|^2/g)T_n^H(b) \ge 0$$

# Preconditioning Schur-Compl.:

$$T_n^{-1}\left(f-\frac{|b|^2}{g}\right)\cdot\left(T_n(f)-T_n^H(b)T_n^{-1}(g)T_n(b)\right)=?$$

It holds $$T_n\left(f-\frac{|b|^2}{g}\right)\le T_n(f)-T_n^H(b)T_n^{-1}(g)T_n(b)$$

and therefore the eigenvalues of the preconditioned Schur Complement are $\geq 1$ (see also Ching,Ng,Wen)

Clustering only for well-conditioned problems!
No clustering for ill-conditioned problems!

# BT systems and Schur Complements

Block-Cholesky:

$$\begin{pmatrix} T_1 & T_2 \\ T_2^H & T_3 \end{pmatrix} = \begin{pmatrix} I & T_2 T_3^{-1} \\ 0 & I \end{pmatrix} \cdot \begin{pmatrix} T_1 - T_2 T_3^{-1} T_2^H & 0 \\ 0 & T_3 \end{pmatrix} \cdot \begin{pmatrix} I & 0 \\ T_3^{-1} T_2^H & I \end{pmatrix}$$

$$= \begin{pmatrix} I & 0 \\ T_2^H T_1^{-1} & I \end{pmatrix} \cdot \begin{pmatrix} T_1 & 0 \\ 0 & T_3 - T_2^H T_1^{-1} T_2 \end{pmatrix} \cdot \begin{pmatrix} I & T_1^{-1} T_2 \\ 0 & I \end{pmatrix}$$

Reduce original LS to solving two systems, one in a Toeplitz matrix and the other in a Schur Complement matrix in Toeplitz matrices.

(see also Ching,Ng,Wen)

# The Inverse of a BT Matrix

$$\begin{pmatrix} T_1 & T_2 \\ T_2^H & T_3 \end{pmatrix}^{-1} = \begin{pmatrix} T_1^{-1} & 0 \\ 0 & T_3^{-1} \end{pmatrix} \cdot \begin{pmatrix} T_1 & -T_2 \\ -T_2^H & T_3 \end{pmatrix} \cdot \begin{pmatrix} S_1^{-1} & 0 \\ 0 & S_3^{-1} \end{pmatrix}$$

with $S_1 = T_1 - T_2 T_3^{-1} T_2^H$

$$S_3 = T_3 - T_2^H T_1^{-1} T_2$$

# Band Preconditioners

$$\begin{pmatrix} P_1 & P_2 \\ P_2^H & P_3 \end{pmatrix} \approx \begin{pmatrix} T_1 & T_2 \\ T_2^H & T_3 \end{pmatrix}$$

with trigonometric polynomials $p_i$ .

e.g. using the Taylor expansion of f , g , and b .

# 4. Special Cases

$$\begin{pmatrix} T_1 & T_2 \\ T_2 & T_1 \end{pmatrix}: \qquad \begin{pmatrix} I & I \\ I & -I \end{pmatrix} \cdot \begin{pmatrix} T_1 & T_2 \\ T_2 & T_1 \end{pmatrix} \cdot \begin{pmatrix} I & I \\ I & -I \end{pmatrix} = 2\begin{pmatrix} T_1 + T_2 & 0 \\ 0 & T_1 - T_2 \end{pmatrix}$$

$$\begin{pmatrix} T_1 & T_3 \\ T_3^H & T_3 \end{pmatrix}: \qquad S = T_1 - T_3^H T_3^{-1} T_3 \; = T_1 - T_3^H$$

$$\begin{pmatrix} T_1 & P \\ P^H & Q \end{pmatrix}$$ with banded P and Q: $S = T_1 - P^H Q^{-1} P = T_1 - T(|\,p\,|^2 / q)$

$$\begin{pmatrix} T_1 & T_2 \\ T_2^H & T_3 \end{pmatrix}$$ with well-conditioned Schur compl. $S = T_3 - T_2^H T_1^{-1} T_2$

e.g. $\begin{pmatrix} 1+|\,t\,| & t \\ t & |\,t\,| \end{pmatrix}$ with $S = T(1+|\,t\,|) - T(t)T(|\,t\,|)^{-1}T(t) \approx I$

# 5. Numerical Examples

$$F_1(t) = \begin{pmatrix} 1+(t/\pi)^2 & t/\pi \\ t/\pi & (t/\pi)^2 \end{pmatrix}$$

$$S_1 = T_1 - T_2 T_3^{-1} T_2^H \rightarrow 1+t^2 - t\cdot t^{-2}\cdot t \rightarrow T(t^2)$$

$$S_3 = T_3 - T_2^H T_1^{-1} T_2 \rightarrow t^2 - t\cdot(1+t^2)^{-1}\cdot t \rightarrow T(t^4)$$

$$\rightarrow T(t)\cdot T(1+t^2)^{-1}\cdot T(t^2)\cdot T(t)$$

Eigenvalues $t^4$ and $O(1)$

---

$$F_2(t) = \begin{pmatrix} 1+(t/\pi)^2 & |t/\pi| \\ |t/\pi| & (t/\pi)^2 \end{pmatrix}$$

# Numerical Examples

## Block Circulant Preconditioner

$C_F$: Block Frobeniusnorm approximation

$C_S$: Blockwise Strang circulant

$C_D$: Blockwise $F_n^H \, diag\left( f(\frac{2\pi j}{n})_{j=1,\ldots,n/2} \right) F_n$

# $F_2(t)$:

| n | 2*20 | 2*40 | 2*80 | 2*160 | 2*320 | 2*640 |
|---|---|---|---|---|---|---|
| $C_F$: | 21 | 33 | 56 | * | * | * |
| $C_D$: | 32 | 52 | 97 | * | * | * |
| $C_S$: | 14 | 20 | 23 | 33 | 50 | 94 |

# $F_1(t)$:

| n | 2*20 | 2*40 | 2*80 | 2*160 | 2*320 | 2*640 |
|---|---|---|---|---|---|---|
| $C_F$: | 28 | 44 | 71 | * | * | * |
| $C_D$: | 39 | 60 | * | * | * | * |
| $C_S$: | 15 | 18 | 24 | 30 | 34 | 51 |

# Compare Toeplitz case $T(t^2)$:

| n | 20 | 40 | 80 | 160 |
|---|---|---|---|---|
| $C_F$: | 11 | 14 | 16 | 20 |
| $C_D$: | 12 | 13 | 13 | 13 |
| $C_S$: | 7 | 7 | 7 | 7 |

# Numerical Examples

## Banded Preconditioner: $F_1(t)$ and $F_2(t)$:

$$P(t) = \begin{pmatrix} 1 + \frac{2}{\pi^2}(1 - \cos(t)) & \sin(t)/\pi \\ \sin(t)/\pi & \frac{2}{\pi^2}(1 - \cos(t) \end{pmatrix}$$

| n | 2*20 | 2*40 | 2*80 | 2*160 | 2*320 | 2*640 |
|---|---|---|---|---|---|---|
| $F_1$: | 20 | 22 | 24 | 25 | 25 | 25 |
| $F_2$: | 64 | 178 | * | * | * | * |

# Numerical Examples

## Schur Complement Preconditioner

Replace in the Cholesky factorization the Schur complement

by a Toeplitz matrix with the corresponding generating function:

$$\begin{pmatrix} T_1 & T_2 \\ T_2^H & T_3 \end{pmatrix} = \begin{pmatrix} I & 0 \\ T_2^H T_1^{-1} & I \end{pmatrix} \cdot \begin{pmatrix} T_1 & 0 \\ 0 & S_3 \end{pmatrix} \cdot \begin{pmatrix} I & T_1^{-1} T_2 \\ 0 & I \end{pmatrix}$$

$$S_3 = T_3 - T_2^H T_1^{-1} T_2 = T(f) - T(b)^H T(g)^{-1} T(b) \to T(f - \frac{|b|^2}{g})$$

# Iteration number for $F_1$ and $F_2$ and right or left Cholesky

$F_1$:

| n | 2*20 | 2*40 | 2*80 | 2*160 | 2*320 | 2*640 |
|---|---|---|---|---|---|---|
| left: | 7 | 8 | 10 | 11 | 11 | 17 |
| right: | 7 | 7 | 7 | 9 | 10 | 10 |

$F_2$:

| n | 2*20 | 2*40 | 2*80 | 2*160 | 2*320 | 2*640 |
|---|---|---|---|---|---|---|
| left: | 7 | 9 | 11 | 15 | 20 | 28 |
| right: | 7 | 9 | 11 | 15 | 19 | 26 |

# 6. Problems with iterative methods for BT

$$T\left(f-\frac{|b|^2}{g}\right)^{-1}\cdot\left(T(f)-T(b)T^{-1}(g)T^H(b)\right)=$$

$$=T\left(f-\frac{|b|^2}{g}\right)^{-1}\cdot\left(T(f-\frac{|b|^2}{g})+\left(T(\frac{|b|^2}{g})-T(b)T^{-1}(g)T^H(b)\right)\right)=$$

$$=I+T\left(f-\frac{|b|^2}{g}\right)\cdot\left(T\left(\frac{|b|^2}{g}\right)-T(b)T^{-1}(g)T^H(b)\right)=$$

$$=I \quad + \quad S \quad \cdot \quad (E \quad + \quad R)$$

= I + ill-cond. * (small norm + low rank)

# 7. Multigrid for BT:

$$E_0 = \begin{pmatrix} 1 & & \\ 2 & & \\ 1 & 1 & \\ & 2 & \\ & 1 & . \\ & & . \end{pmatrix} : \qquad \begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix}^T \begin{pmatrix} T_1 & T_2 \\ T_2^H & T_3 \end{pmatrix} \begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix}$$

| $F_2$: n | 2*20 | 2*40 | 2*80 | 2*160 | 2*320 | 2*640 |
|---|---|---|---|---|---|---|
| Block-Jacobi: | 19 | 21 | 23 | 24 | 24 | 24 |
| Block-GS: | 14 | 14 | 14 | 14 | 14 | 14 |