Preconditioners for ill conditioned (block) Toeplitz systems: facts and ideas

Paris Vassalos

Department of Informatics, Athens University of Economics and Business, Athens, Greece. Email: pvassal@aueb.gr, pvassal@uoi.gr

Joint work with
D. Noutsos
We are interested in the fast and efficient solution of $nm \times nm$ BTTB systems

$$T_{n,m}(f)x = b$$

where $f$ is nonnegative real-valued belonging to $C_{2\pi,2\pi}$ defined in the fundamental domain $Q = (-\pi, \pi]^2$ and is a priori known. The entries of the coefficient matrix are given by

$$t_{j,k} = \frac{1}{4\pi^2} \int_{Q} f(x, y) e^{-i(jx + ky)} dxdy,$$

for $j = 0, \pm 1, \ldots, \pm(n - 1)$ and $k = 0, \ldots, \pm(m - 1)$. 
Connection between $f$ and $T_{nm}(f)$

The main connection between $T_{nm}(f)$ and the generating function is described by the following result:

**Theorem**

If $f \in C[-\pi, \pi]^2$ and $c < C$ the extreme values of $f(x, y)$ on $Q$. Then every eigenvalue $\lambda$ of the block Toeplitz matrix $T$ satisfies the strict inequalities

$$c < \lambda < C$$

Moreover as $m, n \to \infty$ then

$$\lambda_{\min}(T_{nm}(f)) \to c \quad \text{and} \quad \lambda_{\max}(T_{nm}(f)) \to C$$
The basis for the construction of effective preconditioners is described by the following theorem:

**Theorem**

Let $f, g \geq 0 \in C[-\pi, \pi]^2$ (f and g not identically zero). Then for every $m, n$ the matrix $T_{nm}^{-1}(g)T_{nm}(f)$ has eigenvalues in the open interval $(r, R)$, where

$$r = \inf_Q \frac{f}{g} \quad \text{and} \quad R = \sup_Q \frac{f}{g}$$
Negative result for $\tau$ algebra preconditioners

Theorem (Noutsos, Serra, Vassalos, TCS (2004))

Let $f$ be equivalent to

$$p_k(x, y) = (2 - 2 \cos(x))^k + (2 - 2 \cos(y))^k \text{ with } k \geq 2 \text{ and let } \beta \text{ be a fixed positive number independent of } n.$$ 

Then for every sequence $\{P_n\}$ with $P_n \in \tau$, $n = (n_1, n_2)$, and such that

$$\lambda_{\text{max}}(P_n^{-1} T_n(f)) \leq \beta$$

uniformly with respect to $\hat{n}$, we have

(a) the minimal eigenvalue of $P_n^{-1} T_n(f)$ tends to zero.

(b) the number

$$\#\{\lambda(n) \in \sigma(P_n^{-1} T_n(f)) : \lambda(n) \to_{N(n) \to \infty} 0\}$$

tends to infinity as $N(\hat{n})$ tends to infinity.
Negative result for $\tau$ algebra preconditioners

Theorem (Noutsos, Serra, Vassalos, TCS (2004))

Let $f$ be equivalent to

$$p_k(x, y) = (2 - 2 \cos(x))^k + (2 - 2 \cos(y))^k$$

with $k \geq 2$ and let $\alpha$ be a fixed positive number independent of $n = (n_1, n_2)$.

Then for every sequence $\{P_n\}$ with $P_n \in \tau$ and such that

$$\lambda_{\min}(P_n^{-1} T_n(f)) \geq \alpha$$

uniformly with respect to $n$, we have

(a) the maximal eigenvalue of $P_n^{-1} T_n(f)$ tends to $\infty$.

(b) the number

$$\#\{\lambda(n) \in \sigma(P_n^{-1} T_n(f)) : \lambda(n) \to N(n) \to \infty \to \infty\}$$

tends to infinity as $N(n)$ tends to infinity.
Direct methods: Levinson type methods cost $O(n^2m^3)$ ops while superfast methods $O(nm^3 \log^2 n)$. Stability problems. Not optimal.

PCG method with matrix algebra preconditioners. Cost $O(k(\epsilon)nm \log nm)$, where $k(\epsilon)$ is the required number of iterations and depends from the condition number of $T_{nm}$.

PCG method where the preconditioner is band Toeplitz matrix. Under some assumptions, the cost is optimal ($O(nm \log nm)$).

S. Serra-Capizzano (BIT, (1994)) and M. Ng (LAA, (1997)) proposed as preconditioner the band BTTB matrix generated by the minimum trigonometric polynomial $g$ which has the same roots with $f$.

Let $f = g \cdot h$ with $h > 0$. Then D. Noutsos, S. Serra Capizzano and P. Vassalos (Numer. Math. (2006)) proposed as preconditioners the band BTTB matrix generated by $g \cdot \hat{h}$ where $\hat{h}$:

1. is the trigonometric polynomial arises from the Fourier approximation on $h$.
2. arises from the Lagrange interpolation of $h$ at 2D Chebyshev points, or from the interpolation of $h$ using the 2D Fejer kernel.
We assume that the nonnegative function $f$ has isolated zeros $(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)$, on $Q$ each one of multiplicities $(2\mu_1, 2\nu_1), (2\mu_2, 2\nu_2), \ldots, (2\mu_k, 2\nu_k)$. Then, $f$ can be written as

$$f = g \cdot w$$

where

$$g = \prod_{i=1}^{k} [(2 - 2 \cos(x - x_i))^{\mu_i} + (2 - 2 \cos(y - y_i))^{\nu_i}]$$

and $w$, is strictly positive on $Q$. 
For the system

\[ T_{nm}(f)x = b \]

we define and propose as a preconditioner the product of matrices

\[ K_{nm}(f) = A_{nm}(\sqrt{w}) T_{nm}(g) A_{nm}(\sqrt{w}) = A_{nm}(h) T_{nm}(g) A_{nm}(h) \]

with \( h = \sqrt{w} \), \( A_{nm} \in \{ \tau, C, H \} \), where \( \{ \tau, C, H \} \) is the set of matrices belonging to Block \( \tau \), Block Circulant and Block Hartley algebra, respectively.
Construction of 2D algebras

<table>
<thead>
<tr>
<th>$A$</th>
<th>$v^{[n]}$</th>
<th>$Q_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$</td>
<td>$v_i^{[n]} = \frac{\pi i}{n+1}$</td>
<td>$\sqrt{\frac{2}{n+1}} \left( \sin(jv_i^{[n]}) \right)_{i,j=1}^n$</td>
</tr>
<tr>
<td>$C$</td>
<td>$v_i^{[n]} = \frac{2\pi i}{n}$</td>
<td>$F_n = \frac{1}{\sqrt{n}} \left( e^{ijv_i^{[n]}} \right)$</td>
</tr>
<tr>
<td>$H$</td>
<td>$v_i^{[n]} = \frac{2\pi i}{n}$</td>
<td>$\text{Re}(F_n) + \text{Im}(F_n)$</td>
</tr>
</tbody>
</table>

The matrices $C(h), \tau(h), H(h)$ can be written as

$$A_{nm}(h) = Q_{nm} \cdot \text{Diag} \left( f(v^{[nm]}) \right) \cdot Q_n^H$$

where

$$v^{[nm]} = v^{[n]} \times v^{[m]} \quad \text{and} \quad Q_{nm} = Q_n \bigotimes Q_m$$
Obviously $K_{nm}(f)$ has all the properties that a preconditioner must satisfy, i.e,

- is symmetric,
- is positive definite,
- The cost for the solution of the arbitrary system

$$K_{nm}(f)x = b$$

is of order $O(nm \log nm)$. $O(nm \log nm)$ for the “inversion” of $A_{nm}(h)$ by 2D FFT, and $O(nm)$ for the “inversion” of block band Toeplitz matrix $T_{n,m}(g)$ which can be done by multigrid methods. So, the only condition that must be fulfill, in order to be a competitive preconditioner, is the spectrum of $[K_{nm}(f)]^{-1} T_{nm}(f)$ being bounded from above and below.
We say that a function $h$ is a $((k_1, k_2), (x_0, y_0))$-smooth function if

$$\frac{\partial^{l_1+l_2}}{\partial x^{l_1} \partial y^{l_2}} h(x_0, y_0) = 0, \quad l_1 < k_1, \quad l_2 < k_2, \quad and \quad l_1+l_2 < \max\{k_1, k_2\}$$

and

$$\frac{\partial^{l_1+l_2}}{\partial x^{l_1} \partial y^{l_2}} h(x_0, y_0)$$

is bounded for $l_1 = k_1, l_2 = 0$ and $l_2 = k_2, l_1 = 0$, and $l_1 + l_2 = \max\{k_1, k_2\}, l_1 < k_1, l_2 < k_2$. 
Case: Clustering

**Theorem**

Let $f$ belongs to the Wiener class. Then for every $\epsilon$ the spectrum of

$$[K^\tau_{nm}(f)]^{-1} T_{nm}(f)$$

lies in $[1 - \epsilon, 1 + \epsilon]$, for $n, m$ sufficient large, except of $O(m + n)$ outliers. Thus, we have weak clustering around unity of the spectrum of the preconditioned matrix.
Case: Bounds

**Theorem**

Let $f \in C_{2\pi, 2\pi}$ even function on $Q$ with roots $(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)$, each one of multiplicities $(2\mu_1, 2\nu_1), (2\mu_2, 2\nu_2), \ldots, (2\mu_k, 2\nu_k)$, respectively. If $g$ is the trigonometric polynomial of minimal degree that rises the roots of $f$ and $h$ is $(\mu_i - 1, \nu_i - 1)$ smooth function at the roots $(x_i, y_i), i = 1(1)k$, then the spectrum of the preconditioned matrix $P^\tau = [K_{nm}(f)]^{-1} T_{nm}(f)$ is bounded from above as well as bellow:

$$c < \lambda_{\min}(P^\tau) < \lambda_{\max}(P^\tau) < C$$

with $c, C$ constants independent of $n, m$.
Let $f$ belong to the Wiener class on $Q$. Then for every $\epsilon$ the spectrum of

$$\left[ K_{nm}^C(f) \right]^{-1} T_{nm}(f)$$

lies in $[1 - \epsilon, 1 + \epsilon]$ for $n, m$ sufficient large except $O(m + n)$ outliers. Thus, we have a weak clustering of the spectrum of the preconditioned matrix around unity.
Theorem

Let \( f \in C_{2\pi,2\pi} \) even function on \( Q \) with roots \((x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)\), on \( Q \) each one of multiplicities \((2\mu_1, 2\nu_1), (2\mu_2, 2\nu_2), \ldots, (2\mu_k, 2\nu_k)\) respectively. If \( g \) is the trigonometric polynomial of minimal degree that rises the roots of \( f \) and \( h \) is \((\mu_i, \nu_i)\) smooth function at the roots \((x_i, y_i)\), \( i = 1(1)k \) then the spectrum of the preconditioned matrix \( P_C = \left[K_{nm}^C(f)\right]^{-1} T_{nm}(f) \) is bounded above as well as below:

\[ c \leq \lambda_{\text{min}}(P^C) < \lambda_{\text{max}}(P^C) < C \]

with \( c, C \) constants independent of \( n, m \)
Case where $h$ is not smooth enough

Let us suppose that $f$ has roots at $(x_i, y_i)$, $i = 1(1)k$. If $h = \frac{f}{g}$ is not as smooth as the previous Theorems demand, then the spectrum of the preconditioned matrix could be unbound. For that case we make a “smoothing correction” and instead of $h$ we use $\hat{h}$ defined as

$$\hat{h} = \begin{cases} 
    h(x) & (x, y) \in Q/\bigcup \Omega_i \\
    h(x_i, y_i) + \alpha_i g_i(x, y) & (x, y) \in \bigcup \Omega_i
\end{cases}$$

where $\Omega_i = \{(x, y) : \| (x_i, y_i) - (x, y) \|_\infty < \epsilon_i \}$ and $\alpha_i$ is defined such that

$$h(\epsilon_i, \epsilon_i) = h(x_i, y_i) + \alpha_i g(\epsilon_i, \epsilon_i)$$
Figure: Smoothing of \( h(x, y) = \frac{(|x|+|y|+1)(x^4+y^4)}{(2-2\cos(x))^2+(2-2\cos(y))^2} \), by \( \hat{h}(x, y) \).
We compare our proposal with the already known band preconditioners:

- $B$, which is generated by the trigonometric polynomial $g$ that raises the roots of $f$
- $K$ which is the product $g \cdot \hat{h}$ with $\hat{h}$ being the trigonometric polynomial arising from the 2D Fejer kernel on the positive part $h$ of $f$
- $P$ and $F$ which arise: from the Lagrange interpolation of $h$ at $2D$ Chebyshev points, and from the approximation of $h$ using the $2D$ Fourier expansion, respectively.

For all the tests the stopping criterion was $\frac{\|r_k\|_2}{\|r_0\|_2} < 10^{-5}$, the starting vector the zero one and the righthand side vector of the system was $(1 1 \ldots 1)^T$
Table: \( f_1(x, y) = (x^4 + y^2)(|x| + |y|^3 + 1) \)

<table>
<thead>
<tr>
<th>( n = m )</th>
<th>( B )</th>
<th>( K_{4,4} )</th>
<th>( F_{4,4} )</th>
<th>( P_{4,4} )</th>
<th>( \tau )</th>
<th>( C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>17</td>
<td>12</td>
<td>11</td>
<td>9</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>47</td>
<td>21</td>
<td>15</td>
<td>15</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>32</td>
<td>73</td>
<td>28</td>
<td>20</td>
<td>18</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>64</td>
<td>91</td>
<td>31</td>
<td>22</td>
<td>20</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>128</td>
<td>103</td>
<td>33</td>
<td>23</td>
<td>21</td>
<td>13</td>
<td>21</td>
</tr>
</tbody>
</table>
Table: \( f_2(x, y) = (x^2 + y^2)(x^4 + y^4 + .1) \)

<table>
<thead>
<tr>
<th>( n = m )</th>
<th>( B )</th>
<th>( K_{6,6} )</th>
<th>( F_{6,6} )</th>
<th>( P_{6,6} )</th>
<th>( \tau )</th>
<th>( C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>10</td>
<td>11</td>
<td>-</td>
<td>10</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>16</td>
<td>48</td>
<td>17</td>
<td>18</td>
<td>27</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>32</td>
<td>81</td>
<td>21</td>
<td>98</td>
<td>80</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>64</td>
<td>94</td>
<td>23</td>
<td>*</td>
<td>*</td>
<td>13</td>
<td>21</td>
</tr>
<tr>
<td>128</td>
<td>*</td>
<td>24</td>
<td>*</td>
<td>*</td>
<td>14</td>
<td>23</td>
</tr>
</tbody>
</table>

*: The number of iterations exceeds 100
-: The matrix is singular.
Let $f \in C_{2\pi,2\pi}$ even function on $Q$ with roots $(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)$, on $Q$ each one of multiplicities $(2\mu_1, 2\nu_1), (2\mu_2, 2\nu_2), \ldots, (2\mu_k, 2\nu_k)$ respectively. Consider the system

$$(T_{nm}(f) + B)x = b,$$

where $B$ is a symmetric, p.d block band matrix. We can choose as preconditioner the

$$A_{nm}(\sqrt{w})(T_{nm}(g) + B)A_{nm}(\sqrt{w})$$

where $\{\tau, C, H\}$ is the set of matrices belonging to Block $\tau$, Block Circulant and Block Hartley algebra, respectively, and $g(x, y)$ the trigonometric polynomial having the same roots with the same multiplicities with $f$. 
Table: $f(x, y) = (|x||y|)(|x| + |y| + 2)$

<table>
<thead>
<tr>
<th>$n = m$</th>
<th>$I$</th>
<th>$B$</th>
<th>$\tau(f)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>10</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>50</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>32</td>
<td>174</td>
<td>39</td>
<td>7</td>
</tr>
<tr>
<td>64</td>
<td>603</td>
<td>62</td>
<td>8</td>
</tr>
<tr>
<td>128</td>
<td>*</td>
<td>98</td>
<td>9</td>
</tr>
</tbody>
</table>
We define as $P_{nm}$ the matrix

$$Q_{nm} \cdot \text{Diag} \left( f(v^{[nm]}) \right) \cdot Q_{nm}^H$$

where $Q_{nm}$ is the orthogonal matrix diagonalize the $\tau$ algebra and $v^{[nm]}_i = \left( \frac{\pi i}{n+1}, \frac{\pi j}{m+1} \right)$. Then, the following theorem holds true:

**Theorem**

Let $f \in C_{2\pi,2\pi}$ even function on $Q$ with root at $(0,0)$ with multiplicity $(q, r)$ with $q, r \leq 2$. Then, the spectrum of $P_{nm}^{-1} T_{nm}(f)$ is clustered around unity. Moreover, for every $n, m$ it holds that

$$c < \sigma(P_{nm}^{-1} T_{nm}(f)) < C$$

with $c, C > 0$ independent of $n, m$. 
Table: \( f(x, y) = (|x| + |y|)(x^2 + y^2 + 1) \)

<table>
<thead>
<tr>
<th>( n = m )</th>
<th>#I</th>
<th>( \lambda_{\text{min}}(I) )</th>
<th>( \lambda_{\text{max}}(P) )</th>
<th>( \lambda_{\text{min}}(P) )</th>
<th>#P</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>10</td>
<td>0.808</td>
<td>1.363</td>
<td>0.894</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>40</td>
<td>0.329</td>
<td>1.361</td>
<td>0.819</td>
<td>6</td>
</tr>
<tr>
<td>32</td>
<td>89</td>
<td>0.152</td>
<td>1.388</td>
<td>0.757</td>
<td>6</td>
</tr>
<tr>
<td>64</td>
<td>142</td>
<td>0.072</td>
<td>1.405</td>
<td>0.711</td>
<td>7</td>
</tr>
<tr>
<td>128</td>
<td>207</td>
<td>0.037</td>
<td>1.419</td>
<td>0.692</td>
<td>7</td>
</tr>
</tbody>
</table>
Thank you very much for your attention!