REMARKS ON THE SOLUTION OF TOEPLITZ SYSTEMS OF EQUATIONS

by

Seymour V. Parter

Computer Sciences Technical Report #721

October 1987

REMARKS ON THE SOLUTION OF+ TOEPLITZ SYSTEMS OF EQUATIONS

by

Seymour V. Parter(*)

+ Sponsored by the Air Force Office of Scientific Research under Contracts No. AFOSR-86-0163 and AFOSR-86-0126. This report is taken from the author's invited plenary address at the 2nd Internationl Symposium on National Analysis, Prauge, Czechoslovakia, Aug. 24-29, 1987.

(*) Departments of Mathematics and Computer Sciences, University of Wisconsin, Madison, WI 53706.

REMARKS ON THE SOLUTION OF TOEPLITZ SYSTEMS OF EQUATIONS

Seymour V. Parter

Introduction

In this survey we recall some of the theory of Toeplitz matrices which is relevant for the questions which arise in the inversion of Toeplitz systems of equations. In the course of our presentation we discuss some examples and raise some particular questions. Much of our discussion is based on the classical results of G. Szegö [7]; M. G. Krein [4]; Calderón, Spitzer and Widom [3], G. Baxter [1] and I. Gohberg.

A recent paper [2] by J. Bunch provides an excellent background for reasons for the interest in questions discussed here.

A Toeplitz matrix is an $(n + 1) \times (n + 1)$ matrix of the form

$$T_n = \begin{bmatrix} t_0 & t_{-1} & \cdots & & & t_{-n} \\ t_1 & t_0 & t_{-1} & & & \\ t_2 & t_1 & t_0 & t_{-1} & & \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & t_{-1} \\ t_n & t_{n-1} & \cdots & & t_1 & t_0 \end{bmatrix}. \tag{1.1}$$

Most of the methods for the "fast" inversion of T_n rely on the Levinson algorithm (see [2], [5]) which uses the decomposition

$$T_n = \begin{bmatrix} 1 & \alpha^T \\ \beta & T_{n-1} \end{bmatrix} \tag{1.2}$$

with t_0 set equal to "one". Then T_n^{-1} is obtained with the aid of T_{n-1}^{-1}. With this fact in mind Bunch considers the matrices

$$(1.3) \qquad T_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad T_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Observe that T_1 is singular while the condition number of T_2 is 5.828... . However, when one studies Toeplitz systems one is interested in large n. Let us consider the extended Bunch example: T_n is a tridiagonal matrix with "1" on the main diagonal and "1" on both the first subdiagonal and first superdiagonal. All other entries are "0". The eigenvalues of this T_n are

$$(1.4) \qquad \lambda_k^{(n)} = 1 + 2 \cos \frac{\pi k}{n+2}, \qquad k = 1,2,\ldots(n+1) .$$

Thus, T_n is singular if and only if $n \equiv 1 (\mathrm{mod}\ 3)$ [let $\hat{k} = \frac{2}{3}(n+2)$]. However, if $n \not\equiv 1 (\mathrm{mod}\ 3)$ then T_n^{-1} exists but (a) condition number of $T_n \sim 1/n$ and (b) there are very many eigenvalues $\lambda_j^{(n)} = O(1/n)$. So, either T_n is singular or badly conditioned and deflation techniques and/or singular value decomposition techniques will not be easily applicable.

Question 1: Can we construct an example of Toeplitz matrices T_n which for large n possess the "Bunch" property: T_{n-1} is singular and T_n is nonsingular and well conditioned?

2. General Theory:

Let $f(\theta) \sim \Sigma\, t_m e^{im\theta}$ and $\hat{x}(\theta) \sim \Sigma\, x_j e^{ij\theta}$. Then (formally)

$$f(\theta)\hat{x}(\theta) \sim \sum_{(k)} \{ \sum_{(j)} t_{k-j} x_j \} e^{ik\theta} . \tag{2.1}$$

Let $x = \{x_j\}_{j=-\infty}^{\infty}$ and define the operator $T[f]$ by

$$(T[f]x)_k = \sum_j t_{k-j} x_j . \tag{2.2}$$

If, for every formal doubly infinite sequence $x = \{x_j\}$ we define

$$\hat{x}(\theta) = \Sigma\, x_j e^{ij\theta} ,$$

then

$$\widehat{T[f]x}\,(\theta) = f(\theta)\hat{x}(\theta) . \tag{2.3}$$

Hence, the Toeplitz operator $T[f]$ defined on the (formal) doubly infinite sequences is isomorphic to multiplication and $T[f]T[g] = T[fg]$.

Now let us consider two special cases.

<u>Case 1</u>: Let $\ell^2 = \{x = \{x_k\}_{k=-\infty}^{\infty};\ \Sigma\, |x_k|^2 < \infty\}$. Then $T[f]$ is a bounded map with $T[f] : \ell^2 \to \ell^2$ if and only if $f(\theta) \in L_\infty(-\pi,\pi)$. Moreover, $(T[f])^{-1}$ exists as a bounded map on ℓ^2 if and only if $[1/f(\theta)] \in L_\infty(-\pi,\pi)$. And, we have $T[f]^{-1} = T[1/f]$.

<u>Case 2</u>: Let $\ell^\infty = \{x = \{x_k\}_{k=-\infty}^{\infty} : \sup|x_k| < \infty\}$. Then $T[f]$ is a bounded map with $T[f] : \ell^\infty \to \ell^\infty$ if and only if $f(\theta)$ has an absolutely convergent Fourier Series. In this case we write $f \in A$. Moreover, $T[f]^{-1}$ exists (Wiener-Levy theorem) if and only if $f(\theta) \neq 0$. Then $1/f \in A$ and $T[f]^{-1} = T[1/f]$.

We now turn to the <u>semi-infinite case</u>.

Let $\ell_+^\infty = \{x = \{x_k\}_{k=0}^{\infty} : \sup|x_n| < \infty\}$. Consider the map

$$(2.4) \qquad (T_+[f]x)_k = \sum_{j=0}^{\infty} t_{k-j}x_j, \qquad k \geq 0 .$$

Once more $T_+[f]$ is a bounded map of ℓ_+^∞ into ℓ_+^∞ if and only if $f \in A$. The discussion of $T_+[f]^{-1}$ is the result of a beautiful theory essentially discovered simultaneously by M. G. Kreĭn [4] and Calderón, Spitzer and Widom [3].

Define

$$(2.5) \qquad I(f) := \frac{1}{2\pi}\,[\arg f(2\pi) - \arg f(0)] .$$

Then, $T_+[f]^{-1}$ exists as a bounded map from ℓ_+^∞ to ℓ_+^∞ if and only if

$$(2.6a) \qquad f(\theta) \neq 0 ,$$

$$(2.6b) \qquad I(f) = 0 .$$

If (2.6a) holds but $I(f) = k \neq 0$ we have

<u>Case 1</u>: $k < 0$. $T_+[f]$ is invertible on the right and the nullspace of $T_+[f]$ has dimension $|k|$.

<u>Case 2</u>: $k > 0$. $T_+[f]$ is invertible on the left and the factor space $\ell_+^\infty/\text{Range } T_+$ has dimension k.

<u>Example</u>: $f(\theta) = e^{ik\theta}$, $I[f] = k$. If $k > 0$ then $T_+[f]\hat{x}(\theta) \sim \sum_{j=0}^{\infty} x_j e^{i(k+j)\theta}$ and the vectors corresponding to $1, e^{i\theta}, \ldots e^{i(k-1)\theta}$ are not in Range $T_+[f]$. On the other hand, if $k < 0$ these same functions correspond to vectors in the nullspace of $T_+[f]$.

The case of $\ell_+^2 := \{\{x_k\}_{k=0}^\infty, \sum_{k=0}^{\infty} |x_k|^2 < \infty\}$ is not yet described so completely. As before, $T_+[f]$ is a bounded map ℓ_+^2 to ℓ_+^2 if and only if $f(\theta) \in L_\infty(-\pi,\pi)$. But, what about $T_+[f]^{-1}$? While we do not have a complete theory we have some knowledge in special cases.

First: If $f(\theta) \in A$ and (2.6a), (2.6b) hold then $T_+[f]^{-1}$ exists as a bounded map from ℓ_+^2 to ℓ_+^2.

Second: If $f(\theta) \in L_\infty(-\pi,\pi)$ and (a) $f(\theta)$ is real valued with $0 < m = \inf|f(\theta)| \leq \sup|f(\theta)| = M$. Then $T_+[f]^{-1}$ exists as a bounded map from ℓ_+^2 to ℓ_+^2.

Third: If $f(\theta) = |f(\theta)|R(\theta)$ with

$$|f(\theta)| = 1, \quad R(\theta) \in A$$

and $R(\theta)$ satisfies (2.6a) and (2.6b). Then $T_+[f]^{-1}$ exists as a bounded map from ℓ_+^2 to ℓ_+^2.

Finally, let us return to $T_n[f] : \mathbf{C}_{n+1} \to \mathbf{C}_{n+1}$. Observe that we have three possibilities concerning the relationship of $T_n[f]$ to $T[f]$ or $T_+[f]$.

(a) We are given the specific finite-dimensional problem

$$T_n[f]X = Y \ . \tag{2.7}$$

(b) We are given a problem $T[f]X = Y$ and we came upon (2.7) as a Galerkin approximation.

(c) We are given a problem $T_+[f]X = Y$ and we came upon (2.7) as a Galerkin approximation.

Consider the two latter cases (b) and (c) wherein (2.7) is an attempt to approximate an infinite dimensional problem. We must first answer the question: Suppose the infinite dimensional operator - $T[f]$ on ℓ^2 or ℓ^∞ or $T_+[f]$ on ℓ_+^2 or ℓ_+^∞ has a bounded inverse, can we assert that for n large enough $T_n[f]$ will have a (uniformly) bounded inverse? Unfortunately, the answer is: not necessarily!

Consider the doubly infinite cases, ℓ^2 or ℓ^∞. Let $f(\theta) = e^{i\theta}$. Then

$$T_n[f] = \begin{bmatrix} 0 & 0 & \cdots & & & & 0 \\ 1 & 0 & \cdots & & & & 0 \\ 0 & 1 & 0 & \cdots & & & 0 \\ \vdots & & \ddots & & & & \vdots \\ 0 & \cdots & & & 0 & 1 & 0 \end{bmatrix}.$$

Then (i) $T_n[f]$ is singular for all n, (ii) all of its eigenvalues are zero, (iii) exactly one singular value is zero and the remaining n singular values are all "one". This example involves a complex-valued generating function. Suppose we require that $f(\theta)$ be real valued. Then $T_n[f]$ is hermitian. However, if $f(\theta)$ is both real valued and "odd" then

(2.8) $$T_n[f] = i[\text{real, skew-symmetric}], \quad i = \sqrt{-1} .$$

Thus, whenever n is even (and $n + 1$ is odd) $T_n[f]$ is singular. In particular, consider the square wave

(2.9) $$f(\theta) = \pi \operatorname{sgn} \theta, \quad -\pi < \theta < \pi .$$

In this case $T[f]^{-1} = T[\frac{1}{\pi} \operatorname{sgn} \theta]$ is a bounded map on ℓ^2 to ℓ^2. Suppose we acknowledge this obvious difficulty and restrict ourselves to the case of odd n $(n + 1$ even). Is $T_n[f]$ non-singular? At this time, even for this special case, we do not know the answer. There is one tool which can help in the general case.

Theorem (Szegö 1920): Let

$$m = \inf\{f(\theta)\} \leq \sup\{f(\theta)\} \leq M .$$

Let $\lambda_j^{(n)}$, $j = 1,2,\ldots(n + 1)$ be the eigenvalues of $T_n[f]$. Then

(2.10) $$m \leq \lambda_j^{(n)} \leq M ,$$

and, for fixed $j \geq 1$ we have

(2.11) $$\lambda_j^{(n)} \to m, \quad \lambda_{n+2-j}^{(n)} \to M \quad \text{as} \quad n \to \infty .$$

In fact, let $F(\lambda) \in C[m,M]$. Then

(2.12) $$\lim_{n\to\infty} \left\{\frac{1}{n+1} \sum_{j=1}^{n+1} F(\lambda_j^{(n)})\right\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(f(\theta))d\theta .$$

This last result implies that if $f(\theta_0) = 0$ and f is continuous at θ_0 we have: For every $\varepsilon > 0$ there is a $\delta > 0$ and the number of eigenvalues $\lambda_j^{(n)}$ which satisfy $|\lambda_j^{(n)}| < \varepsilon$ is $O(\delta n)$. Thus, in such case, even if $T_n[f]$ is not singular it is (very) badly conditioned and deflation techniques or singular value techniques can be expected to have great difficulties.

However, the special function (2.9) is not continuous at a zero. Hence this theorem sheds no light on this case. Therefore we have computed the singular values of $T_n[f]$ for odd values of $n \leq 299$. We find

$$.903 \leq \sigma_j^{(n)} \leq \pi, \quad n \text{ odd}, \ n \leq 299 .$$

The inclusion theorems assert that in the case n even and $n \leq 298$ we have exactly one singular value $\sigma_{n+1}^{(n)} = 0$ and

$$\sigma_j^{(n)} \geq .903, \quad j \neq n+1 .$$

Question 2: For this special case consider $\text{limit}\{\sigma_{n+1}^{(n)} : n \to \infty, n \text{ odd}\} = \bar{\sigma}$. Is $\bar{\sigma} = 0$? If so, does there exist a finite n_0 such that $\sigma_{n_0+1}^{(n_0)} = 0$?

Question 3: For all cases where $f(\theta)$ and $1/f(\theta)$ belong to $L_\infty(-\pi,\pi)$ find a good method to approximate the solution of $T[f]X = Y$.

When dealing with $T_+[f]$, $f \in A$ we have a beautiful result of Baxter [1].

Theorem: If $f \in A$ and satisfies (2.6a), (2.6b) then there is an n_0 and a $B > 0$ such that, if $n \geq n_0$ then $T_n[f]^{-1}$ exists and

$$\|T_n[f]^{-1}\|_\infty \leq B, \quad \|T_n[f]^{-1}\|_2 \leq B .$$

As a consequence of the Baxter theorem we also have the following additional insights. Let $I(f) = k \neq 0$.

Case 1: $k < 0$. Consider the matrix obtained from $T_n[f]$ by deleting the first k columns and the last k rows. This is $T_{n-k}[e^{-ik\theta}f]$. Applying the Baxter theorem we see that for large n this submatrix is uniformly invertible and $T_n[f]$ has at most k bad singular values. Thus, deflation or singular value techniques are useful to handle this case.

Case 2: $k > 0$. Consider the matrix obtained from $T_n[f]$ by deleting the first k rows and the last k columns. This is $T_{n-k}[e^{-ik\theta}f]$ and the conclusions above apply in this case also. The special function $f(\theta) = e^{ik\theta}$ illustrates the remarks above.

In the second case discussed earlier: $f(\theta)$ real, $f(\theta) \in L_\infty(-\pi,\pi)$, $0 < m = \inf\{f(\theta)\}$ is uniformly invertible in ℓ^2_{n+1} because of (2.10). Unfortunately there is no comparable result in the third case, $f(\theta) = |f(\theta)|R(\theta)$ with $0 < m = \inf|f(\theta)|$ and $R(\theta) \in A$ and $T_+[R]$ invertible. On the other hand we do know that there is an n_0 and a k such that: for $n \geq n_0$ the number of singular values of $T_n[f]$ less than $\frac{m}{2}$ is $\leq k$ (see [6]). Indeed, this result holds when $R(\theta)$ is merely continuous on $-\infty < \theta < \infty$ and periodic with period 2π. No index condition is required.

Question 4: Suppose $f(\theta) = |f(\theta)|R(\theta)$ with $R(\theta) \in A$ and (2.6a), (2.6b) hold. Does the conclusion of the Baxter theorem hold? If not, can one find an effective way to approximate the solution of $T_+[f]X = Y$?

Question 5: Suppose $f(\theta) = |f(\theta)|R(\theta)$ with $R(\theta)$ continuous on $-\infty < \theta < \infty$ and periodic with period 2π. Suppose $0 < m = \inf|f(\theta)|$ and (2.6b) holds. What can we say about $T_+[f]^{-1}$? What can we say about $T_n[f]^{-1}$?

We close this report with the complex version of Szegö's theorem.

Theorem: Let $f(\theta) \in L_\infty(-\pi,\pi)$. Let $M = \sup|f(\theta)|$. Let $\sigma_1^{(n)} \geq \sigma_2^{(n)} \geq \sigma_{n+1}^{(n)} \geq 0$, and let $F(\lambda) \in C[0,M]$. Then

$$\lim_{n\to\infty} \left\{\frac{1}{n+1} \sum_{j=1}^{n+1} F(\sigma_j^{(n)})\right\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(|f(\theta)|)d\theta .$$

Proof: This was first proven for $f(\theta) = |f(\theta)|R(\theta)$ with $R(\theta)$ continuous on $-\infty < \theta < \infty$ and periodic with period 2π [6]. However, Florin Avram showed me a simple proof in the more general case.

This work was supported by the U.S. Air Force under AFOSR contract AFOSR-860163.

References

[1] G. Baxter, A norm inequality for a "finite-section" Wiener-Hopf equation, Illinois J. Math. 7 (1963), 97-103.

[2] J. R. Bunch, Stability of methods for solving Toeplitz systems of equations, SIAM J. Sci. Stat. Comput. 6 (1985), 349-364.

[3] A. Calderón, F. Spitzer and H. Widom, Inversion of Toeplitz matrices, Illinois J. Math. 3 (1959), 490-498.

[4] M. G. Kreĭn, Integral equations on the half-line with a kernel depending on the difference of the arguments, Uspehi Mat. Nauk 13 (1958), 3-120. English transl. Amer. Math. Soc. Transl. 22 (1962), 163-788.

[5] N. Levinson, The Wiener rms (root-mean-square) error in filter design and prediction, J. Math. Phys. 25 (1947), 261-278.

[6] S. V. Parter, On the distribution of the singular values of Toeplitz matrices, Lin. Alg. and its Appl. 80 (1986), 115-130.

[7] G. Szegö, Beiträge zur Theorie der Toeplitzuchen Formen, I, Math. Z. 6 (1920), 167-202.