A New Algorithm for the Nearest Singular Toeplitz Matrix to a Given Toeplitz Matrix

Andrew E. Yagle
Department of EECS, The University of Michigan, Ann Arbor, MI 48109-2122

Abstract— **Several signal processing applications can be formulated as the computation of the null vector of a Hermitian Toeplitz matrix. These include array processing, spectral estimation, and beamforming algorithms applied directly to data rather than to its autocorrelation, and some blind deconvolution algorithms. When the data are noisy, the matrix is nonsingular, and the closest singular Toeplitz matrix (in the mean square norm) to the given matrix must be computed. Two major approaches have been used for this problem: (1) alternatingly subtracting off the outer product of minimum singular vectors and averaging along diagonals; and (2) structured total least squares. Both require many iterations of computationally intensive singular value decompositions. We present a new algorithm that is: (1) non-iterative; and (2) requires only solution of a Toeplitz system of equations. Several interesting linear algebra issues arise. Numerical examples illustrate the new algorithm.**

Keywords— **Toeplitz matrices, spectral estimation.**
**Phone: 734-763-9810. Fax: 734-763-1503.
Email: aey@eecs.umich.edu.** **EDICS: 2-REST.**

I. INTRODUCTION

A. Background

Several problems in signal processing can be formulated as computation of the null vector of a Hermitian Toeplitz matrix. For example, let $\{X_{-M}, \dots X_0 \dots X_M\}$ with $X_{-k} = X_k^*$ be a complex-valued time series modelled by finite Fourier series

$$X_k = \sum_{n=1}^{M} x_n e^{j\omega_n k}, \quad |k| \le M \tag{1}$$

for unknown real $\{x_n\}$ and $\{\omega_n\}$. These unknown model parameters can be computed from the data by computing the null vector of a Hermitian Toeplitz matrix with $(m,n)^{th}$ element X_{m-n} and computing the roots of the polynomial whose coefficients are elements of its Hermitian symmetric null vector $\vec{a}$:

$$\begin{bmatrix} X_0 & \cdots & X_M \\ \ddots & \ddots & \ddots \\ X_{-m} & \cdots & X_0 \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_0^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$a_0 z^M + \ddots + a_0 = 0 \rightarrow z = e^{j\omega_n}. \tag{2}$$

Alternatively, the X_k might be real and even and the Fourier coefficients x_n conjugate symmetric.

This problem arises in 1D array processing, spectral estimation, beamforming, detection of resonance frequencies, and many other applications.

In practice, the data X_k are noisy, so that the above matrix is not singular. A simple likelihood function argument shows that if the noise is additive zero-mean white Gaussian (in either time or frequency), then the likelihood is maximized when the given data $\{X_k\}$ are perturbed as little as possible (in the mean square norm sense) to make the matrix drop rank. The difficulty is that the Hermitian Toeplitz structure of the matrix must be maintained.

B. Other Spectral Estimation Algorithms

Most spectral estimation algorithms, such as Pisarenko method, MUSIC, and ESPRIT, operate not on the data but on the autocorrelation function estimated from the data. This has the advantage that additive white noise tends to be concentrated in the subspace spanned by the singular vectors associated with the minimum singular values, since the autocorrelation of zero-mean white noise is an impulse.

However, autocorrelation-based methods can be inappropriate, for the following three reasons:

- Only a small number of data points are available, not a long time series of data;
- Estimation of autocorrelation from data, which is always inexact due to end effects, is impractical;
- In practice, the additive noise is often neither white nor uncorrelated with the data.

Hence an approach that operates directly on the data, rather than on the autocorrelation, is desirable. TLS Prony method is an example of this. However, TLS Prony uses the TLS solution (minimum singular vector of the matrix), which does not preserver the Hermitian Toeplitz structure of the matrix. The above problem is an improvement on TLS Prony.

Two major approaches are known for this problem. First is an iterative algorithm alternating between:

- Computing the nearest (in Frobenius norm) lower rank matrix using the singular value decomposition, by subtracting the outer product of the minimum singular vectors times the minimum singular value;

• Computing the nearest (in Frobenius norm) Toeplitz matrix by averaging along the diagonals.

The other is structured total least squares, which iteratively perturbs the matrix closer to singularity, averaging along the diagonals in each iteration to preserve Toeplitz structure (standard TLS is one iteration).

Both of these approaches have been applied successfully. However, the enormous computation of repeatedly computing the singular value decomposition of the matrix for possibly many iterations, suggests that a much simpler algorithm is desirable.

C. New Algorithm

This paper proposes a new algorithm for finding the closest reduced-rank Hermitian Toeplitz matrix to a given Hermitian Toeplitz matrix. The error criterion is not the Frobenius norm (sum of squared magnitudes of matrix elements) of the perturbation, which weights lower-indexed values of X_k more than higher-indexed values, since they occur more often in the matrix. For example, X_0 appears M+1 times along the main diagonal, while X_M appears only twice (once as X_M and once as $X_{-M} = X_M^*$).

The new algorithm has three advantages over the previous algorithms mentioned above:

• It requires no singular value decompositions, only the solution of a single Toeplitz system of equations;
• It is non-iterative, hence no convergence issues;
• The least-squares error criterion applied to the $\{X_k\}$, rather than to the Toeplitz matrix, is more suitable in many signal processing applications.

II. Noisy Data: Computation of Nearest Singular Hermitian Toeplitz Matrix

A. Background

The algorithm is based on a result that goes back to Kronecker: A Hankel matrix has rank M if and only if its elements h_{i+j-1} are Fourier coefficients of a strictly proper rational function whose denominator has degree M. Suppose that this rational function has M distinct poles $e^{j\omega_i}$ all on the unit circle. Then the elements h_{i+j-1} can be expanded as

$$h_n = \sum_{i=1}^{M} C_i e^{j\omega_i n}. \tag{3}$$

for some constants C_i. The function $\hat{h}_n$ closest (in the least-squares norm) to h_n that reduces the matrix rank from M to M-1 is determined by discarding the term with the smallest $|C_i|$. This is similar to model order reduction techniques in linear system theory.

More precisely, suppose the ω_i are all rational numbers, and let N be the least common multiple of their denominators. Then h_n is periodic with period N, and the $e^{j\omega_i n}$ are orthogonal functions. Parseval's theorem then proves discarding the term with the smallest $|C_i|$ produces the minimum perturbation in the least-squares norm of a period of h_n.

There are three problems with this approach:

• How to compute poles from a nonsingular matrix;
• There is no reason to believe the poles will lie on the unit circle, as required by the above model;
• Hence the terms are not orthogonal, and discarding the term with the smallest $|C_i|$ may not be optimal.

B. Toeplitz Matrix Extension

We can solve all of these problems by dealing with Hermitian Toeplitz matrices instead of Hankel matrices; this explains the use of Toeplitz matrices throughout this paper. We first fit a model to the Hermitian Toeplitz matrix (2) by extending it to

$$\begin{bmatrix} X_1 & \dots & X_{K+1} & X \\ \ddots & \ddots & \ddots & \ddots \\ X^* & X_{K+1}^* & \ddots & X_1 \end{bmatrix} \tag{4}$$

and choosing X so that this matrix is singular. This clearly fits a model of order K+1 to (2). To reduce the rank of (2) from K+1 to K, we reduce the model order from K+1 to K by discarding the term with smallest $|C_i|$ in the polar expansion.

But there is a problem: setting the determinant of this matrix to zero produces a quadratic equation in X if the X_k are real, and an equation of the form

$$\begin{gathered} XX^* + A^*X + AX^* + B = 0 \rightarrow \\ (X+A)(X+A)^* = |X+A|^2 = |A|^2 - B \end{gathered} \tag{5}$$

for some constants A and B. The quadratic equation has two solutions for X; the above equation has an infinite number of solutions! This is not surprising; there are not enough data points to uniquely determine the poles from the data. What is to be done?

C. Structures of Null Vectors

Some matrix theory: Let J be the exchange matrix with ones on the main antidiagonal and zeros elsewhere; note $J^2 = I$ is the identity matrix. Then a Hankel matrix can be converted to a Toeplitz matrix by pre or post-multiplying by J, which does not affect the rank. However, the Hermitian Toeplitz structure imposes a certain type of Hermitian structure on the null vector. To see this note that for any Toeplitz matrix T we have $JTJ = T^T$, while for Hermitian Toeplitz matrices $JTJ = T^*$. The null vector $\vec{a}$ of a Hermitian Toeplitz matrix has the structure

$$0 = T\vec{a} = JT^*J\vec{a} \to 0 = T(CJ\vec{a}^*) \to \vec{a} = CJ\vec{a}^*. \tag{6}$$

But consistency requires that $\vec{a}$ satisfy

$$\vec{a} = CJ\vec{a}^* \to CJ\vec{a}^* = CC^*\vec{a} \to |C| = 1 \tag{7}$$

so that the constant C must lie on the unit circle. Real symmetric Toeplitz matrices can have a null vector that may be either symmetric or antisymmetric. Hermitian Toeplitz matrices can have a null vector that has the Hermitian structure defined above for any constant $|C| = 1$, so C lies on the unit circle.

This explains the ambiguity in the choice of X making the extended matrix singular. For real symmetric matrices, the two X's lead to symmetric (C=1) or antisymmetric (C=-1) structure in the null vector. For Hermitian matrices, the infinite number of X's from solving (5) lead to an infinite number of choices for the constant $|C| = 1 \to C = e^{j\theta}$.

A more useful way of seeing this ambiguity is to apply the Levinson algorithm to (2). The extended matrix will be singular if and only if the magnitude of the next reflection coefficient is unity (±1 for real symmetric matrices). The "inner product" expression in the Levinson algorithm then provides a quick way to compute the values of X that result in the desired reflection coefficient. This avoids computing the determinant (which is computationally expensive) and solving an equation of the form (5). A simple example of this is given in the next section.

This still doesn't answer the issue of which null vector structure to choose. The answer is to note that the poles are the roots of the polynomial having for its coefficients the null vector elements (although we don't compute the poles this way). If the null vector has *purely Hermitian structure* ($\vec{a} = J\vec{a}^*$), then the poles will all: (1) lie on the unit circle; or (2) occur in reciprocal complex conjugate quadruples $\{p, p^*, \frac{1}{p}, \frac{1}{p^*}\}$. Since the noiseless X_k have all their poles on the unit circle, the noisy X_k will also, if the noise level is low. Hence we choose the reflection coefficient (or X) that yields a purely Hermitian null vector (see the example given in the next section).

More precisely, we have the following summary for various null vector structures:

structure	length	fixed zeros
Hermitian	odd	none
anti-Herm	odd	±1
Hermitian	even	-1
anti-Herm	even	+1

For odd lengths, we clearly prefer the Hermitian structure, since anti-Hermitian constrains two zeros. For even lengths, we cannot avoid a zero constraint. The constrained zero is then the zero that is discarded (see the example in the next section).

For very large noise levels, there may be a reciprocal complex conjugate quadruple. A root locus argument shows that this arises from two poles on the unit circle coalescing and then moving off the unit circle. The poles may be moved to the unit circle. Another possibility is a pair of real-valued poles at reciprocal locations (e.g., 2 and 1/2). In this case it may be desirable to discard the pole inside the unit circle, if it decays fast enough so that its effect on X_k is minimal.

D. Nearest Singular Toeplitz Matrix Micro-Example

The goal is to compute the singular symmetric Toeplitz matrix nearest to the matrix

$$\begin{bmatrix} 5 & 1 & -5 \\ 1 & 5 & 1 \\ -5 & 1 & 5 \end{bmatrix} \quad \text{from} \quad 4\cos(\frac{\pi}{2}n) + \{1, 1, -1\} \tag{8}$$

There are two ways to compute x values that render singular the extended matrix

$$\begin{bmatrix} 5 & 1 & -5 & x \\ 1 & 5 & 1 & -5 \\ -5 & 1 & 5 & 1 \\ x & -5 & 1 & 5 \end{bmatrix}. \tag{9}$$

The first way is to set the determinant to zero and solve the quadratic equation

$$24x^2 + 152x + 224 = 0 \to x = -4, -7/3. \tag{10}$$

The other way is to use the Levinson algorithm, whose last recursion computes

$$\begin{bmatrix} 5 & 1 & -5 \\ 1 & 5 & 1 \\ -5 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ -\frac{5}{12} \\ \frac{13}{12} \end{bmatrix} = \begin{bmatrix} -\frac{10}{12} \\ 0 \\ 0 \end{bmatrix}. \tag{11}$$

The "inner product" for reflection coefficient ρ is

$$\pm 1 = \rho = -\begin{bmatrix} x & -5 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -\frac{5}{12} \\ \frac{13}{12} \end{bmatrix} / [-\frac{10}{12}] \tag{12}$$

which leads to the following results:

ρ	x	null vector	symmetry
+1	$-\frac{7}{3}$	[3 2 2 3]	symmetric
-1	-4	[2 -3 3 -2]	antisymm.

We choose the symmetric null vector from

symmetry	cubic equation	three roots
symmetric	$3z^3$+$2z^2$+2z+3=0	-1, $e^{\pm j1.403}$
antisymm.	$2z^3$–$3z^2$+3z–2=0	$+1$, $e^{\pm j1.318}$

We compute the constants A and B from

$$\begin{bmatrix} 5 \\ 1 \\ -5 \end{bmatrix} = \begin{bmatrix} \cos(1.403(0)) & (-1)^0 \\ \cos(1.403(1)) & (-1)^1 \\ \cos(1.403(2)) & (-1)^2 \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} \tag{13}$$

which has the solution A=5.143 and B=-0.143. This models the original Toeplitz matrix as

$$\begin{bmatrix} 5 & 1 & -5 \\ 1 & 5 & 1 \\ -5 & 1 & 5 \end{bmatrix} \text{ from } 5.143\cos(1.403n) - 0.143(-1)^n. \tag{14}$$

The closest singular symmetric Toeplitz matrix is

$$\begin{bmatrix} 5.143 & 0.857 & -4.857 \\ 0.857 & 5.143 & 0.857 \\ -4.857 & 0.857 & 5.143 \end{bmatrix}. \tag{15}$$

This was computed by keeping $5.143\cos(1.403n)$ and discarding $-0.143(-1)^n$ from the above model. And 1.403 is close to the noiseless value of $\frac{\pi}{2}$=1.571.

For comparison, the iterative Toeplitzation algorithm described above was also run on this example. After a half-dozen iterations the algorithm converged

$$\begin{bmatrix} 5.1392 & 0.9275 & -4.806 \\ 0.9275 & 5.1392 & 0.9275 \\ -4.806 & 0.9275 & 5.1392 \end{bmatrix} \tag{16}$$

which has null vector [1,-.361,1]'. The roots are

$$z^2 - 0.361z + 1 = 0 \rightarrow z = e^{\pm j1.389} \tag{17}$$

and 1.389 is farther than 1.403 from noiseless 1.571.

References