# Efficient Computation of $\mathrm{tr}\{\mathbf{TR}^{-1}\}$ for Toeplitz Matrices

José M. B. Dias and José M. N. Leitão

***Abstract*—An efficient algorithm for the computation of $\mathrm{tr}\{\mathbf{TR}^{-1}\}$, where $\mathbf{T}$ and $\mathbf{R}$ are Toeplitz matrices and $\mathbf{R}$ is also symmetric positive definite, is presented. The method exploits the fact that the trace of $\mathbf{TR}^{-1}$ depends only on the sum of the diagonals of $\mathbf{R}^{-1}$, and not on the whole matrix $\mathbf{R}^{-1}$. To obtain this sum, a fast efficient technique, built upon the Trench algorithm for computing the inverse of a Toeplitz matrix, is developed. The complexity of the algorithm depends on the generation function of matrix $\mathbf{R}$ and is $O(N \ln N)$ for generic functions and $O(p \ln p)$ for AR$(p)$ functions.**

***Index Terms*—Fast algorithm, fast Fourier transform, Toeplitz matrix, trace, Trench algorithm.**

## I. Introduction

The NEED for computing $\mathrm{tr}\{\mathbf{TR}^{-1}\}$, where $\mathbf{T}$ and $\mathbf{R}$ are Toeplitz matrices and $\mathbf{R}$ is also symmetric and positive definite (SPD), appears in many signal processing problems. Relevant examples are estimation of Toeplitz constrained covariance matrices [1], [2], matrix approximation under the *Frobenius norm* [3], functional approximation of Gaussian densities using *Kullback divergence* [4], channel estimation [5], pulse time-of-arrival analysis [6], and computation of the *Fisher information matrix* of zero-mean Gaussian processes [7]. In the latter case, one has to compute terms of the form $\mathrm{tr}\{\mathbf{T}\delta\mathbf{R}^{-1}\}$, where $\delta\mathbf{R}^{-1}$ denotes the derivative of $\mathbf{R}^{-1}$ with respect to a given parameter. These terms can be determined numerically from the knowledge of $\mathrm{tr}\{\mathbf{TR}^{-1}\}$.

Given the Toepliz matrices $\mathbf{T} \equiv [t_{i-j}]$ and $\mathbf{R} \equiv [r_{i-j}]$, for $i, j = 1, \ldots, N$, the obvious way of computing $\mathrm{tr}\{\mathbf{TR}^{-1}\}$ is to invert $\mathbf{R}$ and then determine the trace of $\mathbf{TR}^{-1}$; by using the Trench algorithm (see, e.g., [8]) to compute $\mathbf{R}^{-1}$, the total complexity in computing $\mathrm{tr}\{\mathbf{TR}^{-1}\}$, measured in floating point operations, is $(9/4)N^2 + (3/2)N$ ($\mathbf{R}^{-1}$ takes $(7/4)N^2$ and the remaining operations take $N^2/2 + (3/2)N$).

Porat and Friedlander [9], based on the Levinson–Durbin algorithm for computing the orthogonal polynomials of a Toepliz matrix, proposed an algorithm for the exact computation of the Fisher information matrix. This algorithm can be easily adapted to compute $\mathrm{tr}\{\mathbf{TR}^{-1}\}$. The complexity of the method is still $O(N^2)$.

When $N \to \infty$, and under adequate hypothesis on sequences $\{t_\tau, \tau \in \mathbb{Z}\}$ and $\{r_\tau, \tau \in \mathbb{Z}\}$ ($\mathbb{Z}$ denotes the integer set), we have (see, e.g., [7, p. 140])

$$\lim_{N\to\infty} N^{-1}\mathrm{tr}\{\mathbf{TR}^{-1}\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_t(\omega)S_r^{-1}(\omega)\,d\omega \tag{1}$$

where $S_t(\omega)$ and $S_r(\omega)$ are the discrete Fourier transform of sequences $\{t_\tau, \tau \in \mathbb{Z}\}$ and $\{r_\tau, \tau \in \mathbb{Z}\}$, respectively. Asymptotical result (1) is the basis of Whittle's formula [10], for the asymptotic normalized Fisher information matrix of a zero-mean normal process. This formula (1), despite leading to closed and simple expressions with light complexity, yields frequently a poor approximation for *small* sample sizes (see e.g., [11]).

To our knowledge, there is no general technique for the determination of $\mathrm{tr}\{\mathbf{TR}^{-1}\}$ with complexity lower than $O(N^2)$. In this letter, we introduce a faster technique. We begin by noting that $\mathrm{tr}\{\mathbf{TR}^{-1}\}$ depends only on the sum of the diagonals of $\mathbf{R}^{-1}$ (which we refer to as the *diagonal sum* of $\mathbf{R}^{-1}$). With this fact in mind, it is then proved, based on the Trench algorithm for determining the inverse of a Toeplitz matrix, that the diagonal sum of the referred matrix can be computed with $O(N \ln N)$ complexity. In computing the diagonal sum, it is necessary to solve a Toeplitz system. By using the preconditioned conjugate gradient technique (see [12]–[15]), this step has $O(N \ln N)$ complexity. If matrix $\mathbf{R}$ is generated by a rational function of order $(p, q)$, the methods [16], [17] solve the system with complexity $\max(p, q)O(N)$. Therefore, the total complexity in computing $\mathrm{tr}\{\mathbf{TR}^{-1}\}$ does not exceed $O(N \ln N)$.

## II. Efficient Computation of $\mathrm{tr}\{\mathbf{TR}^{-1}\}$

Let $\mathbf{T}, \mathbf{R} \in \mathbb{R}^{N\times N}$ be Toeplitz matrices[1] of real elements, where $\mathbf{R}$ is SPD and $\mathbf{R}^{-1} \equiv [c_{ij}]$. Noting that $t_{ij} = t_\tau$, where $\tau \equiv i - j$ for $i, j = 1, \ldots, N$, it follows that

$$\begin{aligned}\mathrm{tr}\{\mathbf{TR}^{-1}\} &= \sum_{i,j=1}^{N} t_{ij}c_{ji} = \sum_{\tau=-N+1}^{N-1} t_\tau \sum_{i\in S_\tau} c_{i-\tau,i} \\ &= \sum_{\tau=-N+1}^{N-1} t_\tau \bar{c}_{-\tau}\end{aligned} \tag{2}$$

[1]Sometimes, we use the subscript $N$ meaning that matrices $\mathbf{T}_N$ and $\mathbf{R}_N$ are of dimension $N \times N$.

Manuscript received May 11, 2001; revised December 20, 2001. This work was supported by the Fundação para a Ciência e Tecnologia under Projects POSI/34071/CPS/2000 and 2/2.1.TIT/1580/95. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Shuichi Ohno.

The authors are with the Instituto Superior Técnico, Instituto de Telecomunicações, 1049-001 Lisboa, Portugal (e-mail: bioucas@lx.it.pt; jleitao@lx.it.pt).

Publisher Item Identifier S 1070-9908(02)03404-1.

where $\bar{c}_\tau \equiv \sum_{i\in S_\tau} c_{i+\tau,i}$ is the $\tau$th diagonal sum of $\mathbf{R}^{-1}$ ($\tau \geq 0$ denotes south-west diagonals), and

$$S_\tau = \begin{cases} 1+\tau, \ldots, N & \tau \geq 0 \\ 1, \ldots, N-|\tau| & \tau < 0. \end{cases} \qquad (3)$$

According to (2), $\mathrm{tr}\{\mathbf{TR}^{-1}\}$ depends on $t_\tau$ and on $\bar{c}_\tau$ (sum of the elements of $\mathbf{R}^{-1}$ along diagonal $\tau$).

### A. Sum of Diagonals of $\mathbf{R}^{-1}$

Toeplitz matrices belong to the larger class of *persymmetric matrices* [18]: matrix $\mathbf{R}_N$ is persymmetric if it is symmetric about its northeast–southwest diagonal, i.e., if $r_{ij} = r_{N-j+1,N-i+1}$ for $i, j = 1, \ldots, N$. In an equivalent form $\mathbf{R}_N = \mathbf{E}_N\mathbf{R}_N^T\mathbf{E}_N$, where $\mathbf{E}_N = [\delta_{N-i+1,j}]$ is the $N \times N$ *exchange matrix*. Note that $\mathbf{E}_N^{-1} = \mathbf{E}_N$. Thus, the inverse of a persymmetric matrix is, if it exits, also persymmetric.

Consider the partition

$$\mathbf{R}_N^{-1} = \begin{bmatrix} \mathbf{R}_{N-1} & \mathbf{E}_{N-1}\mathbf{r} \\ \mathbf{r}^T\mathbf{E}_{N-1} & r_0 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{R}' & \boldsymbol{\nu} \\ \boldsymbol{\nu}^T & \gamma \end{bmatrix} \qquad (4)$$

with $\mathbf{r} = [r_{-1}, \ldots, r_{-N+1}]^T$. Vector $[\boldsymbol{\nu}^T, \gamma]^T$, the last column of $\mathbf{R}_N^{-1}$, and matrix $\mathbf{R}'$ are given by (see, e.g., [18, p. 130])

$$\begin{cases} \boldsymbol{\nu} = \gamma\mathbf{E}\boldsymbol{\alpha} \\ \gamma = \dfrac{1}{r_0 + \mathbf{r}^T\boldsymbol{\alpha}} \\ \mathbf{R}' = \mathbf{R}_{N-1}^{-1} + \dfrac{\boldsymbol{\nu}\boldsymbol{\nu}^T}{\gamma} \end{cases} \qquad (5)$$

where $\boldsymbol{\alpha}$ is the solution of the Yule–Walker equation $\mathbf{R}_{N-1}\boldsymbol{\alpha} = -\mathbf{r}$.

Matrix $\mathbf{R}_{N-1}^{-1}$ exists and is persymmetric; thus $[\mathbf{R}_{N-1}^{-1}]_{i,j} = [\mathbf{R}_{N-1}^{-1}]_{N-j,N-i}$ and therefore

$$c_{ij} = \left[\mathbf{R}_{N-1}^{-1}\right]_{i,j} + \frac{1}{\gamma}\nu_i\nu_j \qquad (6)$$

$$= \left[\mathbf{R}_{N-1}^{-1}\right]_{N-j,N-i} + \frac{1}{\gamma}\nu_i\nu_j \qquad (7)$$

for $i, j = 1, \ldots, N-1$. From (6) we see that $[\mathbf{R}_{N-1}^{-1}]_{N-j,N-i} = c_{N-j,N-i} - (1/\gamma)\nu_{N-j}\nu_{N-i}$. Using this equality in (7), we are led to

$$c_{ij} = c_{N-j,N-i} + \frac{1}{\gamma}(\nu_i\nu_j - \nu_{N-j}\nu_{N-i}). \qquad (8)$$

Since $[c_{ij}] = \mathbf{R}^{-1}$ is persymmetric, element $c_{ij}$ satisfies $c_{N-j,N-i} = c_{i+1,j+1}$ for $i, j = 1, \ldots, N-1$; replacing $c_{N-j,N-i}$ by $c_{i+1,j+1}$ in (8), it follows that

$$c_{i+1,j+1} = c_{ij} + \frac{1}{\gamma}(\nu_{N-j}\nu_{N-i} - \nu_i\nu_j) \qquad (9)$$

for $i, j = 1, \ldots, N-1$. Defining $j \equiv i+\tau$, $\nu_N \equiv \gamma$, and noting that $c_{1,i} = \nu_{N+1-i}$ for $i, j = 1, \ldots, N$, it follows that

$$c_{i+1,i+1+\tau} = c_{i,i+\tau} + \frac{1}{\gamma}(\nu_{N-i-\tau}\nu_{N-i} - \nu_i\nu_{i+\tau}) \qquad (10)$$

valid for $|\tau| \leq N-2$ and $i = 1, \ldots, N-\tau-1$, when $\tau \geq 0$, and $i = 1-\tau, \ldots, N-1$, when $\tau \leq 0$. Expressions (9) or (10) generate recursively, from vector $\boldsymbol{\nu}$, elements $c_{ij}$ along each diagonal of $\mathbf{R}^{-1}$. The diagonal sum $\bar{c}_\tau$ is, after simple but lengthy manipulation of (10), given by

$$\gamma\bar{c}_\tau = \sum_{i=1}^{N-\tau}[i\nu_i\nu_{i+\tau} + \nu_i\nu_{i+\tau}(i+\tau) - N\nu_i\nu_{i+\tau}]. \qquad (11)$$

The determination of $\bar{c}_\tau$, for $\tau = 1, \ldots, N-1$, according to (11), has $O(N^2)$ complexity. Notice, however, that each term in the sum (11) defines a convolution, which can be computed using the *fast Fourier transform* (FFT) with $O(N\ln N)$ complexity. For this purpose, define $\mathbf{0}_N^T$ as a zero row vector of dimension $N$ and the sequences $\{u_i'\}$ and $\{\nu_i'\}$, with $i \in \mathbb{Z}$, as periodic extensions, of period $2N$, of sequences $\{1, \ldots, N, \mathbf{0}_N^T\}$ and $\{\nu_1, \ldots, \nu_N, \mathbf{0}_N^T\}$, respectively. Sum (11), using entities $u_i'$ and $\nu_i'$, assumes the form

$$\gamma\bar{c}_\tau = \sum_{i=1}^{2N}\left(u_i'\nu_i'\nu_{i+\tau}' + \nu_i'\nu_{i+\tau}'u_{i+\tau}' - N\nu_i'\nu_{i+\tau}'\right) \qquad (12)$$

$$= \underbrace{(u_{-i}'\nu_{-i}') \star \nu_i' + \nu_{-i}' \star (\nu_i'u_i')}_{A_{N,\tau}} - N\underbrace{\nu_{-i}' \star \nu_i'}_{B_{N,\tau}} \qquad (13)$$

where symbol $\star$ means circular convolution of length $2N$. Denoting $\mathcal{DF}[x]$ as the $2N$-vector containing the discrete time Fourier series of $x$ and $\mathcal{DF}^{-1}$ its inverse (i.e., $x_n = \mathcal{DF}^{-1}[\mathcal{DF}[x]]_{(n)}$), and using the discrete time Fourier series properties (convolution, time symmetry, and conjugation), expression (13) is given by

$$\gamma\bar{c}_\tau = \mathcal{DF}^{-1}\left\{\mathrm{Re}\left(\mathcal{DF}\left[(u_i''\nu_i')\right] \odot \mathcal{DF}^*\left[\nu_i'\right]\right)\right\}\big|_{(-\tau)} \qquad (14)$$

where $u_i'' \equiv 2u_i' - N$ and symbol $\odot$ denotes element-wise multiplication.

The number of floating point operations needed to implement (14) is, approximately, $3N\ln 2N$ (corresponding to three FFTs of size $2N$). Table I summarizes the proposed algorithm.

In the next section, we show that, if $\mathbf{R}$ is the covariance matrix of an AR($p$) process, then the terms $\gamma$, $A_{N,\tau}$ and $B_{N,\tau}$ of expression (13) are constant for $N > p$.

TABLE I
SUMMARY OF THE PROPOSED SCHEME FOR THE DETERMINATION OF THE DIAGONAL SUM OF $\mathbf{R}_N^{-1}$

| **Computation of $\bar{\mathbf{c}}_\tau$** | |
|---|---|
| Step | Complexity |
| 1. Solve $\mathbf{R}_{N-1}\boldsymbol{\alpha} = -\mathbf{r}$ | $O(N\ln N)$ |
| 2. Compute $[\boldsymbol{\nu}^T, \gamma]$ according (5) | $N$ |
| 3. Compute $\mathcal{DF}[u_i''\nu_i']$ and $\mathcal{DF}[\nu_i']$ | $2N\ln_2 2N$ |
| 4. Compute $\bar{c}_\tau$ according (14) | $N\ln_2 2N$ |

## III. AUTOREGRESSIVE PROCESSES

Suppose that $\mathbf{R}$ is a covariance matrix of an AR($p$) process with coefficients $\{a_k, 0 \leq k \leq p\}$ where $a_0 \equiv 1$ and $a_p \neq 0$, By hypothesis, filter $a(z) = \sum_{k=0}^{p} a_k z^{-k}$ is stable [i.e., $a(z)$ does not have zeros for $|z| \geq 1$]. In theses conditions, we have for $p \geq N+1$ (see, e.g., [7, ch. 6])

$$\gamma_{N+1} = \gamma_N \quad (15)$$

$$\boldsymbol{\alpha}_N = (a_1, a_2, \ldots, a_p, 0, \ldots, 0). \quad (16)$$

From the expression for $\boldsymbol{\nu}$ given in (5), we conclude that, for $N > p$

$$(\nu_1, \ldots, \nu_{N-1}, \nu_N) = \gamma(0, \ldots, 0, a_p, \ldots, a_1, a_{\mathbf{0}}). \quad (17)$$

Replacing (17) into (13), assuming that $N > p$, and after some manipulation, we get

$$\gamma\bar{c}_\tau = \underbrace{-\left(\sum_{i=0}^{p-\tau} i a_i a_{i+\tau} + \sum_{i=0}^{p-\tau} a_i a_{i+\tau}(i+\tau)\right)}_{A_{N,\tau}} + N \underbrace{\left(\sum_{i=0}^{p-\tau} a_i a_{i+\tau}\right)}_{B_{N,\tau}}. \quad (18)$$

We conclude then that terms $\gamma$, $A_{N,\tau}$ and $B_{N,\tau}$ do not depend on $N$ for $N > p$. Therefore, the computation of $\mathrm{tr}\{\mathbf{TR}^{-1}\}$, where $\mathbf{R}$ is the covariance matrix of an AR$(p)$ process, according to the proposed method, has $O(p \ln p)$ complexity for $N > p$.

For generic processes, formula (18) does not apply. However, terms $\gamma_N$, $A_{N,\tau}$, and $B_{N,\tau}$ converge to a constant as values of $N \to \infty$. This is a consequence of the Levison–Durbin recursions

$$\boldsymbol{\alpha}_N = [\boldsymbol{\alpha}_{N-1}, 0] - c_N[0, \mathbf{E}_N \boldsymbol{\alpha}_N]$$

where $c_N$ is the $N$th-order reflection coefficient, which satisfy (see, e.g., [9])

$$\lim_{N\to\infty} \sum_{i=N+1}^{\infty} c_i^2 = 0.$$

Hence, for $N_0$ sufficiently large we have, for $N \geq N_0$

$$c_\tau \simeq \gamma_{N_0}^{-1}(A_{N_0,\tau} + N B_{N_0,\tau}).$$

The problem in applying the above extrapolation is, of course, how to determine $N_0$, in order to obtain a good approximation for $c_\tau$. A rule of thumb is to determine $\gamma_N$, $A_{N,\tau}$, and $B_{N,\tau}$ for an increasing sequence $N_1, N_2, \ldots$, for example $N_i = 2^i$, for $i = 1, 2, \ldots$, and estimate an integer $N_i$ above which the former terms are practically constant.

## IV. Concluding Remarks

We presented an efficient algorithm for computing $\mathrm{tr}\{\mathbf{TR}^{-1}\}$, where $\mathbf{T}$ and $\mathbf{R}$ are Toeplitz matrices and $\mathbf{R}$ is also symmetric and positive definite. The complexity of the algorithm depends on the generation function of matrix $\mathbf{R}$ and is $O(N \ln N)$ for generic functions and $O(p \ln p)$ for AR$(p)$ functions. We have also presented an approximated extrapolation formula with $O(N_0 \ln N_0)$ complexity, thus independent of $N$.

The central idea exploited is that $\mathrm{tr}\{\mathbf{TR}^{-1}\}$ depends only on the sum of diagonals of $\mathbf{R}^{-1}$ and not on each single element of this matrix. The computation of the referred sum is carried out efficiently by using the fast Fourier transform.

## References

[1] D. Fuhrmann and M. Miller, "On the existence of positive-definite maximum likelihood estimates of structured covariance matrices," *IEEE Trans. Inform. Theory*, vol. 34, pp. 722–729, 1988.
[2] M. Miller and D. Snyder, "The role of likelihood and entropy in incomplete data problems: Application to estimating point processes intensity and Toeplitz constrained covariances," *Proc. IEEE*, vol. 75, pp. 892–907, July 1987.
[3] L. Scharf, *Statistical Signal Processing. Detection, Estimation and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.
[4] S. Kullback, *Information Theory and Statistics*. New York: Wiley, 1978.
[5] C. Tellambura, Y. Guo, and S. Barton, "Channel estimation using aperiodic binary sequences," *IEEE Commun. Lett.*, vol. 2, pp. 140–142, May 1998.
[6] S. Ray, "A novel pulse TOA analysis technique for radar identification," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, pp. 716–721, July 1998.
[7] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
[8] W. F. Trench, "An algorithm for the inversion of finite Toeplitz matrices," *J. SIAM*, vol. 12, pp. 515–522, 1964.
[9] B. Porat and B. Friedlander, "Computation of the exact information matrix of Gaussian time series with stationary random components," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 118–130, 1986.
[10] P. Whittle, "Estimation and information in stationary time series," *Arkiv Matematick*, vol. B-2, no. 23, pp. 423–434, 1953.
[11] B. Porat and B. Friedlander, "The exact Cramér–Rao bound for Gaussian autoregressive processes," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 537–542, 1987.
[12] T. F. Chan, "Circulant preconditioners for Hermitian Toeplitz systems," *SIAM J. Matrix Anal. Applicat.*, vol. 10, pp. 542–550, 1989.
[13] ——, "Toeplitz equations by conjugate gradients with circulant preconditioner," *SIAM J. Sci. Statist. Comput.*, vol. 10, pp. 104–119, 1989.
[14] T. Ku and C. Kuo, "Design and analysis of optimal Toeplitz preconditioners," *IEEE Trans. Signal Processing*, vol. 40, pp. 129–141, Jan. 1992.
[15] ——, "Spectral properties of preconditioned rational Toeplitz matrices," *SIAM J. Matrix Anal. Applicat.*, vol. 14, pp. 146–165, 1993.
[16] W. E. Trench, "Solution of systems with Toeplitz matrices generated by rational functions," *Lin. Algeb. Applicat.*, vol. 74, pp. 191–211, 1986.
[17] ——, "Toeplitz systems associated with the product of a formal Laurent series and a Laurent polynomial," *SIAM J. Matrix Anal. Applicat.*, vol. 9, pp. 181–193, 1988.
[18] G. H. Golub and C. F. Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1983.