# ON THE INVERSION OF THE SAMPLE COVARIANCE MATRIX IN A STATIONARY AUTOREGRESSIVE PROCESS[1]

By M. M. Siddiqui[2]

*University of North Carolina*

Let $x_1, \cdots, x_N$ be the observations on a variate at times $t = 1, \cdots, N$. It is assumed that the underlying model is an autoregressive scheme of order $k$

$$a_0 x_t + a_1 x_{t-1} + \cdots + a_k x_{t-k} = z_t, \tag{1}$$

where $z$'s are independent $N(0, 1)$ variates and the roots of the equation

$$\sum_{j=0}^{k} a_j y^j = 0$$

lie inside the unit circle $|y| = 1$ in the complex plane. The variate $z_t$ is, then, independent of $x_{t-1}, x_{t-2}, \cdots$ ([2], p. 38). It is further assumed that the process is stationary so that $E\, x_t$, $E\, x_t x_{t+j}$, $j = 0, 1, 2, \cdots$ are independent of $t$. Writing $\sigma_x^2$ for the variance of any $x$, we observe that since $E\, x_t = 0$, $\sigma_x^2 = E\, x_t^2$. We define autocorrelation between $x_t$ and $x_s$ by

$$\gamma_{|t-s|} = E\, x_t x_s / \sigma_x^2 \tag{2}$$

so that $\gamma_t$ satisfies Eq. (1) with $z_t$ replaced by zero and $\gamma_{-t} = \gamma_t$.

Let $X_j$ stand for the column vector of the first $j$ observations and $X_j'$ for its transpose, i.e.,

$$X_j' = (x_1, \cdots, x_j), \qquad j = 1, 2, \cdots, N. \tag{3}$$

Also write $A_j$ for the covariance matrix of the vector $X_j$, i.e.,

$$A_j = \sigma_x^2 \begin{bmatrix} 1 & \gamma_1 & \gamma_2 & \cdots & \gamma_{j-1} \\ \gamma_1 & 1 & \gamma_1 & \cdots & \gamma_{j-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{j-1} & \gamma_{j-2} & \gamma_{j-3} & \cdots & 1 \end{bmatrix}, \tag{4}$$

for $j = 1, 2, \cdots, N$. We note here that the matrix $A_j$ is persymmetric, i.e., symmetric about both the diagonals. This property will be used to obtain $A_N^{-1}$.

The distribution of $X_N$ is given by

$$dF(X_N) = (2\pi)^{-N/2} |A_N|^{-1/2} \exp\left[-\tfrac{1}{2}(X_N' A_N^{-1} X_N)\right] dX_N. \tag{5}$$

J. Wise [1] has given a method of finding $A_N^{-1}$ using the spectral density function. We propose here another method of obtaining $A_N^{-1}$ based on the symmetric property of $A_N$.

Received February 2, 1957; revised November 13, 1957.

[1] Sponsored by the Office of Naval Research under the contract for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] At present with the Boulder Laboratories, National Bureau of Standards.

The distribution of $x_1, \cdots, x_k, z_{k+1}, \cdots, z_N$ is given by

$$dF(x_1, \cdots, x_k, z_{k+1}, \cdots, z_N) = (2\pi)^{-N/2} |A_k|^{-1/2} \exp\left[-\frac{1}{2}\left\{X_k' A_k^{-1} X_k + \sum_{t=k+1}^{N} z_t^2\right\}\right] dX_k \, dz_{k+1} \cdots dz_N.$$

We shall assume here that $N > 2k$. Considering (1) as a transformation from $z_t$ to $x_t$ for $t = k + 1, \cdots, N$, we obtain the distribution of $X_N$ as

$$dF(X_N) = (2\pi)^{-N/2} a_0^{N-k} |A_k|^{-1/2} \exp\left[-\frac{1}{2}\left\{X_k' A_k^{-1} X_k + \sum_{t=k+1}^{N}\left(\sum_{i=0}^{k} a_i x_{t-i}\right)^2\right\}\right] dX_N. \tag{6}$$

Comparing (5) and (6) we have

$$a_0^{2N} |A_N| = a_0^{2k} |A_k| \tag{7}$$

and

$$X_N' A_N^{-1} X_N = X_k' A_k^{-1} X_k + \sum_{t=k+1}^{N}\left(\sum_{i=0}^{k} a_i x_{t-i}\right)^2. \tag{8}$$

Let $C_N$ be the $N \times N$ matrix which has $A_k^{-1}$ in the upper right-hand corner and zeroes elsewhere, i.e.,

$$C_N = \begin{bmatrix} A_k^{-1} & 0 \\ 0 & 0 \end{bmatrix}, \tag{9}$$

and $B_N$ be the matrix of the quadratic form in the second term on the right of Eq. (8), i.e.,

$$\sum_{t=k+1}^{N}\left(\sum_{i=0}^{k} a_i x_{t-i}\right)^2 = X_N' B_N X_N, \tag{10}$$

so that we have

$$A_N^{-1} = B_N + C_N. \tag{11}$$

Denoting by $a_{ij}$, $b_{ij}$, and $c_{ij}$, $i, j = 1, 2, \cdots, N$, the elements in the $i$th row and $j$th column of the matrices $A_N^{-1}$, $B_N$, and $C_N$ respectively, we have

$$a_{ij} = b_{ij} + c_{ij}. \tag{12}$$

But $c_{ij} = 0$ if either $i$ or $j > k$. Hence

$$a_{ij} = b_{ij} \text{ if either } i \text{ or } j > k. \tag{13}$$

Now $B_N$ is completely known. In fact, assuming $j \geqq i$,

$$b_{ji} = b_{ij} = \begin{cases} \sum_{t=k+1}^{k+i} a_{t-i} a_{t-j} & \text{for } j \leqq k, \\ 0 & \text{for } i + k < j \leqq N, \quad i \leqq N - k, \\ \sum_{t=j}^{k+i} a_{t-i} a_{t-j} & \text{for } k + 1 \leqq j \leqq i + k, \quad i \leqq N - k, \\ \sum_{t=j}^{N} a_{t-i} a_{t-j} & \text{for } i \geqq N - k + 1. \end{cases} \tag{14}$$

Thus all the $a_{ij}$, except those for which both $i$ and $j$ are less than or equal to $k$, are known. Now, since $A_N$ is persymmetric, so is $A_N^{-1}$. Therefore

$$a_{ji} = a_{ij} = a_{N-i+1,N-j+1}, \qquad i, j = 1, 2, \cdots, N. \tag{15}$$

Using (13) and remembering that $N > 2k$, we have

$$a_{ij} = a_{ji} = b_{N-i+1,N-j+1} \qquad \text{for } i, j = 1, 2, \cdots, k. \tag{16}$$

Thus $A_N^{-1}$ is completely determined. We now use relations (12) to obtain all the elements of $A_k^{-1}$. Once $A_k^{-1}$ is known, we can find $A_N^{-1}$ for any $N \geqq k$ using (6).

If $k$ is less than 5, we can directly compute $A_k^{-1}$ and then use Eq. (6) to obtain $A_N^{-1}$.

*Illustration.* Let $k = 2$. The distribution of $X_N$ is

$$dF(X_N) = (2\pi)^{-N/2} a_0^{N-2} |A_2|^{-1/2} \cdot \exp\left[-\frac{1}{2}\left\{X_2' A_2^{-1} X_2 + \sum_{t=3}^{N} (a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2})^2\right\}\right] dX_N,$$

so that

$$B_N = \begin{bmatrix} a_2^2 & a_1 a_2 & a_0 a_2 & 0 & \cdots & 0 & 0 \\ a_1 a_2 & a_1^2 + a_2^2 & a_0 a_1 + a_1 a_2 & a_0 a_2 & \cdots & 0 & 0 \\ a_0 a_2 & a_0 a_1 + a_1 a_2 & a_0^2 + a_1^2 + a_2^2 & a_0 a_1 + a_1 a_2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & a_0^2 + a_1^2 & a_0 a_1 \\ 0 & 0 & 0 & 0 & \cdots & a_0 a_1 & a_0^2 \end{bmatrix}.$$

Hence

$$A_N^{-1} = \begin{bmatrix} a_0^2 & a_0 a_1 & a_0 a_2 & 0 & \cdots & 0 & 0 \\ a_0 a_1 & a_0^2 + a_1^2 & a_0 a_1 + a_1 a_2 & a_0 a_2 & \cdots & 0 & 0 \\ a_0 a_2 & a_0 a_1 + a_1 a_2 & a_0^2 + a_1^2 + a_2^2 & a_0 a_1 + a_1 a_2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & a_0^2 + a_1^2 & a_0 a_1 \\ 0 & 0 & 0 & 0 & \cdots & a_0 a_1 & a_0^2 \end{bmatrix},$$

$$A_2^{-1} = \begin{bmatrix} a_0^2 - a_2^2 & a_0 a_1 - a_1 a_2 \\ a_0 a_1 - a_1 a_2 & a_0^2 - a_2^2 \end{bmatrix},$$

$$|A_2^{-1}| = (a_0^2 - a_2^2)^2 - (a_0 - a_2)^2 a_1^2,$$

and

$$a_0^{2N} \mid A_N \mid = a_0^4[(a_0^2 - a_2^2)^2 - (a_0 - a_2)^2 a_1^2]^{-1}.$$

It may be mentioned here that Ulf Gernander and Murray Rosenblatt ([2], pp. 238–239) have considered asymptotic properties of $A_N^{-1}$ as $N$ tends to infinity. They, however, do not attempt to determine the $k^2$ elements standing in the first $k$ rows and the first $k$ columns of $A_N^{-1}$, although they suggest a method of orthogonalization of the vector $X_N$.

## REFERENCES

[1] J. WISE, "The autocorrelation function and the spectral density function," *Biometrika*, Vol. 42 (1955), pp. 151–159.

[2] ULF GERNANDER AND MURRAY ROSENBLATT, *Statistical Analysis of Stationary Time Series*, John Wiley and Sons, New York, 1957.

# A PROBLEM OF BERKSON, AND MINIMUM VARIANCE ORDERLY ESTIMATORS[1]

BY JOHN W. TUKEY

*Princeton University*

**1. Summary.** The distinction between efficiency in the asymptotic sense originally introduced by Fisher ([2], 1925, p. 703), and the finite sample sense sometimes used by others has been recently stressed by various writers (e.g., Berkson [1]). The technique of proof used below was originally developed to provide a simple example where the maximum likelihood estimate of location, though asymptotically efficient, was not of minimum variance for any finite sample size whatever. The (symmetrical) double exponental distribution with known scale, where the sample median is the maximum likelihood estimator of location, could easily be shown to be such an example. (While this result is useful in deflating unwarranted views about minimum variance properties of maximum likelihood estimates, Fisher's ([2], p. 716) results about intrinsic accuracy in the same situation are of more basic interest.)

On examination, however, the technique used to provide this rather isolated and special result was found capable of showing, for a class of distributions with suitable monotony properties (in particular all distributions for which $f'(y)/f(y)$ is monotone decreasing, and all normal, exponential, gamma and beta distributions), that the covariances of the order statistics in a sample of any chosen size are monotone in either index separately.

Received March 22, 1957.

[1] Prepared in connection with research sponsored by the Office of Naval Research. Based, in part, on Memorandum Report 11 of the Statistical Research Group, Princeton University, which was issued 25 October 1948.