A Note on the Computation of the Variances of the Estimated Parameters In Multivariate Autoregression

Georgi N. Boshnakov*
Institute of Mathematics, Bulgarian Academy of Sciences

1. Introduction. Let $X = \{X_t, t \in Z\}$, $X_t = (X_{t1}, \ldots, X_{td})^T$, be a d-variate causal autoregression of order p,

$$X_t = \sum_{i=1}^{p} A_i X_{t-i} + \epsilon_t,$$

where $A_i, i = 1, \ldots, p$, and S are $d \times d$ matrices, $E(\epsilon_t \epsilon_t^T) = S$, $E(\epsilon_t \epsilon_s^T) = 0$, $t \neq s$, $E(\epsilon_t) = 0$.

When the autoregression parameters A_i, $i = 1, \ldots, p$, are estimated, their asymptotic covariance matrix Γ is a measure of the quality of the estimates. Its diagonal entries are of primary interest since they constitute the variances of the estimated parameters. That is the information which most computer programs give as output, even for the univariate autoregression and for the linear regression.

The matrix Γ is the inverse of the tensor product of certain matrices. In general it is necessary to compute the whole inverse matrix even if one needs its diagonal elements only. In this case, however, it is possible to avoid the redundant computations.

The aim of this note is to give formulas for the diagonal entries of Γ. These formulas admit efficient and simple program realization. They can be used when only the diagonal entries of Γ are needed.

We obtain also some identities regarding the "forward" and "backward" autoregression matrices which reveal some interconnections between them (Lemma 1). Although they can be deduced from the more general results obtained by Akaike (1973) the reader may be interested in the direct approach adopted here (see Section 3).

2. The covariance matrix of the estimated parameters. Let $n > 2p$, $X_{(n)} = (X_1^T, \ldots, X_n^T)^T$, $G_n = E(X_{(n)} X_{(n)}^T)$, $C_n = G_n^{-1}$. Let us arrange the elements of the matrices A_i, $i = 1, ..., p$, in a column vector β, so that the $((i-1)dp + (k-1)d + j)$-th element of β equals the (i,j)-th element of A_k (that is we take the first rows of A_i, $i = 1, ..., p$, then their second rows and so on). Let $\hat{\beta}_N$ be the multivariate Yule-Walker estimator of β, based on a sample of length N (Whittle (1963)). Then $\sqrt{N}(\hat{\beta}_N - \beta)$ is asymptotically normal with (asymptotic) covariance matrix $\Gamma = (S \otimes G_p)^{-1}$ (Hannan (1970), Chapter 6, Theorem 1). Noting that $(S \otimes G_p)^{-1} = S^{-1} \otimes G_p^{-1} = S^{-1} \otimes C_p$, since the inverse of a tensor product equals the products of the inverses, we obtain the following proposition.

Proposition. Let s^{ii} be the (i,i)-th entry of S^{-1}. and $C_p(j,j)$ - the (j,j)-th element (scalar!) of C_p. Then

$$var(\hat{\beta}_{(i-1)d+j}) = (1/N) s^{ii} C_p((i-1)d + j, (i-1)d + j).$$

* *Supported by the Ministry of Education and Science under Contract No. I 19/91.*

This proposition shows that the computation of the required variances needs only the inversion of the $p \times p$ matrix S, instead of the $dp \times dp$ matrix Γ. Moreover, only the diagonal elements (which can be computed very efficiently, see the next section) of the matrix C_p enter the formulas.

3. The inverse matrix. The matrix C_n can be written in block form as $[C_n^{(st)}]_{s,t=1}^n$, where $C_n^{(st)}$ are $d \times d$ matrices. It is known (Akaike(1973), Jones(1978)) that X admits also the "backward" representation

$$X_t = \sum_{i=1}^{p} B_i X_{t+i} + a_t, \qquad E(a_t a_t^T) = S_a, \qquad E(a_t a_s^T) = 0, \quad t \neq s, \qquad E(a_t) = 0.$$

where $B_i, i = 1, .., p$, are $d \times d$ matrices. Therefore the quadratic form $X_{(n)}^T C_n X_{(n)}$ can be written in two ways

$$X_{(n)}^T C_n X_{(n)} = X_{(p)}^T C_p X_{(p)} + U = V + Y_{(p)}^T C_p Y_{(p)}, \tag{1}$$

where

$$Y_{(p)} = (X_{n-p+1}^T, \ldots, X_n^T)^T,$$

$$U = \sum_{t=p+1}^{n} \left(X_t - \sum_{i=1}^{p} A_i X_{t-i} \right)^T S^{-1} \left(X_t - \sum_{i=1}^{p} A_i X_{t-i} \right)$$

$$V = \sum_{t=1}^{n-p} \left(X_t - \sum_{i=1}^{p} B_i X_{t+i} \right)^T S_a^{-1} \left(X_t - \sum_{i=1}^{p} B_i X_{t+i} \right)$$

Denote $M_i = A_i^T S^{-1} A_i$, $H_i = B_i^T S_a^{-1} B_i$, for $i = 1, \ldots, p$, $M_0 = S^{-1}$, $H_0 = S_a^{-1}$. We can obtain every desired block of C_n by the comparison of the similar terms in (1). To derive expressions for its diagonal blocks we compare those terms in (1) that have the form $X_s^T(\ldots)X_s$. The comparison gives the following Lemma.

Lemma 1.

$$\begin{aligned} C_n^{(ss)} &= C_p^{(ss)} + \sum_{i=p+1-s}^{p} M_i = \sum_{i=0}^{s-1} H_i, && s = 1, \ldots, p, \\ &= \sum_{i=0}^{p} M_i = \sum_{i=0}^{p} H_i, && s = p+1, \ldots, n-p, \\ &= \sum_{i=0}^{n-s} M_i = C_p^{(s-(n-p), s-(n-p))} + \sum_{i=s}^{n} H_{p-(n-i)}, \\ & && s = n-p+1, \ldots, n. \end{aligned}$$

As a corollary we obtain

Corollary 1. Let $C_p^{(00)} = 0$. Then for $s = 1, \ldots, p$,

$$C_p^{(ss)} = \sum_{i=0}^{s-1} H_i - \sum_{i=p+1-s}^{p} M_i = \sum_{i=0}^{s-1} H_i - \sum_{i=0}^{s-1} M_{p-i} =$$
$$= C_p^{(s-1s-1)} + H_{s-1} - M_{p-s+1}.$$

Corollary 1 is a special case of a formula, obtained under more general setting by Akaike (1973, equation (3.25)).

4. Examples. In the univariate case ($d = 1$) the forward and backward representations coincide (Box, Jenkins (1970)) and the coefficients are scalars. Introducing the notation $\phi_i = A_i = B_i$, $i = 1, \ldots, p$, we obtain from the previous section

$$C_p^{(ss)} = 1 + \sum_{i=1}^{s-1} \phi_i^2 - \sum_{i=0}^{s-1} \phi_{p-i}^2 = 1 - \phi_p^2 + \sum_{i=1}^{s-1} (\phi_i^2 - \phi_{p-i}^2), \qquad s = 1, \ldots, p.$$
$$= C_p^{(s-1s-1)} + \phi_{s-1}^2 - \phi_{p-s+1}^2.$$

This nice formula gives explicitly the variances of the autoregressive estimates in the univariate case. It is not new (Friedlander (1984)) but quite surprisingly it is not easily available in the time series literature. Wellknown textbooks and monographs do not state it (see Box G.P.E., Jenkins G.M. (1970), Hannan E.J.(1970), Priestly M.B.(1981), Brockwell, Davis (1987)). The authors say only that these are elements of an inverse matrix.
In the multivariate case we have, for example

$$C_p^{(11)} = H_o - M_p, \qquad C_p^{(22)} = H_o + H_1 - M_p - M_{p-1}.$$

5. Algorithm The generalization of the Levinson-Durbin algorithm to the solution of the multivariate Yule-Walker equations (Whittle (1963)) gives the matrices A_i, B_i, $i = 1, \ldots, p$, S^{-1}, S_a^{-1}. We summarize the steps necessary to compute the variances of the parameters in the following algorithm. Note that only the diagonal entries of the matrices M_i, H_i, and C_i, $i = 0, \ldots, p$, are computed and used.

1) **Compute** the diagonal elements of M_i, H_i, $i = 0, \ldots, p$.
   Put them in the $(p+1) \times d$ matrices m and h.
2) **Compute** $v[j] = h[0][j] - m[p][j]$, $j = 1, \ldots, d$.
3) **for** $i = 2$ **to** p **do**
   for $j = 1$ **to** d **do**
   $v[(i-1)d + j] = v[(i-2)d + j] + h[i-1][j] - m[p-i+1][j]$
4) **for** $i = 1$ **to** $d^2 p$ **do**
   $v[i] = v[i]/n$.

References

1. Akaike H. (1973) Block Toeplitz matrix inversion. SIAM J.Appli.Math., 24, 234-341.
2. Box G.E.P., Jenkins G.M. (1970) Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.
3. Brockwell P.J., Davis R.A. (1987) Time Series: Theory and Methods, Springer-Verlag.
4. Friedlander B. (1984) On the computation of the Cramer-Rao bound for ARMA parameter estimation, Trans. IEEE, ASSP-32, 4, 721-727.
5. Jones, R.H. (1978) Multivariate autoregression estimation using residuals. In Applied Time Series Analysis, ed. D.F.Findley, pp. 139-162, Academic Press, New York-London.
6. Hannan E.J. (1970) Multiple time series, Wiley.
7. Priestley M.B. (1981) Spectral Analysis and Time Series, Academic Press.
8. Whittle P. (1953) The analysis of multiple stationary time series, J.Roy.Statist.Soc., Ser. B, 15, 125-139.
9. Whittle P. (1963) On the fitting of multivariate autoregressions and the approximate factorization of a spectral density matrix, Biometrika, 50, pp. 129-134.