Ann. Inst. Statist. Math.
30 (1978), Part B, 499-504

COVARIANCE MATRIX COMPUTATION OF THE STATE VARIABLE OF A STATIONARY GAUSSIAN PROCESS

HIROTUGU AKAIKE

(Received Oct. 16, 1978; revised Dec. 7, 1978)

1. Summary

A recursive procedure for the computation of one-step ahead predictions for a finite span of time series data by a Gaussian autoregressive moving average model can be realized by using the Markovian representation of the model. The covariance matrix of the stationary state variable of the Markovian representation is required to implement a computational procedure of the predictions. A simple computational procedure of the covariance matrix which does not need an iterative method is obtained by using a canonical representation of the autoregressive moving average process. The recursive computation of the predictions realized by using this procedure provides a computationary efficient method of exact likelihood evaluation of a Gaussian autoregressive moving average model.

2. Recursive computation of one-step ahead predictions

Assume that a stationary scalar zero-mean Gaussian process $y(n)$ is defined by the relation

$$z(n+1)=Fz(n)+Gx(n+1)\,, \qquad y(n)=Hz(n) \tag{2.1}$$

where $x(n)$ is a scalar white noise independent of $z(n-1), z(n-2), \cdots$, and $z(n)$ is the state variable which is a stationary p-vector process and the matrices F, G, H are $p\times p$, $p\times 1$ and $1\times p$, respectively. When $y(0), y(1), \cdots, y(n-1)$ are given, the one-step ahead prediction of $z(n)$ is defined by

$$z(n|0, n-1)=\text{projection of } z(n) \text{ onto the linear space spanned by the components of } y(0), y(1), \cdots, y(n-1)\,,$$

and the one-step ahead prediction of $y(n)$ is given by $y(n|0, n-1)=Hz(n|0, n-1)$.

The computation of $z(n|0, n-1)$ can be done recursively by using

the relations

$$
\begin{aligned}
&e(n)=y(n)-Hz(n\,|\,0,\,n-1)\,,\\
&z(n+1\,|\,0,\,n)=Fz(n\,|\,0,\,n-1)+K_n r_n^{-1}e(n)\,,\\
&z(0\,|\,0,\,-1)=0\,,
\end{aligned}
\tag{2.2}
$$

where $r_n=\mathrm{E}\,e(n)^2$ and $K_n=\mathrm{E}\,z(n+1)e(n)$, the Kalman gain vector. By the Kalman filtering procedure, which is a standard procedure of recursive computation of the predictions, K_n and R_n are computed by the relations

$$
K_n=FP(n\,|\,n-1)H'\,,\qquad r_n=HP(n\,|\,n-1)H'\,,
\tag{2.3}
$$

where $'$ denotes transpose and $P(n\,|\,n-1)=\mathrm{E}\,(z(n)-z(n\,|\,0,\,n-1))(z(n)-z(n\,|\,0,\,n-1))'$ and is obtained by the relations

$$
\begin{aligned}
&P(n+1\,|\,n)=FP(n\,|\,n-1)F'+GqG'-K_n r_n^{-1}K_n'\,,\\
&P(0\,|\,-1)=P_0\,,
\end{aligned}
\tag{2.4}
$$

where $q=\mathrm{E}\,x(n)^2$ and $P_0=\mathrm{E}\,z(0)z(0)'$, the covariance matrix of the stationary state vector. Thus the numerical evaluation of P_0 forms the starting point of the Kalman filtering procedure.

3. Prediction of ARMA process

When $y(n)$ is a stationary Gaussian autoregressive moving average (ARMA) process defined by

$$
\begin{aligned}
&y(n)+b_1y(n-1)+\cdots+b_My(n-M)\\
&\quad=x(n)+a_1x(n-1)+\cdots+a_Lx(n-L)
\end{aligned}
\tag{3.1}
$$

there is a Markovian representation (2.1) defined by

$$
F=\begin{bmatrix}
0 & 1 & 0 & \cdots & 0\\
0 & 0 & 1 & \cdots & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots\\
0 & 0 & 0 & \cdots & 1\\
-b_K & -b_{K-1} & -b_{K-2} & \cdots & -b_1
\end{bmatrix},\qquad
G=\begin{bmatrix} w_0\\ w_1\\ \vdots\\ w_{K-2}\\ w_{K-1}\end{bmatrix},
\tag{3.2}
$$

$$
H=[1\ \ 0\ \ 0\cdots 0]\,,
$$

where $a_L\neq 0$, $b_M\neq 0$ and $K=\max(M,\,L+1)$ and the w_i's are the impulse responses of the system (3.1), i.e., $(w_0,\,w_1,\cdots,\,w_{K-1})=(y(0),\,y(1),\cdots,\,y(K-1))$ under the assumption that $y(-1)=y(-2)=\cdots=y(-M)=0$ and $x(n)=1$, for $n=0$, 0, otherwise. The corresponding state variable $z(n)$

is defined by

$$z_i(n)=y(n+i\,|\,n) \qquad i=0,1,\cdots,K-1\,, \tag{3.3}$$

where $z_i(n)$ denotes the $i+1$st component of $z(n)$ and $y(n+i\,|\,n)$ the projection of $y(n+i)$ onto the space spanned by the components of $y(n)$, $y(n-1),\cdots$; see, for example, Akaike ([1], [2]).

Thus the recursive procedure for the computation of one-step ahead predictions for a finite span of data can be applied to the autoregressive moving average process (3.1) via the representation (2.1) with the matrices defined by (3.2), if only the covariance matrix P_0 of the state variable $z(n)$ is obtained.

4. Covariance matrix of the state variable

Under the assumption that the characteristic polynomial

$$B(z)=1+b_1z+\cdots+b_Mz^M \tag{4.1}$$

has the zero's outside the unit circle $y(n)$ can be represented in the form

$$y(n)=\sum_{m=0}^{\infty} w_m x(n-m)\,, \tag{4.2}$$

where the summation denotes the limit in the mean square and w_m's are the impulse responses of the system defined by (3.1). By using (4.2) we get for $i\geqq 0$ the representation

$$y(n+i\,|\,n)=\sum_{m=0}^{\infty} w_{i+m}x(n-m)$$

or the relation

$$y(n+i)=\sum_{m=-i}^{-1} w_{i+m}x(n-m)+y(n+i\,|\,n)\,.$$

From this last relation we get for $j\geqq i$ the relation

$$\mathrm{E}\,y(n+i)y(n+j)=\sigma^2\sum_{m=0}^{i-1} w_m w_{m+j-i}+\mathrm{E}\,y(n+i\,|\,n)y(n+j\,|\,n)\,, \tag{4.3}$$

where σ^2 is the variance of $x(n)$. When the ARMA model (3.1) is given the computation of w_n's is straightforward. Thus it is only the computational procedure of $\mathrm{E}\,y(n+i)y(n+j)$ that is required for the evaluation of the covariance matrix of the state variable defined by (3.3).

By multiplying the both sides of (3.1) by $y(n-k)$ and taking the expectations we get the relations ($k=0,1,\cdots,K$)

$$R(-k)+b_1R(-k+1)+\cdots+b_KR(-k+K)$$
$$=a_kw_0+a_{k+1}w_1+\cdots+a_{K-1}w_{K-1-k}\ ,$$

where $b_m=0$ for $m>M$, $a_0=1$ and $a_l=0$ for $l>L$, and $R(k)=\mathrm{E}\,y(n+k)\cdot y(n)$. In the matrix form these relations are given by

$$\text{(4.4)}\qquad \begin{bmatrix} 1 & b_1 & b_2 & \cdots b_{K-1} & b_K \\ b_1 & 1+b_2 & b_3 & \cdots b_K & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{K-1} & b_{K-2}+b_K & b_{K-3}\cdots 1 & & 0 \\ b_K & b_{K-1} & b_{K-2}\cdots b_1 & & 1 \end{bmatrix} \begin{bmatrix} S(1) \\ S(2) \\ \vdots \\ S(K) \\ S(K+1) \end{bmatrix} = \begin{bmatrix} C(1) \\ C(2) \\ \vdots \\ C(K) \\ C(K+1) \end{bmatrix},$$

where $C(m)=a_{m-1}w_0+a_mw_1+\cdots+a_{K-1}w_{K-m}$ and $S(m)=R(m-1)$.

To solve the equation (4.4) for $S(m)$'s we want to transform the matrix on the left-hand side into lower triangular form with 1's on the diagonal. For notational convenience we put $b_i^K=b_i$ $(i=1, 2,\cdots, K)$. We subtract from the ith row b_K^K times the $(K+2-i)$th row and divide the resulting ith row by $1-(b_K^K)^2$ for $i=1, 2,\cdots, [(K+2)/2]$, where $[x]$ denotes the integer part of x. Hereafter the same operation as that applied to the elements of the matrix on the left-hand side is applied to the elements of the vector $[C(1), C(2),\cdots, C(K+1)]'$. By the present operation the b_i^K's in the first $[(K+2)/2]$ rows of the matrix are transformed into b_i^{K-1}'s which are defined by

$$\text{(4.5)}\qquad b_i^{K-1}=\frac{b_i^K-b_K^Kb_{K-i}^K}{1-(b_K^K)^2}\ ,\qquad i=1, 2,\cdots, K-1\ ,$$

and

$$b_K^{K-1}=0\ .$$

From (4.5) we get the relation

$$b_i^{K-1}=b_i^K-b_K^Kb_{K-i}^{K-1}\ ,\qquad i=1, 2,\cdots, K-1\ .$$

Thus by subtracting b_K^K times the $(K+2-i)$th row from the ith row for $i=[(K+2)/2]+1, [(K+2)/2]+2,\cdots, K+1$, the b_i^K's within the matrix are all replaced by b_i^{K-1}'s. When $b_K^K=0$ the above operation is non-effective and can be skipped. At the next stage, since we have $b_K^{K-1}=0$, we can restrict our attention to the first K rows and apply the same type of operation to transform b_i^{K-1}'s into b_i^{K-2}'s with $b_{K-1}^{K-2}=0$. The operation can also be skipped when $b_{K-1}^{K-1}=0$. By repeating the same type of operations K times we get

$$
(4.6)\qquad
\begin{bmatrix}
1 & 0 & 0 & \cdots 0 & 0 \\
b_1^1 & 1 & 0 & \cdots 0 & 0 \\
b_2^2 & b_1^2 & 1 & \cdots 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots \\
b_{K-1}^{K-1} & b_{K-2}^{K-1} & b_{K-3}^{K-1}\cdots 1 & & 0 \\
b_K^K & b_{K-1}^K & b_{K-2}^K\cdots b_1^K & & 1
\end{bmatrix}
\begin{bmatrix} S(1) \\ S(2) \\ S(3) \\ \vdots \\ S(K) \\ S(K+1) \end{bmatrix}
=
\begin{bmatrix} D(1) \\ D(2) \\ D(3) \\ \vdots \\ D(K) \\ D(K+1) \end{bmatrix}.
$$

The desired solution is then obtained by the recursive relations

$$
(4.7)\qquad
\begin{aligned}
&S(m)=D(m)-\sum_{i=1}^{m-1} b_i^{m-1}S(m-i)\,, \qquad m=2, 3, \cdots, K+1\,, \\
&S(1)=D(1)\,.
\end{aligned}
$$

The description of the above procedure shows that the equation (4.4) has a unique solution if only $|b_M^M|=1$ does not hold for $M=1, 2, \cdots, K$. It is already well known that $|b_M^M|<1$ $(M=1, 2, \cdots, K)$ is necessary and sufficient for the characteristic equation (4.1) to have zero's outside the unit circle; see, for example, Szaraniec [6]. The non-singularity of the matrix of (4.4) has been used by Kitagawa [4]. Thus under the present assumption of stationarity the equation (4.4) uniquely determines $R(m)$ $(m=0, 1, \cdots, K)$. From these $R(m)$'s we can get, via (4.3),

$$
P_0(i, j)=R(j-i)-\sigma^2\sum_{m=0}^{i-1} w_m w_{m+j-i}\,, \qquad j\geqq i\,,
$$

where $P_0(i, j)$ denotes the (i, j)-element of the covariance matrix P_0 of the state variable and $P_0(j, i)=P_0(i, j)$.

5. Application to exact likelihood computation

For a set of data $(y(1), y(2), \cdots, y(N))$ the exact likelihood of a zero-mean stationary Gaussian ARMA model (3.1) is given by

$$
\prod_{n=1}^{N}\left(\frac{1}{2\pi\sigma^2 r_n}\right)^{1/2}\exp\left[-\frac{1}{2\sigma^2 r_n}e(n)^2\right]\,,
$$

where $e(n)$ and r_n are obtained by (3.2), (2.2), (2.3) and (2.4) under the assumption that $q=1$ and σ^2 denotes the assumed value of the variance of $x(n)$.

Once P_0 is given a computational procedure of $e(n)$ and r_n given by Morf, Sidhu and Kailath [5], which is more efficient than that by (2.4), can be used for the computation of the exact likelihood. A computer program for the maximum likelihood computation of a Gaussian ARMA model by this procedure is already developed by Akaike et al. [3]. The computational efficiency of the likelihood computation procedure is con-

firmed through the comparison with other procedures.

Acknowledgements

The present author is indebted to T. Ozaki and G. Kitagawa who invited his attention to the subject of exact likelihood computation of an ARMA model.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Akaike, H. (1974). Markovian representation of stochastic processes and its application to the analysis of autoregressive moving-average processes, *Ann. Inst. Statist. Math.*, **26**, 363-387.

[2] Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion, In *System Identification: Advances and Case Studies*, R. K. Mehra and D. G. Lainiotis, eds., Academic Press, New York, 27-96.

[3] Akaike, H., Kitagawa, G., Arahata, E. and Tada, F. (1979). TIMSAC-78, *Computer Science Monographs*, No. 11, The Institute of Statistical Mathematics, Tokyo.

[4] Kitagawa, G. (1977). On a search procedure for the optimal AR-MA order, *Ann. Inst. Statist. Math.*, **29**, B, 319-332.

[5] Morf, M., Sidhu, G. S. and Kailath, T. (1974). Some new algorithms for recursive estimation in constant, linear, discrete-time systems, *IEEE Trans. Automat. Contr.*, AC-**19**, 315-323.

[6] Szaraniec, E. (1973). Stability, instability and aperiodicity tests for linear discrete systems, *Automatica*, **9**, 513-516.