A Note on Autoregressive-Moving Average Identification

E. J. Hannan

Biometrika, Volume 66, Issue 3 (Dec., 1979), 672-674.

Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28197912%2966%3A3%3C672%3AANOAAI%3E2.0.CO%3B2-U

Biometrika (1979), **66**, 3, *pp.* 672–4
Printed in Great Britain

A note on autoregressive-moving average identification

By E. J. HANNAN

Department of Statistics, Australian National University, Canberra

SUMMARY

A discussion is given of the identification and parameterization of autoregressive-moving average systems in relation to the use of certain canonical forms.

Some key words: Autoregressive-moving average system; Canonical form; Identification; McMillan degree; Prior constraint; Time series.

This note refers to the stationary p-variate autoregressive-moving average model,

$$X(t) = \sum_{j=1}^{q} A(j)\,X(t-j) + \varepsilon(t) + \sum_{j=1}^{r} B(j)\,\varepsilon(t-j), \tag{1}$$

$$E\{\varepsilon(t)\} = 0, \quad E\{\varepsilon(s)\,\varepsilon(t)'\} = \delta_{st}\,G,$$

and the discussion of its identification by Tuan (1978). The notation of that paper will be used and the conditions imposed will be maintained, namely that

$$h(z) = I - \Sigma\,A(j)\,z^j, \quad g(z) = I + \Sigma\,B(j)\,z^j$$

are left prime, that $h^{-1}g$ is analytic and nonsingular for $|z| < 1$ and that G is nonsingular. Nevertheless (1) is not uniquely specified. However (Hannan, 1969), if, for prescribed q, r, $[A(q) \vdots B(r)]$ is of rank p then (1) is uniquely specified. This will be called the rank condition. Some structures (1) cannot be brought to a form where this condition is satisfied (Hannan, 1971), so that the condition is overidentifying. The set, $C_{q,r}$, of all structures satisfying the rank condition for given q and r is mapped, by using the elements of $A(j)$ and $B(j)$ and the on and above diagonal elements of G, onto an open set in Euclidean space, if it is required that $h^{-1}g$ is nonsingular for $|z| \leqslant 1$, and hence constitutes an analytic manifold. For q and r fixed the set of structures (1) not in $C_{q,r}$ is evidently of lower dimension than $C_{q,r}$.

Tuan (1978, end of § 1) states a number of objections to the use of $C_{q,r}$ and we wish to discuss these, and the general problem of parameterizing (1). For this last purpose Tuan (1978) used a family of canonical forms, called by him the quasiautoregressive-moving average representation. With any structure (1) is associated a set of integers m_j ($j = 1, \ldots, p$) which determine the form of this representation (Tuan, 1978, p. 101). Thus given any structure (1) there is associated a set of m_j and a matrix of polynomials $u(z)$, with unit determinant, such that ug and uh are in the canonical form. Of course the matrix function u is an extremely complicated function of g and h. Using K as a symbol for the m_j ($j = 1, \ldots, p$), we may also map the set C_K of all structures with these m_j into an open set in Euclidean space, if $h^{-1}g$ is nonsingular for $|z| \leqslant 1$. Now there is no overidentification. However, there is a major problem if there are prior constraints imposed on (1) for it seems almost impossible, in general, to translate these constraints into constraints on the canonical forms because u is such a complicated function of g and h. On the other hand, the set of all structures (1) for given q and r is very complicated to parameterize. However, for the reason mentioned at the end of the previous paragraph, almost all of that set is constituted by $C_{q,r}$, which is easily parameterized. Tuan (1978) objects to $C_{q,r}$ because of overidentification and because constraints may cause the rank condition to fail, and for other reasons to be discussed below. It is most unlikely that constraints would make $[A(q) \vdots B(r)]$ identically of rank less than p, that is, for all parameter values satisfying the constraints, if only because such constraints are unlikely to be applied to the $B(j)$. If that problem did arise then other requirements would be needed (Deistler, Dunsmuir & Hannan, 1978). In any case it seems strange to criticize $C_{q,r}$ on the basis of the possibility of constraints since the canonical forms seem very difficult to use when these arise. It should be mentioned, of course, that the system might originally be built in state space form (Tuan, 1978, formula (2·1)) and that this will be constrained. In that case both $C_{q,r}$ and C_K will be unusable. In that case also it seems likely that the constraints will identify the system so that a unique parameterization is obtained and in particular the constraints will specify the McMillan degree so that no problems will arise. We shall therefore not discuss this case further.

In the unconstrained case the use of the C_K has considerable appeal. However the union, C_d, say, of all C_K, for K such that $\Sigma m_j = d$, itself constitutes an analytic manifold as discussed by J. M. C. Clark in an unpublished paper, and the decomposition of C_d onto the C_K is somewhat arbitrary (Tuan, 1978, p. 102). The integer d is called the McMillan degree. To cover C_d a total of $(d-1)!/\{(p-1)!\,(d-p)!\}$ coordinate neighbourhoods is required. Each of these coordinate neighbourhoods may be chosen to constitute an open submanifold of C_d that is dense in C_d. The submanifold, C_K, for $m_j = m_0$, $j \leqslant j_0$; $m_j = m_0 - 1$, $j > j_0$, which we shall call C_0, is one of these but the remaining C_K are submanifolds of lower dimension. Thus, as Tuan (1978, p. 102) says, a more appealing way to proceed may be to first choose d and then to choose the point in C_d, say by maximum likelihood; see also Hannan (1979).

A second objection raised by Tuan (1978) to $C_{q,r}$ is that if q_0 and r_0 are the true values and $q > q_0$, $r > r_0$, then the standard likelihood ratio tests are not appropriate, because the estimate of the $A(j)$ and $B(j)$ will not then converge in any reasonable fashion. To this author that does not seem a valid objection for two reasons. The first is that the objection applies equally well to the C_K or C_d. Thus if $d > d_0$ or $m_j > m_{0j}$ for some j, then the same phenomenon will occur. Indeed this problem seems to have nothing specially to do with the use of $C_{q,r}$. The second reason is as follows. The use of the tests mentioned is required in order to determine q, r, or d or K. However, instead of proceeding via tests one might set out to directly estimate q, r, or d or K, by forming say

$$\log \det \hat{G}_{qr} + \{\dim (C_{q,r}) \log N\}/N, \tag{2}$$

where $\hat{G}_{qr}$ is the maximum likelihood estimate of G, given q, r, and $\dim (C_{q,r})$ is the dimension of the manifold. Analogous quantities may be formed for C_d or C_K. Under rather general conditions it may be shown that the estimate $(\hat{q}, \hat{r})$ obtained by minimizing (2), subject to $q \leqslant Q$, $r \leqslant R, Q, R$ fixed *a priori*, will converge almost surely to q_0, r_0. The proof of this kind of proposition is long and fairly difficult and will have to be given elsewhere. Thus the difficulty in finding the asymptotic distribution of the likelihood ratio tests does not seem so important.

In favour of the procedure based on the C_K, Tuan (1978) refers to a method introduced by Akaike (1976) for a first determination of the m_j and a first estimate of the parameter point in C_K, given the m_j. That method is not designed to be consistent since a true set of m_j is not postulated. The method is computationally cheaper than a full examination of the C_K but, of course, is also less than fully efficient. A first consistent estimation procedure for the $C_{q,r}$ is also available (Hannan, 1975). Experience suggests that this, inefficient, procedure does not work well unless the sample is large but the same may be true of Akaike's (1976) procedure. The smallest sample size used in simulations there, for a low order system with $p = 2$, is 700.

The problem of constructing a good algorithm for any one of the sequences C_d, C_K and $C_{q,r}$ is considerable. For C_d it manifests itself in the possibly large number of coordinate systems. Full examination of one of them, say C_0, should locate the maximum of the likelihood, for the same kind of reason as was used in relation to $C_{q,r}$, at the end of the first paragraph. Nevertheless this examination would be difficult if that coordinate system was inappropriate, so that the optimum point lay near the boundary. Since the C_K for $m_j = d$ make up C_d it is evident that the problem for the C_K is essentially the same as for C_d. For $C_{q,r}$ the problem may manifest itself by near failure of the rank condition. It is probably for these kinds of reasons that there has been little use so far of these systems for $p > 1$, in the unconstrained case. The dimension of C_d, namely $2pd - \frac{1}{2}p(p-1)$, may be large and the topological structure of C_d is complicated, so that the estimation problem is very difficult unless prior constraints are imposed that confine the possibilities to a manageable set.

References

Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion. In *Systems Identification: Advances and Case Studies*, Eds R. K. Mehra and D. G. Lainiotis, pp. 27–96. New York: Academic Press.

Deistler, M., Dunsmuir, W. T. M. & Hannan, E. J. (1978). Vector linear time series models: corrections and extensions. *Adv. Appl. Prob.* **10**, 360–72.

Hannan, E. J. (1969). The identification of vector mixed autoregressive moving average system. *Biometrika* **56**, 222–5.

Hannan, E. J. (1971). The identification problem for multiple equation systems with moving average errors. *Econometrica* **39**, 751–66.

Hannan, E. J. (1975). The estimation of ARMA models. *Ann. Statist.* **3**, 975–81.

Hannan, E. J. (1979). The statistical theory of linear systems. In *Developments in Statistics*, Vol. II, Ed. P. R. Krishnaiah, pp. 83–121. New York: Academic Press.

Tuan, P.-D. (1978). On the fitting of multivariate processes of the autoregressive-moving average type. *Biometrika* **65**, 99–107.

[*Received November* 1978. *Revised April* 1979]