SHIBATA, R. (1981). An optimal autoregressive spectral estimate. *Ann. Statist.* **9** 300–306.

SOLO, V. (1984). The exact likelihood for a multivariate ARMA model. *J. Multivariate Anal.* **15** 164–173.

TJØSTHEIM, D. (1986). Some doubly stochastic time series models. *J. Time Ser. Anal.* **7** 51–72.

TJØSTHEIM, D. and PAULSEN, J. (1983). Bias of some commonly-used time series estimates. *Biometrika*, **70** 389–399.

WHITTLE, P. (1963a). On the fitting of multivariate autoregressions and the approximate canonical factorisation of a spectral density matrix. *Biometrika* **50** 129–134.

Comment

R. J. Bhansali

I would like to congratulate Ted Hannan on a masterly survey of the current state of the art for fitting multivariate autoregressive moving average models, ARMA(p, q). Hannan is quite correct in emphasizing that there may not be a true (p, q) and thus a fitted ARMA model is at best thought of as an approximation to the generating structure of the observed time series. The question then arises: what are the properties of the order selected by minimizing AIC when viewed in this light rather than as an estimator of an underlying "true" order? The work of Shibata (1980) would suggest that the order selected by AIC is such that the one-step mean square error of prediction is minimized within the class of all order selection procedures. However, the more-than-one-step mean square error of prediction may not be minimized (see also Whittle, 1963b, page 36). Indeed, for autoregressive model fitting, Findley (1983) has advocated that a different order should be selected for each forecast lead, and he has suggested that a criterion introduced by Shibata (1980, page 163) may be used for this purpose. However, a justification for introducing this criterion has not been given. A related but different criterion is suggested by the work of Hannan and Rissanen (1982).

As has already been noted by Franke (1985a) and Chen (1985), at the second stage of the Hannan-Rissanen procedure for ARMA model selection, "autoregressive" estimates of the coefficients $b(u)$, say, in the moving average representation of a univariate stationary nondeterministic process $\{x_t\}$ are obtained as

$$\hat{b}_h(u) = \hat{c}_h(u)/\hat{c}_h(0), \quad u = 0, 1, \cdots,$$

where

$$\hat{c}_h(u) = \sum_{j=0}^{h} \hat{a}_h(j) R^{(T)}(u + j)$$

provides the corresponding estimator of the cross-covariance $c(u)$ say, between x_{t+u} and the linear innovations ε_t; the $\hat{a}_h(j)$ are the hth order "Yule-Walker" estimates of the autoregressive coefficients;

$$R^{(T)}(u) = T^{-1} \sum_{t=1}^{T-u} x_t x_{t+u}, \quad u = 0, 1, \cdots,$$

is a "positive definite" estimator of the covariance function of $\{x_t\}$; and $x_1, \cdots, x_T$ denotes an observed realization of length T of $\{x_t\}$.

Now, if the complete past $\{x_t, t \le 0\}$, say of $\{x_t\}$ is known, the s step mean square error of prediction is given by

$$V(s) = \sigma^2 \sum_{j=0}^{s-1} b^2(j), \quad s = 1, 2, \cdots,$$

where $\sigma^2 = c(0)$ is the variance of ε_t.

Bhansali (1978) and Lewis and Reinsel (1985) consider the effect on the mean square error of prediction of estimating the prediction constants by fitting an autoregressive model of order h, when h is a function of T and tends to infinity simultaneously with it. It is clear from their work that if $o_p(h^{1/2}/T^{1/2})$ terms are ignored and certain additional regularity conditions are satisfied, the resulting mean square error of prediction may be approximated by

$$L(s) = V(s)\left(1 + \frac{h}{T}\right).$$

On adopting an argument similar to that used by Akaike (1970) for deriving his FPE criterion, which as discussed by Bhansali (1986) is closely related to the argument used for deriving AIC, one may consider

R. J. Bhansali is Senior Lecturer, Department of Statistics and Computational Mathematics, University of Liverpool, Liverpool L69 3BX, England.