Biometrika (1984), **71**, 1, *pp.* 207–11
Printed in Great Britain

Duality and other properties of multiplicative seasonal autoregressive-moving average models

By A. I. McLEOD

Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada

Summary

Model duality is defined between four models referred to as the primal, the dual, the autoregressive adjoint and the moving average adjoint. A duality theorem which generalizes the results of Box & Pierce (1970) and Pierce (1970) is presented. Applications of this duality theorem to autoregressive-moving average models and multiplicative seasonal autoregressive-moving average models are discussed. These applications include:

(i) a convenient method for calculating the covariance matrix of the estimated parameters;
(ii) convenient formulae for the variances of the residual autocorrelations;
(iii) the distribution of the inverse partial autocorrelations.

Finally, a useful approximation to the covariance determinant of multiplicative seasonal models is derived.

Some key words: Covariance determinant; Inverse partial autocorrelations; Parameter estimation; Residual autocorrelation.

1. Introduction

The multiplicative seasonal autoregressive-moving average $(p, q)(p_s, q_s)_s$ model,

$$\Phi(B^s)\,\phi(B)\,z_t = \Theta(B^s)\,\theta(B)\,a_t, \tag{1}$$

where

$$\phi(B) = 1-\phi_1 B-\ldots-\phi_p B^p, \quad \theta(B) = 1-\theta_1 B-\ldots-\theta_q B^q,$$

$$\Phi(B^s) = 1-\Phi_1 B^s-\ldots-\Phi_{p_s} B^{sp_s}, \quad \Theta(B^s) = 1-\Theta_1 B^s-\ldots-\Theta_{q_s} B^{sq_s},$$

B is the backshift operator, s the seasonal period and a_t a sequence of independent normal variables with mean zero and variance σ^2, was developed by Box & Jenkins (1976). The regular autoregressive-moving average (p, q) model is obtained by taking $p_s = q_s = 0$ in (1). Alternatively, the $(p, q)(p_s, q_s)_s$ model may be considered as a special case of the (p^*, q^*) model by taking

$$p^* = p+sp_s, \quad q^* = q+sq_s, \quad \phi^*(B) = \Phi(B^s)\,\phi(B), \quad \theta^*(B) = \Theta(B^s)\,\theta(B).$$

It will be assumed that the model is stationary, invertible and not redundant, so that the polynomials $\phi^*(B)$ and $\theta^*(B)$ have no roots in common and all roots are outside the unit circle.

2. Duality

The concept of duality in autoregressive-moving average models has proved useful to various authors. Theorem 1 below generalizes model duality results of Pierce (1970) and Box & Pierce (1970).

Consider the models

$$\Phi(B^s)\,\phi(B)\,z_t = \Theta(B^s)\,\theta(B)\,a_t, \quad \Theta(B^s)\,\theta(B)\,y_t = \Phi(B^s)\,\phi(B)\,a_t, \tag{2}$$

$$\Xi(B^s)\,\xi(B)\,x_t = a_t, \quad w_t = \Xi(B^s)\,\xi(B)\,a_t, \tag{3}$$

where

$$\xi(B) = 1-\Sigma\,\xi_i B^i = \phi(B)\,\theta(B), \quad \Xi(B^s) = 1-\Sigma\,\Xi_i B^{si} = \Phi(B^s)\,\Theta(B^s).$$

These four models may be referred to as the primal, the dual, the autoregressive adjoint and the moving average adjoint respectively. Note that all series are generated from the same innovation series. Let

$$\beta = (\phi_1, \ldots, \phi_p,\ \theta_1, \ldots, \theta_q,\ \Phi_1, \ldots, \Phi_{p_s}, \Theta_1, \ldots, \Theta_{q_s}),$$

$$\alpha = (\xi_1, \ldots, \xi_{p+q},\ \Xi_1, \ldots, \Xi_{p_s+q_s}).$$

Given a series of n observations from each model, let $\hat{\beta}_z$, $\hat{\beta}_y$, $\hat{\alpha}_x$ and $\hat{\alpha}_w$ denote corresponding efficient approximate maximum likelihood estimates and denote the corresponding residuals by $\hat{a}_{z,t}$, $\hat{a}_{y,t}$, $\hat{a}_{x,t}$ and $\hat{a}_{w,t}$.

THEOREM 1. *Apart from a quantity which is $O_p(1/n)$,*

$$\hat{a}_{z,t} = \hat{a}_{y,t} = \hat{a}_{x,t} = \hat{a}_{w,t}, \tag{4}$$

$$\hat{\alpha}_w - \alpha = -(\hat{\alpha}_x - \alpha) = -J'(\hat{\beta}_y - \beta) = J'(\hat{\beta}_z - \beta), \tag{5}$$

where

$$J = (\theta_{i-j} \ \vdots\ -\phi_{i-j} \ \vdots\ \Theta_{i-j} \ \vdots\ -\Phi_{i-j}) \tag{6}$$

and the (i, j)th entry in each partitioned matrix is indicated, and ϕ_i, θ_i, Φ_i and Θ_i are defined more generally for any integer i as minus the coefficient of B^i in their respective polynomials $\phi(B)$, $\theta(B)$, $\Phi(B)$ and $\Theta(B)$.

Note that Pierce's theorem (1970) corresponds to (5) with $q = p_s = q_s = 0$, in which case J is minus the identity matrix. The duality result of Box & Pierce (1970) corresponds to (4) with $p_s = q_s = 0$.

Proof. The proof is given only for $p_s = q_s = 0$ since the method extends directly to the multiplicative seasonal case. Also, let $\sigma^2 = 1$.

Auxiliary processes v_t and u_t are defined by $\phi(B)\,v_t = -a_t$ and $\theta(B)\,u_t = a_t$. Let $A_t = (v_{t-1}, \ldots, v_{t-p}, u_{t-1}, \ldots, u_{t-q})'$ and $D_t = -(x_{t-1}, \ldots, x_{t-p-q})'$. If we use the technique of Box & Jenkins (1976, p. 240) it is easily shown that the information matrix of β in the z_t process is $I_{z,\beta} = E(A_t A_t')$. Similarly, $I_{y,\beta} = E(A_t A_t')$ and $I_{x,\alpha} = E(D_t D_t') = I_{w,\alpha}$.

Consider the x_t-process. Then $\partial a_t/\partial\alpha_i = D_{t,i}$, $\partial a_t/\partial\phi_i = A_{t,i}$ and $\partial a_t/\partial\theta_i = -A_{t,p+i}$. Also, $\partial\alpha_j/\partial\phi_i = -\theta_{j-i}$ and $\partial\alpha_j/\partial\theta_i = -\phi_{j-i}$. It follows that $I_{z,\beta} = J I_{x,\alpha} J'$.

From Lemma 2 of McLeod (1978), it may be shown, to $O_p(1/n)$, that

$$\hat{\beta}_z - \beta = I_{z,\beta}^{-1} s_{z,\beta}, \quad \hat{\beta}_y - \beta = I_{y,\beta}^{-1} s_{y,\beta}, \quad \hat{\alpha}_x - \alpha = I_{x,\alpha}^{-1} s_{x,\alpha}, \quad \hat{\alpha}_w - \alpha = I_{w,\alpha}^{-1} s_{w,\alpha},$$

where $s_{z,\beta} = -\Sigma\, a_t A_t/n$, $s_{y,\beta} = -s_{z,\beta}$, $s_{x,\alpha} = -\Sigma\, a_t D_t/n$ and $s_{w,\alpha} = -s_{x,\alpha}$. Equation (5) follows.

On expanding the residuals in a Taylor series, we have that

$$\hat{a}_{z,t} = a_t + (\hat{\beta}_z - \beta)\,A_t, \quad \hat{a}_{y,t} = a_t - (\hat{\beta}_y - \beta)\,A_t,$$

$$\hat{a}_{x,t} = a_t + (\hat{\alpha}_x - \alpha)\,D_t, \quad \hat{a}_{w,t} = a_t - (\hat{\alpha}_w - \alpha)\,D_t.$$

Equality of the residuals follows.

3. Covariance matrix calculation

As a simple application, consider the calculation of the large-sample covariance matrix of $\hat{\beta}$. For the autoregression, $\phi(B)x_t = a_t$, the covariance matrix of the estimated parameters may be obtained using a result given by Pagano (1973),

$$\operatorname{cov}(\hat{\alpha}) = n^{-1}\left(\sum_{k=1}^{\min(i,j)} \phi_{i-k}\phi_{j-k} - \phi_{p+k-i}\phi_{p+k-j}\right). \tag{7}$$

Hence for (p,q) models, $\operatorname{cov}(\hat{\beta}) = (J^{-1})' \operatorname{cov}(\hat{\alpha}) J^{-1}$. This method is much simpler than that recently proposed by Godolphin & Unwin (1983). The multiplicative $(p,q)(p_s,q_s)_s$ model may be treated as a special case of the (p^*,q^*) model or if s is large enough, say $s \geqslant 12$, then the covariance matrix is approximately block diagonal corresponding to (p,q) and (p_s,q_s) models.

4. Convenient formulae

Simulations reported by Ansley & Newbold (1979) suggest that diagnostic checks using the residual autocorrelations, $\hat{r}_a(1)$ and $\hat{r}_a(s)$, may often be useful in detecting model inadequacy. In an adequate model, the observed values of these residual autocorrelations should not differ from zero by more than about two standard deviations 95 per cent of the time.

Theorem 2. *Provided that s is not too small,*

$$n \operatorname{var}\{\hat{r}_a(1)\} \doteqdot (\phi_p \theta_q)^2, \tag{8}$$

$$n \operatorname{var}\{\hat{r}_a(s)\} \doteqdot (\Phi_{p_s}\Theta_{q_s})^2. \tag{9}$$

For a (p,q) model, (8) holds exactly asymptotically.

Proof. Due to duality, it is sufficient to prove this result for the $(p,0)(p_s,0)_s$ model, $\Phi(B^s)\phi(B)z_t = a_t$. If this model is reparameterized in terms of g_i $(i = 1,\ldots,p)$ and G_j $(j = 1,\ldots,p_s)$, where $\phi(B) = \Pi(1-g_i B)$ and $\Phi(B) = \Pi(1-G_j B)$, then the joint information for g_i and G_j can be shown to be

$$I(g_i, G_j) = g_i^{s-1} G_j/(1-g_i^{s-1} G_j). \tag{10}$$

It follows that provided s is large enough, the information matrix for $(\phi_1,\ldots,\phi_p,$ $\Phi_1,\ldots,\Phi_{p_s})$ is approximately $\operatorname{diag}(I_p, I_{p_s})$, where I_p and I_{p_s} are the information matrices corresponding to autoregressions $\phi(B)v_t = a_t$ and $\Phi(B)V_t = a_t$. If we use (7), the leading entries in the inverse matrix are respectively $1-\phi_p^2$ and $1-\Phi_{p_s}^2$. The formulae now follow directly from McLeod (1978, (44)).

5. Inverse partial autocorrelations

Inverse correlations in a primal (p,q) model may be defined as the correlations in the corresponding dual (Cleveland, 1972; Hipel, McLeod & Lennox, 1977). Inverse correlations can be estimated by using a high-order autoregressive approximation to the primal model, $\Pi(B)z_t = a_t$, where $\Pi(B) = 1-\Pi_1 B - \ldots - \Pi_k B^k$. After fitting this model, the inverse autocorrelations are then estimated by,

$$\operatorname{ri}(l) = (-\hat{\Pi}_l + \Sigma\hat{\Pi}_{l+i}\hat{\Pi}_i)/(1+\Sigma\hat{\Pi}_i^2) \quad (l = 1,\ldots,k). \tag{11}$$

The inverse partial autocorrelation $\theta_{l,l}$ $(l = 1,\ldots,k)$ is estimated using $\operatorname{ri}(l)$ in place of $r(l)$ $(l = 1,\ldots,k)$. The usefulness of inverse correlations in model identification is greatly enhanced by knowledge of their distribution. Hosking (1980) has shown that the distribution of the sample inverse autocorrelations is equivalent to that of the sample autocorrelations in the dual model.

THEOREM 3. *If* z_t $(t = 1, \ldots, n)$ *is generated by a* $(0,q)$ *model,* $\sqrt{n}\hat{\theta}_{k,k}$ $(k = q+1, q+2, \ldots)$ *are asymptotically independent normal with mean zero and variance one.*

Proof. The estimate, $\hat{\theta}_{k,k}$, corresponds to an estimate of the coefficient $\theta_{k,k}$ in the model, $z_t = a_t - \theta_{k,1} a_{t-1} - \ldots - \theta_{k,k} a_{t-k}$, using the method of Durbin (1960). Since, as shown by Durbin, this is asymptotically equivalent to the maximum likelihood estimate, Theorem 3 follows directly from Theorem 1 and the well-known result of Quenouille (1949).

6. COVARIANCE DETERMINANT

The asymptotic determinant, $M(p,q) = \lim |\sigma^2 \Gamma_n^{-1}|$, where Γ_n is the covariance matrix of n successive observations, z_t $(t = 1, \ldots, n)$, from a (p,q) model, can be easily calculated for small values of p and q (McLeod, 1977). For $(p,q)(p_s, q_s)_s$ models, the asymptotic determinant is given by $M(p,q,p_s,q_s,s) = M(p^*, q^*)$.

THEOREM 4. *We have that*

$$M(p,q,p_s,q_s,s) \simeq M(p,q)\{M(p_s,q_s)\}^s. \tag{12}$$

The relative error in the approximation (12) *is* $O(r^s)$, *where* $0 \leqslant r < 1$.

Proof. From the duality property noted by Finch (1960), it suffices to prove the result for $(p,0)(p_s,0)_s$ models. Let $\phi(B) = \Pi(1-g_i B)$ and $\Phi(B^s) = \Pi(1-G_j B^s)$. Then from the result of Finch (1960, (6)),

$$M(p,0) = |g_1 \ldots g_p|^{2p} \prod_{i,j} \frac{|1-g_i\bar{g}_j|}{|g_i\bar{g}_j|},$$

$$M(p_s,0) = |G_1 \ldots G_{p_s}|^{2p_s} \prod_{i,j} \frac{|1-G_i\bar{G}_j|}{|G_i\bar{G}_j|}.$$

Similarly

$$M(p^*,0) = |g_1^* \ldots g_{p^*}^*|^{2p^*} \prod_{i,j} \frac{|1-g_i^*\bar{g}_j^*|}{|g_i^*\bar{g}_j^*|},$$

where the g^*'s denote values of the form g_i $(i = 1, \ldots, p)$ or $G_j^{1/s}\omega^k$ $(j = 1, \ldots, p_s;$ $k = 0, \ldots, s-1)$, where ω is a complex sth root of unity. Simplifying, using the identity

$$\prod_{k=0}^{s-1} (1-\zeta\omega^k) = 1-\zeta^s,$$

we have that $M(p,0,p_s,0,s) = M(p,0)\{M(p_s,0)\}^s T$ where

$$T = \prod_{i,j} |1-g_i^s \bar{G}_j|^2.$$

The approximation (12) follows from the fact that $|g_i| < 1$ $(i = 1, \ldots, p)$.

Fortran subroutines to evaluate $M(p,q)$ and the approximate maximum likelihood estimator (McLeod, 1977) using the approximation (12) are available (McLeod & Holanda Sales, 1983).

Helpful suggestions from a referee are acknowledged.

References

ANSLEY, C. F. & NEWBOLD, P. (1979). On the finite sample distribution of residual autocorrelations in autoregressive-moving average models. *Biometrika* **66**, 547–53.

BOX, G. E. P. & JENKINS, G. M. (1976). *Time Series Analysis: Forecasting and Control,* 2nd edition. San Francisco: Holden-Day.

BOX, G. E. P. & PIERCE, D. A. (1970). Distribution of the residual autocorrelations in autoregressive-integrated moving average time series models. *J. Am. Statist. Assoc.* **65**, 1509–26.

CLEVELAND, W. S. (1972). The inverse autocorrelations of a time series and their applications. *Technometrics* **14**, 277–93.

DURBIN, J. (1960). Efficient estimation of the parameters in moving average models. *Biometrika* **46**, 306–16.

FINCH, P. D. (1960). On the covariance determinants of moving average and autoregressive models. *Biometrika* **47**, 194–6.

GODOLPHIN, E. J. & UNWIN, J. M. (1983). Evaluation of the covariance matrix for the maximum likelihood estimator of a Gaussian autoregressive-moving average process. *Biometrika* **70**, 279–84.

HIPEL, K. W., MCLEOD, A. I. & LENNOX, W. C. (1977). Advances in Box–Jenkins modelling. *Water Resources Res.* **13**, 567–86.

HOSKING, J. R. M. (1980). The asymptotic distribution of the sample inverse autocorrelations of an autoregressive-moving average process. *Biometrika* **67**, 223–6.

MCLEOD, A. I. (1977). Improved Box–Jenkins estimators. *Biometrika* **64**, 531–4.

MCLEOD, A. I. (1978). On the distribution of residual autocorrelations in Box–Jenkins models. *J. R. Statist. Soc.* B **40**, 296–302.

MCLEOD, A. I. & HOLANDA SALES, P. R. (1983). An algorithm for approximate likelihood calculation of ARMA and seasonal ARMA models. *Appl. Statist.* **32**, 211–23.

PAGANO, M. (1973). When is an autoregressive process stationary? *Comm. Statist.* **1**, 533–44.

PIERCE, D. A. (1970). A duality between autoregressive and moving average processes concerning their least squares parameter estimates. *Ann. Math. Statist.* **41**, 722–6.

QUENOUILLE, M. H. (1949). Approximate tests of correlation in time-series. *J. R. Statist. Soc.* B **11**, 68–84.

[*Received September* 1982. *Revised May* 1983]