Faculteit der Economische Wetenschappen en Econometrie

Serie Research Memoranda

ON THE CRITERION FUNCTION FOR ARMA ESTIMATION

D.S.G. Pollock

Research Memorandum 1991-74
December 1991

vrije Universiteit *amsterdam*

ON THE CRITERION FUNCTION FOR ARMA ESTIMATION

by D.S.G. POLLOCK
Vrije Universiteit, Amsterdam,
The Netherlands.

This paper investigates some of the criteria functions which may be used in deriving estimates of the parameters of autoregressive moving-average models. The object is to show how we can guarantee that the estimates will fulfil the conditions of stationarity and invertibility.

1. Introduction

In this paper, we seek to elucidate some of the criteria functions which are employed in estimating autoregressive moving-average (ARMA) models. At the same time, we look for ways of ensuring that the estimates of the parameters of an ARMA process will fulfil the conditions of stationarity and invertibility which we shall assume to be characteristic of the process generating the data.

The problematic nature of the least-squares estimators of moving-average models, which are liable to violate the conditions of invertibility when the sample is small, was demonstrated in a widely read but unpublished paper by Kang [11]. The problem was analysed in the context of a first-order moving-average (MA) model by Osborn [16] who derived expressions for the expected values of various criterion functions. A similar analysis was conducted by Davidson [7] who used the method of Monte-Carlo experiments.

There is also a widespread awareness of the analogous problems with least-squares estimators of autoregressive (AR) models, which are liable to violate the conditions of stationarity; and it is known that one way to avoid the problem is to use the Yule–Walker method of estimation—see, for example, Pagano [17].

It can be show that the exact maximum-likelihood (ML) estimates of ARMA models are bound to fulfil the conditions of stationarity regardless of the size of the sample from which they are derived; and in some quarters it has been argued that they should be used in preference to any other estimators when the sample is small. However, the ML estimates are laborious to compute; and they are inappropriate to real-time signal-processing applications where the data arrives rapidly and in abundance.

The question of the invertibility of the exact ML estimates of an MA model is more complicated. The value of the likelihood function is uniquely determined by the value of the dispersion matrix which contains the estimates of the autocovariances. If a condition is imposed that none of the roots of the MA operator fall inside the unit circle, then the autocovariances correspond to a unique MA process.

However, the condition of invertibility is more stringent than the condition of uniqueness since it requires, in addition, that none of the roots should lie on the perimeter of the unit circle. It has been shown by Cryer and Ledolter [5], in the case of a first-order MA process, that there is indeed a finite probability that the ML estimator will deliver a root which lies on the perimeter. Nevertheless, Anderson and Takemura [1] have shown that this probability tends to zero as the sample size increases. The conclusion is that, at least asymptotically, the exact ML estimates are bound to fulfil the conditions of invertibility and stationarity when the data come from a stationary and invertible process.

In the limit, as the sample size becomes indefinitely large, the estimates derived from a wide variety of criteria—including the least-squares criteria—will tend to the values generated by the exact ML estimator. At the same time, the criterion function of the ML estimator tends to a form which is drastically simplified.

The thought which inspires this paper is as follows. Imagine that we are able to discern, within the simplified asymptotic form of the ML criterion function, the characteristics which guarantee stationarity and invertibility. If we can mimic these characteristics within the criterion function of a finite-sample estimator, then we should be able to find estimates of a simpler sort than the exact ML estimates which also fulfil the conditions of invertibility and stationarity.

The paper also pursues a second theme. ARMA models which are defined in terms of infinite realisations of the stochastic sequences obey a simple algebra of rational functions and infinite series. We loose this simplicity when we are constrained to work with sequences of finite length and with polynomials of finite degree for which there exist matrix representations. In order to attribute the appropriate asymptotic characteristics to our finite-sample criterion function, we need to establish the appropriate correspondence between the infinite-order polynomial algebra and the finite-order matrix algebra. Throughout the paper, we shall be examining the matrix analogues of the polynomial algebra.

2. The Polynomial Algebra of the ARMA Model

Let us begin by representing the ARMA model by the equation

$$\alpha(z)y(z) = \mu(z)\varepsilon(z), \tag{1}$$

where

$$\begin{aligned} \alpha(z) &= 1 + \alpha_1 z + \cdots + \alpha_p z^p, \\ \mu(z) &= 1 + \mu_1 z + \cdots + \mu_q z^q, \\ y(z) &= \{y_{-p}z^{-p} + \cdots + y_0 + y_1 z + \cdots + y_{n-1}z^{n-1} + \cdots\} \quad \text{and} \\ \varepsilon(z) &= \{\varepsilon_{-q}z^{-q} + \cdots + \varepsilon_0 + \varepsilon_1 z + \cdots + \varepsilon_{n-1}z^{n-1} + \cdots\}. \end{aligned} \tag{2}$$

The sample period runs from $t = 0$ to $t = n-1$. We may allow the unknown elements $y_t, \varepsilon_t; t > n-1$, which lie beyond the sample period, to be replaced by zeros.

The polynomials $\alpha(z)$ and $\mu(z)$ are subject to the condition that they have no factors in common. The condition of invertibility is the requirement that the roots of $\mu(z) = 0$ must lie outside the unit circle, whilst the condition of stationarity is the requirement that those of $\alpha(z) = 0$ must lie outside the unit circle.

The autocovariance generating function, which is defined by

$$\begin{aligned}\gamma(z) &= \frac{\mu(z)\mu(z^{-1})}{\alpha(z)\alpha(z^{-1})} \\ &= \{\gamma_0 + \gamma_1(z + z^{-1}) + \gamma_2(z^2 + z^{-2}) + \cdots\},\end{aligned} \tag{3}$$

is a self-reciprocal polynomial with the property that $\gamma(z) = \gamma(z^{-1})$. On setting $z = e^{i\omega}$, this becomes the spectral density function of the ARMA process.

Some useful products are available:

(i) $y(z) = \mu(z)\alpha^{-1}(z)\varepsilon(z)$: forecast function,

(ii) $\varepsilon(z) = \alpha(z)\mu^{-1}(z)y(z)$: prediction error,

(iii) $\partial\varepsilon(z)/\partial\mu_j = -z^j\alpha(z)\mu^{-2}(z)y(z)$: derivatives,

$\partial\varepsilon(z)/\partial\alpha_j = z^j\mu^{-1}(z)y(z)$: derivatives,

(iv) $\gamma(z) = \sigma_\varepsilon^2\mu(z)\mu(z^{-1})$: the Cramér–Wold factorisation of the autocovariance generating function, (assuming that $\alpha(z) = 1$),

(v) $\gamma(z) = \sigma_\varepsilon^2\alpha^{-1}(z)\alpha^{-1}(z^{-1})$: the Yule–Walker factorisation of the autocovariance generating function, (assuming that $\mu(z) = 1$).

An algorithm for obtaining the Cramér–Wold factorisation has been described by Wilson [23] and his implementation of it is to be found amongst the programs described by Box and Jenkins [2, Program 2]. The algorithm has also been implemented by Laurie [13], [14]. The Yule–Walker factorisation is, of course, readily available.

Commentary. In estimating the ARMA model from a data series $\{y_0, \ldots, y_{n-1}\}$, it is common to set the presample elements $\{y_{-p}, \ldots, y_{-1}\}$ and $\{\varepsilon_{-q}, \ldots, \varepsilon_{-1}\}$ to zeros. In general, when $\alpha(z)$ and $\mu(z)$ are specified arbitrarily, the equality of (1) can be maintained only by allowing the residual polynomial $\varepsilon(z)$ to be replaced by an infinite series. There are exceptions.

First, if $\mu(z) = 1$, then the equality can be maintained by allowing the residual polynomial to take the form of $\varepsilon(z) = \varepsilon_0 + \varepsilon_1 z + \cdots + \varepsilon_{n-1+p}z^{n-1+p}$, which is a polynomial of degree $n - 1 + p$.

Secondly, if the polynomial argument z^j is nilpotent of degree n in the index j, such that $z^j = 0$ for all $j \geq n$, then all polynomial products are of degree $n-1$ at most, and the equality may be maintained without the degree of $\varepsilon(z)$ exceeding $n-1$. Making z^j nilpotent of degree n will enable us to construct a correspondence between the algebra of the polynomials of degree $n-1$ and the algebra of the class of $n \times n$ lower-triangular Toeplitz matrices.

Thirdly, if the polynomial argument z^j is an n-periodic function of the index j, such that $z^{j+n} = z^j$ for all j, then, again, all polynomial products are of degree $n-1$ at most. Making z^j an n-periodic function, enables us to construct a correspondence between the algebra of the polynomials of degree $n-1$ and the algebra of the class of $n \times n$ circulant Toeplitz matrices.

A polynomial of degree $n-1$ is completely specified by the values which it assumes at n equally spaced points on the circumference of the unit circle in the complex plane which are $e^{i\omega_j}; j = 0, \ldots, n-1$, where $\omega_j = 2\pi j/n$. In particular, a product $\gamma(z) = \alpha(z)\beta(z)$ of two periodic polynomials is completely specified by $\gamma(e^{i\omega_j}) = \alpha(e^{i\omega_j})\beta(e^{i\omega_j}); j = 0, \ldots, n-1$. Therefore, when the polynomials have an n-periodic argument, the time-consuming business of polynomial multiplication can be circumvented by performing an equivalent but a much simpler set of operations at the frequency points ω_j.

3. A Least-Squares Criterion Function

Using the polynomial algebra, we can define a criterion function for ARMA estimation which takes the form of

$$\begin{aligned} S &= \frac{1}{2\pi i}\oint \varepsilon(z)\varepsilon(z^{-1})\frac{dz}{z} \\ &= \frac{1}{2\pi i}\oint y(z)y(z^{-1})\frac{\alpha(z)\alpha(z^{-1})}{\mu(z)\mu(z^{-1})}\frac{dz}{z} \\ &= \frac{n}{2\pi i}\oint \frac{c(z)}{\gamma(z)}\frac{dz}{z}. \end{aligned} \tag{4}$$

Here

$$\begin{aligned} c(z) &= \frac{y(z)y(z^{-1})}{n} \\ &= c_0 + c_1(z + z^{-1}) + \cdots + c_{n-1}(z^{n-1} + z^{1-n}) \end{aligned} \tag{5}$$

is the empirical autocovariance generating function. When $z = \exp\{-i2\pi j/n\}$, this becomes the periodogram which is defined on the points $j = 0, \ldots, n/2$ when n is even and on the points $j = 0, \ldots, (n-1)/2$ if n is odd.

The value of S is nothing more than the coefficient associated with z^0 in the Laurent expansion of $\varepsilon(z)\varepsilon(z^{-1})$. This is demonstrated in the appendix. It follows that, if the coefficients of $\alpha(z)$ and $\mu(z)$ were to assume their true values and if the presample elements $\{y_{-p},\ldots,y_{-1}\}$ were incorporated in $y(z)$, then S would be equal to the sum of squares of the disturbances $\{\varepsilon_{-q},\ldots,\varepsilon_{n-1}\}$.

In the case of the pure AR model, the criterion function assumes the form of

$$\begin{aligned} S &= \frac{1}{2\pi i}\oint y(z)y(z^{-1})\alpha(z)\alpha(z^{-1})\frac{dz}{z} \\ &= n\sum_{j=0}^{p}\sum_{k=0}^{p}\alpha_k\alpha_j c_{|k-j|} \\ &= \sum_{t=0}^{n}\sum_{s=0}^{n} y_t y_s \lambda_{|t-s|}. \end{aligned} \tag{6}$$

Here $\lambda_{|t-s|}$ is a coefficient of the self-reciprocal polynomial

$$\begin{aligned} \lambda(z) &= \alpha(z)\alpha(z^{-1}) \\ &= \{\lambda_0 + \lambda_1(z+z^{-1}) + \cdots + \lambda_p(z^p + z^{-p})\} \end{aligned} \tag{7}$$

which corresponds to the autocovariance generating function of a synthetic pth-order MA process $y(z) = \alpha(z)\varepsilon(z)$ based on a sequence $\{\varepsilon_t\}$ of unit-variance white-noise disturbances.

In the appendix we prove the following theorem:

Theorem 1. Let $Q = \sum_{j=0}^{p}\sum_{k=0}^{p}\alpha_k\alpha_j c_{|j-k|} = \alpha' C\alpha$ where $C = [c_{|j-k|}]$ is a symmetric positive-definite Toeplitz matrix and $\alpha = \{1, \alpha_1, \ldots, \alpha_p\}$ is a vector. If α is chosen so as to minimise the value of Q, then all the roots of the equation $\alpha(z) = 1 + \alpha_1 z + \cdots + \alpha_p z^p = 0$ will lie outside the unit circle.

In fact, the values which minimise the AR criterion function are simply the Yule–Walker estimates; and thus the theorem serves to prove, in a direct way, that the estimates correspond to a stationary model.

Now let us consider the case of the pure MA model. Let the series expansion of the inverse of $\mu(z)$ be written as $\mu^{-1}(z) = \{\psi_0 + \psi_1 z + \cdots\}$. Then the criterion function in the case of the pure MA model can be expressed as

$$\begin{aligned} S &= \frac{1}{2\pi i}\oint \frac{y(z)y(z^{-1})}{\mu(z)\mu(z^{-1})}\frac{dz}{z} \\ &= n\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_k\psi_j c_{|k-j|} \\ &= \sum_{t=0}^{n}\sum_{s=0}^{n} y_t y_s \delta_{|t-s|}. \end{aligned} \tag{8}$$

Here $\delta_{|t-s|}$ is a coefficient of the self-reciprocal polynomial

$$\begin{aligned} \delta(z) &= \frac{1}{\mu(z)\mu(z^{-1})} \\ &= \{\delta_0 + \delta_1(z+z^{-1}) + \delta_2(z^2+z^{-2}) + \cdots\} \end{aligned} \tag{9}$$

which corresponds to the autocovariance generating function of a synthetic qth-order AR process $\mu(z)y(z) = \varepsilon(z)$ based on a sequence $\{\varepsilon_t\}$ of unit-variance white-noise disturbances.

In the appendix, we prove the following theorem which concerns the invertibility of the estimated MA model:

Theorem 2. Let $Q = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_k\psi_j c_{|k-j|}$ where ψ_k,ψ_j are coefficients in the expansion of the inverse polynomial $\mu^{-1}(z) = \{\psi_0 + \psi_1 z + \cdots\}$. If the coefficients of $\mu(z) = 1 + \mu_1 z + \cdots + \mu_q z^q$ are chosen so as to minimise the value of Q, then all the roots of the equation $\mu(z) = 0$ will lie outside the unit circle.

Indeed this theorem is almost self-evident. For Q will have a finite value if and only if $\sum|\psi_i| < \infty$; and for this to arise, it is necessary and sufficient that the roots $\mu(z) = 0$ lie outside the unit circle.

The MA criterion function in the form of

$$\begin{aligned} Q &= \frac{1}{n}\sum_{\tau}\sum_{t} y_t y_{t-\tau}\delta_\tau \\ &= c_0 + 2\sum_{t=1}^{n-1} c_\tau\delta_\tau \end{aligned} \tag{10}$$

has been considered by Godolphin [9]. He has devised a specialised method for minimising the function which makes use of some approximations to the derivatives $d\delta_\tau/d\mu_j$ which are based on truncated power-series expansions. However, his iterative procedure has only a linear rate of convergence; and it is not clear that the desirable properties of the criterion function survive the process of approximation.

The ARMA criterion function under (4) combines the features of the specialised AR and MA criteria functions of (6) and (8). Thus theorems 1 and 2 together serve to show that the values of $\{\alpha_1,\ldots,\alpha_p\}$ and $\{\mu_1,\ldots,\mu_q\}$ which minimise the function correspond to a model which satisfies conditions both of stationarity and invertibility.

Commentary. Whilst the criterion function presented above has the form of the asymptotic limit of the ML criterion function, it caters for the case where the polynomials $y(z)$ and $\varepsilon(z)$ are finite.

The mere fact that the criterion function has the asymptotic ML form implies that the minimising values of $\{\alpha_1,\ldots,\alpha_p\}$ and $\{\mu_1,\ldots,\mu_q\}$ will correspond to autoregressive and moving-average operators which fulfil the conditions of stationarity and invertibility respectively. The finite-sample criterion function differs from the asymptotic form only by the fact that, beyond the point $t = n-1$, the elements y_t and ε_t assume values of zero instead of taking values which have been generated by a stochastic process. This feature cannot affect the essential properties of the minimising polynomials $\alpha(z)$ and $\mu(z)$.

Although the criterion function leads to estimates which have the sought-after properties, it is not clear, at first, how it should be given a computable mathematical expression. The following sections of the paper are devoted to the task of finding such expressions.

4. Matrix Representations

Now let us look for appropriate matrix representations of the ARMA model and of the criterion function. For a start, we can see that the n realisations of the process from $t = 0$ to $t = n-1$ are comprised in the following system:

$$\begin{bmatrix} y_0 & y_{-1} & \dots & y_{1-n} \\ y_1 & y_0 & \dots & y_{2-n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n-1} & y_{n-2} & \dots & y_0 \end{bmatrix} \begin{bmatrix} 1 \\ \alpha_1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \varepsilon_0 & \varepsilon_{-1} & \dots & \varepsilon_{1-n} \\ \varepsilon_1 & \varepsilon_0 & \dots & \varepsilon_{2-n} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n-1} & \varepsilon_{n-2} & \dots & \varepsilon_0 \end{bmatrix} \begin{bmatrix} 1 \\ \mu_1 \\ \vdots \\ 0 \end{bmatrix}. \tag{11}$$

This may be represented, in summary notation, by

$$Y\alpha = \mathcal{E}\mu. \tag{12}$$

Lower-Triangular Toeplitz Matrices

We can set the presample elements above the principal diagonals in Y and $\mathcal{E}$ to zeros. Then Y and $\mathcal{E}$ are replaced by lower-triangular (LT) Toeplitz matrices $Y = Y(y)$ and $\mathcal{E} = \mathcal{E}(\varepsilon)$ which are completely characterised by their leading vectors which are given by $\mathcal{E}e_0 = \varepsilon = [\varepsilon_0, \dots, \varepsilon_{n-1}]'$ and $Ye_0 = y = [y_0, \dots, y_{n-1}]'$, where e_0 is the leading vector of the identity matrix of order n. On the same principle, we can define lower-triangular Toeplitz matrices $A = A(\alpha)$ and $M = M(\mu)$ which are characterised by their respective leading vectors $\alpha = [1, \alpha_1, \dots, \alpha_p, 0, \dots, 0]'$ and $\mu = [1, \mu_1, \dots, \mu_q, 0, \dots, 0]'$.

Lower-triangular Toeplitz matrices can be represented as polynomial functions of a matrix $L = [e_1, \dots, e_{n-1}, 0]$ which has units on the first subdiagonal and zeros elsewhere and which is formed from the identity matrix $I = [e_0, e_1, \dots, e_{n-1}]$ by deleting the leading vector and appending a zero vector to the end of the array. Thus the matrix $A = A(\alpha)$ can be written as

$$\begin{aligned} A &= \alpha(L) \\ &= I + \alpha_1 L + \dots + \alpha_p L^p. \end{aligned} \tag{13}$$

We may note that L^j is nilpotent of degree n in the index j such that $L^j = 0$ for $j \geq n$.

In some respects, the algebra of the matrices resembles the ordinary algebra of polynomials with a real or complex argument:

(i) The matrices commute such that $AY = YA \rightarrow$ $Ay = (AY)e_0 = (YA)e_0 = Y\alpha$,

(ii) If A, Y are LT Toeplitz, then so is $AY = YA$,

(iii) If A is LT Toeplitz, then so is A^{-1}. In particular, $A^{-1}e_0 = [\omega_0, \omega_1, \ldots, \omega_{n-1}]'$ has the leading coefficients of the expansion of $\alpha^{-1}(z)$ as its elements.

In terms of the ARMA model, we find that

(i) $Y = A^{-1}M\mathcal{E} \rightarrow y = A^{-1}M\varepsilon$: forecast function,

(ii) $\mathcal{E} = M^{-1}AY \rightarrow \varepsilon = M^{-1}Ay$: prediction error,

(iii) $\partial\varepsilon/\partial\mu_j = -AM^{-2}Ye_j, \quad \partial\varepsilon/\partial\alpha_j = M^{-1}Ye_j$: derivatives,

(iv) $G = M'M$: this is not a Toeplitz matrix and the Cramér–Wold factorisation does not apply.

Now consider the matrix representation of the criterion function. As an approximation of the function S of (4), we have

$$S^z = e_0'(Y'A'M'^{-1}M^{-1}AY)e_0 = y'A'M'^{-1}M^{-1}Ay. \tag{14}$$

This is just the coefficient associated with $I = L^0$ in the expansion of

$$\begin{aligned}\varepsilon(L')\varepsilon(L) &= y(L')\alpha(L')\mu^{-1}(L')\mu^{-1}(L)\alpha(L)y(L)\\ &= \frac{y(L')\alpha(L')\alpha(L)y(L)}{\mu(L')\mu(L)},\end{aligned} \tag{15}$$

where $L' = [0, e_0, \ldots, e_{n-2}]$. We can afford to write this function in rational form on account of the commutativity in multiplication of the LT Toeplitz matrices.

The vectors α and μ which minimise the function S^z contain what Box and Jenkins [2] described as the conditional least-squares estimators. These estimators are not guaranteed to satisfy the conditions of stationarity and invertibility.

Consider the specialisation of the criterion function to the case of a pure AR model. Scaling by n^{-1} gives

$$\begin{aligned}\frac{1}{n}S^z &= \frac{1}{n}y'A'Ay\\ &= \frac{1}{n}\alpha'Y'Y\alpha.\end{aligned} \tag{16}$$

The first way of gauging the difference between this function and the model least-squares criterion function of (6) is to compare the matrix $A'A$ with the

Toeplitz matrix $\Lambda = [\lambda_{|t-s|}]$ which, as we have already indicated, corresponds to the dispersion matrix of a synthetic pth-order MA process.

The second way is to compare the matrix

$$\frac{1}{n}Y'Y = \begin{bmatrix} \hat{c}_{00} & \hat{c}_{01} & \ldots & \hat{c}_{0,n-1} \\ \hat{c}_{10} & \hat{c}_{11} & \ldots & \hat{c}_{1,n-1} \\ \vdots & \vdots & & \vdots \\ \hat{c}_{n-1,0} & \hat{c}_{n-1,1} & \ldots & \hat{c}_{n-1,n-1} \end{bmatrix} \tag{17}$$

with the matrix $C = [c_{|j-k|}]$ which contains the usual estimates of the autocovariances of the process. The matrix $Y'Y/n$ does not have the Toeplitz form which is required if the condition of stationarity is to be fulfilled in all cases.

Imagine that p zeros elements are added to the tail of the vector $y = \{y_0, \ldots, y_{n-1}\}'$ and that an LT Toeplitz matrix $\bar{Y}$ of order $n+p$ is formed from the 'padded' vector on the same principle as $Y = Y(y)$ is formed from y. Then we should find that the principal minor of order $p+1$ of the matrix $\bar{Y}\bar{Y}/n$ would coincide with that of the matrix C. Now let $\bar{\alpha}$ be formed from the vector α by the addition of p zeros. Then $\bar{\alpha}'\bar{Y}\bar{Y}\bar{\alpha}/n = \alpha'C\alpha$, and it follows that the criterion function has become equivalent to the function which delivers the stationary Yule–Walker estimates.

Now consider specialising the criterion function to the case of the pure MA model. Then

$$\begin{aligned} S^z &= y'M'^{-1}M^{-1}y \\ &= \psi'Y'Y\psi, \end{aligned} \tag{18}$$

where $\psi = M^{-1}e_0$ contains the first n coefficients of the expansion of $\mu^{-1}(z)$. This function may be compared with the function S of (8) which can be written as $S = y'\Delta y$, where $\Delta = [\delta_{t-s}]$ is the dispersion of a synthetic AR process.

In pursuit of estimates which fulfil the condition of invertibility, we can improve the approximation of $M'^{-1}M^{-1}$ to $\Delta = \Delta(\mu)$ by adding extra rows to the matrix M^{-1} so as to include additional coefficients of the series expansion of $\mu^{-1}(z)$. In practice, this object may be achieved by padding the tail of the vector y with zeros.

Commentary. An advantage of the LT Toeplitz representation of the ARMA criterion function is the ease of computing the analytic derivatives $\partial\varepsilon/\partial\mu_j$ and $\partial\varepsilon/\partial\alpha_j$. This greatly facilitates the use of the Gauss–Newton (GN) iterative procedure in finding the estimates.

However, it is sometimes observed that, whereas the value of the criterion function is rapidly reduced in the initial stages of the computation, the estimates are slow to converge to their final values. The fault usually lies in the fact that matrix $Y'Y$ has been replaced, within the function $S^z = e_0'(A'M'^{-1}Y'YM^{-1}A)e_0$, by the Toeplitz matrix $C = [c_{|j-k|}]$ containing the usual estimates of the autocovariances.

The motive for this substitution is the fact that C contains only n distinct elements which compare with the $(n^2+n)/2$ distinct elements of $Y'Y/n$. The consequence is that the derivatives are no longer appropriate to the function which is to be minimised; and this problem will have its greatest effect on the performance of the GN procedure in the neighbourhood of the minimum. Once more, a cure for the problem is to pad the data vector with zeros.

Circulant Matrices

The following is as example of a circulant matrix:

$$Y = \begin{bmatrix} y_0 & y_3 & y_2 & y_1 \\ y_1 & y_0 & y_3 & y_2 \\ y_2 & y_1 & y_0 & y_3 \\ y_3 & y_2 & y_1 & y_0 \end{bmatrix}. \tag{19}$$

Circulant matrices can be represented as polynomial functions of a matrix $J = [e_1, \ldots, e_{n-1}, e_0]$ which is formed from the identity matrix $I = [e_0, e_1, \ldots, e_{n-1}]$ by moving the leading vector to the back of the array. Thus the circulant matrix $A = A(\alpha)$ can be written as

$$\begin{aligned} A &= \alpha(J) \\ &= I + \alpha_1 J + \cdots + \alpha_p J^p. \end{aligned} \tag{20}$$

We may note that $J^{j+n} = J^j$ for all j.

The algebra of circulant matrices closely resembles that of polynomials:

(i) The matrices commute in multiplication, $AY = YA$,

(ii) If A, Y are circulant, then so is $AY = YA$,

(iii) If A is circulant, then so are A' and A^{-1},

(iv) If M is circulant, then $M'M = MM'$ and $(M'M)^{-1}$ are circulant Toeplitz matrices,

(v) Although $M'M$ is a Toeplitz matrix, the equation $G = M'M$ does not correspond to the Cramér–Wold factorisation.

Let $Y = y(J)$, $A = \alpha(J)$ and $M = \mu(J)$ be circulant matrices constructed from the same vectors y, α, μ as were the corresponding lower-triangular Toeplitz matrices. Then we can construct the following criterion function:

$$S^c = e_0'(Y'A'M'^{-1}M^{-1}AY)e_0 = y'A'M'^{-1}M^{-1}Ay. \tag{21}$$

This is just the coefficient associated with $I = J^0$ in the expansion of

$$\begin{aligned} \varepsilon(J)\varepsilon(J^{-1}) &= \frac{y(J)\alpha(J)\alpha(J^{-1})y(J^{-1})}{\mu(J)\mu(J^{-1})} \\ &= n\frac{c(J)}{\gamma(J)}, \end{aligned} \tag{22}$$

where $J^{-1} = [e_{n-1}, e_0, \ldots, e_{n-2}]$, and where $c(J) = y(J)y(J^{-1})/n$ and $\gamma(J) = \alpha(J)\alpha(J^{-1})/\mu(J)\mu(J^{-1})$.

The role of the matrix J in the above expression is essentially that of an indeterminate algebraic symbol, and it may be replaced by any other quantity which is an n-periodic function of the index j. In particular, we may replace J^j by $e^{i\omega_j} = \exp\{i2\pi j/n\}$. Then we have the result that

$$\begin{aligned} S^c &= \sum_{j=0}^{n-1} \frac{y(e^{i\omega_j})\alpha(e^{i\omega_j})\alpha(e^{-i\omega_j})y(e^{-i\omega_j})}{\mu(e^{i\omega_j})\mu(e^{-i\omega_j})} \\ &= n\sum_{j=0}^{n-1} \frac{c(e^{i\omega_j})}{\gamma(e^{i\omega_j})}. \end{aligned} \tag{23}$$

This follows from the fact that

$$\sum_{j=0}^{n-1} e^{i\omega_j} = \begin{cases} 0 & \text{if } j \neq 0, \\ n & \text{if } j = 0, \end{cases} \tag{24}$$

which indicates that the coefficient associated with $e^{i\omega_0} = 1$ can be isolated by summing over j.

The sum in (23) is manifestly an approximation to the function

$$S = \frac{n}{2\pi}\int_{-\pi}^{\pi} \frac{c(e^{i\omega})}{\gamma(e^{i\omega})} d\omega \tag{25}$$

which is a form of the function of (4). The approximation of S^c to S can be made arbitrarily close by increasing the number of frequency points ω_j at which the function is evaluated or, equivalently, by increasing the order of the matrix J. If the data series is of a fixed length n, then this is achieved by padding the vector y with zeros.

Consider specialising the criterion function to the case of a pure AR model. Then $n^{-1}S^c = n^{-1}y'A'Ay = n^{-1}\alpha'Y'Y\alpha$ has the form of the function of (16) apart from the fact that the matrices $A = \alpha(J)$ and $Y = y(J)$ are circulant matrices instead of LT Toeplitz matrices. The matrix $c(J) = Y'Y/n$ is given by

$$\begin{aligned} c(J) &= c_{n-1}J^{1-n} + \cdots + c_1 J^{-1} + c_0 I + c_1 J + \cdots + c_{n-1}J^{n-1} \\ &= c_0 I + (c_1 + c_{n-1})J + (c_2 + c_{n-2})J^2 + \cdots + (c_{n-1} + c_1)J^{n-1}, \end{aligned} \tag{26}$$

where the elements $c_0, \ldots, c_{n-1}$ come from (5). The equality follows from the fact that $J^{j-n} = J^j$.

Given that $c(J) = Y'Y/n$ is a positive-definite Toeplitz matrix, it follows from Theorem 1 that the values which minimise the AR criterion function will correspond to a model which satisfies the condition of stationarity. The estimates will differ slightly from the Yule–Walker estimates because of the differences between $c(J) = Y'Y/n$ and $C = [c_{|j-k|}]$.

Consider the effect of adding p zeros to the tail of the vector y to create a vector $\tilde{y}$ and a corresponding matrix $\tilde{Y} = \tilde{y}(J)$ where J is now of order $n+p$ and $J^{n+p} = I$. Then, if $\tilde{c}(J) = \tilde{Y}'\tilde{Y}/n$, we have

$$\begin{aligned}\tilde{c}(J) = c_0 I + \cdots + c_p J^p + (c_{p+1} + c_{n-1})J^{p+1} + \cdots \\ + (c_{n-1} + c_{p+1})J^{n-1} + \cdots + c_1 J^{n-1+p}.\end{aligned} \tag{27}$$

It can the seen that the principal minor of order $p+1$ of the matrix $\tilde{C} = \tilde{c}(J)$ is identical to that of the matrix $C = [c_{|k-j|}]$, and it follows that the criterion function has become equivalent to the function which delivers the Yule–Walker estimates.

Finally, let us comment on the specialisation of the criterion function to the case of the pure MA model. The criterion function has the form of the function under (18) albeit with circulant matrices $Y = y(J)$ and $M = \mu(J)$ in place of LT Toeplitz matrices. The conditions of Theorem 2, which guarantee that the MA estimates will satisfy the condition of invertibility, are no longer fulfilled. Nevertheless, if there is any danger that the condition of invertibility may be violated, the simple of expedient of padding the tail of the vector y with a sufficient number of zeros will avert the problem.

Commentary. The representation of the least-squares criterion function which is in terms of circulant matrices is of some practical interest since, in the guise of equation (23), it is the criterion function which is entailed in the frequency-domain estimation of ARMA models. This method of estimation was expounded by Hannan [10] and has been investigated in some detail by Pukkila [18], [19]. Cameron and Turner [4] have show how to implement the method within the context of a flexible regression package.

The distinguishing feature of a frequency-domain ARMA estimation is the use of the fast Fourier transform (FFT) in performing the convolutions which are entailed in the multiplication of the polynomials or in the multiplication of the analogous matrices. There is little, if anything, to be gained from using the FFT when the data sample contains fewer than several hundred points.

Nowadays ARMA models are being used increasingly in signal-processing applications where there may be an abundance of data and where speed of computation is important. In such cases, a well coded frequency-domain method may be may be far superior to a corresponding time-domain method.

5. The Sampling Properties of the Estimators

Commentary. There has been some concern over the small-sample properties of the frequency-domain estimator of a pure AR process. In particular, it appears that, in small samples, the modulii of the roots of the AR operator tend to be underestimated; and the severity of this bias increases as the roots approach the unit circle. The peaks of the estimated AR spectral density function which correspond to these roots assume less prominence that they should, and they may even disappear altogether. In these respects, the exact ML estimator fares rather better.

As we have shown, the frequency-domain AR estimator can be made equivalent to the Yule–Walker estimator by the device of padding the data vector with zeros. Therefore what is known about the properties of the Yule–Walker estimator is also pertinent to the frequency-domain estimator. Thus one may study profitably the evidence gathered by Tjøstheim and Paulsen [21] and by Lysne and Tjøstheim [15].

Pukkila [18], [19], has characterised the properties of the unpadded frequency-domain estimator, and he has suggested some modifications which are aimed at reducing the small-sample biases.

There is less evidence on the small-sample properties of the frequency-domain estimator of a pure MA model. However, the experience of the author suggests that there is a tendency to underestimate the modulii of the roots of the MA operator in small samples; and this is exacerbated when one resorts to the device of padding.

There are alternative ways of reaching an intuitive explanation of the small-sample bias of the Yule–Walker estimates which lead to various suggestions for improving their properties. These explanations make reference either to the sequence of empirical autocovariances or to the sequence of periodogram ordinates which represents the Fourier transform of the autocovariances.

To begin, consider the empirical autocovariance of lag τ which, on the assumption of that $E(y_t) = 0$, is given by

$$c_\tau = \frac{1}{n}\sum_{t=\tau}^{n-1} y_t y_{t-\tau}. \tag{28}$$

The expected value is

$$E(c_\tau) = \gamma_\tau \left(1 - \frac{|\tau|}{n}\right), \tag{29}$$

where γ_τ is the true value. If n is small, then the sequence of the estimated autocovariances is liable to decline more rapidly than it should as the lag value τ increases.

To understand the consequences of the over-rapid decline of the empirical autocovariances, we may consider the fact there is a one-to-one correspondence between the sequence $c_0, \ldots, c_p$ and the Yule–Walker estimates of the parameters $\sigma^2 = V(\epsilon_t), \alpha_1, \ldots, \alpha_p$. In particular, the estimates satisfy the difference equation

$$c_p + \alpha_1 c_{p-1} + \cdots + \alpha_p c_0 = 0. \tag{30}$$

If $\{c_0, \ldots, c_p\}$ is declining too rapidly, then the solution of the difference equation is liable to be overdamped, which means that the roots of the polynomial equation $\alpha(z) = 0$ will be too far from the unit circle.

One way of addressing the problem of bias is to replace the divisor n in the formula for c_τ by a divisor of $n - \tau$ so as to obtain unbiased estimates of the autocovariances. However, the resulting matrix of autocovariances is no longer guaranteed to be positive definite; and this can lead to the violation of the condition of stationarity.

Another recourse it to adopt a two-stage estimation procedure. The initial Yule–Walker estimates can be used in forecasting sequences of postsample and presample values which are added to the sample. The forecast values in both directions will converge or 'taper' to zero as the distance from the sample increases. At the points where the values are judged to be sufficiently close to zero, the sequences may be terminated. The Yule–Walker estimates which are obtained from the supplemented data have less bias than the initial estimates.

Recently, Pukkila [20] has proposed a modified Yule–Walker estimator which is calculated from autocorrelations which are obtained indirectly via the partial autocorrelation function. His sampling experiments suggest that the properties of the estimator are as good as, if not better than, those of the Burg estimator (see, for example, Burg [3], Ulrych and Bishop [22] and Giordano and Hsu [8]) which, in recent years, has provided a benchmark for small-sample performance.

The alternative way of explaining the bias in the Yule–Walker estimates is to consider the expectation of the periodogram which is the Fourier transform of sequence of the expectated values of the empirical autocovariances :

$$E\{c(e^{i\omega})\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \gamma(e^{i\lambda})\kappa(e^{i(\omega-\lambda)})d\lambda. \tag{31}$$

The expected periodogram is also the convolution of the spectral density function $\gamma(e^{i\omega})$ with the Fejér kernel $\kappa(e^{i\omega})$. The former represents the Fourier transform of the sequence of true autocovariances whilst the latter represents the Fourier transform of the triangular sequence of the weights

$$d_\tau = \begin{cases} 1 - \dfrac{|\tau|}{n} & \text{if } |\tau| < n, \\ 0 & \text{if } |\tau| \geq n, \end{cases} \tag{32}$$

which appear in the expression for $E(c_\tau)$ under (29).

The convolution represents a smoothing operation, performed upon spectral density function, which has the Fejér kernel as its weighting function. The effect of the operation is to diffuse the spectral power which spreads from the peaks of the spectrum, where it is concentrated, into the valleys. This is described as spectral leakage.The dispersion of the Fejér kernel diminishes as n increases, and, in the limit, it becomes a Dirac delta function. When the Dirac function replaces the Fejér kernel, the convolution delivers the spectral density function $\gamma(e^{i\omega})$ unimpaired.

An explanation of the presence of the Fejér kernel can be obtained from the notion that the sample values $y_t; t := 0, \ldots, n-1$ are obtained by applying the weights

$$w_t = \begin{cases} 1 & \text{if } 0 \leq t < n, \\ 0 & \text{otherwise,} \end{cases} \tag{33}$$

of a rectangular data window to the elements of an infinite sequence. The triangular weighting function $d_\tau = n^{-1}\sum_t w_t w_{\tau-t} = 1 - |\tau|/n$ of (32), which affects the sequence of autocovariances, and whose transform is the Fejér kernel, is formed from the convolution of two rectangular widows. By modifying the data window, we may alter the kernel function so as to reduce the leakage. In general, the leakage may be reduced by applying a taper to the ends of the rectangular window.

Investigations into the use of data-tapering in autoregressive estimation were pioneered by Pukkila [18] who modified the rectangular window by removing its corners to create a trapezoidal form. More recently, Dahlhaus [6] has investigated the effects upon the leakage of a tapered window obtained by splitting a cosine bell and inserting a sequence of units between the two halves. The sampling experiments of both Pukkila and Dahlhaus reveal dramatic improvements in the bias of the autoregressive estimates and in the resolution of the spectral density function which is inferred from these estimates.

Ideally the degree of tapering—which, in the case of Dahlhaus, is the ratio of the width of the cosine bell to the width of the data window—should be attuned to the values of the roots of $\alpha(z)$. A high degree of tapering is called for when the modulus of the dominant root is close to unity, which is usually the case when there is a prominent peak in the spectral density function.

The emphasis which has been placed, in the literature, upon the sampling properties of AR estimators should not detract from the importance of the MA component in time-series models. Its presence can greatly enhance the flexibility of the model in approximating transfer functions. An example is provided by the case of an AR process which has been corrupted by a white-noise error.

A white-noise corruption, which might arise simply from rounding error in the observations, increases the variance of the data, leaving its autocovariances unaffected. The inflated variance increases the damping of the autocovariances at the start of the sequence. This can lead to a severe underestimation of the modulii of the autoregressive roots. Formally, an AR(p) model with added white noise gives rise to an ARMA(p,p) process. Nevertheless, the noise corruption can often be accommodated by adding just a few moving-average parameters to the model.

Appendix

The Cauchy Integral Theorem. The Cauchy Integral Theorem—see, for example, Kreyszig [12]—indicates that

$$\frac{1}{2\pi i}\oint z^t \frac{dz}{z} = \begin{cases} 1, & \text{if } t = 0; \\ 0, & \text{if } t \neq 0. \end{cases} \tag{A1}$$

A familiar specialisation of this result comes from taking the perimeter of the unit circle as the contour of integration. Then , by setting $z = e^{i\omega}$ and changing the variable of integration from z to $\omega \in (-\pi, \pi]$, we get

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\omega t} d\omega = \frac{1}{\pi}\int_{0}^{\pi} \cos(\omega t) d\omega = \begin{cases} 1, & \text{if } t = 0; \\ 0, & \text{if } t \neq 0. \end{cases} \tag{A2}$$

Consider, for example, the self-reciprocal function

$$\gamma(z) = \gamma_0 + \gamma_1(z + z^{-1}) + \cdots + \gamma_q(z^q + z^{-q}). \tag{A3}$$

It can easily be seen that

$$\gamma_0 = \frac{1}{2\pi i}\oint \gamma(z)\frac{dz}{z}. \tag{A4}$$

Theorem 1. Let $S = \sum_{j=0}^{p}\sum_{k=0}^{p} \alpha_k \alpha_j c_{|j-k|} = \alpha' C \alpha$ where $C = [c_{|j-k|}]$ is a symmetric positive-definite Toeplitz matrix and $\alpha = \{1, \alpha_1, \ldots, \alpha_p\}'$ is a vector. If α is chosen so as to minimise the value of S, then all the roots of the equation $\alpha(z) = 1 + \alpha_1 z + \cdots + \alpha_p z^p = 0$ will lie outside the unit circle.

Proof. If we factorise $\alpha(z)$ as $\alpha(z) = q(z)\phi(z)$ where $q(z) = 1 + q_1 z + q_2 z^2$ is quadratic, then, for a given value of $\phi(z)$, we can write the function as

$$S(q) = \begin{bmatrix} 1 & q_1 & q_2 \end{bmatrix} \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_2 & \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ q_1 \\ q_2 \end{bmatrix}, \tag{A5}$$

where $\gamma_k = \sum_i \sum_j \phi_i \phi_j c_{|i-j+k|}$. At the point of the minimum, we find that

$$\begin{aligned} q_1 &= \frac{\gamma_1\gamma_2 - \gamma_0\gamma_1}{\gamma_0^2 - \gamma_1^2}, \\ q_2 &= \frac{\gamma_1^2 - \gamma_0\gamma_2}{\gamma_0^2 - \gamma_1^2} \quad \text{and} \\ S &= \gamma_0 + q_1\gamma_1 + q_2\gamma_2. \end{aligned} \tag{A6}$$

In terms of these values, we can can express $\gamma_0, \gamma_1, \gamma_2$ as

$$\begin{aligned} \gamma_0 &= \frac{S(1+q_2)}{d}, \\ \gamma_1 &= \frac{-Sq_1}{d} \quad \text{and} \\ \gamma_2 &= \frac{S\{q_1^2 - q_2(1+q_2)\}}{d}, \quad \text{where} \\ d &= (1-q_2)(1+q_2+q_1)(1+q_2-q_1). \end{aligned} \tag{A7}$$

Now the matrix of (A5) is positive definite by virtue of its construction. In particular, the principal minor of order 2 must be positive definite which is equivalent to the conditions that $\gamma_0 > 0$ and $\gamma_0^2 - \gamma_1^2 > 0$. Given that $S > 0$, these imply that

$$\begin{aligned} &(1+q_2)^2 > q_1^2 \quad \text{and} \\ &\frac{1+q_2}{1-q_2} > 0 \quad \text{or , equivalently,} \quad 1 - q_2^2 > 0. \end{aligned} \tag{A8}$$

The latter conditions are necessary and sufficient to ensure that the roots of $q(z) = 0$ lie outside the unit circle. They are equivalent to the conditions listed under (3.2.18) by Box and Jenkins [2, p. 58].

We can repeat this analysis for every other quadratic factor of the polynomial $\alpha(z)$ in order to show that all of the complex roots must lie outside the unit circle when S is minimised. It is also easy to show that the real roots must lie outside the unit circle.

Theorem 2. Let $S = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\psi_k\psi_j c_{|k-j|}$ where ψ_i,ψ_j are coefficients in the expansion of the inverse polynomial $\mu^{-1}(z) = \{1 + \psi_1 z + \cdots\}$. If the coefficients of $\mu(z) = 1 + \mu_1 z + \cdots + \mu_q z^q$ are chosen so as to minimise the value of S, then all the roots of the equation $\mu(z) = 0$ will lie outside the unit circle.

Proof. By setting $\psi_i = 0$ for $i > 1$, we can obtain the inequality $\min S(\mu) \leq c_0$ which shows that there is a finite-valued minimum. But $S(\mu)$ is bounded if and only if $\sum|\psi_i| < \infty$. Therefore $\mu^{-1}(z) = \psi(z)$ converges for $|z| \leq 1$; and so the roots of $\mu(z) = 0$ must lie outside the unit circle. Therefore the estimates fulfil the condition of invertibility.

References

[1] Anderson, T.W. and A.A. Takemura (1986), Why do noninvertible estimated moving averages occur?, Journal of Time Series Analysis, 7, 235–254.

[2] Box G.E.P. and G.M. Jenkins (1970), Time-Series Analysis : Forecasting and Control, Holden-Day, San Francisco.

[3] Burg, J.P. (1978), Maximum entropy spectral analysis, in D.G. Childers (ed.), Modern Spectral Analysis, 34-41, IEEE Press, New York.

[4] Cameron, M.A. and T.R. Turner (1987), Fitting models to spectra using regression packages, Applied Statistics, 36, 47–57.

[5] Cryer, J.D. and J. Ledolter (1981), Small-sample properties of the maximum-likelihood estimator in the first-order moving average model, Biometrika, 68, 691–694.

[6] Dahlhaus, R. (1984), Parameter estimation of stationary processes with spectra containing strong peaks, in Franke, J., W. Hardle and D. Martin (eds.), Robust and Nonlinear Time Series Analysis, Lecture Notes in Statistics, vol. 26, 50-86, Springer Verlag, New York.

[7] Davidson, E.H. (1981), Problems with the estimation of moving-average processes, Journal of Econometrics, 16, 295–310.

[8] Giordano, A.A. and F.M. Hsu (1985), Least Square Estimation with Applications to Digital Signal Processing, John Wiley and Sons, New York.

[9] Godolphin, E.J. (1977), A direct representation for the maximum-likelihood estimator of a Gaussian moving-average process, Biometrika, 64, 375–84.

[10] Hannan, E.J. (1969), The estimation of mixed moving-average autoregressive systems, Biometrika, 56, 579–593.

[11] Kang, K.M. (1975), A comparison of the least-squares and maximum-likelihood estimation for moving-average processes, Bureau of Census and Statistics, Canberra.

[12] Kreyszig, E. (1988), Advanced Engineering Mathematics: sixth edition, John Wiley and Sons, New York.

[13] Laurie, D.P. (1980), Efficient implementation of Wilson's algorithm for factorising a self-reciprocal polynomial. BIT, 20, 257–259.

[14] Laurie, D.P. (1982), Algorithm AS 175 : Cramér-Wold Factorisation, Applied Statistics , 86–93.

[15] Lysne, D. and D. Tjøstheim (1987), Loss of spectral peaks in autoregressive spectral estimation, Biometrika, 74, 200–206.

[16] Osborn, D.R. (1982), On the criteria functions used for the estimation of moving-average processes, Journal of the American Statistical Association, 77, 388–392.

[17] Pagano, M. (1973), When is an autoregressive scheme stationary? Communications in Statistics, 1, 533–544.

[18] Pukkila, T. (1977), Fitting of Autoregressive Moving Average Models in the Frequency Domain, Doctoral Thesis, Department of Mathematical Sciences, University of Tampere, Report A-6.

[19] Pukkila, T.M. (1979), The bias in periodogram ordinates and the estimation of ARMA models in the frequency domain, Australian Journal of Statistics, 21, 121–128.

[20] Pukkila, T.M. (1988), An improved estimation method for univariate autoregressive models, Journal of Multivariate Analysis, 27, 422–433.

[21] Tjøstheim, D. and J. Paulsen (1983), Bias of some commonly-used time series estimates, Biometrika, 70, 389–399. Correction (1984), 71, 656.

[22] Ulrych, T.J. and T.N. Bishop (1975), Maximum entropy spectral analysis and autoregressive decomposition, Review of Geophysics 13, 182-200.

[23] Wilson, G.T. (1969), Factorisation of the covariance generating function of a pure moving-average process. SIAM J. Statist. Comput. Simul, 8, 301–309.

1990-1	B. Vogelvang	Testing For Co-Integration with Spot Prices of Some Related Agricultural Commodities
1990-2	J.C.J.M. van den Bergh P. Nijkamp	Ecologically Sustainable Economic Development Concepts and Model Implications
1990-3	J.C.J.M. van den Bergh P. Nijkamp	Ecologically Sustainable Economic Development in a Regional System: A Case Study in Agri cultural Development Planning in the Netherlands
1990-4	C.Gorter P.Nijkamp P.Rietveld	Employers' Recruitment Behaviour and Re-Employment Probabilities of Unemployed
1990-5	K.Burger	Off-farm income and the farm-household the case of Kenyan smallholders
1990-6	H. Visser	Crowding out and the Government Budget
1990-7	P. Rietveld	Ordinal Data in Multicriteria Decision Making, a Stochastic Dominance Approach to Siting Nuclear Power Plants
1990-8	G. van der Laan P.H.M. Ruys D.J.J. Talman	Signaling devices for the supply of semi-public goods
1990-9	F.A.G. den Butter	Labour Productivity Slowdown and Technical Progress: An empirical analysis for The Netherlands
1990-10	R.W. van Zijp	Neo-Austrian Business Cycle Theory
1990-11	J.C. van Ours	Matching Unemployment and Vacancies: The Efficiency of the Dutch Labour Market
1990-12	B. Vogelvang	Hypotheses Testing Concerning Relationships between Spot Prices of Various Types of Coffee
1990-13	A.F. de Vos I.J. Steyn	Stochastic Nonlinearity: A Firm Basis for the Flexible Functional Form
1990-14	Y.H. van Emmerik D. de Jong W.W.A. Zuurmond D.N. Dukkers-van Emden	Opereren in overleg: geprotocolleerde samen-werking 1e-2e-lijn bij dagchirurgie
1990-15	T.J.J.B. Wolters	Mediation and Collective Bargaining: A Diagnostic Approach
1990-16	E.M.A. Scholten J. Koelewijn	Financieringsproblematiek van startende ondernemingen: een mogelijke verklaring op basis van empirisch onderzoek.
1990-17	E. Hüner H.P. Smit	Saturation and Model Specification of Passen ger car Ownership
1990-18	F.A.G. den Butter	Sociale zekerheid, de wig en economische groei
1990-19	F.A.G. den Butter R.F. v.d. Wijngaert	Who is Correcting the Error ? A Co-inte-gration Approach for Wages, Wage Space and Labour Conflicts in the Netherlands
1990-20	J.P. de Groot R. Ruben	Sistemas de Producción y Transferencia de Teconología en la Economía Cafetalera de Centroamerica
1990-21	R. Ruben	Campesinado y Reforma Agraria en El Salvador
1990-22	J. van Ours G. Ridder	Vacancies and the Recruitment of New Employees
1990-23	A.F. de Vos J.J. de Vries	The Likelihood Function of a Generalized Gravity Model; Handling the Implicit Singularity of a Nonlineair Transformation
1990-24	D. van der Wal	Geloofwaardigheid en het Nederlandse wisselkoersbeleid
1990-25	R.J. Veldwijk M. Boogaard M.V. van Dijk E.R.K. Spoor	EDSOs, implosion and explosion: concepts to automate a part of application maintenance
1990-26	B. Hanzon	The area enclosed by the (oriented) Nyquist diagram and the Hilbert-Schmidt-Hankel norm of a linear system
1990-27	R.W. van Zijp	Why Lucas is not a Hayekian
1990-28	J. Rouwendal	On discrete choice under uncertainty: A generalization of the logit model and its application
1990-29	J. Rouwendal	On the equitable distribution of the housing stock
1990-30	J. Rouwendal	Stochastic market equilibria with efficient rationing (with an application to the Dutch housing market)
1990-31	J.A. Vijlbrief	The effects of unemployment insurance on the labour market
1990-32	J.G.W. Simons H.P. Wansink	Traffic Ban, a means to combat smog?
1990-33	J.C. van Ours T. Zoethout	De Interne Arbeidsmarkt van de Gemeente Amsterdam
1990-34	H.J. Bierens	A Note On the Limiting Distribution of Sample Autocorrela-tions in the Presence of a Unit Root
1990-35	T. Kuhlman	The Economic Integration of Refugees in Developing Countries: A Research Model.
1990-36	T. Kuhlman	Towards a Definition of Refugees.
1990-37	T. Kuhlman	Organized versus Spontaneous Settlement of Refugees in Africa.
1990-38	R. Zuidema	The Neo-Austrian View on Interest
1990-39	G.v.d.Laan	General Equilibrium in a Closed International Trade Model