Applied Mathematical Sciences, Vol. 4, 2010, no. 23, 1129 - 1139

High-Dimensional ARMA Model Identification and Its Application to Healthcare Picture Smoothing Using a Forgetting Factor

Jack Penn

School of Finance and Applied Statistics
College of Business and Economics
The Australian National University
ACT 0200, Australia
Jack.Penm@anu.edu.au

Abstract

In this paper a set of formulations of an N-dimensional (ND) autoregressive-moving average (ARMA) model identification method, and a two-dimensional (2D) forgetting factor approach in time-series modelling, is developed. An optimum estimation and prediction approach in healthcare picture smoothing based on a 2D ARMA modelling, has been implemented; and satisfactory results have been obtained. Our approach indicates the desirability of accurate statistical modelling of high-dimensional or periodic digital data.

Keywords: N-dimensional ARMA Model; Forgetting factor.

1. Introduction

Most recently there has been remarkable progress (see O'Neill *et al*, 2007) in electronic healthcare as well as in informatics. One important benefit of this perspective and its application to healthcare research is that our life expectancy is now significantly greater than it was even a few decades ago. This progress, leading to generally more ageing societies, is of importance to the organisation of healthcare services, technology management and the future development of its information systems, in particular 2D time-series signal processing and smoothing systems applied to services and technology management for digital signal processing.

The latest major advances in powerful computing equipment, which undertake a shift from paper-based to computer-based processing and storage, as well as an increase of data in electronic service and technology settings, have made a tremendous impact on development of 2D time-series smoothing approaches for electronic services and technology monitoring systems. 2D time-series smoothing is a preliminary process in many electronic healthcare applications. It aims to reduce 'noise' in 2D time-series signals. All 2D time-series modelling tasks may benefit from the reduction of 'noise'. Our paper provides a formulation of the high-dimensional forgetting factor approach in N-dimensional (ND) data. We also design and describe an order identification algorithm of a ND ARMA model. The proposed algorithm is a general case of a 2D ARMA identification approach for smoothing of 2D time-series data.

The remaining organisation of this paper is as follows. Section 2 describes 2D ARMA modelling. Section 3 shows an identification algorithm of ND ARMA modelling, which is a general case of 2D ARMA identification procedures. Section 4 presents a high-dimensional forgetting factor approach. Section 5 exhibits estimation recursions of 2D ARMA modelling. Section 6 provides the prediction and estimation of an identified chromosome picture. Section 7 summarises the paper.

2. 2D ARMA modelling

1D ARMA models always specify a factorization of the two-sided z-transform rational function $\phi_0(z)$ of the auto-covariance matrix $R_y(\tau) = E\{y(t+\tau)y(t)\}$. But in the 2D case such a factorization may not exist for $\Phi_0(z_1, z_2)$, which is traditionally identified as follows:

$$\Phi_0(z_1, z_2) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} R_y(i,j) z_1^i z_2^j = (\sum_{i=0}^{M_1}\sum_{j=0}^{M_2} a_{ij} z_1^i z_2^j)^{-1} (\sum_{i=0}^{N_1}\sum_{j=0}^{N_2} f_{ij} z_1^i z_2^j), \tag{1}$$

with $a_{00} = 1$. On combining these last two equations and equating coefficients of identical powers of z_1 and z_2 one obtains

$$\sum_{r=0}^{M_1}\sum_{s=0}^{M_2} a_{rs} u(t_1 - r) u(t_2 - s) R_y(t_1 - r, t_2 - s) = 0 \tag{2}$$

for all $(t_1,t_2) \notin ((t_1,t_2)\colon 0 < t_1 < M_1,\ 0 < t_2 < M_2)$, $u(x)=0$ as $x < 0$; $=1$, as $x \geq 0$. Earlier work has considered the identification of 1D ARMA model based on a previously identified infinite dimensional AR model. This paper extends and modifies this approach to show the 2D ARMA model of equation (2) can be identified from a one-quadrant, 2D spectrum of the form

$$\Phi_0(z_1,z_2)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}c_{i,j}z_1^i z_2^j \text{ with } c_{0,0}=1. \tag{3}$$

In an effort to distinguish between computational and fundamental difficulties, we will make two assumptions. In the first place, to reduce the computational difficulties we have assumed that R(i,j)=R(i,-j). This assumption is necessary for the one - quadrant spectrum of equation (3) to characterize the process.

We also assume, indirectly, that the numerator and denominator polynomials of equation (3) have closed-order property, i.e. $f_{i,j}=0,$ if $f_{i,0}$ or $f_{0,j}=0$; and $a_{i,j}=0,$ if $a_{i,o}$ or $a_{0,j}=0$. The identification procedure developed in this paper automatically truncates the model in such a way that the solution is correct only when the closed-order property is satisfied. Some ideas as to how this truncation can be overcome are considered briefly in the conclusion. It should also be observed that for factorable processes, the closed-order assumption is automatically satisfied.

3. Identification of ND ARMA modelling

If the 2D ARMA model of (2) exists and satisfies (3), then

$$\left(\sum_{i=0}^{M_1}\sum_{j=0}^{M_2}a_{i,j}z_1^i z_2^j\right)\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}c_{i,j}z_1^i z_2^j\right)=\left(\sum_{i=0}^{N_1}\sum_{j=0}^{N_2}f_{i,j}z_1^i z_2^j\right)$$

where $a_{i,j}=0$ for $i>M_1$ or $j>M_2$. Equating the coefficients of like powers of $z_1{}^i, z_2{}^j$, we have $f_{n,k}=\sum_{r=0}^{n}\sum_{s=0}^{k}a_{n-r,k-s}c_{r,s},$ $\text{for } n=1,2,...N_1; k=1,2,...N_2$

For k=0, one obtains $f_{n,0}=\sum_{r=0}^{n}a_{n-r,0}c_{r,0},$, for n=1, 2,·· N_1,

$$\text{and } 0=\sum_{r=m}^{N_1+m}a_{n_1+m-r,0}c_{r,0} \quad \text{for m=1,2,..., } M_1$$

In the N-dimensional case, the measured power spectrum, $\phi_0(z_1,z_2,\cdots z_n)$ will be of the form $\phi_0(z_1,z_2,\cdots z_n)=\sum_{i_1=0}^{\infty}\sum_{i_2=0}^{\infty}\cdots\sum_{i_N=0}^{\infty}c_{i_1,i_2,\dots,i_n}z_1^{i_1}z_2^{i_2}\cdots z_N^{i_N}$ (4)

If the power spectrum $\phi_0(z_1,z_2,\cdots z_n)$ can be represented by

$$\phi_0(z_1, z_2, \cdots z_n) = (\sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} \cdots \sum_{i_N=0}^{M_N} a_{i_1,i_2,\ldots,i_n} z_1^{i_1} z_2^{i_2} \cdots z_N^{i_N})^{-1}$$

$$(\sum_{i_1=0}^{N_1} \sum_{i_2=0}^{N_2} \cdots \sum_{i_N=0}^{N_N} f_{i_1,i_2,\ldots,i_n} z_1^{i_1} z_2^{i_2} \cdots z_N^{i_N})$$

where $N_1 \not\subset M_1 \; N_2 \not\subset M_2, \ldots, N_N \not\subset M_N$, and with closed-order property, then

$$\sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \sum_{i_N=0}^{\infty} \cdots c_{i_1,i_2,\ldots,i_n} z_1^{i_1} z_2^{i_2} \cdots z_N^{i_N} = (\sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} \sum_{i_N=0}^{M_N} \cdots a_{i_1,i_2,\ldots,i_n} z_1^{i_1} z_2^{i_2} \cdots z_N^{i_N})^{-1}$$

$$x(\sum_{i_1=0}^{N_1} \sum_{i_2=0}^{N_2} \sum_{i_N=0}^{N_N} \cdots f_{i_1,i_2,\ldots,i_n} z_1^{i_1} z_2^{i_2} \cdots z_N^{i_N})$$

Cross multiplication and equating the coefficients of like powers of $z_1, z_2, \cdots z_n$ yields:

$$\begin{bmatrix} c_{N_1,0\ldots0} & c_{N_1-1,0\ldots0} \cdots & c_{N_1-M_1+1,0\ldots0} \\ c_{N_1+1,0\ldots0} & c_{N_1,0\ldots0} \cdots & c_{N_1-M_1+2,0\ldots0} \\ . & \cdots & . \\ . & \cdots & . \\ c_{N_1+M_1-1,0\ldots0} & \cdots & c_{N_1,0\ldots0} \end{bmatrix} \begin{bmatrix} a_{1,0\ldots0} \\ a_{2,0\ldots0} \\ . \\ . \\ a_{M_1,0\ldots0} \end{bmatrix} =$$

$$\underset{\sim}{A}_{1,0\ldots0} \times \begin{bmatrix} a_{1,0\ldots0} \\ a_{2,0\ldots0} \\ . \\ . \\ a_{M_1,0\ldots0} \end{bmatrix} = - \begin{bmatrix} c_{N_1+1,0\ldots0} \\ c_{N_1+2,0\ldots0} \\ . \\ . \\ c_{N_1+M_1,0\ldots0} \end{bmatrix};$$

$$\begin{bmatrix} c_{0,..N_1,..0} & c_{0,..N_1-1,..0} \cdots & c_{0,\ldots N_1-M_1+1,..0} \\ c_{0,..N_1+1,..0} & c_{0,..N_1,..0} \cdots & c_{0,\ldots N_1-M_1+2,..0} \\ . & \cdots & . \\ . & \cdots & . \\ c_{0,..N_1+M_1-1,..0} & \cdots & c_{0,..N_1,..0} \end{bmatrix} \begin{bmatrix} a_{0,..1,..0} \\ a_{0,..2,..0} \\ . \\ . \\ a_{0,..M_1,..0} \end{bmatrix} =$$

$$\underset{\sim}{A}_{0,\ldots 1_i,\ldots 0} \times \begin{vmatrix} a_{0,..1,..0} \\ a_{0,..2,..0} \\ . \\ . \\ . \\ a_{0,..M_1,..0} \end{vmatrix} = - \begin{vmatrix} c_{0,..N_1+1,..0} \\ c_{0,..N_1+2,..0} \\ . \\ . \\ . \\ c_{0,..N_1+M_1,..0} \end{vmatrix} . \qquad (5)$$

The square matrix $\underset{\sim}{A}_{0,\ldots 1_i \ldots 0}$ in (5) must be non-singular for a unique solution $a_{0,..j_i,..0}$. Such will be the case if the M_i and N_i specified are of minimal order. Therefore the order of the model can be estimated by examining the determinants of the $\underset{\sim}{A}_{0,\ldots 1_i \ldots 0}$ test matrices. If the coefficients $c_{i_1,i_2,\ldots,i_n}$ in equation (3) are known, we can decide the orders $\{M_1,N_1\},\ldots,\{M_i,N_i\}, \ldots, \{M_N,N_N\}$ by testing, and the coefficients ($a_{i_1,i_2,\ldots,i_n}$) and ($f_{i_1,i_2,\ldots,i_n}$) in (1) can be uniquely determined.

Let us consider the shortest path method for determining the order (M_1, N_1, M_2, N_2) of a higher-order two dimensional ARMA model.

Step 1: Find the spike in the plot of $DR_i(I') = \dfrac{\det(A_{j_{(j-1)i}(J-1)i})}{\det(A_{i_{j_i - i_i}})}, J = 1,2\ldots\ i = 1,2.$

Step 2: Find the spike in the plot of $DR_i(I'') = \dfrac{\det(A_{i_{k_i},J_i})}{\det(A_{i_{(k+1)i},\,j_i})}, k = 0,1\ldots\ j-1, i = 1,2.$

Step 3: Determine the order M_i, N_i for i=1,2. For example, if the spike location from step 1 occurs at $\dfrac{\det(A_{1,3,3})}{\det(A_{1,4,4})}, \dfrac{\det(A_{2,4,4})}{\det(A_{2,5,5})}$, and from step 2 a spike occurs at $\dfrac{\det(A_{1,2,4})}{\det(A_{1,3,4})}, \dfrac{\det(A_{2,3,5})}{\det(A_{2,4,5})}$, then an ARMA model with order $(M_1 = 2, N_1 = 3, M_2 = 3, N_2 = 4)$ is identified.

4. The 2D forgetting factor

The forgetting factor is a data weighting process that gives more weight to recent observations in time and space and less weight to earlier data. The use of the forgetting factor in time series analysis has attracted considerable interest in recent years. For example, Chen *et al* (2006) utilise a forgetting factor in subset autoregressive modelling of the spot aluminium and nickel prices on the London

Metal Exchange. In 2-D cases, no previous work investigates the forgetting factor approach in combination with applications of time-series modelling in data analysis. Our new 2-D and high-D forgetting factor approaches are new and unique and conceptually well-advanced. In particular, our new techniques allow detailed assessment of how interactions between variables evolve over time by appropriately discounting irrelevant distant interactions, thus giving more weight to relevant near interactions. This research is at the cutting edge of the application of statistics to the data analysis field and yet we are confident it has some practical and tangible outcomes.

Consider a 2-D ARMA model of the following form:

$$y(x_{t,r})+\sum_{i=0}^{p}\sum_{j=0}^{q}a_{i,j}y(x_{t,r}-x_{t-i,r-j})+\sum_{\substack{i=0\\ \begin{cases}i\neq j\\ i=j\neq 0\end{cases}}}^{p}\sum_{j=0}^{q}b_{i,j}\varepsilon(x_{t-i,r-j})=\varepsilon(x_{t,r}).$$

The coefficients $a_{i,j}$ are obtained by minimising

$$\sum_{t=1}^{T}\sum_{r=1}^{R}\kappa\left[y(x_{t,r})+\sum_{i=0}^{p}\sum_{j=0}^{q}a_{i,j}y(x_{t,r}-x_{t-i,r-j})+\sum_{\substack{i=0\\ \begin{cases}i\neq j\\ i=j\neq 0\end{cases}}}^{p}\sum_{j=0}^{q}b_{i,j}\varepsilon(x_{t-i,r-j})\right]^2,$$

where κ is the forgetting profile. Following O'Neill *et al* (2007), a strategy for determining κ is as follows: $\kappa=\lambda^{\sqrt{(T-t)^2+(R-r)^2}}$ if $1\leq t\leq T$ or $1\leq r\leq R$; and $=1$ if T-t=0 and R-r=0. (7)

In (7) λ denotes the forgetting factor. Therefore $\kappa = \lambda^{R-r}$ if T-t=0; $\kappa=\lambda^{T-t}$ if R-r=0. Equation (7) means that 'forgetting' of the past or distant occurs from time point [T, R]. If λ=1 for every [t,r] then we obtain the ordinary least squares solution. If $0<\lambda<1$, the past and distant are weighted down geometrically from [t,r]. In theory, the value of λ could be different at different [t, r] (a so-called variable forgetting factor). For simplicity, we only consider the fixed forgetting factor case in which the value of λ is constant for different [t, r].

5. Recursive estimations of ARMA modelling

We seek a recursive algorithm for the estimation of a zero-mean, bivariate process x(m,n) embedded in y(m,n). The observation noise v(m,n) defined by

$$y(m,n) = x(m,n) + v(m,n)$$

is assumed, be additive, white, with power V, and independent of x(m,n). The autocorrelation matrix of x(m,n) is assumed to be stationary and known. The natural "past" of each observation y(m,n) consists of all y(i,j) with (i,j) in the set S(m,n) defined below and depicted in Figure 1.

$S(m,n) = \{(i,j) : 0 \leq i \leq m-1)\ 0 \leq j\}\ U\ \{(i,j) : i = m,\ 0 \leq j \leq n-l\}$

The optimum, mean-square, one-step prediction of y(m,n) is denoted by $\hat{y}$ (m,n):

$\hat{y}(m,n) = E\ \{y(m,n)/y(i.j),\ (i,j) \in S(m,n)\}$

The optimum, mean-square, filtered estimate of x(m,n) given both past and present observations will be denoted by $\hat{x}(m,n) := E\ \{x(m,n)/y(m,n)\ U\ y(i,j),\ (i,j) \in S(m.n)\}$ The form of recursive algorithms reduce to $\hat{x}(m,n) = a(m,n)\,\hat{y}(m,n) + b(m,n)\,y(m,n)$. We wish to consider algorithms where $\hat{y}$(m,n) is based on a limited portion of the past and we will compare only the steady state version of these algorithms (i.e. determine and consider only the steady state values of a(m,n), b(m,n) and the estimation parameters of $\hat{y}$(m,n)). By using (2), we now defines, for $m > N_1,\ n > N_2$

$$w(m,n) \underline{\Delta} \sum_{r=0}^{M_1}\sum_{s=0}^{M_2} a_{rs} u(m-r)u(n-s)y(m-r,n-s). \tag{8}$$

With this formulation one can see that $\hat{y}(m,n) = -\sum_{\substack{i=0 \\ i+j\neq 0}}^{N_1}\sum_{j=0}^{N_2} a_{ij}y(m-i,n-j) + \hat{w}(m,n).$

This is a convenient formulation because R_w and R_{wy} are non-zero in a limited region and $\hat{w}$(m,n) becomes relatively easy to determine. It has been shown that:

$$\hat{w}(m,n) = \sum_{T(m,n)} \frac{R_{we}(m,i;n,j)}{R_{we}(i,i;j,j)}[y(i,j) - \hat{y}(i,j)], \tag{9}$$

where ε (i,j) is the error (y(i,j) - $\hat{y}$(.i,j)] or equivalently [w(i,j) - $\hat{w}$(i,j)]. The cross-correlation between w(m,n) and ε(p,q) can be found from

$$R_{w\varepsilon}(m,p;n,q) = R_w(m,p;n,q) - \sum_{T(m,n)\cap T(p,q)} \frac{R_{w\varepsilon}(m,i;n,j)}{R_{w\varepsilon}(i,i;j,j)} R^T_{w\varepsilon}(p,i;q,j),$$

for $T(m,n) = S(m,n) - \Lambda(m,n)$, where

$\Lambda(m,n) = \{(i,j) : 0 \leq i < m - N_1, 0 \leq j\} \cup \{(i.j) : i = m - N_1, 0 \leq j < n - N_2\}$,

and $T(m,n)$ is shown in Figure 2. R(m,p; n,q) is readily determined from equation (8). The mean square error e(i,j) is given by $e(i,j) = R_{w\varepsilon}(i,i;j,j) = R_{y\varepsilon}(i,i;j,j)$ (10)

Thus, equations (8 – 10) defines a recursive algorithm based on previous estimates and values limited to T(m,n). This algorithm converges rather rapidly enabling one to find the steady state weight of $\hat{w}$(m,n) and the steady state mean square error e(i,j). Usually the number of "significant" weights are far fewer than the number of values in T(m,n) resulting in a very near optimum algorithm that is easily implemented.

6. An application of 2D ARMA to a healthcare chromosome picture

A noisy 128 by 96 chromosome picture is processed with the 2D ARMA model. White Gaussian zero-mean noise is added of unit variance to produce a SNR of 1. We apply a forgetting factor with the value 0.998 to the system involved. A (1, 1; 1, 1) ARMA model has been identified for the system. To start recursions, we set the initial condition region covered by the upper and left hand edges. It is necessary to set the initial condition region; otherwise it will take a very long time to reach the steady state. Finally convergence does take place after 8 lines or so. We have $w(m,n) = y(m,n) - 0.776y(m,n-1) - 0.506y(m-1,n) + 0.346y(m-1,n-1)$; and $\hat{w}(m,n) = -0.61\varepsilon(m,n-1) - 0.04\varepsilon(m,n-2) + 0.28\varepsilon(m-1,n-1) - 0.42\varepsilon(m-1,n) + 0.03\varepsilon(m-1,n+1)$ respectively. The signal's steady state estimate is $\hat{x}(m,n) = 0.31y(m,n) + 0.69\hat{y}(m,n)$, where $\hat{y}(m,n) = 0.16y(m,n-1) - 0.04y(m,n-2) + 0.09y(m-1,n) +$

$$0.06y(m-1,n-1) + 0.03y(m-1,n+1) + 0.61\hat{y}(m,n-1) + 0.04\hat{y}(m,n-2) +$$
$$+0.42\hat{y}(m-1,n) + 0.42\hat{y}(m-1,n) - 0.28\hat{y}(m-1,n-1) - 0.03\hat{y}(m-1,n+1).$$

The mean square error is 0.30 and the filter performance is given by N = 5.05 db.

7. Conclusions

A procedure for identifying the ND closed-order ARMA model has been designed and presented. Further, a formulation of the 2D forgetting factor approach is established. Modelling healthcare data such as chromosome pictures with finite order ARMA models have been demonstrated. The prediction and estimation algorithm has been processed, and a steady state estimate of the signal has been easily reached.

References

A. Chen, J. Penm and R.D. Terrell, An evolutionary recursive algorithm in selecting statistical subset neural network/VDL filtering. Journal of Applied Mathematics and Decision Sciences, Article ID 46592(2006), 1-12.

T.J. O'Neill, J. Penm and J. Penm, a subset polynomial neural networks approach for breast cancer diagnosis. International Journal of Electronic Healthcare, 3, (2007), 293–302.

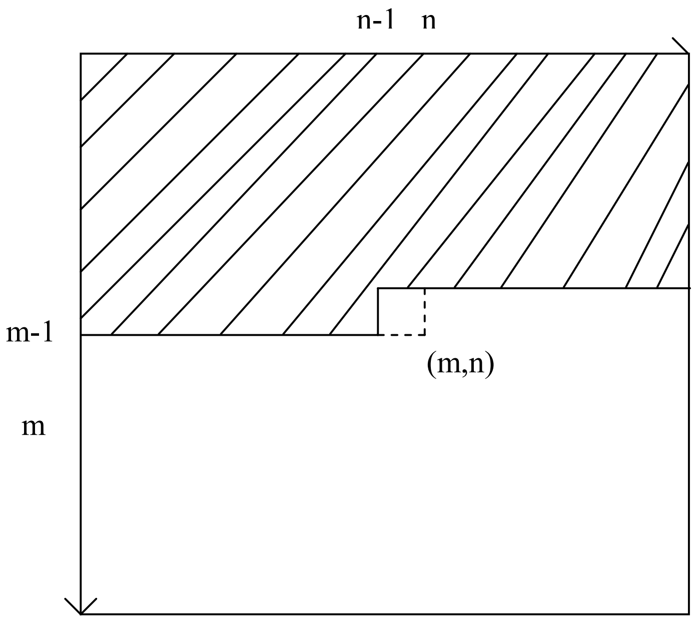

Figure 1

$$S(m,n) = [(i,j): 0 \le i \le m-1; 0 \le j] \bigcup [(i,j): i = m; 0 \le j \le n-1]$$

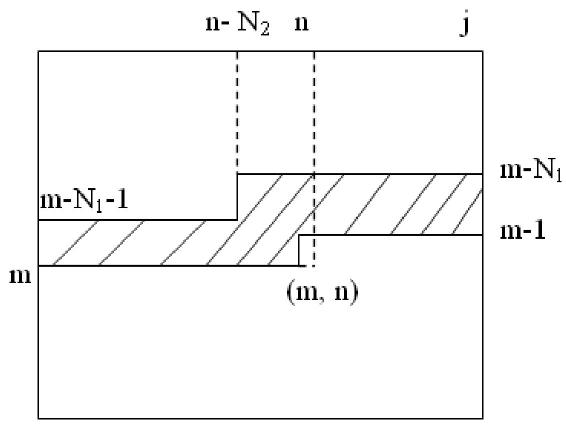

Figure 2 T(m,n)

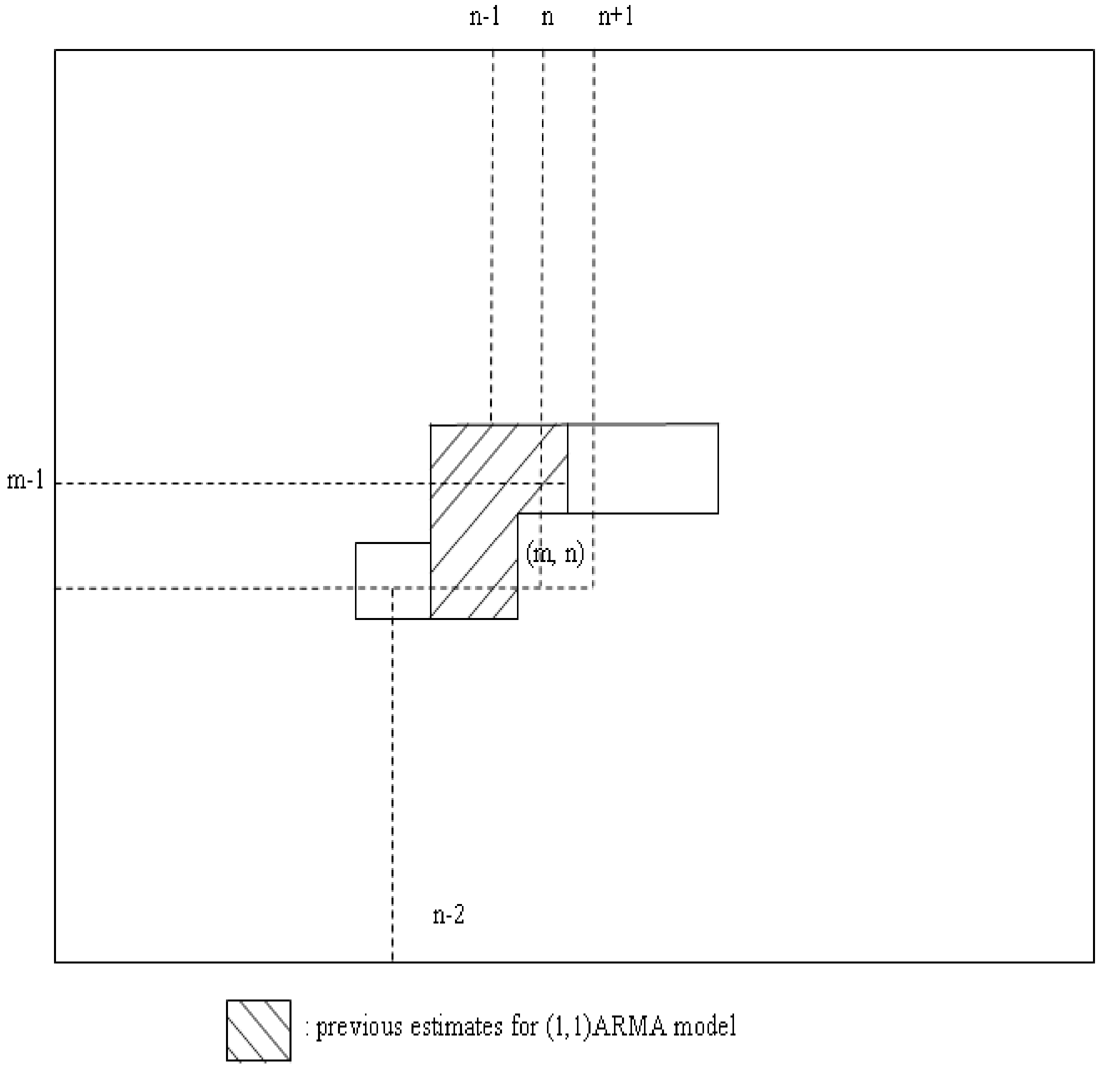

Figure 3 steady state estimates

Received: November, 2009