Research Article

Parameter Estimation for p-Order Random Coefficient Autoregressive (RCA) Models Based on Kalman Filter

Mohammed Benmoumen ⓘD, Jelloul Allal, and Imane Salhi

LaMSD, Faculty of Sciences, Mohammed Premier University, Oujda, Morocco

Correspondence should be addressed to Mohammed Benmoumen; med.benmoumen@gmail.com

In this paper we elaborate an algorithm to estimate p-order Random Coefficient Autoregressive Model (RCA(p)) parameters. This algorithm combines quasi-maximum likelihood method, the Kalman filter, and the simulated annealing method. In the aim to generalize the results found for RCA(1), we have integrated a subalgorithm which calculate the theoretical autocorrelation. Simulation results demonstrate that the algorithm is viable and promising.

1. Introduction

Random Coefficient Autoregressive (*RCA*) processes have been widely studied in the literature for modeling time series exhibiting nonlinear behavior. The *RCA* process was introduced by Andel (1976) who also studied its properties. He had obtained conditions for the existence of a singly infinite process which is second-order stationary satisfying:

$$X(t) = \sum_{k=1}^{p} \left(\phi_k + \beta_k(t) \right) X(t-k) + \varepsilon(t) \tag{1}$$

where ϕ_k's are fixed coefficients, $(\beta_k(t), t = 1, \ldots, p)$ is a sequence of i.i.d. random vectors with mean zero and constant covariance matrix $C = (\sigma_{i,j})$, $\varepsilon(t)$ is a white noise process with mean zero and variance σ_ε, and $\beta_k(t)$ and $\varepsilon(t)$ are assumed to be mutually independent.

Multiple studies have emerged after Conlisk (1974, 1976) has derived conditions for the stability of RCA models, Robinson (1978) has considered statistical inference for the RCA model, and Nicholls and Quinn [1] have extended the results of Andel to the multivariate RCA model.

Many authors have investigated the estimation of parameters in the random coefficient autoregressive. Among them, we cite Nicholls and Quinn who obtained the least squares and the maximum likelihood estimates. Under certain assumptions, they established the strong consistency as well

as the asymptotic normality for both estimates. For a detailed early study, we refer to Nicholls and Quinn [1]. On their side, Thavaneswaran and Abraham [2] apply Godambe's theorem (1985) to obtain optimal estimates for *RCA* models.

Recently several authors are interested in the *RCA* model. Aue and Horvath (2011) propose a unified quasi-likelihood procedure for the estimation of the unknown parameters of RCA(1) models that works for both stationary and nonstationary processes. They also establish the weak consistency and the asymptotic normality for this procedure.

Liang et al. (2012) have described moment properties for *RCA* processes and the corresponding squared processes; they have given also a joint prediction study of the mean and volatility.

A new algorithm was proposed by Allal and Benmoumen [3] to estimate first-order RCA's parameters. This algorithm combines quasi-maximum likelihood method, the Kalman filter, and the Powell's method. Our contribution aims to extend previous method to higher-order of RCA process.

As we shall see, the proposed algorithm exacts initial values. To provide the concerned values we are obliged to calculate the theoretical autocorrelation of *p*-order RCA models. To ward off tedious computation we implement an algorithm, more details in Section 4, who calculate the autocorrelations in numerical way.

This paper is organized as follows. In Section 2, we present definition and some basic properties of RCA models.

In Section 3 we recall Kalman's algorithm and apply it to calculate likelihood function for RCA models. In Section 4, we describe our estimating algorithm. The performance of this algorithm is examined by Monte Carlo simulations and compared to the quasi-maximum likelihood method, in Section 5. Finally, we achieve with a conclusion.

2. Stationarity and Moment Properties

Let $(X(t), t \in \mathbb{Z})$ be a $RCA(p)$ model. In order to derive stationarity conditions for the processes $X(t)$, Nicholls and Quinn proposed the following vectorial presentation.

Define the p random vector $Y(t)$ by $Y(t) = [X(t + 1 - p), \ldots, X(t)]'$, equation (1) may be rewritten in terms of $Y(t)$ by

$$Y(t) = (M + D(t)) Y(t - 1) + \eta(t) \quad (2)$$

where the $p \times p$ matrix M is given by $M = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & 0 & \cdots & 1 \\ \phi_p & \phi_{p-1} & \cdots & \phi_1 \end{pmatrix}$, $D(t) = L' \otimes B(t)$, where $B(t) = [\beta_p(t), \ldots, \beta_1(t)]$ and $L = [0, \ldots, 0, 1] \in \mathscr{M}_{p,1}$ and $\eta(t) = L' \otimes \varepsilon(t)$.

Theorem 1. *If $\mathscr{F}_t$ is the σ−field generated by $(\epsilon(s), \beta(s)), s \leq t$ and $C > 0$ then there exists a unique $\mathscr{F}_t$−measurable second-order stationary solution to (1) if and only if M has all its eigenvalues within the unit circle and $(\text{vec } C)' A < 1$, where $C = E[B(t) B'(t)]$ and A is the last column of the matrix $(Id − M \bigotimes M)^{-1}$ (see Andel (1976) or Nicholls and Quinn [1]).*

In the next theorem, Lian et al. [4] give us autocorrelation structure and marginal variance of RCA(p) models.

Theorem 2. *Consider the stationary RCA(p) in (1).*

(i) *The processes have the same autocorrelation structure as the AR(p) processes:*

$$\rho_k = \sum_{j=1}^{p} \phi_j \rho_{k-j}, \quad k \geq 1 \quad (3)$$

(ii) *The marginal variance of a stationary RCA(p) process is given by*

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \sum_{j=1}^{p} \sum_{k=1}^{p} \left(\phi_j \phi_k + \sigma_{jk} \right) \rho_{j-k}} \quad (4)$$

In practice, this theorem will enable us to deduce the starting conditions for Kalman filter. Hence, we develop in Section 3 a recursive algorithm called variance to calculate the autocovariance function for RCA models.

3. Quasi-Maximum Likelihood and Kalman Filter

Assuming the joint normality of $(\epsilon(t), t \in \mathbb{Z})$ and $(B(t), t \in \mathbb{Z})$, the conditional log-likelihood is given by

$$L(x; \theta) = \log \left(\prod_{t=1}^{n} f\left(x_t \mid \chi_{t-1} \right) \right) = \sum_{t=1}^{n} \log \left(f\left(x_t \chi_{t-1} \right) \right) \quad (5)$$

where θ is the vector of unknown parameters, $\chi_t = (x_t, \ldots, x_t)$ is a sample of observations at $t = 1, \ldots, n$, and $f(x_t \mid \chi_{t-1})$ indicate the normal density function of x_t given χ_{t-1}, with mean $\widehat{x}_{t|t-1} = E[X_t \mid \chi_{t-1}]$ and variance $\widehat{M}_{t|t-1} = E[(x_t - \widehat{x}_{t|t-1})^2]$. The quasi-maximum likelihood may be put in the following form:

$$L(x; \theta) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^{n} \log \left(\widehat{M}_{t|t-1} \right)$$
$$- \frac{1}{2} \sum_{t=1}^{n} \frac{\left(x_t - \widehat{x}_{t|t-1} \right)^2}{\widehat{M}_{t|t-1}} \quad (6)$$

Given this formula the quasi-maximum likelihood can be calculated using the Kalman filter; see Hamilton [5].

In order to apply Kalman filter, we consider the appropriate state-space representation for RCA model in (1).

$$Z_{t+1} = FZ_t + G\nu_t + e_t; \quad \text{state equation,}$$
$$X_t = HZ_t; \quad \text{observation equation} \quad (7)$$

where $Z_t = (X_t, \ldots, X_{t-p+1})' \in \mathscr{M}_{p,1}(\mathbb{R})$, $H = (1, 0, \ldots, 0) \in \mathscr{M}_{1,p}(\mathbb{R})$, $e_t = (\varepsilon_t, 0, \ldots, 0)' \in \mathscr{M}_{p,1}(\mathbb{R})$, $\nu_{t-1} = (X_{t-1}\beta_1(t), \ldots, X_{t-p}\beta_p(t)) \in \mathscr{M}_{p,1}(\mathbb{R})$

$$F = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \in \mathscr{M}_{p,p}(\mathbb{R}),$$

$$G = \begin{pmatrix} 1 & \cdots & 1 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathscr{M}_{p,p}(\mathbb{R}), \quad (8)$$

$$e_t \sim \mathscr{N}(0, Q) \quad \text{where } Q = \begin{pmatrix} \sigma_\epsilon^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

We should point out that the previous representation was proposed by Benmoumen in his M.Sc. Thesis [6].

Now, we describe the Kalman filter in the aim to build the log-likelihood function. The Kalman filter is a recursive algorithm derived by Kalman [7] to provide an optimal forecast $\widehat{Z}_{t+1|t} = E[Z_{t+1} \mid X_1, \ldots, X_t]$ of Z_{t+1} given $X_1, \ldots, X_t$, with the mean square error $P_{t+1|t} = V[Z_{t+1} - \widehat{Z}_{t+1|t}], t = 1, \cdot, n$. Given starting values $\widehat{Z}_{1|0}$ and $P_{1|0}$ which are derived from Theorem 2, the recursive procedure is as follows:

(i) Calculate the forecasting $\widehat{X}_{t|t-1}$ of the observation X_t, and the error $\widehat{M}_{t|t-1}$ of this forecast.

(ii) Update the state vector $\widehat{Z}_{t|t}$. Compute $P_{t|t}$ the MSE of this updated projection.

(iii) Calculate the forecasting $\widehat{Z}_{t+1|t}$ and the MSE $P_{t+1|t}$ of this forecast.

Thus, we could construct the log-likelihood function using Kalman filter, so as to obtain the maximum likelihood estimators and in order to avoid the fastidious computation of partial derivatives of $L(x;\theta)$ we used the simulated annealing method (see Corana et al. [8]) which is a global optimization algorithm.

4. Estimating Algorithm for p-Order Random Coefficient Autoregressive Model parameters

The algorithm proposed here is a generalization of a procedure developed by Benmoumen and al. (2013) to estimate first-order RCA's parameters.

Recently, the same idea has been developed for parameters estimation in GARCH(1,1), ARCH(1), and ARCH(p) models by Benmoumen et al. (2011, 2014, and 2015).

Before describing our algorithm MLKF (quasi-maximum likelihood and Kalman filter estimation), it is worthwhile to provide a subalgorithm which tests if parameters fulfill the conditions of stationarity; we will denote it by Test. The second subalgorithm, which we must provide, concerns the computation of $L(x;\theta)$ by Kalman filter; we will denote it by KF. These two subalgorithms will be implemented in our global estimating algorithm.

Herein, we are interested in minimizing $-L(x;\theta)=\ell(x;\theta)$.

Subalgorithm Test(θ)

if The eigenvalues of the matrix M have modul less than unity and $CA < 1$

  then
   Then go to next
else
   Take the last point as starting point

end if
End Subalgorithm

Subalgorithm variance(θ)

Solve the equations: $A[\rho_1, \cdots, \rho_p]' = [0, \cdots, 0]'$

  Where

A_{ij}

$$= \begin{cases} \phi_{i-1} & \text{si } j = 1; \ i = 2, \cdots, p+1, \\ \phi_{i-j} + \phi_{i+j-2} & \text{si } j = 2, \cdots, p+1; \ i = 2, \cdots, p+1. \end{cases} \quad (9)$$

And

$$\phi_k = \begin{cases} -1 & k = 0, \\ 0 & k > p \end{cases} \quad (10)$$

Calculate:

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \sum_{j=1}^{p} \sum_{k=1}^{p} \left(\phi_j\phi_k + \sigma_{jk}\right)\rho_{j-k}} \quad (11)$$

End Subalgorithm

Subalgorithm KF(θ)

Given the starting conditions $\widehat{Z}_{0|1}$ and $P_{0|1}$

Calculate: $\widehat{X}_{0|1} = H\widehat{Z}_{0|1}$ and $\widehat{M}_{0|1} = HP_{0|1}H'$

for $t = 1$ to n **do**

  $K_t, \widehat{Z}_{t|t}, P_{t|t}, \widehat{Z}_{t+1|t}$ and $P_{t+1|t}$
  $\widehat{X}_{t+1|t}$ and $\widehat{M}_{t+1|t}$

end for

$$som = \frac{1}{2}\log(2\pi) \quad (12)$$

for $t = 1$ to n **do**

$$som = som + \frac{1}{2}\log\left(\widehat{M}_{t+1|t}\right) + \frac{1}{2}\frac{\left(X_t - \widehat{X}_{t+1|t}\right)^2}{\widehat{M}_{t+1|t}} \quad (13)$$

end for

$$\ell(x;\theta) = som \quad (14)$$

End Subalgorithm

Indeed, as we shall see later our algorithm is an iterative process requiring initial estimates of the parameters to commence the iterations. The consistent least squares estimates are suitable for this purpose.

MLKF Algorithm

Step 1: Initialize: the vector parameters θ the step vector ν and the temperature T.

Step 2: Starting from the point θ_i, generate a random point θ along the direction h: $\theta = \theta_i + r\nu_{m_h}e_h$, where r is a random number generated in the range $[-1, 1]$ by a pseudorandom generator; e_h is the vector of the hth coordinate direction; and ν_{m_h} is the component of the step vector ν along the same direction.

Step 3: Call sub algorithm $Test(\theta)$

Step 4: Call sub algorithm $KF(\theta)$

  Compute $KF(\theta_i)$ and $KF(\theta)$

  If $KF(\theta) \leq KF(\theta_i)$ accept the new point

Else accept or reject the new point with acceptance probability p:

$$p = \exp\left(\frac{KF(\theta_i) - KF(\theta)}{T}\right) \quad (15)$$

generate a uniformly distributed random number p' in the range $[0, 1]$

If $p' < p$, the point is accepted otherwise it is rejected.

Step 5: Steps 2 to 4 are repeated for each coordinate direction i, $i = 1, \ldots, m$ (m is the dimension of the vector parameter).

Step 6: Steps 2 to 5 are repeated N_s times (N_s is the number of step variation) and the step vector v is adjusted.

Step 7: Steps 2 to 6 are repeated N_T times (N_T is the number of temperature reduction) the temperature is reduced following the rule: $T' = r_T T$ with $r_T \in [0, 1]$.

Step 8: Steps 2 to 7 are repeated until a termination criterion is satisfied.

End Algorithm

5. Simulations

To examine the performance of our algorithm, we have carried out series of simulation experiments. In this study, we consider two examples of models RCA.

(1)

$$X(t) = \beta_1(t) X(t-1) + (0.36 + \beta_2(t)) X(t-2)$$
$$+ \varepsilon(t), \quad t \in \mathbb{R} \quad (16)$$

with $(\varepsilon(t), \beta(t)) \sim \mathcal{N}((0, 0, 0), \Omega)$, where $\Omega = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.2176 & 0 \\ 0 & 0 & 0.2176 \end{pmatrix}$.

(2)

$$X(t) = (0.8 + \beta_1(t)) X(t-1)$$
$$+ (-0.15 + \beta_2(t)) X(t-2) + \varepsilon(t), \quad t \in \mathbb{R} \quad (17)$$

with $(\varepsilon(t), \beta(t)) \sim \mathcal{N}((0, 0, 0), \Omega)$, where $\Omega = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.0919 & 0.0919 \\ 0 & 0.0919 & 0.1838 \end{pmatrix}$

For the models mentioned earlier, we generated 1000 replications of sample sizes $n = 100$ and 200.

The results of this experiment are displayed in Tables 1–4 where for each estimator we give the mean and MSE, where we used notation QMLE for the quasi-maximum likelihood estimators and MLKF for the estimation by our algorithm.

Remark 3.

(i) The least squares estimate of $\theta = (\phi, \gamma, \sigma_\varepsilon^2)$, where $\phi = [\phi_1, \ldots, \phi_p]'$ and $\gamma = vech(C)$ is given by

TABLE 1: Mean and MSE of estimated parameters for Example 1; $n = 100$.

Parameters	MLKF Mean	MLKF MSE	QMLE Mean	QMLE MSE
$\phi_1 = 0$	-0.00032	0.01491	-0.00051	0.03021
$\phi_2 = 0.36$	0.32874	0.01265	0.29532	0.03127
$\sigma_{11} = 0.2176$	0.20847	0.01506	0.17879	0.03414
$\sigma_{12} = 0$	0.00416	0.00635	0.00631	0.00827
$\sigma_{22} = 0.2176$	0.23316	0.01197	0.18413	0.03698
$\sigma_\varepsilon = 1$	1.05658	0.05597	1.10734	0.07859

TABLE 2: Mean and MSE of estimated parameters for Example 1; $n = 200$.

Parameters	MLKF Mean	MLKF MSE	QMLE Mean	QMLE MSE
$\phi_1 = 0$	0.00021	0.00523	0.00038	0.02003
$\phi_2 = 0.36$	0.34857	0.00556	0.31857	0.02114
$\sigma_{11} = 0.2176$	0.19859	0.00902	0.17016	0.02487
$\sigma_{12} = 0$	0.00363	0.00412	0.00578	0.00723
$\sigma_{22} = 0.2176$	0.21072	0.01041	0.19189	0.02865
$\sigma_\varepsilon = 1$	1.02176	0.03397	1.09754	0.06682

TABLE 3: Mean and MSE of estimated parameters for Example 2; $n = 100$.

Parameters	MLKF Mean	MLKF MSE	QMLE Mean	QMLE MSE
$\phi_1 = 0.8$	0.82356	0.01564	0.75590	0.03258
$\phi_2 = -0.15$	-0.15886	0.00982	-0.18164	0.02042
$\sigma_{11} = 0.0919$	0.08326	0.01013	0.13584	0.04782
$\sigma_{12} = 0.0919$	0.03735	0.01126	0.02104	0.03152
$\sigma_{22} = 0.1838$	0.21214	0.01342	0.24065	0.03356
$\sigma_\varepsilon = 1$	0.93012	0.05263	0.89702	0.08237

TABLE 4: Mean and MSE of estimated parameters for Example 2; $n = 200$.

Parameters	MLKF Mean	MLKF MSE	QMLE Mean	QMLE MSE
$\phi_1 = 0.8$	0.79036	0.00523	0.76137	0.02278
$\phi_2 = -0.15$	-0.17777	0.00358	-0.17776	0.00935
$\sigma_{11} = 0.0919$	0.10248	0.00647	0.08250	0.00838
$\sigma_{12} = 0.0919$	0.06556	0.00512	0.05217	0.00912
$\sigma_{22} = 0.1838$	0.21036	0.00788	0.23481	0.01014
$\sigma_\varepsilon = 1$	0.97054	0.01463	0.94957	0.03327

$$\widehat{\phi} = \sum_{t=1}^{N}\left(Y(t-1)Y'(t-1)\right)^{-1}\sum_{t=1}^{N}Y(t-1)X(t)$$

$$\widehat{\sigma}_\varepsilon^2 = \frac{1}{N}\sum_{t=1}^{N}\widehat{u}^2(t) - \widehat{\gamma}\overline{z} \quad (18)$$

$$\widehat{\gamma} = \left[\sum_{t=1}^{N}(z(t)-\overline{z})(z(t)-\overline{z})'\right]^{-1}\sum_{t=1}^{N}\widehat{u}^2(t)(z(t)-\overline{z})$$

where $\widehat{u}(t) = X(t) - \widehat{\phi}'Y(t-1), t = 1, \ldots, N$, $z(t) = K_n \text{vec}(Y(t-1)Y'(t-1))$, and $\overline{z}(t) = (1/n)\sum_{t=1}^{n}z(t)$.

(ii) The maximum likelihood estimate of θ is defined by minimizing the following function:

$$
\begin{aligned}
\widehat{\ell}_n(\theta) = {} & \left(\frac{1}{n}\right) \sum_{t=1}^{n} \ln\left(\sigma_\epsilon^2 + \gamma' z(t)\right) \\
& + \left(\frac{1}{n}\right) \sum_{t=1}^{n} \frac{\left(X(t) - \beta' Y(t-1)\right)^2}{\sigma_\varepsilon^2 + \gamma' z(t)}
\end{aligned}
\tag{19}
$$

Under certain assumptions, both of the estimates, the least squares and the maximum likelihood, are strongly consistent and obey a central limit theorem. In fact, condition $E(X^4(t)) < \infty$ is required for the strong consistency of the least squares estimates, and $E(X^8(t)) < \infty$ for the asymptotic normality of these estimators. Respectively, in order that a central limit theorem exist for maximum likelihood estimators, conditions $E(\varepsilon^4(t)) < \infty$ and $E(\beta_k^4(t)) < \infty$, $k = 1, \ldots, p$ will be required. For more details see Nicholls and Quinn [1].

5.1. Comparison with Quasi-Maximum Likelihood Estimators. In this series of simulation we compare our algorithm (MLKF) versus the quasi-maximum likelihood (QMLE). In each method we use simulated annealing algorithm for optimization and the least squares estimators for initiation.

As is to be expected, the MLKF estimation procedure has performed better, as is seen from the fact that the sample mean square errors (MSE) are generally smaller than for the quasi-maximum likelihood estimators (QMLE). Hence, we can conclude that the performance of our estimation procedure is promising.

6. Conclusion

In this paper, we constructed an algorithm for calculating the covariance matrix of RCA(p) models and we have generalized a procedure developed by Benmoumen et al. (2013) to estimate first-order RCA's parameters.

The log-likelihood function is constructed using the Kalman filter and is numerically maximized applying simulated annealing method. The results of our simulation study show that our estimation approach succeeds and it performs better than the competitor.

Data Availability

Only computer-generated data have been used so all researchers can find our results from the application of our algorithms and computer-simulated data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] D. F. Nicholls and B. G. Quinn, *Random Coefficient Autoregressive Models: An Introduction*, Springer-Verlag, New York, NY, USA, 1982.

[2] A. Thavaneswaran and B. Abraham, "Estimation for nonlinear time series models using estimating equations," *Journal of Time Series Analysis*, vol. 9, no. 1, pp. 99–108, 1988.

[3] J. Allal and M. Benmoumen, "Parameter estimation for first-order Random Coefficient Autoregressive (RCA) Models based on Kalman Filter," *Communications in Statistics—Simulation and Computation*, vol. 42, no. 8, pp. 1750–1762, 2013.

[4] Y. Lian and A. Thavaneswaran, "RCA models: joint prediction of mean and volatility," *Statistics & Probability Letters*, vol. 83, no. 2, pp. 527–533, 2013.

[5] J. D. Hamilton, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey, 1994.

[6] M. Benmoumen, *Algorithmes Pour L'Inférence Statistique Dans Les Modèles RCA et GARCH basée sur le Filtre de Kalman, thèse De Doctorat*, Université Mohammed Premier, 2012.

[7] R. E. Kalman, "A new approach to linear filtering and prediction problems; transactions of the ASME," *Journal of Basic Engineering*, vol. 82, pp. 34–45, 1960.

[8] A. Corana, M. Marchesi, C. Martini, and S. Ridella, "Minimizing Multimodal functions of continuous variables with simulated annealing Algorithm," *ACM Transactions on Mathematical Software*, vol. 35, pp. 262–280, 1987.