KYBERNETIKA — VOLUME 47 (2011), NUMBER 4, PAGES 501–518

ON A CLASS OF ESTIMATORS IN A MULTIVARIATE RCA(1) MODEL

Zuzana Prášková and Pavel Vaněček

This work deals with a multivariate random coefficient autoregressive model (RCA) of the first order. A class of modified least-squares estimators of the parameters of the model, originally proposed by Schick for univariate first-order RCA models, is studied under more general conditions. Asymptotic behavior of such estimators is explored, and a lower bound for the asymptotic variance matrix of the estimator of the mean of random coefficient is established. Finite sample properties are demonstrated in a small simulation study.

Keywords: multivariate RCA models, parameter estimation, asymptotic variance matrix

Classification: 60F05, 60G10, 60G46, 62M10

1. INTRODUCTION

Random coefficient autoregressive models belong to a broad class of conditional heteroscedastic time series models because of their varying conditional variance and as such may be used in various applications.

We say that a process of random vectors $\boldsymbol{X}_t = (X_t^1, \dots, X_t^m)' \in \mathbb{R}^m, t \in \mathbb{Z}$, follows the multivariate first-order random coefficient autoregressive model, abbreviated as RCA(1), if $\boldsymbol{X}_t$ for each $t \in \mathbb{Z}$ satisfies

$$\boldsymbol{X}_t = (\boldsymbol{\beta} + \boldsymbol{B}_t)\boldsymbol{X}_{t-1} + \boldsymbol{Y}_t, \tag{1}$$

where $\boldsymbol{\beta}$ is an $m \times m$ matrix of (unknown) parameters, $\{\boldsymbol{B}_t, t \in \mathbb{Z}\}$ is a sequence of $m \times m$ random matrices and $\{\boldsymbol{Y}_t, t \in \mathbb{Z}\}$ is an $m \times 1$ random error process.

Equation (1) can be rewritten into

$$\boldsymbol{X}_t = \boldsymbol{\beta}\boldsymbol{X}_{t-1} + \boldsymbol{u}_t \tag{2}$$

with the new error process $\boldsymbol{u}_t = \boldsymbol{B}_t\boldsymbol{X}_{t-1} + \boldsymbol{Y}_t = (\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I}) \cdot \operatorname{vec}(\boldsymbol{B}_t) + \boldsymbol{Y}_t$, where $\boldsymbol{I}$ is the $m \times m$ identity matrix, $\otimes$ denotes the Kronecker product and vec is the matrix column operator (for definitions and properties of matrices operators see the Appendix and Lemmas A.2 – A.4 there).

To specify model (1) in detail, we introduce the following assumptions.

A1: The random coefficient matrix process $\{\boldsymbol{B}_t, t \in \mathbb{Z}\}$ is a centered *iid* sequence with a finite positive definite matrix $\boldsymbol{\Sigma} = \mathrm{E}[\mathrm{vec}\,(\boldsymbol{B}_0) \cdot \mathrm{vec}\,'(\boldsymbol{B}_0)]$.

A2: All the eigenvalues of the matrix $\mathrm{E}(\boldsymbol{B}_0 \otimes \boldsymbol{B}_0) + (\boldsymbol{\beta} \otimes \boldsymbol{\beta})$ are less than unity in modulus.

A3: The error process $\{\boldsymbol{Y}_t\}$ is an ergodic and strictly stationary martingale difference sequence with respect to the filtration $\mathcal{F}_t = \sigma(\boldsymbol{B}_s, \boldsymbol{Y}_s; s \le t)$, such that $\mathrm{E}[\boldsymbol{Y}_t \boldsymbol{Y}_t' | \mathcal{F}_{t-1}] = \boldsymbol{G}$ *a.s.* for all t, where $\boldsymbol{G}$ is a finite positive definite matrix.

A4: $\{\boldsymbol{B}_t, t \in \mathbb{Z}\}$ and $\{\boldsymbol{Y}_t, t \in \mathbb{Z}\}$ are mutually independent.

The following theorem is of the fundamental importance.

Theorem 1.1. Under the assumptions A1 to A4, there exists a unique solution to the stochastic difference equation (1) that is $\mathcal{F}_t$-measurable, strictly stationary, ergodic, and is of the form

$$\boldsymbol{X}_t = \boldsymbol{Y}_t + \sum_{j=1}^{+\infty} \left[\prod_{i=0}^{j-1} (\boldsymbol{\beta} + \boldsymbol{B}_{t-i}) \right] \cdot \boldsymbol{Y}_{t-j}, \tag{3}$$

where the sum in (3) converges (component-wise) in the quadratic mean and also absolutely with probability one. Further, for all $t \in \mathbb{Z}$,

$$\mathrm{E}\boldsymbol{X}_t = \mathbf{0}, \ \mathrm{E}\boldsymbol{X}_t \boldsymbol{X}_t' = \mathbf{M},$$

where

$$\begin{aligned} \mathrm{vec}\,(\mathbf{M}) &= \sum_{j=0}^{\infty} [\mathrm{E}(\boldsymbol{B}_0 \otimes \boldsymbol{B}_0) + (\boldsymbol{\beta} \otimes \boldsymbol{\beta})]^j \mathrm{vec}\,(\boldsymbol{G}) \\ &= [\mathbf{I} - (\mathrm{E}(\boldsymbol{B}_0 \otimes \boldsymbol{B}_0) + (\boldsymbol{\beta} \otimes \boldsymbol{\beta}))]^{-1} \mathrm{vec}\,(\boldsymbol{G}). \end{aligned}$$

Proof. The assertion summarizes results that were proved in [11], Chapter 2 (see Theorem 2, Corollaries 2.2.1 and 2.2.2, and Theorem 2.7 there) with errors $\boldsymbol{Y}_t$ being *iid* random vectors. However, it can be shown that all the crucial steps in the proofs remain valid under the assumptions A1 to A4. See also [16] for some details. □

Remark 1.2. A strictly stationary solution of (1) can be obtained under weaker or modified assumptions then those considered in Theorem 1.1. Under the assumptions that $\{(\boldsymbol{B}_t, \boldsymbol{Y}_t), t \in \mathbb{Z}\}$ is strictly stationary and ergodic, that both $\mathrm{E}\log^+(||\boldsymbol{\beta}+\boldsymbol{B}_0||)$ and $\mathrm{E}\log^+(||\boldsymbol{Y}_0||_m)$ are finite and $\mathrm{E}\log(||\boldsymbol{\beta}+\boldsymbol{B}_0||) < 0$, where $x^+ = \max(x, 0)$, $||\cdot||_m$ is any norm in $\mathbb{R}^m$, and $||\cdot||$ denotes an operator norm defined for an $m \times m$ matrix $\boldsymbol{A}$ by

$$||\boldsymbol{A}|| = \sup\{||\boldsymbol{A}\boldsymbol{x}||_m / ||\boldsymbol{x}||_m, \boldsymbol{x} \in \mathbb{R}^m, \boldsymbol{x} \neq \mathbf{0}\},$$

we can prove that $\boldsymbol{X}_t$ in (3) converges absolutely almost surely and $\{\boldsymbol{X}_t, t \in \mathbb{Z}\}$ is the unique strictly stationary solution to (1), see [5] and [6]. In [6], necessary and sufficient conditions for the existence of a non-anticipative, strictly stationary solution of (1) were stated provided that $(\boldsymbol{B}_t, \boldsymbol{Y}_t)$ are *iid* pairs. Considering a univariate RCA(1) process and the same assumptions as in [6], the authors of [1] established minimal conditions to obtain finite moments of $\{\boldsymbol{X}_t\}$ of order $\nu \geq 1$. For $\nu = 2$, their conditions coincide with the moment assumptions in A1 − A3. In a multivariate case, sufficient conditions for the existence of finite moments of $\{\boldsymbol{X}_t\}$ of any even order ν were established in [8]. Conditions for the existence of finite moments of order $\nu \geq 1$ can be established analogously as in [1], but will not be considered here, see also Remark 2.2 below.

Usually, $\boldsymbol{\beta}, \boldsymbol{\Sigma}$, and $\boldsymbol{G}$ are unknown parameters of the model. Estimators of parameters in univariate RCA models have been considered by many authors. The least-squares estimators, their weighted versions, maximum and quasi-maximum likelihood estimators were studied, e. g., in [1, 2, 3, 9, 11] under various model assumptions. Adaptive estimators of β in a univariate RCA(1) were studied in [10]. Schick in [12] proposed a class of modified least-squares estimators in a univariate RCA(1) model, indexed by a family of bounded measurable functions. The best estimator in that class minimizing the asymptotic variance is asymptotically equivalent to the conditionally weighted least-squares estimator, which coincides with the quasi-maximum likelihood estimator. Statistical properties of the Schick-type estimators in univariate RCA models were further developed and extended in [14] and [16]. These estimators are computationally simpler than the conditionally weighted least-squares estimators, do not require any prior knowledge of additional parameters, and, as numerical studies performed in [16] show, they behave well under more general conditions and better than the least-squares estimators.

Estimation in multivariate RCA models is more complicated due to increasing number of unknown parameters. Asymptotic properties of the least-squares estimators in multivariate RCA models with *iid* errors were studied in [11]. Maximum likelihood estimating procedures were only briefly mentioned there without further details.

In this paper, we study properties of the Schick-type estimators of parameters in a multivariate RCA(1) model that satisfies assumptions A1 to A4 and generalize and extend results published recently in [15].

The paper is further organized as follows. In Section 2 we introduce Schick-type estimators and prove their strong consistency and asymptotic normality. In Section 3 we will continue with consistent variance matrices estimators of the parameter $\boldsymbol{\beta}$ and find the lower bound for the asymptotic variance matrices. A small simulation study is included in Section 4. Some necessary results from matrix theory and some auxiliary assertions are given in the Appendix.

2. MODIFIED LEAST-SQUARES ESTIMATOR

Let $\boldsymbol{X}_0, \ldots, \boldsymbol{X}_n$ be observations of process (1) that satisfy assumptions A1 to A4. Then we have

$$\mathrm{E}[\boldsymbol{X}_t|\mathcal{F}_{t-1}] = \boldsymbol{\beta}\boldsymbol{X}_{t-1} = (\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\mathrm{vec}\,(\boldsymbol{B}_t),$$

$$\begin{aligned}\mathrm{var}\,[\boldsymbol{X}_t|\mathcal{F}_{t-1}] &= \mathrm{E}\Big[\,(\boldsymbol{X}_t - \mathrm{E}[\boldsymbol{X}_t|\mathcal{F}_{t-1}])\,(\boldsymbol{X}_t - \mathrm{E}[\boldsymbol{X}_t|\mathcal{F}_{t-1}])'\,|\mathcal{F}_{t-1}\Big] = \mathrm{E}[\boldsymbol{u}_t\boldsymbol{u}'_t|\mathcal{F}_{t-1}] \\ &= \mathrm{E}\Big[\big((\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\mathrm{vec}\,(\boldsymbol{B}_t) + \boldsymbol{Y}_t\big)\,\big((\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\mathrm{vec}\,(\boldsymbol{B}_t) + \boldsymbol{Y}_t\big)'\,|\mathcal{F}_{t-1}\Big] \\ &= \mathrm{E}\big[(\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\mathrm{vec}\,(\boldsymbol{B}_t)\mathrm{vec}\,'(\boldsymbol{B}_t)(\boldsymbol{X}_{t-1} \otimes \boldsymbol{I})|\mathcal{F}_{t-1}\big] + \mathrm{E}\big[\boldsymbol{Y}_t\boldsymbol{Y}'_t|\mathcal{F}_{t-1}\big] \\ &= (\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I}) \cdot \boldsymbol{\Sigma} \cdot (\boldsymbol{X}_{t-1} \otimes \boldsymbol{I}) + \boldsymbol{G}. \end{aligned} \tag{4}$$

Using operators vec and vech and their properties as given in Lemmas A.3 and A.4, we further get after some computations

$$\mathrm{vech}\,\left(\mathrm{var}\,[\boldsymbol{X}_t|\mathcal{F}_{t-1}]\right) = \mathrm{vech}\,\left(\mathrm{E}[\boldsymbol{u}_t\boldsymbol{u}'_t|\mathcal{F}_{t-1}]\right) = \boldsymbol{A}'_{t-1} \cdot \mathrm{vech}\,(\boldsymbol{\Sigma}) + \mathrm{vech}\,(\boldsymbol{G}) \tag{5}$$

where $\boldsymbol{A}'_{t-1} = \boldsymbol{H}_m(\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I}) \otimes (\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\boldsymbol{K}'_{m^2}$ and $\boldsymbol{H}_m$ and $\boldsymbol{K}_{m^2}$ are the elimination and duplication matrices from Lemma A.4. The least-squares estimator of the parameter vec $(\boldsymbol{\beta})$ (and $\boldsymbol{\beta}$, respectively) can be obtained by minimizing

$$\sum_{t=1}^{n} \boldsymbol{u}'_t\boldsymbol{u}_t = \sum_{t=1}^{n} [\boldsymbol{X}_t - (\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\mathrm{vec}\,(\boldsymbol{\beta})]'[\boldsymbol{X}_t - (\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\mathrm{vec}\,(\boldsymbol{\beta})]$$

with respect to vec $(\boldsymbol{\beta})$, which leads to the normal equation

$$\sum_{t=1}^{n} (\boldsymbol{X}_{t-1} \otimes \boldsymbol{I})[\boldsymbol{X}_t - (\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\mathrm{vec}\,(\boldsymbol{\beta})] = \boldsymbol{0}$$

and to the estimator

$$\mathrm{vec}\,(\widehat{\boldsymbol{\beta}}_{LS}) = \left(\sum_{t=1}^{n} [\boldsymbol{X}_{t-1}\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I}]\right)^{-1} \sum_{t=1}^{n} (\boldsymbol{X}_{t-1} \otimes \boldsymbol{I})\boldsymbol{X}_t, \tag{6}$$

respectively,

$$\widehat{\boldsymbol{\beta}}_{LS} = \left(\sum_{t=1}^{n} \boldsymbol{X}_t\boldsymbol{X}'_{t-1}\right) \cdot \left(\sum_{t=1}^{n} \boldsymbol{X}_{t-1}\boldsymbol{X}'_{t-1}\right)^{-1}. \tag{7}$$

In [11], Section 7.2, the strong consistency and asymptotic normality of vec $(\widehat{\boldsymbol{\beta}}_{LS})$ are proved under assumptions that the 4^{th} moments of the components of the vector $\boldsymbol{X}_t$ are finite.

The other parameters of the model, variance matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{G}$, can be estimated from the regression equation (5), when we use estimated residuals $\widehat{\boldsymbol{u}}_t =$

$\boldsymbol{X}_t - \widehat{\boldsymbol{\beta}}_{LS}\boldsymbol{X}_{t-1}$. The least-squares estimators are then as follows:

$$\mathrm{vech}\,(\widehat{\boldsymbol{\Sigma}}_{LS}) = \left(\sum_{t=1}^{n}\left(\boldsymbol{A}_{t-1}-\overline{\boldsymbol{A}}\right)\left(\boldsymbol{A}_{t-1}-\overline{\boldsymbol{A}}\right)'\right)^{-1}\left(\sum_{t=1}^{n}\left(\boldsymbol{A}_{t-1}-\overline{\boldsymbol{A}}\right)\cdot \mathrm{vech}\,\left(\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_t'\right)\right),$$

$$\mathrm{vech}\,(\widehat{\boldsymbol{G}}_{LS}) = \frac{1}{n}\sum_{t=1}^{n}\mathrm{vech}\,\left(\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_t'\right)-\overline{\boldsymbol{A}}'\cdot \mathrm{vech}\,(\widehat{\boldsymbol{\Sigma}}_{LS}), \tag{8}$$

where $\overline{\boldsymbol{A}} = \frac{1}{n}\sum_{t=1}^{n}\boldsymbol{A}_{t-1}$. Under higher moment conditions on the process $\{\boldsymbol{X}_t\}$ such estimators are strongly consistent (one requires finite 4^{th} moments of the components of the vector $\boldsymbol{X}_t$) and asymptotically normal (finite 8^{th} moments), see [11], Theorem 7.2.

In [12] Schick considered estimators of parameters in a univariate RCA(1) model with *iid* errors as a solution of modified normal equations and studied asymptotic properties of these estimators without additional moment assumptions. We generalize the Schick method and propose an extension of the least-squares estimators of $\boldsymbol{\beta}$ into a class of estimators that solve the equation

$$\sum_{t=1}^{n}(\phi(\boldsymbol{X}_{t-1})\otimes\boldsymbol{I})[\boldsymbol{X}_t-(\boldsymbol{X}_{t-1}'\otimes\boldsymbol{I})\mathrm{vec}\,(\boldsymbol{\beta})]=\boldsymbol{0} \tag{9}$$

where ϕ is a measurable function $\phi:\mathbb{R}^m\to\mathbb{R}^m$. Thus, we define estimator

$$\mathrm{vec}\,(\widehat{\boldsymbol{\beta}}_\phi) = \left(\sum_{t=1}^{n}\phi(\boldsymbol{X}_{t-1})\boldsymbol{X}_{t-1}'\otimes\boldsymbol{I}\right)^{-1}\cdot\sum_{t=1}^{n}(\phi(\boldsymbol{X}_{t-1})\otimes\boldsymbol{I})\boldsymbol{X}_t, \tag{10}$$

respectively,

$$\widehat{\boldsymbol{\beta}}_\phi = \left(\sum_{t=1}^{n}\boldsymbol{X}_t\phi(\boldsymbol{X}_{t-1})'\right)\cdot\left(\sum_{t=1}^{n}\boldsymbol{X}_{t-1}\phi(\boldsymbol{X}_{t-1})'\right)^{-1}. \tag{11}$$

The following theorem reveals basic properties of this estimator.

Theorem 2.1. Consider a multivariate RCA(1) model as in (1) that satisfies assumptions A1 to A4. Let $\phi:\mathbb{R}^m\to\mathbb{R}^m$ be a measurable function. Denote $\boldsymbol{P}=\phi(\boldsymbol{X}_0)\boldsymbol{X}_0'\otimes\boldsymbol{I}$ and $\boldsymbol{Q}=\phi(\boldsymbol{X}_0)\otimes\boldsymbol{I}$, and assume that the matrix $\mathrm{E}(\phi(\boldsymbol{X}_0)\boldsymbol{X}_0')$ is finite and invertible, and $\mathrm{E}(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}'+\boldsymbol{Q}\boldsymbol{G}\boldsymbol{Q}')$ is finite. Then $\mathrm{vec}\,(\widehat{\boldsymbol{\beta}}_\phi)$ defined by (10) is a strongly consistent and asymptotically normal estimator of the parameter $\mathrm{vec}\,(\boldsymbol{\beta})$. The asymptotic variance matrix of $\sqrt{n}\cdot\mathrm{vec}\,\left(\widehat{\boldsymbol{\beta}}_\phi-\boldsymbol{\beta}\right)$ is given by

$$\boldsymbol{V}(\phi) = (\mathrm{E}\boldsymbol{P})^{-1}\cdot\mathrm{E}\left(\boldsymbol{Q}\boldsymbol{G}\boldsymbol{Q}'+\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}'\right)\cdot(\mathrm{E}\boldsymbol{P}')^{-1}. \tag{12}$$

Remark 2.2. The choice $\phi(\boldsymbol{x})=\boldsymbol{x}$ leads to the least-squares estimator of $\boldsymbol{\beta}$ and fulfills the finite matrices assumptions provided that $\{\boldsymbol{X}_t\}$ has finite fourth moments. If ϕ is bounded, this assumption reduces to the finiteness of the second moments of $\{\boldsymbol{X}_t\}$.

Proof. According to definition (10) and using Lemma A.3, we have

$$\operatorname{vec}(\widehat{\boldsymbol{\beta}}_\phi - \boldsymbol{\beta}) = \left[\left(\frac{1}{n} \sum_{t=1}^{n} \phi(\boldsymbol{X}_{t-1}) \boldsymbol{X}'_{t-1} \right)^{-1} \otimes \boldsymbol{I} \right] \left[\frac{1}{n} \sum_{t=1}^{n} \Big(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I} \Big) \boldsymbol{u}_t \right]. \quad (13)$$

Strict stationarity and ergodicity of $\boldsymbol{X}_t$ guarantee strict stationarity and ergodicity of both sequences $\{\phi(\boldsymbol{X}_{t-1})\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I}, t \in \mathbb{Z}\}$ and $\{(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I})\boldsymbol{u}_t, t \in \mathbb{Z}\}$. Moreover, the components of the latter sequence form a martingale difference sequence with zero mean value, which can be seen by choosing an arbitrary $\boldsymbol{\alpha} \in \mathbb{R}^{m^2}$ and noticing that

$$\mathrm{E}[\boldsymbol{\alpha}'(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I})\boldsymbol{u}_t | \mathcal{F}_{t-1}] = \boldsymbol{\alpha}'(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I})\mathrm{E}[\boldsymbol{u}_t | \mathcal{F}_{t-1}] = 0.$$

The ergodic theorem (see, e. g., [7], Theorem 13.12) tells us that, almost surely, the first term on the r.h.s. in (13) converges to $(\mathrm{E}(\phi(\boldsymbol{X}_0)\boldsymbol{X}'_0))^{-1} \otimes \boldsymbol{I} = (\mathrm{E}\boldsymbol{P})^{-1}$ and the second term converges to zero, which implies that $\operatorname{vec}(\widehat{\boldsymbol{\beta}}_\phi - \boldsymbol{\beta}) \to \mathbf{0}$ almost surely as $n \to +\infty$.

Further notice that

$$\begin{aligned}
\operatorname{var}\Big(\boldsymbol{\alpha}'(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I})\boldsymbol{u}_t\Big) &= \mathrm{E}\Big(\mathrm{E}\Big[(\boldsymbol{\alpha}'(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I})\boldsymbol{u}_t)^2 \,|\mathcal{F}_{t-1}\Big]\Big) \\
&= \mathrm{E}\Big(\mathrm{E}[\boldsymbol{\alpha}'(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I})\boldsymbol{u}_t\boldsymbol{u}'_t(\phi(\boldsymbol{X}_{t-1})' \otimes \boldsymbol{I})\boldsymbol{\alpha}|\mathcal{F}_{t-1}]\Big) \\
&= \mathrm{E}\Big(\boldsymbol{\alpha}'(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}) \cdot \mathrm{E}[\boldsymbol{u}_t\boldsymbol{u}'_t|\mathcal{F}_{t-1}] \cdot (\phi(\boldsymbol{X}_{t-1})' \otimes \boldsymbol{I})\boldsymbol{\alpha}\Big) \\
&= \mathrm{E}\Big(\boldsymbol{\alpha}'(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}) \cdot \big[(\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\boldsymbol{\Sigma}(\boldsymbol{X}_{t-1} \otimes \boldsymbol{I}) + \boldsymbol{G}\big] \cdot (\phi(\boldsymbol{X}_{t-1})' \otimes \boldsymbol{I})\boldsymbol{\alpha}\Big) \\
&= \boldsymbol{\alpha}'\mathrm{E}\Big((\phi(\boldsymbol{X}_{t-1})\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I})\boldsymbol{\Sigma}(\boldsymbol{X}_{t-1}\phi(\boldsymbol{X}_{t-1})' \otimes \boldsymbol{I})\Big)\boldsymbol{\alpha} \\
&\quad + \boldsymbol{\alpha}'\mathrm{E}\Big((\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I})\boldsymbol{G}(\phi(\boldsymbol{X}_{t-1})' \otimes \boldsymbol{I})\Big)\boldsymbol{\alpha} \\
&= \boldsymbol{\alpha}'\mathrm{E}(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}' + \boldsymbol{Q}\boldsymbol{G}\boldsymbol{Q}')\boldsymbol{\alpha}.
\end{aligned} \quad (14)$$

We know that, due to Lindeberg–Lévy theorem for martingales, see [4], for any $\boldsymbol{\alpha} \in \mathbb{R}^{m^2}$, $\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \boldsymbol{\alpha}' \Big(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I} \Big) \boldsymbol{u}_t$ has an asymptotic normal distribution with zero mean and variance (14). Since

$$\sqrt{n} \cdot \operatorname{vec}(\widehat{\boldsymbol{\beta}}_\phi - \boldsymbol{\beta}) = \left[\left(\frac{1}{n} \sum_{t=1}^{n} \phi(\boldsymbol{X}_{t-1}) \boldsymbol{X}'_{t-1} \right)^{-1} \otimes \boldsymbol{I} \right] \cdot \left[\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Big(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I} \Big) \boldsymbol{u}_t \right],$$

the rest of the proof easily follows by using the previous considerations and the Cramér–Wold device. □

To obtain estimators of the remaining parameters $\boldsymbol{\Sigma}$ and $\boldsymbol{G}$, we can use the regression (5) again, now with residuals $\widehat{\boldsymbol{u}}_t = \boldsymbol{X}_t - \widehat{\boldsymbol{\beta}}_\phi \boldsymbol{X}_{t-1}$, and define modified least-squares estimators

$$\operatorname{vech}(\widehat{\boldsymbol{\Sigma}}_h) = \left(\sum_{t=1}^{n}\left(h(\boldsymbol{A}_{t-1})-\overline{\boldsymbol{h}}\right)\left(\boldsymbol{A}_{t-1}-\overline{\boldsymbol{A}}\right)'\right)^{-1}\left(\sum_{t=1}^{n}\left(h(\boldsymbol{A}_{t-1})-\overline{\boldsymbol{h}}\right)\cdot \operatorname{vech}\left(\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_t'\right)\right),$$

$$\operatorname{vech}(\widehat{\boldsymbol{G}}_h) = \frac{1}{n}\sum_{t=1}^{n}\operatorname{vech}\left(\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_t'\right) - \overline{\boldsymbol{A}}'\cdot\operatorname{vech}(\widehat{\boldsymbol{\Sigma}}_h), \tag{15}$$

where

$$h(\boldsymbol{A}_{t-1}) = [\boldsymbol{H}_m(h(\boldsymbol{X}_{t-1})'\otimes\boldsymbol{I})\otimes(h(\boldsymbol{X}_{t-1})'\otimes\boldsymbol{I})\boldsymbol{K}'_{m^2}]',$$

$\overline{\boldsymbol{h}} = \frac{1}{n}\sum_{t=1}^{n} h(\boldsymbol{A}_{t-1})$, and h is a measurable function $h:\mathbb{R}^m\to\mathbb{R}^m$. The choice $h(\boldsymbol{x}) = \boldsymbol{x}$ leads to the least-squares estimators (8).

Theorem 2.3. Consider a multivariate RCA(1) model as in (1) that satisfies assumptions A1 to A4. Let ϕ, h be bounded measurable functions from $\mathbb{R}^m$ to $\mathbb{R}^m$ and let the matrix $\mathrm{E}[(h(\boldsymbol{A}_0)-\mathrm{E}h(A_0))(\boldsymbol{A}_0-\mathrm{E}\boldsymbol{A}_0)']$ be invertible. Then the estimators $\operatorname{vech}(\widehat{\boldsymbol{\Sigma}}_h)$ and $\operatorname{vech}(\widehat{\boldsymbol{G}}_h)$ defined in (15) are strongly consistent estimators of $\operatorname{vech}(\boldsymbol{\Sigma})$ and $\operatorname{vech}(\boldsymbol{G})$, respectively.

Proof. We give only the main steps of the proof. Obviously, both $\{h(\boldsymbol{A}_t)\}$ and $\{\boldsymbol{A}_t\}$ are strictly stationary and ergodic and thus, as $n\to\infty$,

$$\frac{1}{n}\sum_{t=1}^{n}(h(\boldsymbol{A}_{t-1})-\overline{\boldsymbol{h}})(\boldsymbol{A}_{t-1}-\overline{\boldsymbol{A}})' \to \mathrm{E}h(\boldsymbol{A}_0)\boldsymbol{A}_0' - \mathrm{E}h(\boldsymbol{A}_0)(\mathrm{E}\boldsymbol{A}_0)'$$

almost surely. Further,

$$\begin{aligned}\operatorname{vech}(\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_t') &= \operatorname{vech}(\boldsymbol{u}_t\boldsymbol{u}_t') + \operatorname{vech}[(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_\phi)\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-1}'(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_\phi)'] \\ &\quad + \operatorname{vech}((\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_\phi)\boldsymbol{X}_{t-1}\boldsymbol{u}_t') + \operatorname{vech}[(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_\phi)\boldsymbol{X}_{t-1}\boldsymbol{u}_t')]'\end{aligned}$$

and

$$\operatorname{vech}[(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_\phi)\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-1}'(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_\phi)'] = \boldsymbol{A}_{t-1}'\operatorname{vech}[\operatorname{vec}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_\phi)(\operatorname{vec}(\boldsymbol{\beta}-\widehat{\boldsymbol{\beta}}_\phi))'].$$

Then, from the strict stationarity and ergodicity, strong consistency of $\widehat{\boldsymbol{\beta}}_\phi$, and the martingale difference properties of $\{\boldsymbol{X}_{t-1}\boldsymbol{u}_t'\}$ we conclude that

$$\frac{1}{n}\sum_{t=1}^{n}\operatorname{vech}(\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_t') \to \mathrm{E}\operatorname{vech}(\boldsymbol{u}_0\boldsymbol{u}_0') = \mathrm{E}\boldsymbol{A}_0\operatorname{vech}(\boldsymbol{\Sigma}) + \operatorname{vech}(\boldsymbol{G})$$

almost surely. In the same way, we get

$$\frac{1}{n}\sum_{t=1}^{n} h(\boldsymbol{A}_{t-1})\operatorname{vech}(\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_t') \to \mathrm{E}h(\boldsymbol{A}_0)\boldsymbol{A}_0'\operatorname{vech}(\boldsymbol{\Sigma}) + \mathrm{E}h(\boldsymbol{A}_0)\operatorname{vech}(\boldsymbol{G})$$

almost surely. Combining all these results we complete the proof. □

3. ASYMPTOTIC VARIANCE MATRIX

In this section we deal with the asymptotic variance matrix of the estimator vec $(\widehat{\boldsymbol{\beta}}_\phi)$. First, ve suggest a consistent estimator of $\boldsymbol{V}(\phi)$.

Theorem 3.1. Consider a multivariate RCA(1) model as in (1) that satisfies assumptions A1 to A4. Let $\phi : \mathbb{R}^m \to \mathbb{R}^m$ be a measurable function such that the assumptions of Theorem 2.1 are satisfied. Let $\widehat{\boldsymbol{G}}_n$ and $\widehat{\boldsymbol{\Sigma}}_n$ be strongly consistent estimators of $\boldsymbol{G}$ and $\boldsymbol{\Sigma}$, respectively. Denote $\boldsymbol{P}_t = \phi(\boldsymbol{X}_{t-1})\boldsymbol{X}'_{t-1} \otimes \boldsymbol{I}$ and $\boldsymbol{Q}_t = \phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}$.
Then

$$\begin{aligned}\widehat{\boldsymbol{V}}_n(\phi) = n \left(\sum_{t=1}^n \boldsymbol{P}_t\right)^{-1} \cdot \sum_{t=1}^n \left(\boldsymbol{Q}_t \widehat{\boldsymbol{G}}_n \boldsymbol{Q}'_t\right) \cdot \left(\sum_{t=1}^n \boldsymbol{P}'_t\right)^{-1} \\ + n \left(\sum_{t=1}^n \boldsymbol{P}_t\right)^{-1} \cdot \sum_{t=1}^n \left(\boldsymbol{P}_t \widehat{\boldsymbol{\Sigma}}_n \boldsymbol{P}'_t\right) \cdot \left(\sum_{t=1}^n \boldsymbol{P}'_t\right)^{-1}\end{aligned}$$

is a strongly consistent estimator of the asymptotic variance matrix $\boldsymbol{V}(\phi)$ given by (12).

Proof. The process $\{\boldsymbol{X}_t, t \in \mathbb{Z}\}$ is strictly stationary and ergodic which implies that $\{\boldsymbol{P}_t, t \in \mathbb{Z}\}$, $\{\boldsymbol{Q}_t, t \in \mathbb{Z}\}$, $\{\boldsymbol{P}_t \boldsymbol{\Sigma} \boldsymbol{P}'_t, t \in \mathbb{Z}\}$, and $\{\boldsymbol{Q}_t \boldsymbol{G} \boldsymbol{Q}'_t, t \in \mathbb{Z}\}$ are strictly stationary and ergodic. According to the ergodic theorem, $\frac{1}{n}\sum_{t=1}^n \boldsymbol{P}_t \xrightarrow{\text{a.s.}} \mathrm{E}\boldsymbol{P}$ and $\frac{1}{n}\sum_{t=1}^n \boldsymbol{Q}_t \xrightarrow{\text{a.s.}} \mathrm{E}\boldsymbol{Q}$ as $n \to +\infty$, where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are defined in Theorem 2.1. Consistency of the estimators $\widehat{\boldsymbol{G}}_n$, $\widehat{\boldsymbol{\Sigma}}_n$, and Lemma A.5 complete the proof. □

We are interested in an optimal choice of the generating function ϕ for the estimator $\widehat{\boldsymbol{\beta}}_\phi$. Optimality of an estimator could be defined using its asymptotic variance matrix. We say that the estimator $\widehat{\boldsymbol{\beta}}(\psi)$ defined by equation (10) is optimal if its asymptotic variance matrix $\boldsymbol{V}(\psi)$ defined by (12) satisfies $\boldsymbol{V}(\phi) - \boldsymbol{V}(\psi) \geq \boldsymbol{0}$ for any estimator $\widehat{\boldsymbol{\beta}}_\phi$ with variance matrix $\boldsymbol{V}(\phi)$, i. e., the difference of the variance matrices is a positively semi-definite matrix.

We have performed an analysis of the asymptotic variance matrix for the univariate higher-order RCA models, see [16]. In that case, the optimal choice of the function ϕ leads to the estimator of $\boldsymbol{\beta}$ that is formally equivalent to the conditionally weighted least-squares estimator when the remaining parameters are known. In the multivariate first-order case, we can only establish a lower bound for the asymptotic variance matrix of the estimators defined by (10).

Theorem 3.2. Let $\{\boldsymbol{X}_t, t \in \mathbb{Z}\}$ be a multivariate RCA(1) model as in (1) that satisfies assumptions A1 to A4 such that the matrix

$$\boldsymbol{J} = \mathrm{E}\left((\boldsymbol{X}_0 \otimes \boldsymbol{I}) \cdot [\boldsymbol{w}(\boldsymbol{X}_0)]^{-1} \cdot \left(\boldsymbol{X}'_0 \otimes \boldsymbol{I}\right)\right) \tag{16}$$

is nonsingular, where for any $\boldsymbol{z} \in \mathbb{R}^m$,

$$\boldsymbol{w}(\boldsymbol{z}) = (\boldsymbol{z}' \otimes \boldsymbol{I}) \cdot \boldsymbol{\Sigma} \cdot (\boldsymbol{z} \otimes \boldsymbol{I}) + \boldsymbol{G}. \tag{17}$$

Then the matrix $\boldsymbol{J}^{-1}$ is a lower bound of the asymptotic variance matrix for all estimators $\widehat{\boldsymbol{\beta}}_\phi$ such that $\mathrm{E}\left(\phi(\boldsymbol{X}_0)\boldsymbol{X}_0'\right)$ is finite and invertible, and $\mathrm{E}(\boldsymbol{P\Sigma P}' + \boldsymbol{QGQ}')$ is finite.

P r o o f. Consider the m^2-dimensional random vectors

$$\boldsymbol{T}_1 = \sum_{t=1}^{n} \left(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}\right) \cdot (\boldsymbol{X}_t - \boldsymbol{\beta}'\boldsymbol{X}_{t-1}) = \sum_{t=1}^{n} \left(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}\right) \cdot \boldsymbol{u}_t,$$

$$\begin{aligned}\boldsymbol{T}_2 &= \sum_{t=1}^{n} \left(\boldsymbol{X}_{t-1} \otimes \boldsymbol{I}\right) \cdot [\boldsymbol{w}(\boldsymbol{X}_{t-1})]^{-1} \cdot (\boldsymbol{X}_t - \boldsymbol{\beta}'\boldsymbol{X}_{t-1}) \\ &= \sum_{t=1}^{n} \left(\boldsymbol{X}_{t-1} \otimes \boldsymbol{I}\right) \cdot [\boldsymbol{w}(\boldsymbol{X}_{t-1})]^{-1} \cdot \boldsymbol{u}_t,\end{aligned}$$

where $\boldsymbol{w}(\boldsymbol{z})$ is defined by (17).

Since both sequences $\{(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}) \cdot \boldsymbol{u}_t\}$ and $\{(\boldsymbol{X}_{t-1} \otimes \boldsymbol{I}) \cdot \boldsymbol{w}(\boldsymbol{X}_{t-1})^{-1} \cdot \boldsymbol{u}_t\}$ are martingale differences w.r.t. $\mathcal{F}_t$, it immediately follows that $\mathrm{E}\boldsymbol{T}_1 = \boldsymbol{0}$, $\mathrm{E}\boldsymbol{T}_2 = \boldsymbol{0}$ and the variance matrix of the vector $\boldsymbol{T}_1$ equals

$$\begin{aligned}\mathrm{E}\boldsymbol{T}_1\boldsymbol{T}_1' &= \mathrm{E}\left(\left(\sum_{t=1}^{n} \left(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}\right) \boldsymbol{u}_t\right) \cdot \left(\sum_{s=1}^{n} \boldsymbol{u}_s' \left(\phi(\boldsymbol{X}_{t-1})' \otimes \boldsymbol{I}\right)\right)\right) \\ &= \sum_{t=1}^{n} \mathrm{E}\left(\left(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}\right) \boldsymbol{u}_t\boldsymbol{u}_t' \left(\phi(\boldsymbol{X}_{t-1})' \otimes \boldsymbol{I}\right)\right) \\ &= \sum_{t=1}^{n} \mathrm{E}\left(\left(\phi(\boldsymbol{X}_{t-1}) \otimes \boldsymbol{I}\right) \cdot \mathrm{E}[\boldsymbol{u}_t\boldsymbol{u}_t' | \mathcal{F}_{t-1}] \cdot \left(\phi(\boldsymbol{X}_{t-1})' \otimes \boldsymbol{I}\right)\right) \\ &= n \cdot \mathrm{E}\left(\left(\phi(\boldsymbol{X}_0) \otimes \boldsymbol{I}\right) \cdot \boldsymbol{w}(\boldsymbol{X}_0) \cdot \left(\phi(\boldsymbol{X}_0)' \otimes \boldsymbol{I}\right)\right) \\ &= n \cdot \mathrm{E}\left(\left(\phi(\boldsymbol{X}_0) \otimes \boldsymbol{I}\right) \cdot \left((\boldsymbol{X}_0' \otimes \boldsymbol{I}) \cdot \boldsymbol{\Sigma} \cdot (\boldsymbol{X}_0 \otimes \boldsymbol{I}) + \boldsymbol{G}\right) \cdot \left(\phi(\boldsymbol{X}_0)' \otimes \boldsymbol{I}\right)\right) \\ &= n \cdot \mathrm{E}\left(\left(\phi(\boldsymbol{X}_0)\boldsymbol{X}_0' \otimes \boldsymbol{I}\right) \cdot \boldsymbol{\Sigma} \cdot \left(\boldsymbol{X}_0\phi(\boldsymbol{X}_0)' \otimes \boldsymbol{I}\right)\right) \\ &+ n \cdot \mathrm{E}\left(\left(\phi(\boldsymbol{X}_0) \otimes \boldsymbol{I}\right) \cdot \boldsymbol{G} \cdot \left(\phi(\boldsymbol{X}_0)' \otimes \boldsymbol{I}\right)\right) \\ &= n \cdot \mathrm{E}\left(\boldsymbol{P\Sigma P}' + \boldsymbol{QGQ}'\right), \end{aligned} \tag{18}$$

as follows from Lemma A.2 and the notation defined in Theorem 2.1. As a direct analogue we can infer that

$$\mathrm{E}\boldsymbol{T}_2\boldsymbol{T}_2' = n \cdot \mathrm{E}\left(\left(\boldsymbol{X}_0 \otimes \boldsymbol{I}\right) \cdot [\boldsymbol{w}(\boldsymbol{X}_0)]^{-1} \cdot \left(\boldsymbol{X}_0' \otimes \boldsymbol{I}\right)\right). \tag{19}$$

The cross-covariance matrix of $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ can be computed as follows:

$$\begin{aligned}
\mathrm{E}\boldsymbol{T}_1\boldsymbol{T}_2' &= \mathrm{E}\left(\left(\sum_{t=1}^{n}\Big(\phi(\boldsymbol{X}_{t-1})\otimes\boldsymbol{I}\Big)\cdot\boldsymbol{u}_t\right)\cdot\left(\sum_{s=1}^{n}\boldsymbol{u}_s'\cdot[\boldsymbol{w}(\boldsymbol{X}_{s-1})]^{-1}\cdot\Big(\boldsymbol{X}_{s-1}'\otimes\boldsymbol{I}\Big)\right)\right)\\
&= \sum_{t=1}^{n}\mathrm{E}\left(\Big(\phi(\boldsymbol{X}_{t-1})\otimes\boldsymbol{I}\Big)\cdot\mathrm{E}[\boldsymbol{u}_t\boldsymbol{u}_t'|\mathcal{F}_{t-1}]\cdot[\boldsymbol{w}(\boldsymbol{X}_{t-1})]^{-1}\cdot\Big(\boldsymbol{X}_{t-1}'\otimes\boldsymbol{I}\Big)\right)\\
&= \sum_{t=1}^{n}\mathrm{E}\left(\Big(\phi(\boldsymbol{X}_{t-1})\otimes\boldsymbol{I}\Big)\cdot\boldsymbol{w}(\boldsymbol{X}_{t-1})\cdot[\boldsymbol{w}(\boldsymbol{X}_{t-1})]^{-1}\cdot\Big(\boldsymbol{X}_{t-1}'\otimes\boldsymbol{I}\Big)\right)\\
&= n\cdot\mathrm{E}\left(\Big(\phi(\boldsymbol{X}_0)\otimes\boldsymbol{I}\Big)\cdot\Big(\boldsymbol{X}_0'\otimes\boldsymbol{I}\Big)\right) = n\cdot\mathrm{E}\Big(\phi(\boldsymbol{X}_0)\boldsymbol{X}_0'\otimes\boldsymbol{I}\Big) = n\cdot\mathrm{E}\boldsymbol{P}.
\end{aligned} \tag{20}$$

The variance matrix of the $2m^2$-dimensional random vector $(\boldsymbol{T}_1', \boldsymbol{T}_2')'$ is equal to

$$\begin{pmatrix} \mathrm{E}\boldsymbol{T}_1\boldsymbol{T}_1', & \mathrm{E}\boldsymbol{T}_1\boldsymbol{T}_2' \\ \mathrm{E}\boldsymbol{T}_2\boldsymbol{T}_1', & \mathrm{E}\boldsymbol{T}_2\boldsymbol{T}_2' \end{pmatrix}$$

where the block elements were computed in (18) – (20). Lemma A.6 tells us that, if the block elements $\mathrm{E}\boldsymbol{T}_1\boldsymbol{T}_2'$ and $\mathrm{E}\boldsymbol{T}_2\boldsymbol{T}_2'$ are invertible matrices, then

$$\begin{aligned}
&\Big(\mathrm{E}\boldsymbol{T}_1\boldsymbol{T}_2'\Big)^{-1}\cdot\big(\mathrm{E}\boldsymbol{T}_1\boldsymbol{T}_1'\big)\cdot\big(\mathrm{E}\boldsymbol{T}_2\boldsymbol{T}_1'\big)^{-1} - \big(\mathrm{E}\boldsymbol{T}_2\boldsymbol{T}_2'\big)^{-1} \geq \boldsymbol{0}\\
&\Longleftrightarrow (\mathrm{E}\boldsymbol{P})^{-1}\cdot\mathrm{E}\Big(\boldsymbol{P\Sigma P}' + \boldsymbol{QGQ}'\Big)\cdot(\mathrm{E}\boldsymbol{P}')^{-1} - \big(\mathrm{E}\boldsymbol{T}_2\boldsymbol{T}_2'\big)^{-1} \geq \boldsymbol{0}\\
&\Longleftrightarrow \boldsymbol{V}(\phi) - \big(\mathrm{E}\left((\boldsymbol{X}_0\otimes\boldsymbol{I})\cdot[\boldsymbol{w}(\boldsymbol{X}_0)]^{-1}\cdot\left(\boldsymbol{X}_0'\otimes\boldsymbol{I}\right)\right)\big)^{-1} \geq \boldsymbol{0},
\end{aligned}$$

where $\boldsymbol{V}(\phi)$ is the asymptotic variance matrix of the general estimator $\widehat{\boldsymbol{\beta}}_\phi$. Thus

$$\big(\mathrm{E}\left((\boldsymbol{X}_0\otimes\boldsymbol{I})\cdot[\boldsymbol{w}(\boldsymbol{X}_0)]^{-1}\cdot\left(\boldsymbol{X}_0'\otimes\boldsymbol{I}\right)\right)\big)^{-1} = \boldsymbol{J}^{-1}$$

is a lower bound of the asymptotic variance matrix for all functional $\widehat{\boldsymbol{\beta}}_\phi$ such that $\mathrm{E}\Big(\phi(\boldsymbol{X}_0)\boldsymbol{X}_0'\otimes\boldsymbol{I}\Big)$ and $\boldsymbol{J}$ are invertible. □

There are a few special cases when we can compute the optimal estimator explicitly. If $\boldsymbol{\Sigma} = \boldsymbol{0}$ for instance, which corresponds to the classical AR model, and $\boldsymbol{G} = \sigma^2\boldsymbol{I}$ for some $\sigma^2 > 0$, we have $\boldsymbol{w}(\boldsymbol{z}) = \sigma^2\boldsymbol{I}$. Then the lower bound equals

$$\boldsymbol{J}^{-1} = \big(\mathrm{E}\left((\boldsymbol{X}_0\otimes\boldsymbol{I})\cdot[\boldsymbol{w}(\boldsymbol{X}_0)]^{-1}\cdot\left(\boldsymbol{X}_0'\otimes\boldsymbol{I}\right)\right)\big)^{-1} = \sigma^2\Big(\mathrm{E}\left(\boldsymbol{X}_0\boldsymbol{X}_0'\otimes\boldsymbol{I}\right)\Big)^{-1},$$

and the asymptotic variance matrix of the estimator $\widehat{\boldsymbol{\beta}}_\phi$ with $\phi(\boldsymbol{z}) = \boldsymbol{z}$ attains this lower bound, whereas such estimator corresponds to the least-squares estimator. In a general RCA(1) model, however, the lower bound for the asymptotic variance of the least-squares estimator is not attained.

If $\boldsymbol{\Sigma} = \widetilde{\boldsymbol{\Sigma}} \otimes \boldsymbol{I}$, where $\widetilde{\boldsymbol{\Sigma}}$ is an $m \times m$ positive definite matrix, and $\boldsymbol{G} = \sigma^2 \boldsymbol{I},\ \sigma^2 > 0$, then $w(\boldsymbol{z}) = (\boldsymbol{z}'\widetilde{\boldsymbol{\Sigma}}\boldsymbol{z} + \sigma^2)\boldsymbol{I}$,

$$\boldsymbol{J}^{-1} = \left[\mathrm{E}\left(\frac{\boldsymbol{X}_0 \boldsymbol{X}_0'}{\boldsymbol{X}_0' \widetilde{\boldsymbol{\Sigma}} \boldsymbol{X}_0 + \sigma^2} \right) \right]^{-1} \otimes \boldsymbol{I}$$

and this lower bound is attained with the function $\phi(\boldsymbol{z}) = \boldsymbol{z}(\boldsymbol{z}'\widetilde{\boldsymbol{\Sigma}}\boldsymbol{z} + \sigma^2)^{-1} = w(\boldsymbol{z})^{-1}\boldsymbol{z}$ which corresponds to the conditionally weighted estimator. Like in the univariate case, such choice, of course, depends heavily on the matrix $\widetilde{\boldsymbol{\Sigma}}$ and the parameter σ^2 that are usually unknown.

If we assume that $\boldsymbol{\Sigma} = \boldsymbol{I}, \boldsymbol{G} = \boldsymbol{I}$, then the function $\phi(\boldsymbol{z}) = \boldsymbol{z}(1 + \boldsymbol{z}'\boldsymbol{z})^{-1}$ leads to the optimal estimator with the asymptotic variance

$$\boldsymbol{V}(\phi) = \boldsymbol{J}^{-1} = \left(\mathrm{E}((\boldsymbol{X}_0 \boldsymbol{X}_0' \otimes \boldsymbol{I})(1 + \boldsymbol{X}_0' \boldsymbol{X}_0)^{-1}) \right)^{-1}.$$

In a simulation study we show that this choice of ϕ provides a reasonable estimator even with other values of variance matrices $\boldsymbol{G}$ and $\boldsymbol{\Sigma}$. The advantage of this estimator is that it does not depend on nuisance parameters and seems to be more stable then the least-squares estimator.

Remark 3.3. Notice that any one-dimensional RCA process of order p,

$$X_t = \sum_{i=1}^{p} (\beta_i + B_{ti}) X_{t-i} + Y_t, \tag{21}$$

where β_i and $B_{ti}, i = 1, \ldots, p$, are constant and random components of the vector random coefficient, can be written as a multivariate RCA(1) model of form (1) with $\boldsymbol{X}_t = (X_t, X_{t-1}, \ldots, X_{t-p+1})', \boldsymbol{Y}_t = (Y_t, 0, \ldots, 0)'$, and

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 & \beta_2 & \ldots & \beta_{p-1} & \beta_p \\ 1 & 0 & \ldots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}, \quad \boldsymbol{B}_t = \begin{pmatrix} B_{t1} & \ldots & B_{tp} \\ 0 & \ldots & 0 \\ \vdots & & \\ 0 & \ldots & 0 \end{pmatrix}. \tag{22}$$

Then the role of $\boldsymbol{\Sigma}$ and $\boldsymbol{G}$ is played by $\mathrm{E}(B_t B_t')$ and $\mathrm{E}[Y_t^2 | \mathcal{F}_{t-1}]$, respectively, where $B_t = (B_{t1}, \ldots, B_{tp})'$. The least-squares estimator of the parameter $\beta = (\beta_1, \ldots, \beta_p)'$ is

$$\widehat{\beta}_{LS} = \left(\sum_{t=1}^{n} \boldsymbol{X}_{t-1} \boldsymbol{X}_{t-1}' \right)^{-1} \sum_{t=1}^{n} \boldsymbol{X}_{t-1} X_t$$

and the Schick-type estimator is of the form

$$\widehat{\beta}_{\phi} = \left(\sum_{t=1}^{n} \phi(\boldsymbol{X}_{t-1}) \boldsymbol{X}_{t-1}' \right)^{-1} \sum_{t=1}^{n} \phi(\boldsymbol{X}_{t-1}) X_t$$

where ϕ denotes a measurable function $\phi : \mathbb{R}^p \to \mathbb{R}^p$. The asymptotic behavior of $\widehat{\beta}_{\phi}$ was studied in detail in [16].

4. SIMULATION STUDY

In this short study we compare the least-squares estimator to a particular choice of the Schick-type estimator. We simulated 100 observations from a 2-dimensional RCA(1) model given by

$$\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \left(\begin{pmatrix} 0.2 & 0.1 \\ 0.3 & 0.4 \end{pmatrix} + \begin{pmatrix} B_t^{11} & B_t^{12} \\ B_t^{21} & B_t^{22} \end{pmatrix} \right) \cdot \begin{pmatrix} X_{t-1}^1 \\ X_{t-1}^2 \end{pmatrix} + \begin{pmatrix} Y_t^1 \\ Y_t^2 \end{pmatrix} \tag{23}$$

where both random coefficients $\boldsymbol{B}_t$ and error process $\boldsymbol{Y}_t$ are mutually independent and identically normally distributed with $\boldsymbol{\Sigma} = \operatorname{var}(\operatorname{vec} \boldsymbol{B}_t) = 0.2 \cdot \boldsymbol{I}$ and $\boldsymbol{G} = \operatorname{var} \boldsymbol{Y}_t = \boldsymbol{I}$. Notice that assumption A2 is fulfilled, because the matrix $\mathrm{E}(\boldsymbol{B}_0 \otimes \boldsymbol{B}_0) + (\boldsymbol{\beta} \otimes \boldsymbol{\beta})$ has eigenvalues $0.583, 0.064 \pm 0.03i, 0.050$ that are less than one in the modulus.

Then we estimated the parameter $\boldsymbol{\beta}$ using both $\widehat{\boldsymbol{\beta}}_{LS}$ given by equation (7) and $\widehat{\boldsymbol{\beta}}_{\phi}$ given by (10) with $\phi(\boldsymbol{z}) = \frac{\boldsymbol{z}}{1+\boldsymbol{z}'\boldsymbol{z}}$. We ran the simulation 1000 times. The results of the simulation study are displayed in Figure 1 and in Table below. The estimators are compared using sample means and density estimations of the 1000 estimated values of the true parameters β_{11}, β_{12}, β_{21} and β_{22} (we used the default density estimation procedure in the R programming language). We can see that the least-squares estimator $\widehat{\boldsymbol{\beta}}_{LS}$ always underestimates the true value, especially for β_{11} and β_{21} whereas the estimator $\widehat{\boldsymbol{\beta}}_{\phi}$ is closer to the true values. The density estimation also reveals bias for the least-squares estimator. These results are in accordance with previous simulations made for univariate RCA processes, see [16].

parameter	**true value**	**LS est.**	$\phi(\boldsymbol{z}) = \frac{\boldsymbol{z}}{1+\boldsymbol{z}'\boldsymbol{z}}$
β_{11}	0.2	0.163	0.182
β_{12}	0.1	0.088	0.098
β_{21}	0.3	0.257	0.278
β_{22}	0.4	0.391	0.401

Tab. 1. Average values of estimated parameters in simulated model (23).

We also simulated 500 observations of the one-dimensional RCA process of second order given by

$$X_t = (0.4 + B_{t1}) X_{t-1} + (0.2 + B_{t2}) X_{t-2} + Y_t. \tag{24}$$

The random coefficients $B_t = (B_{t1}, B_{t2})'$ were independent identically normally distributed with zero mean and diagonal variance matrix $\mathrm{Diag}[\sigma_1^2, \sigma_2^2]$, where σ_1^2, σ_2^2 will be specified below, and $\{Y_t\}$ was a sequence of *iid* random errors with standard normal distribution, independent of $\{B_t\}$.

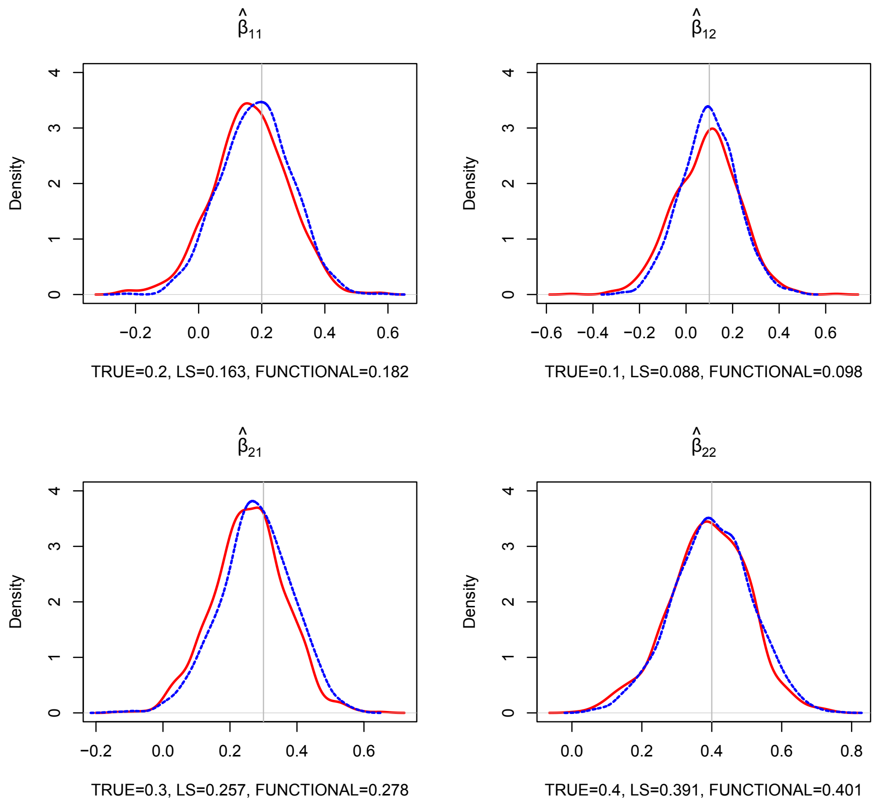

Fig. 1. Density of estimated values of the parameters (solid line for the least-squares estimator, dotted line for Schick-type estimator) in simulated 2-dimensional RCA(1) process (23).

In our case,

$$\boldsymbol{\beta} = \begin{pmatrix} 0.4 & 0.2 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{B}_t = \begin{pmatrix} B_{t1} & B_{t2} \\ 0 & 0 \end{pmatrix}$$

and for $\sigma_1^2, \sigma_2^2 \in (0; 0.05; 0.10; \ldots; 0.35)$, all the eigenvalues of $\mathrm{E}(\boldsymbol{B}_0 \otimes \boldsymbol{B}_0) + \boldsymbol{\beta} \otimes \boldsymbol{\beta}$ are in the unit circle and thus the assumption A2 is satisfied. We estimated the parameter $\beta = (\beta_1, \beta_2)'$ using $\widehat{\beta}_{LS}$ and $\widehat{\beta}_\phi$ with $\phi(\boldsymbol{z}) = \frac{\boldsymbol{z}}{1+\boldsymbol{z}'\boldsymbol{z}}$, respectively, and compared the mean square errors. The results are displayed in Figures 2 and 3 as functions of σ_1^2 for a fixed level of σ_2^2. The estimators $\widehat{\beta}_{LS}$ and $\widehat{\beta}_\phi$ with ϕ as above behave similarly when the random coefficients have low variances, however, $\widehat{\beta}_{LS}$ underestimates the true value of the parameter when the variances increase. On the other hand, $\widehat{\beta}_\phi$ is stable no matter how large the variances of the random coefficients are and the estimated values do not vary so much even if only 100 repetitions are used.

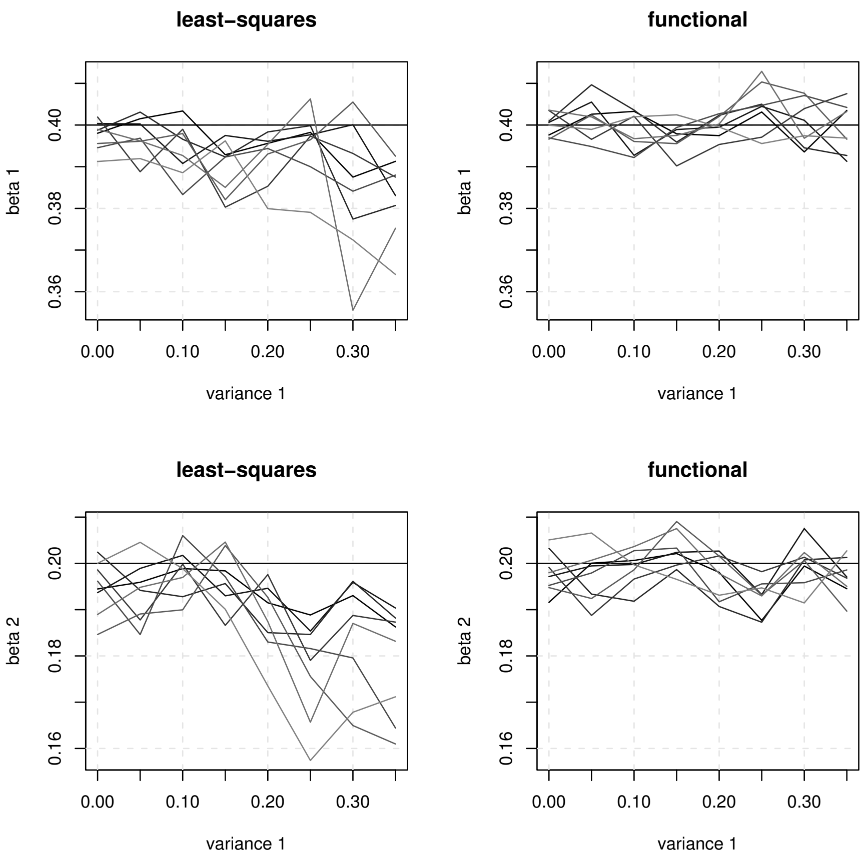

Fig. 2. Estimates of true parameter $\beta = (0.4, 0.2)'$ in simulated model (24). Upper panels stand for estimation of β_1, lower for β_2, left for the least-squares estimator, right for the Schick-type estimator. Horizontal axis stands for variance σ_1^2, curves correspond to various choices of variance σ_2^2.

A. APPENDIX

Definition A.1. We introduce the following notions.

1. Let $\boldsymbol{A}$ be an $m \times n$ matrix. Then the mn-component vector $vec(\boldsymbol{A})$ is defined as stacking the columns of $\boldsymbol{A}$, one on top of the other in order from left to right.

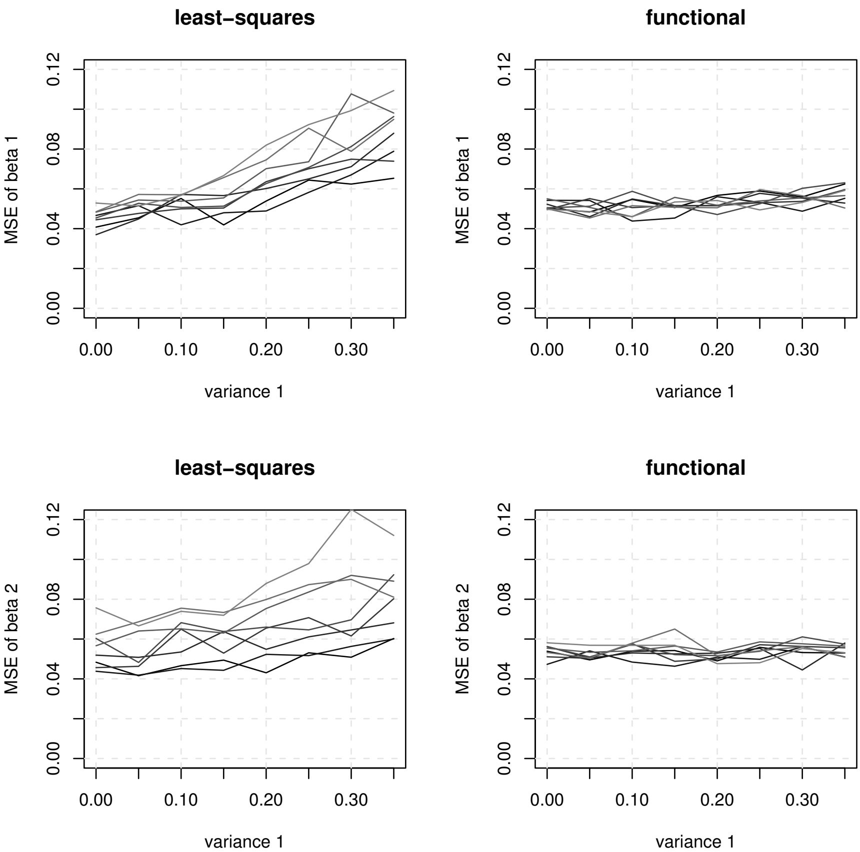

Fig. 3. MSE of estimates of parameters β in simulated model (24). Upper panels for β_1, lower for β_2, left panels for least-squares estimator, right for the Schick-type estimator. Horizontal axis stands for variance σ_1^2, curves correspond to various choices of variance σ_2^2.

2. Let $\boldsymbol{A}$ be an $n \times n$ symmetric matrix. Then the $n(n+1)/2$-component vector $vech(\boldsymbol{A})$ is defined as stacking those parts of the columns of $\boldsymbol{A}$ on and below the main diagonal, one on top of the other in order from left to right.

3. Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{B}$ be a $p \times q$ matrix. Then the *Kronecker product* $\boldsymbol{A} \otimes \boldsymbol{B}$ of $\boldsymbol{A}$ and $\boldsymbol{B}$ is the $mp \times nq$ matrix whose (i,j)th block is the $p \times q$ matrix $a_{i,j} \cdot \boldsymbol{B}$, where $a_{i,j}$ is the (i,j)th element of $\boldsymbol{A}$.

Lemma A.2. Let $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$, $i = 1, 2, \dots, n$, be any matrices. Then

$$\left(\prod_{i=1}^{n} \boldsymbol{A}_i\right) \otimes \left(\prod_{j=1}^{n} \boldsymbol{B}_j\right) = \prod_{i=1}^{n} (\boldsymbol{A}_i \otimes \boldsymbol{B}_i) .$$

Proof. See [13], Paragraph 3 in Chapter 7. □

Lemma A.3. Let $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ be any matrices and $\boldsymbol{u}$, $\boldsymbol{v}$ be any vectors such that the expressions below are well defined. Then

(a) $\operatorname{vec}(\boldsymbol{ABC}) = (\boldsymbol{C}' \otimes \boldsymbol{A}) \cdot \operatorname{vec}(\boldsymbol{B})$

(b) $\boldsymbol{A}\mathbf{u} = (\mathbf{u}' \otimes \boldsymbol{I}) \cdot \operatorname{vec}(\boldsymbol{A})$

(c) $\operatorname{vec}(\boldsymbol{A} \cdot \boldsymbol{B}) = (\boldsymbol{B}' \otimes \boldsymbol{I}) \cdot \operatorname{vec}(\boldsymbol{A})$

(d) $\operatorname{vec}(\boldsymbol{u}\boldsymbol{v}') = (\boldsymbol{v} \otimes \boldsymbol{I}) \cdot \boldsymbol{u}$.

Proof. Properties (a) and (b) are proved in [13], Paragraph 5 in Chapter 7. Properties (c) and (d) are direct applications of property (a). □

Lemma A.4. Let $\boldsymbol{A}$ be an $n \times n$ symmetric matrix. Then there exist constant $(n(n+1)/2) \times n^2$ matrices $\boldsymbol{K}_n$ and $\boldsymbol{H}_n$ for which

$$\begin{aligned} \operatorname{vech}(\boldsymbol{A}) &= \boldsymbol{H}_n \operatorname{vec}(\boldsymbol{A}) \\ \operatorname{vec}(\boldsymbol{A}) &= \boldsymbol{K}_n' \operatorname{vech}(\boldsymbol{A}) \end{aligned}$$

such that $\boldsymbol{H}_n \boldsymbol{K}_n' = \boldsymbol{I}_{n(n+1)/2}$. Matrices $\boldsymbol{H}_n$ and $\boldsymbol{K}_n$ are sometimes called the elimination and duplication matrices, respectively.

Proof. See Theorem A.1.3 in [11]. □

Lemma A.5. Let $\{\boldsymbol{A}_n, n \in \mathbb{N}\}$ and $\{\boldsymbol{B}_n, n \in \mathbb{N}\}$ be random $r \times r$-dimensional matrix processes. Assume that $\{\boldsymbol{B}_n\}$ is ergodic with finite second moments and that $\boldsymbol{A}_n \xrightarrow{\text{a. s.}} \boldsymbol{A}$ as $n \to +\infty$ where $\boldsymbol{A} \in \mathbb{R}^{r \times r}$ is a finite constant matrix. Then

$$\frac{1}{n} \sum_{t=1}^{n} \boldsymbol{B}_t \boldsymbol{A}_n \boldsymbol{B}_t' \xrightarrow{\text{a. s.}} \mathrm{E}(\boldsymbol{B}_1 \boldsymbol{A} \boldsymbol{B}_1') \text{ as } n \to +\infty.$$

Proof. We prove the convergence of the correspondent components using ergodicity and vec operator:

$$\begin{aligned} \operatorname{vec}\left(\frac{1}{n} \sum_{t=1}^{n} \boldsymbol{B}_t \boldsymbol{A}_n \boldsymbol{B}_t'\right) &= \frac{1}{n} \sum_{t=1}^{n} \operatorname{vec}\left(\boldsymbol{B}_t \boldsymbol{A}_n \boldsymbol{B}_t'\right) = \frac{1}{n} \sum_{t=1}^{n} \boldsymbol{B}_t \otimes \boldsymbol{B}_t \cdot \operatorname{vec}(\boldsymbol{A}_n) \\ &\xrightarrow{\text{a. s.}} \mathrm{E}\left(\boldsymbol{B}_1 \otimes \boldsymbol{B}_1\right) \cdot \operatorname{vec}(\boldsymbol{A}) = \mathrm{E}\left(\operatorname{vec}\left(\boldsymbol{B}_1 \boldsymbol{A} \boldsymbol{B}_1'\right)\right) = \operatorname{vec}\left(\mathrm{E}\left(\boldsymbol{B}_1 \boldsymbol{A} \boldsymbol{B}_1'\right)\right). \end{aligned}$$

□

Lemma A.6. Consider block matrix

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}' & \boldsymbol{C} \end{pmatrix},$$

where $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ are $p \times p$-dimensional matrices. Let $\boldsymbol{M} \geq \boldsymbol{0}$, $\boldsymbol{B}$ be a invertible matrix, and $\boldsymbol{C}$ be a symmetric invertible matrix. Then

$$\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}'^{-1} - \boldsymbol{C}^{-1} \geq \boldsymbol{0}.$$

Proof. We will prove that $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}' \geq \boldsymbol{0}$ from which the result could be obtained by multiplication of matrices $\boldsymbol{B}^{-1}$ and $\boldsymbol{B}'^{-1}$, respectively. Choose arbitrary $\boldsymbol{x} \in \mathbb{R}^p$ and $\boldsymbol{y} \in \mathbb{R}^p$. Matrix $\boldsymbol{M}$ is positively semi-definite which means that

$$(\boldsymbol{x}', \boldsymbol{y}')\boldsymbol{M} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} + \boldsymbol{y}'\boldsymbol{B}'\boldsymbol{x} + \boldsymbol{x}'\boldsymbol{B}\boldsymbol{y} + \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y} = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} + \boldsymbol{y}'\boldsymbol{C}\boldsymbol{y} + 2\boldsymbol{x}'\boldsymbol{B}\boldsymbol{y} \geq 0.$$

We want to prove that for any $\boldsymbol{x} \in \mathbb{R}^p$,

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}'\boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}'\boldsymbol{x} \geq 0.$$

Comparing the previous two inequalities we conclude that it suffices to find $\boldsymbol{y} \in \mathbf{R}^p$ such that

$$\boldsymbol{y}'\boldsymbol{C}\boldsymbol{y} + 2\boldsymbol{x}'\boldsymbol{B}\boldsymbol{y} = -\boldsymbol{x}'\boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}'\boldsymbol{x}.$$

Denote $\boldsymbol{z} = \boldsymbol{B}'\boldsymbol{x}$. Then $\boldsymbol{y}$ has to satisfy the elliptical equation

$$\boldsymbol{y}'\boldsymbol{C}\boldsymbol{y} + 2\boldsymbol{z}'\boldsymbol{y} + \boldsymbol{z}'\boldsymbol{C}^{-1}\boldsymbol{z} = 0$$

and $\boldsymbol{y} = -\boldsymbol{C}^{-1}\boldsymbol{z}$ solves the latter equation which completes the proof. □

ACKNOWLEDGEMENT

This work was supported by grants GAČR 201/09/0775, GAČR 201/09/J006 and of the Ministry of Education, Youth and Sports of the Czech Republic 0021620839.

(Received January 3, 2011)

REFERENCES

[1] A. Aue, L. Horváth, and J. Steinebach: Estimation in random coefficient autoregressive models. J. Time Ser. Anal. *27* (2006), 60–67.

[2] S. Y. Hwang and I. V. Basawa: Parameter estimation for generalized random coefficient autoregressive processes. J. Statist. Plann. Inference *68* (1998), 323–337.

[3] I. Berkes, L. Horváth, and S. Ling: Estimation in nonstationary random coefficient autoregressive models. J. Time Ser. Anal. *30* (2009), 395–416.

[4] P. Billingsley: The Lindeberg–Lévy theorem for martingales. Proc. Amer. Math. Soc. *12* (1961), 788–792.

[5] A. Brandt: The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. Adv. Appl. Probab. *18* (1986), 211–220.

[6] P. Bougerol and N. Picard: Strict stationarity of generalized autoregressive processes. Ann. Probab. *20* (1992), 1714–1730.

[7] J. Davidson: Stochastic Limit Theory. Advanced Texts in Econometrics. Oxford University Press, Oxford 1994.

[8] P. D. Feigin and R. L. Tweedie: Random coefficient autoregressive processes: A Markov chain analysis of stationarity and finiteness of moments. J. Time Ser. Anal. *6* (1985), 1–14.

[9] H. Janečková and Z. Prášková: CWLS and ML estimates in a heteroscedastic RCA(1) model. Statist. Decisions *22* (2004), 245–259.

[10] H. L. Koul and A. Schick: Adaptive estimation in a random coefficient autoregressive model. Ann. Statist. *24* (1996), 1025–1052.

[11] D. F. Nicholls and B. G. Quinn: Random coefficient autoregressive models: An introduction. Lecture Notes in Statistics *11*, Springer, New York 1982.

[12] A. Schick: $\sqrt{n}$-consistent estimation in a random coefficient autoregressive model. Austral. J. Statist. *38* (1996), 155–160.

[13] J. Schott: Matrix Analysis for Statistics. Wiley Series in Probability and Statistics, Wiley, New York 1996.

[14] P. Vaněček: Rate of convergence for a class of RCA estimators. Kybernetika *6* (2006), 698–709.

[15] P. Vaněček: Estimators of multivariate RCA models. In: Bull. Internat. Statistical Institute LXII (M. I. Gomes at al., eds.), Instituto Nacional de Estatística, Lisbon 2007, pp. 4027–4030.

[16] P. Vaněček: Estimation of Random Coefficient Autoregressive Models. PhD Thesis, Charles University, Prague 2008.

Zuzana Prášková, Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Praha 8. Czech Republic.

e-mail: praskova@karlin.mff.cuni.cz

Pavel Vaněček, TNS AISA s.r.o., Budějovická 1518/13B, 140 00 Praha 4. Czech Republic.

e-mail: pavel.vanecek@tns-global.com