Statistical Inference in Infinite-Order Cointegrated Vector Autoregressive Processes Under Uncorrelated but Dependent Errors

Chafik Bouhaddioui*

Abstract
The concept of cointegration processes is one of the most used concepts in economics and finance. Mainly, researchers are interested in behavior of the estimators of the model parameters. In this paper, we will investigate the asymptotic behavior of the estimators of an infinite-order cointegrated vector autoregressive series under nonindependent errors by showing its asymptotic distribution. Using this result, we will construct a Likelihood Ratio (LR) test of the cointegration rank. One can also develop a method under unrestrictive assumptions to select the autoregressive order.

Key Words: Cointegrated Processes, infinite order, Weak errors, LR tests, QML estimation.

1. Introduction

Multivariate time series are widely used in economics since a substantial part of economic theory generally deals with long-run equilibrium relationships generated by market forces and behavioral rules. In order to study the long run relationship, Engle & Granger (1987) introduced the concept of cointegration which is used in many recent studies across several fields. One can say that time series variables are cointegrated if they have a common stochastic trend, or simply, a linear combination of these variables can be represented by a stationary process. The number of independent linear combinations is the cointegrating rank and is an important parameter in analyzing economic data. However, if cointegrating relations are present in a system of variables, the vector autoregressive (VAR) form is not the most convenient model setup. In that case, it is useful to consider specific parametrization that supports the analysis of the cointegration structure known as vector error correction models (VECM). For the VECM, most studies suppose two important assumptions. The first assumption is related to the order of the VECM representation which is supposed finite and known. Of course, this assumption is unrealistic in practice for various reasons. For instance, the true data-generation processes (DGP) may not be a finite order process. If it is a finite order VAR process then the true order is not likely to be known. Therefore it is of interest to know the consequences of a violation of the assumption that the DGP is a VAR process with known finite order. The second condition is related to the innovations process which are supposed to be independent and identically distributed (i.i.d). This assumption is too restrictive when economic or financial data is to study. Most of the macroeconomic time series exhibit a conditional heteroscedasticity or any nonlinear form. To introduce the IVAR(∞), Saikkonen (1992) proposed a different way to write a VECM representation than the well known one proposed by Johansen (1988), see Juselius (2006).
To clarify details, let consider the following d-dimensional process $Y = \{Y_t\ ,\ t \in \mathbb{Z}\}$. The data generating process has the form:

$$\Delta Y_t \quad = \quad J\Theta' Y_{t-1} + u_t \tag{1}$$

*UAE University, Department of Statistics, P.O.Box 15551, Al Ain UAE

where, $\Delta Y_t = Y_t - Y_{t-1}$, J and Θ are matrices containing the model parameters and the process u_t is assumed to have an infinite-order VAR representation

$$\sum_{l=0}^{\infty} G_l u_{t-l} = \epsilon_t, \qquad G_0 = \mathcal{I}_d, \tag{2}$$

where ϵ_t is usually assumed i.i.d with mean zero and positive definite covariance matrix Σ_ϵ. Under some regular conditions on G_l, the process u_t and hence the process Y_t can be approximated by a finite order autoregression. In the literature, studying the behavior of the parameter estimators of an IVAR(∞) with i.i.d errors was done by Saikkonen (1992) and Saikkonen & Lütkepohl (1996). They showed the asymptotic properties of the estimated coefficients of the autoregressive error correction model (VECM) and the pure vector autoregressive (VAR) representations derived under the assumption that the autoregressive order goes to infinity with the sample size. Under the same strong assumptions on the innovations process, Saikkonen & Lütkepohl (1996) constructed a test for linear (zero) restrictions which arise in exogeneity or Granger causality analyses. In Bouhaddioui & Dufour (2008), the rate of convergence of the least square (LS) estimators was found and used to test the non-correlation between two IVAR(∞). In this project, under the more general model which is represented by an infinite-order cointegrated process with uncorrelated but dependent errors, denoted by WIVAR(∞), we will address two main aims.

The First aim is to study the behavior of the estimators of the cointegration. The difficulty of this problem comes from the fact that we have to consider the approximation of the infinite-order cointegrated process by a finite-order autoregressive process where the order of the fitted autoregression is a function of the sample size and the fact that the errors are uncorrelated but dependent which needs the use of a more general result than the central limit theorem which assume that the errors are independent. In the case of i.i.d errors, Saikkonen & Lütkepohl (1996) showed that the least square (LS) estimators of the coefficients representing the short-run dynamics in the VECM and the pure VAR are seen to have asymptotic normal distribution when suitable normalizations are used. Seo (2007) explored the asymptotic distribution of the maximum likelihood estimator (MLE) of a finite-order cointegrating vector with conditionally heteroskedastic errors. He showed that the MLE of the cointegrating vector follows mixture normal, and its asymptotic distribution depends on the conditional heteroskedasticity and the kurtosis of the innovations. However, the conditional heteroskedasticity does not preclude other forms of dependence, see Francq & Zakoian (1998) and Francq, Roy & Zakoian (2005).

The second aim is to study the validity the likelihood ratio (LR) tests of the cointegration rank where the process is WIVAR(∞) . This aim is mainly related to the first one, since the investigation of the cointegration properties at early stage (before the estimation stage) of the analysis has become standard practice by now. The problem is more complicate than the regular LR tests proposed by Johansen (1988), Johansen (1991) and Reinsel & Ahn (1992) where the tests are performed conditionally on the order being the true one. In some studies, it was shown that the choice of the lag order or truncation lag has an important impact on the unit root and cointegration tests, see Ng & Perron (1995) and Haug (1996). Lütkepohl & Saikkonen (1997) extended the LR tests proposed by Ng & Perron (1995) to the case of the IVAR(∞). Rahbek, Hansen & Dennis (2002) studied the impact of ARCH innovations on the LR test. An important output of their work is that the LR test remains valid when the error process is a martingale difference. Also, In the context of uncorrelated but dependent errors, Raïssi (2009) showed that the LR test remains valid where the order of the VECM is predetermined and finite.

2. Preliminaries and Notations

Following the notations of Saikkonen (1992) and Saikkonen & Lütkepohl (1996), we consider a d-dimensional process $\boldsymbol{X} = \{\boldsymbol{X}_t\ ,\ t \in \mathbb{Z}\}$ partitioned into two subprocesses $\boldsymbol{X}_i = \{\boldsymbol{X}_{it}\ ,\ t \in \mathbb{Z}\}$, $i = 1, 2$, with d_1 and d_2 components respectively ($d_1 + d_2 = d$). The data generating process has the form:

$$\boldsymbol{X}_{1t} = \boldsymbol{C}_1 \boldsymbol{X}_{2t} + \boldsymbol{\varepsilon}_{1t}, \quad (3)$$

$$\Delta \boldsymbol{X}_{2t} = \boldsymbol{\varepsilon}_{2t}, \quad (4)$$

where $\boldsymbol{C}_1$ is a given $d_1 \times d_2$ matrix, Δ is the usual difference operator, and $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}'_{1t}, \boldsymbol{\varepsilon}'_{2t})'$ is a stationary process with zero mean and continuous spectral density matrix which is positive definite at zero frequency. $\boldsymbol{X}_{2t}$ is an integrated vector process of order one (with no cointegrating relationship), while $\boldsymbol{X}_{1t}$ and $\boldsymbol{X}_{2t}$ are cointegrated. By taking first differences in (3), the above system can be written in the form

$$\Delta \boldsymbol{X}_t = \begin{bmatrix} -\mathbb{I}_{d_1} & \boldsymbol{C}_1 \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{X}_{t-1} + \boldsymbol{b}_t = \boldsymbol{J}\boldsymbol{\Theta}' \boldsymbol{X}_{t-1} + \boldsymbol{b}_t \quad (5)$$

where $\mathbb{I}_d$ represents the $d \times d$ identity matrix, $\boldsymbol{J}' = [-\mathbb{I}_{d_1} : \boldsymbol{0}]$, $\boldsymbol{\Theta}' = [\mathbb{I}_{d_1} : -\boldsymbol{C}_1]$, $\boldsymbol{b}_t = [\boldsymbol{b}'_{1t} : \boldsymbol{b}'_{2t}]'$ is nonsingular transformation of $\boldsymbol{\varepsilon}_t$ defined by

$$\boldsymbol{b}_{1t} = \boldsymbol{\varepsilon}_{1t} + \boldsymbol{C}_1 \boldsymbol{\varepsilon}_{2t}\ ,\ \boldsymbol{b}_{2t} = \boldsymbol{\varepsilon}_{2t}\,. \quad (6)$$

The notation $\boldsymbol{A} = [\boldsymbol{A}_1 : \boldsymbol{A}_2]$ means that the matrix $\boldsymbol{A}$ is partitioned into a matrix $\boldsymbol{A}_1$ consisting of the first d_1 columns and a matrix $\boldsymbol{A}_2$ with d_2 columns.

We suppose also that the process $\boldsymbol{b}_t$ (and hence $\boldsymbol{\varepsilon}_t$) has an infinite-order autoregressive representation

$$\sum_{l=0}^{\infty} \boldsymbol{G}_l \boldsymbol{b}_{t-l} = \boldsymbol{a}_t, \qquad \boldsymbol{G}_0 = \mathbb{I}_d, \quad (7)$$

where $\boldsymbol{a}_t$ is independent and identically distributed white noise process with $\mathbb{E}(\boldsymbol{a}_t) = \boldsymbol{0}$ and $\mathbb{E}(\boldsymbol{a}_t \boldsymbol{a}'_t) = \boldsymbol{\Sigma}_{\boldsymbol{a}}$ is a definite positive matrix. Setting $\boldsymbol{G}(z) = \boldsymbol{I}_d - \sum_{l=1}^{\infty} \boldsymbol{G}_l z^l$, the stationarity hypothesis of the process $\boldsymbol{b}_t$ implies that the zeros of the equation $\det\{\boldsymbol{G}(z)\} = 0$ all lie outside the unit circle $|z| = 1$, where $\det\{\boldsymbol{A}\}$ denotes the determinant of the square matrix $\boldsymbol{A}$. A further assumption is that the coefficient matrices $\boldsymbol{G}_l$ satisfy the summability condition

$$\sum_{l=1}^{\infty} l^n \|\boldsymbol{G}_l\| < \infty \quad (8)$$

for some $n \geq 1$ and $\|.\|$ is the Euclidean matrix norm defined by $\|\boldsymbol{A}\|^2 = tr(\boldsymbol{A}'\boldsymbol{A})$. This is a standard condition for weakly stationary processes, which ensures that the process be well defined. Depending on n, it imposes weak restrictions on the autocorrelation structure of the process $\boldsymbol{b}_t$. Also, it implies that the process $\boldsymbol{b}_t$ and, consequently, $\boldsymbol{X}_t$ can be approximate by a finite-order autoregression. The order p_N of the fitted autoregression is a function of the sample size; *i.e.*, $p_N = p(N)$. In order to reduce approximation errors, we allow the maximal order p_N to increase to infinity, at some rate, simultaneously with realization length N, see Burnham & Anderson (2002). In the sequel, we assume the following assumption on the finite autoregressive order.

Assumption 2.1 *$N^{-1/3}p_N \to 0$ and $\sqrt{p_N}\sum_{l=p_N+1}^{\infty}\|\boldsymbol{G}_l\| \to 0$ as $N \to \infty$.*

The condition $p_N = o(N^{1/3})$ for the rate of increase of p_N ensures that enough sample information is asymptotically available for estimators to have standard limiting distributions. The condition $\sqrt{p_N}\sum_{j=p_N+1}^{\infty}\|\boldsymbol{G}_j\| \to 0$ imposes a lower bound on the growth rate of p_N, which ensures that the approximation error of the true underlying model by a finite-order autoregression gets small when the sample size increases. A more detailed discussion of these conditions is available in Burnham & Anderson (2002) and Lütkepohl (2005).

Using the equations (5) - (7) and rearranging terms, we obtain the autoregressive *error correction model* (ECM) representation

$$\Delta \boldsymbol{X}_t = \boldsymbol{\Psi}\boldsymbol{\Theta}'\boldsymbol{X}_{t-1} + \sum_{l=1}^{p_N}\boldsymbol{\Pi}_l\boldsymbol{\Delta X}_{t-l} + \boldsymbol{e}_t, \qquad t = p_N+1, p_N+2, \ldots \tag{9}$$

where $\boldsymbol{e}_t = \boldsymbol{a}_t - \sum_{l=p_N+1}^{\infty}\boldsymbol{G}_l\boldsymbol{b}_{t-l}$, $\boldsymbol{\Psi} = -\sum_{l=0}^{p_N}\boldsymbol{G}_l\boldsymbol{J}$, and the $d \times d_1$ matrix $\boldsymbol{\Psi}$ is of full column rank (at least for p_N large enough). Details for this derivation can be found in Saikkonen & Lütkepohl (1994). Note that the coefficient matrices $\boldsymbol{\Pi}_l (l = 1, \ldots, p_N)$ are functions of $\boldsymbol{\Theta}$ and $\boldsymbol{G}_l (l = 1, 2, ...)$, and they depend on p_N. Furthermore, the sequence $\boldsymbol{\Pi}_l (l = 1, \ldots, p_N)$ is absolutely summable as $p_N \to \infty$.

The autoregressive ECM in (9) can also be rewritten in a pure vector autoregressive (VAR) form

$$\boldsymbol{X}_t = \sum_{l=1}^{p_N+1}\boldsymbol{\Phi}_l\boldsymbol{X}_{t-l} + \boldsymbol{e}_t \tag{10}$$

where $\boldsymbol{\Phi}_1 = \mathbb{I}_d + \boldsymbol{\Psi}\boldsymbol{\Theta}' + \boldsymbol{\Pi}_1$, $\boldsymbol{\Phi}_l = \boldsymbol{\Pi}_l - \boldsymbol{\Pi}_{l-1}$, $l = 2, \ldots, p_N$ and $\boldsymbol{\Phi}_{p_N+1} = -\boldsymbol{\Pi}_{p_N}$. Although the $\boldsymbol{\Pi}_l$ depend on p_N, the same is not true for the $\boldsymbol{\Phi}_l$ except for $\boldsymbol{\Phi}_{p_N+1}$.

Saikkonen & Lütkepohl (1996) derived the asymptotic properties of the multivariate least square (LS) estimators of the VAR coefficients under a standard assumption. Let $\boldsymbol{\Phi}(p_N) = (\boldsymbol{\Phi}_1, \ldots, \boldsymbol{\Phi}_{p_N})$ be the matrix of the first p_N autoregressive parameter matrices in the representation (10) and denote by $\hat{\boldsymbol{\Phi}}(p_N) = (\hat{\boldsymbol{\Phi}}_1, \ldots, \hat{\boldsymbol{\Phi}}_{p_N})$ the corresponding LS estimator. Let the process $\boldsymbol{e}_t$ satisfy the following assumptions:

Assumption 2.2 (Uncorrelated but Dependent Errors) *The error process $\boldsymbol{e}_t$ is strictly stationary and such that $Cov(\boldsymbol{e}_t, \boldsymbol{e}_{t-h}) = 0$ for all $t \in \mathbb{Z}$ and all $h \neq 0$.*

Such error processes are commonly named *weak white noise*. Now, let consider the α-mixing coefficients

$$\alpha_a(l) = \sup_{A\in\sigma(a_u,u\leq t),B\in\sigma(a_u,u\geq t+l)} |P(A\cap B) - P(A)P(B)| \tag{11}$$

that measure the temporal dependence of the stationary process $\{a_t\}$ and the following assumption on the structure of the errors.
Let consider also the following two assumptions:

Assumption 2.3 *The process $\boldsymbol{e}_t$ satisfies $\|\boldsymbol{e}_t\|_{2+\nu+\eta} < \infty$, and the mixing coefficients of the process $\boldsymbol{e}_t$ are such that $\sum_{l=0}^{\infty}\{\alpha_{\boldsymbol{e}}(l)\}^{\nu/(2+\nu)} < \infty$ for some $\nu > 0$ and $\eta > 0$.*

Note that this kind of assumptions are very common and considered as a mild assumption for the process $\boldsymbol{e}_t$. In addition, the folowing assumption is needed to strength 2.3

Assumption 2.4 *The process $\boldsymbol{e}_t$ satisfies $\|\boldsymbol{e}_t\|_{4+2\nu} < \infty$, and the mixing coefficients of the process $\boldsymbol{e}_t$ are such that $\sum_{l=0}^{\infty}\{\alpha_{\boldsymbol{e}}(l)\}^{\nu/(2+\nu)}$ for some $\nu > 0$.*

3. Consistency of QMLE Estimators

In the case of *i.i.d* error process, Saikkonen & Lütkepohl (1996) derived the asymptotic properties of the multivariate least square (LS) estimators of the VAR coefficients. The following theorem is a generalization of this result to the case of weak error process. First, let $\mathbf{\Phi}(p_n) = (\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_{p_n})$ be the matrix of the first p_n autoregressive parameter matrices in the representation (10) and denote by $\hat{\mathbf{\Phi}}(p_n) = (\hat{\mathbf{\Phi}}_1, \ldots, \hat{\mathbf{\Phi}}_{p_n})$ the corresponding Quasi Maximum Likelihood Estimators (QMLE).

Theorem 3.1 *Let $\{\boldsymbol{X}_t\}$ a process given by* (10) *and assume that $\mathbf{e}_t$ satisfies Assumptions 2.2-2.4. Then,*

$$\sqrt{n/p_n^3}\left(\hat{\mathbf{\Phi}}(p_n) - \mathbf{\Phi}(p_n)\right) \Rightarrow \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\phi}) \tag{12}$$

where $\mathbf{\Sigma}_{\phi}$ is well defined covariance matrix.

The proof and the details of $\hat{\mathbf{\Phi}}(p_n)$ and $\mathbf{\Sigma}_{\phi}$ can be found in the technical report Bouhaddioui (2013) upon request from the author. This theorem is more general than the one proved by Raïssi (2009) which shows the consistency for the decomposed parameters in (10) $\mathbf{\Psi}$ and $\mathbf{\Theta}$.

4. LR test under weak error process

Now, we will study the validity the likelihood ratio (LR) tests of the cointegration rank where the process is WIVAR(∞). This aim is mainly related to the first one, since the investigation of the cointegration properties at early stage (before the estimation stage) of the analysis has become standard practice by now. The problem is more complicate than the regular LR tests proposed by Johansen (1988), Johansen (1991) and Reinsel & Ahn (1992) where the tests are performed conditionally on the order being the true one. In some studies, it was shown that the choice of the lag order or truncation lag has an important impact on the unit root and cointegration tests, see Ng & Perron (1995).
Lütkepohl & Saikkonen (1997) extended the LR tests proposed by Ng & Perron (1995) to the case of the IVAR(∞). Rahbek et al. (2002) studied the impact of ARCH innovations on the LR test. If r_0 is the cointegrated rank, let consider the test where, for some r ($0 \le r \le d$), the null hypothesis

$$\mathcal{H}_0 :\ r_0 = r \ \text{ vs. } \ \mathcal{H}_1 :\ r_0 > r.$$

Now, in order to test the null hypothesis, we consider the LR test statistic

$$\mathcal{R} \ = \ -2\log \mathcal{Q}_r = -T \sum_{i=r+1}^{d} \log(1-\hat{\lambda}_i),$$

where $\hat{\lambda}_1 \ge \ldots \ge \hat{\lambda}_d$ are the d largest solutions of the eigenvalue for a well defined matrix. The main result of this part is given by the following theorem.

Theorem 4.1 *Let $\{\boldsymbol{X}_t\}$ a process given by* (10) *and assume that $\mathbf{e}_t$ satisfies Assumptions 2.2-2.4. Then, the LR test statistic, denoted by $\mathcal{R}$, has the same asymptotic distribution as in the iid Gaussian case and true VAR order is known, i.e.*

$$\mathcal{R}_0 \Rightarrow tr\left(\left[\int_0^1 \mathbb{K}(dB)'\right]' \left[\int_0^1 \mathbb{K}\mathbb{K}' du\right]^{-1} \left[\int_0^1 \mathbb{K}(dB)'\right]\right) \tag{13}$$

where $\mathbb{K}$ is a function of standard Brownian motion.

A detailed proof of this theorem can be found in the technical report.

References

Bouhaddioui, C. (2013), Statistical inference in infinite-order cointegrated vector autoregressive processes under uncorrelated but dependent errors, Technical report.

Bouhaddioui, C. & Dufour, J.-M. (2008), 'Tests for non-correlation of two infinite-order cointegrated vector autoregressive series', *Journal of Applied Probability and Statistics* **3**(1), 78–94.

Burnham, K. P. & Anderson, D. R. (2002), *Model selection and multimodel inference : a practical information-theoretic approach*, Springer-Verlag, New York.

Engle, R. & Granger, C. (1987), 'Co-integration and error correction: Representation, estimation and testing', *Econometrica* **55**, 251–276.

Francq, C., Roy, R. & Zakoian, J.-M. (2005), 'Diagnostic checking in ARMA models with uncorrelated errors', *Journal of the American Statistical Analysis* **100**, 532–544.

Francq, C. & Zakoian, J.-M. (1998), 'Estimating linear representations of nonlinear processes', *Journal of Statistical Planning and Inference* **68**, 145–165.

Haug, A. A. (1996), 'Tests for cointegartion: A monte carlo comparison', *Journal of Econometrics* **71**, 89–115.

Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control* **12**(2-3), 231–254.

Johansen, S. (1991), 'Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models', *Econometrica* **59**(6), 1551–80.

Juselius, K. (2006), *The Cointegrated VAR model*, Oxford University Press.

Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Lütkepohl, H. & Saikkonen, P. (1997), 'Impulse response analysis in infinite order cointegrated vector autoregressive processes', *Journal of Econometrics* **81**, 127–157.

Ng, S. & Perron, P. (1995), 'Unit root tests ARMA models with data dependent methods for the selection of the truncation lag', *Journal of the American Statistical Association* **90**, 268–281.

Rahbek, A., Hansen, E. & Dennis, J. (2002), ARCH innovations and their impact on cointegration rank testing, Technical report.

Raïssi, H. (2009), 'Testing the cointegration rank with uncorrelated but dependent errors', *Stochastic Analysis and Applications* **27**(1), 24–50.

Reinsel, G. C. & Ahn, S. K. (1992), 'Vector autoregressive models with unit roots and reduces rank structure: Estimation, likelihood ratio test, adn forecasting', *Journal of Time Series Analysis* **13**, 353–375.

Saikkonen, P. (1992), 'Estimation and testing of cointegrated systems by an autoregressive approximation', *Econometric Theory* **8**, 1–27.

Saikkonen, P. & Lütkepohl, H. (1994), Infinite order cointegrated autoregressive processes: Estimation and inference, Technical Report 5, Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin.

Saikkonen, P. & Lütkepohl, H. (1996), 'Infinite-order cointegrated vector autoregressive processes: Estimation and inference', *Econometric Theory* **12**, 814–844.

Seo, B. (2007), 'Asymptotic distribution of the cointegrating vector estimator in error correction models with conditional heteroskedasticity.', *Journal of Econometric* **137**, 532–544.