Asymptotic normality of least squares estimators in autoregressive linear regression models
van der Genugten, B.B.

Publication date:
1985

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright, please contact us providing details, and we will remove access to the work immediately and investigate your claim.
ASYMPTOTIC NORMALITY OF LEAST SQUARES ESTIMATORS IN AUTOREGRESSIVE LINEAR REGRESSION MODELS

by B.B. van der Genugten
ASYMPTOTIC NORMALITY OF LEAST SQUARES ESTIMATORS IN AUTOREGRESSIVE LINEAR REGRESSION MODELS

by B.B. van der Genugten

For the linear regression model $y_t = \beta' x_t + \epsilon_t$ $(t = 1, \ldots, n)$ asymptotic normality of the least squares estimator of $\beta$ is proved in the case that the $\epsilon_t$ are mutually independent with finite second moments and that $\epsilon_t$ is independent of $x_1, \ldots, x_t$ for each $t$. The results obtained are applied to autoregressive models with nonstochastic, possibly unbounded regressors.

1) AMS 1970 subject classifications. Primary 62J05; Secondary 60F05.
Key words and phrases: Least squares estimators, linear regression, asymptotic normality, weak consistency, autoregressive models.
1. Introduction

In this paper we derive central limit theorems for LS-estimators in linear models with stochastic explanatory variables (regressors). Conditions for these variables bring together the purely deterministic case and the mixed autoregressive case in which some of these variables may be lagged dependent variables.

In the existing literature for the mixed case strong conditions on the behaviour of the regressors are imposed, which are not satisfied for very simple cases. The conditions in this paper are rather weak and cover a wide range of models of practical importance.

The analysis is based on some recently obtained central limit theorems for martingales and some newly developed weak laws of large numbers for linear combinations of random variables. By a special method of normalization the usual but rather unnatural and strong conditions for convergence of the regressors (in some sense) to non-trivial values are circumvented.

We restrict ourselves to univariate models. Generalizations for the multivariate case are easily obtained. So, we shall be concerned with the model

\[ y_t = \beta' x_t + \epsilon_t, \quad t \in \mathbb{N} \]

Here, \((x_t)_{t \in \mathbb{N}}\) is a sequence of random \(k \times 1\)-vectors of explanatory variables and \((\epsilon_t)_{t \in \mathbb{N}}\) an error sequence of random variables, both defined on some probability space \((\Omega, \mathcal{F}, P)\). The \(k \times 1\) vector \(\beta\) is the nonstochastic vector of regression coefficients.

The assumptions are precisely stated in section 2. We will discuss the main features here.

We assume that the errors are independent and that they are of the same order of magnitude (formally expressed by the Eicker conditions (2.1) below).

The assumptions of independent instead of uncorrelated errors is made to avoid conditions about the existence of moments higher than the second order. Especially for i.i.d. errors assumptions about higher order moments are superfluous and we do not want to lose this special case.
Since the case of lagged dependent variables has to be included, we cannot assume that the error process is independent of the regressor process. It is only reasonable to assume that the errors at time \( t \) are independent of the regressors before or at time \( t \) (conditions (2.2) below). Furthermore we will only consider the (weak) non-collinear case, i.e. the assumptions are always such that they imply \( P\{S_n > 0\} \to 1, n \to \infty \), where

\[
S_n = \sum_{t=1}^{n} x_t x'_t.
\]

A least squares (LS) estimator \( b_n \) of \( \beta \), based on the first \( n \) observations, is defined by

\[
b_n = S_n^{-1} \sum_{t=1}^{n} x_t y_t,
\]

where \( S_n^{-1} \) is some pseudo-inverse of \( S_n \). In the non-collinear case the probability of the set for which \( S_n^{-1} \) is not defined tends to zero. So there can be no confusion in writing \( S_n^{-1} \) instead of \( S_n^{-1} \) since we deal with convergence in distribution. Therefore we will do this from now on.

Theorem 2.1 below gives conditions for the \( x_t \) and \( \epsilon_t \) under which \( b_n \) is asymptotically normal, i.e. there exists a sequence \( C_n \) of non-stochastic positive definite \( k \times k \) matrices, not depending on \( \beta \), such that

\[
\mathcal{L}(C_n^{-1}(b_n - \beta)) \to N_k(0, I), \quad n \to \infty.
\]

Here \( N_k(0, I) \) denotes the \( k \)-dimensional standard normal distribution. For \( C_n \) we do not take \( \text{Cov}(S_n^{-1} \sum_{t=1}^{n} x_t \epsilon_t) \) but

\[
C_n = \text{Cov}(S_n^{-1} \sum_{t=1}^{n} x_t \epsilon_t),
\]

where \( S_n = E(S_n) \). This particular choice admits the use of central limit theorems for martingales. The theorem generalizes the results for non-stochastic \( x_t \) of Eicker [4]; theorem 3.1 or [5], section 2, theorem. Furthermore, the conditions are such that they apply to autoregressive models of the form
(1.2) \[ \sum_{g=0}^{p} \alpha_g y_{t-g} = \gamma' w_t + \varepsilon_t, \quad t \in \mathbb{N}, \quad (a_0 = 1). \]

Here, \( \alpha = (\alpha_1, \ldots, \alpha_p)' \) and \( \gamma \) are nonstochastic \( p \times 1 \) and \( q \times 1 \) vectors, respectively, of regression coefficients, \( w_t \) is a nonstochastic \( q \times 1 \) vector, and the initial values \( y_{1-p}, \ldots, y_0 \) are random variables. By taking \( \beta' = (\gamma', -\alpha') \) and \( x_t' = (w_t', y_{t-1}', \ldots, y_{t-p}') \) we see that (1.2) is a special case of (1.1).

Theorem 2.2 below states that for this autoregressive model the conditions for the \( x_t \) in theorem 2.1 are implied by certain conditions for the \( w_t \) and the \( \alpha, \gamma \). The proof of this result forms the hard and technical part of the paper. The result is attractive because in these conditions the \( \alpha, \gamma \) appear in a simple way. In particular there is no need to solve the difference equation (1.2) explicitly.

Finally, theorem 2.3 below is added to simplify the verification of the conditions of theorem 2.2. The result in the first part of example 2.2 given after this theorem is simpler but can be compared with results in the literature up till now such as Anderson [1], theorem 5.5.14 or Schönfeld [10]. The example 2.3 there after shows clearly the advantages of the methods used in this paper. Simple practical problems with unbounded \( (w_t) \) can be solved too. So we can say that an old problem which went back to Mann and Wald [8] has found a satisfactory solution now in the theorems 2.1 and 2.2.

From a statistical point of view it is an interesting question if \( C_n \) can be estimated while preserving asymptotic normality of \( b_n \). This can be done along the lines of Eicker [4]. Here \( \sigma_t^2 \) is replaced by the square of the least-squares-residual at \( t \). We will not go into details.
2. Statement of the results

In the following we suppose that the $\varepsilon_t$ are mutually independent with $E[\varepsilon_t] = 0$, $\sigma_t^2 = E[\varepsilon_t^2] < \infty$, $t \in \mathbb{N}$. Furthermore, we assume that they satisfy the Eicker conditions (see e.g. Eicker [4], theorem 3.1):

\[(2.1) \quad \inf_{t \in \mathbb{N}} \sigma_t^2 > 0, \quad \sup_{t \in \mathbb{N}} E[\varepsilon_t^2 I(\varepsilon_t^2 > \delta)] \to 0, \quad \delta \to \infty.\]

In the following conditions it is assumed that $S_n > 0$ for some (and therefore all) $n$ sufficiently large whenever $S_n^{-1}$ appears.

Theorem 2.1 (general model (1.1)). If

\[(2.2) \quad \varepsilon_t \text{ is independent of } x_1, \ldots, x_t \text{ and } E[x_t^2] < \infty \text{ for each } t \in \mathbb{N},\]

\[(2.3) \quad \frac{S_n^{-1}}{n} \left( \sum_{t=1}^{n} \lambda_t x_t x_t' - E \left( \sum_{t=1}^{n} \lambda_t x_t x_t' \right) \right) S_n^{-1} \overset{P}{\to} 0 \]

for $\lambda = 1$ and $\lambda_t = \sigma_t^2$,

\[(2.4) \quad \max_{1 \leq t \leq n} x_t' S_n^{-1} x_t \overset{P}{\to} 0,\]

then

$$\chi^2 \left( C_n^{-1}(b_n - \beta) \right) \to N_k(0, I).$$

Remark. The condition (2.3) for $\lambda_t = 1$ gives $S_n^{-1} S_n^{-1} \overset{P}{\to} I$, implying weak noncollinearity.

Example 2.1 (nonstochastic $x_t$). For such $x_t$ the conditions (2.2), (2.3) are fulfilled in a trivial way and (2.4) reduces to

$$\max_{1 \leq t \leq n} x_t' S_n^{-1} x_t \to 0$$

or, equivalently (lemma 3.5 below),...
This implies that $x_n$ is non-exponentially increasing (i.e. $\rho^n x_n \to 0$ for any $\rho$ with $|\rho| < 1$). However, polynomial trends are included. For a discussion of the necessity of these conditions and for some special cases see Eicker [5].

For the general model (1.1) the appearance of (2.4) in theorem 2.1 is motivated by example 2.1. The condition (2.3) for $\lambda_t = \sigma_t^2$ admits the use of a central limit theorem for martingales.

For the model (1.2) we assume that the starting values $y_{-p}, \ldots, y_0$ are independent of $(\epsilon_t)_{t \in \mathbb{N}}$ and have finite second moments. Let the $(p+q) \times 1$ vector $v_t$ be defined by

$$v_t' = (\sum_{g=0}^{p} \alpha g^t \gamma_{t-g}, \gamma_{t-1}, \ldots, \gamma_{t-p}), \quad t = p + 1, p + 2, \ldots$$

and the $(p+q) \times (p+q)$ matrix $Z_n$ by

$$Z_n = \sum_{t=p+1}^{n} v_t v_t' + n I_0.$$

Here $I_0$ is a $(p+q) \times (p+q)$ matrix with all elements equal to zero except the last $p$ elements of its diagonal which are equal to one. Finally, let

$$A(z) = \sum_{g=0}^{p} \alpha g z^g, \quad z \in \mathbb{C}.$$

We have:

**Theorem 2.2** (autoregressive model (1.2)). If

(2.5) $A(z) \neq 0, \quad |z| < 1,$

(2.6) $\|Z_n^{-1}\| = O(1/n)$,

(2.7) $v_n' Z_n^{-1} v_n \to 0$. 


then the conditions (2.2) - (2.4) of theorem 2.1 are satisfied.

For the model (1.2) the condition (2.5) is the usual stability condition. The conditions (2.6) and (2.7) are more difficult to explain. Roughly spoken, in some sense, $v_t$ can serve as an approximation for $\mu_t = E(x_t)$. Then $Ev'_t$ can be considered to be an approximation of the part of $S_n$ coming from the $w_t$. The other part of $S_n$ coming from products of errors is approximated by $nI_n$. So $Z_n$ is an approximation of $S_n$. In fact we have that $cZ_n \leq S_n \leq c'Z_n$ for some $c, c' > 0$ and $n$ sufficiently large. So we may hope that (2.4) can be replaced by the condition $\max_v v'_t z t^{-1} v_t + 0$ or equivalently, using lemma 3.5,b), $Z_t^{-1} + 0$ and (2.7). The condition (2.6) is slightly stronger, and is needed to justify the approximations as well as to verify the weak law of large numbers as stated in (2.3).

For many interesting sequences $(w_t)$ the verification of the conditions (2.6), (2.7) of theorem 2.2 is not easy. Therefore we add theorem 2.3 below to simplify this verification.

Suppose for some integer $s > q$ there exists a sequence $(q_t)^{t \geq p+1}$ of nonstochastic $s \times 1$ vectors such that

$$v_t = \Phi q_t, \quad t = p + 1, p + 2, \ldots$$

for some nonstochastic $(p+q) \times s$ matrix $\Phi$. Let $(D_n)^{n \geq p+1}$ be a sequence of $s \times s$ nonstochastic positive definite diagonal matrices. Consider the normed $s \times 1$ vectors

$$\tilde{q}_t(n) = D_n^{-\frac{1}{2}} q_t, \quad t = p + 1, \ldots, n.$$

For some integer $r$ such that $q \leq r \leq \min(s, p + q)$ let $\Phi_{00}$ be the $r \times r$ submatrix of $\Phi$ formed by the first $r$ rows and $r$ columns of $\Phi$, let $D_{n0}$ be a similar $r \times r$ submatrix of $D_n$, and for the case that $r < s$ let $D_{n1}$ be the $(s-r) \times (s-r)$ submatrix of $D_n$ formed by the last $s-r$ rows and $s-r$ columns of $D_n$. We have:
Theorem 2.3. Suppose

(2.8) \[ \det(\phi_{00}) \neq 0, \]

(2.9) \[ \|D_{n0}^{-1}\| = O(1/n), \]

(2.10) \[ \text{if } r < s \text{ then } \|D_{ni}\| = O(1/\|D_{n0}^{-1}\|), \]

(2.11) \[ \left\| \sum_{p+1}^{n} q_t(n)q_t'(n)^{-1} \right\| = O(1), \]

(2.12) \[ \tilde{q}_n(n) \to 0, \]

then the conditions (2.6), (2.7) of theorem 2.2 are satisfied.

Example 2.2. We apply theorem 2.3 for the simple case \( s = (p+1)q \) and \( q_t' = (w'_t, \ldots, w'_{t-p}). \) Then \( v_t = \phi q_t \) with

\[
\phi = \begin{bmatrix}
\alpha_0 I & \alpha_1 I & \cdots & \alpha_p I \\
0 & \gamma' & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \gamma'
\end{bmatrix}
\]

Take \( D_n = nI. \) Then the conditions (2.11), (2.12) are satisfied if

(2.13) \[ \left\| \sum_{p+1}^{n} q_tq_t' \right\| = O(1/n), \quad q_n = o(\sqrt{n}) \]

Take \( r = q. \) Then (2.9), (2.10) are fulfilled. Since \( \phi_{00} = \alpha_0 I \) the condition (2.8) holds as well. Therefore, with theorem 2.3 we see that (2.13) leads to the asymptotic normality of the LS-estimator in the autoregressive model (1.2) for all \( \gamma \) and all \( \alpha \) satisfying (2.5). The condition (2.13) is implied by

\[
\frac{1}{n} \sum_{p+1}^{n} q_tq_t' \to Q
\]
for some \( q > 0 \).

Note that (2.13) will not be satisfied if the model contains a constant term. However, a slight modification of the derivation above leads again to a condition of the type (2.13). We write \( \mathbf{w}_t' = (1, \tilde{w}_t'), \gamma' = (\gamma_0', \tilde{\gamma}') \). Choose \( s = pq + q - p \) and \( \mathbf{q}_t' = (1, \tilde{w}_t', \ldots, \tilde{w}_{t-p}) \). Then \( \mathbf{v}_t = \mathbf{q}_t \) with

\[
\phi = \begin{bmatrix}
\mathbf{P} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{0} & \mathbf{g} & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{0} & \mathbf{a}_0 \mathbf{I} & \mathbf{a}_1 \mathbf{I} & \cdots & \mathbf{a}_p \mathbf{I} \\
\mathbf{Y}_0 & 0 & \tilde{\gamma}' & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\mathbf{Y}_0 & 0 & 0 & \cdots & \tilde{\gamma}'
\end{bmatrix}
\]

Take \( r = q \) then again \( \det(\phi_{00}) \neq 0 \) since \( \sum \mathbf{a} \neq 0 \) because of (2.5).

**Example 2.3.** Let \( \mathbf{w}_t' = (t_1', \ldots, t_p') \) for integers \( a_1 > a_2 > \ldots > a_q \geq 0 \). If \( a_1 > 0 \) then (2.13) of example 2.2 cannot be applied. For \( q \leq r \leq \min (a_1 + 1, p + q) \) we can write

\[
\begin{bmatrix}
\mathbf{w}_t \\
\mathbf{w}_{t-1} \\
\vdots \\
\mathbf{w}_{t-p}
\end{bmatrix} = \begin{bmatrix}
\psi_0 & \ast \\
\psi_1 & \ast \\
\vdots & \vdots \\
\psi_p & \ast
\end{bmatrix} \begin{bmatrix}
a_1 \\
t \\
a_1 - 1 \\
t \\
\vdots \\
1
\end{bmatrix}
\]

Here \( \psi_0', \ldots, \psi_p \) are \( q \times r \) upper triangular matrices. Not interesting matrices are denoted by stars. Choose \( s = a_1 + 1 \) and \( \mathbf{q}_t' = (t_1', t_1', \ldots, t_1') \). Then \( \mathbf{v}_t = \mathbf{q}_t \) with
Take $D_n = \text{diag}(n, \ldots, n)$. Then

$$
\sum_{t=1}^{n} \tilde{q}_t(n)\tilde{q}_t(n) + H
$$

where $H = \{(2a_1+3-i-j)^{-1}; i,j = 1, \ldots, a_1 + 1\}$ is the so-called Hilbert-matrix. Note that $H > 0$ and $\tilde{q}_t(n) \to 0$ for fixed $t$. Therefore the condition (2.11) is satisfied. Since

$$
\tilde{q}_n(n) = D^{-\frac{1}{2}}_n q_n = n^{-\frac{1}{2}}(1, 1, \ldots, 1)'
$$

we see that (2.12) holds. Furthermore,

$$
D_{n0} = \text{diag}(n, \ldots, n)
$$

$D_{n1} = \text{diag}(n, \ldots, n)$ if $r < a_1 + 1$,

and so the conditions (2.9), (2.10) are fulfilled also. Therefore it remains to verify the condition (2.8). In general this cannot be done for all $\gamma$ and all $\alpha$ satisfying (2.5).

1°) Consider the case of successive powers

$$
a_q = a_1 - q + 1
$$

Then all $\psi_j$ have diagonal elements 1. From (2.5) it follows that $\sum_{j=1}^{p} \alpha_j \neq 0$. Hence, if we take $r = q$ then $\phi_{00} = \sum_{j=1}^{p} \alpha_j \psi_j$ is nonsingular, and consequently the condition (2.8) is fulfilled for all $\gamma$ and all $\alpha$ satisfying (2.5).

2°) Consider the particular case of non-successive powers $p = 1, q = 2$, $a_1 = 2, a_2 = 0$. Set $\gamma' = (\gamma_1, \gamma_2)$. The choice $r = 3$ leads to

$$
\phi = \begin{bmatrix}
\alpha_0 \psi_0 & \alpha_1 \psi_1 & \ldots & \alpha_p \psi_p \\
0 & \gamma' \psi_0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \gamma' \\
\end{bmatrix}
$$

$$
= \phi_{00} = \begin{bmatrix}
\sum_{j=1}^{p} \alpha_j \psi_j \\
0 & \gamma' \psi_1 \\
\vdots & \vdots & \ddots & \vdots \\
\gamma' \psi_{r-q} \\
\end{bmatrix}
$$

$$
\begin{bmatrix}
2a_1 + 1 & 2a_1 - 1 \\
0 & 1 \\
\vdots & \vdots \\
0 & 0 \\
\end{bmatrix}
$$

$$
\begin{bmatrix}
\psi_0 \\
\psi_1 \\
\vdots \\
\psi_p \\
\end{bmatrix}
$$

$$
\begin{bmatrix}
\psi_0 \\
\psi_1 \\
\vdots \\
\psi_p \\
\end{bmatrix}
$$
\[ \phi_{00} = \begin{bmatrix} a_0 + a_1 & -2a_1 & a_1 \\ 0 & 0 & a_0 + a_1 \\ \gamma_1 & -2\gamma_1 & \gamma_1 + \gamma_2 \end{bmatrix} \Rightarrow \det(\phi_{00}) = 2a_0\gamma_1(a_0 + a_1). \]

So the condition (2.8) is fulfilled for all \( \gamma \) with \( \gamma_1 \neq 0 \) and for all \( a \) satisfying (2.5). This makes clear that in this example it is difficult to prove the asymptotic normality for all \( \gamma \) and all \( a \) satisfying (2.5).

A special case of example 2.3, (10) is the model

\[ y_t + a_1 y_{t-1} = \gamma_1 + \gamma_2 t + \varepsilon_t \]

which can serve as an empirical description of the log of the Dutch national income during the years 1949-1968. As far as I know the available asymptotic results on literature to justify the usual estimation and testing procedures are not applicable to this simple model. This was a part of the motivation to write this paper.
3. Proofs of the theorems

We mention the following notations and conventions. For the norm \( \|A\| \) of an arbitrary matrix \( A \) we take \( \lambda_{\max}^\frac{1}{2} (A'A) \). We write \( A \geq 0 \) if \( A \) is positive semidefinite and \( A > 0 \) if \( A \) is positive definite. We write \( A \geq B \) if \( A \geq 0 \), \( B > 0 \) and \( A - B \geq 0 \). Note that \( A \geq 0 \) implies \( \|A\| = \lambda_{\max}^\frac{1}{2} (A) \) and \( A > 0 \) that \( \|A^{-1}\| = \lambda_{\min}^\frac{1}{2} (A) \).

The random sequence \( (x_n)_{n \in \mathbb{N}} \) is called \( P \)-bounded if for every \( \varepsilon > 0 \) there exists a number \( M \) such that \( P\{|x_n| > M\} < \varepsilon \) for all \( n \). If \( (x_n)_{n \in \mathbb{N}} \) and \( (y_n)_{n \in \mathbb{N}} \) are sequences of random variables such that \( (x_n) \) is \( P \)-bounded and \( y_n \xrightarrow{P} 0 \) then \( x_n y_n \xrightarrow{P} 0 \). If \( \mathbb{E}|x_n| \) is bounded then \( (x_n) \) is \( P \)-bounded.

The proof of theorem 2.1 is based on a central limit theorem for martingale triangular arrays. Let

\[
1 = \inf_{t \in \mathbb{N}} \sigma_t^2 \quad \text{and} \quad m = \sup_{t \in \mathbb{N}} \sigma_t^2
\]

then (2.1) implies

\[
(3.1) \quad 0 < 1 < m < \infty
\]

**Proof of theorem 2.1.** We take

\[
C_n = \text{Cov} \left( \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t \varepsilon_{nt} \right).
\]

Then \( C_n \) does not depend on \( \beta \). We have

\[
C_n (\beta - \beta) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t \varepsilon_t = \left[ \sqrt{n} (S_n S_n^{-1} S_n^{-1} S_n^{-1} I) S_n \right] \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t \varepsilon_t.
\]

From (2.2) and (3.1) we get

\[
(3.2) \quad 1 - C_n \leq \frac{1}{\sqrt{n}} \mathbb{E} \left( \sum_{t=1}^{n} x_t x_t' \right) \frac{1}{\sqrt{n}} \mathbb{E} \leq m.
\]
This implies that $\|C_n^{-\frac{1}{2}} S_n^{-1} \|$ and $\|C_n^{-\frac{1}{2}} S_n^{-1} P_n \|$ are bounded. From (2.3) for $\lambda_t = 1$ it follows that $S_n^{-\frac{1}{2}} S_n^{-1} P_n \overset{p}{=} I$. Therefore in the relation above (3.2) the expression between square brackets tends to $I$ in probability. Thus it remains to prove that

$\mathcal{L}\{C_n^{-\frac{1}{2}} S_n^{-1} \Sigma x \varepsilon_t \} \rightarrow N_k(0, I)$.

Let $a \in \mathbb{R}^k$ with $|a| = 1$ and set $x_{nt} = a^t C_n^{-\frac{1}{2}} S_n^{-1} x_t \varepsilon_t$. Then it suffices to prove

\[ (3.3) \quad \mathcal{L}\{\Sigma x_{nt} \} \rightarrow N(0, 1). \]

Let $\mathcal{F}_n = \mathcal{F}(x_1^t)$, $\mathcal{F}_{nt} = \mathcal{F}(x_1^t, \ldots, x_{t+1}^t, \varepsilon_1^t, \ldots, \varepsilon_t^t)$ for $1 \leq t \leq n$. Then the $\mathcal{F}_{nt}$ are sub-$\sigma$-fields of $\mathcal{F}$ with $\mathcal{F}_{n,t-1} \subset \mathcal{F}_{nt}$. Furthermore, $x_t$ is $\mathcal{F}_{nt}$-measurable and $\varepsilon_t$ is independent of $\mathcal{F}_{n,t-1}$ as follows from (2.2). This gives $E(x_{nt} \mid \mathcal{F}_{n,t-1}) = 0$ a.s., and so $(x_{nt})_{1 \leq t \leq n, n \in \mathbb{N}}$ is a martingale triangular array (MTA). Furthermore,

\[ E \left\{ \Sigma x_{nt} \left\{ |x_{nt} \mid > \varepsilon \right\} \mid \mathcal{F}_{n,t-1} \right\} \rightarrow 0, \quad \text{for every } \varepsilon > 0. \]

Using (2.3) for $\lambda_t = \sigma_t^2$ and the boundedness of $\|C_n^{-\frac{1}{2}} S_n^{-1} \|$.

Then (3.3) follows from a central limit theorem for MTA's (see e.g. Mcleish [9], Corollory (3.8) or Gaenssler [6], theorem 2), provided that we can verify the conditional Lindeberg-Feller condition.

\[ (3.4) \quad E \left\{ x_{nt}^2 I(\{\|x_{nt} \| \geq \varepsilon\}) \right\} \rightarrow 0, \quad \text{for every } \varepsilon > 0. \]

Set

\[ r_{nt} = x_t^t S_n^{-1} x_t, \quad r_n = \max_{1 \leq t \leq n} r_{nt} \]

\[ Q_t(\delta) = E(\varepsilon_t^2 I(\{|\varepsilon_t| \geq \delta\})) \quad \text{and} \quad R_n(\delta) = \max_{1 \leq t \leq n} Q_t(\delta) \]

Then (2.3) gives

\[ x_{nt}^2 \leq (r_{nt} / m) \cdot \varepsilon_t^2 \leq (r_{nt} / m) \cdot \varepsilon_t^2. \]
Since $\epsilon_t$ is independent of $\mathcal{F}_{n,t-1}$ it follows that
\[
\sum_{i=1}^{n} E\left( x_{n,t}^2 I\left( |x_{n,t}| > \epsilon \right) | \mathcal{F}_{n,t-1} \right) < \\
\sum_{i=1}^{n} r_{n,t} E\left( \epsilon_t^2 I\left( |\epsilon_t| > \epsilon \sqrt{m/r_{n,t}} \right) | \mathcal{F}_{n,t-1} \right) = \\
\sum_{i=1}^{n} r_{n,t} Q_t(\epsilon \sqrt{m/r_{n,t}}) \\
\leq R_n(\epsilon \sqrt{m/r_{n,t}}) \sum_{i=1}^{n} r_{n,t}.
\]

From (2.1) we get that $R_n(\delta_n) \to 0$ for any nonstochastic sequence $(\delta_n)$ with $\delta_n \to \infty$, and from (2.4) that $r_n \to 0$. By considering the a.s. convergence of subsequences it follows that $R_n(\epsilon \sqrt{m/r_{n,t}}) \to 0$. Furthermore, $E\left( \sum_{i=1}^{n} r_{n,t} \right) = k$ and $r_{n,t} \to 0$ imply that $\sum_{i=1}^{n} r_{n,t}$ is P-bounded. Together this shows that the right-hand side of (3.5) tends to 0 in probability, completing the proof.

The proof of theorem 2.2 is rather tedious. We make some preliminary remarks and formulate some intermediate results as lemma's. Let
\[
m_n = \max_{1 \leq i < n} \sigma_{n,t}^2, \quad s_n = \sum_{i=1}^{n} \sigma_{n,t}^2
\]
From (2.1) it follows that $s_n \geq n\ell$ and that for every $\epsilon > 0$
\[
(3.6) \quad s_n^{-1} \sum_{i=1}^{n} E\left( \epsilon_t^2 I\left( |\epsilon_t| > \epsilon \sqrt{s_n} \right) \right) \leq 1^{-1} \sup_{t \in \mathbb{N}} E\left( \epsilon_t^2 I\left( |\epsilon_t| > \epsilon \sqrt{s_n} \right) \right) \to 0.
\]
So the Eicker conditions (2.1) imply the well-known Lindeberg-Feller (LF) condition for the sequence $(\epsilon_{n,t})_{t \in \mathbb{N}}$. This leads to the corollary of lemma 3.1 below. The lemma itself is a slight generalization of a theorem of Raikov (see Gnedenko [7], § 28, theorem 4). Its proof is kept short and added for the sake of completeness.

Lemma 3.1. Let $(x_{n,j})_{1 \leq j \leq k_n}$, $n \in \mathbb{N}$ be a triangular array of random variables such that $x_{n,1}, \ldots, x_{n,k_n}$ are mutually independent for each fixed $n$. 
Suppose $E\{x_{nj}\} = 0$, $\sigma_{nj}^2 = E\{x_{nj}^2\} < \infty$ for all $n,j$.

If the LF condition
$$\sum_{j=1}^{k_n} E\{x_{nj}^2 I(|x_{nj}| > \varepsilon)\} \to 0,$$
for every $\varepsilon > 0$,
holds, and if $\sum_{j=1}^{k_n} \sigma_{nj}^2$ is bounded in $n$, then

$$\max_{1 \leq j < k_n} |x_{nj}| \overset{P}{\to} 0$$

$$\max_{1 \leq j < k_n} \sigma_{nj}^2 \to 0,$$

and for any bounded nonstochastic array $(\lambda_{nj})$ we have

$$\sum_{j=1}^{k_n} \lambda_{nj} (x_{nj}^2 - \sigma_{nj}^2) \overset{P}{\to} 0.$$

**Proof.** The first relation follows from

$$P\{\max |x_{nj}| \geq \varepsilon\} \leq \sum_{j=1}^{k_n} P\{|x_{nj}| \geq \varepsilon\} \leq \varepsilon^{-2} \sum_{j=1}^{k_n} E\{x_{nj}^2 I(|x_{nj}| > \varepsilon)\}$$

and the second one from

$$\sigma_{nj}^2 \leq \varepsilon^2 + \max E\{x_{nj}^2 I(|x_{nj}| \geq \varepsilon^2\} \leq \varepsilon^2 + \sum_{j=1}^{k_n} E\{x_{nj}^2 I(|x_{nj}| > \varepsilon)\}.$$

The last relation will be proved first for $\lambda_{nj} = 1$.

Let $\varphi_{nj}(u) = E\{\exp(\iu u x_{nj}\})$ then it suffices to prove that

$$\sum_{j=1}^{k_n} (\log \varphi_{nj}(u) - \iu \sigma_{nj}^2) \to 0.$$

Since $|\varphi_{nj}(u) - 1| \leq |u| \sigma_{nj}^2$ we have

$$\max_{1 \leq j < k_n} |\varphi_{nj}(u) - 1| \to 0,$$

since $|\log(1+z) - z| \leq |z|^2$, $|z| \leq 1/2$, this gives

$$\sum_{j=1}^{k_n} |\log \varphi_{nj}(u) + 1 - \varphi_{nj}(u)| \to 0.$$
\[ |\varphi_{n_j}(u) - 1 - iu\sigma_{n_j}^2| \leq \frac{1}{2} |u|^2 \varepsilon g_{n_j}^2 + 2|u|E(x_{n_j}^2 I(x_{n_j}^2 \geq \varepsilon)) \]

we obtain

\[ \sum_{k \leq n_j} |\varphi_{n_j}(u) - 1 - iu\sigma_{n_j}^2| \to 0 \]

Together this gives \( \sum_{k \leq n_j} (\log \varphi_{n_j}(u) - iu\sigma_{n_j}^2) \to 0 \), proving the result for \( \lambda_{n_j} \).

For non-negative \( \lambda_{n_j} \) the result follows from this. Replace only \( x_{n_j} \) by \( \lambda_{n_j} x_{n_j} \).

For arbitrary \( \lambda_{n_j} \) the result follows by splitting up the sum for positive and negative \( \lambda_{n_j} \).

**Corollary.** By taking \( x_{n_j} = \varepsilon_j / \sqrt{s_n} \) we see that the LF-condition for \( (x_{n_j}) \) follows from (3.6). Hence,

\[ (3.7) \quad s_n^{-1/2} \max_{1 \leq t \leq n} |\varepsilon_t| \mathbb{P} 0 \]

\[ (3.8) \quad m_n / s_n \to 0 \]

and for bounded nonstochastic \( (\lambda_{nt})_{0 \leq t \leq n}, n \in \mathbb{N} \):

\[ (3.9) \quad s_n^{-1/2} \sum_{t=1}^{n} \lambda_{nt} (\varepsilon_t - \sigma_t^2) \mathbb{P} 0 \]

**Lemma 3.2.** For nonstochastic \( \phi_h \) and \( \psi_h \) \((h = 0, 1, \ldots)\) let

\[ y_t = \sum_{h=0}^{\infty} \phi_h \varepsilon_{t-h}, \quad z_t = \sum_{h=0}^{\infty} \psi_h \varepsilon_{t-h}, \quad t \in \mathbb{Z}, \]

where \( \varepsilon_t = 0 \) if \( t \leq -p, \sigma_t^2 = E(\varepsilon_t^2) < \infty \) if \( 1 - p < t \leq 0. \)
If $\Sigma_{h}z^h$, $\Sigma_{h}z^h$, have convergence radii larger than 1, then for any bounded nonstochastic array $\left(\lambda_{nt}\right)_{1 \leq t \leq n}$, $n \in \mathbb{N}$ we have

\begin{equation}
(3.10) \quad \sum_{t=1}^{n} \left( \lambda_{nt} z_{t-i} z_{t-j} - \mathbb{E} \left[ \sum_{t=1}^{n} \lambda_{nt} z_{t-i} z_{t-j} \right] \right) \leq 0, \quad i, j \in \mathbb{Z}
\end{equation}

**Proof.** Note that $\varphi_h = o(p^h)$, $\psi_h = o(p^h)$ for some $\rho$ with $0 < \rho < 1$.

Set $\varphi_r = \psi_r = 0$ if $r < 0$. Then

\begin{equation}
(3.11) \quad \sum_{t=1}^{n} \lambda_{nt} z_{t-i} z_{t-j} = \sum_{t=1}^{n} \lambda_{nt} \left( \varphi_{r} \varepsilon_{t-r-i} \psi_{s} \varepsilon_{t-j-s} \right) = \sum_{t=1}^{n} \lambda_{nt} \varphi_{r} \psi_{s} \varepsilon_{r} \varepsilon_{s}
\end{equation}

where

$$a_{nrs} = \sum_{t=1}^{n} \lambda_{nt} \varphi_{t-i-r} \psi_{t-j-s}$$

Note that $a_{nrs} = 0$ if $r > n - i$ or $s > n - j$. For any $a \geq b$ we have

$$|\sum_{t=1}^{n} \lambda_{nt} \varphi_{t-a} \psi_{t-b}| \leq c'. \sum_{t=a}^{n} |\varphi_{t-a} \psi_{t-b}| \leq c'' \sum_{t=a}^{n} \rho^{t-a} \rho^{t-b} \leq c''' \rho^{a-b}$$

for some constants $c'$, $c''$, $c'''$. Therefore there exists a constant $c$ not depending on $n$, $r$, $s$ such that

$$|a_{nrs}| \leq c \rho^{|r-s|}.$$

In particular, $a_{nrs}$ is bounded in $n$, $r$, $s$. We split up

$$\sum_{r,s}^{n} a_{nrs} \varepsilon_{r} \varepsilon_{s} = \sum_{k=1}^{5} A_{k}(n)$$

according to the ranges $(r, s \leq 0)$, $(r > 1, s \leq 0)$, $(r < 0, s \geq 1)$, $(r, s > 1$ and $r \neq s)$, $(r, s > 1$ and $r = s)$, respectively.

The number of terms in $A_{1}(n)$ is finite and the terms themselves are bounded in $n$. Furthermore,

$$E|A_{2}(n)| < \sum_{s \leq 0}^{n-i} \left| a_{nrs} \right| E|\varepsilon_{r} \varepsilon_{s}| \leq c' \sum_{s \leq 0}^{n-i} \sum_{r=1}^{n-i} \rho^{r-s} \rho^{s-r} \leq c'' \sum_{s \leq 0}^{n-i} \left(1 - \rho^{r-s} \rho^{s-r} \right) \rho^{n-i},$$

for some constants $c'$, $c''$. A similar relation holds for $A_{3}(n)$. Hence, with (3.8) we get
\( (3.12) \quad s_n^{-1} E |A_k(n)| = 0 , \quad k = 1,2,3. \)

Since \( E[A_4(n)] = 0 \) and

\[
V[A_4(n)] = \sum_{r \neq s} \sum_{a_{nrs} + a_{nrs}} \sigma_r^2 \sigma_s^2 \leq 2c^2 \sum_{r \neq s} \rho^2 |r-s| \sigma_r^2 \sigma_s^2 = 2c^2 (\sum_{r \neq s} \rho^2 (s-r) + \sum_{r \neq s} \rho^2 (r-s)) \sigma_r^2 \sigma_s^2 < \]

\[
\leq c' \left( \sum_{n-i-s} m_{n-j} s_{n-i} + s_{n-i} s_{n-j} \right)
\]

for some \( c' \), we get with (3.8) that

\( (3.13) \quad s_n^{-1} (A_4(n) - E[A_4(n)]) \xrightarrow{P} 0. \)

Finally, with (3.9) we get

\( (3.14) \quad s_n^{-1} (A_5(n) - E[A_5(n)]) = s_n^{-1} \max(i,j) \sum_{t=1} \sum_{a_{ntt} (\epsilon^2 - \sigma_t^2) \xrightarrow{P} 0. \)

Combining (3.11) - (3.14) gives (3.10).

Lemma 3.3. Let \( \{x_t\}_{t \in \mathbb{N}} \) be a sequence of \( k \times 1 \) vectors and \( \{\varphi_h\}_{h \in \mathbb{Z}} \) a sequence of scalars. Let

\[
y_{nj} = \sum_{t=1}^n x_t \varphi_{t-j} , \quad n \in \mathbb{N}, \quad j \in \mathbb{Z}
\]

If \( \sum |\varphi_h| < \infty \) then

\[
\sum_{j} y_{nj} y_{nj}^* \leq c \sum_{t=1}^n x_t x_t^* , \quad n \in \mathbb{N}
\]

for some constant \( c \) not depending on \( n \).

Proof. Set

\[
\varphi(\lambda) = \sum_{j} \varphi_j e^{i j \lambda} , \quad x_n(\lambda) = \sum_{t=1}^n x_t e^{i t \lambda} , \quad y_n(\lambda) = \sum_{j} y_{nj} e^{i j \lambda}
\]
then \( Y_n(\lambda) = X_n(\lambda) \phi(-\lambda) \). With

\[ c = \sup_{|\lambda| < \pi} |\phi(1-\lambda)|^2 \]

this gives

\[
\sum_{j \neq n} Y_j Y_n^* = \frac{1}{2\pi} \int_{-\pi}^{\pi} Y_n(\lambda) Y_n^*(\lambda) d\lambda = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(-\lambda)^2 X_n(\lambda) X_n^*(\lambda) d\lambda \leq c \frac{1}{2\pi} \int_{-\pi}^{\pi} X_n(\lambda) X_n^*(\lambda) d\lambda = c \sum_{i=0}^{n} x_i x_i^* \]

Lemma 3.4. Let \((x_t)_{t \in \mathbb{Z}}\) be a sequence of \(k \times 1\) vectors with \(x_t = 0\) if \(t = -p\) for some \(p \geq 0\), and let \((\phi_h)_{h \geq 0}\) be a sequence of scalars. Let

\[
y_t = \sum_{h=0}^{\infty} \phi_h x_{t-h}, \quad t \in \mathbb{Z}.
\]

If \(\phi(z) = \sum_{h=0}^{\infty} \phi_h z^h\) has convergence radius larger than 1 and \(\phi(z) \neq 0\) if \(|z| < 1\), then there exist constants \(c_1, c_2 > 0\), only depending on \((\phi_h)\), such that

\[
c_1 \sum_{t=1-p}^{n} x_t x_t^* \leq \sum_{t=1-p}^{n} y_t y_t^* \leq c_2 \sum_{t=1-p}^{n} x_t x_t^*, \quad n \in \mathbb{N}.
\]

Proof. Let \(A(z) = 1/\phi(z) = \sum_{h=0}^{\infty} \alpha_h z^h\), then \(\sum_{h=0}^{\infty} \alpha_h \phi_{h-m} = 5_0 h\) for \(h = 0,1,\ldots\).

Note that

\[
\sum_{h=0}^{\infty} |\phi_h| < \infty, \quad \sum_{h=0}^{\infty} |\alpha_h| < \infty.
\]

Let \(L_n\) be the \((n+p) \times (n+p)\) lag matrix with unit elements just below the diagonal and zero elements elsewhere. Then \(L_n^m = [x_{1-p}, \ldots, x_n]^t = [x_{1-p-m}, \ldots, x_{n-m}]^t\) with the convention \(L_0^n = I\). Note that \(L_n^m = 0\) for \(m \geq n + p\). Introduce

\[
\phi_n = \sum_{m=0}^{n+p-1} \phi_m L_n^m = \sum_{m=0}^{\infty} \phi_m L_n^m
\]

then
Lemma 3.5. Let \((x_t)_{t \in \mathbb{N}}\) be a nonstochastic sequence of \(k \times 1\) vectors.

a) Let \(S_n := \sum_{t=1}^{n} x_t x_t'\) with \(S_n > 0\) for some \(n\). We have

\[
\max_{1 \leq t \leq n} x_t S_n^{-1} x_t = S_n^{-1} \rightarrow 0, \quad x_t S_n^{-1} x_t \rightarrow 0
\]

b) Let \(Z_n\) be a sequence of nonstochastic \(k \times k\) matrices with \(0 \preceq Z_1 \preceq Z_2 \preceq \ldots\)

and \(Z_n > 0\) for some \(n\). We have

\[
Z_n^{-1} \rightarrow 0, \quad x_t Z_n^{-1} x_t \rightarrow 0 = \max_{1 \leq t \leq n} x_t Z_n^{-1} x_t \rightarrow 0
\]

Proof. a) Take some fixed \(c \in \mathbb{R}_k\). Let \(N\) be such that \(S_N > 0\). Then \(c = \sum_{t=1}^{N} \alpha_t x_t\)

for some \(\alpha_1, \ldots, \alpha_N\) not depending on \(n\). Hence, for \(n \geq N\):

\[
c' S_n^{-1} c \leq \sum_{t=1}^{N} |\alpha_t| |x_t S_n^{-1} x_t| \leq (\sum_{t=1}^{N} |\alpha_t|)^2 \max_{1 \leq t \leq n} x_t S_n^{-1} x_t \rightarrow 0
\]

This holds for any \(c\), implying \(S_n^{-1} \rightarrow 0\).

b) For given \(n\) let \(\tau_n\) be the largest index for which the maximum is attained. If \((\tau_n)\) is bounded then \((x_{\tau_n})\) is bounded. So,

\[
x_{\tau_n} Z_n^{-1} x_{\tau_n} \leq |x_{\tau_n}|^2 \cdot \|Z_n^{-1}\| \rightarrow 0.
\]
If \((\tau_n)\) is unbounded then \(\tau_n \to \infty\). Since \(\tau_n < n\) we have \(Z_{\tau_n}^{-1} > Z_n^{-1}\). So,
\[
x_{\tau_n} Z_{\tau_n}^{-1} x_{\tau_n} < x_\tau Z_{\tau_n}^{-1} x_\tau \to 0.
\]

**Proof of theorem 2.2.** The condition (2.2) is fulfilled. So we have only to prove that (2.5) - (2.7) imply (2.3) - (2.4).

From (2.6) and \(s_n < n\) we get

\[
(3.15) \quad \|Z_n^{-1}\| = 0(s_n^{-1})
\]

In particular \(Z_n^{-1} \to 0\). With (2.7) and lemma 3.5b this leads to

\[
(3.16) \quad \max_{1 \leq t \leq n} |Z_n^{-1} v_t| \to 0
\]

We derive a suitable expression for \(x_t\). Set \(y_t = 0\) if \(t \leq -p\), \(w_t = 0\) if \(t \leq 0\). Define \(e_t\) for \(t \leq 0\) by the relation (1.2). Then this relation holds for all \(t \in Z\). In particular \(e_t = 0\) of \(t \leq -p\). Let \(\varphi(z) = 1/A(z) = \sum_{m=0}^{\infty} \varphi_m z^m\).

Then \(\varphi(z)\) has convergence radius larger than 1 and \(\varphi(z) \neq 0\) if \(|z| < 1\). In particular \(\sum_{m=0}^{\infty} |\varphi_m| < \infty\). The solution of (1.2) can be written as

\[
y_t = \sum_{h=0}^{\infty} \varphi_h (y_{t-h} + e_{t-h}), \quad t \in Z.
\]

Furthermore, since \(\sum \varphi_j \alpha_{j-k} = \delta_{0k}\), we have

\[
w_t = \sum_{j=0}^{\infty} k=0 \varphi_k \alpha_{j-k} w_{t-j} = \sum_{j=0}^{\infty} k=0 \varphi_k \alpha_{j-k} w_{t-j} = \sum_{k=0}^{\infty} p j=0 \varphi_j \alpha_{j-k} w_{t-j-k}
\]

Therefore,

\[
x'_t = (w'_t, y_{t-1}, \ldots, y_{t-p}) = \sum_{k=0}^{\infty} \varphi_k [(\sum_{j=0}^{\infty} \alpha_{j-k} w_{t-k-j}, y_{t-k-1}, \ldots, y_{t-k-p}) + (0, e_{t-k-1}, \ldots, e_{t-k-p})].
\]

This gives
(3.17) \[ x_t = \mu_t + \xi_t = \sum_{k=0}^{\infty} \varphi_k (v_{t-k} + \eta_{t-k}), \quad t \in \mathcal{I}, \]

where \( \eta_t' = (0, \epsilon_{t-1}, \ldots, \epsilon_{t-p}) \) and

\[ \mu_t = \sum_{k=0}^{\infty} \varphi_k v_{t-k}, \quad \xi_t = \sum_{k=0}^{\infty} \varphi_k \eta_{t-k}, \quad t \in \mathcal{I}. \]

With (3.17) we can derive a lower bound for \( \bar{S}_n \). From lemma 3.4 it follows that

\[ \sum_{t=1}^{n} x_t x_t' \geq c_1 \sum_{p+1}^{n} (v_t + \eta_t)(v_t + \eta_t)' \]

Since

\[ \sum_{t=1}^{n} E\{x_t x_t'\} = \text{diag} (0, \sigma_{t-1}^2, \ldots, \sigma_{t-p}^2) \geq (n-p) I_0 \]

we get, by taking expectations,

\[ 0 \geq \sum_{t=1}^{n} E\{x_t x_t'\} + \bar{S}_n - c'.Z_n \]

for some \( c' > 0 \). Since \( Z_n^{-1} \to 0 \) this implies

(3.18) \[ \bar{S}_n \geq cZ_n \]

for some \( c > 0 \).

With these preparations we can prove (2.3), (2.4). First we prove (2.4). From (3.17) and (3.18) we get

\[ (c \max_{t} x_t' \bar{S}_n^{-1} x_t)^{\frac{1}{2}} \leq \left\{ \max_{t} x_t' \bar{S}_n^{-1} x_t \right\}^{\frac{1}{2}} = \max_{t} \left| Z_n^{-1} x_t \right| \]

\[ \leq \max_{t} \left| Z_n^{-1} \mu_t \right| + \max_{t} \left| Z_n^{-1} \xi_t \right| \]

\[ \leq \sum_{k=0}^{\infty} |\varphi_k| \left( \max_{t} \left| Z_n^{-1} v_t \right| + \max_{t} \left| Z_n^{-1} \eta_t \right| \right). \]

According to (3.16) the first term on the right-hand side tends to 0, and with (3.15) and (3.7) we see that the second term on the right-hand side tends to 0 in probability. Together this proves (2.4).

Finally we prove (2.3). From (3.15) and (3.17) we get
for some constants $c', c''$. So it suffices to prove that

\[(3.19) \quad \sum_{n}^{\infty} (\sum_{t}^{n} \lambda_{t+i} \xi_{t-j} - \sum_{t}^{n} \lambda_{t-i} \xi_{t-j}) \to 0, \quad i, j \in \mathcal{Z} \]

\[(3.20) \quad \sum_{n}^{\infty} (\sum_{t}^{n} \lambda_{t+i} \xi_{t-j} - \sum_{t}^{n} \lambda_{t-i} \xi_{t-j}) \to 0, \quad i, j \in \mathcal{Z}, \]

where

\[\xi_{t} = \sum_{h=0}^{\infty} \phi_{h} \epsilon_{t-h} \]

The relation (3.19) follows from lemma 3.2 and the fact that $(\lambda_{t})$ is bounded. For the proof of (3.20) we write

\[(3.21) \quad \sum_{1}^{n} \lambda_{t+i} \epsilon_{t-j} = \sum_{1}^{n} \lambda_{t+i} (\sum_{h=0}^{\infty} \phi_{h} \epsilon_{t-h-i}) = \sum_{1}^{n} a_{nj} \epsilon_{j-i} \]

where

\[a_{nj} = \sum_{1}^{n} \lambda_{t+i} \phi_{t-j} \]

From (3.16) we get for some $c > 0$ that

\[|Z_{n}^{-b} a_{nj}| \leq \sum_{h=0}^{\infty} |\phi_{h}| \cdot \max |Z_{n}^{-b} \mu_{t}| \leq c \cdot \max |Z_{n}^{-b} v_{t}| \to 0 \]

and so
Since \((\lambda_t)\) is bounded and \(Z_n^{-1} \to 0\) we get from lemma 3.3 and lemma 3.4 for some \(c, c' > 0\) that

\[
\|Z_n^{-\frac{1}{2}} (\sum_{j \leq i} a_{nj} e_{j-i}) Z_n^{-\frac{1}{2}}\| \leq c' \|Z_n^{-\frac{1}{2}} (\sum_{t \leq i} \lambda_t \nu_t' \nu_t') Z_n^{-\frac{1}{2}}\|
\]

\[
\leq c \|Z_n^{-\frac{1}{2}} (\sum_{t \leq i} \nu_t' \nu_t') Z_n^{-\frac{1}{2}}\| \leq c,
\]

where the last inequality follows from

\[
\sum_{t \leq i} \nu_t' \nu_t \leq Z_n.
\]

This gives

\[
(3.23) \quad s_n^{-\frac{1}{2}} Z_n^{-\frac{1}{2}} \sum_{j \leq i} a_{nj} e_{j-i} \to 0.
\]

Combining (3.21) - (3.23) yields (3.20), which completes the proof.

Proof of theorem 2.3. We restrict ourselves to the general case \(r = \min(s, p+q)\). (The proof for other special cases follows along similar lines.) We write

\[
\Phi_D = \begin{bmatrix} P_n & Q_n \end{bmatrix} (r), \quad \Phi = \begin{bmatrix} \phi_{00} & \phi_{01} \end{bmatrix} (r)
\]

\[
\begin{bmatrix} Q_n' & R_n \end{bmatrix} (p+q-r), \quad \phi = \begin{bmatrix} \phi_{10} & \phi_{11} \end{bmatrix} (p+q-r)
\]

\[
(3.23) \quad s_n^{-\frac{1}{2}} Z_n^{-\frac{1}{2}} \sum_{j \leq i} a_{nj} e_{j-i} \to 0.
\]
At first we derive some properties of $P_n$ and $Q_n$. We have

$$P_n = \phi_{00}D_n\phi_{00} + \phi_{01}D_n\phi_{01} = \phi_{00}D_n\phi_{00},$$

and so, with (2.8), (2.9) it follows that

(3.24) \[ \|P_n^{-1}\| = O(1/n). \]

Furthermore,

$$Q_n = \phi_{00}D_n\phi_{00} + \phi_{01}D_n\phi_{01},$$

and with (2.8) this gives

$$P_n^{-1}Q_n = (\phi_{00}D_n\phi_{00})^{-1}Q_n = (\phi_{00}^{-1})'(\phi_{00}^{-1} + D_n\phi_{00}^{-1} + D_n\phi_{01}^{-1}).$$

Using (2.10) this gives

(3.25) \[ \|P_n^{-1}Q_n \| = O(1). \]

Next we derive a lower bound for $Z_n^{-1}$. We have for some constants $c, c' > 0$:

$$Z_n = \sum_{p=1}^{n} v^t v + nI_0 = \sum_{p=1}^{n} \phi_{00}^2 (\sum \tilde{q}_t(n)\tilde{q}'_t(n))D_n\phi' + nI_0$$

or

(3.26) \[ Z_n \geq c'A_n \]

where

$$A_n = \begin{bmatrix} P_n & Q_n \\ Q_n' & R_n + nI \end{bmatrix}$$

Note that
where \( T_n = P_n^{-1} Q_n \), \( \Delta_n = nI + R_n - Q_n P_n^{-1} Q_n \). Since \( \pmb{D}' \neq 0 \) we have

\[
R_n - Q_n P_n^{-1} Q_n \geq 0 \quad \text{and so } \|\Delta_n^{-1}\| = O(1/n). \]

Then from (3.24), (3.25) it follows that \( \|A_n^{-1}\| = O(1/n) \) and with (3.26) this proves (2.6).

Finally, with (3.26) we get

\[
v' Z_n^{-1} v_n = \tilde{q}'(n) \cdot D_n^{-1} \hat{\phi}'(n) \phi D_n^{-1} \hat{\phi}'(n)
\]

\[
\leq \frac{1}{c} |q_n(n)|^2 \cdot \|D_n^{-1} \hat{\phi}' A_n^{-1} \phi D_n^{-1} \hat{\phi}'\|
\]

\[
\leq \frac{1}{c} |q_n(n)|^2 \cdot \text{tr}(A_n^{-1} \phi D_n \phi').
\]

Therefore (2.7) follows from (2.12) if we can prove that \( \text{tr}(A_n^{-1} \phi D_n \phi') = O(1) \).

However, this follows from

\[
\text{tr}(A_n^{-1} \phi D_n \phi') = \text{tr}\{(P_n^{-1} + T_n \Delta_n^{-1} T_n') P_n - T_n \Delta_n^{-1} Q_n\} +
\]

\[
+ \text{tr}\{-\Delta_n^{-1} T_n' Q_n + \Delta_n^{-1} R_n\} = \text{tr}\{I + \Delta_n^{-1} (R_n - Q_n P_n^{-1} Q_n)\} =
\]

\[
p + \text{tr}(\Delta_n^{-1} (R_n - Q_n P_n^{-1} Q_n) \Delta_n^{-1}) \leq 2p + q - r,
\]

completing the proof.
4. Weak consistency

We make some final remarks about weak consistency of the LS-estimator in the autoregressive model (1.2).

The conditions of theorem 2.2 imply $b_n \xrightarrow{P} \beta$. This follows from a short inspection of the proofs of theorems 2.1 and 2.2. The condition (2.6) implies $Z_n^{-1} \rightarrow 0$ and therefore also $C_n \rightarrow 0$. Then the convergence in distribution of $C_n^{-1}(b_n - \beta)$ gives $b_n \xrightarrow{P} \beta$.

In fact, for weak consistency the conditions of theorem 2.2 can be weakened considerably because we need only to prove that $S_n^{-1} S_n \xrightarrow{P} I$ and $C_n \rightarrow 0$.

This can be done under the stability assumption (2.5), the assumption that the $\varepsilon_t$ satisfy the LF condition instead of the Eicker conditions (see (3.6)), and the assumption $\|Z_n^{-1}\| = o(s^{-1})$, where in the definition of $Z_n$ the term $nI_0$ is replaced by $s I_0$. The behaviour of $Z_n$ can be investigated as in theorem 2.3. For other conditions about weak consistency, see e.g. Drygas [2], remark 4.5, and Eicker [3], Satz 1.
References


IN 1984 REEDS VERSCHENEN

138 G.J. Cuypers, J.P.C. Kleijnen en J.W.M. van Rooyen
Testing the Mean of an Asymmetric Population:
Four Procedures Evaluated

139 T. Wansbeek en A. Kapteyn
Estimation in a linear model with serially correlated errors when
observations are missing

140 A. Kapteyn, S. van de Geer, H. van de Stadt, T. Wansbeek
Interdependent preferences: an econometric analysis

141 W.J.H. van Groenendaal
Discrete and continuous univariate modelling

142 J.P.C. Kleijnen, P. Cremers, F. van Belle
The power of weighted and ordinary least squares with estimated
unequal variances in experimental design

143 J.P.C. Kleijnen
Superefficient estimation of power functions in simulation
experiments

144 P.A. Bekker, D.S.G. Pollock
Identification of linear stochastic models with covariance
restrictions.

145 Max D. Merbis, Aart J. de Zeeuw
From structural form to state-space form

146 T.M. Doup and A.J.J. Talman
A new variable dimension simplicial algorithm to find equilibria on
the product space of unit simplices.

147 G. van der Laan, A.J.J. Talman and L. Van der Heyden
Variable dimension algorithms for improper labellings.

148 G.J.C.Th. van Schijndel
Dynamic firm behaviour and financial leverage clienteles

149 M. Plattel, J. Peil
The ethico-political and theoretical reconstruction of contemporary
economic doctrines

150 F.J.A.M. Hoes, C.W. Vroom
an exercise in sociological demystification

151 T.M. Doup, G. van der Laan and A.J.J. Talman
The \((2^{n+1}-2)\)-ray algorithm: a new simplicial algorithm to compute
economic equilibria
IN 1984 REEDS VERSCHIJNEN (vervolg)

152 A.L. Hempenius, P.G.H. Mulder
Total Mortality Analysis of the Rotterdam Sample of the Kaunas-Rotterdam Intervention Study (KRIS)

153 A. Kapteyn, P. Kooreman
A disaggregated analysis of the allocation of time within the household.

154 T. Wansbeek, A. Kapteyn
Statistically and Computationally Efficient Estimation of the Gravity Model.

155 P.F.P.M. Nederstigt
Over de kosten per ziekenhuisopname en levensduurmodellen

156 B.R. Meijboom
An input-output like corporate model including multiple technologies and make-or-buy decisions

157 P. Kooreman, A. Kapteyn
Estimation of Rationed and Unrationed Household Labor Supply Functions Using Flexible Functional Forms

158 R. Heuts, J. van Lieshout
An implementation of an inventory model with stochastic lead time

159 P.A. Bekker
Comment on: Identification in the Linear Errors in Variables Model

160 P. Meys
Functies en vormen van de burgerlijke staat
Over parlementarisme, corporatisme en autoritair etatisme

161 J.P.C. Kleijnen, H.M.M.T. Denis, R.M.G. Kerckhoffs
Efficient estimation of power functions

162 H.L. Theuns
The emergence of research on third world tourism: 1945 to 1970; An introductory essay cum bibliography

163 F. Boekema, L. Verhoef
De "Grijze" sector zwart op wit
Werklozenprojecten en ondersteunende instanties in Nederland in kaart gebracht

164 G. van der Laan, A.J.J. Talman, L. Van der Heyden
Shortest paths for simplicial algorithms

165 J.H.F. Schilderinck
Interregional structure of the European Community
IN (1984) REEDS VERSCHENEN (vervolg)

166 P.J.F.G. Meulendijks
   An exercise in welfare economics (I)

167 L. Elsner, M.H.C. Paardekooper
   On measures of nonnormality of matrices.
IN 1985 REEDS VERSCHENEN

168  T.M. Doup, A.J.J. Talman
     A continuous deformation algorithm on the product space of unit simplices

169  P.A. Bekker
     A note on the identification of restricted factor loading matrices

170  J.H.M. Donders, A.M. van Nunen
     Economische politiek in een twee-sectoren-model

171  L.H.M. Bosch, W.A.M. de Lange
     Shift work in health care