Tilburg University

The asymptotic behaviour of the estimated generalized least squares method in the linear regression model

van der Genugten, B.B.

Publication date:
1982

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

van der Genugten, B. B. (1982). *The asymptotic behaviour of the estimated generalized least squares method in the linear regression model*. (Research memorandum / Tilburg University, Department of Economics; Vol. FEW 111). Unknown Publisher.

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.

CBM
R

7626
1982
111

FEW

subfaculteit der econometrie

RESEARCH MEMORANDUM

Bestemming	TIJDSCRIFTENBUREAU BIBLIOTHEEK KATHOLIEKE HOGESCHOOL TILBURG	Nr.
---	--	--

TILBURG UNIVERSITY
DEPARTMENT OF ECONOMICS

Postbus 90153 - 5000 LE Tilburg
Netherlands

The asymptotic behaviour of the estimated
generalized least squares method in the
linear regression model

by

B.B. van der Genugten

Summary

In the linear regression model the generalized least squares (GLS) method is only applicable if the covariance matrix of the errors is known but for a scalar factor. Otherwise an estimator for this matrix has to be used. Then we speak of the estimated generalized least squares (EGLS) method. In this paper the asymptotic behaviour of both methods is compared. Results are applied to some standard models commonly used in econometrics.

§ 1. Introduction

We consider the linear regression model¹⁾

$$(1.1) \quad \underline{y} = \underline{X}_n \underline{\beta} + \underline{\varepsilon},$$

where $\underline{y}$ is an (observable) random $n \times 1$ -vector, $\underline{X}_n$ an (observable) non-random $n \times k$ -matrix, $\underline{\beta}$ an unknown non-random $k \times 1$ -vector and $\underline{\varepsilon}$ a (non-observable) random $n \times 1$ -vector. We suppose

$$(1.2) \quad E\{\underline{\varepsilon}\} = 0, \text{Cov}\{\underline{\varepsilon}\} = \sigma^2 \underline{Q}_n = \underline{\Omega}_n,$$

where $\sigma^2 > 0$ and $\underline{Q}_n > 0$ (strictly positive definite).

The generalized least squares (GLS) method leads to the GLS-estimator $\tilde{\underline{b}}_n$ of $\underline{\beta}$, defined by

$$\tilde{\underline{b}}_n = \tilde{\underline{S}}_n^{-1} \underline{X}_n' \underline{Q}_n^{-1} \underline{y}, \quad \tilde{\underline{S}}_n = \underline{X}_n' \underline{Q}_n^{-1} \underline{X}_n,$$

(where $\underline{A}^{-}$ denotes a pseudo-inverse of $\underline{A}$ i.e. $\underline{A}\underline{A}^{-}\underline{A} = \underline{A}$), and the GLS-residual $\tilde{\underline{s}}_n^2$, defined by

$$\tilde{\underline{s}}_n^2 = \tilde{\underline{e}}' \underline{Q}_n^{-1} \tilde{\underline{e}}/n, \quad \tilde{\underline{e}} = \underline{y} - \underline{X}_n \tilde{\underline{b}}_n.$$

Often the matrix $\underline{Q}_n$ is not fully known and is estimated by some random $n \times n$ -matrix $\hat{\underline{Q}}_n$. Then we speak of the estimated generalized least squares (EGLS) method (see e.g. Judge (1980), p. 113). This leads to the EGLS-estimator $\hat{\underline{b}}_n$ of $\underline{\beta}$, defined by

$$\hat{\underline{b}}_n = \hat{\underline{S}}_n^{-1} \underline{X}_n' \hat{\underline{Q}}_n^{-1} \underline{y}, \quad \hat{\underline{S}}_n = \underline{X}_n' \hat{\underline{Q}}_n^{-1} \underline{X}_n,$$

and the EGLS-residual $\hat{\underline{s}}_n^2$, defined by

1) Random variables are underlined.

$$\underline{s}_n^2 = \underline{\hat{e}}' \underline{\hat{Q}}_n^{-1} \underline{\hat{e}}/n, \underline{\hat{e}} = \underline{y} - \underline{x}_n \underline{\hat{b}}_n$$

(where $\underline{\hat{Q}}_n^{-1}$ is defined arbitrarily for singular $\underline{\hat{Q}}_n$).

In theorem 2.1 below we give conditions under which the asymptotic behaviour of $\underline{\hat{b}}_n$, $\underline{\hat{s}}_n^2$ is the same as that of $\underline{\tilde{b}}_n$, $\underline{\tilde{s}}_n^2$. The theorem 2.2 thereafter takes in account the fact that in general for approximations based on the normal distribution the covariance matrix of $\underline{\hat{b}}_n$ is also unknown and has to be estimated.

In many cases the choice of the matrix $\underline{\hat{Q}}_n$ is based on the ordinary least squares (OLS) method. The OLS-estimator $\underline{b}_n$ of β is defined by

$$\underline{b}_n = \underline{S}_n^{-1} \underline{x}_n' \underline{y}, \underline{S}_n = \underline{x}_n' \underline{x}_n$$

and the OLS-residual $\underline{s}_n^2$ by

$$\underline{s}_n^2 = |\underline{e}|^2/n, \underline{e} = \underline{y} - \underline{x}_n \underline{b}_n.$$

In order to facilitate the verification of the conditions of theorem 2.1 with respect to $\underline{\hat{Q}}_n$ we add theorem 2.3 below, which describes the behaviour of quadratic forms in $\underline{e}$.

In section 3 we give the proofs of these theorems. In section 4-6 the obtained results are applied to some standard models commonly used on econometrics.

The general approach as sketched above can also be found in Theil (1971), Chapter 8, or Schmidt (1976), p. 64-72. However, the conditions in this paper are much weaker, and also more convenient to deal with, especially in the specific standard models mentioned above.

§ 2. Statement of the results

For sake of simplicity we restrict ourselves to the non-collinear case that $S_n > 0$ for all n sufficiently large (although for some special results connected with residuals this condition is not needed at all).

In the following we write $\|A\| = \|A\|_2 = \lambda_{\max}^{1/2}(A'A)$. In particular, $\|A\| = \lambda_{\max}(A)$ if $A \geq 0$ and $\|A^{-1}\| = \lambda_{\min}^{-1}(A)$ if $A > 0$.

For theorem 2.1 we formulate the following conditions:

$$(2.1) \quad \|\tilde{S}_n^{-1}\| \rightarrow 0,$$

$$(2.2) \quad \left\| Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I_n \right\| \xrightarrow{P} 0,$$

$$(2.3) \quad \|\tilde{S}_n^{-1}\| = O(n^{-1}),$$

$$(2.4) \quad \tilde{S}_n^{-1/2} X_n' (\hat{Q}_n^{-1} - Q_n^{-1}) \underline{\varepsilon} \xrightarrow{P} 0.$$

We have

Theorem 2.1

$$a) \quad \tilde{S}_n^2 - \underline{\varepsilon}' Q_n^{-1} \underline{\varepsilon} / n \xrightarrow{P} 0,$$

$$b) \quad (2.1) \Rightarrow \tilde{b}_n \xrightarrow{P} \beta,$$

$$c) \quad (2.2) \Rightarrow \hat{S}_n^2 - \tilde{S}_n^2 \xrightarrow{P} 0,$$

$$d) \quad (2.2), (2.3) \Rightarrow \hat{b}_n - \tilde{b}_n \xrightarrow{P} 0,$$

$$e) \quad (2.2), (2.4) \Rightarrow \tilde{S}_n^{1/2} (\hat{b}_n - \tilde{b}_n) \xrightarrow{P} 0,$$

$$f) \quad (2.1), (2.2), (2.4) \Rightarrow \hat{b}_n - \tilde{b}_n \xrightarrow{P} 0.$$

Remark. From the proof of c) we get also

$$(2.2) \Rightarrow \tilde{S}_n^{-1/2} \hat{S}_n \tilde{S}_n^{-1/2} \xrightarrow{P} I.$$

The theorem 2.1 can be applied in the following way. For many models we have

$$(2.5) \quad \underline{\varepsilon}' \Omega_n^{-1} \underline{\varepsilon} / n \xrightarrow{P} 1.$$

Then a) gives that $\underline{\hat{S}}_n^2 \xrightarrow{P} \sigma^2$, and c) gives conditions under which $\underline{\hat{S}}_n^2 \xrightarrow{P} \sigma^2$. The combination of b) with d) or f) gives conditions under which $\underline{\hat{b}}_n \rightarrow \beta$. In many cases $\underline{\tilde{b}}_n$ is even asymptotically normal i.e.

$$(2.6) \quad \underline{\tilde{S}}_n^{1/2} (\underline{\tilde{b}}_n - \beta) \xrightarrow{L} N_k(0, \sigma^2 I).$$

Then e) gives conditions under which the limit distributions of $\underline{\hat{b}}_n$ and $\underline{\tilde{b}}_n$ are the same, i.e.

$$\underline{\tilde{S}}_n^{1/2} (\underline{\hat{b}}_n - \beta) \xrightarrow{L} N_k(0, \sigma^2 I)$$

or, equivalently,

$$\underline{D}_n^{-1} (\underline{\hat{b}}_n - \beta) \xrightarrow{L} N_k(0, \sigma^2 I)$$

for any non-random $k \times k$ -matrix $\underline{D}_n$ with $\underline{D}_n \underline{D}_n' = \underline{\tilde{S}}_n^{-1}$.

From a practical point of view this result is insufficient because generally $\underline{\tilde{S}}_n$ will be unknown. To justify approximations with the normal distribution we must show that $\underline{D}_n$ can be replaced by some random $k \times k$ -matrix $\underline{\hat{D}}_n$ such that $\underline{\hat{D}}_n \underline{\hat{D}}_n' = \underline{\hat{S}}_n^{-1}$. The theorem 2.2 below shows that without further conditions such a decomposition of $\underline{\hat{S}}_n^{-1}$ can be found, although the matrix $\underline{\hat{D}}_n$ is not equal to the most obvious choice $\underline{\hat{S}}_n^{-1/2}$. However, in practice the specific $\underline{\hat{D}}_n$ is never needed. It is only its existence that matters.

Theorem 2.2

$$(2.2), (2.4), (2.6) \Rightarrow \underline{\hat{D}}_n^{-1} (\underline{\hat{b}}_n - \beta) \xrightarrow{L} N_k(0, \sigma^2 I),$$

where $\underline{\hat{D}}_n = \underline{\hat{S}}_n^{-1/2} (\underline{\tilde{S}}_n^{1/2} \underline{\hat{S}}_n^{-1} \underline{\tilde{S}}_n^{1/2})^{1/2}$

For the verification of the conditions of the foregoing theorems when using the OLS-method for estimating Q , the theorem 2.3 below is useful.

Theorem 2.3. Let

$$(2.7) \quad \|Q_n\| = o(n), \quad \text{tr}(Q_n) = O(n),$$

then for any non-random sequence of $n \times n$ -matrices A_n such that $\|A_n\|$ is bounded we have

$$\frac{1}{n} (\underline{e}' A_n \underline{e} - \underline{\varepsilon}' A_n \underline{\varepsilon}) \xrightarrow{P} 0$$

Corollary. Under (2.7) we get

$$\frac{s_n^2}{n} - |\underline{\varepsilon}|^2/n \xrightarrow{P} 0$$

§ 3. Proofs of the theorems

In the following we say that a sequence of random variables $\{\underline{x}_n\}$ is P-bounded if for every $\varepsilon > 0$ there exists a M such that $P\{|\underline{x}_n| \geq M\} < \varepsilon$ for all n . If $\{\underline{x}_n\}$ and $\{\underline{y}_n\}$ are sequences of random variables such that $\{\underline{x}_n\}$ is P-bounded and $\underline{y}_n \xrightarrow{P} 0$ then $\underline{x}_n \underline{y}_n \xrightarrow{P} 0$. If $E|\underline{x}_n|^\alpha$ is bounded in n for some $\alpha > 0$ then $\{\underline{x}_n\}$ is P-bounded.

Generalizations for random vectors are immediate. We will use these results belows.

Proof of theorem 2.1

a) Let $M_n = X_n' \tilde{S}_n^{-1} X_n Q_n^{-1}$ then $\tilde{\underline{\varepsilon}} = (I_n - M_n)\underline{\varepsilon}$. We have

$$\begin{aligned} \underline{\varepsilon}' Q_n^{-1} \underline{\varepsilon} - n \tilde{S}_n^{-2} &= \underline{\varepsilon}' Q_n^{-1} \underline{\varepsilon} - \tilde{\underline{\varepsilon}}' Q_n^{-1} \tilde{\underline{\varepsilon}} = \\ &= (M_n \underline{\varepsilon})' Q_n^{-1} (M_n \underline{\varepsilon}) = \underline{\varepsilon}' Q_n^{-1} M_n \underline{\varepsilon} \geq 0, \end{aligned}$$

and $E[\underline{\varepsilon}' Q_n^{-1} M_n \underline{\varepsilon}] = \sigma^2 \text{tr}(Q_n^{-1} M_n Q_n) = k\sigma^2$. So, $\underline{\varepsilon}' Q_n^{-1} \underline{\varepsilon} - n \tilde{S}_n^{-2}$ is non-negative and P-bounded. This implies the result of a).

b) $E\{\tilde{\underline{b}}_n\} = \beta$, $\text{Cov}\{\tilde{\underline{b}}_n\} = \sigma^2 \tilde{S}_n^{-1}$ implying $\tilde{\underline{b}}_n \xrightarrow{P} \beta$ and therefore also $\tilde{\underline{b}}_n \xrightarrow{P} \beta$.

c), d) and e) The condition (2.2) implies that $P\{\hat{Q}_n > 0\} \rightarrow 1$ if $n \rightarrow \infty$. So in the following we may assume without loss of generality that $\hat{Q}_n$ is non-singular. Since

$$(\tilde{S}_n^{-1/2} X_n' Q_n^{-1/2}) (\tilde{S}_n^{-1/2} X_n' Q_n^{-1/2})' = I$$

we have $\|\tilde{S}_n^{-1/2} X_n' Q_n^{-1/2}\| = 1$. With (2.2) this gives

$$\begin{aligned} \|\tilde{S}_n^{-1/2} \tilde{S}_n \tilde{S}_n^{-1/2} - I\| &= \|\tilde{S}_n^{-1/2} X_n' Q_n^{-1/2} (Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I_n) Q_n^{-1/2} X_n \tilde{S}_n^{-1/2}\| \\ &\leq \|Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I_n\| \xrightarrow{P} 0, \end{aligned}$$

proving the remark of theorem 2.2. This implies that $P\{\hat{S}_n^{-1} > 0\} \rightarrow 1$ if

$n \rightarrow \infty$. Since k does not depend on n this yields

$$\|\tilde{S}_n^{-1/2} \hat{S}_n^{-1} \tilde{S}_n^{-1/2} - I\| \xrightarrow{P} 0. \text{ From}$$

$$E|\tilde{S}_n^{-1/2} X_n' Q_n^{-1} \underline{\varepsilon}|^2 = \text{tr}[E\{\tilde{S}_n^{-1/2} X_n' Q_n^{-1} \underline{\varepsilon} \underline{\varepsilon}' Q_n^{-1} X_n \tilde{S}_n^{-1/2}\}] = k\sigma^2,$$

$$E|Q_n^{-1/2} \tilde{\underline{\varepsilon}}|^2 = E\{\tilde{\underline{\varepsilon}}' Q_n^{-1} \tilde{\underline{\varepsilon}}\} \leq E\{\underline{\varepsilon}' Q_n^{-1} \underline{\varepsilon}\} = n\sigma^2$$

we see that $\tilde{S}_n^{-1/2} X_n' Q_n^{-1} \underline{\varepsilon}$, $Q_n^{-1/2} \tilde{\underline{\varepsilon}}/\sqrt{n}$ and $Q_n^{-1/2} \underline{\varepsilon}/\sqrt{n}$ are P -bounded. With (2.2) this gives

$$\frac{1}{\sqrt{n}} \tilde{S}_n^{-1/2} X_n' (\hat{Q}_n^{-1} - Q_n^{-1}) \underline{\varepsilon} = \tilde{S}_n^{-1/2} X_n' Q_n^{-1/2} (Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I) \cdot \frac{1}{\sqrt{n}} Q_n^{-1/2} \underline{\varepsilon} \xrightarrow{P} 0.$$

From the definition of $\tilde{\underline{b}}_n$ and $\hat{\underline{b}}_n$ we get

$$\begin{aligned} \tilde{S}_n^{1/2} (\hat{\underline{b}}_n - \tilde{\underline{b}}_n) &= \tilde{S}_n^{1/2} \{(\hat{\underline{b}}_n - \beta) - (\tilde{\underline{b}}_n - \beta)\} = \tilde{S}_n^{1/2} (\hat{S}_n^{-1} X_n' \hat{Q}_n^{-1} - \tilde{S}_n^{-1} X_n' Q_n^{-1}) \underline{\varepsilon} = \\ &= (\tilde{S}_n^{1/2} \hat{S}_n^{-1} \tilde{S}_n^{-1/2} - I) \cdot \tilde{S}_n^{-1/2} X_n' Q_n^{-1} \underline{\varepsilon} + \tilde{S}_n^{1/2} \hat{S}_n^{-1/2} \tilde{S}_n^{-1/2} X_n' (\hat{Q}_n^{-1} - Q_n^{-1}) \underline{\varepsilon}. \end{aligned}$$

Therefore, under the condition (2.4) we get that $\tilde{S}_n^{1/2} (\hat{\underline{b}}_n - \tilde{\underline{b}}_n) \xrightarrow{P} 0$ proving e), and without this condition that $\tilde{S}_n^{1/2} (\hat{\underline{b}}_n - \tilde{\underline{b}}_n)/\sqrt{n} \xrightarrow{P} 0$. With (2.3) this shows $\hat{\underline{b}}_n - \tilde{\underline{b}}_n \xrightarrow{P} 0$, the result of d). Finally,

$$\begin{aligned} \|Q_n^{-1/2} (\tilde{\underline{\varepsilon}} - \underline{\varepsilon})/\sqrt{n}\| &= \|Q_n^{-1/2} X_n' (\hat{\underline{b}}_n - \tilde{\underline{b}}_n)/\sqrt{n}\| \leq \\ &\leq \|Q_n^{-1/2} X_n' \tilde{S}_n^{-1/2}\| \cdot \|\tilde{S}_n^{1/2} (\hat{\underline{b}}_n - \tilde{\underline{b}}_n)/\sqrt{n}\| \xrightarrow{P} 0 \end{aligned}$$

and so

$$\begin{aligned} |\hat{s}_n^2 - \tilde{s}_n^2| &= |\hat{\underline{\varepsilon}}' \hat{Q}_n^{-1} \hat{\underline{\varepsilon}} - \tilde{\underline{\varepsilon}}' Q_n^{-1} \tilde{\underline{\varepsilon}}|/n = |\underline{\varepsilon}' \hat{Q}_n^{-1} \underline{\varepsilon} - \underline{\varepsilon}' Q_n^{-1} \underline{\varepsilon}|/n = \\ &= |(\underline{\varepsilon}' Q_n^{-1/2}/\sqrt{n}) \cdot \{Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - Q_n^{-1/2} (\hat{\underline{\varepsilon}} - \tilde{\underline{\varepsilon}})/\sqrt{n} + \\ &+ (Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I_n) Q_n^{-1/2} \tilde{\underline{\varepsilon}}/\sqrt{n}\}| \leq \end{aligned}$$

$$\leq \|Q_n^{-1/2} \underline{\varepsilon}/\sqrt{n}\| \cdot \{\|Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2}\| \cdot \|Q_n^{-1/2}(\underline{\varepsilon} - \tilde{\underline{\varepsilon}})/\sqrt{n}\| + \\ + \|Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I_n\| \|Q_n^{-1/2} \tilde{\underline{\varepsilon}}/\sqrt{n}\|\} \xrightarrow{P} 0.$$

f) Follows immediately from e).

Proof of theorem 2.2

Set $\underline{G}_n = \tilde{S}_n^{-1/2} \hat{S}_n \tilde{S}_n^{-1/2}$, then $\underline{G}_n \xrightarrow{P} I$ according to theorem 2.1, remark.

Note that $\hat{D}_n^- = \underline{G}_n^{-1/2} \tilde{S}_n^{1/2}$. So the result follows from (2.6) provided that we can show that $\underline{G}_n^{1/2} \xrightarrow{P} I$. However,

$$\|\underline{G}_n^{1/2} - I\| = \|(\underline{G}_n - I)(\underline{G}_n^{1/2} + I)^{-1}\| \leq \|\underline{G}_n - I\| \xrightarrow{P} 0.$$

concluding the proof.

Proof of theorem 2.3

Let $M_n = X_n S_n^{-1} X_n'$ then $\underline{e} = (I_n - M_n)\underline{\varepsilon}$. This gives

$$E|\underline{\varepsilon} - \underline{e}|^2 = E|M_n \underline{\varepsilon}|^2 = \sigma^2 \text{tr}(M_n Q_n M_n') \leq \sigma^2 \|Q_n\| \text{tr}(M_n M_n') = \sigma^2 k \|Q_n\|,$$

$$E|\underline{e}|^2 \leq E|\underline{\varepsilon}|^2 = \sigma^2 \text{tr}(Q_n).$$

So (2.7) implies that $|\underline{\varepsilon} - \underline{e}|^2/n \xrightarrow{P} 0$ and that $|\underline{e}|^2/n$ and $|\underline{\varepsilon}|^2/n$ are P -bounded. This gives

$$\frac{1}{n} (\underline{\varepsilon}' A_n \underline{\varepsilon} - \underline{e}' A_n \underline{e}) = \frac{1}{n} |\underline{\varepsilon}' A_n (\underline{\varepsilon} - \underline{e}) + (\underline{\varepsilon} - \underline{e})' A_n \underline{e}| \leq \\ \leq \|A_n\| \cdot \left\{ \frac{2}{n} |\underline{\varepsilon} - \underline{e}|^2 \cdot (|\underline{\varepsilon}|^2/n + |\underline{e}|^2/n) \right\}^{1/2} \xrightarrow{P} 0,$$

proving the result.

§ 4. Autocorrelation

We consider the univariate linear model

$$y_t = \beta' x_t + \varepsilon_t, \quad t = 1, 2, \dots$$

where for time t y_t is an observable random variable with values taken by the dependent variable, x_t is an observable non-random k -vector of k -explanatory variables, β a k -vector of regression coefficients and ε_t a non-observable random error. We consider estimators based on the first n observations. Then this model can be written in the form (1.1) by taking $\underline{y}' = (y_1, \dots, y_n)$, $\underline{x}'_n = [x_1, \dots, x_n]$, $\underline{\varepsilon}' = (\varepsilon_1, \dots, \varepsilon_n)$. Let $S_n > 0$ for some n .

We assume that the ε_t satisfy an autoregressive scheme of the first order, i.e.

$$(4.1) \quad \begin{cases} \varepsilon_t = \rho \varepsilon_{t-1} + \eta_t, & t = 2, 3, \dots \\ \varepsilon_1 = \eta_1 / \sqrt{1-\rho^2} \end{cases}$$

with $|\rho| < 1$. Here the η_t are i.i.d. (independent identically distributed) with $E\{\eta_t\} = 0$ and $V\{\eta_t\} = \sigma^2 > 0$. From (4.1) we get that (1.2) holds with

$$Q_n = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & \rho & \dots & \rho^{n-1} \\ \rho & 1 & \dots & \rho^{n-2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \rho^{n-1} & \rho^{n-2} & \dots & 1 \end{bmatrix}, \quad Q_n^{-1} = \begin{bmatrix} 1 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \dots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}$$

From these expressions it follows that

$$(4.2) \quad (1 + |\rho|)^{-2} \leq \lambda_{\min}(Q_n) \leq \lambda_{\max}(Q_n) \leq (1 - |\rho|)^{-2}.$$

We have:

Theorem 4.1. Under the assumption (4.1) the condition (2.5) holds.

Moreover, if

$$(4.3) \quad S_n^{-1} \rightarrow 0, \quad x_n' S_n^{-1} x_n \rightarrow 0, \quad ,$$

then the condition (2.6) is also fulfilled.

Proof. The condition (2.5) follows from the weak law of large numbers since

$$\frac{1}{n} \varepsilon' \Omega_n^{-1} \varepsilon = \frac{1}{n} \sum_{t=1}^n (\eta_t / \sigma)^2 \xrightarrow{P} 1.$$

Let $\tilde{x}_1 = \sqrt{1-\rho^2} x_1$, $\tilde{x}_t = x_t - \rho x_{t-1}$, $t = 2, 3, \dots$. Then $\tilde{S}_n = \sum_{t=1}^n \tilde{x}_t \tilde{x}_t'$.

Using the lemma in the appendix with x_t , ε_t replaced by $\tilde{x}_t$, $\tilde{\varepsilon}_t$ this gives

$$\tilde{S}_n^{-1/2} (\tilde{b}_n - \beta) = \tilde{S}_n^{-1/2} x_n' \Omega_n^{-1} \varepsilon = \tilde{S}_n^{-1/2} \sum_{t=1}^n \tilde{x}_t \tilde{\varepsilon}_t \xrightarrow{L} N_k(0, \sigma^2 I),$$

provided that

$$\max_{(1 \leq t \leq n)} \tilde{x}_t' \tilde{S}_n^{-1} \tilde{x}_t \rightarrow 0.$$

With the relation (A1) of the appendix and (4.2) it is easily seen that this condition follows from (4.3). This proves (2.6).

Let $\hat{\rho}_n$ be an estimator for ρ based on the first n observations and let $\hat{Q}_n$ be obtained from Q_n by substitution of $\hat{\rho}_n$ for ρ . We have:

Theorem 4.2. Under the assumption (4.1) we have:

$$a) \quad \|S_n^{-1}\| \rightarrow 0 \Rightarrow (2.1),$$

$$b) \quad \|S_n^{-1}\| = O(1/n) \Rightarrow (2.3),$$

$$c) \quad \hat{\rho}_n \xrightarrow{P} \rho \Rightarrow (2.2), (2.4).$$

Proof. a) and b) follow immediately from (4.2).

c) We can write

$$\hat{Q}_n^{-1} - Q_n^{-1} = (\hat{\beta}_n - \rho) \{-C_n + (\hat{\beta}_n + \rho) D_n\}$$

Here C_n is a $n \times n$ -matrix with unit elements just below and above its diagonal and zero elements elsewhere, and D_n is a diagonal matrix with $\text{diag}(D_n) = (0, 1, \dots, 1, 0)$. Since $\|C_n\| < 2$, $\|D_n\| = 1$ we get

$$\|Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I\| \leq (1 - |\rho|)^{-2} \cdot |\hat{\beta}_n - \rho| \{ \|C_n\| + |\hat{\beta}_n + \rho| \cdot \|D_n\| \} \xrightarrow{P} 0,$$

proving (2.2). Furthermore,

$$\begin{aligned} \tilde{S}_n^{-1/2} X_n' (\hat{Q}_n^{-1} - Q_n^{-1}) \underline{\varepsilon} &= (\hat{\beta}_n - \rho) \left[-\tilde{S}_n^{-1/2} \sum_{t=1}^n x_{t-1} \varepsilon_t - \tilde{S}_n^{-1/2} \sum_{t=2}^n x_t \varepsilon_{t-1} + \right. \\ &\quad \left. + (\hat{\beta}_n + \rho) \tilde{S}_n^{-1/2} (x_1 \varepsilon_1 + \sum_{t=1}^n x_{t-1} \varepsilon_{t+1}) \right] \end{aligned}$$

All terms between brackets are P-bounded since their covariance matrices are bounded, e.g. for the first term we get with (2.4) that

$$\text{Cov} \left\{ \sum_{t=1}^n x_{t-1} \varepsilon_{t-1} \right\} = X_{n-1}' Q_{n-1}^{-1} X_{n-1} \leq (1 - |\rho|)^{-2} S_n$$

or

$$\text{Cov} \left\{ \tilde{S}_n^{-1/2} \sum_{t=1}^n x_{t-1} \varepsilon_{t-1} \right\} \leq (1 - |\rho|)^{-2} \tilde{S}_n^{-1/2} S_n \tilde{S}_n^{-1/2} \leq \left(\frac{1 + |\rho|}{1 - |\rho|} \right)^2 I.$$

Since $\hat{\beta}_n \xrightarrow{P} \rho$ this gives (2.4).

Theorems 4.1 and 4.2 above in combination with theorem 2.1 describe the asymptotic behaviour of the EGLS-method. The discussion is complete if we can construct a consistent estimator $\hat{\beta}_n$ for ρ . Let $\underline{e}' = (\underline{e}_1, \dots, \underline{e}_n)$ and let the Durbin-Watson statistic $\underline{d}_n$ be defined by

$$\underline{d}_n = \frac{\sum_{t=1}^{n-1} (\underline{e}_{t+1} - \underline{e}_t)^2}{\sum_{t=1}^n \underline{e}_t^2}.$$

Theorem 4.3 below states that $\underline{r}_n = 1 - \frac{1}{2} \underline{d}_n$ is a consistent estimator for ρ . We add an elementary proof of this well-known result for sake of completeness.

Theorem 4.3. Under the assumption (4.1) we have

$$\underline{r}_n \xrightarrow{P} \rho$$

Proof. We have to prove that $\underline{d}_n \xrightarrow{P} 2(1-\rho)$. We can write $\underline{d}_n = \underline{e}' A_n \underline{e} / \underline{e}' \underline{e}$, where A_n is the $n \times n$ -matrix which follows from the expression for Q_n^{-1} by formal substitution of $\rho = 1$. So it suffices to prove that

$$\frac{1}{n} \underline{e}' \underline{e} \xrightarrow{P} \sigma^2 / (1-\rho^2), \quad \frac{1}{n} \underline{e}' A_n \underline{e} \xrightarrow{P} 2\sigma^2 / (1+\rho)$$

or, using theorem 2.3 since (2.7) is fulfilled, that

$$\frac{1}{n} \underline{\varepsilon}' \underline{\varepsilon} \xrightarrow{P} \sigma^2 / (1-\rho^2), \quad \frac{1}{n} \underline{\varepsilon}' A_n \underline{\varepsilon} \xrightarrow{P} 2\sigma^2 / (1+\rho).$$

Let $\underline{\eta}' = (\eta_1, \dots, \eta_n)$. Then $\underline{\varepsilon} = B_n \underline{\eta}$, where

$$B_n = \begin{bmatrix} 1/\sqrt{1-\rho^2} & 0 & 0 & \dots & 0 & 0 \\ \rho/\sqrt{1-\rho^2} & 1 & 0 & \dots & 0 & 0 \\ \rho^2/\sqrt{1-\rho^2} & \rho & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho^{n-1}/\sqrt{1-\rho^2} & \rho^{n-2} & \rho^{n-3} & \dots & \rho & 1 \end{bmatrix}$$

Let $C_n = \{c_{ts}(n)\} = B_n' B_n$. Then

$$\text{diag}(C_n) = (1-\rho^2)^{-1} ((1-\rho^{2n})/(1-\rho^2), 1-\rho^{2(n-1)}, \dots, 1-\rho^2)$$

$$\text{tr}(C_n' C_n) = \text{tr}(Q_n^2) \leq \lambda_{\max}(Q_n) \cdot \text{tr}(Q_n) = O(n) \dots$$

This gives

$$\frac{1}{n} \varepsilon' \varepsilon = \frac{1}{n} \underline{\eta}' C_n \underline{\eta} = \frac{1}{n(1-\rho^2)} \sum_1^n \underline{\eta}_t^2 + \underline{R}_{n1} + \underline{R}_{n2},$$

where

$$\underline{R}_{n1} = \frac{1}{n(1-\rho^2)} \{ \rho^2 (1-\rho^{2n}) / (1-\rho^2) \underline{\eta}_1^2 - \sum_1^n \rho^{2(n+1-t)} \underline{\eta}_t^2 \} \xrightarrow{P} 0$$

$$\underline{R}_{n2} = \frac{1}{n} \sum_{t \neq s} \sum c_{ts}^2(n) \underline{\eta}_t \underline{\eta}_s \xrightarrow{P} 0$$

since

$$E|\underline{R}_{n1}| \leq \frac{\sigma^2}{n(1-\rho^2)} \{ \rho^2 (1-\rho^2)^{-1} + \rho^2 (1-\rho^2)^{-1} \} = \frac{2\rho^2 \sigma^2}{n(1-\rho^2)^2} \rightarrow 0$$

$$E\{\underline{R}_{n2}\} = 0, \quad V\{\underline{R}_{n2}\} = \frac{2\sigma^4}{n^2} \sum_{t \neq s} c_{ts}^2(n) \leq \frac{2\sigma^4}{n^2} \text{tr}(C_n' C_n) \rightarrow 0.$$

The weak law of large numbers gives $\sum_1^n \underline{\eta}_t^2 / n \xrightarrow{P} \sigma^2$ and together this proves

$$\frac{1}{n} \varepsilon' \varepsilon = \frac{1}{n} \sum_1^n \underline{\varepsilon}_t^2 \xrightarrow{P} \sigma^2 / (1-\rho^2).$$

Finally

$$\frac{1}{n} \varepsilon' A_n \varepsilon = \frac{1}{n} \sum_1^{n-1} (\underline{\varepsilon}_{t+1} - \underline{\varepsilon}_t)^2 = \frac{1}{n} \sum_1^{n-1} \underline{\varepsilon}_{t+1}^2 + \frac{1}{n} \sum_1^{n-1} \underline{\varepsilon}_t^2 - \frac{2}{n} \sum_1^{n-1} \underline{\varepsilon}_{t+1} \underline{\varepsilon}_t$$

and so the proof is complete if we can show that

$$\frac{2}{n} \sum_1^{n-1} \underline{\varepsilon}_{t+1} \underline{\varepsilon}_t \xrightarrow{P} 2\rho\sigma^2 / (1-\rho^2).$$

For $\rho \neq 0$ this follows from the relation $\underline{\varepsilon}_{t+1} - \rho \underline{\varepsilon}_t = \underline{\eta}_{t+1}$ which gives

$$\frac{2\rho}{n} \sum_{i=1}^{n-1} \varepsilon_{t+1} \varepsilon_t = \frac{1}{n} \sum_{i=1}^{n-1} \varepsilon_{t+1}^2 + \frac{\rho^2}{n} \sum_{i=1}^{n-1} \varepsilon_t^2 - \frac{1}{n} \sum_{i=1}^{n-1} \eta_{t+1}^2 \xrightarrow{P} \frac{2\rho^2}{1-\rho^2} \sigma^2.$$

For $\rho = 0$ we have $\varepsilon_t = \eta_t$ and so the result follows from

$$E\{\eta_{t+1} \eta_t\} = 0, \quad V\{\frac{1}{n} \sum_{i=1}^{n-1} \eta_{t+1} \eta_t\} = \sigma^4 (n-1)/n^2 \rightarrow 0.$$

§ 5. Heteroskedasticity

We consider the univariate linear model of section 4. However, in stead of (4.1) we suppose

$$(5.1) \quad \varepsilon_t = \eta_t / \sqrt{\lambda_t}, \quad t = 1, 2, \dots,$$

where for the weighting coefficients λ_t we restrict ourselves to the case that

$$(5.2) \quad 0 < \inf \lambda_t \leq \sup \lambda_t < \infty.$$

From (4.1) we see that (1.2) holds with

$$Q_n = \text{diag}(1/\lambda_1, \dots, 1/\lambda_n).$$

We have:

Theorem 5.1. Under the assumptions (5.1), (5.2) the condition (2.5) holds. Moreover, if (4.2) holds then the condition (2.6) is fulfilled.

Proof. Compare the proof of theorem 4.1. In this case take $\tilde{x}_t = \sqrt{\lambda_t} \cdot x_t$.

For this model we will only give an illustration of the use of theorem 4.2. Let $\beta' = (\beta_1, \dots, \beta_k)$, $x_t' = (x_{t1}, \dots, x_{tk})$ and assume

$$(5.3) \quad \lambda_t = 1 + 1/(\beta_1 x_{t1})^2, \quad t = 1, 2, \dots,$$

where $\beta_1 \neq 0$, $\inf |x_{t1}| > 0$. The assumption (5.3) implies that $E\{\varepsilon_t^2\}$ is an increasing function of the amount $\beta_1 x_{t1}$ to the expectation $E\{y_t\} = \beta' x_t$ belonging to the first explanatory variable. This assumption seems to be reasonable if the first variable is dominant.

Let $\hat{\beta}_{n1}$ be an estimator for β_1 based on the first n observations, let $\hat{\lambda}_{nt}$ be obtained from λ_t by substitution of $\hat{\beta}_{n1}$ for β_1 , and let $\hat{Q}_n$ be obtained from Q_n by substitution of $\hat{\lambda}_{nt}$ for λ_t .

Theorem 5.2. Under the assumptions (5.1) - (5.3) we have

- a) $\|S_n^{-1}\| \rightarrow 0 \Rightarrow (2.1),$
- b) $\|S_n^{-1}\| = O(1/n) \Rightarrow (2.3),$
- c) $\hat{\beta}_{n1} \xrightarrow{P} \beta_1 \Rightarrow (2.2), (2.4).$

Proof. a) and b) follow immediately from (5.2).

c) Note that

$$\hat{\lambda}_{nt}/\lambda_t - 1 = (\beta_1^2/\hat{\beta}_{n1}^2 - 1)/(1 + \beta_1^2 x_{t1}^2).$$

Hence,

$$\|Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I\| = \max_{(1 \leq t \leq n)} |\hat{\lambda}_{nt}/\lambda_t - 1| \leq |\beta_1^2/\hat{\beta}_{n1}^2 - 1| \xrightarrow{P} 0,$$

proving (2.2). Furthermore,

$$\begin{aligned} \tilde{S}_n^{-1/2} x_n' (\hat{Q}_n^{-1} - Q_n^{-1}) \underline{\varepsilon} &= \tilde{S}_n^{-1/2} \sum_1^n x_t (\hat{\lambda}_{nt} - \lambda_t) \underline{\varepsilon}_t = \\ &= (\beta_1^2/\hat{\beta}_{n1}^2 - 1) \cdot \tilde{S}_n^{-1/2} \sum_1^n \sqrt{\lambda_t} x_t / (1 + \beta_1^2 x_{t1}^2) \cdot \underline{\eta}_t \xrightarrow{P} 0 \end{aligned}$$

since

$$\text{Cov}(\tilde{S}_n^{-1/2} \sum_1^n \sqrt{\lambda_t} x_t / (1 + \beta_1^2 x_{t1}^2) \cdot \underline{\eta}_t) \leq \sigma^2 I.$$

This proves (2.4).

As in section 4 the discussion is complete if we can construct a consistent estimator $\hat{\beta}_{n1}$ for β_1 . Let $\underline{b}'_n = (\underline{b}_{n1}, \dots, \underline{b}_{nk})$ be the OLS-estimator for $\beta' = (\beta_1, \dots, \beta_k)$. Theorem 5.3 below states that $\underline{b}_n$ is consistent for β . Then also $\underline{b}_{n1}$ is consistent for β_1 .

Theorem 5.3. Under the assumption (5.1), (5.2) we have

$$\|S_n^{-1}\| \Rightarrow \underline{b}_n \xrightarrow{P} \beta$$

Proof. Follows from $E\{\underline{b}_n\} = \beta$ and

$$\text{Cov}\{\underline{b}_n\} = S_n^{-1} X_n' Q_n^{-1} X_n S_n^{-1} \leq \sup \lambda_t \cdot S_n^{-1} \rightarrow 0.$$

§ 6. Multivariate linear regression

We consider the m-variate linear model

$$(6.1) \quad y_{ti} = \beta'_i x_t(i) + \varepsilon_{ti}, \quad t = 1, 2, \dots; i = 1, \dots, m,$$

where for time t and the i^{th} -equation y_{ti} is an observable random variable with values taken by the i^{th} dependent variable, $x_t(i)$ is an observable non-random k_i -vector of k_i explanatory variables, β_i a k_i -vector of regression coefficients and ε_{ti} a non-observable random error. We consider estimators based on the first T observations.

Let $k = \sum_{i=1}^m k_i$, $y'_t = (y_{t1}, \dots, y_{tm})$, $\varepsilon'_t = (\varepsilon_{t1}, \dots, \varepsilon_{tm})$, $\beta' = (\beta'_1, \dots, \beta'_m)$ and let the $k \times m$ -matrix $X(t)$ be defined by $X(t) = \text{diag}(x_t(1), \dots, x_t(m))$. Then we can write

$$y_t = X'(t)\beta + \varepsilon_t, \quad t = 1, 2, \dots$$

Therefore this model can be written in the form (1.1) with $n = mT$,

$$y' = (y'_1, \dots, y'_T), \quad \varepsilon' = (\varepsilon'_1, \dots, \varepsilon'_T), \quad X_n = [X(1), \dots, X(T)]'.$$

Let $S_{ni} = \sum_{t=1}^T x_t(i) x'_t(i)$ and suppose $S_{ni} > 0$ for some n for all i (non-collinearity for each equation). Then $S_n = X'_n X_n = \text{diag}(S_{n1}, \dots, S_{nm}) > 0$ for some n .

We assume that the ε_t are i.i.d. with $\text{Cov}\{\varepsilon_t\} = \Sigma > 0$. Then (1.2) holds with $\sigma^2 = 1$ and

$$Q_n = I_n \otimes \Sigma.$$

Note that

$$(6.2) \quad \|\Sigma\|^{-1} \cdot S_n \leq \tilde{S}_n \leq \|\Sigma^{-1}\| \cdot S_n$$

We have:

Theorem 6.1. Consider the model (6.1). If for each $i = 1, \dots, m$ we have that

$$(6.3) \quad S_{ni}^{-1} \rightarrow 0, \quad x'_n(i) S_{ni}^{-1} x_n(i) \rightarrow 0$$

then the condition (2.6) is fulfilled.

Proof. Note that $\tilde{S}_n = \sum_{t=1}^T X(t) \Sigma^{-1} X'(t)$. Using the lemma in the appendix we get

$$\begin{aligned} \tilde{S}_n^{-1/2} (\underline{b}_n - \beta) &= \tilde{S}_n^{-1/2} X'_n Q_n^{-1} \underline{\varepsilon} = \\ &= \tilde{S}_n^{-1/2} \sum_{t=1}^T (X(t) \Sigma^{-1/2}) (\Sigma^{-1/2} \underline{\varepsilon}_t) \xrightarrow{L} N_k(0, I), \end{aligned}$$

provided that

$$\max_{(1 \leq t \leq n)} \text{tr}(\Sigma^{-1/2} X'(t) S_n^{-1} X(t) \Sigma^{-1/2}) \rightarrow 0$$

or, equivalently according to (6.2), that

$$\max_{(1 \leq t \leq n)} \text{tr}(X'(t) S_n^{-1} X(t)) \rightarrow 0.$$

With the relation (A1) in the appendix it is easily seen that this condition is equivalent to (6.3).

Let $\tilde{\Sigma}_n$ be an estimator of Σ based on the first T observations and let $\hat{Q}_n$ be obtained from Q_n by substitution of $\tilde{\Sigma}_n$ for Σ . We have:

Theorem 6.2. For the model (6.1) we have:

- a) $\|S_{ni}^{-1}\| \rightarrow 0$ for all $i \Rightarrow (2.1)$,
- b) $\|S_{ni}^{-1}\| = O(1/n)$ for all $i \Rightarrow (2.3)$,
- c) $\tilde{\Sigma}_n \xrightarrow{P} \Sigma \Rightarrow (2.2), (2.4)$.

Proof. a) and b) follow immediately from (6.2).

c) The relation (2.2) follows from

$$\|Q_n^{1/2} \hat{Q}_n^{-1} Q_n^{1/2} - I_n\| = \|I_T \otimes (\Sigma^{1/2} \tilde{\Sigma}_n^{-1} \Sigma^{1/2} - I)\| = \|\Sigma^{1/2} \tilde{\Sigma}_n^{-1} \Sigma^{1/2} - I\| \xrightarrow{P} 0.$$

For the proof of (2.4) we may replace $\tilde{S}_n$ by S_n according to (6.2). So we have to prove

$$S_n^{-1/2} \sum_1^T x(t) (\tilde{\Sigma}_n^{-1} - \Sigma^{-1}) \varepsilon_t \xrightarrow{P} 0.$$

The i^{th} part of the expression on the left-hand side equals

$$S_{ni}^{-1/2} \sum_t \sum_j x_t(i) (\tilde{\Sigma}_n^{-1} - \Sigma^{-1})_{ij} \varepsilon_{tj}.$$

Therefore, using $\tilde{\Sigma}_n \xrightarrow{P} \Sigma$, it suffices to prove that

$$S_{ni}^{-1/2} \sum_1^T x_t(i) \varepsilon_{tj}$$

is P-bounded. However, this follows immediately since its covariancematrix equals $(\Sigma)_{jj} \cdot I$.

As in section 4.5 we construct a consistent estimator $\tilde{\Sigma}_n$ for Σ . We split up the OLS-residual according to $\underline{e}' = (\underline{e}'_1, \dots, \underline{e}'_T)$ and take

$$\hat{\Sigma}_n = \frac{1}{T} \sum_1^T \underline{e}_t \underline{e}_t'$$

Theorem 6.3. For the model (6.1) we have:

$$\hat{\Sigma}_n \xrightarrow{P} \Sigma$$

Proof. The weak law of large numbers gives that $(\sum_1^T \underline{e}_t \underline{e}_t')/T \xrightarrow{P} \Sigma$. Therefore it suffices to prove that

$$\frac{1}{n} \sum_1^T (\underline{e}_{ti} \underline{e}_{tj} - \varepsilon_{ti} \varepsilon_{tj}) \xrightarrow{P} 0$$

for all i, j . Let I_{ij} be the $m \times m$ -matrix with value 1 for the $(i, j)^{\text{th}}$ element

and zero elements elsewhere. Let A_n be the $n \times n$ -matrix of the form $A_n = \text{diag}(I_{ij}, \dots, I_{ij})$. Since $\|Q_n\|$ and $\|A_n\|$ are bounded, theorem 2.3 gives

$$\frac{1}{n} \sum_1^T (e_{ti} e_{tj} - \varepsilon_{ti} \varepsilon_{tj}) = \frac{1}{n} (\underline{e}' A_n \underline{e} - \underline{\varepsilon}' A_n \underline{\varepsilon}) \xrightarrow{P} 0,$$

completing the proof.

Appendix

The following lemma is basically contained in Eicker (1966).

Lemma. Let $(X_t)_{t \in \mathbb{N}}$ be a sequence of non-random $k \times m$ -matrices and $(\underline{\varepsilon}_t)_{t \in \mathbb{N}}$ a sequence of i.i.d. random m -vectors with $E\{\underline{\varepsilon}_t\} = 0, \text{Cov}\{\underline{\varepsilon}_t\} = I, t \in \mathbb{N}$.

Set $S_n = \sum_{t=1}^n X_t X_t'$. We have:

$$\max_{(1 \leq t \leq n)} \text{tr}(X_t' S_n^{-1} X_t) \rightarrow 0 \Rightarrow S_n^{-1/2} \sum_{t=1}^n X_t \underline{\varepsilon}_t \xrightarrow{d} N_k(0, I)$$

Proof. Let $c \in \mathbb{R}_k$ with $|c| = 1, \underline{x}_{nt} = c' S_n^{-1/2} X_t \underline{\varepsilon}_t, \rho_n = \max \text{tr}(X_t' S_n^{-1} X_t)$. Then it suffices to prove that

$$\sum_{t=1}^n \underline{x}_{nt} \xrightarrow{d} N(0, 1).$$

However, this follows by verifying the conditions of the Lindeberg-Feller (LF) central limit theorem.

We have $E\{\underline{x}_{nt}\} = 0, \sum_{t=1}^n V\{\underline{x}_{nt}\} = 1$ and, using

$$\underline{x}_{nt}^2 \leq \text{tr}(X_t' S_n^{-1} X_t) \cdot |\underline{\varepsilon}_t|^2 \leq \rho_n |\underline{\varepsilon}_t|^2$$

$$\sum_{t=1}^n \text{tr}(X_t' S_n^{-1} X_t) = k$$

that

$$\sum_{t=1}^n E\{\underline{x}_{nt}^2 I(|\underline{x}_{nt}| \geq \varepsilon)\} \leq \sum_{t=1}^n \text{tr}(X_t' S_n^{-1} X_t) \cdot E\{|\underline{\varepsilon}_t|^2 I(|\underline{\varepsilon}_t| \geq \varepsilon/\sqrt{\rho_n})\}$$

$$\leq k \cdot \max_{(1 \leq t \leq n)} E\{|\underline{\varepsilon}_t|^2 \cdot I(|\underline{\varepsilon}_t| \geq \varepsilon/\sqrt{\rho_n})\} \rightarrow 0.$$

Remark. Consider the special case $m = 1$ and write $x_t = X_t$. The following equivalence is easy to prove and often useful:

$$(A1) \quad \max_{(1 \leq t \leq n)} x_t' S_n^{-1} x_t \rightarrow 0 \Leftrightarrow x_n' S_n^{-1} x_n \rightarrow 0, S_n^{-1} \rightarrow 0.$$

We mention some properties connected with (A1).

Let $x_t' = (x_{t1}, \dots, x_{tk})$, $\Delta_n = \text{diag}(\sum_1^n x_{t1}^2, \dots, \sum_1^n x_{tk}^2)$, $\bar{x}_t(n) = \Delta_n^{-1/2} x_t$ and $\bar{S}_n = \sum_1^n \bar{x}_t(n) \bar{x}_t'(n)$. If for some $\phi > 0$ we have that

$$(A2) \quad \bar{S}_n \rightarrow \phi, |\bar{x}_n(n)| \rightarrow 0, \sum_1^n x_{tj}^2 \rightarrow \infty \text{ for } j = 1, \dots, k,$$

then the condition in (A1) is fulfilled. If for some $\psi > 0$ we have

$$(A3) \quad \frac{1}{n} S_n \rightarrow \psi$$

then the condition (A2) is fulfilled.

The condition (A3) appears in many econometric textbooks but is very restrictive. For polynomials $x_t' = (t^{a_1}, \dots, t^{a_k})$ with $a_1 > \dots > a_k \geq 0$ the condition (A3) is not fulfilled if $a_1 > 0$ but (A2) holds. If $k = 2$, $x_{t1} = 1$, $x_{t2} = 1 + 1/\sqrt{t}$ then (A2) is not fulfilled but the condition in (A1) holds.

The condition in (A1) implies that x_t is non-exponentially increasing (i.e. $\rho^n x_n \rightarrow 0$ for all ρ with $|\rho| < 1$).

References

Eicker, F. (1966) - A multivariate central limit theorem for random linear vector forms - Ann. Math. Stat. 37, 1825-1828.

Judge, G.G.; Griffiths, W.E.; Hill, R.C. and Tsoung-Chao Lee (1980) - The theory and practice of econometrics - Wiley, New York.

Schmidt, P. (1976) - Econometrics - Dekker, New York.

Theil, H. (1971) - Principles of econometrics - Wiley, New York.

IN 1981 REEDS VERSCHENEN:

- 96 B.B. Van der Genugten
Central limit theorems for least squares estimators in linear regression models including lagged variables.
- 97 A.J.M. Swanenberg
Rationing and price dynamics in a simple market-game.
- 98 A.J. Hendriks
Theoretische en praktische problemen van het distributieplanologisch onderzoek.
Referaten themadag RSA op 14 mei 1981.
- 99 A.J.J. Talman en L. Van der Heyden
Algorithms for the linear complementarity problem which allow an arbitrary starting point.
- 100 J.P.C. Kleijnen
Cross-validation using the t statistic.
- 101 J.P.C. Kleijnen
Statistical aspects of simulation: an updated survey. Version 1.
- 102 Dr. Wim G.H. van Hulst
On the concept of divergence in the theory of industrial organization.
- 103 H. Gremmen, T. van Bergen, J. Hotterbeekx
Waar liggen de Nederlandse comparatieve voordelen? Voorlopige versie.
- 104 A.L. Hempenius en P.G.H. Mulder
Multiple Failure Rates and Observations of Time Dependent Covariables.
(Part 1: Theory.)
- 105 M.H.C. Paardekooper
A Newton-like method for error analysis. Applied to linear continuous systems and eigenproblems.
- 106 G. van der Laan en A.J.J. Talman
Note of the path following approach of equilibrium programming.

IN 1982 REEDS VERSCHENEN:

107 Aart J. de Zeeuw

Hierarchical decentralized optimal control in econometric
policy models.

108 Arie Kapteyn en Tom Wansbeek

Identification in Factor Analysis.

109 G. van der Laan en A.J.J. Talman

Simplicial Algorithms for finding Stationary Points, a unifying
description

110 Pieter Boot

Ekonomiese betrekkingen tussen Oost en West Europa.

Bibliotheek K. U. Brabant

17 000 01059805 1