A central limit theorem with applications in regression analysis
van der Genugten, B.B.

Publication date:
1978

Link to publication

Citation for published version (APA):
REEKS "TER DISCUSSIE"
A central limit theorem with applications in regression analysis

B.B. van der Genugten

FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
Contents.

1. Introduction .................................................. 1
2. Double sequences of random variables and vectors ............. 2
3. Multivariate linear regression with independent errors .......... 14
4. Multivariate linear regression with dependent errors ............ 38
References .................................................................. 43
1. Introduction.

This paper deals with a central limit theorem and its applications in regression analysis. A little part of it is a generalization and reformulation of van der Genugten (1977), to which we refer for a discussion of the related work of Eicker (1963\textsuperscript{a}), (1963\textsuperscript{b}), (1965), (1966).

In section 2 the basic central limit theorem 2.1. is proved. From this another theorem 2.2, useful in estimating problems, is derived.

In section 3 these results are applied to the multivariate linear model with independent (but not necessarily identically distributed) errors. At first, the OLS-estimator is considered. Its consistency and asymptotic normality are proved (theorem 3.1) under conditions that deal separately with the behaviour of the errors and that of the explanatory variables. Furthermore it is shown that the results maintain to hold if the covariance matrices of the errors are replaced by suitable estimators (theorem 3.3). Secondly, similar results are derived for the GLS-estimator under exactly the same conditions (theorems 3.4 and 3.6). Finally, the analysis is extended to the case of stochastic explanatory variables which are assumed to be independent of the errors.

In section 4 the basic central limit theorem of section 2 is applied to the multivariate linear model with dependent errors, where it is assumed that these errors form a moving average with independent (but not necessarily identically distributed) components. Again the consistency and asymptotic normality of the OLS-estimator is deduced (theorem 4.1).
2. Double sequences of random variables and vectors

We start with the following well-known result:

**Lemma 2.1.** For each \( n = 1, 2, \ldots \) let there be \( j_n \) random variables \( \{\xi_{nj}; j = 1, 2, \ldots, j_n\} \) where \( j_n \to \infty \) if \( n \to \infty \), and suppose that for each fixed \( n \) the random variables \( \xi_{nj}, \ldots, \xi_{n_{n-1}} \) are mutually independent with

\[
E[\xi_{nj}] = 0, \quad \sigma_{nj}^2 = V[\xi_{nj}] < \infty \quad \text{and normed in such a way that}
\]

\[
E \sum_{j=1}^{j_n} V[\xi_{nj}] = 1 \tag{2.1}
\]

If

\[
E \int_{|x|>\delta} x^2 dF_{nj}(x) \to 0, \quad n \to \infty, \quad \text{for all } \delta > 0, \tag{2.2}
\]

where \( F_{nj} \) denotes the distribution function of \( \xi_{nj} \), then

\[
\max_{1 \leq j \leq j_n} P\{|\xi_{nj}| \geq \varepsilon\} \to 0, \quad n \to \infty, \quad \text{for all } \varepsilon > 0 \tag{2.3}
\]

\[
\sum_{j=1}^{j_n} \xi_{nj}^2 \to 1, \quad n \to \infty \tag{2.4}
\]

\[
\sum_{j=1}^{j_n} \xi_{nj} \xrightarrow{L} N(0,1), \quad n \to \infty \tag{2.5}
\]

**Proof:** The Lindeberg-Feller theorem (e.g. see Chung (1974), theorem 7.2.1) implies that (2.3), (2.5) follow from (2.1) and (2.2). A part of a special case of a theorem of Raikov (see Gnedenko and Kolmogorov (1954), § 28, theorem 4) gives that (2.4) follows from (2.3) and (2.5). For a more straightforward proof of this lemma see also van der Genugten (1977).

For applications it is often desirable to extend lemma 2.1 to double sequences of random variables containing infinitely many elements.

Consider the double sequence of random variables \( \{\xi_{nj}; n = 1, 2, \ldots; j = \ldots-1, 0, 1, \ldots\} \). We suppose that for each fixed
The random variables \( \xi_{nj} \), \( j = -1, 0, 1, \ldots \) are mutually independent with
\[ E(\xi_{nj}) = 0, \quad V(\xi_{nj}) < \infty \]
and normed in such a way that \( \sum_j V(\xi_{nj}) = 1 \).

Then for each \( n = 1, 2, \ldots \) the series \( \sum_j \xi_{nj} \) converges in mean square and
the sum has expectation 0 and variance 1. Furthermore, the series \( \sum_j \xi_{nj}^2 \)
converges almost sure and this sum has expectation 1.

For such double sequences we have the following extension of lemma 2.1:

**Lemma 2.2.** If

\[ E \int x^2 \, dF_{nj}(x) \to 0, \quad n \to \infty, \quad \text{for all} \quad \delta > 0, \]

where \( F_{nj} \) denotes the distribution function of \( \xi_{nj} \), then

\[ \sup_j P(\{ |\xi_{nj}| > \epsilon \}) \to 0, \quad n \to \infty \]

\[ \sum_j \xi_{nj}^2 \to 1, \quad n \to \infty \]

\[ \sum_j \xi_{nj} \to N(0, 1), \quad n \to \infty \]

**Proof.** Let \( \sigma_{nj}^2 = V(\xi_{nj}) \). Since for all \( \delta > 0 \):

\[ \sigma_{nj}^2 = \int \frac{x^2}{|x| < \delta} dF_{nj}(x) + \int \frac{x^2}{|x| \geq \delta} dF_{nj}(x) \]

we have

\[ \sup_j \sigma_{nj}^2 \leq \delta^2 + \sup_j \int \frac{x^2}{|x| \geq \delta} dF_{nj}(x) \leq \delta^2 + \sum_j \int \frac{x^2}{|x| \geq \delta} dF_{nj}(x) \]

Take \( n \to \infty \) and use (2.6). By taking \( \delta > 0 \) arbitrary small this gives

\[ \sup_j \sigma_{nj}^2 \to 0, \quad n \to \infty \]
The relation (2.7) follows from (2.10) and

\[ P\left[ \left| \xi_{n,j} \right| \geq \varepsilon \right] \leq \frac{E(\xi_{n,j}^2)}{\varepsilon^2} = \frac{\sigma_{n,j}^2}{\varepsilon^2} \]

Let \( \{j_n\} \) be a sequence of positive integers with \( j_n \to \infty \) if \( n \to \infty \). Then

\[ \tau_n = \sum_{|j| \leq j_n} \sigma_{n,j}^2 + 1 \quad , \quad n \to \infty . \]

Set \( n_{nj} = \xi_{n,j}/\tau_n \) and let \( G_{nj} \) be the distribution function of \( n_{nj} \). Then for \( n \) sufficiently large and all \( \delta > 0 \):

\[ \sum_{|j| \leq j_n} \frac{\int \frac{x^2}{|x| \geq \delta} G_{nj}(x)}{\tau_n} \frac{\int \frac{y^2}{|y| \geq \delta} dF_{nj}(y)}{\tau_n} \leq 2 \sum_{|j| \leq j_n} \frac{\int \frac{y^2}{|y| \geq \delta/2} dF_{nj}(y)}{\tau_n} \to 0 \quad , \quad n \to \infty \]

by (2.6). Therefore the conditions of lemma (2.1) are fulfilled for the double sequence \( \{n_{nj}, j = -j_n, \ldots, j_n\} \). This gives

\[ \sum_{|j| \leq j_n} \frac{\xi_{n,j}^2}{|j| \leq j_n} \to 1 \quad , \quad \sum_{|j| \leq j_n} \frac{\xi_{n,j}}{|j| \leq j_n} \to N(0,1) \]

and therefore also

\[ \sum_{|j| \leq j_n} \frac{\xi_{n,j}^2}{|j| \leq j_n} \to 1 \quad , \quad \sum_{|j| \leq j_n} \frac{\xi_{n,j}}{|j| \leq j_n} \to N(0,1) \]

Since

\[ E\left( \sum_{|j| > j_n} \frac{\xi_{n,j}^2}{|j| > j_n} \right) = V\left( \sum_{|j| > j_n} \frac{\xi_{n,j}}{|j| > j_n} \right) = 1 - \tau_n \to 0 \quad , \quad n \to \infty \]

we have

\[ \sum_{|j| > j_n} \frac{\xi_{n,j}^2}{|j| > j_n} \to 0 \quad , \quad \sum_{|j| > j_n} \frac{\xi_{n,j}}{|j| > j_n} \to 0 \quad , \quad n \to \infty \]

Combining (2.11) and (2.12) the relations (2.8) and (2.9) follow.
From lemma 2.2 we can derive similar results for double arrays of random vectors \( \xi_{nj} \in \mathbb{R}^k \) \((k = 1, 2, \ldots)\). Suppose that for each fixed \( n \) the random vectors \( \xi_{nj}, j = \ldots -1, 0, 1, \ldots \) are mutually independent with expectation vector \( E(\xi_{nj}) = 0 \) and finite covariance matrix \( V(\xi_{nj}) \), normed in such a way that \( E V(\xi_{nj}) = I \). Then (element by element) the series \( \sum_j \xi_{nj} \) converges in mean square and the sum has expectation vector 0 and covariance matrix \( I \). Furthermore, (element by element) \( \sum_j \xi_{nj} \xi'_{nj} \) is a.s. absolutely convergent and this sum has expectation matrix \( I \).

For such double sequences of random vectors we have the following extension of lemma 2.2:

**Lemma 2.3.** If

\[
(2.13) \quad \sum_j \int |z|^2 \, dF_{nj}(z) \to 0, \ n \to \infty, \text{ for all } \delta > 0,
\]

where \( F_{nj} \) denotes the distribution function of \( \xi_{nj} \), then

\[
(2.14) \quad \sup_j P(|\xi_{nj}| \geq \epsilon) \to 0, \ n \to \infty, \text{ for all } \epsilon > 0
\]

\[
(2.15) \quad \sum_j \xi_{nj} \xi'_{nj} \to I, \quad n \to \infty
\]

\[
(2.16) \quad \sum_j \xi_{nj} \overset{L}{\to} \mathcal{N}_k(0, I), \quad n \to \infty
\]

**Proof:** Let \( S_k = \{z: |z| = 1\} \) be the unit sphere in \( \mathbb{R}^k \). Then \( |c'z| \leq |z| \) and \( \{z: |c'z| \geq \delta\} \subset \{z: |z| \geq \delta\} \) for all \( c \in S_k \) and \( \delta > 0 \). This gives

\[
\int_{|c'z| \geq \delta} |c'z|^2 \, dF_{nj}(z) \leq \int_{|z| \geq \delta} |z|^2 \, dF_{nj}(z), \quad c \in S_k
\]

and so (2.13) implies

\[
(2.17) \quad \sum_j \int_{|c'z| \geq \delta} |c'z|^2 \, dF_{nj}(z) \to 0, \ n \to \infty, \text{ for all } \delta > 0 \text{ and } c \in S_k
\]
For fixed \( c \in S_k \) consider the double array \( \{c'\xi_{nj}\} \) of random variables. Then for each fixed \( n \) the random variables \( c'\xi_{nj} \), \( j = \ldots, -1, 0, 1, \ldots \) are mutually independent with \( \mathbb{E}(c'\xi_{nj}) = 0 \) and
\[
\mathbb{E} \, V(c'\xi_{nj}) = c' (\mathbb{E} \, V(\xi_{nj})) = c' I c = 1
\]
Furthermore, if \( F_{nj}(c) \) denotes the distribution function of \( c'\xi_{nj} \) then
\[
\int_{|c'z| \geq \delta} |c'z|^2 \, dF_{nj}(z) = \int_{|x| \geq \delta} x^2 \, dF_{nj}(x)
\]
It follows from (2.17) that theorem 2.2 can be applied to the double sequence \( \{c'\xi_{nj}\} \). This gives:

\[
(2.18) \quad \sup_j P(\{|c'\xi_{nj}| \geq \varepsilon\}) \to 0, \quad n \to \infty, \quad \text{for all } \varepsilon > 0 \text{ and } c \in S_k
\]

\[
(2.19) \quad c'(\mathbb{E} \, \xi_{nj} \xi_{nj}') c \overset{P}{\to} 1, \quad n \to \infty, \quad \text{for all } c \in S_k
\]

\[
(2.20) \quad c'(\mathbb{E} \, \xi_{nj}) \overset{L}{\to} N(0, 1), \quad n \to \infty, \quad \text{for all } c \in S_k
\]

Let \( c_i' = (0, \ldots, 0, 1, 0, \ldots, 0) \) be the unit vector with 1 on the \( i \)-th place and 0 elsewhere, \( i = 1, \ldots, k \) then the \( i \)-th component \( \xi_{nji} \) of \( \xi_{nj} \) equals \( c_i' \xi_{nj} \). With (2.18) this gives
\[
\sup_j P(\{|\xi_{nj}| \geq \varepsilon\}) \leq \sup_j \bigcup_{i=1}^k \{\{|\xi_{nji}| \geq \varepsilon/\sqrt{k}\}\}
\]
\[
< \frac{1}{k} \sup_{i=1}^k \sup_j P(\{|c_i' \xi_{nj}| \geq \varepsilon/\sqrt{k}\}) \to 0, \quad n \to \infty
\]
and (2.14) follows.

Furthermore, by taking \( c = c_i \) in (2.19) for all \( i \) and \( c = (c_i + c_1)/\sqrt{2} \) for all \( i \neq 1 \) it follows that
\[
\mathbb{E} \xi_{nji}^2 \overset{P}{\to} 1, \quad \frac{1}{2} \mathbb{E} (\xi_{nji} + \xi_{njl})^2 \overset{P}{\to} 1
\]
and this gives
or (2.15).

Finally, from (2.20) it easily follows that

\[ c'(E \xi_{nj}) \overset{L}{\to} N(0, c'c), \text{ for all } c \in \mathbb{R}_k \]

and it is well-known that this is equivalent to (2.16).

We will use lemma 2.3 for certain arrays of not-normed random vectors appearing in regression analysis.

For fixed \( k, p = 1, 2, \ldots \) let there be a double array

\[ \{\xi_{nj}, A_{nj} ; n = 1, 2, \ldots ; j = \ldots -1, 0, 1, \ldots \} \]

of random \( k \)-vector \( \xi_{nj} \) and \( k \times p \) matrices \( A_{nj} \). We assume that for each fixed \( n \) the random vectors \( \xi_{nj}, j = \ldots -1, 0, 1, \ldots \), are mutually independent with \( E(\xi_n) = 0 \) and \( \Sigma_{nj} = V(\xi_{nj}) \) exists.

We are interested in the asymptotic distribution of \( \sum_j A_{nj} \xi_{nj} \) for \( n \to \infty \). Under certain regularity conditions this series converges in mean square. Then the sum has expectation vector 0 and covariance matrix

\[
(2.21) \quad C_n^2 = \sum_j A_{nj} \Sigma_{nj} A_{nj}'
\]

If \( C_n^2 \) is non-singular we can introduce the normed vectors

\[
(2.22) \quad \xi_{nj} = C_n^{-1} A_{nj} \xi_{nj}.
\]

Then \( \sum_j \xi_{nj} \) converges in mean square with

\[
(2.23) \quad \sum_j \xi_{nj} = C_n^{-1} \sum_j A_{nj} \xi_{nj}, \quad \text{a.s.}
\]

For such normed series we can apply theorem 2.3. We want to have our conditions in such a form that they deal separately with the behaviour of the random vectors \( \xi_{nj} \) and the coefficient matrices \( A_{nj} \).

The distributions of the \( \xi_{nj} \) are supposed to be comparable with each other. More precisely, we introduce the conditions
\[ m = \inf_{n,j} \min (\Sigma_{nj}) > 0 \]

\[ \sup_{n,j} \int_{|z| > \delta} |z|^2 dG_{nj}(z) = 0, \quad \delta \to \infty, \]

where \( G_{nj} \) denotes the distribution function of \( \varepsilon_{nj} \). Note that

\[ \lambda_{\max}(\Sigma_{nj}) \leq \text{tr}(\Sigma_{nj}) = \int |z|^2 dG_{nj}(z) - \int |z|^{2 \delta} dG_{nj}(z) + \delta^2 \]

and therefore the condition (2.25) implies

\[ M = \sup_{n,j} \lambda_{\max}(\Sigma_{nj}) < \infty \]

The appropriate rather weak conditions for the coefficient matrices \( A_{nj} \) are

\[ \Sigma \text{tr}(A_{nj} A_{nj}) < \infty \]

\[ Z_n = \Sigma A_{nj} A_{nj} \text{ non-singular for all } n \]

\[ \sup_{j} \text{tr}(A_{nj} Z_n^{-1} A_{nj}) = 0, \quad n \to \infty \]

Here, and in what follows, "for all n sufficiently large" is abbreviated to "for all n".

Theorem 2.1. If the distributions of the \( \{\varepsilon_{nj}\} \) satisfy (2.24), (2.25) and the coefficient matrices \( \{A_{nj}\} \) the conditions (2.27), (2.28), (2.29), then \( \Sigma A_{nj} \varepsilon_{nj} \) converges in mean square and \( C_n^2 \) is non-singular for all \( n \), and

\[ C_n^{-1} \Sigma A_{nj} \varepsilon_{nj} \overset{L}{\to} N(0, I), \quad n \to \infty \]

\[ C_n^{-1}(\Sigma A_{nj} \varepsilon_{nj} \varepsilon_{nj} A_{nj}) C_n^{-1} \overset{P}{\to} I \]
Proof. With (2.24), (2.26) we get:
\[(2.32) \quad 0 < m Z_n \leq C_n^2 \leq M Z_n\]

Hence, $E_{A_n} \xi_{n_j}$ converges in mean-square and the covariance matrix $C_n^2$ of this sum is non-singular.

Since (2.15), (2.16) correspond to (2.30), (2.31) we have only to prove that
\[(2.13) \quad \sum \int |z|^2 \, dF_{n_j}(z) \to 0, \quad n \to \infty, \quad \text{for all } \delta > 0, \quad |z| \geq \delta\]

where $F_{n_j}$ is the distribution function of $\xi_{n_j}$.

With (2.22) we get
\[(2.33) \quad \int |z|^2 \, dF_{n_j}(z) = \int |C_n^{-1} A_{n_j} z|^2 \, dG_{n_j}(z)\]

and so (2.13) can be replaced by
\[(2.34) \quad a_n : = \sum \int |C_n^{-1} A_{n_j} z|^2 \, dG_{n_j}(z) \to 0, \quad n \to \infty, \quad \text{for all } \delta > 0, \quad |z| \geq \delta\]

With (2.32) it follows that
\[(2.35) \quad |C_n^{-1} A_{n_j} z|^2 = z'A_{n_j} C_n^{-2} A_{n_j} z \leq \frac{1}{m} z'A_{n_j} Z_n^{-1} A_{n_j} z\]

\[\leq \frac{1}{m} |z|^2 \|A_{n_j}^t Z_n^{-1} A_{n_j}\| \leq \frac{1}{m} |z|^2 \text{tr}(A_{n_j}^t Z_n^{-1} A_{n_j})\]

Let
\[\rho_n : = \sup \text{tr}(A_{n_j}^t Z_n^{-1} A_{n_j}),\]

then $\rho_n \to 0, \quad n \to \infty$, because of (2.29). From (2.35) it follows that
\[(2.36) \quad \sup_j |C_n^{-1} A_{n_j} z| \leq |z| \cdot \sqrt{\frac{\rho_n}{m}}\]
From (2.35), (2.36) we get:

\[(2.37)\quad 0 \leq a_n \leq \frac{1}{m} \sum_{j} \text{tr}(A'_{n_j} Z^{-1}_{n} A_{n_j}) \int_{|z| \geq \delta \sqrt{m}/\rho_n} |z|^2 \, dG_{n_j}(z)\]

Note that

\[(2.38)\quad \sum_{j} \text{tr}(A'_{n_j} Z^{-1}_{n} A_{n_j}) = \text{tr}(Z^{-1}_{n} \sum_{j} A_{n_j} A'_{n_j}) = \text{tr}(I) = k\]

Using (2.25) we can for any \(\varepsilon > 0\) find a \(\delta_0 = \delta_0(\varepsilon) > 0\) such that

\[\sup_{n,j} \int_{|z| \geq \delta_0} |z|^2 \, dG_{n_j}(z) \leq \frac{m}{k} \varepsilon\]

Since \(\rho_n \rightarrow 0\) if \(n \rightarrow \infty\) we have \(\delta \sqrt{m}/\rho_n \geq \delta_0\) for all \(n\) sufficiently large. This gives for such \(n\) and all \(j\):

\[(2.39)\quad \int_{|z| \geq \delta \sqrt{m}/\rho_n} |z|^2 \, dG_{n_j}(z) \leq \sup_{n,j} \int_{|z| \geq \delta_0} |z|^2 \, dG_{n_j}(z) \leq \frac{m}{k} \varepsilon\]

Combining (2.37), (2.38), (2.39) we see

\[0 \leq a_n \leq \frac{1}{m} \sum_{j} \text{tr}(A'_{n_j} Z^{-1}_{n} A_{n_j}) \cdot \frac{m}{k} \varepsilon = \varepsilon\]

for all \(n\) sufficiently large, and this proves (2.34).

**Remark.** In applications it is often desired that

\[(2.40)\quad \sum_{j} A_{n_j} \frac{\varepsilon}{\rho_n} \rightarrow 0, \quad n \rightarrow \infty\]

Under this conditions of the theorem and the additional condition

\[(2.41)\quad Z_n \rightarrow 0, \quad n \rightarrow \infty\]

the relation (2.40) holds. For, (2.32) and (2.41) imply \(C^n_2 \rightarrow 0, \quad n \rightarrow \infty\), and this in turn implies (2.40).
In applications the coefficient matrices $A_{nj}$ and the covariance matrices $\Sigma_{nj}$ are often replaced by estimators $\hat{A}_{nj}$ and $\hat{\Sigma}_{nj}$. Suppose that

$$
\Sigma \hat{A}_{nj} \hat{\varepsilon}_{nj} = \Sigma \alpha_{nj} \hat{\Sigma}_{nj} \alpha'_{nj}
$$

and

$$
(\Sigma \hat{\alpha}_{nj} \hat{\varepsilon}_{nj} A'_{nj})_2 = (\Sigma \alpha_{nj} \Sigma_{nj} A'_{nj})_2
$$

converge in some sense (a.s., in mean or in probability) for all $n$. The corresponding estimator for $C_n$ is given by $\tilde{C}_n$. We wonder if (2.30) remains true if $C_n$ is replaced by the square root $C_n^{1/2}$ of $C_n$. In general this does not hold. To get simple conditions it is even better to take another decomposition $\tilde{C}_n$ of $C_n^{1/2}$ satisfying $\tilde{C}_n C'_n = C_n^{1/2}$.

Let $H_n$ be a non-singular $k \times k$-matrix and define

$$
(2.43) \quad \tilde{C}_n = H_n^{-1} (H_n C_n H_n')^{1/2}
$$

Then $\tilde{C}_n C'_n = C_n^{1/2}$ and we have:

**Theorem 2.2.** Let

$$
(2.44) \quad \min_{\lambda n} (H_n Z_n H_n') > 0
$$

$$
(2.45) \quad \max_{\lambda n} (H_n Z_n H_n') < \infty
$$

Under the conditions of theorem 2.1 and the additional conditions

$$
(2.46) \quad C_n^{-1} \{\Sigma (\hat{A}_{nj} - A_{nj}) \hat{\varepsilon}_{nj}\} \xrightarrow{P} 0, \quad n \to \infty
$$

$$
(2.47) \quad Z_n^{-1/2} (\tilde{C}_n^{1/2} - C_n^{1/2}) Z_n^{-1/2} \xrightarrow{P} 0
$$

we have

$$
(2.48) \quad \tilde{C}_n^{-1} \Sigma \alpha_{nj} \hat{\varepsilon}_{nj} \xrightarrow{L} \mathcal{N}_k(0, I)
$$
Proof. From (2.30) and (2.46) we get immediately
\[
C^{-1}_n \sum_j A_{nj} \xi_{nj} \xi_{nk}(0, I)
\]
Set \(G_n^2 = H_n C_n^2 H_n\), \(G_n^2 = H_n \hat{C}_n^2 H_n\) then \(\hat{G}_n = H_n \tilde{C}_n\) by (2.43). Since
\[
\hat{C}_n^{-1} = (\hat{G}_n^{-1} G_n)(G_n^{-1} H_n C_n) C_n^{-1}
\]
\[
(G_n^{-1} H_n C_n)(G_n^{-1} H_n C_n)' = I
\]
the relation (2.48) follows if we can prove that
\[
(2.49) \quad G_n^{-1} \tilde{G}_n \to I
\]
From (2.45) it follows that \(H_n^2 Z_n^{1/2}\) is bounded and so with (2.47):
\[
H_n (\tilde{C}_n^2 - C_n^2) H_n \to 0
\]
or
\[
\tilde{G}_n^2 - G_n^2 \to 0
\]
or
\[
\tilde{G}_n - G_n \to 0
\]
or
\[
G_n (G_n^{-1} \tilde{G}_n - I) \to 0
\]
From (2.32) and (2.44) it follows that \(\lambda_{\min}(G_n)\) is bounded away from zero, or that \(G_n^{-1}\) is bounded. This gives
\[
G_n^{-1} \tilde{G}_n - I \to 0,
\]
proving (2.49).
Remark 1). In applications it is often desired that

$$\sum_j \bar{A}_{nj} \frac{\varepsilon_{nj}}{n} \rightarrow 0$$

Under the conditions of the theorem and the additional condition (2.41) the relation (2.50) holds. Since $C_n \rightarrow 0$, $n \rightarrow \infty$, this follows from (2.40) and (2.46).

Remark 2). If we choose

$$H_n = z_n^{-1/2}$$

then the conditions (2.44), (2.45) are fulfilled.
3. Multivariate linear regression with independent errors

We consider a system of $p$ linear equations $(p = 1, 2, \ldots)$ in which the explanatory variables may differ from equation to equation. It can be written in the form

\[(3.1) \quad y_{ti} = x_{ti}' \beta_i + \varepsilon_{ti}, \quad t = 1, 2, \ldots \quad ; \quad i = 1, \ldots, p.\]

For time $t$ and the $i$th equation $y_{ti}$ is the observable random variable with values taken by the $i$th dependent variable, $x_{ti}$ is an observable non-random $k_i$-vector of $k_i$ explanatory variables $(k_i = 1, 2, \ldots)$, $\beta_i$ a $k_i$-vector of regression coefficients and $\varepsilon_{ti}$ a non-observable random error.

The equations (3.1) can be written in the form

\[(3.2) \quad y_t = X_t' \beta + \varepsilon_t, \quad t = 1, 2, \ldots\]

with

\[
\begin{align*}
Y_t &= \begin{pmatrix} Y_{t1} \\ \vdots \\ Y_{tp} \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{t1} \\ \vdots \\ \varepsilon_{tp} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \quad X_t = \begin{pmatrix} x_{t1} & 0 & \ldots & 0 \\ 0 & x_{t2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & x_{tp} \end{pmatrix}
\end{align*}
\]

We consider observations belonging to the time period \{1, \ldots, n\}, $n = 1, 2, \ldots$. Let $k = \Sigma_{i} k_i$, $X'_{ni} = [x_{1i}, \ldots, x_{ni}]$, $Y'_n = (y'_1, \ldots, y'_n)$, $X'(n) = [X_1, \ldots, X_n]$, $\varepsilon'(n) = (\varepsilon'_1, \ldots, \varepsilon'_n)$. Then for the given time period (3.2) can be written as

\[(3.3) \quad Y(n) = X(n)' \beta + \varepsilon(n), \quad n = 1, 2, \ldots\]

Furthermore, let

\[
\begin{align*}
S_{ni} &:= X'_{ni} X_{ni} = \Sigma_{t=1}^n x_{ti} x_{ti}' \\
S_n &:= X'(n) X(n) = \Sigma_{t=1}^n X_t X_t' = \text{diag}(S_{n1}, \ldots, S_{np})
\end{align*}
\]
Suppose that $e_1, e_2, \ldots$ are mutually independent with $E(e_t) = 0$, $E_t: = V(e_t)$ exists for all $t$. Then

\[
\begin{cases}
E(e(n)) = 0 \\
V(e(n)) = \Sigma(n) = \text{diag}(\Sigma_1, \ldots, \Sigma_n)
\end{cases}
\]

Let $G_t$ denote the distribution function of $e_t$.

We are interested in asymptotic normality of the OLS-estimator $b_n$ and the GLS-estimator $\hat{\beta}_n$ for $\beta$.

**OLS-estimators**

If $S_n$ is non-singular then

\[
b_n - \beta = (X'(n)X(n))^{-1} X'(n)e(n) = S_n^{-1} \sum_{t=1}^{n} X_t e_t
\]

\[
V_n = V(e(n)) = S_n^{-1} X'(n)e(n)X(n)S_n^{-1} = S_n^{-1} \left( \sum_{t=1}^{n} X_t e_t X_t' \right) S_n^{-1}
\]

From (3.6) we see that theorem 2.1 can be applied if we take

$e_{nt} = e_t$ and

$A_{nt} = S_n^{-1} X_t$, \quad $t = 1, \ldots, n$

$A_{nt} = 0$, \quad elsewhere.

The conditions (2.24), (2.25) for the distributions of the $\{e_t\}$ become

\[
m = \inf_{t} \lambda_{\min}(\Sigma_t) > 0
\]

\[
\sup_{t} \int_{|z| > \delta} |z|^2 dG_t(z) \to 0 \quad , \quad \delta \to \infty
\]
and the property (2.26) based on it becomes

\[(3.10) \quad M := \sup_t \lambda_{\max}(E_t) < \infty\]

The explanatory variables \{x_{ti}\} must have the property that \(S_n\) is non-singular for all \(n\). From (3.4) it follows that this is equivalent to

\[(3.11) \quad r(X_{ni}) = k_i \text{ for all } i \text{ and some } n\]

From (2.28) we get \(Z_n = S_n^{-1}\). The condition

\[(3.12) \quad \max_{1 \leq t \leq n} x_{ti}^t S_{ni}^{-1} x_{ti} \to 0, \quad n \to \infty, \quad \text{for all } i\]

is equivalent to

\[(3.13) \quad \max_{1 \leq t \leq n} \text{tr}(X_t^t S_n^{-1} X_t) \to 0, \quad n \to \infty\]

This immediately follows from

\[X_t^t S_n^{-1} X_t = \text{diag}(x_{ti}^t S_{ni}^{-1} x_{ti}; i = 1, \ldots, p)\]

The condition (2.29) corresponds to (3.13) and therefore also to (3.12). It should be noted that the condition (3.12) implies (2.41) or equivalently

\[(3.14) \quad S_n^{-1} \to 0, \quad n \to \infty\]

This follows from the lemma 3.1 below:

**Lemma 3.1.** Let \(\{x_t, t = 1, 2, \ldots\}\) be a sequence of vectors in \(R^k\) with

\[S_n := \sum x_t x_t^t > 0 \text{ for some } n \text{ and let } p = 0, 1, \ldots. \text{ Then we have (for } n \to \infty):\]

\[
\max_{1 \leq t \leq n} x_t^t S_n^{-1} x_t \to 0 \Leftrightarrow \begin{cases} 
S_n^{-1} \to 0 \\
x_{n+p}^t S_n^{-1} x_{n+p} \to 0
\end{cases}
\]
Proof.

\((\ast)\): Take some fixed \( c \in \mathbb{R}^k \) and consider \( c \mathbf{S}_n^{-1} \). Since for some \( N \) sufficiently large \( r(S_n) = k \) we can write \( c = \sum a_t x_t \) for some \( a_1, \ldots, a_N \) not depending on \( n \). Then for \( n \geq N \):

\[
\begin{align*}
\mathbf{c}' \mathbf{S}_n^{-1} \mathbf{c} &= (\sum a_t x_t) \mathbf{S}_n^{-1} (\sum a_s x_s) \\
&\leq |a_t a_s| |x_t' \mathbf{S}_n^{-1} x_s| \\
&\leq \left( \sum |a_t| \sqrt{x_t' \mathbf{S}_n^{-1} x_t} \right)^2 \\
&\leq \left( \sum |a_t|^2 \max_{1 \leq t \leq n} x_t' \mathbf{S}_n^{-1} x_t \right)^2.
\end{align*}
\]

Therefore, \( c \mathbf{S}_n^{-1} c \to 0 \) and since this holds for all \( c \) we have \( \mathbf{S}_n^{-1} \to 0 \).

Furthermore, for \( p = 1, 2, \ldots \):

\[
\max_{1 \leq t \leq n+p} x_t' \mathbf{S}_n^{-1} x_t \to 0
\]

and so

\[
\sum_{j=1}^{p} x_{n+j}' \mathbf{S}_n^{-1} x_{n+j} \to 0
\]

From

\[
\mathbf{S}_{n+q} = \mathbf{S}_n + \sum_{j=1}^{p} x_{n+j}' x_{n+j}
\]

it follows that

\[
\mathbf{I} = \mathbf{S}_n^{-1/2} n^{+p} \mathbf{S}_n^{-1/2} + \mathbf{S}_n^{-1/2} (\sum_{j=1}^{p} x_{n+j}' x_{n+j}) \mathbf{S}_n^{-1/2}
\]

and with

\[
\text{tr}(\mathbf{S}_n^{-1/2} (\sum_{j=1}^{p} x_{n+j}' x_{n+j}) \mathbf{S}_n^{-1/2}) = \sum_{j=1}^{n} x_{n+j}' \mathbf{S}_n^{-1} x_{n+j}
\]

this gives

\[
\mathbf{S}_n^{-1/2} n^{+p} \mathbf{S}_n^{-1/2} \to \mathbf{I}.
\]
Therefore,

$$S_{n+p}^{1/2} S_n^{-1} S_{n+p}^{1/2} \to I$$

Since

$$\sum_{j=1}^{p} x'_{n+j} S_n^{-1} x_{n+j} = \text{tr}(S_n^{-1} \sum_{j=1}^{p} x_{n+j} x'_{n+j}) = \text{tr}(S_n^{-1}(S_{n+p} - S_n)) =$$

$$= \text{tr}(S_n^{-1} S_{n+p} - I) = \text{tr}((S_n^{-1} - S_n^{-1} S_{n+p}) S_{n+p}) =$$

$$= \text{tr}(S_n^{1/2}(S_{n+p} - S_{n+p} S_{n+p}^{1/2} S_{n+p}) S_n^{1/2}) = \text{tr}(S_{n+p}^{-1} S_{n+p}^{1/2} - I)$$

this gives

$$\sum_{j=1}^{p} x'_{n+j} S_n^{-1} x_{n+j} \to 0$$

or

$$x'_{n+p} S_n^{-1} x_{n+p} \to 0,$$

completing the first part of the proof.

($\Leftarrow$): Suppose the assertion is wrong. Then there exist $\epsilon > 0$ and a sequence $\{n_m\}$ with $n_m \to \infty$ if $m \to \infty$ such that for all $n_m$:

$$\max_{1 \leq t < n_m} x'_t S_n^{-1} x_t \geq \epsilon$$

Let $\tau_m$ be a value of $t$ for which the maximum is attained. Then $\tau_m \leq n_m$ and

$$\epsilon \leq x'_{\tau_m} S_n^{-1} x_{\tau_m} \leq \|x_{\tau_m}\| \|S_n^{-1}\|$$

Since $S_n^{-1} \to 0$ we have $\|x_{\tau_m}\| \to \infty$, $m \to \infty$, and therefore also $\tau_m \to \infty$, $m \to \infty$. 
Since \( S_{n+1} > S_n \) we have \( S_{n+1}^{-1} \leq S_n^{-1} \) and so

\[
x' \tau_m S_{n-1}^{-p} x \geq x' \tau_m S_{n-1}^{-p} x > \varepsilon,
\]

contradicting

\[
x' \left( \begin{array}{c} S_{n+1}^{-1} \end{array} \right) x_n \left( \begin{array}{c} \tau_m \end{array} \right) + 0.
\]

This completes the proof.

Hence we can apply theorem 2.1 and its remark.

**Theorem 3.1.** If the distributions of the \( \{e_t\} \) satisfy (3.3), (3.9) and the explanatory variables \( \{x_{ti}\} \) the conditions (3.11), (3.12) then

\[
(3.15) \quad b_n \overset{P}{\to} \beta, \quad n \to \infty
\]

\[
(3.16) \quad V_n^{-1}(b_n - \beta) \overset{P}{\to} N_k(0, I), \quad n \to \infty
\]

\[
(3.17) \quad V_n^{-1}(\sum_{t=1}^{n} x_t e_t e'_t x'_t) V_n^{-1} \overset{P}{\to} I, \quad n \to \infty
\]

Since in general the \( \Sigma_t \) are unknown it makes sense in (3.16) to replace them by estimators \( \hat{\Sigma}_{nt} \). Since in this case \( \hat{A}_{nt} = A_{nt} \) the expressions corresponding to (2.12) and (2.43) are

\[
(3.18) \quad \hat{V}_n^{-1} = S_n^{-1}(\sum_{t=1}^{n} x_t e_t e'_t x'_t) S_n^{-1}
\]

\[
(3.19) \quad \overline{V}_n = H_n^{-1}(H_n V_n^{-1} H'_n)^{1/2}
\]

Then from theorem 2.2 we get immediately:

**Theorem 3.2.** Suppose

\[
(3.20) \quad \inf \lambda_n \min_n \left( H_n S_n^{-1} H_n' \right) > 0
\]
\begin{align*}
(3.21) \quad & \sup_{n} \lambda_{\text{max}}(H_n S_n^{-1} H_n') < \infty \\
\text{Under the conditions of theorem 3.1 and the additional condition} \\
(3.22) \quad & S_n^{-1/2} \{ \sum_{t=1}^{n} X_t (\hat{\Sigma}_n - \Sigma_t) X_t' \} S_n^{-1/2} \to 0, \ n \to \infty \\
\text{we have} \\
(3.23) \quad & \frac{v_n^{-1}(b_n - \beta)}{n} \xrightarrow{d} N_k(0, I), \quad \sigma \to \infty \\
\text{Remark. The conditions (3.20), (3.21) are satisfied if we choose} \ H_n = S_n^{1/2}. \\
\text{We will investigate two estimators} \ \hat{\Sigma}_{nt}, \ \text{both connected with the} \\
\text{OLS-residual vector} \\
\epsilon(n) = x(n) - X(n) b_n \\
\text{Let} \\
\xi(n) = \epsilon(n) - \hat{\epsilon}(n) = X(n) S_n^{-1} X'(n) \epsilon(n) \\
\text{denote the difference with the error vector and set} \ \epsilon(n) = (\epsilon_1'(n), \ldots, \epsilon_n'(n)), \\
\hat{\epsilon}(n) = (\hat{\epsilon}_1'(n), \ldots, \hat{\epsilon}_n'(n)), \ \text{then} \\
(3.24) \quad & \xi_t(n) = \epsilon_t - \hat{\epsilon}_t(n) = X_t' S_n^{-1} X'(n) \epsilon(n) = X_t' S_n^{-1} \sum_{\tau=1}^{n} X_{\tau} \epsilon_{\tau} \\
\text{A usual estimator (for the case of identically distributed errors) is} \\
(3.25) \quad & \hat{\Sigma}_{nt} = \frac{\Sigma_n}{n} = \frac{1}{n} \sum_{t=1}^{n} \epsilon_t(n) \epsilon_t'(n) \\
\text{For the behaviour of} \ \hat{\Sigma}_{nt} \ \text{the following lemma is useful:} \\
\text{Lemma 3.2. Let} \ \Sigma_n \ = \ \frac{1}{n} \sum_{t=1}^{n} \Sigma_t. \ \text{If} \\
(3.26) \quad & \text{tr} \{ \sum_{t=1}^{n} (\Sigma_t - \Sigma_n)^2 \} \text{is bounded in} \ n 
\end{align*}
and if (3.12) holds, then

\[(3.27) \quad \frac{\hat{E}_n - E_n}{n} \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty \]

**Proof.** We have

\[(3.28) \quad \frac{\hat{E}_n - E_n}{n} = \frac{1}{n} \sum_{t=1}^{n} \left( e_t(n) \xi_t(n) - E_t \right) = A_n + B_n - C_n - C', \]

where

\[
A_n := \frac{1}{n} \sum_{t=1}^{n} \left( e_t - E_t \right) \xi_t(n) \xi_t'(n) \\
B_n := \frac{1}{n} \sum_{t=1}^{n} \xi_t(n) \xi_t'(n) \\
C_n := \frac{1}{n} \sum_{t=1}^{n} e_t \xi_t'(n) \\
C' := \frac{1}{n} \sum_{t=1}^{n} \xi_t(n) \xi_t'(n) \\
D_n := \frac{1}{n} \sum_{t=1}^{n} e_t \xi_t \\
E_n := \frac{1}{n} \sum_{t=1}^{n} \xi_t \xi_t'.
\]

We use theorem 2.1 with \(A_n \equiv 1\), \(t = 1, \ldots, n\). Then (2.31) gives

\[(2.31) \quad \frac{\hat{E}_n}{n} = \frac{1}{n} \sum_{t=1}^{n} \left( e_t - E_t \right) \xi_t(n) \xi_t'(n) \to I, \quad n \to \infty
\]

Since \(\hat{E}_n\) is bounded this implies \(A_n \xrightarrow{p} 0\) and

\[D_n = \frac{1}{n} \sum_{t=1}^{n} e_t \xi_t \to 0, \quad n \to \infty.
\]

bounded in probability. From (3.24) we have

\[(3.29) \quad E\{\xi_t(n) \xi_t'(n)\} \leq M \cdot X_t S_n^{-1} X_t
\]

and therefore with (3.12):

\[E\{B_n\} \leq M \cdot \max_{1 \leq t \leq n} X_t S_n^{-1} X_t \to 0, \quad n \to \infty
\]

Since \(B_n \geq 0\) this implies \(B_n \xrightarrow{p} 0\).

Finally, for \(a, b \in \mathbb{R}_p\) we get with Cauchy-Schwarz:
\[(a' C_n b)^2 \leq (a' B_n a)(b' D_n b) \implies 0, \quad n \to \infty,
\]
since \(B_n \implies 0\) and \(D_n\) is bounded in probability. This gives \(C_n \implies 0\).
Therefore (3.27) follows from (3.28) and \(A_n, B_n, C_n \implies 0\).

A more remarkable estimator of \(\Sigma_t\) is given by

\[(3.30) \quad \hat{\Sigma}_{nt} = \tilde{\Sigma}_t = e_t(n) e_t'(n)\]

The following theorem shows that (3.29) is more easier to work with than (3.25). We have:

**Theorem 3.3. (conditions of theorem 3.1).**

a) Let \(\hat{\Sigma}_{nt} = \tilde{\Sigma}_t\) be given by (3.30). Then (without further conditions) the relation (3.22) holds.

Hence, if (3.20), (3.21) are satisfied then (3.23) holds.

b) Let \(\hat{\Sigma}_{nt} = \Sigma_n\) be given by (3.25). Then the condition (3.22) holds provided that (3.26) is true. Hence, if (3.20), (3.21), (3.26) are satisfied then (3.23) holds.

**Proof.**

a) Since

\[
\hat{\Sigma}_{nt} - \Sigma_t = (e_t(n) e_t'(n) - e_t e_t') + (e_t e_t' - \Sigma_t)
\]

we can write

\[(3.22) \quad S_n^{-1/2} \left( \sum_{t=1}^{n} X_t \left( \hat{\Sigma}_{nt} - \Sigma_t \right) X_t' \right) S_n^{-1/2} = A_n + B_n - C_n - C_n',
\]

where

\[
A_n := S_n^{-1/2} \left( \sum_{t=1}^{n} X_t \left( e_t e_t' - \Sigma_t \right) X_t' \right) S_n^{-1/2}
\]

\[
B_n := S_n^{-1/2} \left( \sum_{t=1}^{n} X_t \tilde{\Sigma}_t(n) \tilde{\Sigma}_t'(n) X_t' \right) S_n^{-1/2}
\]
\[ C_n = S_n^{-1/2} \left( \sum_{t=1}^{n} X_t \Sigma_t (n) X'_t \right) S_n^{-1/2} \]

From (3.17) we see that
\[ V_n^{-1} \left( \sum_{t=1}^{n} X_t (\Sigma_t - \Sigma) X'_t \right) V_n^{-1} \xrightarrow{P} 0 \]

and since \( S_n^{-1/2} V_n \) is bounded this gives \( A_n \xrightarrow{P} 0 \) and also, since \[ \text{tr}(S_n^{-1/2} \left( \sum_{t=1}^{n} X_t \Sigma_t X'_t \right) S_n^{-1/2}) \leq M \kappa, \]

\[ D_n = S_n^{-1/2} \left( \sum_{t=1}^{n} X_t \Sigma_t X'_t \right) S_n^{-1/2} \]

is bounded in probability. With (3.29) and (3.13):
\[ E(B_n) \leq M \max_{1 \leq t \leq n} X_t' S_n^{-1} X_t \rightarrow 0 \quad n \rightarrow \infty \]

and so \( B_n \xrightarrow{P} 0 \), because \( B_n \geq 0 \). Finally, for all \( a, b \in \mathbb{R}^k \) we get with Cauchy-Schwarz:
\[ (a' C_n b)^2 \leq (a' B_n a)(b' D_n b) \rightarrow 0 \quad n \rightarrow \infty \]

or \( C_n \xrightarrow{P} 0 \). Therefore (3.22) holds and so part a) is proved.

b) Since
\[ \hat{\Sigma}_n - \Sigma_t = (\hat{\Sigma}_n - \overline{\Sigma}_n) + (\overline{\Sigma}_n - \Sigma_t) \]

the condition (3.28) holds if
\[ S_n^{-1/2} \left( \sum_{t=1}^{n} X_t (\overline{\Sigma}_n - \Sigma_t) X'_t \right) S_n^{-1/2} \rightarrow 0 \]
\[ S_n^{-1/2} \left( \sum_{t=1}^{n} X_t (\hat{\Sigma}_n - \overline{\Sigma}_n) X'_t \right) S_n^{-1/2} \xrightarrow{P} 0 \]

At first consider (3.31). The matrix forming the \((i, j)^{th}\) part of this term equals
For any $c,d \in \mathbb{R}^k_i$ we have

$$
S_{n+1}^{-1/2} \sum_{t=1}^{n} \left( (\bar{F}_n - \bar{E}_t) x_{t_i} x_{t_j} \right) S_{n+1}^{-1/2}
$$

Hence, with (3.12) and (3.26) we see that (3.31) holds.

Secondly, consider (3.32). The matrix forming the $(i,j)^{th}$ part of this term equals

$$
S_{n+1}^{-1/2} \sum_{t=1}^{n} \left( (\bar{F}_n - \bar{E}_t) x_{t_i} x_{t_j} \right) S_{n+1}^{-1/2}
$$

For any $c,d \in \mathbb{R}^k_i$ we have with Cauchy-Schwarz:

$$
\left\{ c' S_{n+1}^{-1/2} \sum_{t=1}^{n} \left( (\bar{F}_n - \bar{E}_t) x_{t_i} x_{t_j} \right) S_{n+1}^{-1/2} d \right\}^2 \leq \prod_{1 \leq t \leq n} \left| (\bar{F}_n - \bar{E}_t) x_{t_i} x_{t_j} \right|\left| c' S_{n+1}^{-1/2} x_{t_i} \right|^2 \left| d' S_{n+1}^{-1/2} x_{t_j} \right|^2
$$

since

$$
\sum_{t=1}^{n} \left( c' S_{n+1}^{-1/2} x_{t_i} \right)^2 = \sum_{t=1}^{n} \left( c' S_{n+1}^{-1/2} x_{t_i} x_{t_i} S_{n+1}^{-1/2} c \right) = c' S_{n+1}^{-1/2} \sum_{t=1}^{n} x_{t_i} x_{t_i} S_{n+1}^{-1/2} c = c' c = |c|^2.
$$

Therefore, (3.32) holds if we can prove that
\[ \hat{\beta}_n - \beta \xrightarrow{P} 0 \quad , \quad n \to \infty \]

However, this is precisely the result of lemma 3.1, and so the proof of part b) is complete.

**GLS-estimators**

If \( X'(n) \Sigma(n) X(n) \) is non-singular then

\[ (3.33) \quad \hat{\beta}_n - \beta = (X'(n)\Sigma^{-1}(n) X'(n))^{-1} X'(n)\Sigma^{-1}(n) \Sigma(n) = W_n \sum_{t=1}^{n} X_t \Sigma_t^{-1} \Sigma_t \]

where

\[ (3.34) \quad W_n = V(\hat{\beta}_n) = (X'(n)\Sigma^{-1}(n) X'(n))^{-1} = \sum_{t=1}^{n} (X_t \Sigma_t^{-1} X_t')^{-1} \]

From (3.33) we see that theorem 2.1 can be applied if we take

\[ e_{nt} = \epsilon_t \quad \text{and} \]

\[ A_{nt} = W_n X_t \Sigma_t^{-1} \quad , \quad t = 1, \ldots, n \]

\[ A_{nt} = 0 \quad , \quad \text{elsewhere.} \]

It is reasonable and appropriate to impose the same conditions for the errors and the explanatory variables as in the OLS-case. Hence we assume (3.8), (3.9) hold for the distributions of the \( \epsilon_t \) and (3.11), (3.12) for the explanatory variables. Then

\[ (3.35) \quad \frac{1}{M} S_n - \frac{W_n^2}{n} = \sum_{t=1}^{n} X_t \Sigma_t^{-1} X_t' < \frac{1}{m} S_n \]

and so \( W_n^2 \) is well-defined. From (2.28) we get

\[ Z_n = W_n^2 (\sum_{t=1}^{n} X_t \Sigma_t^{-2} X_t') W_n \]

and so by (3.14):
\[
\frac{1}{M^2} S_n^{-1} \leq \frac{1}{M} W_n^2 < Z_n \leq \frac{1}{m} W_n^2 \leq \frac{1}{M^2} S_n^{-1} \to 0, \quad n \to \infty
\]

Therefore the condition (2.28) that \( Z_n \) is non-singular and the condition (2.41) that \( Z_n \to 0 \) are satisfied. Moreover,

\[
A_{nt}^t Z_n^{-1} A_{nt} \leq M. \Sigma_t^{-1} X_t W_n^2. W_n^2. W_n X_t \Sigma_t^{-1} =
\]

\[
= M. \Sigma_t^{-1} X_t W_n^2 X_t \Sigma_t^{-1} \leq M^2. \Sigma_t^{-1} X_t S_n^{-1} X_t \Sigma_t^{-1}
\]

and therefore with (3.13):

\[
\sup_{n,t} \text{tr}(A_{nt}^t Z_n^{-1} A_{nt}) \leq k. M^2 \max_{1 \leq t \leq n} \lambda_{\max}(\Sigma_t^{-1} X_t S_n^{-1} X_t \Sigma_t^{-1})
\]

\[
\leq k^2. M^2 \max_{1 \leq t \leq n} \text{tr}(X_t S_n^{-1} X_t) \to 0, \quad n \to \infty,
\]

showing that (2.29) is satisfied. Hence we can again apply theorem 2.1 and its remark:

**Theorem 3.4. (conditions of theorem 3.1).**

(3.36)  \[ \beta_n \xrightarrow{p} \beta, \quad n \to \infty \]

(3.37)  \[ W_n^{-1}(\hat{\beta}_n - \beta) \xrightarrow{L} N_k(0, I) \]

As in the OLS-case we replace the \( \Sigma_t \) by estimators \( \hat{\Sigma}_{nt} \). Suppose that the \( \hat{\Sigma}_{nt} \) are non-singular a.s. and also that

(3.38)  \[ \hat{W}_n = (\Sigma_{t=1}^n X_t \hat{\Sigma}_{nt}^{-1} X_t')^{-1} \]

exists a.s. Take

\[ \hat{A}_{nt} = \hat{W}_n X_t \hat{\Sigma}_{nt}^{-1} \]

then (2.42) holds for \( \hat{C}_n = \hat{W}_n \). The expression corresponding to (2.43) for non-singular \( H_n \) becomes
(3.39) \[ \frac{\hat{w}}{n} = H_n^{-1}(H_n \hat{w}_n^2 H_n^t)^{1/2} \]

Denoting the corresponding estimator for \( \beta \) by \( \hat{b}_n \) we have

(3.40) \[ \frac{\hat{b}_n}{n} - \beta = \frac{\hat{w}_n^2}{n} \sum_{t=1}^{n} X_t \hat{e}_t \hat{e}_t^t \]

Then from theorem 2.2 and its remark 1) we get:

**Theorem 3.5.** Suppose that (3.20), (3.21) hold.

Under the conditions of theorem 3.1 and the additional conditions

(3.41) \[ S_n^{-1/2} \sum_{t=1}^{n} X_t (\hat{e}_t - E_t) e_t \rightarrow 0 \]

(3.42) \[ S_n^{-1/2} \{ \sum_{t=1}^{n} X_t (\hat{e}_t - E_t)^2 \} S_n^{-1/2} \rightarrow 0 \]

we have

(3.43) \[ \frac{\hat{b}_n}{n} \rightarrow \beta, \quad n \rightarrow \infty \]

(3.44) \[ W_n^{-1}(\frac{\hat{b}_n}{n} - \beta) \rightarrow N_k(0, I), \quad n \rightarrow \infty \]

(3.45) \[ W_n^{-1}(\frac{\hat{b}_n}{n} - \beta) \rightarrow N_k(0, I), \quad n \rightarrow \infty \]

**Proof.** Since

\[ \frac{1}{M^2} S_n^{-1} \leq \hat{Z}_n \leq \frac{1}{M^2} S_n^{-1} \]

the conditions (2.44), (2.45) follow from (3.20), (3.21). The conditions (2.46), (2.47) read as

(3.46) \[ W_n^{-1} \sum_{t=1}^{n} (\hat{w}_n^2 X_t \hat{e}_t^t - \hat{w}_n^2 X_t E_t \hat{e}_t^t) e_t \rightarrow 0 \]

(3.47) \[ S_n^{1/2} \{ \hat{w}_n - \hat{w}_n^2 \} S_n^{1/2} \rightarrow 0 \]
We will verify these conditions for proving (3.44), (3.45). From (3.41) and the fact that $W_n S_n^{1/2}$ is bounded we get

\begin{equation}
W_n \sum_{t=1}^{n} X_t (\Sigma_{nt}^{-1} - \Sigma_t^{-1}) \xi_t \xrightarrow{P} 0
\end{equation}

and with (3.37) this gives

\begin{equation}
W_n \sum_{t=1}^{n} X_t \Sigma_{nt}^{-1} \xi_t \xrightarrow{L} N_k(0, I)
\end{equation}

From (3.42) we get

\begin{equation}
W_n \{ \sum_{t=1}^{n} X_t (\Sigma_{nt}^{-1} X_t' - \Sigma_t^{-1} X_t') W_n \} \xrightarrow{P} 0
\end{equation}

or

\begin{equation}
W_n \frac{\hat{W}^{-2} - W_n^{-2}}{W_n} \xrightarrow{P} 0
\end{equation}

or

\begin{equation}
W_n \frac{\hat{W}^{-2}}{W_n} \xrightarrow{P} I
\end{equation}

or

\begin{equation}
W_n^{-1} \frac{\hat{W}^{-2}}{W_n} W_n^{-1} \xrightarrow{P} I
\end{equation}

This in turn implies

\begin{equation}
W_n^{-1} (\frac{\hat{W}^{-2}}{W_n} - W_n^{-2}) W_n^{-1} \xrightarrow{P} 0
\end{equation}

and since $S_n^{1/2} W_n$ is bounded this gives (3.47). The left-hand side of (3.46) can be written as

\begin{equation}
W_n \sum_{t=1}^{n} X_t (\Sigma_{nt}^{-1} - \Sigma_t^{-1}) \xi_t + (W_n^{-1} \frac{\hat{W}^{-2}}{W_n} W_n^{-1} - I) W_n \sum_{t=1}^{n} X_t \Sigma_{nt}^{-1} \xi_t
\end{equation}

and so with (3.48), (3.49), (3.50) we see that (3.46) holds. Finally, (3.43) follows from $W_n \rightarrow 0$ and (3.44).
For the GLS-case we will only work out the estimator \( \hat{\Sigma}_{nt} = \hat{\Sigma} \)
given by (3.25). We have:

**Theorem 3.6. (conditions of theorem 3.1)**

Let \( \hat{\Sigma}_{nt} = \hat{\Sigma} \) be given by (3.25). Then the conditions (3.41), (3.42) hold provided that (3.26) is true. Hence, if (3.20), (3.21), (3.26) are satisfied then (3.43), (3.44), (3.45) hold.

**Proof.** We have to verify the conditions (3.41), (3.42). Therefore it is sufficient to prove

\[
(3.51) \quad S_{ni}^{-1/2} \sum_{t=1}^{n} X_t (\hat{\Sigma}_{nt}^{-1} - \Sigma_t^{-1}) E_t \xrightarrow{p} 0
\]

\[
(3.52) \quad S_{ni}^{-1/2} \left\{ \sum_{t=1}^{n} X_t (\hat{\Sigma}_{nt}^{-1} - \Sigma_t^{-1}) X_t' \right\} S_{ni}^{-1/2} \xrightarrow{p} 0
\]

\[
(3.53) \quad S_{ni}^{-1/2} \sum_{t=1}^{n} X_t (\hat{\Sigma}_{nt}^{-1} - \Sigma_t^{-1}) E_t \xrightarrow{p} 0
\]

\[
(3.54) \quad S_{ni}^{-1/2} \left\{ \sum_{t=1}^{n} X_t (\hat{\Sigma}_{nt}^{-1} - \Sigma_t^{-1}) X_t' \right\} S_{ni}^{-1/2} \xrightarrow{p} 0
\]

The \( i \)th-part of (3.51) equals

\[
S_{ni}^{-1/2} \sum_{j=1}^{n} \sum_{t=1}^{n} x_{ti} (\hat{\Sigma}_{nt}^{-1} - \Sigma_t^{-1})_{ij} E_t
\]

and the \((i,j)\)th-part of (3.52) equals

\[
S_{ni}^{-1} \left\{ \sum_{t=1}^{n} x_{ti} (\hat{\Sigma}_{nt}^{-1} - \Sigma_t^{-1})_{ij} x_{tj} \right\} S_{nj}^{-1/2}
\]

Similarly for (3.53) and (3.54). Therefore these conditions can be replaced by

\[
(3.55) \quad A_n = S_{ni}^{-1/2} \sum_{t=1}^{n} X_t (\hat{\Sigma}_{nt}^{-1} - \Sigma_t^{-1})_{ij} E_t \xrightarrow{p} 0
\]

\[
(3.56) \quad B_n = S_{ni}^{-1/2} \left\{ \sum_{t=1}^{n} X_t (\hat{\Sigma}_{nt}^{-1} - \Sigma_t^{-1})_{ij} x_{tj} \right\} S_{nj}^{-1/2} \xrightarrow{p} 0
\]
(3.57) \[ s^{-1/2}_{ni} \sum_{t=1}^{n} x_{ti} (\hat{\Sigma}^{-1}_n - \bar{\Sigma}_n^{-1})_{ij} \Sigma_{tj} \to 0 \]

(3.58) \[ s^{-1/2}_{ni} \{ \sum_{t=1}^{n} x_{ti} (\hat{\Sigma}^{-1}_n - \bar{\Sigma}_n^{-1})_{ij} x_{ti} S^{-1/2}_{nj} \to 0 \]

Since

\[ \Sigma_t^{-1} - \bar{\Sigma}_n^{-1} = \Sigma_t^{-1} (\bar{\Sigma}_n - \Sigma_t) \bar{\Sigma}_n^{-1} \]

and \( \Sigma_t^{-1}, \bar{\Sigma}_n^{-1} \) are bounded we have that

\[ |(\Sigma_t^{-1} - \bar{\Sigma}_n^{-1})_{ij}|^2 \leq C \text{tr}(\Sigma_t - \bar{\Sigma}_n)^2 \]

for some \( C > 0 \) and so with (3.26) we see that

\[ n \sum_{t=1}^{n} |(\Sigma_t^{-1} - \bar{\Sigma}_n^{-1})_{ij}|^2, n \sum_{t=1}^{n} |(\bar{\Sigma}_n^{-1} - \Sigma_t^{-1})_{ij}| \]

are bounded in \( n \). Therefore for some \( C > 0 \):

\[ E\{|\Sigma_n^{-1} \Sigma |^2 \} = \text{tr} E[\Sigma_n^{-1} \Sigma_n^2] = \text{tr}(S_n^{-1/2} \sum_{t=1}^{n} |(\Sigma_t^{-1} - \bar{\Sigma}_n^{-1})_{ij}|^2 (\Sigma_t^{-1} x_{ti} x_{ti}^t) S_n^{-1/2}) \]

\[ \leq C \max_{1 \leq t \leq n} x_{ti}^t S_n^{-1} x_{ti} \to 0, \quad n \to \infty \]

or \( A_n \to 0 \) and so (3.55) holds. Furthermore, for any \( c, d \in \mathbb{R}_{k_i} \) we have

for some \( C_1 > 0 \):

\[ |c' B_n d| = \sum_{t=1}^{n} (c' S_n^{-1/2} x_{ti})(\bar{\Sigma}_n^{-1} - \Sigma_t^{-1})_{ij}(d' x_{tj} S_n^{-1/2}) | \]

\[ \leq C_1 |c|^2 |d|^2 \{ \max_{1 \leq t \leq n} x_{ti}^t S_n^{-1} x_{ti}, \max_{1 \leq t \leq n} x_{tj}^t S_n^{-1} x_{tj} \}^{1/2} \to 0, \quad n \to \infty \]

and so \( B_n \to 0 \), proving (3.56).
The left-hand side of (3.57) can be written as
\[
(\hat{\Sigma}^{-1} - \overline{\overline{\Sigma}}^{-1})_{ij} \cdot \sum_{t=1}^{n} x_{t_i} e_{t_j}
\]

According to theorem 3.1, (3.16), we see that
\[
\Sigma_{n}^{-1} \sum_{t=1}^{n} x_{t} e_{t} \sim N_{k}(0, I)
\]
and therefore
\[
\Sigma_{n}^{-1/2} \sum_{t=1}^{n} x_{t} e_{t}
\]
is bounded in probability. Since \(\Sigma_{n}^{-1}\) and \(\Sigma_{n}^{-1}\) are bounded it follows from lemma 3.2 that
\[
\Sigma_{n}^{-1} \Sigma_{n} \overset{p}{\to} I
\]
and therefore also
\[
(3.59) \quad \Sigma_{n}^{-1} - \Sigma_{n}^{-1} \overset{p}{\to} 0
\]
Together this proves (3.57). Finally, the left-hand side of (3.58) equals
\[
(\hat{\Sigma}^{-1} - \overline{\overline{\Sigma}}^{-1})_{ij} \cdot \sum_{t=1}^{n} (\Sigma_{n}^{-1/2} x_{t_i})(\Sigma_{n}^{-1/2} x_{t_j})'
\]
Since
\[
\sum_{t=1}^{n} |\Sigma_{n}^{-1/2} x_{t_i}|^2 = \text{tr}(\sum_{t=1}^{n} \Sigma_{n}^{-1/2} x_{t_i} x_{t_i} S_{n}^{-1/2}) = k_i
\]
and similarly for the index \(j\), we see with Cauchy-Schwarz that
\[
\sum_{t=1}^{n} (\Sigma_{n}^{-1/2} x_{t_i})(\Sigma_{n}^{-1/2} x_{t_j})'
\]
is bounded in \(n\). With (3.59) this gives (3.58), completing the proof.
Stochastic explanatory variables

The model of this section as treated so far is a special case of the multivariate linear regression model with stochastic explanatory variables of the form

\[(3.60) \quad y_{ti} = x_{ti}' \beta_i + \varepsilon_{ti}, \quad t = 1, 2, \ldots; i = 1, \ldots, p.\]

where it is supposed that \(x_i = \{x_{ti}, t = 1, 2, \ldots; i = 1, 2, \ldots, p\}\), the process of the explanatory variables, is independent of \(\varepsilon_i = \{\varepsilon_{ti}, t = 1, 2, \ldots; i = 1, \ldots, p\}\), the error process.

The analysis of this model can immediately be reduced to the corresponding model with non-random explanatory variables because the conditional distribution of \(\varepsilon_i\) given \(x = x\) does not depend on \(x\) and is in fact the marginal distribution of \(\varepsilon_i\).

In particular, it is easy to see that all foregoing theorem maintain to hold if the conditions (3.11), (3.12) for the non-random explanatory variables are replaced by

\[(3.61) \quad r(x_{ni}) = k_i \text{ a.s.}, \text{ for all } i \text{ and some } n\]

\[(3.62) \quad \max_{1 \leq t \leq n} x_{ti}' S_{ni}^{-1} x_{ti} \overset{a.s.}{=} 0, \quad n \to \infty, \text{ for all } i\]

Note that these conditions deal separately with each equation. Therefore dependency of explanatory variables in different equations is allowed. This is important because in practice the same explanatory variable will often appear in more equations.

The question under which conditions (3.61), (3.62) hold can be transformed to the following problem:

**Problem.** Let \(x_1, x_2, \ldots\) be a sequence of random vectors in \(\mathbb{R}_k\). Find conditions in terms of the distributions of \(\{x_t\}\) under which

\[(3.63) \quad S_n \overset{a.s.}{=} \frac{1}{n} \sum_{t=1}^{n} x_t x_t' > 0 \text{ a.s. for some } n\]

\[(3.64) \quad \max_{1 \leq t \leq n} x_t' S_n^{-1} x_t \overset{a.s.}{=} 0, \quad n \to \infty\]
If the $x_t$ are i.i.d. (independent and identically distributed) with $E(x_t x_t') = \Gamma > 0$ then (3.63) and (3.64) are satisfied. For, the strong law of large numbers gives

$$\frac{1}{n} \sum_{t=1}^{n} x_t x_t' = \frac{1}{n} \sum_{t=1}^{n} x_t x_t' a_s. \Gamma > 0$$

Then we have also

$$\frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} |x_t| a_s. 0,$$

and together this implies (3.63), (3.64).

However, if the $x_t$ are not i.i.d. then the solution of this problem is not trivial. We will give a first step to the solution of the problem. (A definite solution will be a point for further research).

Suppose $\mu_t = E(x_t)$ exists and let $x_t = \mu_t + z_t$. In order to cover the non-random case as well it would be fine to have conditions for the expectation $\mu_t$ of the same type as (3.63), (3.64) for the random vectors $x_t$. The following theorem is appropriate:

**Theorem 3.7.** Suppose that

(3.65) $M_n : = \sum_{t=1}^{n} \mu_t \mu_t' > 0$ for some $n$

(3.66) $\max_{1 \leq t \leq n} \mu_t' M_n^{-1} \mu_t \rightarrow 0$, $n \rightarrow \infty$

Then (3.63), (3.64) hold provided that

(3.67) $M_n^{-1/2} (\sum_{t=1}^{n} \mu_t z_t') M_n^{-1/2} a_s. 0$, $n \rightarrow \infty$

(3.68) $z_t' M_n^{-1} z_t a_s. 0$, $n \rightarrow \infty$, for some $p = 0, 1, \ldots$

**Proof.** According to lemma 3.1 the conditions (3.65), (3.66) are equivalent to

(3.69) $M_n^{-1} \rightarrow 0$, $n \rightarrow \infty$
and (3.63), (3.62) are equivalent to

\[(3.71) \quad S_n^{-1} a.s. \rightarrow 0, \quad n \rightarrow \infty \]

\[(3.72) \quad x_{n+p} \rightarrow a.s. \Rightarrow 0, \quad n \rightarrow \infty \]

Therefore it suffices to prove (3.71), (3.72) from (3.67) - (3.70). We can write

\[S_n = \sum_{i=1}^{n} \left( \mu_t + \xi_t \right) \left( \mu_t + \xi_t \right)' = M_n^{1/2} \left( I + A_n \right) M_n^{1/2} + \sum_{i=1}^{n} \xi_t \xi_t', \]

where \(A_n\) is symmetric and defined by

\[A_n = M_n^{-1/2} \sum_{i=1}^{n} \left( \mu_t \xi_t + \mu_t \xi_t' \right) M_n^{-1/2}.\]

From (3.67) it follows that \(A_n \rightarrow 0\), a.s. Therefore, for some (random) \(n\) sufficiently large \(I + A_n > 0\) and so

\[S_n > M_n^{1/2} \left( I + A_n \right) M_n^{1/2}, \quad \text{a.s.} \]

or

\[(3.73) \quad S_n^{-1} \leq M_n^{-1/2} \left( I + A_n \right)^{-1} M_n^{1/2}, \quad \text{a.s.} \]

This gives with (3.69):

\[\|S_n^{-1}\| \leq \|M_n^{-1/2}\| \|I + A_n\|^{-1} \|M_n^{-1}\| \left( 1 - \|A_n\|^{-1} \right)^{-1} a.s. \rightarrow 0\]

and so (3.71) holds.

Furthermore, with (3.73) we have a.s.:

\[x_{n+p} \rightarrow \frac{x_{n+p}}{S_n} \leq \frac{x_{n+p} M_n^{-1/2} \left( I + A_n \right)^{-1} M_n^{-1/2} x_{n+p}}{x_{n+p}^2 \left( 1 - \|A_n\|^{-1} \right)}, \]

\[\leq \|M_n^{-1/2} \left( I + A_n \right)^{-1} M_n^{-1/2}\| x_{n+p}^2 \left( 1 - \|A_n\|^{-1} \right)^{-1} \]
and so (3.72) is proved if

\[ x_{n+p}^\prime \frac{M^{-1}}{n} x_{n+p} \overset{\text{a.s.}}{\to} 0 \]

However,

\[ x_{n+p}^\prime \frac{M^{-1}}{n} x_{n+p} = \mu_{n+p}^\prime \frac{M^{-1}}{n} \mu_{n+p} + \xi_{n+p}^\prime \frac{M^{-1}}{n} \xi_{n+p} + \]

\[ + \mu_{n+p}^\prime \frac{M^{-1}}{n} \xi_{n+p} + \xi_{n+p}^\prime \frac{M^{-1}}{n} \mu_{n+p} \]

and so this follows from (3.68), (3.70) and Cauchy-Schwarz.

We conclude this section with a theorem that gives sufficient conditions for (3.67), (3.68). Basic are the following two lemma's:

Lemma 3.2. Let \( \xi_1, \xi_2, \ldots \) be mutually independent random variables with \( E(\xi_n^2) < \infty \) and \( E(\xi_n) = 0 \) for all \( n = 1, 2, \ldots \), and let \( \{a_n\} \) be a sequence with \( 0 < a_n \downarrow 0, n \to \infty \). If

\[ \sum_{n=1}^{\infty} E a_n^2 E(\xi_n^2) < \infty \]

then

\[ \sum_{n=1}^{\infty} a_n \xi_n < \infty \text{ a.s.} \quad \text{and} \quad \sum_{n=1}^{\infty} \xi_t a_t^2 \overset{\text{a.s.}}{\to} 0 \]

Proof. See Feller (1966), VII.8, theorem 2 or Chung (1974), theorem 5.4.1, corollary, particular case (i).

Lemma 3.3. Let \( \{E_n\} \) be a sequence random vectors in \( \mathbb{R}^k \) and \( \{a_n\} \) a sequence with \( 0 < a_n \downarrow 0, n \to \infty \). If

(i) the distribution \( G_n \) has a density \( g_n \) with respect to some measure \( G \) such that

\[ g_n(x) \leq C_1, \quad x \in \mathbb{R}^k, \quad n = 1, 2, \ldots \]

\[ \int |x|^2 \, dG(x) < \infty \]
then

$$a_n |\xi_n|^2 \overset{a.s.}{\to} 0, \quad n \to \infty$$

Proof. The assertion holds if for all $\varepsilon > 0$ we have that $P\{a_n |\xi_n|^2 \geq \varepsilon \text{ i.o.}\} = 0$.

According to the Borel-Cantelli law it suffices to prove that

$$\sum_{n=1}^{\infty} P\{|\xi_n| > c_n\} < \infty,$$

where $c_n = \sqrt{\varepsilon/a_n}$. Note that $0 < c_n \to \infty$, $n \to \infty$.

We have:

$$\sum_{n=1}^{\infty} P\{|\xi_n| > c_n\} = \sum_{n=1}^{\infty} \sum_{j=n}^{\infty} P\{c_j \leq |\xi_n| < c_{j+1}\} = $$

$$= \sum_{j=1}^{\infty} \sum_{n=1}^{j} P\{c_j \leq |\xi_n| < c_{j+1}\} = \sum_{j=1}^{\infty} \sum_{n=1}^{j} \int_{c_j \leq |x| < c_{j+1}} dG_n(x)$$

$$\leq \sum_{j=1}^{\infty} \int_{c_j \leq |x| < c_{j+1}} dG(x) \leq \sum_{j=1}^{\infty} \int_{c_j \leq |x| < c_{j+1}} x^2 dG(x) =$$

$$\leq C_1 C_2 \sum_{j=1}^{\infty} \int_{c_j \leq |x| < c_{j+1}} x^2 dG(x) =$$

$$\leq \frac{1}{\varepsilon} C_1 C_2 \sum_{j=1}^{\infty} \int_{c_j \leq |x| < c_{j+1}} x^2 dG(x) \leq \frac{1}{\varepsilon} C_1 C_2 \int |x|^2 dG(x) < \infty,$$

concluding the proof.

Theorem 3.8. Let $\xi_1, \xi_2, \ldots$ be random vectors with $E\{\xi_t\} = 0$, $V\{\xi_t\} = S_t$ exists.

a) If $\xi_1, \xi_2, \ldots$ are independent with

$$\sum_{n=1}^{\infty} \lambda_{n}^{-2} (H_n, u_n, S_n, u_n) < \infty$$

then (3.6.7) holds.

b) If the distribution $F_n$ of $\xi_n$ has a density $f_n$ with respect to some measure $F$ such that $f_n(x)$ is bounded in $n,x$ and
\[ \int |x|^2 \, dF(x) < \infty \]

and if \( n \lambda_{\min}^{-1}(M_n) \) is bounded in \( n \) then (3.68) holds.

**Proof.**

a) Follows with lemma 3.2 from

\[ \|M_n^{-1/2} (\sum_{t=1}^{n} \nu_t \xi_t') M_n^{-1/2}\| \leq C. \|M_n^{-1}\| \|\sum_{t=1}^{n} \nu_t \xi_t\| \]

by taking \( \xi_t' = \nu_t' \xi_t \) and \( a_n = \|M_n^{-1}\| = \lambda_{\min}^{-1}(M_n) \)

b) Follows with lemma 3.3. from

\[ \|\xi_n' M_n^{-1} \xi_n\| \leq \|M_n^{-1}\| |\xi_n|^2. \]

**Corollary.** From lemma 3.1, theorem 3.7 and theorem 3.8 we get immediately that (3.63), (3.64) are satisfied under the following conditions:

\[ |\mu_n|^2 \lambda_{\min}^{-1}(M_n) \to 0 \]

\[ |n \lambda_{\min}^{-1}(M_n)| \text{ bounded in } n \]

\[ \sum_{n=1}^{\infty} \lambda_{\min}^{-2}(M_n) |\mu_n|^2 \|S_n\| < \infty \]

with \( \xi_1, \xi_2, \ldots \) independent with \( E(\xi_t) = 0, \sigma(\xi_t) = S_t \) and such that the distribution \( F_n \) of \( \xi_n \) has a density \( f_n \) with respect to some measure \( F \) with

\[ \int |x|^2 \, dF(x) < \infty \]

and \( f_n \) uniformly bounded.

**Remark.** In most applications we will have \( dF(x) = g(x) \, du(x) \) with \( u \) Lebesgue measure (continuous case) or counting measure (discrete case). If \( g_n \) is the density of \( \xi_n \) with respect to \( u \) then \( g_n(x) \leq g(x) \) for all \( n \) is equivalent to the condition that \( f_n \) is uniformly bounded.
4. Multivariate linear regression with dependent errors

We consider the same system of linear equations as in section 3. However, we will no longer assume that $\varepsilon_1, \varepsilon_2, \ldots$ are mutually independent. In stead of this we suppose that the $\varepsilon_t$ are generated by a moving average of the form

$$
\varepsilon_t = \sum_{h=-\infty}^{\infty} c_h n_{t-h} = \sum_{j=-\infty}^{\infty} c_{t-j} n_j
$$

Here, it is supposed that the $p$-vectors $n_j$ are mutually independent with $E[n_j] = 0$, $\Sigma_j = V[n_j]$ exists for all $j$. Under certain regularity conditions the series defining $\varepsilon_t$ will converge in mean square. Let $G_j$ denote the distribution function of $n_j$.

We are interested in asymptotic normality of the OLS-estimator $b_n$ for $\beta$. If $S_n$ is non-singular then

$$
b_n - \beta = S_n^{-1} \sum_{t=1}^{n} X_t \varepsilon_t = S_n^{-1} \sum_{j=-\infty}^{\infty} \left( \sum_{t=1}^{n} X_t c_{t-j} \right) n_j
$$

From (4.2) we see that again theorem 2.1 can be applied if we take $\varepsilon_{nj} = n_j$ and $A_{nj} = S_n^{-1} B_{nj}$ where

$$
B_{nj} = \sum_{t=1}^{n} X_t c_{t-j}
$$

Then

$$
V_n^2 = V(b_n) = S_n^{-1} \left( \sum_{j=1}^{n} B_{nj} B_{nj}' \right) S_n^{-1}
$$

The conditions (2.24), (2.25) for the distributions of the $\{n_j\}$ become

$$
m = \inf \lambda_{\text{min}}(\Sigma_j) > 0
$$

$$
\sup_{j} \int_{|z| \geq \delta} |z|^2 dG_j(z) \to 0 \quad , \quad \delta \to \infty
$$

implying
(4.7) \[ M: = \sup_{j} \lambda_{\max} (\Sigma_{j}) < \infty \]

It is reasonable and appropriate to impose the same conditions on the explanatory variables as in section 3. So we assume that (3.11) and (3.12) are satisfied.

Finally we have to assume some regularity conditions with respect to the coefficient matrices \( C_{h} \). If (4.7) holds then convergence in mean square in (4.1) is guaranteed if

\[ \sum_{h=-\infty}^{\infty} \text{tr}(C_{h} C_{h}^{'}) < \infty \]

Then

\[ C(\lambda) = \sum_{h=-\infty}^{\infty} C_{h} e^{i\lambda h} \]

converges in mean square with respect to Lebesgue measure. We will need the conditions

\[ a_{1} = \text{ess} \inf_{\lambda} \lambda \min (C(\lambda)C^{*}(\lambda)) > 0 \]

\[ a_{u} = \text{ess} \sup_{\lambda} \lambda \max (C(\lambda)C^{*}(\lambda)) < \infty \]

With these conditions it turns out that theorem 2.1 and its remark can be applied:

**Theorem 4.1.** If the distributions of the \( \{x_{ni}\} \) satisfy (4.5), (4.6), the coefficient matrices \( \{C_{h}\} \) the conditions (4.8), (4.10), (4.11) and the explanatory variables \( \{x_{ti}\} \) the conditions (3.11), (3.12) then

\[ b_{n} \xrightarrow{p} \beta, \quad n \to \infty \]

\[ V_{n}^{-1}(b_{n} - \beta) \xrightarrow{L} N_{k}(0, I), \quad n \to \infty \]

**Proof.** Set for \( \lambda \in (-\infty, \infty) \):

\[ X_{n}(\lambda) = \sum_{t=1}^{n} X_{t} e^{it\lambda} \]


\[ B_n(\lambda) = \sum_{j=-\infty}^{\infty} B_{nj} e^{ij\lambda} \]  
(convergence in mean square)

Since the sequence \( \{B_{nj}\} \) is the convolution of the (finite) sequence \( \{X_t\} \) and the sequence \( \{C_{-j}\} \) we have

\[ B_n(\lambda) = X_n(\lambda) C(-\lambda) \]

Therefore, with the Parseval-relation:

\[ \sum_j B_{nj} B_{nj}^* = \frac{1}{2\pi} \int_{-\pi}^{\pi} B_n(\lambda) B_n^*(\lambda) d\lambda = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_n(\lambda) C(-\lambda) C^*(-\lambda) X_n^*(\lambda) d\lambda, \]

and with (4.10), (4.11) and

\[ S_n = \sum_t X_t X_t^* = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_n(\lambda) X_n^*(\lambda) d\lambda \]

this gives

\[ (4.14) \quad \alpha_1 S_n \leq \sum_j B_{nj} B_{nj}^* \leq \alpha_u S_n \]

Hence, for

\[ Z_n = \sum_j A_{nj} A_{nj}^* = S^{-1}(\sum_j B_{nj} B_{nj}^*) S^{-1} \]

we have that

\[ (4.15) \quad \alpha_1 S_n^{-1} \leq Z_n \leq \alpha_u S_n^{-1} \]

proving the conditions (2.28) that \( Z_n \) is non-singular and the condition (2.41) that \( Z_n \to 0, n \to \infty. \)

Furthermore,

\[ A_{nj}^* Z_n^{-1} A_{nj} = B_{nj}^* (\sum_j B_{nj} B_{nj}^*)^{-1} B_{nj} \leq \frac{1}{\alpha_1} B_{nj} S^{-1} B_{nj} \]

and so the condition (2.29) is fulfilled if we can prove that
(4.16) \[ \sup_j \text{tr}(B_j^t S_n^{-1} B_n j) \to 0 \quad , \quad n \to \infty \]

We have
\[
B_j^t S_n^{-1} B_n j = \sum_t \sum_s X_t^t S_n^{-1} X_s C_s j
\]
and so for any \( a \in \mathbb{R}_p \):
\[
a^t B_j^t S_n^{-1} B_n j a = \left| \sum_t S_n^{-1/2} X_t^t C_{t-j} a \right|^2
\]
Since \( \text{tr}(C_j C_j^i) \to 0 \) if \( j \to \pm \infty \) there exists a \( j_n \) such that
\[
\sup_j \left| \sum_t S_n^{-1/2} X_t^t C_{t-j} a \right| = \left| \sum_t S_n^{-1/2} X_t^t C_{t-j} a \right|
\]
and so
\[
\{ \sup_j a^t B_j^t S_n^{-1} B_n j a \}^{1/2} = \left| \sum_t S_n^{-1/2} X_t^t C_{t+j_n} C_t a \right|
\]
\[
\leq \sum_{t=1-j_n}^{n-j_n} \left| S_n^{-1/2} X_{t+j_n} C_t a \right| = \sum_{t=1-j_n}^{n-j_n} \left( a^t C_t^t X_{t+j_n} S_n^{-1} X_{t+j_n} C_t a \right)^{1/2}
\]
\[
\leq |a| \sum_{t=1-j_n}^{n-j_n} \{ \text{tr}(C_t^t C_t) \}^{1/2} \{ \text{tr}(X_{t+j_n}^t S_n^{-1} X_{t+j_n}) \}^{1/2}
\]
There exists a sequence \( m_n \) of positive integers such that \( m_n < n, m_n \to \infty \) and
\[
(4.17) \quad m_n^2 \max_{1 < t < n} \text{tr}(X_t^t S_n^{-1} X_t) \to 0 \quad , \quad n \to \infty
\]
Then we can split up the sum on the right-hand side of the last inequality according to
\[
\sum_{t=1-j_n}^{n-j_n} = \sum_{t=1-j_n}^{n-j_n} + \sum_{t=1-j_n}^{n-j_n} |t| < \lfloor m_n/2 \rfloor \quad |t| > \lfloor m_n/2 \rfloor
\]
The first part is dominated by
\[ m_n |a| \left( \sup_{j \leq j} \text{tr}(C_j C_j) \right)^{1/2} \left( \max_{1 \leq t \leq n} X_t^t S_n^{-1} X_n \right)^{1/2} \]
and this tends to zero because of (4.17).

With Cauchy-Schwarz we see that the second part is dominated by
\[ n-j \sum_{|t|>\lceil m_n/2 \rceil} |a| \text{tr}(C_t C_t) \cdot \sum_{t=1-j}^{n-j} \text{tr}(X_t^t X_t^t S_n^{-1} X_n X_n) = \]
\[ = k \cdot |a| \sum_{|t|>\lceil m_n/2 \rceil} \text{tr}(C_t C_t) \]
and this tends to zero with (4.8) since \( m_n \to \infty \). This proves (4.16) concluding the proof.

We can ask in the same way as this is done in section 3 if it is possible to estimate \( V_n \) by \( \tilde{V}_n \) such that \( \tilde{V}_n^{-1} (b_n - \beta) \to^d N_k(0, I) \). This the will be a point for further research. For the univariate case and identically distributed \( \{n_j\} \) work on this has been done in Eicker (1965).
References

Eicker, F. (1963\textsuperscript{a}) - Central limit theorems for families of sequences of random variables - Ann. Math. Stat. 34, 439-446.


Genugten, B.B. van der (1977) - Limit theorems for LS-estimators in linear regression models with independent errors, Series "Ter Discussie" nr. 77.054, Department of Econometrics, Catholic University, Tilburg.
In de Reeks ter Discussie zijn verschenen:

1. H. H. Tigelaar  Spectraalanalyse en stochastische lineaire differentievergelijkingen.  juni '75
2. J. P. C. Kleijnen  De rol van simulatie in de algemene econometrie.  juni '75
3. J. J. Kriens  A stratification procedure for typical auditing problems.  juni '75
4. L. R. J. Westermann  On bounds for Eigenvalues  juni '75
5. W. van Hulst  Investment/financial planning with endogenous lifetimes: a heuristic approach to mixed integer programming.  juli '75
J. Th. van Lieshout  Distribution of errors among input and output variables.  augustus '75
6. M. H. C. Paardekooper  Design and analysis of simulation practical statistical techniques.  september '75
7. J. P. C. Kleijnen  Accountantscontrole met behulp van steekproeven.  september '75
8. J. Kriens  A note on the regula falsi  november '75
9. L. R. J. Westermann  Analoge simulatie van economische modellen.  november '75
10. B. C. J. van Velthoven  Het economisch nut van nauwkeurige informatie: simulatie van onder- nemingsbeslissingen en informatie.  december '75
11. J. P. C. Kleijnen  Theory change, incompatibility and non-deductibility.  januari '76
12. F. J. Vandamme  De arbeidswaardeleer onderbouwd?  februari '76
13. A. van Schaik  Input-outputanalyse en gelaagde planning.  februari '76
14. J. van Lieshout  Robustness of multiple ranking procedures: a Monte Carlo experiment illustrating design and analysis techniques.  februari '76
J. Roemen  Computers and operations research: a survey.  april '76
15. J. P. C. Kleijnen  Statistical problems in the simulation of computer systems.  mei '76
16. J. P. C. Kleijnen  Towards a more natural deontic logic.  juni '76
17. J. P. C. Kleijnen  Design and analysis of simulation: practical, statistical techniques.  juli '76
18. F. J. Vandamme  Identifiability in models with lagged variables.  augustus '76
19. J. P. C. Kleijnen  Quantile estimation in regenerative simulation: a case study.  september '76
20. H. H. Tigelaar  Principles of Economics for computers.  augustus '76
21. J. P. C. Kleijnen  Inleiding tot econometrische modellen van landen van de E.E.G.  september '76
22. W. Derks  Econometrisch model van België.  augustus '76
23. B. Diedereren  Principele van de economische modellen.  augustus '76
Th. Reijis  van Velthoven  W. Derks  Hybride simulaties van economische modellen.
<table>
<thead>
<tr>
<th>Number</th>
<th>Author</th>
<th>Title</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>26.</td>
<td>F. Cole</td>
<td>Forecasting by exponential smoothing, the Box and Jenkins procedure and spectral analysis. A simulation study.</td>
<td>september '76</td>
</tr>
<tr>
<td>27.</td>
<td>R. Heuts</td>
<td>Some reformulations and extensions in the univariate Box-Jenkins time series analysis.</td>
<td>juli '76</td>
</tr>
<tr>
<td>28.</td>
<td>W. Derks</td>
<td>Vier econometrische modellen.</td>
<td>oktober '76</td>
</tr>
<tr>
<td>29.</td>
<td>J. Frijns</td>
<td>Estimation methods for multivariate dynamic models.</td>
<td>oktober '76</td>
</tr>
<tr>
<td>30.</td>
<td>P. Meulendijks</td>
<td>Keynesianse theorieën van handelsliberalisatie.</td>
<td>oktober '76</td>
</tr>
<tr>
<td>31.</td>
<td>W. Derks</td>
<td>Structuuranalyse van econometrische modellen met behulp van Grafentheorie. Deel I: inleiding in de Grafentheorie.</td>
<td>september '76</td>
</tr>
<tr>
<td>32.</td>
<td>W. Derks</td>
<td>Structuuranalyse van econometrische modellen met behulp van Grafentheorie. Deel II: Formule van Mason. Een direct verband tussen economische veroudering en bezettingsgraadverliesen.</td>
<td>september '76</td>
</tr>
<tr>
<td>33.</td>
<td>A. van Schaik</td>
<td>Een direct verband tussen economische veroudering en bezettingsgraadverliesen.</td>
<td>september '76</td>
</tr>
<tr>
<td>34.</td>
<td>W. Derks</td>
<td>Structuuranalyse van Econometrische Modellen met behulp van Grafentheorie. Deel III. De graaf van dynamische modellen met één vertraging.</td>
<td>oktober '76</td>
</tr>
<tr>
<td>35.</td>
<td>W. Derks</td>
<td>Structuuranalyse van Econometrische Modellen met behulp van Grafentheorie. Deel IV. Formule van Mason en dynamische modellen met één vertraging.</td>
<td>oktober '76</td>
</tr>
<tr>
<td>36.</td>
<td>J. Roemen</td>
<td>De ontwikkeling van de omvangsverdeling in de levensmiddelenindustrie in de D.D.R.</td>
<td>oktober '76</td>
</tr>
<tr>
<td>37.</td>
<td>W. Derks</td>
<td>Structuuranalyse van Econometrische modellen met behulp van grafentheorie. Deel V. De graaf van dynamische modellen met meerdere vertragingen. Een direct verband tussen economische veroudering en bezettingsgraadverliesen.</td>
<td>oktober '76</td>
</tr>
<tr>
<td>38.</td>
<td>A. van Schaik</td>
<td>Een direct verband tussen economische veroudering en bezettingsgraadverliesen.</td>
<td>oktober '76</td>
</tr>
<tr>
<td>39.</td>
<td>W. Derks</td>
<td>Structuuranalyse van Econometrische modellen met behulp van Grafentheorie. Deel VI. Model I van Klein, statistisch.</td>
<td>december '76</td>
</tr>
<tr>
<td>40.</td>
<td>J. Kleijnen</td>
<td>Information Economics: Inleiding en kritiek</td>
<td>december '76</td>
</tr>
<tr>
<td>41.</td>
<td>M. v.d. Tillaart</td>
<td>De spectrale representatie van multivariate zwak-stationaire stochastische processen met discrete tijdp parameter.</td>
<td>november '76</td>
</tr>
<tr>
<td>42.</td>
<td>W. Groenendaal</td>
<td>Een econometrisch model van Engeland</td>
<td>december '76</td>
</tr>
<tr>
<td>43.</td>
<td>R. Heuts</td>
<td>Capital market models for portfolio selection</td>
<td>september '76</td>
</tr>
<tr>
<td>Number</td>
<td>Authors</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>44.</td>
<td>J. Kleijn en P. Rens</td>
<td>A critical analysis of IBM's inventory package impact.</td>
<td></td>
</tr>
<tr>
<td>45.</td>
<td>J. Kleijn en P. Rens</td>
<td>Computerized inventory management: A critical analysis of IBM's impact system.</td>
<td></td>
</tr>
<tr>
<td>46.</td>
<td>A. Willemstein</td>
<td>Evaluatie en foutenanalyse van econometrische modellen.</td>
<td></td>
</tr>
<tr>
<td>47.</td>
<td>W. Derks</td>
<td>Een identificatie methode voor een lineair discreet systeem met storingen op input, output en structuur.</td>
<td></td>
</tr>
<tr>
<td>48.</td>
<td>L. Westermann</td>
<td>Structuuranalyse van econometrische modellen met behulp van grafentheorie.</td>
<td></td>
</tr>
<tr>
<td>49.</td>
<td>W. Derks</td>
<td>Deel I. On systems of linear inequalities over R^n.</td>
<td></td>
</tr>
<tr>
<td>51.</td>
<td>J. Kleijn en P. Rens</td>
<td>Een econometrisch model van het Verenigd Koninkrijk.</td>
<td></td>
</tr>
<tr>
<td>52.</td>
<td>J.J.A. Moors</td>
<td>Estimation in truncated parameter-spaces.</td>
<td></td>
</tr>
<tr>
<td>53.</td>
<td>R.M.J. Heuts</td>
<td>Dynamic transfer function-noise modelling (Some theoretical considerations)</td>
<td></td>
</tr>
<tr>
<td>55.</td>
<td>P.A. Verheyen</td>
<td>Economische interpretatie in modellen betreffende levensduur van kapitaalgoederen.</td>
<td></td>
</tr>
<tr>
<td>56.</td>
<td>W.v.den Bogaard en J. Kleijn</td>
<td>Minimizing wasting times using priority classes.</td>
<td></td>
</tr>
<tr>
<td>57.</td>
<td>W. Derks</td>
<td>Structuuranalyse van Econometrische Modellen met behulp van Grafentheorie. Deel IX. Model van landen van de E.E.G.</td>
<td></td>
</tr>
<tr>
<td>58.</td>
<td>R. Heuts</td>
<td>Capital market models for portfolio selection (a revised version)</td>
<td></td>
</tr>
<tr>
<td>59.</td>
<td>A.P. Willemstein</td>
<td>Evaluatie en foutenanalyse van econometrische modellen. Deel II.</td>
<td></td>
</tr>
<tr>
<td>60.</td>
<td>Th. Dunnewijk en W. van Groenendaal</td>
<td>Het Model I van L.R. Klein.</td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>----------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>62.</td>
<td>A. Hempenius</td>
<td>Over een maat voor de juistheid van voorspellingen</td>
<td>aug.'77</td>
</tr>
<tr>
<td>63.</td>
<td>R.M.J. Heuts</td>
<td>Some reformulations and extensions in the univariate Box-Jenkins time series analysis approach (a revised version)</td>
<td>sept.'77</td>
</tr>
<tr>
<td>64.</td>
<td>R.M.J. Heuts</td>
<td>Applications of univariate time series modelling of U.S. monetary and business indicator data</td>
<td>sept.'77</td>
</tr>
<tr>
<td>65.</td>
<td>A. Hempenius en J. Frijns</td>
<td>Soorten van prijsheteroskedasticiteit in marktvraagfuncties.</td>
<td>okt.'77</td>
</tr>
<tr>
<td>66.</td>
<td>H.H. Tigelaar</td>
<td>Identifiability in Multiple Time Series</td>
<td>okt.'77</td>
</tr>
<tr>
<td>67.</td>
<td>J. van Lieshout en P. Verheyen</td>
<td>Levensduur in een jaargangenmodel</td>
<td>nov.'77</td>
</tr>
<tr>
<td>68.</td>
<td>Pieter J.F.G. Meulendijks</td>
<td>&quot;De macro-economische betekenis van geïnduceerde technische ontwikkeling: een meer-sektoren model met jaargangentheorie&quot;</td>
<td>nov.'77</td>
</tr>
<tr>
<td>69.</td>
<td>R.M.J. Heuts en P.J. Rens</td>
<td>A Monte Carlo study to obtain the percentage points of some goodness of fit tests in testing normality, when observations satisfy a certain low order ARMA-scheme</td>
<td>dec.'77</td>
</tr>
<tr>
<td>70.</td>
<td>J.C.P. Kleijnen</td>
<td>Generalizing Simulation Results through Metamodels</td>
<td>dec.'77</td>
</tr>
<tr>
<td>71.</td>
<td>Th. van de Klundert</td>
<td>Winstmaximalisatie in het Jaargangenmodel met vaste technische coefficients; een inventarisatie van de problematiek</td>
<td>dec.'77</td>
</tr>
</tbody>
</table>