CONSISTENCY OF PARAMETER-ESTIMATES IN A LINEAR TIME-SERIES MODEL

by

Friedhelm Eicker

University of North Carolina

This research was supported by the Office of Naval Research under Contract No. Nonr-855(06) for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States Government.
Consistency of Parameter-Estimates in a
Linear Time-Series Model

by

Friedhelm Eicker

University of North Carolina

1. Introduction and results.

The equation of linear regression under consideration is

\[ y_t = \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \beta_1 x_{1t} + \ldots + \beta_q x_{qt} + \varepsilon_t, \]

\[ t = 1, 2, 3, \ldots . \]

The quantities \( y_0, y_{-1}, \ldots, y_{-p} \) as well as the regression
variables are considered as given constants in a practical problem.

The errors \( \varepsilon_t \) are independently distributed and have mean 0.

Their second and fourth moments are assumed to be included between
two positive constants independent of \( t \). For the unknown constants
\( \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q \) consistent estimates are wanted.

It is assumed that the process described by (1) is stable, i.e.,
that all the roots are in modulus smaller than one in the equation

\[ \xi^p - \alpha_1 \xi^{p-1} - \ldots - \alpha_{p-1} \xi - \alpha_p = 0. \]

*On leave from University of Mainz.*

1. This research was supported by the Office of Naval Research under
Contract No. Nonr-855(06) for research in probability and statistics at
Chapel Hill. Reproduction in whole or in part is permitted for any pur-
pose of the United States Government.
The equations (1) can be described in matrix notation, writing for any \( N, \ t = 1, 2, \ldots, \ N \):

\[
\mathbf{y}_N = (y_1, y_2, \ldots, y_N)\quad \text{and} \quad \mathbf{\varepsilon}_N = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N)\;
\]

(3) \( \mathbf{x}_{iN} = (x_{i1}, x_{i2}, \ldots, x_{iN})' \), \( i = 1, 2, \ldots, q \),

\[
\mathbf{\beta} = (\beta_1, \beta_2, \ldots, \beta_q)'
\]

and as regression matrix

(4) \( \mathbf{x}_N = (\mathbf{x}_{1N}, \mathbf{x}_{2N}, \ldots, \mathbf{x}_{qN}) = (N \times q) \).

The subscript \( N \) on vectors and matrices will be omitted where confusion is unlikely; the prime denotes the transpose.

Equation (1) has been treated by several authors under different assumptions and with different methods. Mann and Wald (\( \sqrt{117} \), 1943) proved the consistency of the least squares estimates \( \mathbf{a} \) of \( \mathbf{a} = (a_1, a_2, \ldots, a_p)' \) in the special case \( q = 1 \), \( X' = (1, 1, \ldots, 1) \).

T. W. Anderson gives in his notes (\( \sqrt{117} \), 1949, p. 82) some conditions that assure consistency. For instance the disturbances \( \varepsilon_t \) are assumed to be identically and normally distributed and the matrix \( \mathbf{N} \), in the notes (188), in our notation:

(5) \( \lim_{N \to \infty} N^{-1} \mathbf{x}' \mathbf{x} \),

is assumed to exist. This implies

\( X_{iN}^2 = o(N) \). A similar assumption concerning the exogenous variables \( X_{iN} \) has been made in a paper by Koopmans, Rubin and Leipnik (\( \sqrt{107} \),
1950, p. 134, eq. 3.68). In our notation it is required that $x_{iN}^{\delta} I_{2N} = o(N)$; $L$ is given by (9) and $\delta = 0, 1, \ldots$. The existence of (5) is also assumed in a paper by T. W. Anderson and H. Rubin ($\sqrt{2}$, 1950, p. 572, Assumption F).

While all papers mentioned so far do not make use of spectral theory, this method leads to a number of conditions of a different nature for the exogenous variables. So for instance, Grenander ($\sqrt{3}$, 1954) and Grenander and Rosenblatt ($\sqrt{8}$, $\sqrt{9}$, 1956 and 1957) consider properties of estimates in linear stochastic difference equations under the conditions that

$$\phi(r) = \phi(r)^2 \rightarrow \infty \text{ as } N \rightarrow \infty, \phi(r) / \phi(r) \rightarrow 1 \text{ and}$$

$$\phi_n^{(s)} \frac{h(r)}{\phi_n^{(s)} / \phi_n^{(s)}}^{1/2} \text{ has a limit. Here } \phi_n^{(s)} =$$

$$(\phi_1^{(r)}, \phi_2^{(r)}, \ldots, \phi_n^{(r)}) \text{ is assumed to be real and equals in our notation } (C_{2N})^1, \text{ (see (13)). Both } r \text{ and } s \text{ run from 1 to } q \text{ and } h \text{ is any positive integer. A simple comparison with our conditions given below is not possible because in our model the } \phi \text{'s would involve parameters } \alpha_i \text{ which are assumed to be unknown. On the other hand a process, with autoregressive disturbances like (1) which is stable, can asymptotically also be considered as a stationary process, if the errors are identically distributed.}$$

The method and notation used in this paper is similar to that of Durbin ($\sqrt{4}$, $\sqrt{5}$, 1959 and 1960), who considers various models.
and their properties, applied in time series analysis.

Section 2 of this paper contains some lemmas helpful to prove the main theorem in Section 3, which states the basic conditions under which consistency can be assured. Some narrower conditions are derived from these, which may be easier to verify in some applications. It may be emphasized that these conditions show the existence of consistent least squares estimates also for some exponentially increasing regression vectors. This case seems not to have been handled so far. It is not included in the conditions of Grenander and Rosenblatt, which seem to be the broadest given up to now.

In Section 4 applications are made to polynomial, exponential, and trigonometric regression.

All results presented here are obtained by means of elementary theorems of matrix theory. It will be noted that all the proofs are simple and straightforward. This fact suggests that by refinement and extension of the method further results may be obtainable. Thus further asymptotic properties can be treated for the model studied here as well as for related models, e.g., asymptotic distributions. But also the explosive autoregression model, or considerations about the goodness of the asymptotic approximations for finite samples can be handled. Work in these directions is under way.
Let \( L = \begin{pmatrix} 0 & \ldots & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ & \ldots & \ldots \\ 0 & \ldots & 0 \end{pmatrix} \) (\( N \times N \)).

Then (11) can be written

\[
0 = \left( \sum_{y=0}^{\mathcal{D}} \alpha_y L^y \right) y + \hat{\beta} + x_0 + \varepsilon,
\]

understanding \( \alpha_0 \) as \( = -1 \). The solution with respect to \( y \) leads to

\[
y = C (\hat{\beta} + x_0 + \varepsilon),
\]

where

\[
C = -\left( \sum_{y=0}^{\mathcal{D}} \alpha_y L^y \right)^{-1} = \sum_{y=0}^{N-1} \mu_y L^y.
\]

The \( \mu_y \)'s are determined from \( CC^{-1} = I \).
The homogeneous equations are satisfied by the unique real solution

\[ \mu_v = \sum_{\xi_i} p_i(v) \xi_i^v, \quad v \geq p + 1, \]

where the summation is to be taken over all different roots \( \xi_i \) of (2). If \( m_i \) is the multiplicity of \( \xi_i \), then \( p_i(v) \) is an arbitrary polynomial of degree \( \leq m_i - 1 \) in \( v \). Together they involve \( p \) constants which are used to adapt the solution of (15) to that of (14). It is seen that \( \mu_i \) is a uniquely determined polynomial of degree \( i \) in the \( a_j \)'s for \( i, j = 1, 2, \ldots, p \). It follows further from our assumption of stability (all \( |\xi_i| \leq \xi < 1 \) that*

\[ |\mu_v| < \text{const} \xi^v \]

for all \( v \).

*The symbol \( \text{const} \) in mathematical expressions denotes a certain constant.
and \( \mu_v \to 0 \) as \( v \to \infty \).

This has the immediate consequence that \( C^2 \to 0 \) as \( N \to \infty \), i.e.,

the initial values \( y_0, \ldots, y_{t-p} \) are damped out and hence there is

no loss of generality in putting \( \Sigma = 0 \).

In passing we remark that addition of all the equations (14) and

(15) leads to

\[
\frac{1}{1 - a_1 - a_2 \ldots - a_p} = \sum_{v=0}^{\infty} \mu_v.
\]

The least squares estimates of \( z = (a_1, a_2, \ldots, a_p)' \) and \( \hat{e} \) are
denoted by \( \hat{a} \) and \( \hat{b} \) respectively. It is convenient to introduce

the matrix

\[
Y = (L^1, L^2, \ldots, L^P)\).
\]

Then the normal equations, derived from the minimum condition for \( \hat{e}^2 \),

namely \( \frac{\partial}{\partial a_i} (\hat{e}^2) = \frac{\partial}{\partial b_j} (\hat{e}^2) = 0 \); \( i = 1, 2, \ldots, p; j = 1, 2, \ldots, q \),

can be written for any \( N \)

\[
(20) \quad Y'Y = Y'a + Y'b
\]

\[
(21) \quad X'Y = X'a + X'b.
\]

Here (21) has been used in the form

\[
(22) \quad \hat{e} = y - Ya - Xb.
\]
Multiplying (22) by \( Y^\prime \) resp. \( X^\prime \) and adding leads to

\[
(23) \quad Y^\prime \hat{e} = Y^\prime Y(a - \gamma) + Y^\prime X(b - \beta),
\]

and

\[
(24) \quad X^\prime \hat{e} = X^\prime Y(a - \gamma) + X^\prime X(b - \beta).
\]

These two matrix equations are the subject of the lemmas in the next section.

As a matter of fact, rank \( X \) has to be assumed equal to \( q \), for almost all \( N \) otherwise the estimator \( \hat{b} \) is not even uniquely determined.

Abbreviating,

\[
X^\prime X = P_{(QxQ)},
\]

(24) can hence be written for all \( N > N_0 \), say,

\[
(26) \quad P^{-1}_Q X^\prime \hat{e} = P^{-1}_Q X^\prime Y(a - \gamma) + (b - \beta).
\]

In the next paragraph the consistency of the left sides and of the matrices of the coefficients of \( a - \gamma \), \( b - \beta \) in (26) and (23) (the latter after multiplication with a suitable factor) will be shown.

The left sides will be shown to converge in probability to \( 0 \). The minimum characteristic value of the coefficient matrix will be shown later on to be bounded away from \( 0 \) uniformly for \( N > N_0 \) under suitable conditions, which finally proves the consistency of \( a \) and \( b \).

2. Consistency of expressions in the normal equations.

For later purposes we need the following lemmas.
Lemma 1: \( \text{E}(P^{-1/2}X'_t \xi) = 0 \) \( \text{Var}(P^{-1/2}X'_t \xi) < \text{const} \) where the constant is independent of \( N \). It is assumed that \( Ee_t = 0 \) and \( E(e_t^2) < \text{const} \) for all \( t \).

Remark: This lemma applies also to ordinary linear regression without lagged variables \( (p = 0) \). The necessary and sufficient condition for the consistency of the least squares estimates of the regression parameters is easily seen to be \( \lambda_{\min}(P) \rightarrow \infty \).

Proof: (I) Clearly \( \text{E}(P^{-1/2}X'_t \xi) = P^{-1/2}X'_t \text{E}\xi = 0 \).

(II) The variance of each component of the vector \( P^{-1/2}X'_t \xi \) is bounded if and only if \( E(e'_tX'P^{-1}X_e) < \text{const} \). Because of the assumption \( E(e_t^2) < \text{const} \), that expression is essentially \( \text{tr} \ xP^{-1}X = \text{tr} \ I_q = q \), which is constant.

This completes the proof.

Remark:

If \( q \) is small, \( \lambda_{\min}(P) \) can be written down explicitly. The same may hold for larger \( q \)'s if \( P \) has a simple pattern or if something is known about its limiting matrix for \( N \rightarrow \infty \). Necessary conditions for \( \lambda_{\min}(P) \rightarrow \infty \) are for instance that \( x_i^2 \rightarrow \infty \) and

\[
\begin{vmatrix} |x_i| & |x_j| \\ |x_i| & |x_j| \end{vmatrix} \rightarrow \infty \text{ for all } i, j = 1, \ldots, q, i \neq j.
\]

Qualitatively speaking, \( \lambda_{\min}(P) \rightarrow \infty \) means that the columns of \( X \) tend in modulus to infinity and that they must not be at the same...
In order to find the asymptotic behavior of $P^{-1/2}X'Y$ it is helpful to introduce the following matrices. Let

$$z = C \varepsilon, \quad w = C \varepsilon \beta.$$  

Then

$$y = z + w, \quad z = C (L^2 \varepsilon, L^2 \varepsilon, \ldots, L^2 \varepsilon), \quad w = C (L^2 \varepsilon, L^2 \varepsilon, \ldots, L^2 \varepsilon),$$

and

$$y = z + w.$$

**Lemma 2:** $E(P^{-1/2}X'Y) = P^{-1/2}X'W; \text{ var } (P^{-1/2}X'Y) < \text{ const.}$

**Proof:** Clearly $E(P^{-1/2}X'Y) = P^{-1/2}X'W$. The variance of each matrix element is bounded if and only if the expectation of the modulus of each column vector is bounded. Now, we have, using $E(\varepsilon^2) < \text{ const}$ for all $\varepsilon$

that $E((L')^k C^1 XP^{-1}X' CLk) < \text{ const } \text{ tr}((L')^k C^1 XP^{-1}X' CLk)$

$< \text{ const } \lambda_{\text{max}}((L')^k C^1 XP^{-1}X' CLk) < \text{ const}$, for any $k$ out of 1, 2, $\ldots$, $p$. Here we use the fact that for any
square matrices $A$ and $B$, where $B$ is symmetric we have

$$\lambda_{\max} (A^TBA) < \lambda_{\max} (A^TA) \lambda_{\max} (B);$$

$$\lambda_{\min} (A^TBA) > \lambda_{\min} (A^TA) \lambda_{\min} (B).$$

Further $\lambda_{\max} (L^k(L^T)^k) = 1$ and lemma 3 is used:

**Lemma 3:** For any $N$

$$0 < \text{const} < \lambda (CC^T) < \text{const} < \infty.$$

**Proof:** There is a vector $x$ of modulus 1 such that

$$\lambda_{\max} (CC^T) = x^TCCx = \sum_{n,m=0}^{N-1} \mu_{nm} x^T (L^T)^n L^m x.$$

This has the upper bound

$$\text{const} \sum_{n,m=0}^{\infty} \xi^{nm} < \text{const},$$

by making use of (17) and the Schwartz inequality. Similarly

$$\frac{1}{\lambda_{\min} (CC^T)} = \lambda_{\max} (C^{-1}(C^T)^{-1}) < \left( \sum_{v=0}^{p} |a_v|^2 \right)^2.$$

As a consequence of lemmas 1 and 2 we have the

**Remark:** As $N \to \infty$ the set of the normal equations given by (26) tends to
\[(32) \quad 0 = P^{-1}\chi W(a - a) + b - \beta = 0.\]

(This of course does not yet state the existence of the p-limits of \( a \) and \( b \)).

We also have

\textbf{Lemma 4:} \( E(Y'_{\xi}) = 0 \). The variance of each component of this vector satisfies with a constant independent of \( N \) and \( i \) the inequality

\[ E((Y'(L')L\xi)) < \text{const} \cdot (\xi^2 + N) < \text{const} \cdot (N + \lambda_1(P)), \quad i = 1, 2, \ldots, p.\]

Here \( \lambda_1(P) \) is the largest characteristic root of \( P \).

\textbf{Proof:} The \( i \)th component of \( Y'_{\xi} \) is \( Y'(L')L\xi = \xi L W + \xi L C \).

The first term is linear in the \( \xi \)'s and hence has 0 expectation.

The second one is a sum of terms containing \( \xi L^k \xi \), \( k = 1, 2, \ldots \), which are free of pure quadratic terms \( \xi^2 \) and hence also have 0 expectation.

As to the variance, we make use of the uniform boundedness of \( E(\xi^2) \).

Hence \( E(\xi L W)^2 < \text{const} \cdot (\xi^2 + N) < \text{const} \cdot \xi^2 \). Further, \( \xi L W \xi L C \) is a cubic form all whose terms contain at least one \( \xi \) in the first power; hence the expectation is zero. Finally, in \( E(\xi L W)^2 \) only the squares of the terms of \( \xi L C \) give a contribution whose sum is smaller than a constant times the sum of squares of elements of \( L^C \), i.e., < \text{const} \cdot \text{tr} \cdot C'(L')L C < \text{const} \cdot \text{tr} \cdot C \cdot C = \text{const} \cdot \text{times} \)

\[ (N + (N-1)\mu_1^2 + (N-2)\mu_2^2 + \ldots) < \text{const} \cdot N(1 + \xi + \xi^2 + \ldots) < \text{const} \cdot N. \]
Lemma 5. $E(Y'X) = W'X$. The variance of each matrix element is smaller than $\text{const} \cdot \text{tr} \; P$ or $\text{const} \cdot \max_{k=1, \ldots, q} \frac{x_k^2}{x_k^2}$. The constants do not depend on $N$.

Proof: The variance of the $(j, k)$ element in $Z'X$ is $E(g^j(L')^jC_iX_k) < \text{const} \cdot (L_i^j)^2 (C_iX_k) < \text{const} \cdot \frac{x_k^2}{x_k^2}$, using lemma 3.

Lemma 6: $E(Y'Y) = E(Z'Z) + W'W$. For the $(j, k)$-th matrix element holds:

$$E(Z'Z)_{j,k} = \sum_{s=1}^{N} E(\varepsilon_s^2) \sum_{r=0}^{N-s} \mu_r \cdot \mu_{r+j-k}$$

whose modulus is $< \text{const} \cdot N$. The variance of the $(j, k)$-th element is

$$\text{var} \ (Y'Y)_{j,k} < \text{const} \cdot \left\{ (L_{\mu_k}^j)^2 + (L_{\mu}^j)^2 + N(L_{\mu_k}^j + L_{\mu_k}^j) + N \right\} < \text{const} \cdot \left\{ \mu_j^2 + N |\mu_k| + N \right\}$$

with constants independent of $N$, $j$, and $k$.

Proof: Any element in $Z'W$ has 0 expectation. This proves the first statement.

It is helpful to introduce $\mu_r$'s with negative indices and define them to be equal to 0. Then $(L^r C)_r,s = \mu_{r-s-n}$, any $n = 1, 2, \ldots$, and

$$E(Z'Z)_{j,k} = E(g^j(L')^jC_iX_k) = \sum_{s=1}^{N} E(\varepsilon_s^2) \sum_{r=1}^{N} \mu_{r-s-j} \mu_{r-s-k}$$
which proves the second statement. The modulus of the second sum is

\[ < \text{const} \xi |j-k| \sum_{v=0}^{\infty} \xi^{2v} < \text{const}. \]

Furthermore, \( \mathbb{E}(\xi^2) < \text{const} \) for all \( s \).

To prove the last statement it is convenient to write \( w_j \) for \( B_j \),

\[ z_j = I_j^{12}, \quad m_{j,k} = \mathbb{E}(Z_{j2})_{j,k}. \]

Then

\[ \text{var} (Y_j)_{j,k} = \mathbb{E}(z_j^{12} - m_{j,k})^2 = \]

\[ \mathbb{E} \left\{ (z_{j-k}^{12} - m_{j,k})^2 + (z_{j-k}^{12} - m_{j-k})^2 + 2(z_{j-k}^{12} - m_{j-k})(m_{j-k}^{12} - m_{j-k}) + 2 (w_j^{12} - z_{j-k}^{12}) \right\}. \]

The first term becomes

\[ \mathbb{E}(z_{j-k}^{12} - m_{j,k})^2 = \mathbb{E} \left\{ \sum_{r,s=1}^{N} \mu_{r-s-j} \mu_{r-s-k} (\xi^2 - \mathbb{E}(\xi^2)) + \right. \]

\[ + \sum_{r,s,t=1; s \neq t}^{N} \mu_{r-s-j} \mu_{r-t-k} \xi_s \xi_t \right\}^2 \]

\[ = \sum_{s=1}^{N} (\mathbb{E}(\xi^2) - (\mathbb{E}(\xi^2))^2) (\sum_{r=1}^{N} \mu_{r-s-j} \mu_{r-s-k})^2 + \]

\[ + \sum_{s \neq t=1}^{N} \{ \mathbb{E}(\xi^2) \mathbb{E}(\xi^2) \sum_{r=1}^{N} \mu_{r-s-j} \mu_{r-t-k} \sum_{u=1}^{N} \mu_{u-s-j} \mu_{u-t-k} \mu_{u-t} \mu_{u-j} \mu_{u-s-k} \} \]

\[ < \text{const} (N + \sum_{s>t} \xi^2(s-t)) < \text{const} N, \]

by making use of the assumed uniform boundedness of the first \( \mu \) moments.
For any positive integers $a, b, c, d$ it follows from the Schwartz inequality that
\[ E((v_{-b}^2 + c_d) = \frac{w^H L C \text{ diag} \{E(e_s^2)\} (N x N)}{c_l (L')} w_c, \]

the modulus being $< \text{const} |v_a| |v_d| < \text{const} (\frac{w^2}{v_a} + \frac{w^2}{v_d}).$

The remaining term contributes
\[ E((z_{-b}^2 - k_{-b}^2) (v_k^2 + v_j^2) z_{-d}^2) = \sum_{s=1}^{N} \sum_{r=1}^{N} \mu_{r-s-j} \mu_{r-s-k} \]
\[ \times \sum_{t=1}^{N} (w_k t_{t-s-j} + w_j t_{t-s-k}), \]

the modulus being $< \text{const} \left\{ \left| \frac{w_k}{k} \right| + \left| \frac{w_j}{j} \right| \right\}$

by using $(\sum_{t} w_k t_{t-s-j})^2 < \frac{w^2}{w_k} \text{const} \sum_{t} \xi_{2t}^2.$

Hence finally
\[ \text{var} (Y_i Y_j)_{j,k} < \text{const} \left\{ N \frac{w^2}{w_k} + \frac{w^2}{w_j} + N \left( \left| \frac{w_k}{k} \right| + \left| \frac{w_j}{j} \right| \right) \right\}, \]

which completes the proof.

3. The consistency of the estimates $a$ and $b$.

We now prove the basic

Theorem: The least squares estimates $a$ and $b$ of the parameters in

equation (1), where $y_0, y_{-1}, \ldots, y_{-p}, x_{1}; j = 1, 2, \ldots, p; t = 1, 2, \ldots$ are given constants, are consistent if the following
conditions are satisfied:

(A) All \( \varepsilon_t \)'s are independently distributed with 0 means and such that uniformly in \( t \) holds

\[
E(\varepsilon_t^2) > \text{const} > 0, \quad E(\varepsilon_t^1) < \text{const};
\]

(B) the modulus of the roots of (2) are smaller than 1;

(C1) if the set \( \{N\} \) of those \( N \) for which

\[
\lambda_1(\mathbf{P})/N^2 \geq g(N), \quad \text{with some } g(N) \to 0 \quad \text{for } N = 1, 2, \ldots,
\]

is infinite, then

\[
(40) \quad (N + N \sqrt{\lambda_1(\mathbf{P})} + \lambda_1(\mathbf{P}))^{-1/2} \quad \lambda_{\min}(K^TK) \to \infty, \quad N \in \{N\},
\]

\[K = (X, LX, \ldots, L^pX).\]

Here \( \lambda_1(\mathbf{P}) \) is the largest root of \( \mathbf{P} = X'X \).

(C2) If the complementary set \( \{\bar{N}\} \) to \( \{N\} \) is infinite, \( \lambda_{\min}(\mathbf{P}) \to \infty \) is sufficient for \( a \to \alpha \) i.p. on \( \{\bar{N}\} \). For\( b \to \beta \) on \( \{\bar{N}\} \) it is sufficient to have in addition that the elements of \( \mathbf{P}^{-1} X' L^j X \) are uniformly bounded for \( j = 1, 2, \ldots \).

Remark: The subdivision into conditions like (C1) and (C2) is desirable to prove the consistency in such cases where \( X \) is a constant vector for which (40) does not hold. Compare also the remark at the
of example (5) in Section 4.

Proof: Case \( \beta = 0 \). The equations

\[
\begin{align*}
(23) & \quad \begn{y_{i\varepsilon}} = \begn{y'_i y : y'_i x} \begn{a - \varepsilon : b - \beta} \\
(2h) & \quad \begn{x_{i\varepsilon}} = \begn{x'_i y : x'_i x} \begn{a - \varepsilon : b - \beta}
\end{align*}
\]

can be put into the form

\[
(4h) \quad f(n) \begn{N^{-1/2} y_{i\varepsilon} : \ldots} = \begn{N^{-1/2} y'_i y : \ldots} \begn{\ldots} \begn{f(n) N^{1/2} (a - \varepsilon) : \ldots} \begn{f(n) P^{1/2} (b - \beta)}
\]

where \( f(n) \to 0 \) as \( n \to \infty \).

The left side tends to the \( 0 \) vector i.p. according to lemmas 1 and \( 4 \), thereby using \( y = CX\beta = 0 \). We note by the way that \( y^2 = y^2 \leq \beta^2 \lambda_1(P) \lambda_1(C^t C) < \text{const } \lambda_1(P) |\beta| \).

Lemmas 2 and 6 indicate that the matrix in (4h) tends i.p. to

\[
\begn{N^{-1} E(z'z) : \ldots} = \begn{0 : \ldots} \begn{\ldots : I_q}
\]

thereby using \( W = 0 \). This matrix is symmetric, and it will be shown to have a positive minimum characteristic value \( \mu_{\min} \), say. If \( u \) is the unimodular characteristic vector to \( \lambda_{\min}(E(z'z)) \) and \( \varepsilon_t = (\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-p+1}) \), where \( \varepsilon_0 = \varepsilon_{-1} = \ldots = 0 \) then

\[
(4h5) \quad \lambda_{\min}(E(z'z)) = E(u'z'z) > \text{const } \sum_{t=1}^{N-1} E \{ (\varepsilon_t u)^2 \} > \text{const } N > 0.
\]
Here assumption (A) of the uniform boundedness of the var $\epsilon_t$ from below has been applied. Now

\[(46) \quad \mu_{\text{min}} = \min_{\lambda} (1, N^{-1} \lambda \{E(Z'Z)\} > \text{const} > 0).\]

If we now square (43) and let $N \to \infty$ then, using Slutsky's theorem (Cramer, 2nd Ed., p. 255), the left side tends to zero and the right side is greater than or equal to

\[\mu_{\text{min}}^2 \left\{ f^2 H(x-a)^2 + f^2 \int_{-\infty}^{\infty} (p-b)^2 \right\}.\]

If we choose $f$ slowly enough decreasing then because of (46)

\[(a-a)^2 \to 0, \text{i.e., } a \text{ is consistent and } E(\hat{a}) = a. \text{ Similarly } b \to \beta \text{ i.p., because with a suitable vector } \mathbf{x} \lambda_{\text{min}}(P) = \mathbf{x}' P \mathbf{x} = (\mathbf{y}', 0^2) K'K \left( \begin{array}{c} \mathbf{y} \\ 0 \end{array} \right) \geq \lambda_{\text{min}}(K'K), \text{ and hence with assumption (C1):}

\[f^2 \lambda_{\text{min}}(P) > \text{const } N^{1/2} f^2 > \text{const} > 0;\]

respectively from assumption (C2) follows for a suitable $f(N)$

\[f^2 \lambda_{\text{min}}(P) > \text{const} > 0.\]

Case $\beta \neq 0, \{N\}$ infinite.

\[(47) \quad \text{Let } \kappa = f(N) (N+1)^{-1/2} \lambda_1 (P) + \lambda_1 (P) - 1/2, f(N) \to 0 \text{ as } N \to \infty.\]
Squaring the normal equations, multiplied by \( \kappa \) gives

\[
\kappa^2 \varepsilon^t \begin{pmatrix} Y'Y & XX' \end{pmatrix} \varepsilon = \kappa^2 \begin{pmatrix} a' & b' \end{pmatrix} \begin{pmatrix} Y \kappa \end{pmatrix}^2 \begin{pmatrix} a' \ b' \end{pmatrix} = \kappa^2 \begin{pmatrix} (a,a') & (a,b') \end{pmatrix} \begin{pmatrix} a \ b \end{pmatrix} \begin{pmatrix} (a,a') & (a,b') \end{pmatrix}
\]

\[19\]

\( (M) \) \( (a-a)^2 + (b-b)^2 \); as the left side tends in probability to zero as can be seen from lemmas 1 and 4 and from Slutsky's theorem, the right side must do the same. \( M \) means the matrix

\[
M = M' = \begin{pmatrix} Y'Y & Y'X \\ X'Y & X'X \end{pmatrix}
\]

Again by Slutsky's theorem (Cramer [37], p. 255) one has

\[
k^2 \lambda_{\min} (M) \xrightarrow{\text{I.p.}} k^2 \lambda_{\min} \begin{pmatrix} E(Z'Z) + W'W & W'X \\ X'W & X'X \end{pmatrix}
\]

\[50\]

by using lemmas 5 and 6. The latter matrix can be written as

\[
\begin{pmatrix} W'W & W'X \\ X'W & X'X \end{pmatrix} = \begin{pmatrix} V' & X' \end{pmatrix} C G (V, C^{-1} X)
\]

where

\[
V = (LX_0, L^2X_0, \ldots, L^pX_0)
\]
With (11): \( C^{-1} = \sum_{v=0}^{\infty} a_v L^v \), it comes

\[
(v, C^{-1}x) = K \begin{pmatrix}
0 & \cdots & -I_q \\
\beta & 0 & a_1 I_q \\
& \beta & a_2 I_q \\
& & \ddots & \ddots \\
& & & \beta & a_p I_q 
\end{pmatrix}.
\]

The second matrix on the right is of the form \((p+1) q \times (p+q)\) and has full rank if and only if \( \beta \neq 0 \). We thus have with lemma 3 and a constant independent of \( N \)

\[
\lambda_{\min} \begin{pmatrix} W W & W X \\ X W & P \end{pmatrix} > \text{const} \ \lambda_{\min} (K'K).
\]

If we multiply with \( \kappa \) and choose \( f(N) \) appropriately, the last expression and therefore (50) remains above a positive constant. Hence in (48)

\[
a \rightarrow \mathbf{a}, \ b \rightarrow \mathbf{b} \ \text{i.p.}
\]

Case \( \lceil N \rceil \) infinite.

Lemmas 4 to 6 state that the normal equations (23) and (26) after multiplication with \( \kappa \) given in (47) tend for \( N \rightarrow \infty \) with probability:

one to the system of equations

\[
0 = \kappa (E (Z^T Z) + W W) (a - \mathbf{a}) + \kappa W X (b - \mathbf{b})
\]

(32) \( 0 = P^{-1} X W (a - \mathbf{a}) + b - \mathbf{b} \).

The second set can be used for elimination of \( b - \mathbf{b} \), thus obtaining
\[ 0 = \kappa \left( E(z'z) + W'(I_N^{-1}X'X)^{-1}W \right) (a - a) \]

\((I_N\) being the \(N\)th identity matrix).

Squaring gives

\[ \kappa^2 (a - a)^2 \left( E(z'z) + W'(I_N^{-1}X'X)^{-1}W \right)^2 (a - a) \geq \]

\[ (a - a)^2 \lambda_{\min} \left( \kappa E(z'z) + \kappa W'(I_N^{-1}X'X)^{-1}W \right) \geq \]

\[ (a - a)^2 \kappa^2 \lambda_{\min} (E(z'z)) \geq \text{const} \ (a - a)^2 \]

with a positive constant coming from \((46)\); further, as \(\lambda_1(P) N^{-2} < g(N)\),

\[ \kappa^2 n^2 > r^2(N) \left( \frac{1}{N} + \sqrt{g(N)} + g(N)^{-1} \right) \geq \text{const} > 0 \]

for a suitable \(f(N)\).- Thus follows \(a \rightarrow a\) if \(\lambda_{\min}(P) \rightarrow \infty\).

Further follows from \((C2)\) that for any fixed \(\beta\) the expressions

\[ P^{-1}X_iL_j^iX_\beta \]

are uniformly bounded vectors for all \(j\), hence because of \((17)\):

\[ P^{-1}X_iL_k^iX_\beta, \quad k = 1, 2, \ldots, p, \]

and \(P^{-1}X'W\) are bounded. Thus as \(a - a \rightarrow 0\) in \((32)\) also

\(b \rightarrow \beta\) i.p. This completes the proof.
Modified forms of conditions (C1) and (C2).

Let

\( P_{r,s}^{(r,s)} = \frac{x_r^t, N}{x_s, N-h} \) \]

\( D = \text{diag} \left( x_{1,N}, x_{2,N}, \ldots, x_{q,N-p} \right) \).

With a matrix \( R \) whose elements are of the form (55) we have

\[ KIK = DRD, \]

Hence:

\[ \lambda_{\min}(KIK) \geq \min_{i=1,\ldots,q} x_{i,N-p}^2 \lambda_{\min}(R). \]

Using \( \lambda_1(P) \leq \text{tr} P \) we may then replace assumption (C1) by the following stronger one on the set \( \{N\} \):

(C1)'  (a) \( \lim_{N \to \infty} P_{r,s}^{(r,s)} \) exists for all possible \( r, s, h, \)

(\( \beta \)) the limiting matrix \( R \) is nonsingular,

(\( \gamma \)) \[ \min_{i} x_{i,N-p}^2 \frac{(N+N \sqrt{\text{tr} P} + \text{tr} P)^{-1/2}}{\text{tr} P} \to \infty \]

or stronger:

\[ \min_{i} x_{i,N-p}^2 (\text{tr} P)^{-1/2} \to \infty \text{ plus } \text{tr} P \to \infty. \]
The conditions (a) and (b) are close to some conditions already known (see e.g. Grenander and Rosenblatt [9], p. 233), but condition (c) allows for a faster growth of the regression vectors than that admitted before.

Another set of sufficient conditions (Cl)', (C2)' can easily be derived using the Geršgorin method as in the remarks to lemma 1. These do not require the existence of any limit like (50), just like this existence is not required by the original assumption (Cl), (C2). So one has for instance (Cl)'

\[ (2 - \frac{\varepsilon}{s} \sum \frac{r(s)}{h_i N} \frac{x_{i,N-p}}{N + \text{tr } P})^{1/2} \to \infty \]

for all \( r, i = 1, \ldots, q \).

The following simple case may illuminate the necessity of a distinction between different rates of growth of \( \lambda_1(P) \) in some sort like that measured by (39). If \( q = 1 \) and \( x_{1,t} = \text{const} \) for all \( t \) (the case considered by Mann and Wald [11], 1943), (Cl) is not satisfied for any choice of \( g(N) \). But there is a sequence \( g(N) \) such that (C2) holds.

Condition (C2) holds if the following condition is satisfied

\[ (C2) \quad \lambda_1(P) \neq \lambda_{\min}(P) < \text{const} \text{ for all } N, \quad \lambda_{\min}(P) \to \infty. \]
Some applications.

Regression on polynomials.

(1) The question may be raised as to which powers in a regression on polynomials are allowed such that condition (Cl) of our theorem is fulfilled. Assume $x_{it} = t^{j_i}$, $i = 1, 2, \ldots, q$; $t = 1, 2, \ldots, N$; $j_1 > j_2 > \ldots > j_q$.

Then

$$\begin{pmatrix}
1^{j_1} & 1^{j_2} & \ldots & 1^{j_q} \\
2^{j_1} & 2^{j_2} & \ldots & 2^{j_q} \\
\vdots & \vdots & \ddots & \vdots \\
N^{j_1} & N^{j_2} & \ldots & N^{j_q}
\end{pmatrix}_{N \times q} $$

If we write $t^{(v)} = (1^v, 2^v, \ldots, (N-p)^v)$, $v = 0, 1, \ldots$, then $K^*$ which is obtained from $K = (X, IX, \ldots, IPX)$ by omission of the first $p$ rows and a simple reordering of the columns, can be written

$$K^* = \begin{pmatrix}
1^{j_1} & (1+1)^{j_1} & \ldots & (1+p)^{j_1} & 1^{j_2} & \ldots & \ldots & (1+p)^{j_q} \\
2^{j_1} & (2+1)^{j_1} & \ldots & (2+p)^{j_1} & 2^{j_2} & \ldots & \ldots & (2+p)^{j_q} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(N-p)^{j_1} & (N-p+1)^{j_1} & \ldots & (N-p)^{j_1} & (N-p)^{j_2} & \ldots & \ldots & (N-p)^{j_q}
\end{pmatrix}_{(N-p) \times (p+1)q}$$
\[
\begin{pmatrix}
1 & (j_1)_{10} & (j_1)_{20} & \cdots & (j_1)_{p0} & 0 & 0 & \cdots \\
0 & (j_{1-1})_{11} & (j_{1-1})_{21} & \cdots & (j_{1-1})_{p1} & & & \\
& & & & & & & \\
& & & & & & & \\
0 & (j_1)_{11} & (j_1)_{21} & \cdots & (j_1)_{p1} & 0 & (j_2)_{11} & \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
\cdots & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
\cdots & (j_2)_{p0} & \cdots & (j_3)_{p0} \\
\cdots & & & \\
\cdots & (j_2)_{j_2} & \cdots & (j_3)_{j_3} \\
\cdots & & & \\
\cdots & (0)_{p_2} & \cdots & (j_0)_{p_1} \\
\end{pmatrix}
\]

\[
(j_1+1)x(p+q).
\]
In (61) $\lambda_{\min}(K'K) \to \infty$ is required for consistency. As $|\lambda_{\min}(K'K)| < \text{const independent of } N$, for that requirement to be fulfilled it is necessary that $K^*$ has full rank, i.e., $(p+1)q$. In the factorization (60) of $K^*$ the first matrix for sufficiently large $N$ always contains a non-singular van der Monde determinant. For full rank of the second matrix it is obviously necessary that

$$j_q \geq (q+1-q)(p+1)-1, \quad q = 1, 2, \ldots, q.$$  

As $K^*$ has full rank if and only if this holds for the second matrix in (61) we may state:

**Lemma 7:** For rank $K^* = (p+1)q$ and for $\lambda_{\min}(K'K) \to \infty$ equations (61) must be satisfied.

We are unable to give sufficient conditions for the $j_q$ such that $K^*$ has full rank. The literature seems to offer little help for this problem. One may rewrite the problem observing that $K^*$ has full rank if and only if this is true for, say
which is of the form \((j_1+1-q) \times qp\) and contains a number of zeros in the lower rows. A similarity to a Wronskian is apparent. A Wronskian, submatrix of \(M^*\), could have been obtained also starting directly from \(K^*\) as a criteria for independence of the polynomials \(x^j_1, (x+1)^j_1, \ldots, (x+p)^j_1, x^j_2, \ldots, (x+p)^j_q\).

In some special cases one can determine rank \(M^*\):

(2) If \(p = 1, j_1 > j_2 > \ldots > j_q\) and if at least one of the following inequalities is true:
\[ j_1 - j_2 \geq q + 1, \quad j_2 - j_3 \geq q + 1, \ldots, \quad j_q - j_{q+1} \geq q + 1, \]
where
\[ j_{q+1} = 0, \]
then \( \mathbf{M}^* \) contains a non-singular submatrix according to general formulae (see e.g. Muir \( \textit{J}, 1930, \) p. 688, \( \S 739 \)).

With
\[ \sum_{t=1}^{N} t^n = \frac{N^{n+1}}{n+1} + \frac{N^n}{2} + o(N^n) = o(N^{n+1}) \]

one has \( \lambda_1(p) = o(N^{2j_p+1}) \) and \( \lambda_{\text{min}}(K^1K) = o(N^{2j_p-2q+1}) \),

if \( k \) is the smallest number out of 1, 2, ..., \( q \) such that

\[ j_p - j_{p+1} \geq q + 1. \]
Condition (C2) is hence satisfied if

\[ o\left(\frac{N^{2j_p}-2q}{j_p+1/2}\right) \rightarrow \infty \]
or

\[ j_p \geq q + j_p / 2 \]

One observes that necessarily \( j_p \geq 2q \).

(3) If \( q = 1 \) and \( j_1 \geq p \) one sees easily from (60) that

\[ \lambda_{\text{min}}(K^1K) \equiv o(N^{2(j_1-p)+1}). \]
(C2) is fulfilled if

\[ N^{2(j_1-p)+1} - j_1 - j_1 / 2 \rightarrow \infty, \text{ i.e. } j_1 \geq 2p. \]

(4) A general rule to obtain a lower bound for \( \lambda_{\text{min}}(K^1K) \) is given in
Lemma 8: \( \lambda_{\min}(K^TK) \) is \( o(N^{2S+1}) \) as \( N \to \infty \) where \( S \) is the smallest subscript occurring in the \( (pq \times pq) \) submatrix of \( M^* \) which is a) of full rank, and b) whose first row has the highest possible row number within \( M^* \). Thus condition (C1) holds for \( 2S > j_1 \).

For \( S \) holds the inequality \( S \leq \min \frac{j_2 - (q+1-r)(p+1)+1}{r=1, \ldots, q} \) say.

(5) If \( p = q = 2 \) one finds that in the inequality for \( S \) the equality sign holds. To prove this we distinguish between the cases:

\begin{itemize}
  \item[(a)] \( j_2 = 2 \)
  \item[(b)] \( j_2 = 3 \), (b1) \( j_1 \geq 6 \), (b2) \( j_1 = 5 \)
  \item[(c)] \( j_2 \geq 4 \), (1) \( j_1 = j_2+1 \), (2) \( j_1 = j_2+2 \), (3) \( j_1 \geq j_2+3 \).
\end{itemize}

Because of (61) we need only to consider \( j_1 \geq 5 \), \( j_2 \geq 2 \). There are always submatrices in \( M^* \) of full rank and in the required shape:

\[
\begin{pmatrix}
  1 & 16 & 1 & 16 \\
  i_{j_1} & 8j_1 & 2 & 16 \\
 (j_1)_3 & 2(j_1)_3 & 0 & 0 \\
 (j_1)_4 & (j_1)_4 & 0 & 0
\end{pmatrix} \neq 0, \quad S = 0 = S_{\max}.
\]
\[
\begin{array}{c|cccc}
(\beta_1) & j_1 & 6j_1 & 3 & 6 \\
(3j_1) & 8(3j_1) & 1 & 6 \\
(3j_1)_4 & 2(3j_1)_4 & 0 & 0 \\
(3j_1)_5 & (3j_1)_5 & 0 & 0 \\
\end{array}
\]

\(\neq 0, \ s = 1 = s_{\text{max}}.\)

\[
\begin{array}{c|cccc}
(\beta_2) & 1 & 16 & 1 & 4 \\
(3j_1) & 16 & 3 & 6 \\
(3j_1)_4 & 80 & 6 & 6 \\
(3j_1)_5 & 120 & 0 & 0 \\
\end{array}
\]

\(\neq 0, \ s = 0 = s_{\text{max}}.\)

\[
\begin{array}{c|cccc}
(\gamma_1) & j_2 & 8j_2 & 1 & 8 \\
(3j_2) & 8(3j_2) & 1 & 8 \\
(3j_2)_3 & 4(3j_2)_3 & 4 & 16 \\
(3j_2)_2 & 2(3j_2)_2 & 4 & 3 \\
(3j_2)_1 & (3j_2)_1 & 4 & 3 \\
\end{array}
\]

\(s = j_2 = s_{\text{max}}.\)

\[
\begin{array}{c|cccc}
(\gamma_2) & j_1 & 8j_1 & 7 & 1 \\
(3j_1) & 8(3j_1) & 7 & 1 \\
(3j_1)_3 & 4(3j_1)_3 & 7 & 1 \\
(3j_1)_4 & 2(3j_1)_4 & 7 & 1 \\
(3j_1)_5 & (3j_1)_5 & 7 & 1 \\
\end{array}
\]

\(s = j_1 = s_{\text{max}}.\)
From the preceding lemma follows that condition (Cl) is not fulfilled in cases (a) and (b) and thus consistency of the LS-estimates cannot be assured from our theorem in the given form. But condition (Cl) is fulfilled in cases (γ) if

\[ j_1 = j_2 + 1 \geq 10 \]

or

\[ j_1 = j_2 + 2 \geq 10 \]

or

\[ 2j_2 - 4 \geq j_1 \geq j_2 + 3 \geq 10 . \]

It is possible to continue this table for different values of p and q which are generally in statistical applications confined to small values. Unfortunately, the theorem in its present form requires considerably fast increasing regression vectors though the necessity of this is not always plausible. But one may expect that for certain combinations of j's no consistency can be assured. So one might at least conjecture that it is necessary that K has full rank.
Cases of exponential regression.

(6) If \( x_{it} = c_i^t \), \( c_i > 1 \), \( i = 1, 2, \ldots, q \), the matrix \( K^* \) formed as in (1) contains the submatrix

\[
\begin{pmatrix}
    c_i^p & c_i^{p-1} & \cdots & c_i \\
    c_i^N & c_i^{N-p} & \cdots & c_i \\
    \vdots & \vdots & \ddots & \vdots \\
    c_i^{N+p} & c_i^{N+p-1} & \cdots & c_i \\
\end{pmatrix}
\]

which has only rank 1 such that condition (C1) is not fulfilled. But if we consider for \( q = 1 \) \( x_t = t^n c_i^t \), \( t = 1, 2, \ldots \) the typical row of \( K^* \) is

(70) \( (t + p)^n c_i^{t+p}, (t + p - 1)^n c_i^{t+p-1}, \ldots, t^n c_i^t \).

\( K^* \) has a rank defect if there exists a linear combination between its columns, i.e. between the functions of \( t \) given in (70) for \( t = 1, 2, \ldots \). As these functions are analytic this condition is identical with the vanishing of the linear combination for all \( t \). Introducing similarly to (1)

\[
t^{(v)}_c = (1^v c^1, 2^v c^2, \ldots, N^v c^N),
\]

we have like in (60):

\[
K^* = (t^{(n)}_c c^t, t^{(n-1)}_c c^t, \ldots, t^{(0)}_c c^t)_{N \times (n+1)}
\]

\[
\begin{pmatrix}
    \binom{n}{0} p^n & \binom{n}{0}(p-1)^n & \cdots & 1 \\
    \binom{n}{1} p^{n-1} & \binom{n}{1}(p-1)^{n-1} & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    \binom{n}{n} & \binom{n}{n} & \cdots & 0
\end{pmatrix} \times \text{diag} \left( c_p, c_{p-1}, \ldots, 1 \right)
\]
where all three factor matrices, and hence $K^*$, have full rank if
$n \geq p$ (compare (3)). Now as for any $\nu$ approximately
\[ \sum_{t=1}^{N} t^\nu c_t \sim \text{const} \exp(N \log c) \]
we have condition (C1):
\[ \lambda_{\min}(K^TK) (\lambda_1(\nu))^{-1/2} = o \left( e^{\frac{1}{2} N \log c} \right) \rightarrow \infty \]
fulfilled if only $c > 1$ and $n \geq p$.

(7) We now consider $x_{it} = t^{n_i} c_i^t$, the integers $n_i \geq p$. If several
of the $c_i$ are equal but the $n_i$s are different the matrix $K$ contains
a submatrix in which only this particular $c_i$ occurs and which re-
sembles $K^*$ in example (1) apart from multiplication from the left
and the right with non-singular diagonal matrices. Hence for this
submatrix the considerations of (1) apply. Because of even greater
complications in $K$ we cannot make general statements.

If now $c_1 > c_2 > \ldots > c_q > 1$ is assumed ($n_i \geq p$), $K$ is seen to be
of full rank under utilization of (6), and
\[ \lambda_{\min}(K^TK) = o \left( c_q^{2N} \right) \]
With $\lambda_1(\nu) = o \left( c_1^{2N} \right)$, (C1) is seen to be fulfilled if
\[ c_q^2 > c_1 \]

(8) For a simultaneous regression on a polynomial and an exponential,
(C1) is not fulfilled. We have, with $c > 1$, $\lambda_{\min}(K^TK) < o \left( \text{poly.} \right.$
nominal in $N$), while $\lambda_1(\nu) = o \left( c_N \right)$.
Cases of trigometric regression.

Let \( x_{it} \) be \( \cos w_i t \) or \( \sin w_i t \). As long as no scale factors appear condition (C2) has to be fulfilled for consistancy.

(9) If \( q = 1 \) and \( w < 2\pi \), the scalar \( P \) is \( o(N) \). Also \( X^T L^4 X \) is at most \( o(N) \), hence (C2) is fulfilled, both for \( x_t = \cos w t \) or \( \sin w t \).

(10) If \( q = 2 \) and \( x_{it} = \cos w t, x_{2i, t} = \sin w t, w < 2\pi \), \( P \) tends to a diagonal matrix whose non-zero elements are \( o(N) \). The elements of \( X^T L^3 X \) are also at most \( o(N) \), hence (C2) is fulfilled. Here \( P \) as in (9) is unrestricted.

(11) If \( p \) and \( q \) are any positive integers, if further in

\[
x_{2i-1, t} = \cos w_i t, \quad x_{2i, t} = \sin w_i t
\]

never \( \bar{w}_i = 2\pi - w_j \) or \( w_i = w_j \) for \( i \neq j \) and all \( w_i < 2\pi \) then again \( P \) tends to a diagonal matrix whose non-zero elements are \( o(N) \), and again (C2) is fulfilled.

Acknowledgements.

The author is highly indebted to Professor James Durbin for drawing his interest and giving an introduction to the problems discussed above. He is also very grateful to Professor Harold Hotelling for continued encouragement and stimulation. He finally expresses his thanks to Dr. K. Murthy for hints to related literature, and to Mr. John Adams for going through the manuscript.
Literature:


