Title	Weak and strong consistency of Simple Least Squares Estimators in regression models and uniform strong consistency of residual sample spectral density in the error process
Sub Title	
Author	Toyooka, Yasuyuki
Publisher	慶応義塾大学工学部
Publication year	1978
Jtitle	Keio engineering reports Vol.31, No.10 (1978. 5) ,p.107- 121
JaLC DOI	
Abstract	Some conditions for weak and strong consistency of Simple Least Squares Estimators for regression parameters are given in the cases where a sequence of the error random variables is (i) uncorrelated, (ii) independent, (iii) independentlly identically Normally distributed and (iv) weakly stationary. A set of sufficient conditions for uniform strong consistency of a proposed estimator for the spectral density function of the strictly stationary and ergodic error process is given.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00310010-0107

KEIO ENGINEERING REPORTS
VOL. 31, NO. 10, pp. 107-121 1978

WEAK AND STRONG CONSISTENCY OF SIMPLE LEAST SQUARES ESTIMATORS IN REGRESSION MODELS AND UNIFORM STRONG CONSISTENCY OF RESIDUAL SAMPLE SPECTRAL DENSITY IN THE ERROR PROCESS

YASUYUKI TOYOOKA

Dept. of Administration Engineering, Keio University, Yokohama 223, Japan

(*Received Dec. 21, 1977*)

ABSTRACT

Some conditions for weak and strong consistency of Simple Least Squares Estimators for regression parameters are given in the cases where a sequence of the error random variables is (i) uncorrelated, (ii) independent, (iii) independentlly identically Normally distributed and (iv) weakly stationary. A set of sufficient conditions for uniform strong consistency of a proposed estimator for the spectral density function of the strictly stationary and ergodic error process is given.

1. Introduction

In many practical statistical procedures, we often use statistical linear model as an approximate model for the statistical phenomenon. Especially in the time domain analysis of the model which contains the mean functions (regression functions), we can reduce the original time series to the statistical linear model, $y_t = \boldsymbol{\beta}'\boldsymbol{x}_t + u_t, t=1, 2, \cdots$, where $\{y_t\}$ is the original time series, $\boldsymbol{\beta}$ a $p\times 1$ vector of unknown parameters, $\boldsymbol{x}_t$ a $p\times 1$ vector of explanatory fixed variates and $\{u_t\}$ the error stochastic process which is specified by many probabilistic properties.

This paper deals with two problems in time and frequency domain analysis of the time series.

The first is to give conditions for weak and strong consistency of Simple Least Squares Estimator $\hat{\boldsymbol{\beta}}_T$ for the regression parameter $\boldsymbol{\beta}$ in the case where a sequence of error random variables $\{u_t\}$ is (i) uncorrelated, (ii) independent, (iii) independently identically normally distributed and (iv) weakly stationary. Eicker, $F.$([8]) gave a necessary and sufficient condition for weak consistency of $\hat{\boldsymbol{\beta}}_T$ under the assumption that $\{u_t\}$ is a sequence of independently normally distributed random variables. We can get the same necessary and sufficient condition under the milder assumption of (i) than his. Anderson, T.W. and J.B. Taylor ([2]) gave a necessary and sufficient condition for strong consistency of $\hat{\boldsymbol{\beta}}_T$ by using a martingale limit theorem in the case of (iii). Here we give an another proof of the result and show the same result given in the case of (iii) is still valid under the case of (ii) if the parameter is one dimensional. While Hannan, E.J. ([9]) gave a set of sufficient conditions for strong consitency of one dimensional Simple Least Squares Estimator $\hat{\beta}_T$ in the case of (iv), we show that his result can be naturally extended to the multi-dimensional parameter case under the modified so-called Grenander's condition for explanatory variates.

The second is to discuss uniform strong consistency of the estimate of the spectral density function. Anderson, T.W. ([1]) discussed weak consistency of estimates of the spectral density function based on residuals from the mean function estimated by least squares without assuming the structure of $\{u_t\}$. In this paper, we give a set of sufficient conditions for uniform strong consistency of the residual sample spectral density function under the assumption that $\{u_t\}$ constitutes Autoregressive Process of given order. The residual sample spectral density function is constructed as a function of Simple Least Squares Estimators for the autoregressive parameters by using the information of the estimated residuals.

2. Conditions for Weak Consistency of $\hat{\beta}_T$ in the Case of Uncorrelated Error Random Variables

We shall treat the following statistical linear model,

$$y_t=\boldsymbol{\beta}'\boldsymbol{x}_t+u_t \qquad t=1,2,\cdots, \tag{2.1}$$

where $\{y_t\}$ is the original time series, $\boldsymbol{\beta}$ a $p\times 1$ vector of unknown parameters, $\boldsymbol{x}_t$ a $p\times 1$ vector of explanatory fixed variates and $\{u_t\}$ a sequence of random variables specified by some probabilistic properties. In this section, we assume that $\{u_t\}$ is a sequence of uncorrelated random variables and $A_p=\sum_{t=1}^{p}\boldsymbol{x}_t\boldsymbol{x}_t'$ is a nonsingular matrix.

At first, we shall consider the case where $\{u_t\}$ is a sequence of uncorrelated random variables with mean 0 and an equal variance σ^2.

Let us review the following procedure in an ordinary regression analysis. (Anderson, T.W. and J.B. Taylor ([2])) Let, for $j=1,2,\cdots,p$,

$$\boldsymbol{\beta}=\begin{bmatrix}\beta_j\\ \boldsymbol{\beta}^{(2)}\end{bmatrix},\ \boldsymbol{x}_t=\begin{bmatrix}x_{jt}\\ \boldsymbol{x}_t^{(2)}\end{bmatrix}$$

$$\hat{\boldsymbol{\beta}}_T=\begin{bmatrix}\hat{\beta}_{jT}\\ \hat{\boldsymbol{\beta}}_T^{(2)}\end{bmatrix},\ A_T=\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'=\begin{bmatrix}a_{jjT}\boldsymbol{a}_{12T}'\\ \boldsymbol{a}_{21T}\ A_{22T}\end{bmatrix},$$

where $\boldsymbol{\beta}^{(2)} = (\beta_1, \beta_2, \cdots, \beta_{j-1}, \beta_{j+1}, \cdots, \beta_p)'$, $\boldsymbol{x}_t^{(2)} = (x_{1t}, x_{2t}, \cdots, x_{j-1t}, x_{j+1t}, \cdots, x_{pt})'$, $\boldsymbol{a}'_{12T} = (a_{j1T}, a_{j2T}, \cdots, a_{jj-1T}, a_{jj+1T}, \cdots, a_{jpT})$, $\boldsymbol{a}_{12T} = (a_{1jT}, a_{2jT}, \cdots, a_{j-1jT}, a_{j+1jT}, \cdots, a_{pjT})'$ and

$$A_{22T}=\begin{bmatrix} a_{11T} & a_{12T} & \cdots & a_{1j-1T} & a_{1j+1T} & \cdots & a_{1pT} \\ a_{21T} & a_{22T} & \cdots & a_{2j-1T} & a_{2j+1T} & \cdots & a_{2pT} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{j-11T} & a_{j-12T} & \cdots & a_{j-1j-1T} & a_{j-1j+1T} & \cdots & a_{j-1pT} \\ a_{j+11T} & a_{j+12T} & \cdots & a_{j+1j-1T} & a_{j+1j+1T} & \cdots & a_{j+1pT} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{p1T} & a_{p2T} & \cdots & a_{pj-1T} & a_{pj+1T} & \cdots & a_{ppT} \end{bmatrix}.$$

Then

$$\hat{\beta}_{jT}-\beta_j=\frac{Y_{jT}}{S_{jT}}.$$

where

$$Y_{jT}=\sum_{t=1}^{T}(x_{jt}-\boldsymbol{a}'_{12T}A_{22T}^{-1}\boldsymbol{x}_t)u_t, \tag{2.2}$$

$$S_{jT}=\sum_{t=1}^{T}(x_{jt}-\boldsymbol{a}'_{12T}A_{22T}^{-1}\boldsymbol{x}_t)^2 \tag{2.3}$$

for $j=1, 2, \cdots, p$. Define $\gamma_{j1}^2=S_{jp}$, $v_1=Y_{jp}/\gamma_{j1}$ and

$$\gamma_{jt}^2=S_{jt+p-1}-S_{jt+p-2},\ t=2, 3, \cdots$$

$$v_t=\frac{Y_{jt+p-1}-Y_{jt+p-2}}{\gamma_{jt}},\ t=2, 3, \cdots.$$

Then $v_1, v_2, \cdots$ constitute a sequence of uncorrelated random variables with mean zero and variance σ^2. And

$$\sum_{t=1}^{T-p+1}\gamma_{jT}^2=S_{jT}=\sum_{t=1}^{T}(x_{jT}-\boldsymbol{a}'_{12T}A_{22T}^{-1}\boldsymbol{x}_t)^2=a_{jjT}-\boldsymbol{a}'_{12T}A_{22T}^{-1}\boldsymbol{a}_{21T}, \tag{2.4}$$

which is the reciprocal of the (j,j)-th element of A_T^{-1}. From the above discussion,

$$\hat{\beta}_{jT}-\beta_j=\frac{\sum_{t=1}^{T-p+1}\gamma_{jt}v_t}{\sum_{t=1}^{T-p+1}\gamma_{jt}^2}\quad \text{for } j=1, 2, \cdots, p. \tag{2.5}$$

We remark that $A_T^{-1}\to 0$ as $T\to\infty$ if and only if every diagonal elements of A_T^{-1} converge to 0 as $T\to\infty$ from the positive definiteness of A_T^{-1} if $T\geqq p$.

Then we have the next proposition with respect to a necessary and sufficient condition for weak consistency of Simple Least Squares Estimator $\hat{\boldsymbol{\beta}}_T$ for $\boldsymbol{\beta}$ by using the representation of (2.5).

THEOREM 1. Let $\{u_t\}$ be a sequence of uncorrelated random variables with an equal variance σ^2 and mean zero. And we assume that A_p is a nonsingular matrix. Then

$\hat{\boldsymbol{\beta}}_T \to \boldsymbol{\beta}$ in probability as $T \to \infty$ if and only if $A_T^{-1} \to 0$ as $T \to \infty$.

Proof. (Sufficiency) A sufficient condition for $\hat{\boldsymbol{\beta}}_T \to \boldsymbol{\beta}$ in probability as $T \to \infty$ is

$$\text{Cov}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}) = \sigma^2 A_T^{-1} \to 0 \text{ as } T \to \infty$$

from Chebychev's inequality. So we obtain that if $A_T^{-1} \to 0$ as $T \to \infty$ then $\hat{\boldsymbol{\beta}}_T \to \boldsymbol{\beta}$ in probability as $T \to \infty$.

(Necessity) The contrapositive proposition of the necessary part is the following; if there exists an (i, j)-th element such that $a_T^{ij} \not\to 0$ as $T \to \infty$, then there exists a k-th element such that $\hat{\beta}_{kT} - \beta_k \not\to 0$ in probability as $T \to \infty$, where a_T^{ij} denotes the (i, j)-th element of A_T^{-1}.

From the positive definiteness of A_T^{-1} for $T \geqq p$, the above proposition is equivalent to the following; if there exists a (j, j)-th element such that $a_T^{jj} \not\to 0$ as $T \to \infty$, then there exists a j-th element such that $\hat{\beta}_{jT} - \beta_j \not\to 0$ in probability as $T \to \infty$ by choosing j as k. This is equivalent to that

(i) if $a_T^{jj} \to K \neq 0$ as $T \to \infty$, then $\hat{\beta}_{jT} - \beta_j \not\to 0$ in probability as $T \to \infty$, or

(ii) if $\{a_T^{jj}\}$ is oscilating, then $\hat{\beta}_{jT} - \beta_j \not\to 0$ in probability as $T \to \infty$. On the other hand, from (2.5),

$$a_T^{jj} = \frac{1}{\sum_{t=1}^{T-p+1} \gamma_{jt}^2} \quad \text{for } T \geqq p, \tag{2.6}$$

so $\{a_T^{jj}, T = p, p+1, \cdots\}$ is a nonincreasing sequence with respect to T for all $j = 1, 2, \cdots, p$. Therefore there does not happen the situation of (ii).

Now we shall prove only the proposition (i). From (2.5), we shall investigate only

$$\hat{\beta}_{jT} - \beta_j = \frac{\sum_{t=1}^{T-p+1} \gamma_{tj} v_t}{\sum_{t=1}^{T-p+1} \gamma_{tj}^2} \not\to 0 \text{ in probability as } T \to \infty.$$

Generally if $\sum_{t=1}^{\infty} \gamma_{tj}^2 < \infty$, then $\sum_{t=1}^{T-p+1} \gamma_{tj} v_t$ converges in probability as $T \to \infty$. However

$$E\left\{\left(\frac{\sum_{t=1}^{\infty} \gamma_{tj} v_t}{\sum_{t=1}^{\infty} \gamma_{tj}^2}\right)^2\right\} = \frac{\sum_{t=1}^{\infty} \gamma_{tj}^2 \sigma^2}{\left(\sum_{t=1}^{\infty} \gamma_{tj}^2\right)^2} = \frac{\sigma^2}{\sum_{t=1}^{\infty} \gamma_{tj}^2} > 0.$$

So we can obtain the desired result. Q.E.D.

In (2.1), let $\{u_t\}$ be a sequence of uncorrelated random variables with variances which are bounded above and away from zero uniformly in t. Then we have the next theorem.

THEOREM 2. Let $\{u_t\}$ be a sequence of uncorrelated random variables with mean zero and variance σ_t^2 such that $0 < M_1 < \sigma_t^2 < M_2 < \infty$ for all $t = 1, 2, \cdots$, where

M_1 and M_2 are absolute constants and A_p be a nonsingular matrix.
Then

$$\hat{\boldsymbol{\beta}}_T \longrightarrow \boldsymbol{\beta} \text{ in probability as } T\to\infty \text{ if and only if } A_T^{-1}\to 0 \text{ as } T\to\infty.$$

Proof. By transforming the model (2.1) to

$$\frac{y_t}{\sigma_t}=\frac{\boldsymbol{\beta}'\boldsymbol{x}_t}{\sigma_t}+\frac{u_t}{\sigma_t} \quad t=1,2,\cdots,$$

the desired conditions are reduced to those in the case of a sequence of uncorrelated random variables with an equal variance.

Generally if A and B are the same order $(p\times p)$ square matrices such that $A\geqq B$ and $B>0$, where $A\geqq B$ denotes the positive semidefiniteness of $A-B$ and $B>0$ denotes the positive definiteness of B, then $B^{-1}\geqq A^{-1}>0$. Remark that

$$B_T=\sum_{t=1}^{T}\frac{1}{\sigma_t^2}\,\boldsymbol{x}_t\boldsymbol{x}_t'\geqq\frac{1}{M_2}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'=\frac{1}{M_2}\,A_T>0 \quad \text{for } T\geqq p.$$

By the above discussion, we can get

$$\left(\frac{1}{M_2}A_T\right)^{-1}\geqq B_T^{-1}.$$

So if $A_T^{-1}\to 0$ as $T\to\infty$, then $B_T^{-1}\to 0$ as $T\to\infty$.

Similarly by the fact that $B_T^{-1}\geqq((1/M_1)A_T)^{-1}$ for $T\geqq p$, if $B_T^{-1}\to 0$ as $T\to\infty$, then $A_T^{-1}\to 0$ as $T\to\infty$.

Therefore we can get the desired result. Q.E.D.

3. Conditions for Strong Consistency of $\hat{\boldsymbol{\beta}}_T$ in the Case of Independent Error Random Variables

In this section, at first we shall give an another simple proof of ANDERSON, T.W. and J.B. TAYLOR's results ([2]) for the necessary and sufficient condition for $\hat{\boldsymbol{\beta}}_T\to\boldsymbol{\beta}$ a.e. as $T\to\infty$ without using a martingale limit theorem in the case where $\{u_t\}$ is a sequence of independently identically normally distributed random variables. In the single parameter case, we shall prove that the necessary and sufficient condition for strong consistency of $\hat{\beta}_T$ is $\left(\sum_{t=1}^{T}x_t^2\right)^{-1}\longrightarrow 0$ as $T\to\infty$ without assuming identicality and normality of $\{u_t\}$. This result seems slightly different from Drygas's ([7]) in the point that we do not assume the boundedness of variances of $\{u_t\}$ away from zero.

Before proving the theorems, we shall refer the following two lemmata in the theory of the asymptotic behaviour of a sequence of sum of independent random variables without proof.

LEMMA 1. Let $\{u_t\}$ be a sequence of independent random variables on a basic probability space $(\Omega, \mathfrak{A}, P)$ such that $Eu_t=0$ and $Eu_t^2=\sigma_t^2<\infty$ for all $t=1, 2, \cdots$. Let $b_t(\geqq 0)$ converge up to ∞. Then if $\sum_{t=1}^{\infty}(\sigma_t^2/b_t^2)<\infty$, then $\sum_{j=1}^{T}u_t/b_T \to 0$ a.e. as $T\to\infty$.

Proof. (See. Breiman, L([4]))

LEMMA 2. Under the same assumptions on $\{u_t\}$ as those in Lemma 1, if $\sum_{t=1}^{\infty}\sigma_t^2<\infty$, $\sum_{t=1}^{T}u_t$ converges a.e. as $T\to\infty$.

Proof. (See. Lukacs, E ([12]))

THEOREM 3. Let A_p be nonsingular and $\{u_t\}$ be a sequence of independently identically normally distributed random variables. Then

$$\hat{\boldsymbol{\beta}}_T \to \boldsymbol{\beta} \text{ a.e. as } T\to\infty \text{ if and only if } A_T^{-1}\to 0 \text{ as } T\to\infty.$$

Proof. (Sufficiency) From the discussion of the section 2, $A_T^{-1}\to 0$ as $T\to\infty$ if and only if any diagonal elements of $A_T^{-1}\to 0$ as $T\to\infty$, and

$$\hat{\beta}_{jT}-\beta_j=\frac{\sum_{t=1}^{T-p+1}\gamma_{jt}v_t}{\sum_{t=1}^{T-p+1}\gamma_{jt}^2}, \tag{3.1}$$

where $\sum_{t=1}^{T-p+1}\gamma_{jt}^2$ is the reciprocal of the (j, j)-th element of A_T^{-1} and $\{v_t\}$ is a sequence of independently identically normally distributed random variables with mean zero and variance σ^2.

Setting $b_{jT}=\sum_{t=1}^{T}\gamma_{jt}^2$ in Lemma 1, a sufficient condition for $\hat{\beta}_{jT}-\beta_j\to 0$ a.e. as $T\to\infty$ for $j=1, 2, \cdots, p$ is

$$\sum_{T=1}^{\infty}\frac{\gamma_{jT}^2\sigma^2}{\left(\sum_{t=1}^{T}\gamma_{jt}^2\right)^2}<\infty \quad \text{for } j=1, 2, \cdots, p. \tag{3.2}$$

On the other hand,

$$\begin{aligned}\sum_{T=1}^{\infty}\frac{\gamma_{jT}^2}{\left(\sum_{t=1}^{T}\gamma_{jt}^2\right)^2} &\leqq 1+\sum_{T=2}^{\infty}\gamma_{jT}^2\left(\frac{1}{\sum_{t=1}^{T-1}\gamma_{jt}^2}-\frac{1}{\sum_{t=1}^{T}\gamma_{jt}^2}\right)\frac{1}{\gamma_{jT}^2}\\ &=1+\sum_{T=2}^{\infty}\left(\frac{1}{\sum_{t=1}^{T-1}\gamma_{tj}^2}-\frac{1}{\sum_{t=1}^{T}\gamma_{jt}^2}\right)\end{aligned} \tag{3.3}$$

for each $j=1, 2, \cdots, p$.

So the r.h.s. of (3.3) is convergent because $\sum_{t=1}^{T}\gamma_{jt}^2\to\infty$ as $T\to\infty$ for $j=1, 2, \cdots, p$.

Then we can obtain the desired result.

(Necessity) Similarly to the necessary part of Theorem 1., we can show the contrapositive proposition by using Lemma 2. From the representation of (3.1), it is only sufficient to prove that for some $j=1,2,\cdots,p$, if $\sum_{t=1}^{\infty}\gamma_{jt}^2=K(\neq 0)<\infty$,

$$\frac{\sum_{t=1}^{T-p+1}\gamma_{jt}v_t}{\sum_{t=1}^{T-p+1}\gamma_{jt}^2}\to L(\neq 0) \text{ a.s. as } T\to\infty.$$

By Lemma 2,

$\sum_{t=1}^{\infty}\gamma_{jt}v_t$ converges a.s. and consequently converges in $L_2(\Omega)$-norm.

So

$$\frac{\sum_{t=1}^{T-p+1}\gamma_{jt}v_t}{\sum_{t=1}^{T-p+1}\gamma_{jt}^2}\to\frac{\sum_{t=1}^{\infty}\gamma_{jt}v_t}{\sum_{t=1}^{\infty}\gamma_{jt}^2} \text{ a.s. as } T\to\infty \text{ and in } L_2(\Omega)\text{-norm as } T\to\infty,$$

while

$$E\left(\frac{\sum_{t=1}^{\infty}\gamma_{jt}v_t}{\sum_{t=1}^{\infty}\gamma_{jt}^2}\right)^2=\frac{\sigma^2}{\sum_{t=1}^{\infty}\gamma_{jt}^2}>0.$$

Hence we can get the desired result. Q.E.D.

In the case of $p=1$, it is not necessary to assume normality and identicality except for independence of a sequence of random variables $\{u_t\}$. ANDERSON, T. W. and J.B. TAYLOR ([2]) showed the sufficient part of the next theorem in the case where the error random variable is martingale difference.

THOREM 4. Let $\{u_t\}$ be a sequence of independent random variables with variances σ_t^2 bounded above uniformly in $t=1,2,\cdots$. And let assume that there exists at least one t such that $x_t\neq 0$, $t=1,2,\cdots$.

Then

$$\hat{\beta}_T\to\beta \text{ a.e. as } T\to\infty \text{ if and only if } \left(\sum_{t=1}^{T}x_t^2\right)^{-1}\to 0 \text{ as } T\to\infty.$$

Proof. The logical aspect of this theorem is the same as that of Theorem 3.

(Sufficiency) From the definition of $\hat{\boldsymbol{\beta}}_T$,

$\hat{\beta}_T-\beta=\frac{\sum_{t=1}^{T}x_tu_t}{\sum_{t=1}^{T}x_t^2}$. By using Lemma 1, a sufficient condition for $\hat{\beta}_T-\beta\to 0$ a.e. as

$T \to \infty$ is

$$\sum_{T=1}^{\infty} \frac{x_T^2}{\left(\sum_{t=1}^{T} x_t^2\right)^2} < \infty, \tag{3.4}$$

where we consider $\sum_{t=1}^{T} x_t^2$ as b_T in Lemma 1 and use the boundedness above of σ_t^2 uniformly in t. From the discussion of the sufficient part in Theorem 3, a sufficient condition for (3.4) is $\sum_{t=1}^{T} x_t^2 \to \infty$ as $T \to \infty$.

(Necessity) We shall prove the contrapositive proposition. By applying Lemma 2, if $\sum_{t=1}^{\infty} \mathrm{Var}(x_t u_t) < M \sum_{t=1}^{\infty} x_t^2 < \infty$, then $\sum_{t=1}^{T} x_t u_t$ converges a.e., where M is an absolute constant.

On the other hand,

$$E\left(\frac{\sum_{t=1}^{\infty} x_t u_t}{\sum_{t=1}^{\infty} x_t^2}\right)^2 = \frac{\sum_{t=1}^{\infty} x_t^2 \sigma_t^2}{\left(\sum_{t=1}^{\infty} x_t^2\right)^2} > 0.$$

This contradicts $\dfrac{\sum_{t=1}^{T} x_t u_t}{\sum_{t=1}^{T} x_t^2} \to 0$ a.e. as $T \to \infty$. Q.E.D.

4. Uniform Strong Consistency of the Sample Spectral Density Function Based on the Estimated Residuals

In this section we shall treat the model (2.1) in the case where the error random variable $\{u_t\}$ is a sequence of a weakly stationary stochastic process.

At first, we shall investigate a set of sufficient conditions for strong consistency of Simple Least Squares Estimator $\hat{\boldsymbol{\beta}}_T$ for $\boldsymbol{\beta}$ and make the estimated residuals $u_t^* = y_t - \tilde{\boldsymbol{\beta}}'_T \boldsymbol{x}_t$ $t = 1, 2, \cdots, T$, where $\tilde{\boldsymbol{\beta}}_T$ is a strongly consistent estimator for $\boldsymbol{\beta}$ (for example $\hat{\boldsymbol{\beta}}_T$).

Second, under the assumption that $\{u_t\}$ is an Autoregressive Process of given order K, we shall estimate the autoregressive parameters by the solutions of the system of Normal equations which is constructed by the information of the estimated residuals.

In the last, we shall show that the sample spectral density function which is given as the function of these solutions is uniformly strongly consistent for the true spectral density function under the some regularity conditions for $\{u_t\}$ and $\boldsymbol{x}_t$.

Now let us assume that $\{u_t\}$ is a real valued and weakly stationary stochastic process with mean zero and has a covariance function $\sigma(h)$, $h=0, \pm 1, \pm 2, \cdots$ and a spectral density function $f(\lambda)$.

Then we can get a set of sufficient conditions for strong consistency of Simple Least Squares Estimator $\hat{\boldsymbol{\beta}}_T$ for $\boldsymbol{\beta}$. In the case of $p=1$, HANNAN, E.J. ([10]) gave a set of sufficient conditions for this. But in the case of $p>1$, a set of these conditions is naturally extended under the modified Grenander's condition for the explanatory variates. The modified Grenander's condition reduces the multidimensional case to the single parameter case, so Hannan's method of proof can be straightforward applied.

THEOREM 5. Suppose that the following three conditions hold:

R.1 For any $i, j=1, 2, \cdots, p$, there exists $\lim_{T\to\infty} \frac{\sum_{t=1}^{T} x_{it}x_{jt}}{||x_i||_T ||x_j||_T}$ and we shall put this limit as ρ_{ij}, where $||x_i||_T=\left(\sum_{t=1}^{T} x_{it}^2\right)^{1/2}$ $i=1, 2, \cdots, p$. And R is a nonsingilar matrix, where $R=[\rho_{ij}, i, j=1, 2, \cdots, p]$ is a matrix of order $p\times p$.

R.2 $0<\liminf_{T\to\infty} \frac{\sum_{t=1}^{T} x_{it}^2}{T^\alpha} \leqq \limsup_{T\to\infty} \frac{\sum_{t=1}^{T} x_{it}^2}{T^\alpha} <\infty$ for $i=1, 2, \cdots, p$ and $\alpha>1/2$.

R.3 $0\leqq f(\lambda)<C<\infty$ for $-\pi\leqq\lambda\leqq\pi$, where C is an absolute constant.

Then Simple Least Squares Estimator $\hat{\boldsymbol{\beta}}_T$ is a strongly consistent estimator for $\boldsymbol{\beta}$.

Proof. From the definition of $\hat{\boldsymbol{\beta}}_T$,

$$\begin{aligned}\hat{\boldsymbol{\beta}}_T-\boldsymbol{\beta} &=\left(\sum_{t=1}^{T} \boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1} \sum_{t=1}^{T} \boldsymbol{x}_t u_t \\ &= D_T^{-1}\left(D_T^{-1} \sum_{t=1}^{T} \boldsymbol{x}_t\boldsymbol{x}_t' D_T^{-1}\right)^{-1} D_T^{-1} \sum_{t=1}^{T} \boldsymbol{x}_t u_t, \end{aligned} \tag{4.1}$$

where $D_T=\mathrm{diag}(||x_1||_T, ||x_2||_T, \cdots, ||x_p||_T)$.

By R.1,

$$\left(D_T^{-1} \sum_{t=1}^{T} \boldsymbol{x}_t\boldsymbol{x}_t' D_T^{-1}\right)^{-1} \to R^{-1} \text{ as } T\to\infty,$$

so, for a sufficient large T, $K_T^{-1}=\left(D_T^{-1} \sum_{t=1}^{T} \boldsymbol{x}_t\boldsymbol{x}_t' D_T^{-1}\right)^{-1}$ is a matrix of which all elements are bounded. Therefore the structure of the r.h.s. of (4.1) is

$$\left[\frac{k_T^{(i,j)}}{||x_i||_T||x_j||_T}\ ;\ i, j=1, 2, \cdots, p\right] \sum_{t=1}^{T} \boldsymbol{x}_t u_t, \tag{4.2}$$

where $k_T^{(i,j)}$ denotes the (i, j)-th element of a matrix K_T^{-1}. By R.2, the order of the increase of $||x_i||_T$ is the same as that of $||x_i||_T$ for all $i, j=1, 2, \cdots, p$.

From this fact we may only investigate the asymptotic behaviour of the scalor statistic $\frac{1}{T^\alpha} \sum_{t=1}^{T} x_t u_t$ in the almost everywhere sense.

Then the rest of the proof of this theorem is the same as that of HANNAN, E.J. ([10]). Q.E.D.

REMARK 1: In fact, this result is valid in the case where $\{u_t\}$ belongs to a class of nonstationary stochastic processes such that

$$u_t=\int_{-\pi}^{\pi} e^{it\lambda}c(t)dZ(\lambda),$$

where $c(t)$ is a bounded function of t and $Z(\lambda)$ is a stochastic process with orthogonal increments. But in the case of

$$u_t=\int_{-\pi}^{\pi} e^{it\lambda}c(t,\lambda)dZ(\lambda),$$

where $c(t,\lambda)\in L_2(dF(\lambda))$ for each t such that $dF(\lambda)=E|dZ(\lambda)|^2$, we can not prove this proposition because of the influence of the non-orthogonality of $c(t,\lambda)dZ(\lambda)$. (See. PRIESTLEY, M.B. ([13])).

It is meaningful to investigate a set of sufficient conditions for strong consistency of the residual sample mean because we shall not assume the ergodicity of the original error time series $\{u_t\}$.

Let define the estimated residuals as

$$u_t^*=y_t-\tilde{\boldsymbol{\beta}}_T'\boldsymbol{x}_t,\ t=1,2,\cdots,T,$$

where $\tilde{\boldsymbol{\beta}}_T$ (of course $\hat{\boldsymbol{\beta}}_T$ in Theorem 5) is a strongly consistent estimator for $\boldsymbol{\beta}$. The residual sample mean $\frac{1}{T}\sum_{t=1}^{T}u_t^*$ is naturally an estimator for $Eu_t=0$.

LEMMA 3. Suppose that the following three conditions hold:

1. All element of $\boldsymbol{x}_t=(x_{1t},x_{2t},\cdots,x_{pt})'$ are the bounded functions of t.
2. $\tilde{\boldsymbol{\beta}}_T$ is a strongly consistent estimator for $\boldsymbol{\beta}$.
3. There exists $K>0$ such that $\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\sigma(t-s)\leq\frac{K}{T^\alpha}$, where $\alpha>0$.

Then

$$\frac{1}{T}\sum_{t=1}^{T}u_t^*\to 0 \text{ a.e. as } T\to\infty.$$

Proof. The condition 3 for $\frac{1}{T}\sum_{t=1}^{T}u_t\to 0$ a.e. as $T\to\infty$ is given by DOOB, J.L. ([6]). Of course, if $f(\lambda)$ satisfies R.3 in Theorem 5, then the condition 3 is automatically satisfied.

Now

$$\overline{u_T^*}=\frac{1}{T}\sum_{t=1}^{T}u_t^*=\frac{1}{T}\sum_{t=1}^{T}(y_t-\tilde{\boldsymbol{\beta}}_T'\boldsymbol{x}_T)$$

$$=\frac{1}{T}\sum_{t=1}^{T}u_t+(\boldsymbol{\beta}-\tilde{\boldsymbol{\beta}}_T)'\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_t$$

$$=o(1)+o(1)0(1) \text{ as } T\to\infty$$

$$=o(1) \text{ as } T\to\infty,$$

where $o(1)$ denotes $\lim_{T\to\infty} o(1)=0$ a.e.. Q.E.D.

Next we shall consider a set of sufficient conditions for strong consistency of the sample autocovariance function based on the estimated residuals for the true autocovariance function. The sample autocovariance function based on the estimated residuals is defined as

$$C^*(h)=\frac{1}{T-h}\sum_{t=1}^{T-h}(u_t{}^*-\overline{u_T^*})(u_{t+h}^*-\overline{u_T^*}),$$

where $\overline{u_T^*}=\frac{1}{T}\sum_{t=1}^{T}u_t{}^*$.

LEMMA 4. Suppose that the following three conditions hold:

1. All elements of $\boldsymbol{x}_t=(x_{1t}, x_{2t}, \cdots, x_{pt})'$ are the bounded functions of t.
1. $\tilde{\boldsymbol{\beta}}_T$ is a strongly consistent estimator for $\boldsymbol{\beta}$.
3. $\{u_t\}$ is an ergodic and strictly stationary process.

Then

$$C^*(h)\to\sigma(h) \text{ a.e. as } T\to\infty \text{ for } h=0, \pm1, \pm2, \cdots.$$

Proof. By Condition 3, the asymptotic behaviour of $C^*(h)$ is equivalent to that of

$$C^{**}(h)=\frac{1}{T-h}\sum_{t=1}^{T-h}u_t{}^*u_{t+h}^*$$

in almost everywhere sense.

On the other hand,

$$C^{**}(h)=\frac{1}{T-h}\sum_{t=1}^{T-h}\{u_t+(\tilde{\boldsymbol{\beta}}_T-\boldsymbol{\beta})'\boldsymbol{x}_t\}\{u_{t+h}+(\tilde{\boldsymbol{\beta}}_T-\boldsymbol{\beta})'\boldsymbol{x}_{t+h}\}$$

$$=\frac{1}{T-h}\sum_{t=1}^{T-h}\{u_tu_{t+h}+(\tilde{\boldsymbol{\beta}}_T-\boldsymbol{\beta})'\boldsymbol{x}_tu_{t+h}+(\tilde{\boldsymbol{\beta}}_T-\boldsymbol{\beta})'\boldsymbol{x}_{t+h}u_t$$

$$+(\tilde{\boldsymbol{\beta}}_T-\boldsymbol{\beta})'\boldsymbol{x}_t(\tilde{\boldsymbol{\beta}}_T-\boldsymbol{\beta})'\boldsymbol{x}_{t+h}\}. \tag{4.3}$$

By Condition 1 and Condition 2, the asymptotic behaviour of the r.h.s. of (4.3) is equivalent to that of

$$C(h)=\frac{1}{T-h}\sum_{t=1}^{T-h}u_tu_{t+h}$$

in almost everywhere sense.

Since $C(h)\to\sigma(h)$ a.e. as $T\to\infty$ for $h=0, \pm 1, \pm 2, \cdots$, we obtain

$$C^*(h)\to\sigma(h) \text{ a.e. as } T\to\infty \text{ for } h=0, \pm 1, \pm 2, \cdots,$$

from the fact of the asymptotic equivalence of $C^*(h)$ to $C(h)$. Q.E.D.

Now suppose that $\{u_t\}$ satisfies

$$u_t-\sum_{j=1}^{K} a_j u_{t-j}=\varepsilon_t(a_0=-1), \tag{4.4}$$

where the innovation $\{\varepsilon_t\}$ is a sequence of independently identically distributed random variables with mean zero and variance σ^2. Further suppose that the associated polynomial $Z^K-\sum_{j=1}^{K} a_j Z^{K-j}=0$ has all roots inside the unit circle in the complex plane. Under these assumptions, Condition 3 in Lemma 4 is automatically satisfied, that is, $\{u_t\}$ is an ergodic, strictly stationary stochastic process.

We shall consider the estimation procedure for $\{a_1, a_2, \cdots, a_K\}$ and σ^2 by using the information of the estimated residuals $\{u_t^*\}$. The criterion is the least squares principle, that is, $\{\hat{a}_1^*, \hat{a}_2^*, \cdots, \hat{a}_K^*\}$ is obtained by minimizing

$$\sum_{t=k+1}^{T}\left\{u_t^*-\sum_{j=1}^{K} a_j^* u_{t-j}^*\right\}^2 \tag{4.5}$$

with respect to $\{a_1^*, a_2^*, \cdots, a_K^*\}$, and $\hat{\sigma}^{*2}=\dfrac{1}{T}\sum_{t=1}^{T}\left(u_t^*-\sum_{j=1}^{K}\hat{a}_j^* u_{t-j}^*\right)^2$. (4.6)

In the following theorem, we shall shall show that

$$\hat{f}_T(\lambda)=\frac{\hat{\sigma}^{*2}}{2\pi\left|\sum_{j=0}^{K}\hat{a}_j^* e^{i\lambda j}\right|^2} \text{ converges}$$

to

$$f(\lambda)=\frac{\sigma^2}{2\pi\left|\sum_{j=0}^{K} a_j e^{i\lambda j}\right|^2} \text{ a.e. as } T\to\infty \text{ in } C[-\pi, \pi],$$

where $C[-\pi, \pi]$ is the space of continuous functions with the topology of uniform convergence.

THEOREM 6. Suppose that the following three conditions hold:

1. All elements of $\boldsymbol{x}_t=(x_{1t}, _{2t}, \cdots, x_{pt})'$ are the bounded functions of t.
2. $\tilde{\boldsymbol{\beta}}_T$ is a strongly consistent estimator for $\boldsymbol{\beta}$.
3. $\{u_t\}$ is the stationary Autoregressive Process of given order where the innovation is a sequence of independently identically distributed random variables. Then

$$\hat{f}_T(\lambda)\longrightarrow f(\lambda) \text{ a.s. as } T\to\infty \text{ in } C[-\pi, \pi].$$

Proof. The system of normal equations is given by

$$\frac{1}{T}\sum_{t=k+1}^{T}\left(u_t^* - \sum_{j=0}^{K}\hat{a}_j^* u_{t-j}^*\right) u^*_{t-k}=0,\ k=1, 2, \cdots, K. \tag{4.7}$$

Let

$$\hat{C}^*=[\hat{C}^*(h, h'), h, h'=1, 2, \cdots, K],$$

$$\hat{\boldsymbol{C}}_0^*=[\hat{C}^*(0, 1), \hat{C}^*(0, 2), \cdots, \hat{C}^*(0, K)]',$$

and

$$\hat{\boldsymbol{a}}^*=[\hat{a}_1^*, \hat{a}_2^*, \cdots, \hat{a}_K^*]',$$

where $\hat{C}^*(h, h')=\frac{1}{T}\sum_{t=K+1}^{T} u^*_{t-h}u^*_{t-h'}$. Then (4.7) is represented by

$$\hat{C}^*\hat{\boldsymbol{a}}^*=\hat{\boldsymbol{C}}_0^*. \tag{4.8}$$

And let

$$C=[\sigma(|i-j|), i, j=1, 2, \cdots, K],$$

$$\boldsymbol{C}_0=[\sigma(1), \sigma(2), \cdots, \sigma(K)]',$$

$$\boldsymbol{a}=[a_1, a_2, \cdots, a_K]',$$

then, in the relation of parameters,

$$C\boldsymbol{a}=\boldsymbol{C}_0. \tag{4.9}$$

From Lemma 4,

$$\hat{C}^*\to C \text{ a.e. as } T\to\infty$$

and

$$\hat{\boldsymbol{C}}_0^*\to \boldsymbol{C}_0 \text{ a.e. as } T\to\infty,$$

so we obtain

$$\hat{\boldsymbol{a}}^*\to \boldsymbol{a} \text{ a.e. as } T\to\infty. \tag{4.10}$$

On the other hand,

$$\sigma^2=E\varepsilon_t^2=E\left(u_t-\sum_{j=1}^{K} a_j u_{t-j}\right)^2$$

$$=\sum_{j=0}^{K}\sum_{l=0}^{K} a_j a_l \sigma(j-l),$$

where $a_0=-1$. From (4.6),

$$\hat{\sigma}^{*2}=\frac{1}{T}\sum_{t=1}^{T} u_t^{*2}-2\sum_{j=1}^{K}\hat{a}_j^*\frac{1}{T}\sum_{t=1}^{T} u^*_{t-j}u_t^* +\frac{1}{T}\sum_{j=1}^{K}\sum_{l=1}^{K}\hat{a}_j^*\hat{a}_l^*\sum_{t=1}^{T} u^*_{t-j}u^*_{t-l},$$

where the first term converges to $\sigma(0)$ a.e. as $T\to\infty$, the second term to $-2\sum_{j=1}^{K} a_j\sigma(j)$ a.e. as $T\to\infty$ and the last term to $\sum_{j=1}^{K}\sum_{l=1}^{K} a_ja_l\sigma(j-l)$ a.e. as $T\to\infty$. Then

$$\hat{\sigma}^{*2}\to\sigma^2 \text{ a.e. as } T\to\infty. \tag{4.11}$$

Because of (4.10), (4.11) and the fact that the associated polynomial of $\{u_t\}$ has all roots inside the unit circle,

$$P[\sup_{\lambda}|\hat{f}_T(\lambda)-f(\lambda)|\to 0 \text{ as } T\to\infty]$$

$$=P\left[\sup_{\lambda}\left|\frac{\hat{\sigma}^{*2}}{2\pi\left|\sum_{j=0}^{K}\hat{a}_j^*e^{i\lambda j}\right|^2}-\frac{\sigma^2}{2\pi\left|\sum_{j=0}^{K}a_je^{i\lambda j}\right|^2}\right|\to 0 \text{ as } T\to\infty\right]$$

$$\geq P\left[\sup_{\lambda}\left\{\frac{1}{2\pi\left|\sum_{j=0}^{K}\hat{a}_j^*e_i^{\lambda j}\right|^2}|\hat{\sigma}^{*2}-\sigma^2|+\sigma^2\left|\frac{1}{2\pi\left|\sum_{j=0}^{K}\hat{a}_j^*e^{i\lambda j}\right|^2}-\frac{1}{2\pi\left|\sum_{j=0}^{K}a_je^{i\lambda j}\right|^2}\right|\right\}\right.$$

$$\left.\to 0 \text{ as } T\to\infty\right]=1. \qquad \text{Q.E.D.}$$

REMARK 2. In the case where $\{u_t\}$ is a finite moving average process, that is, $u_t=\sum_{j=0}^{K} b_j\varepsilon_{t-j}$, where $\{\varepsilon_t\}$ is a sequence of independently identically random variables with mean zero and variance 1, we can show, from the result of CLEVENSON, M. ([5]), that

$$\hat{f}_T^*(\lambda)=\frac{1}{2\pi}\sum_{h=-K}^{K} C^*(h)e^{i\lambda h}\to f(\lambda) \text{ a.e. as } T\to\infty \text{ in } C[-\pi,\pi],$$

where $f(\lambda)=\frac{1}{2\pi}\sum_{h=-K}^{K}\sigma(h)e^{i\lambda h}$.

Acknowledgements

The author would like to thank Prof. YASUTOSHI WASHIO, Keio University for his many conscientious and constructive criticisms, suggestions and encouragements. Thanks are also extended to Prof. MITSUAKI HUZII, Tokyo Institute of Technology for his suggestions and to Prof. TAKUJI ONOYAMA, Keio University of suggesting the motivation of the proof of Theorem 3 and to Dr. NOBUO SHINOZAKI, Keio University for discussions.

The author wishes to thank Prof. Theodore W. ANDERSON, Stanford University for his conscientiousness of indicating the reference ([11]).

REFERENCES

[1] ANDERSON, T.W. (1971): The Statistical Analysis of Time Series. John Wiley & Sons

[2] ANDERSON, T. W. and J. B. TAYLOR (1976): Strong Consistency of Least Squares Estimates in Normal Linear Regression. Ann. Statis. **4**, 788–790.

[3] ANDERSON, T.W. and J.B. TAYLOR (1976): Conditions for Strong Consistency of Least Squares Estimates in Linear Models. Technical Report No. 213 The Economic Series. Stanford Univ.

[4] BREIMAN, L. (1968): Probability. Addison-Wesley.

[5] CLEVENSON, M. (1970): Asymptotically Efficient Estimates of the Parameters of a Moving Average Time Series. Ph. D. Thesis, Stanford Univ.

[6] DOOB, J.L. (1953): Stochastic Processes. John Wiley & Sons

[7] DRYGAS, H. (1976): Weak and Strong Consistency of the Least Squares Estimators in Regression Models. Z. Wahrscheinlichkeits theorie verw. Gebiete **34**, 119–127.

[8] EICKER, F. (1963): Asymptotic Normality and Consistency of the Least Squares Estimates for Families of Linear Regressions. Ann. Math. Stat. **34**, 447–456.

[9] HANNAN, E.J. (1970): Multiple Time Series. John Wiley & Sons.

[10] HANNAN, E.J. (1971): Non-linear Time Series Regression. J. Appl. Prob. **8**, 767–780.

[11] LAI, T. L. and H. ROBBINS (1977): Strong Consistency of Least Squares Estimates in Regression Models. Proc. Natl. Acad. Sci. U.S.A. 2667–2669.

[12] LUKACS, E. (1975): Stochastic Convergence. 2nd edition. Academic Press.

[13] PRIESTLEY, M.B. (1965): Evolutionary Spectra and Non-stationary Processes. J.R. statist Soc. B. 27, 204–237.