# Consistency of non-linear least-squares estimators

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics

STATISTICS AND OPERATIONS RESEARCH GROUP

Memorandum COSOR 74-07

# Consistency of non-linear least-squares estimators

by

J.G. Baaij

Eindhoven, June 1974

## 1. Introduction

For additive regression models with i.i.d. (independent identically distributed) errors least squares is an obvious criterion for estimation. For most non-linear models it is impossible to give an analytic expression for the l.s. (least squares) estimator in terms of the data, and its distribution is usually intractable. Therefore, the study of the asymptotic properties of this estimator is of interest.

The model we consider is:

$$Y_t = f_t(\theta^0) + E_t ,$$

where $f_1, f_2, \ldots$ are known, real functions on a parameter space $\Theta$, and $E_1, E_2, \ldots$ are random variables.

Unless otherwise specified, the following three assumptions are valid throughout the paper:

I. $E_1, E_2, \ldots$ are i.i.d. with expectation 0 and variance $\sigma^2$, with $0 < \sigma^2 < \infty$.

II. $\Theta$ is a compact subset of $\mathbb{R}^p$ (the p-dimensional Euclidean space).

III. $f_t(\theta)$ is continuous in $\theta$ for all t.

We now introduce some notation:

$P_n^E$ and $P^E$ denote the probability distributions of $E_1, E_2, \ldots, E_n$, and $E_1, E_2, \ldots$, respectively and similarly for $P_n^Y$ and $P^Y$.

$v_t(\theta)$ is defined by $v_t(\theta) := f_t(\theta^0) - f_t(\theta)$.

A l.s. estimator for $\theta^0$ based on n observations $Y_1, Y_2, \ldots, Y_n$, is a measurable function $\hat{\theta}_n : \mathbb{R}^n \to \Theta$ such that

$$\sum_{t=1}^{n} [y_t - f_t(\hat{\theta}_n(y))]^2 = \inf_{\theta \in \Theta} \sum_{t=1}^{n} (y_t - f_t(\theta))^2$$

for all $y = (y_1, y_2, \ldots, y_n)' \in \mathbb{R}^n$.

Jennrich [2] proved that assumptions I, II and III are sufficient for the existence of a l.s. estimator.

An l.s. estimator is called consistent if

$$\lim_{n\to\infty} P_n^Y(|\hat{\theta}_n(Y_1,Y_2,\ldots,Y_n) - \theta^0| > \varepsilon) = 0$$

for all $\varepsilon > 0$, and strongly consistent if

$$P^Y(\lim_{n\to\infty} \hat{\theta}_n(Y_1,Y_2,\ldots,Y_n) = \theta^0) = 1 .$$

In his paper Jennrich [2] gives sufficient conditions for strong consistency and asymptotic normality of the l.s. estimator. His conditions for strong consistency are the assumptions I, II and III together with:

IV. $\frac{1}{n}\sum_{t=1}^{n} f_t(\alpha)f_t(\beta)$ converges uniformly on $\Theta \times \Theta$ and

V. $Q(\theta) = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} v_t(\theta)^2$ has a unique minimum for $\theta = \theta^0$

(condition IV guarantees the existence of $Q(\theta)$).

Other sufficient conditions for strong consistency are given by Malinvaud [3]. He considers a model that differs notationally from Jennrich's:

$$Y_t = g(x_t,\theta^0) + E_t ,$$

the design parameters $x_t \in \mathbb{R}^m$ for $t = 1,2,\ldots$ .
An objection to this notation is the fact that for certain models Malinvaud's theorem applies only after a transformation of these parameters (reparametrization).
The conditions of Malinvaud are assumption I together with

VI. The vectors $x_t$ are contained in a compact set $Z \subset \mathbb{R}^m$,

VII. If $\mu_n$ is the probability measure on the Borel sets of $\mathbb{R}^m$ defined by $\mu_n(\{x_j\}) = \frac{1}{n}$ for $j = 1,2,\ldots,n$ then the sequence $\mu_n$ converges weakly to a probability measure $\mu$ on the Borel sets of $\mathbb{R}^m$,

VIII. $g(x,\alpha)$ is continuous on $Z \times \Theta$,

IX. if $\alpha, \beta \in \Theta$ and $\alpha \neq \beta$ then

$$\mu(\{x \mid g(x,\alpha) \neq g(x,\beta)\}) > 0 ,$$

X. for all $G > 0$ there exists an $n_0$ such that the set

$$\{\alpha \mid \frac{1}{n} \sum_{t=1}^{n} g(x_t,\alpha)^2 \leq G\}$$

is uniformly bounded for $n > n_0$.

In section 2 we give some examples, and in section 3 some new conditions for strong consistency which are less restrictive than those given by Jennrich and Malinvaud.

## 2. Examples

We first prove

Assertion 1.

Let $E_1, E_2, \ldots$ satisfy assumption I and let $x_1, x_2, \ldots$ be real numbers with $x_1 \neq 0$. If $E_t$ is not degenerate then

$$\frac{\sum_{t=1}^{n} x_t E_t}{\sum_{t=1}^{n} x_t^2} \rightarrow 0 \text{ a.s.} \Leftrightarrow \sum_{t=1}^{n} x_t^2 \rightarrow \infty .$$

Proof.

$\Rightarrow$) If $\sum_{t=1}^{\infty} x_t^2 < \infty$ then $\sum_{t=1}^{n} x_t E_t \rightarrow 0$ a.s.

From this we have $\sum_{t=1}^{n} x_t E_t \rightarrow 0$ in distribution and if $\varphi(u)$ denotes the characteristic function of $E_t$ then $\prod_{i=1}^{n} \varphi(x_t u) \rightarrow 1$ and consequently $\prod_{i=1}^{n} |\varphi(x_t u)| \rightarrow 1$ for all u.

But then $|\varphi(x_t u)| \equiv 1$ and $E_t$ is degenerate.

$\Leftarrow$) By the strong law of large numbers as formulated in Breimann [1], a sufficient condition for the left-hand side of the equivalence is

$$(*) \qquad \sum_{k=1}^{\infty} \frac{x_k^2 \sigma^2}{\left[\sum_{t=1}^{k} x_t^2\right]^2} < \infty, \text{ together with } \sum_{t=1}^{n} x_t^2 \to \infty.$$

If $n > 1$ then

$$\frac{x_n^2}{\left[\sum_{t=1}^{n} x_t^2\right]^2} \le \frac{x_n^2}{\left(\sum_{t=1}^{n-1} x_t^2\right)\left(\sum_{t=1}^{n} x_t^2\right)} = \frac{1}{\sum_{t=1}^{n-1} x_t^2} - \frac{1}{\sum_{t=1}^{n} x_t^2}$$

so

$$\sum_{k=1}^{n} \frac{x_k^2}{\left[\sum_{t=1}^{k} x_t^2\right]^2} \le \frac{1}{x_1^2} + \sum_{k=2}^{n} \left[\frac{1}{\sum_{t=1}^{n-1} x_t^2} - \frac{1}{\sum_{t=1}^{n} x_t^2}\right] \le \frac{2}{x_1^2}$$

and condition (*) holds.

Example 1 is the linear model $Y_t = \theta^0 x_t + E_t$, where $x_1, x_2, \ldots$ and $\theta$ are real numbers.
It is well known that the l.s. estimator

$$\hat{\theta}_n = \frac{\sum_{t=1}^{n} x_t Y_t}{\sum_{t=1}^{n} x_t^2} = \theta^0 + \frac{\sum_{t=1}^{n} x_t E_t}{\sum_{t=1}^{n} x_t^2}$$

and assertion 1 shows that strong consistency is equivalent with $\sum_{t=1}^{n} x_t^2 \to \infty$.

Jennrich's conditions IV and V are in this case equivalent to $\frac{1}{n} \sum_{t=1}^{n} x_t^2$ converges to a positive limit.
The conditions of Malinvaud are less explicit. For this model they are equivalent with:

VI. the sequence $x_1, x_2, \ldots$ is bounded,

VII. $\frac{1}{n} \sum_{t=1}^{n} f(x_t)$ converges for all continuous (and bounded) functions $f : \mathbb{R} \to \mathbb{R}$. This because $\mu_n$ converges weakly to a probability measure $\mu$ on $\mathbb{R}$.

IX. $\mu(\{0\}) < 1$. Note that this condition does not hold if $x_t \to 0$.

X. $\{\alpha \mid \frac{\alpha^2}{n} \sum_{t=1}^{n} x_t^2 < G\}$ is uniformly bounded for all $n \geq n_0(G)$.

This means $\frac{1}{n} \sum_{t=1}^{n} x_t^2 > \varepsilon$ for $n \geq n_0$ and an $\varepsilon > 0$.

Example 2: $Y_t = \alpha^0 e^{\beta^0 x_t} + E_t$.
Jennrich's condition IV requires the convergence of $\frac{1}{n} \sum_{t=1}^{n} \alpha' \alpha e^{(\beta'+\beta)x_t}$, uniformly in $\alpha$, $\beta$, $\alpha'$ and $\beta'$.
For convergence it is not sufficient that $x_1, x_2, \ldots$ are contained in a bounded set Z: the sequence may have more than one limit point. Also condition VIII of Malinvaud is not necessarily satisfied if $x_1, x_2, \ldots$ are in Z: $\mu_n$ converges weakly to a probability measure on $\mathbb{R}$ implies that $\frac{1}{n} \sum_{t=1}^{n} f(x_t)$ converges for each continuous function f on Z, and this is not necessarily for $x_t \in Z$, $t = 1,2,\ldots$ .
Later on we shall see that a condition like Jeunrich's IV together with the condition that $x_1, x_2, \ldots$ be bounded, in this case is sufficient for strong consistency.

Example 3: $Y_t = \cos \theta t + E_t$ with $\theta \in [\varepsilon, \pi-\varepsilon]$ and $\varepsilon > 0$.
From

$$\frac{1}{n} \sum_{t=1}^{n} \cos \theta t = \begin{cases} 1 & \text{if } \theta = k\pi \\ \dfrac{\sin\frac{n}{2}\theta \cos \frac{n+1}{2}\theta}{n \sin \frac{\theta}{2}} \to 0 \text{ for } n \to \infty & \text{if } \theta \neq k\pi \end{cases}$$

we see that $\frac{1}{n}\sum_{t=1}^{n}\cos\alpha t\cos\beta t = \frac{1}{n}\sum_{t=1}^{n}\frac{1}{2}[\cos(\alpha-\beta)+\cos(\alpha+\beta)]$ converges on $[\varepsilon,\pi-\varepsilon]^2$ but the convergence is not uniform: the partial sums are continuous but the limit function is not. So this model does not satisfy the condition IV of Jennrich. Neither does the model satisfy the conditions of Malinvaud, simply because the sequence 1,2,... is not bounded.
After the reparametrization $Y_t = \cos(\frac{\theta}{x_t}) + E_t$ with $x_t = \frac{1}{t}$ we have $\mu(\{0\}) = 1$ and the model does not satisfy condition IX of Malinvaud.

R. Potharst [4] proved that the l.s. estimator in this case is even "consistent of order n", that is

$$P^E(\{\lim_{n\to\infty} n(\hat{\theta}_n - \theta^0) = 0\}) = 1 .$$

In the linear case, we saw that under the assumptions I, II and III a sufficient condition for consistency is that $\sum_{t=1}^{n} v_t(\theta)^2 \to \infty$ for $\theta \neq \theta^0$.
This condition, as condition IV of Jennrich and the conditions IX and X of Malinvaud, is made to guarantee the identifiability of the model.
The next example shows that in general more conditions are necessary.

Example 4: $Y_t = f_t(\theta^0) + E_t$.
The parameterspace $\Theta = [-1,1]$, $\theta^0 = -\frac{1}{2}$ and $P^E(E_t = -\frac{1}{2}) = P^E(E_t = \frac{1}{2}) = \frac{1}{2}$.
Let $m_k = (k+1)! - k!$ and $.d_1(\theta)d_2(\theta)d_3(\theta)\ldots$ the binary expansion of $\theta$ for $\theta \in [0,1]$.
For $t = k!+1, k!+2, \ldots, (k+1)!$ the functions $f_t$ are defined in the points $\theta = -1$, $\theta = 0$ and $\theta_j = j/2^{m_k}$ for $j = 1,2,\ldots,m_k$ as follows:

$$f_t(-1) = -a,\ f_t(0) = a \text{ and } f_t(\theta_j) = \begin{cases} a \text{ if } d_t(\theta_j) = 0 \\ b \text{ if } d_t(\theta_j) = 1 \end{cases}$$

and the functions are linear between these points.

The model satisfies the conditions I, II and III and if $0 < a < b$ then Jennrich's condition IV holds. Note that $\theta^0 = -\frac{1}{2}$ implies $Y_t = E_t$.
Let

$$Z_k = \sum_{t=k!+1}^{(k+1)!} (Y_t + \tfrac{1}{2}) \quad \text{and} \quad \tilde{\theta}_{(k+1)!} = \sum_{t=1}^{(k+1)!} \frac{Y_t + \frac{1}{2}}{2^t}$$

then $Z_k$ is the number of positive $Y_t$ among $Y_{k!+1}, Y_{k!+2}, \dots, Y_{(k+1)!}$ thus $\frac{1}{m_k} Z_k \to \frac{1}{2}$ a.s. and

$$f_t(\tilde{\theta}_{(k+1)!}) = \begin{cases} a & \text{if } Y_t = -\frac{1}{2} \\ b & \text{if } Y = \frac{1}{2} \end{cases} \qquad \text{for } t = k!+1, k!+2, \dots, (k+1)! \;.$$

From this

$$\frac{1}{(k+1)!} \sum_{t=1}^{(k+1)!} (Y_t - f_t(\tilde{\theta}_{(k+1)!}))^2 \le$$

$$\le \frac{k!}{(k+1)!} (b + \tfrac{1}{2})^2 + \frac{1}{(k+1)!} \sum_{t=k!+1}^{(k+1)!} (Y_t - f_t(\tilde{\theta}_{(k+1)!}))^2 =$$

$$= \frac{1}{k+1} (b + \tfrac{1}{2})^2 + \frac{m_k}{(k+1)!} \left[ \frac{1}{m_k} Z_k (\tfrac{1}{2} - b)^2 + \frac{m_k - Z_k}{m_k} (-\tfrac{1}{2} - a)^2 \right]$$

$$\longrightarrow \tfrac{1}{2}(a^2 + b^2 + \tfrac{1}{2} + a - b) \quad \text{a.s.} \quad \text{if } k \to \infty.$$

This is less than $\sigma^2 = \frac{1}{4}$ for appropriate a,b (for example $a = \frac{1}{10}$, $b = \frac{1}{2}$) and it is an immediate consequence of Jennrich's theorem 4 ([2] page 636) that $P^Y\left(\frac{1}{n} \sum_{t=k}^{n} (Y_t - f_t(\theta))^2 \to \sigma^2 + (2a\theta + a)^2 \text{ uniformly for } \theta \in [-1,0]\right) = 1.$

Thus there exists with probability one a $T = T(Y_1, Y_2, \dots)$ such that

$$\frac{1}{(k+1)!} \sum_{t=1}^{(k+1)!} (Y_t - f_t(\tilde{\theta}_{(k+1)!}))^2 < \frac{1}{(k+1)!} \sum_{t=1}^{(k+1)!} (Y_t - f_t(\theta))^2$$

for all $\theta \in [-1,0]$ and $(k+1)! > T$ and consequently the l.s. estimator $\hat{\theta}_{(k+1)!}$ is positive.

## 3. New conditions for strong consistency

Let $U(\theta^0,\varepsilon) = \{\theta \in \Theta \mid |\theta^0-\theta| < \varepsilon\}$

Theorem 1.

If the model $Y_t = f_t(\theta^0) + E_t$ satisfies the assumptions I, II and III, and if

$$\forall_{\varepsilon>0}\ \exists_{n_0}\ \forall_{\theta\in\Theta\backslash U(\theta^0,\varepsilon)}\ [n \geq n_0 \Rightarrow \frac{1}{n}\sum_{t=1}^{n} v_t(\theta)^2 > 4\sigma^2]$$

then the l.s. estimator is strongly consistent.

Proof. The l.s. estimator minimizes

$$Q_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}(Y_t - f_t(\theta))^2 = \frac{1}{n}\sum_{t=1}^{n}(v_t(\theta)+E_t)^2 ,$$

or equivalently

$$R_n(\theta) := \sum_{t=1}^{n} v_t(\theta)^2 + 2\sum_{t=1}^{n} v_t(\theta)E_t .$$

According to Schwarz's inequality is

$$|\sum_{t=1}^{n} v_t(\theta)E_t| \leq \left[\left(\sum_{t=1}^{n} v_t(\theta)^2\right)\left(\sum_{t=1}^{n} E_t^2\right)\right]^{\frac{1}{2}}$$

and thus

$$R_n(\theta) \geq \sum_{t=1}^{n} v_t(\theta)^2 - 2\left[\left(\sum_{t=1}^{n} v_t(\theta)^2\right)\left(\sum_{t=1}^{n} E_t^2\right)\right]^{\frac{1}{2}} .$$

Suppose $\hat{\theta}_n$ is not strongly consistent.

Then there exists an $\varepsilon > 0$ and a realisation $(e_t)$ such that

$$\frac{1}{n}\sum_{t=1}^{n} e_t^2 \to \sigma^2 \quad \text{and} \quad |\theta^0 - \hat{\theta}_{n'}| > \varepsilon$$

for a subsequence $(\hat{\theta}_{n'})$ of the sequence l.s. estimators corresponding to $(e_t)$.

It follows that

$$\frac{1}{n'} \sum_{t=1}^{n'} v_t(\hat{\theta}_{n'})^2 > 0$$

and

$$\frac{\frac{1}{n'} R_{n'}(\hat{\theta}_{n'})}{\frac{1}{n'} \sum_{t=1}^{n'} v_t(\hat{\theta}_{n'})^2} \geq 1 - \frac{2\left(\frac{1}{n'} \sum_{t=1}^{n'} e_t^2\right)^{\frac{1}{2}}}{\left(\frac{1}{n'} \sum_{t=1}^{n'} v_t(\hat{\theta}_{n'})^2\right)^{\frac{1}{2}}}$$

and thus

$$\liminf_{n' \to \infty} \frac{\frac{1}{n'} R_{n'}(\hat{\theta}_{n'})}{\frac{1}{n'} \sum_{t=1}^{n'} v_t(\hat{\theta}_{n'})^2} > 0 ,$$

but this contradicts $R_{n'}(\hat{\theta}_{n'}) \leq R_{n'}(\theta^0) = 0$ .

Note that Theorem 1 applies to example 3 for the case that $\sigma^2 < \frac{1}{4}$.

<u>Lemma 1</u>. If $(a_t)$ and $(b_t)$ are sequences of real numbers such that $|a_t - b_t| < \varepsilon$ for $t = 1,2,\ldots$ and $\frac{1}{n} \sum_{t=1}^{n} b_t^2 > \delta > 0$ for $n \geq n_0$ then

$$\frac{\sum_{t=1}^{n} a_t^2}{\sum_{t=1}^{n} b_t^2} < 1 + (1 + \tfrac{1}{\delta})(4\varepsilon + \varepsilon^2) \text{ for } n \geq n_0 .$$

<u>Proof</u>. Let $n \geq n_0$, $N_1 = \{t \in \mathbb{N} \mid t \leq n, |b_t| < 2\}$ and $N_2 = \{t \in \mathbb{N} \mid t \leq n, |b_t| \geq 2\}$ , then

$$\frac{\sum_{t=1}^{n} a_t^2}{\sum_{t=1}^{n} b_t^2} = 1 + \frac{\frac{1}{n} \sum_{t \in N_1} (a_t - b_t)(a_t + b_t)}{\frac{1}{n} \sum_{t=1}^{n} b_t^2} + \frac{\sum_{t \in N_2} (a_t - b_t)(a_t + b_t)}{\sum_{t=1}^{n} b_t^2} \leq$$

$$\leq 1 + \frac{\varepsilon(4+\varepsilon)}{\delta} + \frac{\varepsilon \sum_{t \in N_2} (2|b_t| + \varepsilon)}{\sum_{t=1}^{n} b_t^2} \leq 1 + \varepsilon[\varepsilon + \frac{4+2\varepsilon}{\delta}].$$

Assertion II.

$E_1, E_2, \ldots$ and $\Theta$ satisfy assumption I resp. II and $(v_t(\theta))$ is a sequence of real functions on $\Theta$ such that

i) $(v_t(\theta))$ is equicontinuous, that is

$$\forall_\theta \ \forall_{\varepsilon>0} \ \exists_{\delta>0} \ \forall_t \ [|\theta-\theta'| < \delta \Rightarrow |v_t(\theta) - v_t(\theta')| < \varepsilon]$$

ii) $\exists_{\delta>0} \ \exists_{n_0} \ \forall_\theta \ [n > n_0 \Rightarrow \frac{1}{n} \sum_{t=1}^{n} v_t(\theta)^2 > \delta]$ .

Then

$$P^E\left(\left\{ \frac{\sum_{t=1}^{n} v_t(\theta)E_t}{\sum_{t=1}^{n} v_t(\theta)^2} \to 0 \text{ uniformly in } \theta\right\}\right) = 1 .$$

Proof. Let $\{\theta_1, \theta_2, \ldots\}$ be dense in $\Theta$ and let

$$\mathcal{E}_0 = \{(e_t) \mid \frac{1}{n} \sum_{t=1}^{n} e_t^2 \to \sigma^2\} \text{ and } \mathcal{E}_i = \{(e_t) \mid \frac{\sum_{t=1}^{n} v_t(\theta_i)e_t}{\sum_{t=1}^{n} v_t(\theta_i)^2} \to 0\},$$

then $P^E(\mathcal{E}_0) = 1$, and $P^E(\mathcal{E}_i) = 1$, $i = 1,2,\ldots$ because of assertion I, thus if $\mathcal{E} = \bigcap_{i=0}^{\infty} \mathcal{E}_i$ then $P^E(\mathcal{E}) = 1$.

Let $\varepsilon$ be positive and $U_i = \{\theta \mid |v_t(\theta) - v_t(\theta_i)| < \varepsilon, \ t = 1,2,\ldots\}$, $i = 1,2,\ldots$ .
Then $(U_i)$ is an open covering of $\Theta$ and there exists a finite subcovering $\{U_k \mid k = k_1, k_2, \ldots, k_r\}$.

If $(e_t) \in \mathcal{E}$ then there exists an $n_0$ such that for $n \geq n_0$

$$\frac{1}{n} \sum_{t=1}^{n} e_t^2 < 2\sigma^2 \quad \text{and} \quad \frac{\sum_{t=1}^{n} v_t(\theta_k)e_t}{\sum_{t=1}^{n} v_t(\theta_k)^2} < \varepsilon \quad (k = k_1, k_2, \ldots, k_r) .$$

If now $\theta \in U_k$, then

$$\frac{\sum_{t=1}^{n} v_t(\theta)e_t}{\sum_{t=1}^{n} v_t(\theta)^2} = \frac{\sum_{t=1}^{n} v_t(\theta_k)^2}{\sum_{t=1}^{n} v_t(\theta)^2} \left[\frac{\sum_{t=1}^{n} v_t(\theta_k)e_t + \sum_{t=1}^{n} (v_t(\theta_k)-v_t(\theta))e_t}{\sum_{t=1}^{n} v_t(\theta_k)^2}\right] \leq$$

$$\leq [1 + \varepsilon(\varepsilon + \frac{4+2}{\delta})][\varepsilon + \frac{2\varepsilon\sigma^2}{\delta}] ,$$

by Lemma 1 and this is arbitrarily small, independent of $\theta$.

Theorem 2. If the model $Y_t = f_t(\theta^0) + E_t$ satisfies the assumptions I, II and III and if

i) the sequence $(f_t(\theta))_t$ is equicontinuous on $\Theta$

ii) $\forall_{\varepsilon>0}\ \exists_{\delta>0}\ \exists_{n_0}\ \forall_{n\geq n_0}\ \forall_{\theta\in\Theta\setminus U(\theta^0,\varepsilon)}\ [\frac{1}{n}\sum_{t=1}^{n} v_t(\theta)^2 > \delta]$,

then the l.s. estimator $\hat{\theta}_n$ is strongly consistent.

Proof. If $\hat{\theta}_n$ is not strongly consistent, then there exists a realisation $(e_t)$ such that

a) $\exists_{n_0}\quad n \geq n_0 \Rightarrow \forall_{\theta\in\Theta\setminus U(\theta^0,\varepsilon)} \left[\left|\frac{\sum_{t=1}^{n} v_t(\theta)e_t}{\sum_{t=1}^{n} v_t(\theta)^2}\right| < \frac{1}{4}\right]$ (assertion 2);

b) $\hat{\theta}_{n'} \in \Theta\setminus U(\theta^0,\varepsilon)$ for a subsequence $(\hat{\theta}_{n'})$ of the sequence of l.s. estimators corresponding to $(e_t)$ and an $\varepsilon > 0$.

It follows that

$$\frac{R_{n'}(\hat{\theta}_{n'})}{\sum_{t=1}^{n'} v_t(\hat{\theta}_{n'})^2} = 1 + \frac{2\sum_{t=1}^{n'} v_t(\hat{\theta}_{n'})e_t}{\sum_{t=1}^{n'} v_t(\hat{\theta}_{n'})^2} > \tfrac{1}{2} \text{ if } n' \geq n_0 .$$

This contradicts $R_{n'}(\hat{\theta}_{n'}) \leq R_{n'}(\theta^0) = 0$.

References

[1] Breiman, L., Probability.
Addison Wesley Publishing Company (1968).

[2] Jennrich, R.I., Asymptotic properties of non-linear least-squares estimators.
Ann. Math. Statist. 40 (1969), 633-643.

[3] Malinvaud, E., The consistency of non-linear regressions.
Ann. Math. Statist. 41 (1970), 956-969.

[4] Potharst, R., An approximate confidence interval for the frequency of a harmonic disturbed by random noise.
Unpublished paper (1971).