Asymptotic properties of the complex valued non-linear regression model
Debasis Kundu

Department of Mathematics, Indian Institute of Technology, Kanpur, Kanpur, PIN, India

Online Publication Date: 01 January 1991
To cite this Article: Kundu, Debasis (1991) 'Asymptotic properties of the complex valued non-linear regression model', Communications in Statistics - Theory and Methods, 20:12, 3793 - 3803
To link to this article: DOI: 10.1080/03610929108830741

URL: http://dx.doi.org/10.1080/03610929108830741

PLEASE SCROLL DOWN FOR ARTICLE
ASYMPTOTIC PROPERTIES OF THE COMPLEX VALUED
NON-LINEAR REGRESSION MODEL

Debasis Kundu

Department of Mathematics
Indian Institute of Technology, Kanpur
Kanpur, PIN 208016, India

Key Words and Phrases: Non-linear regression, Consistent Estimator
AMS 1987 Subject Classification: Primary 62J02, 62F05

ABSTRACT

The non-linear regression model, when the parameters are complex valued is considered here. Jennrich (1969) considered the non-linear regression model when the parameters are real valued. He first rigorously proved the existence of the least square estimator and showed its consistency properties and asymptotic normality. In this paper we generalise the idea for the complex parameters case. Large sample properties of the proposed estimator has been studied.

1. INTRODUCTION

In the past few years, non-linear regression has paid an important role in statistical science because of its wide scale applicability and due to the availability of the high speed computers. Although the statistical theory of parameter estimation in the linear model has almost completely been de-
developed, in the non-linear model many problems are still unsolved. Due to non-linearity, many statistical concepts such as unbiasedness or completeness break down for finite samples. Thus most of the theoretical results for non-linear problems are asymptotic. See for example, Prakasa Rao (1987) for a selection of references.

The least square method plays an important role in drawing inferences about the parameters in the non-linear regression model. Much of the previous work was done by assuming the existence and the consistency of the least square estimator and then proving the asymptotic normality. See for example, Hartley and Booker (1965). Jennrich (1969) first rigorously proved the existence of the least squares estimator and showed its consistency. He considered the following model:

\[ y_t = f(x_t, \theta_0) + \varepsilon_t, t = 1, 2, \ldots, n \]  

(1-1)

where \( x_t \) is the \( t \) th fixed independent input vector which gives rise to the observations \( y_t, f_t(\theta) = f(x_t, \theta) \) are known continuous functions in \( \theta \), when \( \Theta \) is a compact subset of the Euclidean space, \( \theta_0 \) is an interior point of \( \Theta \), and \( \varepsilon_t \) are i.i.d. errors with mean zero and finite variance \( \sigma^2 > 0 \). Any vector \( \hat{\theta} \) which minimizes the residual sum of squares,

\[ R_n(\theta) = \sum_{i=1}^{n} (y_i - f(x_i, \theta))^2 \]

will be called the least squares estimator of \( \theta_0 \). Jennrich in his paper proved the existence of the least squares estimator of \( \theta_0 \) and the strong consistency of \( \theta_0 \) under the following assumptions: \( F_n(\theta_1, \theta_2) \) converges uniformly to a continuous function \( F(\theta_1, \theta_2) \) for all \( \theta_1, \theta_2 \) and \( F(\theta_1, \theta_2) = 0 \) if and only if \( \theta_1 = \theta_2 \), where

\[ F_n(\theta_1, \theta_2) = \sum_{i=1}^{n} (f(x_i, \theta_1) - f(x_i, \theta_2))^2 \]

Under some stronger assumptions, asymptotic normality has been proved in the same paper. Wu (1981) has given sufficient conditions under which the sequence of least squares estimators converges to \( \theta_0 \), almost surely when the growth rate requirement of \( F_n \) is replaced by Lipschitz condition on the sequence \( f_t \). Wu’s assumptions are not comparable with those of Jennrich when the growth rate is of order \( n \). Recently, a lot of attention has been given to numerical methods for different non-linear models. For example, see Bates and Watts (1988), Bunke et al. (1977), Osborne (1975). Some of the testing problems has been answered by Schmidt (1982).
In this paper we have considered the complex valued non-linear regression model of the following form:

\[ y_t = f(x_t, \theta_0) + \epsilon_t \quad (1.2) \]

when the function \( f \) and the error random variable \( \epsilon_t \) might be complex valued. This is an important problem in the field of signal processing where the data may be complex valued. Recently this problem has received considerable attention (see, Rao 1988). So far nobody has studied the asymptotic properties of the least norm square estimator, which is the main aim of this paper.

We define the least norm squares estimator (LNSE) in Section 2. We generalize the concept of tail product and tail cross product to the case of complex valued sequences and complex valued random variables in Section 3. We prove the consistency of LNSE and show that is asymptotically normal in Section 4. In Section 5, we consider one particular example which is very common in signal processing literature.

**LEAST NORM SQUARE ESTIMATOR (LNSE)**

We are considering the model:

\[ y_t = f(x_t, \theta_0) + \epsilon_t, t = 1, \ldots, n \quad (2.1) \]

where \( f(x, \theta) \) is a known complex valued continuous function defined on \( \mathcal{X} \times \Theta \to \mathbb{C}^1 \). \( \mathcal{X} \) is the sample space, \( \Theta \) is a compact subset of \( \mathbb{C}^p \) denotes \( p \)-dimensional complex Euclidean space. The \( \epsilon \)'s can be written explicitly as follows:

\[ \epsilon_t = Re(\epsilon_t) + j Im(\epsilon_t) \]

where \( Re(\epsilon_t) \) represents the real part of \( \epsilon_t \), \( Im(\epsilon_t) \) denotes the imaginary part of \( \epsilon_t \), and \( j = \sqrt{-1} \). We are assuming that

\[
\begin{align*}
E\{Re(\epsilon_t)\} &= E\{Im(\epsilon_t)\} = 0 \\
Var\{Re(\epsilon_t)\} &= \sigma_1^2, Var\{Im(\epsilon_t)\} = \sigma_2^2, \sigma_1^2 < \infty, \sigma_2^2 < \infty
\end{align*}
\]

where \( \sigma_1^2, \sigma_2^2 \) are unknown and \( \theta_0 \) is an interior point of \( \Theta \). The real and imaginary part of \( \epsilon_t \) are independently distributed. Any vector which minimizes
\[ R_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} |y_i - f(x_i, \theta)|^2 \]

will be called the least norm squares estimator of \( \theta_0 \)

3. TAIL PRODUCT AND TAIL CROSS PRODUCT

In this section we extend the definition of tail product and tail cross product for the complex valued sequence and complex valued random variables, which we will need to prove the consistency of LNSE.

Let \( \tilde{x} = \{x_t\}, t=1,2,... \) and \( \tilde{y} = \{y_t\}, t=1,2,... \) be two sequences of complex numbers and let

\[ [\tilde{x}, \tilde{y}]_n = \frac{1}{n} \sum_{t=1}^{n} x_t y_t \]

If \( \lim_{n \to \infty} [\tilde{x}, \tilde{y}]_n \) exists, its limit \( [\tilde{x}, \tilde{y}] \) will be called the tail product of \( \tilde{x} \) and \( \tilde{y} \). Let \( g \) and \( h \) be two sequences of complex valued functions on \( \Theta \), such that \( [g(\alpha), h(\beta)]_n \) converges uniformly to \( [g(\alpha), g(\beta)] \) for all \( \alpha, \beta \in \Theta \). If \( (g, h) \) denotes the function on \( \Theta \times \Theta \) which takes \( (\alpha, \beta) \) into \( [g(\alpha), h(\beta)] \), then this function will be called the tail cross product of \( g \) and \( h \).

We next develop sufficient conditions under which the tail cross product of the two functions \( f, g \) exist. We shall show that if \( f(x, \theta) \) and \( g(x, \theta) \) are continuous function of \( \theta \), further more if both \( f \) and \( g \) are uniformly bounded by a bounded function of \( x \), then the tail product of \( f \) and \( g \) exists.

To show that we need the following result.

**Lemma 3.1**

Let \( A \subset \mathbb{R} \) and let \( B \) be a closed subset of \( \mathbb{C}^p \). If \( f: A \times B \to \mathbb{C} \) is continuous, and if \( B_1 \) is a bounded subset of \( B \), then \( \sup \text{Re}(f(x,y)) \) and \( \sup \text{Im}(f(x,y)) \) for \( y \in B_1 \) are continuous functions of \( x \).

**Proof:**

The proof is similar to the proof of Lemma 1 of Jennrich(1969). Details are given in Kundu(1989).
The following theorem is a simple extension of the Helly Bray theorem to the case of complex valued functions.

**Theorem 3.1**

Let $X$ be a Euclidean space and let $\Theta$ be a compact set $\mathbb{C}^p$. Let $f: X \times \Theta \rightarrow \mathbb{C}$ where $f$ is bounded and continuous. If $F_1, F_2, \ldots$ are distribution functions on $X$ which converges in distribution to $F$, then

$$\int f(x, \theta) dF_n(x) \rightarrow \int f(x, \theta) dF(x)$$

uniformly for all $\theta \in \Theta$.

**Proof:**

The proof follows in the similar manner as the proof of Theorem 1 of Jennrich (1969). The readers are referred to Kundu (1989).

**Lemma 3.2**

Let $f: Y \times \Theta \rightarrow \mathbb{C}$, where $\Theta$ is a closed subset of a complex Euclidean space $\mathbb{C}^p$ and $Y$ is a measurable space. $f(y, \theta)$ is a continuous function of $\theta$ for each $y \in Y$ and it is a measurable function of $Y$ for each $\theta \in \Theta$. If $\Theta_1$ is a bounded subset of $\Theta$ then $\sup_{\theta \in \Theta_1} \text{Re}(f(y, \theta))$ and $\sup_{\theta \in \Theta_1} \text{Im}(f(y, \theta))$ are measurable functions of $y$.

**Proof:**

$$\sup_{\theta \in \Theta_1} \text{Re}(f(y, \theta)) = \sup_{\theta \in \Theta_1} \text{Re}(f(y, \theta)),$$

where $\Theta_1$ is the closure of $\Theta_1$. Since $\Theta_1 \subset \mathbb{C}^p$, there exists a countable collection, of $\{\Theta_1^n\}$ of subsets of $\Theta_1$ such that $\Theta_1^n \uparrow \Theta_1$. Therefore

**Theorem 3.1**

Let $X$ be a Euclidean space and let $\Theta$ be a compact subset of $\mathbb{C}^p$. Let $f: X \times \Theta \rightarrow \mathbb{C}$, where for each fixed $x$, $f$ is a continuous function of $\theta$, and for each fixed $\theta$ both the real and the imaginary parts of $f$ are measurable.
function of $x$. Assume that there exists a bounded continuous function $g$ such that $\int g(x) dF(x) \leq \infty$ and $\|f(x, \theta)\| \leq g(x)$ for all $x$ and $\theta$. Here $F$ is some distribution function on $X$. If $x_1, x_2, \ldots$ is a random sample from $F$, then for almost every $x = \{x_1, x_2, \ldots\}$

$$\frac{1}{n} \sum_{i=1}^{n} f(x_i, \theta) - \int f(x, \theta) dF(x)$$

uniformly for all $\theta \in \Theta$.

**Proof:**

The proof is very similar to the proof of Theorem 3.1, therefore it is omitted here. The readers are referred to Kundu (1989).

Let $X$ be a sample space and $f, g$ be two complex valued, bounded, continuous functions defined on $X \times \Theta$. Let $x_1, x_2, \ldots$ be a random sample from a distribution function $F$ on $X$. Then it is well known that the sample distribution function $F_n$ converges to $F$ weakly.

Let $f_k(\theta) = f(x_k, \theta)$ and $g_k(\theta) = g(x_k, \theta)$. Then with the help of Theorem 3.1, we can conclude

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f_k(\theta_1) g_k(\theta_2) = \lim_{n \to \infty} \int f(x, \theta_1) g(x, \theta_2) dF_n(x)$$

$$= \int f(x, \theta_1) g(x, \theta_2) dF(x)$$

and this convergence is uniform over $\theta_1, \theta_2 \in \Theta$.

Suppose $f$ and $g$ are continuous functions of $\theta$ for each fixed $x$ and measurable functions of $x$ for each fixed $\theta$. If both $f$ and $g$ are uniformly bounded by a bounded continuous function of $x$, then we can conclude (3.1) with the help of Theorem 3.2.

4. EXISTENCE AND CONSISTENCY OF THE LNSE

Jennrich (1969) first proved the existence of the least squares estimator when the parameter space was real. For the complex parameter space, the existence of the LNSE follows from Prakasa Rao (1987).

Now we would like to establish sufficient conditions under which the LNSE is strongly consistent. These sufficient conditions are similar to those
given by Jennrich (1969) and Bunke and Schmidt (1980). Our conditions coincide with those of Jennrich in the real parameter case.

Let's consider,

Assumption 1. \( \{ \epsilon_t \}, \ t=1,2, \ldots \) is a sequence of i.i.d. complex-valued random variables with \( E(\text{Re}(\epsilon_t)) = 0 = E(\text{Im}(\epsilon_t)) \), \( \text{Var}(\text{Re}(\epsilon_t)) = \sigma_1^2 \), \( \text{Var}(\text{Im}(\epsilon_t)) = \sigma_2^2 \), \( \sigma_1, \sigma_2 > 0 \). \( \text{Re}(\epsilon_t) \) and \( \text{Im}(\epsilon_t) \) are independent of each other.

Assumption 2. The tail cross product of the function \( f_t(\theta) = f(x_t, \theta) \) with itself exists, and

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} |f_t(\theta) - f_t(\theta_0)|^2 = R(\theta)
\]

has a unique minimum of \( \theta = \theta_0 \) where \( \theta_0 \) is an interior point of \( \Theta \).

Note that under the assumption of the existence of the tail cross product of a sequence of functions \( \{ f_t(\theta) \} \), the above limit always exists.

The following theorem establishes the strong consistency of the LNSE under assumptions 1 and 2.

**Theorem 4.1**

Let \( \{ \hat{\theta}_n \} \) be a sequence of LNSE's of the model (1.2). Let \( \{ \epsilon_t \} \) and \( \{ f_t \} \) satisfy Assumptions 1 and 2 respectively. Then \( \hat{\theta}_n \) and \( \sigma_n^2 = R_n(\hat{\theta}_n) \) are strongly consistent estimators of \( \theta_0 \) and \( \sigma_1^2 + \sigma_2^2 \) respectively.

**Proof:**

It follows from the strong law of large numbers that

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} |\epsilon_t|^2 = \sigma_1^2 + \sigma_2^2 \quad (4.1)
\]

Let \( \{ \epsilon_t \}_{t=1}^{\infty} \) denote a sequence that satisfies (4.1). Then

\[
R_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} |f_t(\theta_0) - f_t(\theta) + \epsilon_t|^2
\]
Since the tail cross product of $f$ with itself exists, the second and the third term go to zero uniformly. Therefore

$$
\lim_{n \to \infty} R_n(\theta) = R(\theta) + (\sigma_1^2 + \sigma_2^2)
$$

uniformly for $\theta \in \Theta$. Let $\theta_n = \hat{\theta}_n(\epsilon_t)$ and $\theta'$ be a limit point of the sequence $\theta_n$. Then there exists a subsequence $(n_k)$ such that $\lim_{n \to \infty} \theta_n = \theta'$. By the continuity of $R$ and the uniform convergence of $R_n$,

$$
\lim_{n \to \infty} R_n(\theta_n) = R(\theta) + (\sigma_1^2 + \sigma_2^2)
$$

Since $\theta_n$ is a LNSE, we have

$$
R_n(\theta_n) \leq R_n(\theta_0) = \frac{1}{n} \sum_{t=1}^{n} |\epsilon_t|^2
$$

Therefore, it follows by letting $n \to \infty$ that

$$
R(\theta') + \sigma_1^2 + \sigma_2^2 \leq \sigma_1^2 + \sigma_2^2
$$

This implies that $R(\theta') = 0$. Since $R(\theta)$ has a unique minimum at $\theta_0$, this implies that $\theta = \theta_0$.

Therefore, $\theta_n$ converges to $\theta_0$ almost surely. Replacement of $\hat{\theta}_n$ in $R_n$, implies that $R_n(\hat{\theta}_n)$ converges to $R(\theta_0) + \sigma_1^2 + \sigma_2^2$. Since $R(\theta_0) = 0$, this implies $R_n(\hat{\theta}_n)$ a strongly consistent estimator of $\sigma_1^2 + \sigma_2^2$.

We next establish the asymptotic normality of $\hat{\theta}_n$. To this end, we need some differentiability condition on the real and imaginary parts of the function $f$.

For notational convenience, we write

$$
f_t(\theta) = g_t(\theta) + j h_t(\theta) t = 1, 2, \ldots, n
$$

where $g_t, h_t$ denote the real and the imaginary parts of $f$. Similarly we write

$$
y_t = u_t + j v_t
$$
where $u_t, v_t$ denote the real and the imaginary parts of $y_t$ respectively.

In this section, we shall think of the parameter space $\Theta \subset \mathbb{C}^p$ as a real-valued parameter space $\Theta^* \subset \mathbb{R}^{2p}$. Observe that $\gamma \in \Theta^*$ iff $\theta \in \Theta$ where $\gamma = \begin{bmatrix} \theta_1^T, \theta_2^T \end{bmatrix}^T$ and $\theta = \theta_1 + j\theta_2$.

We need the following derivatives:

$$
g_{ii}(\gamma) = \frac{\partial}{\partial \gamma_i} g_i(\gamma); \quad h_{ii}(\gamma) = \frac{\partial}{\partial \gamma_i} h_i(\gamma); \quad i = 1, 2, \ldots, 2p$$

$$
g_{ij}(\gamma) = \frac{\partial}{\partial \gamma_i \partial \gamma_j} g_i(\gamma); \quad g_{ij}(\gamma) = \frac{\partial}{\partial \gamma_i \partial \gamma_j} h_i(\gamma); \quad i = 1, 2, \ldots, 2p$$

Assumption 3: The derivatives $g_{ii}, g_{ij}, h_{ii}, h_{ij}$ all exist for $i, j = 1, 2, \ldots, 2p$ and are continuous on $\Theta^*$. All possible tail cross products between $f$ and $g$ exist. For each $\gamma \in \Theta^*$, define

$$
o_{ij}(\gamma) = \frac{1}{n} \sum_{i=1}^n g_i(\gamma) g_j(\gamma) + \frac{1}{n} \sum_{i=1}^n h_i(\gamma) h_j(\gamma); i, j = 1, 2, \ldots, 2p$$

and let $A_n(\gamma)$ denote the matrix whose $(i,j)$th element is $A_{ij}(\gamma)$. Let $A_{ij}(\gamma) = \lim_{n \to \infty} o_{ij}(\gamma)$ and let $A(\gamma) = \lim_{n \to \infty} A_n(\gamma)$.

Assumption 1.

$A(\gamma_0)$ is nonsingular, where

$$\gamma_0 = \begin{bmatrix} \theta_1^T, \theta_2^T \end{bmatrix}^T, \theta_0 = \theta_1 + j\theta_2 \text{ and } \sigma_\gamma^2 = \sigma^2$

**Theorem 4.2**

Let $\hat{\theta}_n$ be a sequence of LNSE's of $\theta_0$. Let us write $\hat{\gamma}_n = \begin{bmatrix} \hat{\theta}_1^T, \hat{\theta}_2^T \end{bmatrix}^T$, where $\hat{\theta}_1 + j\hat{\theta}_2$. Then under the assumptions 1-4

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \xrightarrow{d} N_{2p}(0, A^{-1}(\gamma_0))$$

**Proof:**

Since $\hat{\gamma}_n \to \gamma_0$ a.s., therefore for sufficiently large $n$, $\hat{\gamma}_n$ lies in a convex, compact neighbourhood of $\gamma_0$, Wu(1981). Consider the following
where \( \gamma_n, \gamma_n^*, \gamma_n, \cdots, \gamma_n \) are measurable functions lying on the line segment joining \( \gamma_n \) and \( \gamma_0 \) (Jennrich, 1969),

\[
\dot{\gamma}_n = [\dot{\gamma}_{n1}, \ldots, \dot{\gamma}_{np}]^T, \quad \gamma_0 = [\gamma_{n1}, \ldots, \gamma_{np}]^T.
\]

Observe that \( g_{ij}(\gamma_n) = h_{ij}(\gamma_n) = 0 \) for \( i = 1, 2, \ldots, 2p \). If we multiply both sides of the above equation by \( \sqrt{n} \), we get

\[
\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} \left[ g_{i1}(\gamma_n)(u_i - g_{i1}(\gamma_0)) + h_{i1}(\gamma_0)(v_i - h_{i1}(\gamma_0)) \right] \right) \xrightarrow{d} \mathcal{N}_{2p}(0, \sigma^2 A(\gamma_0))
\]

[See Rao (1973), p.387]

Since \( A(\gamma_0) \) is non-singular, we have

\[
\sqrt{n}(\dot{\gamma}_n - \gamma_0) \xrightarrow{d} \mathcal{N}_{2p}(0, \sigma^2 A^{-1}(\gamma_0))
\]

5. EXAMPLE

Consider the model

\[
y_t = \sum_{k=1}^{M} \alpha_k \exp[2\pi i k t] + \varepsilon_t, \quad t = 1, 2, \ldots, n
\]
Here $a_i$'s are bounded complex numbers and $\beta_i$'s are real numbers lying between 0 and 1. With the help of Theorem 3.1, it is easy to see that it satisfies Assumptions 1 and 2. Therefore the LNSE's obtained from here will be consistent.

Acknowledgements: The author was a Ph. D. student at Pennsylvania State University under the supervision of Professor C.R. Rao. The author wishes to thank him for all his guidance in the preparation of this paper.

BIBLIOGRAPHY


Received December 1990; Revised July 1991.

Recommended by D. B. Owen, Southern Methodist University, Dallas, TX.