The Annals of Statistics
1982, Vol. 10, No. 1. 316–319

THE CONSISTENCY OF NONLINEAR REGRESSION MINIMIZING THE L_1-NORM

BY WALTER OBERHOFER

Institute für Volkswirtschaftslehre

We consider conditions in a nonlinear regression model for the consistency of the estimator obtained by minimizing the L_1-norm, i.e. the sum of absolute deviations.

1. Introduction. We consider the following nonlinear regression model for a univariate response y

(1) $$y_t = g(x_t, b_0) + e_t, \qquad 1 \le t \le T,$$

where $x_t \in \mathbb{R}^m$ denotes the tth fixed input vector, $b_0 \in \mathbb{R}^p$ is the vector of parameters, e_t is the tth error residual and g is a function $g: \mathbb{R}^{m+p} \to \mathbb{R}$. We denote the vector of all responses by $y = (y_1, \cdots, y_T)'$.

Given an observation y, any vector b minimizing

(2) $$Q_T(y, b) = \frac{1}{T} \sum_{t=1}^{T} | y_t - g(x_t, b) |$$

will be called an L_1-norm estimate of b_0 based on the observation vector y, and will be denoted by $b(y)$.

In Section 2, we establish conditions which will guarantee the existence and consistency of $b(y)$. Problems of L_1-norm estimation are discussed in Bassett and Koenker (1978) and Taylor (1974). In the former, the authors show that the relative efficiency of the L_1-norm estimator to the least squares estimator is the same as that of the sample median to the sample mean. Results for nonlinear least squares estimation are given by Malinvaud (1970) and Jennrich (1969). A relationship also exists between the problems treated in this paper and the so-called monotone median regression; see Cryer et al. (1972).

2. Existence and consistency of L_1-norm estimator. It is necessary to minimize the function (2) as a function of b over a given domain $K \subset \mathbb{R}^p$. In order to guarantee the existence of $b(y)$ minimizing (2) we assume

A1: $b_0 \in K \subset \mathbb{R}^p$ and K is compact
A2: For all t, the function $g(x_t, b)$ is defined for all $b \in K$ and continuous in K as a function of b.

Existence is then an immediate corollary of Lemma 2 of Jennrich (1969).

In addition to A1 and A2, we will assume
A3: b_0 is an inner point of K.

Furthermore, by $K_0 \subset K$, we denote a closed set not containing b_0. Due to the fact that

(3) $$Q_T(y, b(y)) - Q_T(y, b) \le 0 \quad \text{for every} \quad b \in K,$$

the (weak) consistency of $b(y)$ ensues provided we can demonstrate that for any K_0 and every $b \in K_0$

(4) $$\lim_{T\to\infty} P\{\inf_{b \in K_0} Q_T(y, b) - Q_T(y, b_0) > 0\} = 1.$$

Received September 1978; revised May 1981.
AMS 1980 *subject classifications.* Primary 62F12, 62J07.
Key words and phrases. Nonlinear regression, L_1-norm, consistency.

If we define $h_t(b) = g(x_t, b_0) - g(x_t, b)$, then by (1) and (2) we obtain.

(5) $$Q_T(y, b) - Q_T(y, b_0) = \frac{1}{T}\sum_t \{|e_t + h_t(b)| - |e_t|\}.$$

For the proof of (4) we require the following lemmas.

LEMMA 1. *If we assume*
A4: (e_t) *is a sequence of independent random variables,*
A5: $\lim_{T\to\infty} \frac{1}{T^2}\sum_t \{h_t(b)\}^2 = 0$ *for all* $b \in K$, *then we have*

(6) $$p\lim_{T\to\infty} \frac{1}{T}\sum_t[|e_t + h_t(b)| - |e_t| - E\{|e_t + h_t(b)| - |e_t|\}] = 0.$$

PROOF. It is obvious that

$$|e_t + h_t(b)| - |e_t| \leq |h_t(b)|.$$

Therefore, the random variable in (6) has mean zero and variance bounded by

$$T^{-2}\sum_t h_t^2(b),$$

which converges to 0 as $T \to \infty$ from the assumption. The lemma follows from Chebychev's inequality. □

LEMMA 2. *Denote the distribution function of* e_t *by* $F_t(x)$. *We assume*
A6: $F_t(0) = \frac{1}{2}$.

Define the function

(7) $$G_t(z) = \begin{cases} z\{1 - 2F_t(-z)\} - 2\int_{(0,-z]} x dF_t(x) & \text{for } z \leq 0 \\ z\{1 - 2F_t(-z)\} + 2\int_{(-z,0]} x dF_t(x) & \text{for } z > 0. \end{cases}$$

Then, by integration of the piecewise linear function $|e + h_t(b)| - |e|$, we derive

$$\frac{1}{T}\sum_{t=1}^T E(|e_t + h_t(b)| - |e_t|) = \frac{1}{T}\sum_{t=1}^T G_t(h_t(b)),$$

and using A6

$$\frac{1}{T}\sum_t E\{|e_t + h_t(b)| - |e_t|\} = 2\frac{1}{T_{h_t}}\sum_{(b)\leq 0}\int_{(0,-h_t(b)]} \{|h_t(b)| - x\}\, dF_t(x)$$
$$+ 2\frac{1}{T_{h_t}}\sum_{(b)>0}\int_{(-h_t(b),0]} \{|h_t(b)| + x\}\, dF_t(x).$$

If we restrict the domains of integration to $(0, -h_t(b)/2]$, resp. to $(-h_t(b)/2, 0]$, and again use A6, we obtain

$$\frac{1}{T}\sum_t E\{|e_t + h_t(b)| - |e_t|\}$$
$$\geq \frac{1}{T_{h_t}}\sum_{(b)\leq 0}|h_t(b)|\left\{F_t(-h_t(b)/2) - \frac{1}{2}\right\} + \frac{1}{T_{h_t}}\sum_{(b)>0}|h_t(b)|\left\{\frac{1}{2} - F_t(-h_t(b)/2)\right\}$$
$$\geq \frac{1}{T}\sum_{t=1}^T |h_t(b)| \min\left\{F_t(|h_t(b)|/2) - \frac{1}{2}, \frac{1}{2} - F_t(-|h_t(b)|/2)\right\}.$$ □

LEMMA 3. *For all b, $b^* \in K$ we have*

$$\frac{1}{T}\sum_t \{ |e_t + h_t(b)| - |e_t + h_t(b^*)| \} \leq \frac{1}{T}\sum_t |g(x_t, b) - g(x_t, b^*)|$$

due to the inequality

$$|e_t + h_t(b)| - |e_t + h_t(b^*)| \leq |h_t(b) - h_t(b^*)| = |g(x_t, b) - g(x_t, b^*)|. \qquad \square$$

We shall now prove the consistency of $b(y)$ under the following assumptions, the first five of which imply A1 through A6.

B1: b_0 is an inner point of a compact set $K \subset \mathbb{R}^p$.

B2: For all t, the function $g(x_t, b)$ is defined for $b \in K$ and is continuous in K.

B3: (e_t) is a sequence of independent random variables.

B4: The distribution function of e_t is denoted by $F_t(x)$ and $F_t(0) = ½$.

B5: The function

$$\frac{1}{T}\sum_{t=1}^{T} \{g(x_t, b) - g(x_t, b^*)\}^2$$

is continuous in $b \in K$ for every $b^* \in K$ and uniformly for T.

B6: If we define $h_t(b) = g(x_t, b) - g(x_t, b_0)$ and if $K_0 \subset K$ is a closed set not containing b_0, then there exist numbers $\varepsilon > 0$ and T_0 such that for all $T \geq T_0$

$$\inf_{K_0} \frac{1}{T}\sum_{t=1}^{T} |h_t(b)| \min\{F_t(|h_t(b)|/2) - \frac{1}{2}, \frac{1}{2} - F_t(-|h_t(b)|/2)\} \geq \varepsilon.$$

THEOREM. *For the nonlinear regression model* (1), *under assumptions* B1 *through* B6, *the estimator $b(y)$, minimizing the L_1-norm of the errors, is weakly consistent.*

PROOF. We have to prove inequality (4). We take an arbitrary $b^* \in K_0 \subset K$, where K_0 is a closed set not containing b_0. Under B6, according to Lemma 2, there exist an $\varepsilon > 0$ and a T_0, such that for all $T \geq T_0$

$$\frac{1}{T}\sum_t E\{ |e_t + h_t(b^*)| - |e_t| \} \geq \varepsilon.$$

Due to B5 and Lemma 1

$$\lim_{T\to\infty} P\left[\frac{1}{T}\sum_t \{ |e_t + h_t(b^*)| - |e_t| \} \geq \frac{\varepsilon}{2}\right] = 1.$$

Finally, using this fact with B5, Lemma 3 and the Cauchy-Schwarz inequality for all $b \in V$, where V is a suitable neighborhood of b^*, we obtain

$$\lim_{T\to\infty} P\left[\inf_V \frac{1}{T}\sum_t \{ |e_t + h_t(b)| - |e_t| \} \geq \frac{\varepsilon}{4}\right] = 1. \tag{8}$$

Accordingly, a neighborhood V is associated with every $b^* \in K_0$. Because of the compactness of K_0, there exists a finite set of such neighborhoods $(V_1, V_2, \cdots, V_n)$ covering K_0. Therefore we have proved (8) with V replaced by K_0. The proof is thus complete.

3. Discussion of assumptions. Assumptions B1, B2 and B3 are also made in nonlinear least squares estimation. See Malinvaud (1970) and Jennrich (1969). The normalisation $F_t(0) = ½$ (characteristic also in the monotone regression, see Cryer et al, 1972) occurs in place of the least squares regression normalisation $Ee_t = 0$. The Assumption B5 corresponds with Assumption 4 in Malinvaud's (1970) paper. Of paramount interest is Assumption B6. As we will demonstrate, this basically is an identifiability assumption. We obtain $G_t(0) = 0$. If *a* b^* exists, such that $g(x_t, b^*) = g(x_t, b_0)$ for all t, then B6 is not fulfilled and b_0 is not identified. In the following we seek conditions which will guarantee B6.

LEMMA 4. *Assumption* B6 *is fulfilled if the following two assumptions hold.*
B7: For every closed set K_0 not containing b_0, there exist numbers $c > 0$, $d > 0$ and $T_0 > 0$ such that for all $b \in K_0$ and all $T \geq T_0$

$$|\{t \mid t \leq T, |h_t(b)| \geq c\}|/T \geq d > 0.$$

B8: For every $c > 0$, there exists a real number $f > 0$, such that for all t

$$\min[F_t(c) - 1/2, 1/2 - F_t(-c)] \geq f > 0.$$

PROOF. Given a closed set K_0 not containing b_0, due to B7 we obtain numbers $c > 0$, $d > 0$ and T_0; further according to B8 we have a $f > 0$ such that for all $T \geq T_0$

$$\inf_{K_0} \frac{1}{T} \sum_{t=1}^{T} |h_t(b)| \min\{F_t(|h_t(b)|/2) - \frac{1}{2}, \frac{1}{2} - F_t(-|h_t(b)/2)\}$$
$$\geq \inf_{K_0} \frac{1}{T} \sum_{|h_t(b)| \geq c} |h_t(b)| f \geq \text{cdf} > 0.$$

Therefore, Lemma 4 is proved if ε is replaced by cdf > 0.

Assumption B7 occurs in place of the least squares regression assumption that, for every closed set $K_0 \subset K$ not containing b_0, there exist numbers $\varepsilon > 0$ and T_0 such that for $T \geq T_0$

$$\inf_{K_0} \frac{1}{T} \sum_t \{h_t(b)\}^2 \geq \varepsilon.$$

It can easily be shown that if B6 is fulfilled, then due to the fact that $\min\{F_t(c) - 1/2, 1/2 - F_t(-c)\}$ is bounded above, for every closed set $K_0 \subset K$ not containing b_0 there exist numbers $\varepsilon > 0$ and T_0 such that for $T \geq T_0$

$$\inf_{K_0} \frac{1}{T} \sum_t |h_t(b)| \geq \varepsilon.$$

In the linear case and for $m = 1$, i.e. when

$$g(x_t, b) = b_1 + b_2 x_t,$$

a sufficient condition for B7 is that for any non-empty interval I of arbitrarily small length

$$\lim_T \inf|\{t \mid t \leq T, x_t \text{ outside of } I\}|/T > 0.$$

Acknowledgment. The author wishes to thank a referee for several comments and suggestions.

REFERENCES

BASSETT, G. and KOENKER, R. (1978). Asymptotic theory of least absolute error regression. *J. Amer. Statist. Assoc.* **73** 618–622.

CRYER, J. D. et al. (1972). Monotone median regression. *Ann. Math. Statist.* **43** No. 5, 1459–1469.

JENNRICH, R. I. (1969). Asymptotic properties of nonlinear least squares estimators. *Ann. Math. Statist.* **40** No. 2, 633–643.

MALINVAUD, E. (1970). The consistency of nonlinear regression. *Ann. Math. Statist.* **41** No. 3, 956–969.

TAYLOR, L. D. (1974). Estimation by minimizing the sum of absoute errors. In *Frontiers in Econometrics* (ed. Zarembka) Academic, New York.

UNIVERSITÄT REGENSBURG
FACHBEREICH WIRTSCHAFTSWISSENSCHAFT
INSTITUTE FÜR VOLKSWIRTSCHAFTSLEHRE
8400 REGENSBURG
UNIVERSITÄTSSTRAßE 31
POSTFACH, WEST GERMANY