Statistica Sinica 3(1993), 9-18

ON CONDITIONS OF CONSISTENCY OF ML_1N ESTIMATES

Xiru Chen, Lincheng Zhao and Yuehua Wu

University of Science and Technology of China and York University

Abstract: Chen and Wu (1992) proposed a conjecture on the consistency of the Minimum L_1-Norm (ML_1N) estimator in linear regression models. Here we prove two theorems showing that the conjecture is true under some circumstances.

Key words and phrases: Minimum L_1-norm estimate, consistency, linear regression model.

1. Introduction

Consider the linear regression model

$$Y_i = x_i'\beta_0 + e_i, \quad i = 1, \ldots, n, \ldots \tag{1.1}$$

where $Y_1, Y_2, \ldots$ are observations of the dependent variable Y and $x_1, x_2, \ldots$ are known p-vectors, β_0 is the unknown p-vector of regression coefficients, and $e_1, e_2, \ldots$ are unobservable random errors. For simplicity, we shall assume throughout this paper that $e_1, e_2, \ldots$ are i.i.d. with a common distribution function F. This condition will not be mentioned in the sequel.

The ML_1N estimate of β_0, to be denoted by $\hat{\beta}_n$, is defined as a Borel-measurable solution of the minimization problem

$$\min \sum_{i=1}^{n} \left| Y_i - x_i'\beta \right|. \tag{1.2}$$

The usual Least Squares (LS) estimate of β_0 is denoted by $\tilde{\beta}_n$.

Some general sufficient conditions about the (weak and strong) consistency of $\hat{\beta}_n$ are known (see Chen, Bai, Zhao and Wu (1990, 1992)), but not necessary conditions. Neither are the necessary and sufficient conditions. Chen and Wu (1992) proposed the following conjecture about this problem.

Conjecture. Assume that

(A1) $\text{med}(e_1) = 0$, and

(A2) in some neighborhood of 0, say $(-\delta, \delta)$, the derivative f of F exists and is bounded away from 0 and ∞.

Then the condition

$$\lim_{n\to\infty} S_n^{-1} = 0 \quad (S_n = \sum_{i=1}^{n} x_i x_i') \tag{1.3}$$

is both necessary and sufficient for the weak consistency of the ML_1N estimate $\hat{\beta}_n$.

Condition (1.3) plays a prominent role in the asymptotic theory of the LS estimate $\tilde{\beta}_n$. For example, a well known result (see Drygas (1976) and Lai, Robbins and Wei (1979)) is that under the conditions

$$E(e_1) = 0, \quad 0 < \mathrm{Var}(e_1) < \infty, \tag{1.4}$$

(1.3) is necessary and sufficient for both weak and strong consistency of $\tilde{\beta}_n$. It also plays an important role in the ML_1N method (see Chen, Bai, Zhao and Wu (1990, 1992)).

We hit upon the idea of this Conjecture for the following reasons. First, S_n plays an important role in the consistency problems mentioned above. Second, in the study of the consistency of $\hat{\beta}_n$, we gain the impression that the two conditions in (1.4) in the theory of $\tilde{\beta}_n$ are comparable to the two conditions (A1) and (A2) in the Conjecture. What is essentially important in the asymptotic theory of $\hat{\beta}_n$ is the local behavior of the distribution of e_1 in the vicinity of 0, while the tail-probability of e_1 plays a decisive role in the asymptotic theory of $\tilde{\beta}_n$.

Chen and Wu (1992) stated some results concerning this conjecture without proof. In this paper we give detailed proofs of the following theorems, showing that the conjecture is true under some circumstances.

Theorem 1. *In addition to conditions* (A1) *and* (A2), *assume that* $\{x_i\}$ *is bounded, and* f *is absolutely continuous in* $(-\delta, \delta)$ *for some* $\delta > 0$, *and*

$$\int_{-\delta}^{\delta} \left(f'(x)\right)^2 dx < \infty. \tag{1.5}$$

Then (1.3) *is necessary and sufficient* (*n.* & *s.*) *for the weak consistency of the* ML_1N *estimate* $\hat{\beta}_n$ *of* β_0.

To state Theorem 2, we introduce the following weaker condition (A3) instead of (A2):

(A3) There exist positive constants δ, c and C such that

$$c|x| \le p(x) \le C|x|, \quad \text{for } |x| \le \delta, \tag{1.6}$$

where

$$p(x) = \begin{cases} P(0 \le e_1 < x), & \text{for} \quad x \ge 0, \\ P(x < e_1 \le 0), & \text{for} \quad x < 0. \end{cases} \tag{1.7}$$

Note that (A3) is equivalent to the following:

(A3′) For any constant $\Delta > 0$, there exist positive constants c and C such that

$$c|x| \le p(x) \le C|x|, \quad \text{for } |x| \le \Delta. \tag{1.8}$$

Theorem 2. *Assume dimension $p = 1$ in the model* (1.1) *and that* (A1) *and* (A3) *are satisfied. Then* (1.3) *is a n. & s. condition for the weak consistency of the ML_1N estimate $\hat{\beta}_n$ of β_0.*

Theorem 2 is an improved version of Theorem 6 in Chen and Wu (1992). We would also like to point out an important fact that under conditions (A1) and (A2), the n. & s. conditions for weak and strong consistency of $\hat{\beta}_n$ cannot be the same. This is illustrated by an example in Chen and Wu (1992). This is a major difference between the two estimates $\hat{\beta}_n$ and $\tilde{\beta}_n$.

2. Proof of Theorem 1

The sufficiency follows immediately from Theorem 2.1, 1° of Chen, Bai, Zhao and Wu (1990), which states that under (A1) and another condition weaker than (A2), we have

$$d_n \longrightarrow 0 \Longrightarrow \hat{\beta}_n \longrightarrow \beta_0 \qquad \text{in Pr.,} \tag{2.1}$$

where $d_n = \max_{1 \le i \le n} x_i S_n^{-1} x_i$. When $\{x_i\}$ is bounded, $S_n^{-1} \longrightarrow 0$ implies $d_n \longrightarrow 0$ and sufficiency follows by (2.1).

To prove the necessity, we use a result of Li (1984). To begin with, set $\beta_0 = 0$ without loss of generality, and then restrict the parameter space to

$$H = \{\beta : \|\beta\| \le \varepsilon_0\}, \tag{2.2}$$

where ε_0 is some positive constant to be specified later. This means that in (1.2) the minimum is taken only over H. This can also be done without loss of generality, since it is obvious that if $\hat{\beta}_n$ is not consistent when the parameter space is restricted to H, certainly it cannot be consistent when β is free to run over R^p.

First, consider a special case. Suppose that e_1 only takes values in a finite interval (A, B), $A < 0 < B$ by (A1), and on (A, B) e_1 has a density f which is positive and absolutely continuous. The δ in (1.5) can be chosen so that $A < -\delta < \delta < B$. By (1.5) we have

$$\int_{-\delta}^{\delta} \frac{(f'(x))^2}{f(x)} dx < \infty. \tag{2.3}$$

Define two functions, h_1 on (δ, ∞) and h_2 on $(-\infty, -\delta)$, which satisfy the following set of conditions:

1° $h_1(\delta) = f(\delta)$, $h_2(-\delta) = f(-\delta)$,

2° h_1, h_2 are positive and h_1', h_2' exist and continuous everywhere on their respective intervals of definition,

3° $\int_\delta^\infty h_1(x)dx = \int_\delta^B f(x)dx$, $\int_{-\infty}^{-\delta} h_2(x)dx = \int_A^{-\delta} f(x)dx$,

4° $\int_\delta^\infty \left[(h_1'(x))^2/h_1(x)\right]dx + \int_{-\infty}^{-\delta} \left[(h_2'(x))^2/h_2(x)\right]dx < \infty$,

5° $\int_\delta^\infty x^2 h_1(x)dx + \int_{-\infty}^{-\delta} x^2 h_2(x)dx < \infty$,

6° $\int_{-\infty}^{-\delta} x h_2(x)dx + \int_{-\delta}^{\delta} x f(x)dx + \int_\delta^\infty x h_1(x)dx = 0$.

The existence of such functions is trivial. For example, we can take $\exp(ax^2 + bx + c)$ with appropriate a, b and c as such a function. Now define function $g_1(x)$ on (δ, B) and function $g_2(x)$ on $(A, -\delta)$, such that $g_1(\delta) = \delta$, $g_2(-\delta) = -\delta$ and

$$\begin{aligned} \int_\delta^{g_1(x)} h_1(u)du &= \int_\delta^x f(u)du, \quad \delta < x < B, \\ \int_{g_2(x)}^{-\delta} h_2(u)du &= \int_x^{-\delta} f(u)du, \quad A < x < -\delta, \end{aligned}$$

and define a sequence of random variables $\{\tilde{e}_i\}$ as follows:

$$\tilde{e}_i = \begin{cases} e_i, & \text{when } |e_i| \le \delta, \\ g_1(e_i), & \text{when } e_i > \delta, \\ g_2(e_i), & \text{when } e_i < -\delta. \end{cases} \qquad i = 1, 2, \ldots \tag{2.4}$$

Note that g_1 and g_2 are strictly increasing in their respective domain, due to the fact that f is positive on (A, B). Now $\tilde{e}_1$, $\tilde{e}_2$, ... are i.i.d., and simple calculations show that $\tilde{e}_1$ has a density function

$$\tilde{f}(x) = \begin{cases} f(x), & \text{when } |x| \le \delta, \\ h_1(x), & \text{when } x > \delta, \\ h_2(x), & \text{when } x < -\delta. \end{cases}$$

From the assumptions on f and definitions of h_1 and h_2, it is seen that $\tilde{e}_i$ and $\tilde{f}$ satisfy the following conditions:

$$\text{med}(\tilde{e}_1) = 0,\ E(\tilde{e}_1) = 0,\ 0 < \text{Var}(\tilde{e}_1) < \infty, \tag{2.5}$$

$$e_i > \delta \Longleftrightarrow \tilde{e}_i > \delta, \quad e_i < -\delta \Longleftrightarrow \tilde{e}_i < -\delta, \tag{2.6}$$

$$\int_{-\infty}^{\infty} \left[(\tilde{f}'(x))^2/\tilde{f}(x)\right] dx < \infty. \tag{2.7}$$

Denote $M = \sup_i \|x_i\|$. Choose $\varepsilon_0 = \delta/(2M)$ for the definition of H (see (2.2)). Introduce the linear model

$$\tilde{Y}_i = x_i'\tilde{\beta}_0 + \tilde{e}_i, \quad i = 1, \dots, n, \dots \tag{2.8}$$

with $\tilde{\beta}_0 = 0$ and the parameter space H. Consider the ML$_1$N estimate β_n^* of $\tilde{\beta}_0$ in this model, which is a solution of the minimization problem

$$\min \sum_{i=1}^{n} |\tilde{Y}_i - x_i'\beta| = \min \sum_{i=1}^{n} |\tilde{e}_i - x_i'\beta|. \tag{2.9}$$

Write

$$\sum_{i=1}^{n} |\tilde{e}_i - x_i'\beta| = \sum_{(1)} |\tilde{e}_i - x_i'\beta| + \sum_{(2)} |\tilde{e}_i - x_i'\beta|, \tag{2.10}$$

where $\sum_{(1)}$ contains those terms for which $|\tilde{e}_i| \le \delta$ and $\sum_{(2)}$ contains the remaining terms in $\sum_{i=1}^{n}$. From (2.4) and (2.6), and the fact that $|x_i'\beta| \le M\varepsilon_0 < \delta$ for $\beta \in H$, it follows that

$$\sum_{(1)} |\tilde{e}_i - x_i'\beta| = \sum_{(1)} |e_i - x_i'\beta|, \tag{2.11}$$

$$\sum_{(2)} |\tilde{e}_i - x_i'\beta| = \sum_{(2)} |e_i - x_i'\beta| + \sum_{(2)} (\tilde{e}_i - e_i)\text{sgn}(\tilde{e}_i), \tag{2.12}$$

where $\text{sgn}(a) = a/|a|$ for $a \ne 0$ and $\text{sgn}(0) = 0$. From (2.10)-(2.12), we have

$$\sum_{i=1}^{n} |\tilde{e}_i - x_i'\beta| = \sum_{i=1}^{n} |e_i - x_i'\beta| + \sum_{(2)} (\tilde{e}_i - e_i)\text{sgn}(\tilde{e}_i). \tag{2.13}$$

The second term of the right hand side of (2.2) does not depend on β. Hence, the solution β_n^* of (2.9) is the same as the solution of (1.2), i.e.,

$$\beta_n^* = \hat{\beta}_n. \tag{2.14}$$

Therefore, if $\hat{\beta}_n$ is a weak consistent estimate of β_0 in model (1.1), then β_n^* is a weak consistent estimate of $\tilde{\beta}_0$ in model (2.8). But by the result proved in Li (1984), under conditions (2.5) and (2.7), there is no consistent estimate of $\tilde{\beta}_0$ in model (2.8) when (1.3) is not true. Therefore, when (1.3) is not true $\hat{\beta}_n$ cannot be a weak consistent estimate of β_0 in model (1.1).

The general case can be reduced to the one discussed above. For this purpose let δ in (1.5) be such that $f(\pm\delta) > 0$. Put

$$G_1 = \frac{P(e_1 > \delta)}{f(\delta)}, \quad G_2 = \frac{P(e_1 < -\delta)}{f(-\delta)},$$

and define a density f^* as follows

$$f^*(x) = \begin{cases} f(x), & \text{when } |x| \le \delta, \\ f(\delta), & \text{when } \delta < x < \delta + G_1, \\ f(-\delta), & \text{when } -\delta - G_2 < x < -\delta, \\ 0, & \text{otherwise.} \end{cases}$$

Then define a sequence of i.i.d. variables $\{e_i^*\}$ with a common density f^*, such that

$$e_i = e_i^* \text{ when } |e_i| \le \delta;\ e_i > \delta \Longleftrightarrow e_i^* > \delta;\ e_i < -\delta \Longleftrightarrow e_i^* < -\delta,$$

and form the linear model $Y_i^* = x_i'\beta_0^* + e_i^*$, $i = 1, \ldots, n$. The same argument used earlier shows that the $\mathrm{M}L_1\mathrm{N}$ estimator of β_0^* in this model is consistent if and only if the $\mathrm{M}L_1\mathrm{N}$ estimator $\hat{\beta}_n$ of the original linear model is consistent. Hence we return to the case considered already, and Theorem 1 is proved.

Remark. A glance at the proof might give the impression that we have proved the nonexistence of consistent estimate of β_0 under (A1) and (A2) when (1.3) is false. This is not so. Indeed, there may exist consistent estimates of β_0, other than the $\mathrm{M}L_1\mathrm{N}$ or LS estimate. For details, see Chen (1992).

3. Proof of Theorem 2

Without loss of generality, assume $\beta_0 = 0$, and $x_i \ne 0$ for all $i \ge 1$.

Sufficiency. As (1.3) is true, we have $\sum_{i=1}^{\infty} x_i^2 = \infty$. Define random variable ξ_n as follows: for given $e^n = (e_1, \ldots, e_n)$, the conditional distribution of ξ_n is given by

$$P^*(\xi_n = e_i/x_i) = |x_i|/D_n \text{ with } D_n = \sum_{i=1}^{n} |x_i|,\ i = 1, \ldots, n. \tag{3.1}$$

In this paper, notation such as P^* and the following $m^*_-(\cdot)$ and $m^*_+(\cdot)$ refer to the probability calculations under the condition that e^n is given. Since

$$\sum_{i=1}^{n}|Y_i - x_i\beta| = D_n \sum_{i=1}^{n}|e_i/x_i - \beta| \cdot |x_i|/D_n, \tag{3.2}$$

we see that an ML_1N estimator of β_0 is a median of the conditional distribution (3.1), and vice versa. Let $m^*_-(\xi_n)$ and $m^*_+(\xi_n)$ be the infimum and supremum of such medians, respectively. To prove that $\hat{\beta}_n$ is consistent weakly, it is sufficient to verify that for any $\varepsilon > 0$,

$$P\Big(m^*_+(\xi_n) \geq \varepsilon\Big) \to 0, \text{ as } n \to \infty, \tag{3.3}$$

$$P\Big(m^*_-(\xi_n) \leq -\varepsilon\Big) \to 0, \text{ as } n \to \infty. \tag{3.4}$$

Take (3.3) for instance. Let $I(\cdot)$ be the indicator and define

$$Z_n = \sum_{i=1}^{n} I(e_i/x_i \geq \varepsilon)|x_i|/D_n = P^*(\xi_n \geq \varepsilon). \tag{3.5}$$

It is easily seen that

$$m^*_+(\xi_n) \geq \varepsilon \Longleftrightarrow Z_n \geq 1/2. \tag{3.6}$$

Hence, to prove (3.3), it suffices to show that

$$P(Z_n \geq 1/2) \to 0, \text{ as } n \to \infty. \tag{3.7}$$

Define

$$p_i = P(0 \leq e_i/x_i < \varepsilon), \tag{3.8}$$

$$q_i = P(\varepsilon \leq e_i/x_i). \tag{3.9}$$

Then, by (A1) and (A3),

$$p_i + q_i = 1/2. \tag{3.10}$$

Therefore,

$$1/2 - EZ_n = A_n/D_n \text{ with } A_n = \sum_{i=1}^{n}|x_i|p_i, \tag{3.11}$$

$$\mathrm{Var}(Z_n) = B_n^2/D_n^2 \text{ with } B_n^2 = \sum_{i=1}^{n} x_i^2 q_i(1-q_i). \tag{3.12}$$

Now we proceed to prove that

$$B_n^2/A_n^2 \to 0, \text{ as } n \to \infty. \tag{3.13}$$

For any $\eta \in (0, 1/4)$, there exists $\delta_1 = \delta_1(\eta, \varepsilon) > 0$ such that

$$|x_i| \geq \delta_1 \text{ implies } q_i < \eta \text{ and hence } p_i > 1/4. \tag{3.14}$$

By (A3′), there are positive constants c_1 and C_1 such that

$$c_1|x_i| \leq p_i \leq C_1|x_i|, \text{ if } |x_i| \leq \delta_1. \tag{3.15}$$

Since $\sum_{i=1}^{\infty} x_i^2 = \infty$, at least one of the following two assertions is true:

(i) $\sum_{i=1}^{\infty} x_i^2 I(|x_i| < \delta_1) = \infty$,

(ii) $\sum_{i=1}^{\infty} |x_i| I(|x_i| \geq \delta_1) = \infty$.

Hence, by (3.15) and (3.14), it follows that $A_n \to \infty$.

By (3.15) and (3.14), we also have

$$\begin{aligned} B_n^2/A_n^2 &\leq A_n^{-2}\Big\{ \sum_{|x_i|<\delta_1} c_1^{-1}|x_i|p_i + 16\eta \sum_{|x_i|\geq\delta_1} x_i^2 p_i^2 \Big\} \\ &\leq c_1^{-1}A_n^{-1} + 16\eta \to 16\eta, \text{ as } n \to \infty. \end{aligned}$$

Since η is arbitrary, (3.13) follows immediately from the above result.

From (3.11)-(3.13), it follows that

$$\begin{aligned} &P(Z_n \geq 1/2) = P(Z_n - EZ_n \geq 1/2 - EZ_n) \\ &\leq \text{Var}(Z_n)/(1/2 - EZ_n)^2 = B_n^2/A_n^2 \to 0, \text{ as } n \to \infty. \end{aligned}$$

This proves (3.7) and, in turn, (3.3). The (3.4) can be shown in the same way. This concludes the proof of sufficiency.

Necessity. Assume that $\sum_{i=1}^{\infty} x_i^2 < \infty$ and proceed to show that $\hat{\beta}_n$ is not consistent. Without loss of generality, we may assume that $x_1 > 0$. Define

$$\tilde{Z}_n = P^*(\xi_n > \varepsilon) = \sum_{i=1}^{n} \alpha_i |x_i|/D_n \text{ with } \alpha_i = I(e_i/x_i > \varepsilon). \tag{3.16}$$

We see that if $\tilde{Z}_n > 1/2$, then $m_-^*(\xi_n) > \varepsilon$. Since $\hat{\beta}_n$ is a median of ξ_n given e^n, we have

$$\tilde{Z}_n > 1/2 \text{ implies that } \hat{\beta}_n > \varepsilon. \tag{3.17}$$

Put

$$\tilde{p}_i = \tilde{p}_i(\varepsilon) = P(0 \leq e_i/x_i \leq \varepsilon), \quad \tilde{q}_i = P(e_i/x_i > \varepsilon), \tag{3.18}$$

we have

$$E\alpha_i = \tilde{q}_i = 1/2 - \tilde{p}_i \quad \text{and} \quad D_n(1/2 - E\tilde{Z}_n) = \sum_{i=1}^{n} |x_i|\tilde{p}_i. \tag{3.19}$$

Since $\sum_{i=1}^{\infty} x_i^2 < \infty$, we have $x_i \to 0$. From this and (A3′), there exist positive constants c_2 and C_2 such that

$$c_2|x_i| \le \tilde{p}_i \le C_2|x_i|, \quad i = 1, 2, \ldots \tag{3.20}$$

It follows that

$$\sum_{i=1}^{\infty} |x_i|\tilde{p}_i(\varepsilon) < \infty \tag{3.21}$$

and $\sum_{i=1}^{\infty} |x_i|\tilde{p}_i(\varepsilon)$ tends to 0 monotonically as $\varepsilon \downarrow 0$. Take $\varepsilon > 0$ such that

$$\sum_{i=1}^{\infty} \tilde{p}_i|x_i| < x_1/4. \tag{3.22}$$

Put

$$\zeta_n = \sum_{i=2}^{n} |x_i|(\alpha_i - E\alpha_i) = \sum_{i=2}^{n} |x_i|(\alpha_i - \tilde{q}_i).$$

Since

$$\sum_{i=2}^{\infty} \operatorname{Var}(|x_i|\alpha_i) = \sum_{i=2}^{\infty} x_i^2 \tilde{q}_i(1 - \tilde{q}_i) < \infty,$$

there is a random variable ζ such that as $n \to \infty$,

$$\zeta_n \to \zeta \quad \text{a.s. with} \quad P(\zeta \ge 0) > 0. \tag{3.23}$$

By (3.19), (3.22), (3.23), and noting that $D_n(\tilde{Z}_n - E\tilde{Z}_n) = x_1(\alpha_1 - \tilde{q}_1) + \zeta_n$, we have

$$\begin{aligned} &\liminf_{n\to\infty} P(\tilde{Z}_n > 1/2) = \liminf_{n\to\infty} P\Big(D_n(\tilde{Z}_n - E\tilde{Z}_n) > \sum_{i=1}^{n} |x_i|\tilde{p}_i\Big) \\ \ge\ & P\Big(x_1(\alpha_1 - \tilde{q}_1) + \zeta > \sum_{i=1}^{\infty} |x_i|\tilde{p}_i\Big) \ge P\Big(x_1(\alpha_1 - \tilde{q}_1) + \zeta \ge x_1/4\Big) \\ \ge\ & P(\alpha_1 - \tilde{q}_1 \ge 1/4)P(\zeta \ge 0) > 0. \end{aligned} \tag{3.24}$$

By (3.17), for the chosen ε we have

$$\liminf_{n\to\infty} P(\hat{\beta}_n > \varepsilon) > 0. \tag{3.25}$$

Therefore, $\hat{\beta}_n$ is not consistent. This proves Theorem 2.

Acknowledgement

The authors would like to thank the referees for their detailed suggestions, which considerably improved the quality of the paper.

This work is supported by the National Natural Science Foundation of China. The research of Wu is also partially supported by the Natural Science and Engineering Research Council of Canada and Faculty Research Grant, Faculty of Arts, York University.

References

Chen, X. R. (1992). On a problem of existence of consistent estimate. In *The Development of Statistics: Recent Contributions from China, Pitman Research Notes in Math. Ser.* **258**, 237–248. Longman Scientific & Technical.

Chen, X. R., Bai, Z. D., Zhao, L. C. and Wu, Y. H. (1990). Asymptotic normality of minimum L_1-norm estimates in linear models. *Chinese Sci. Ser.A* **33**, 449–463.

Chen, X. R., Bai, Z. D., Zhao, L. C. and Wu, Y. H. (1992). Consistency of minimum L_1-norm estimates in linear models. In *The Development of Statistics: Recent Contributions from China, Pitman Research Notes in Math. Ser.* **258**, 249–260. Longman Scientific & Technical.

Chen, X. R. and Wu, Y. H. (1992). On consistency of minimum L_1-norm estimates in linear models. In *Probability and Statistics: Proceedings of the Special Program at Nankai Institute of Mathematics*, Tianjin, August 1988 - May 1989 (Edited by Ze-pai Jiang, Shi-jian Yan, Ping Cheng and Rong Wu), 32–38. World Scientific, Singapore.

Drygas, H. (1976). Weak and strong consistency of the least squares estimators in regression models. *Z. Wahrsch. verw. Gebiete* **34**, 119–127.

Lai, T. L., Robbins, H. and Wei, C. Z. (1979). Strong consistency of least squares estimates in multiple regression II. *J. Multivariate Anal.* **9**, 343–361.

Li, K. C. (1984). Regression models with infinitely many parameters: Consistency of bounded linear functionals. *Ann. Statist.* **12**, 601–611.

Departmemt of Mathematics, Graduate School, University of Science and Technology of China, Beijing.
Departmemt of Mathematics, University of Science and Technology of China, Hefei.
Departmemt of Mathematics and Statistics, York University, North York, Canada.

(Received December 1990; accepted June 1992)