The Strong Consistency Problem of Least Square Estimate of Linear Regression Coefficients

Jin Mingzhong

Guizhou College for Nationalities,
Guizhou, Huaxi 550025, CHINA

1. Introduction

Consider the linear regression model

\[ Y_i = \alpha^{\top} x_i + \varepsilon_i, \quad i = 1, 2, \ldots, (1.1) \]

where \( x_i \) and \( Y_i \) are the \( i \)-th observation of the \( p \)-dimensional independent variable \( X \) and one-dimensional dependent variable \( Y \). Unless otherwise stated, \( x_i \) will be viewed as known constant \( p \)-vector, \( \alpha \) the \( p \)-vector of regression coefficients to be estimated, and \( \varepsilon_i \) is the random error of the \( i \)-th observation.

Consistency of estimates of \( \hat{\beta} \) is a problem which aroused the interest of many statisticians in recent decades. Roughly speaking, the problem has two aspects. One is concerned with conditions guaranteeing various well-know estimates of, such as the least squares, the minimum \( L_1 \)-norm and Huber’s \( \mathcal{M} \)-estimates. The other aspect is about the existence or non-existence of the consistent estimate of \( \beta \) under various conditions. This problem is studied only recently and not many results are known. This paper considers some problems in this direction. In order to concentrate our efforts on \( \hat{\beta} \), we shall suppose that \( \varepsilon_i \) has a good behavior. Specifically we assume that there exist a positive definite matrix \( \Sigma \) such that

\[ S_n = n^{-1} \sum_{i=1}^{n} \varepsilon_i \varepsilon_i^{\top} \]

or more generally, for large \( n \)

\[ S_n = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i \varepsilon_i^{\top} \]

where \( \varepsilon_i \) are two positive matrices independent of \( n \).

For our purpose it is convenient to rewrite (1.1) as

\[ Y_i = \alpha^{\top} x_i + \varepsilon_i, \quad i = 1, 2, \ldots, (1.4) \]

where \( \hat{\beta} \) satisfies the Gauss-Markov condition

\[ \text{E}(\varepsilon_i | Y_i) = 0, \quad i = 1, 2, \ldots, (1.5) \]

Also, of course \( \text{E}(\varepsilon_i) = 0 \) for \( i = 1, 2, \ldots, (1.4) \). Further, \( \hat{\beta} \) is non-decreasing with \( i: 0 \leq \hat{\beta} \leq 1 \).

Section 2 of this paper deals with the case where \( \hat{\beta} \) is known. Theoretically speaking, this case has already been solved. For, by the transform

\[ Y_i = Y_i | X_i | \hat{\beta} = 1, 2, \ldots, (1.7) \]

one brings model (1.1) into a new model satisfying the Gauss-Markov condition. According to Drygas [1] (see also [2]), the best unbiased linear (BUL) estimate, that with the minimum variance, is consistent if and only if

\[ Q_n^{-1} = (\hat{\beta}^{\top} x_i, x_i^{\top})^{-1} \hat{\beta} \quad \text{as} \quad n \to \infty \]  

(1.8)

If (1.8) not true, then according to [1]-[2], sometimes we can infer that \( \hat{\beta} \) has no linear consistent estimate, or no consistent estimate at all.

Condition (1.8) involves both \( \hat{\beta} \) and \( \hat{\beta} \) simultaneously, and does not fit the purpose of this paper. In section 2 we shall replace it by a very simple condition involving only \( \hat{\beta} \).

Based on this result, we discuss in Section 3 the case where \( \hat{\beta} \) are unknown. There we shall prove that \( \hat{\beta} \) is no linear consistent estimate when \( \hat{\beta} \) are not known, but a nonlinear consistent estimate may exit under certain circumstances. Finally, in Section 4 we shall point out some related open question.
2. The Case Where $\hat{\Sigma}$ is Known

**Theorem 1.** Suppose that in model (1.1), $\hat{\Sigma}$ are known, and condition (1.3), (1.5) are met. Then the necessary and sufficient condition for the consistency of BUL estimate $\hat{\theta}$ is that
\[
\hat{\Sigma}^{-\frac{1}{2}} = \sigma(2.1)
\]

3. The Case Where $\hat{\Sigma}$ are Unknown

Write $H = (h_1, h_2, \ldots, h_k)$ and let $H_1$ be a subset of $H$. Consider the model
\[
(1.4), (1.5), E(e_i) = 0,
\]
\[
(\hat{h}_1, \hat{h}_2)^T H_2 (3.1)
\]

Therefore, the interesting case is that
\[
H_1 = (h_1, h_2, \ldots, h_k, \hat{h}_i, \hat{h}_j)^T (3.2)
\]

**Theorem 2.** If in model (3.1) the set is defined by (3.2), then for any $p$-vector $c \theta$ no linear estimate of $c \theta$ which is consistent in the MS sense exists.

Further, if random errors $e_1, e_2, \ldots$ are mutually independent and contain no subsequence $e_k$ which is asymptotically degenerate, i.e., $e_k b_0 \theta$ (in pr.) for some constant sequence $b_k$, then no weakly consistent linear estimate of $c \theta$ can exist either.

4. Some Open Problems

(1). If in model (3.1) we choose $H_1$ to be some subset of (3.2), what will happen? One specific situation allows for an easy answer, that is, when there exists an element $(h_{11}, h_{20})^T H$ such that for any $(h_1, h_2, \ldots, h_k)^T H$, we have $h_{10}, h_{20}$. Indeed, in this case the estimate
\[
\hat{\theta} = (\sum x_i x_i h_{10})^{-1} (\sum x_i y_i h_{20}^{-1} Y_{x_i})
\]

is an MS-consistent estimate. The problem is whether this is the only case for which a linear consistent estimate for $\theta$ exists.

(2). In practical applications, when the variances $\hat{\Sigma}_{1,2}, \hat{\Sigma}_{2,3}$ are unknown, we may take repeated independent observations $Y_{i1}, Y_{i2}$ for the same value $x_1$ of $x$. This furnishes an estimate of the variance $\sigma$. Can we make use of these estimates of variance to construct a consistent estimate of $\theta$ under what conditions?

We have obtained some results in the simplest case of estimating a location parameter. The general case is still unsolved.

*Remark: This work is supported by the NNSF of Guizhou Province and China.*

**References**
