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

Jin Mingzhong

Guizhou College for Nationalities,

Guiyang,Huaxi.550025, CHINA

1.Introduction

Consider the linear regression model

$Y_i=x_i'\beta+e_i, i=1,2,\cdots$ (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. β is the p-vector of regression coefficients to be estimated, and e_i is the random error of the i-th observation.

Consistency of estimates of β 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 β, such as the least squares, the minimum L_1-norm and Huber's M-estimates. The other aspect is about the existence or non-existence of the consistent estimate of β under various conditions. This problem is studied only recently and not many results are know. This paper considers some problems in this direction. In order to concentrate our efforts on (σ_i^2), we shall suppose that $\{x_i\}$ has a good behavior. Specifically we assume that there exist a positive definite matrix Λ, such that

$S_n/n\to\Lambda$, as $n\to\infty$, $S_n=x_1x_1'+\cdots+x_nx_n'$ (1.2)

or more generally, for large n

$\Lambda_1\le S_n/n\le\Lambda_2$ (1.3)

where Λ_1,Λ_2 are two positive matrices independent of n.

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

$Y_i=x_i'\beta+\sigma_i^2 e_i, i=1,2,\ldots$ (1.4)

where e_i satisfies the Gauss-Markov condition

$E(e_i e_j)=\delta_{ij}$, $i,j=1,2,\cdots$ (1.5)

Also, of course $E(e_i)=0, i=1,2,\ldots$ Further, σ_i^2 is non-decreasing with i:

$0<\sigma_1^2\le\sigma_2^2\le\cdots$ (1.6)

Section 2 of this paper deals with the case where $\sigma_i^2, i=1,2,\ldots$ are know. Theoretically speaking, this case has already been solved. For, by the transform

$\tilde{Y}_i=Y_i/\sigma_i$, $\tilde{x}_i=x_i/\sigma_i$, $i=1,2,\cdots$ (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) $\hat{\beta}_n$ is consistent if and only if

$Q_n^{-1}=(\sum_{i=1}^n x_i x_i')^{-1}\to 0$, as $n\to\infty$ (1.8)

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

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

.Based on this result, we discuss in Section 3 the case where $\sigma_1^2,\sigma_2^2,\cdots$ are unknown. There we shall prove that β has no linear consistent estimate when σ_i^2's 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 σ_i^2's are Known

Theorem 1.suppose that in model(1.1) σ_i^2's are Known, and condition (1.3),(1.5) are met. Then the necessary and sufficient condition for the consistency of BUL estimate $\hat{\beta}_n$ of β is that

$$\sum_{i=1}^{\infty}\sigma_i^{-2}=\infty \quad (2.1)$$

3.The Case Where σ_i^2's are Unknown

Write $H=\{(h_1,h_2,\dots):0<h_1,h_2,\dots<\infty\}$ and let H_1 be a subset of H. Consider the model

(1.4),(1.5),(1.3), $E(e_i)=0$,

$$(\sigma_1^2,\sigma_2^2,\dots)\in H_1 \quad (3.1)$$

Therefore, the interesting case is that

$$H_1=\{(h_1,h_2,\dots):0<h_1,h_2,\dots<\infty,\ \sum_{i=1}^{\infty}h_i^{-2}=\infty\} \quad (3.2)$$

Theorem 2. If in model (3.1)the set is defined by (3.2),then for any p-vector $c\neq 0$ no linear estimate of $c'\beta$ which is consistent in the MS sense exists.

Further, if random errors $e_1,e_2,\dots$ are mutually independent and contain no subsequence $\{e_{nk}\}$ which is asymptotically degenerate, i.e. $e_{nk}-b_k\to 0$ (in pr.) for some constant sequence $\{b_k\}$, then no weakly consistent linear estimate of $c'\beta$ 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 $(h_{10},h_{20},\dots)\in H_1$ such that for any $(h_1,h_2,\dots)\in H_1$ we have $h_i\le h_{i0}, i\ge 1$. Indeed,in this case the estimate

$$\hat{\beta}_n=\left(\sum_{i=1}^{n}x_ix_i'h_{i0}^{-2}\right)^{-1}\sum_{i=1}^{n}h_{i0}^{-1}Y_ix_i$$

is a MS-consistent estimate. The problem is whether this is the only case for which a linear consistent estimate for β exists.

(2). In practical applications, when the variances $\sigma_1^2,\sigma_2^2,\dots$ are unknown, we may take repeated independent observations $Y_{i1},\dots,Y_{in}$ for the same value x_i of x. This furnishes an estimate of the variance σ_i^2. Can we make use of these estimates of variances to construct a consistent estimate of β and 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

[1].Drags., weak and strong consistency of the LS Estimates in regression model, Z. Wahrsch.verw.Gebiete.,34(1976),119-127

[2].Chen xiru, Estimations Theory of Parameter in Linear Model, Science Press,1981