On asymptotics of t-type regression estimation in multiple linear model

CUI Hengjian

Department of Mathematics, Beijing Normal University, Beijing 100875, China (email: hjcui@bnu.edu.cn)

Received October 28, 2003; revised March 10, 2004

Abstract We consider a robust estimator (t-type regression estimator) of multiple linear regression model by maximizing marginal likelihood of a scaled t-type error t-distribution. The marginal likelihood can also be applied to the de-correlated response when the within-subject correlation can be consistently estimated from an initial estimate of the model based on the independent working assumption. This paper shows that such a t-type estimator is consistent.

Keywords: t-type regression estimator, M-estimator, one-step estimate, consistency, asymptotic normality.

DOI: 10.1360/03ys0020

Consider the multiple regression model

\[ Y_i = Z_i^T \beta_0 + \epsilon_i, \quad (i = 1, 2, \cdots, n), \] (0.1)

where \( Z_i = (a, X_i^T)^T \) are given by \((p + 1) \times q\) design points, \( a \in R^q \) is fixed, \( X_i = (x_{i1}, \cdots, x_{iq}) \) with \( x_{ij} \in R^p \), \( \beta_0 \in R^{p+1} \) is a parameter, the \( \epsilon_i \)'s are independent random errors with common density \( f(x) = |\Sigma_0|^{-1/2} f_0(x^T \Sigma_0^{-1} x) \), and \( \Sigma_0 \) is a \( q \times q \) positive definite matrix.

Linear regression model is the most important and popular model in the statistical literature, which attracts many statisticians to estimate the coefficients of the regression. In the past three decades, the robust estimation methods of those coefficients have grown up rapidly, some of which are based on some bounded influencing functions (see Hampel, Ronchetti, Rousseeuw and Stahel[1], and Rousseeuw and Yohai[2]); He[3] proposed the definition of local and global robustness of those estimators. Meanwhile, some GM estimators are obtained due to considering the efficiency of estimator (see Maronna and Yohai[4], Simpson, Ruppert and Carroll[5], Simpson and Yohai[6] and so on). More recently, He, Simpson and Wang[7], and He, Cui and Simpson[8] proposed the weighted t-type regression estimator for linear model; that is, they viewed the components of the error vector \( \{ \epsilon_i \} \) as of independent and identically distributed with common t-distribution whose scale parameter and degrees of freedom are assumed to be \( \tau > 0 \) and \( \nu \) respectively, then the t-type regression estimator \( \hat{\beta}_0 \) of \( \beta_0 \) is obtained by maximizing marginal likelihood of such a scaled t-type error t-distribution. For specific model (0.1), the t-type estimator \( \hat{\beta}_0 \) is
\[ (\hat{\beta}, \hat{\tau}) = \arg \min_{\beta \in \mathbb{R}^p, \tau > 0} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{q} \rho(\tau(y_{ij} - z_{ij}^T \beta)w_{ij}) - \log(\tau), \] (0.2)

where \( \rho(x) = \frac{1}{2}(\nu + 1) \log(1 + x^2/\nu) \), \( w_{ij} = w_i(x_{ij}) \), \( w_n(\cdot) \) is a weighted function, \( \tau \) is a scale parameter, \( y_{ij} \) and \( z_{ij} \) are \( j \)th components of \( Y_i \) and \( Z_i \), respectively. Taking derivative with respect to \( \beta \) and \( \tau \) in (0.2), respectively, we get the following equations:

\[
\begin{align*}
\sum_{i=1}^{n} \sum_{j=1}^{q} \psi(\tau(y_{ij} - z_{ij}^T \beta)w_{ij})w_{ij}z_{ij} &= 0, \\
\sum_{i=1}^{n} \sum_{j=1}^{q} \chi(\tau(y_{ij} - z_{ij}^T \beta)w_{ij}) &= 0,
\end{align*}
\]

(0.3)

where \( \psi(\cdot) = \rho^{(1)}(\cdot) \) and \( \chi(x) = x\psi(x) - 1 \).

The t-distributed errors in multiple linear model can be motivated for a latent hierarchical model as an extension of Gaussian mixed model (see Dempster, Laird and Rubin\[9\], Little\[10\], Large, Little and Taylor\[11\], and He, Cui and Simpson\[8\]). If \( \epsilon_i \) has t-distribution with degrees of freedom \( \nu \), the t-type regression estimator is just the maximum likelihood estimator (MLE) of \( \beta_0 \), which is corresponding to the MLE of error \( \epsilon_i \sim N(0, I_\nu \tau^2) \). When \( \epsilon_i \) is not t-distribution, the t-type regression estimator is corresponding to the ordinary least square estimator. If we take some other suitable \( \rho(\cdot) \) in (0.2), the t-type regression estimator is corresponding to the ordinary M-estimator. The weight function \( w_n(\cdot) \) can be chosen as

\[ w_n(x) = \{1 + (x - M_n)^T C_n^{-1} (x - M_n)/p\}^{-1/2}, \]

or Mallows weight function

\[ w_n(x) = \min \left[ 1, \left\{ b \left( x - M_n \right)^T C_n^{-1} (x - M_n) \right\}^{\alpha/2} \right], \]

where \( a \) and \( b \) are some positive constants, \( (M_n, C_n) \) can be the minimum volume ellipsoid estimator (MVE) of multivariate location and scatter or S estimator.

It is worth noting that there are three advantages of the t-type regression estimator over the ordinary M-estimator.

(i) The ordinary M-estimator provides a robust estimator of \( \beta_0 \) only based on data of the dependent variables \( \{Y_i, \ i = 1, \cdots, n\} \). The t-type estimator, however, gives that based on not only the \( Y \)'s, but also the design points, and its breakdown point depends only on the values of \( \nu \) and \( \omega_n(\cdot) \) (see He, Simpson and Wang\[7\]).

(ii) Usually, M estimation equations can overcome the computing complexity of minimizing the objection function in seeking M-estimator. But those equations have misleading or not unique solution sometimes. The t-type estimator, however, can be obtained rapidly and stably by using EM algorithms directly to optimize the objection function. Meanwhile, such a solution is unique even by using equation (0.3).

(iii) Under the normal model, the t-type regression estimator is still quite efficient, and has a good robust property even if the model is far from the