Asymmetric robust quasi-likelihood

Yoon Dong Lee[1] and Hyemi Choi[2]

ABSTRACT

The robust quasi-likelihood (RQL) proposed by Cantoni & Ronchetti (2001) is a robust version of quasi-likelihood. They adopted Huber function to increase the resistance of the RQL estimator to the outliers. They considered the Huber function only of symmetric type. We extend the class of Huber function to include asymmetric types, and derived a method to find the optimal asymmetric one.

1. Introduction

The quasi-likelihood (QL) is popular, because of their convenience in use and the generality in application. The robust version of the QL was suggested by Cantoni & Ronchetti (2001). The robust quasi-likelihood (RQL) adopts two statistical devices for robustness. One is Huber function to increase the resistance to the outliers, and the other is weighting function to reduce the effect of influential data. The RQL estimator is in Mallow (1975) class.
The QL enable us efficient estimation as much as the (original) likelihood for the distributions of exponential family. Many distributions in the exponential family are highly skewed, differently with Gaussian case. The symmetric Huber function $\psi_c(r)=\min(c, \max(-c,r))$, $c\geq 0$ was originally devised for Gaussian data. The constant c in the Huber function is a balancing parameter between efficiency and robustness. In this article, we consider an extension of the RQL by adopting asymmetric Huber function of the type $\psi_{c_1,c_2}(r)=\min(c_2, \max(c_1,r))$, $c_1<0<c_2$.
To avoid using two parameters c_1 and c_2 simultaneously to keep the balance between efficiency and robustness, we consider a method using only the parameter c_2 and taking the optimal parameter c_1 according to the given parameter c_2, i.e. $c_1=c_1(c_2)$. For the criterion to delete a redundant parameter, we used the method to minimize the asymptotic variance of the RQL estimator. In deriving the asymptotic variance, we mention the corrected form of a formula which is erroneously given in Cantoni & Ronchetti (2001).
The arguments of asymmetric robust quasi-likelihood is possible for general types of distributions in GLM family. However, as an example, we mainly consider the case of gamma distribution, which was the case applied for the estimation of spatial dependence in Lee and Choi (2004). To concentrate to the main topic, we do not consider the matters related to the weighting function of the RQL in Cantoni & Ronchetti (2001).

2. Robust quasi-likelihood

1) Applied Statistics, Konkuk University, Seoul 143-701, poisson@dreamwiz.com
2) Applied Mathematics, KAIST, Daejeon 305-701, hyemichoi@yahoo.com

When the data y_i, $i=1,2,\cdots,n$ follow distributions having mean μ_i, variance function $V(\mu_i)$ and dispersion parameter ϕ_i for each $i=1,2,\cdots,n$. The QL is defined as follows when we assume y_i's are independent;

$$Q(\mu, V, \phi)=\sum_{i=1}^{n}\int_{y_i}^{\mu_i}\frac{y_i-t}{\phi V(t)}\,dt$$

where $\mu=(\mu_1,\cdots,\mu_n)'$ and $\phi=(\phi_1,\cdots,\phi_n)'$.

Recently Cantoni & Ronchetti (2001) extended the QL and suggested RQL by adopting the Huber function of symmetric type, and weighting values $w_i=w(x_i)$ derived from corresponding covariates x_i for each i. When we define

$$Q_i^R(\mu, V, \phi)=\int_{y_i}^{\mu_i}\psi_c\left(\frac{y_i-t}{\sqrt{\phi_i V(t)}}\right)\frac{w_i}{\sqrt{\phi_i V(t)}}\,dt-(1/n)\sum_{i=1}^{n}\int_{y_i}^{\mu_i}E\left[\psi_c\left(\frac{y_i-t}{\sqrt{\phi_i V(t)}}\right)\right]\frac{w_i}{\sqrt{\phi_i V(t)}}\,dt$$

the RQL, Q^R is defined to be the sum of Q_i^R, $i=1,2,\cdots,n$; that is, $Q^R=\sum_{i=1}^{n}Q_i^R$.

When the mode $\mu=X\beta$ is assumed and we denote $r_i=\frac{y_i-\mu_i}{\sqrt{\phi_i V(\mu_i)}}$, $i=1,2,\cdots,n$ and $r=(r_1,\cdots,r_n)'$, the estimation equation $U_R(\beta)\equiv\dot{\mu}'WV^{-1/2}\psi_c(r)-n\alpha=0$ is derived from the RQL. Here $\dot{\mu}$ means the derivatives of μ with respect to β. The constant vector $\alpha\equiv\dot{\mu}'WV^{-1/2}E\psi_c(r)/n$ is directly obtained from Fisher consistency condition $EU_R(\beta)=0$. The matrix W and V are the matrices having w_i's and $\phi_i V(\mu_i)$'s on its diagonal elements, respectively.

3. Asymmetric robust quasi-likelihood

Just by replacing $\psi_c(\cdot)$ with $\psi_{c_1,c_2}(\cdot)$ in RQL, the asymmetric RQL and the corresponding estimating equation are directly obtained. Since there is no meaningful change of notations, we keep to use the same notation for the symmetric and asymmetric cases as long as there is no possibility of confusion. From the basic property of the M-estimator that the solution of the estimating equation, the RQL estimator is asymptotically normal with mean 0 and variance

$$\lim_{n\to\infty} n(E\dot{U}_R)^{-1}E[U_R U_R'](E\dot{U}_R)^{-1}$$

regardless of symmetry of adopted Huber function. Here we have,

$$E\dot{U}_R=-\dot{\mu}'WV^{-1/2}\text{diag}(E[\psi(r)\circ r])V^{-1/2}\dot{\mu},$$

and $$E[U_R U_R']=\dot{\mu}'WV^{-1/2}\Sigma_\psi V^{-1/2}W\dot{\mu}.$$

The term Σ_ψ means the diagonal matrix $\text{diag}(var(\psi(r)))$.

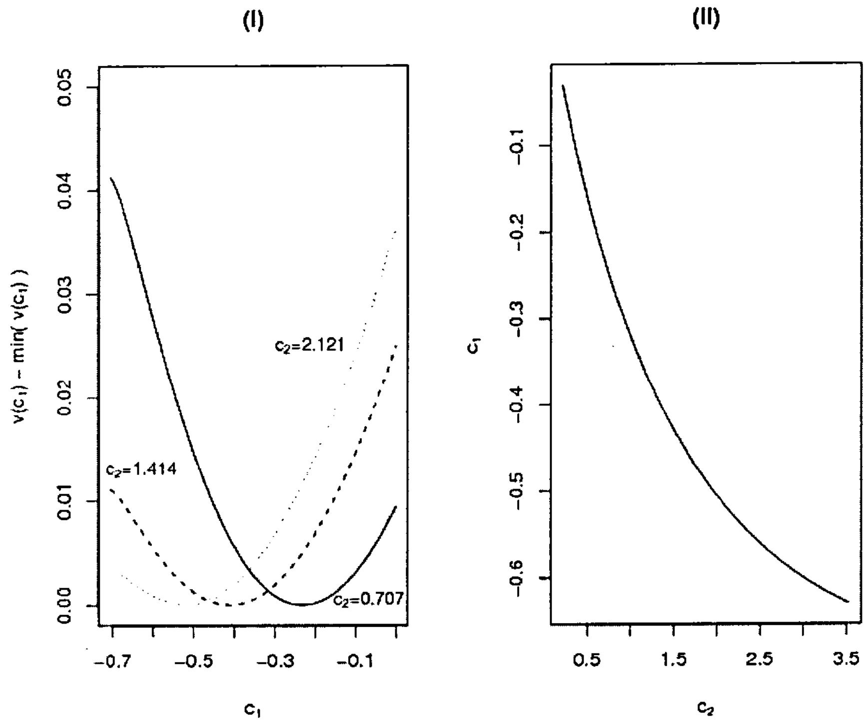

Figure I. The parameters (c_1, c_2) of the asymmetric Huber function ψ. (I): curves of $v(c_1) - \min v(c_1)$ with c_2=0.707, 1.414 and 2.121. (II): the optimal parameter c_1 as a function of the parameter c_2.

Cantoni & Ronchetti (2001) obtained $E[U_R U_{R'}] = \dot{\mu}' W V^{-1/2} \Sigma_{\psi} V^{-1/2} W \dot{\mu} - n\, a\, a'$. In their notation $Q(\psi, F) = (1/n) X^T A X - a\, a'$. This is a miscalculated result coming from the wrong relation of

$$E\{\sum_i^n (g_i - a) \sum_j (g_j - a)'\} = \sum_i^n E(g_i - a)(g_i - a)'$$

for $g_i = \psi_c(r_i) w(x_i) V^{-1/2}(\mu_i) \dot{\mu}_i$. This relation cannot be satisfied for a obtained from $\sum_i^n E(g_i - a) = 0$.

When we consider, as an example, the case $y_i \sim \mu_i \chi^2(1)$, $i = 1, 2, \cdots, n$, the specific terms used to evaluate the asymptotic varinace of the RQL estimator are given as follows;

$$E\psi(r) = \frac{1}{\sqrt{2}} \{(a-1)F_1(a) + (b-1)\overline{F_1}(b) + F_3(a,b) - F_1(a,b)\},$$

$$var(\psi(r)) = \int_a^b (b-t) F_1(t)\, dt - \frac{1}{2}\left\{\int_a^b F_1(t)\, dt\right\}^2,$$

$$E[r\psi(r)] = F_3(a, b),$$

where $F_i(\cdot)$ denotes the distribution function of $\chi^2(1)$, and $\overline{F_i} = 1 - F_i$, $F_i(a, b) = F_i(b) - F_i(a)$, $a = \max(0, 1 + \sqrt{2}c_1)$ and $b = 1 + \sqrt{2}c_2$. Note that any term on the right-hand sides of the above three expectations does not depend on the subscript i of r.

For the balancing parameters c_1 and c_2 of ψ, we suggest to choose $c_1^*(c_2)$ minimizing

$$v(c_1) \equiv var(\psi(r))/(E[r\psi(r)])^2$$

for given c_2. This is motivated from the observation that the asymptotic mean squared error of the RQL estimator depend on $c_1^*(c_2)$ only through the function $v(\cdot)$. That is, such c_1^* can be viewed as a minimizer of the asymptotic mean squared error for fixed c_2. This criterion may determine the value of c_1 uniquely for given c_2 as (I) of Figure 1 indicates, where the minimum of the curve $v(c_1)$ is 1.410, 1.222, 1.126 respectively in each case of c_2=0.707, 1.414 and 2.121.

Moreover, it is worth to note that as c_2 goes to infinity $v(c_1)$ is found to be decreasing to 1. This confirms that the RQL estimator under such selection rule has almost the same asymptotic efficiency with the QL estimator for large value of c_2. Therefore, the $v(c_1)$ can be regarded as a measure of the relative efficiency with respect to the QL estimator of the RQL estimator based on ψ with $(c_1^*(c_2), c_2)$. The right panel (II) of Figure 1 shows the values of c_1 minimizing $v(c_1)$ as a function of $c_2 \in [0, \infty)$. The smaller value of c_1 corresponds to the larger value of c_2.

4. Discussion

In this article, we proposed a asymmetric version of robust quasi-likelihood method. We show that asymmetric version of RQL has advantage that taking not only robustness but also efficiency relative to the symmetric version of RQL. To select a value of appended parameter of asymmetric version, we use the criterion minimizing asymptotic variance (that is, mean squared error, for consistency). We demonstrate the superiority of the asymmetric version to the symmetric version with the example of the regression model of scale parameter of chi-squared distribution.

REFERENCES

Cantoni, E. and Ronchetti, E. (2001), Robust inference for generalized linear models, Journal of the American Statistical Association, 96, 1022-1030.

Lee, Y. D. and Choi, H. (2004), Estimation of spatial dependence by quasi-likelihood method. Korean Journal of Applied Statistics, 17, 519-533.

Mallows, C. L. (1975), On some topics in robustness, Technical Memorandum, Bell Telephone Laboratories, Murray Hill.