行政院國家科學委員會專題研究計畫　成果報告

具有 AR(1)誤差非線性迴歸模型的廣義截斷平均數

計畫類別：個別型計畫
計畫編號：NSC93-2118-M-009-009-
執行期間：93 年 08 月 01 日至 94 年 07 月 31 日
執行單位：國立交通大學統計學研究所

計畫主持人：陳鄰安

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 94 年 10 月 28 日

Generalized Trimmed Means for the Nonlinear Regression With AR(1) Error Model

Abstract

In this work, we apply the idea of trimmed mean of Welsh (1987) to introduce generalized and feasible generalized trimmed means for the nonlinear regression with AR(1) error model. We show that these estimators are asymptotically more efficient than the trimmed means. These results then extend the concept of generalized and feasible generalized least squares estimators for linear regression with AR(1) error model to the robust estimators for nonlinear regression models.

Key words: Generalized Trimmed Mean; nonlinear regression; regression quantile; trimmed least squares; trimmed mean;

中文摘要
在這篇文章裡我們應用 Welsh(1987)的截斷平均數來介紹兩種廣義截斷平均數，用以處理非線性迴歸且有 AR(1)誤差的估計問題。我們證明這兩個廣義估計量皆較截斷平均數具有有效性。這個結果把迴歸中 BLUE 的理論延伸到這種非線性迴歸模型的穩健性理論上。

關鍵詞：廣義截斷平均；非線性迴歸；迴歸分位向量；截斷平均。

1. Introduction

We consider the general nonlinear regression model

$$y_i = g(x_i, \beta) + \epsilon_i, i = 1, ..., n \tag{1.1}$$

where y_i and x_i are, respectively, the response variables and vectors of independent variables, and ϵ_i are error variables. Concerning with estimating regression parameter vector β in the nonlinear regression model with i.i.d. errors, the least squares estimators have been extensively studied. For instance, Hartley and Booker (1965), Jennrich (1969) and Wu (1981) demonstrated the asymptotic normality property. Whereas Ivanov (1976), and Ivanov and Zwanzig (1983) derived an asymptotic expansion for its distribution. Under some regularity conditions and assuming that the errors have comon mean 0 and variance σ^2, then, if we let $\hat{\beta}_{lse}$ represents the least squares estimator (LSE), it has the asymptotic covariance matrix of the form

$$\sigma^2(X(\beta)'X(\beta))^{-1}$$

where $X(b) = (d_1(b), ..., d_n(b))'$ with $d_i(b) = \frac{\partial g(x_i,b)}{\partial b}$. However, outliers or heavy tail error distribution may makes σ^2 large that heavily decreases the efficiency of the LSE.

For increasing the efficiency of the nonlinear least squares estimator, robust estimation aims to develop estimators that have asymptotic covariance matrices of the form

$$\delta(X(\beta)'X(\beta))^{-1}$$

where δ is positive and bounded in error distribution. Among the robust approaches, several authors have proposed and studied some L-estimators. Oberhofer (1982), Richardson and Bhattacharyya (1987) and Wang (1995) studied the ℓ_1-norm estimators, whereas Liese and Vajda (1994) studied the theory of M-estimator. Additionally, from a computational aspect, Procházka (1988) and Koenker and Park (1992) studied the trimmed least squares estimator based on regression quantiles of Koenker and Bassett

(1978). From a theoretical aspect, Jurečková and Procházka (1994) studied it for that model (1.1) includes an intercept term. This trimmed mean is nice to have representation of the form of location trimmed mean. Recently, Huang, Yang and Chen (2003) studied a trimmed mean of Welsh (1987) that has the advantage of easy computation but has the representation in Jurečková and Procházka (1994).

Suppose that the error vector $\epsilon = (\epsilon_1, ..., \epsilon_n)'$ has the covariance matrix structure

$$\sigma^2 \Omega \tag{1.2}$$

where Ω is a positive definite matrix. From the regression theory of the estimation of β, it is known that any estimator having an (asymptotic) covariance matrix of the form

$$\delta(X(\beta)'\Omega^{-1}X(\beta))^{-1} \tag{1.3}$$

is more efficient than the estimator having (asymptotic) covariance matrix of the form

$$\delta(X(\beta)'X(\beta))^{-1}X(\beta)'\Omega^X(\beta)(X(\beta)'X(\beta))^{-1}. \tag{1.4}$$

In the linear regression model with error structure of (1.2), Aitken (1935) call estimators with covariance matrices of the form of (1.3) the generalized estimators. The question we may be interesting is if we have a robust estimator for the nonlinear regression model of (1.1) with error structure of (1.2) that has asymptotic covariance matrix in the form of (1.3).

In this article, we consider the nonlinear regression model of (1.1) with AR(1) errors in the sense that ϵ_i follows

$$\epsilon_i = \rho\epsilon_{i-1} + e_i \tag{1.5}$$

where $e_1, ..., e_n$ are i.i.d. random variables, is one of the most popular models. Suppose that $|\rho| < 1$ and e_i has a distribution function F. We introduce a generalized trimmed mean and feasible generalized trimmed mean. We also derive their asymptotic properties for the regression parameter vector β. We observe that these two estimators are asymptotically more efficient than the Welsh's trimmed means. A data analysis has also been performed.

2. Generalized Trimmed Means

Consider the nonlinear regression model (1.1) where its errors follows the structure of (1.5).

For simplification, denote $D_i(b) = \frac{\partial^2 g(x_i,b)}{\partial b \partial b'}$, the second order partial derivative of the regression function with respect to vector b. The least squares estimate, using the quadratic approximation, is defined as the convergent estimator of the sequence defined by

$$b_j = b_{j-1} + [\sum_{i=1}^{n}(d_i(b_{j-1})d_i'(b_{j-1}) - (y_i - g(x_i, b_{j-1}))D_i(b_{j-1}))]^{-1} \sum_{i=1}^{n} d_i(b_{j-1})(y_i - g(x_i, b_{j-1})) \tag{2.1}$$

where b_0 is a fixed vector.

It is seen that $\text{Cov}(\epsilon) = \sigma^2\Omega$ with

$$\Omega = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{n-1} \\ \rho & 1 & \rho & \ldots & \rho^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \ldots & 1 \end{pmatrix}. \tag{2.2}$$

Define the half matrix of Ω^{-1} as

$$\Omega^{-1/2'} = \begin{pmatrix} (1-\rho^2)^{1/2} & 0 & 0 & \ldots & 0 & 0 \\ -\rho & 1 & 0 & \ldots & 0 & 0 \\ 0 & -\rho & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -\rho & 1 \end{pmatrix}. \tag{2.3}$$

Define matrix $Z(b) = \Omega^{-1/2'}X(b)$. Let $\hat{\beta}_I$ be a predetermined estimator of β. We also define vector $V = \Omega^{-1/2'}(y - G_n(\hat{\beta}_I))$ where $G_n(b) = (g(x_1, b), ..., g(x_n, b))'$ and residuals $e_i = y_i - g(x_i, \hat{\beta}_I), i = 1, ..., n$. Denote the α-th residual quantile as $\eta_n(\alpha)$ and we let $z_i'(b)$ and v_i be, respectively, $i - th$ row of $Z(b)$ and $i - th$ element of V. Combining the quadratization method (2.2) with the construction of Welsh's trimmed mean allows us to define the generalized trimmed mean as the convergent estimator of the sequence defined in the following.

Definition 2.1. The generalized trimmed mean for the nolinear regression model is

$$L_G(\alpha_1, \alpha_2) = \hat{\beta}_I + [\sum_{i=1}^{n}(z_i(\hat{\beta}_I)z_i'(\hat{\beta}_I) - v_i M_i(\hat{\beta}_I))I(\eta_n(\alpha_1) \le v_i \le \eta_n(\alpha_2))]^{-1} \tag{2.4}$$

$$\sum_{i=1}^{n} z_i(\hat{\beta}_I)[v_i I(\eta_n(\alpha_1) \le v_i \le \eta_n(\alpha_2)) + \eta_n(\alpha_1)(I(v_i \le \eta_n(\alpha_1)) - \alpha_1) + \eta_n(\alpha_2)$$
$$(I(v_i \ge \eta_n(\alpha_2)) - (1 - \alpha_2))].$$

After the development of the generalized trimmed mean, the next interesting problem is whether when the parameter ρ is unknown, the trimmed mean of (2.4) with ρ replaced by a consistent estimator $\hat{\rho}$, will have the same asymptotic behavior as displayed by $L_G(\alpha_1, \alpha_2)$. If yes, the theory of generalized least squares estimation is then carried over to the theory of robust estimatimation in this specific nonlinear regression model. Let $\hat{\Omega}$ be the matrix of Ω with ρ replaced by its consistent estimator $\hat{\rho}$. Define matrices $\hat{Z}(b) = \hat{\Omega}^{-1/2'}X(b)$ and $\hat{V} = \hat{\Omega}^{-1/2'}(y - G_n(\hat{\beta}))$. We also let $\hat{z}_i'(b)$ and $\hat{v}_i$ be, respectively, $i - th$ row of $\hat{Z}(b)$ and $i - th$ element of $\hat{V}$.

Definition 2.2. The feasible generalized trimmed mean for the nolinear regression model is

$$L_{FG}(\alpha_1,\alpha_2) = \hat{\beta}_I + [\sum_{i=1}^{n}(\hat{z}_i(\hat{\beta}_I)\hat{z}_i'(\hat{\beta}_I) - \hat{v}_i M_i(\hat{\beta}_I))I(\eta_n(\alpha_1) \le \hat{v}_i \le \eta_n(\alpha_2))]^{-1} \quad (2.5)$$

$$\sum_{i=1}^{n} \hat{z}_i(\hat{\beta}_I)[\hat{v}_i I(\eta_n(\alpha_1) \le \hat{v}_i \le \eta_n(\alpha_2)) + \eta_n(\alpha_1)(I(\hat{v}_i \le \eta_n(\alpha_1)) - \alpha_1) + \eta_n(\alpha_2)$$
$$(I(\hat{v}_i \ge \eta_n(\alpha_2)) - (1-\alpha_2))].$$

3. Large Sample Properties of Generalized Trimmed Mean

We state a set of assumptions (a1-a5) related to the design matrix X and the distribution of the error variable e in the Appendix that are assumed to be true throughout the paper.

Theorem 3.1. The generalized trimmed mean has the following representation

$$n^{1/2}(L_G(\alpha_1,\alpha_2) - (\beta+\gamma)) = n^{-1/2}(\alpha_2-\alpha_1)^{-1}Q_\rho^{-1}\sum_{i=1}^{n} z_i(\beta)(\phi(e_i) - E(\phi(e_i)))$$
$$+ o_p(1)$$

where $\gamma = \lambda(\alpha_2-\alpha_1)^{-1}Q_\rho^{-1}\theta$ with $\lambda = \frac{1-\rho}{\alpha_2-\alpha_1}\int_{F^{-1}(\alpha_1)}^{F^{-1}(\alpha_2)} ef(e)de$, and

$$\phi(e) = \begin{cases} F^{-1}(\alpha_1) & \text{if } e < F^{-1}(\alpha_1) \\ e & \text{if } F^{-1}(\alpha_1) \le e \le F^{-1}(\alpha_2) \\ F^{-1}(\alpha_2) & \text{if } e > F^{-1}(\alpha_2) \end{cases}.$$

For statistical inference, we need an asymptotic distribution of the generalized trimmed mean which is stated in the following.

Corollary 3.2 (a)

$$n^{1/2}(L_G(\alpha_1,\alpha_2) - (\beta+\gamma)) \to N(0, \sigma^2(\alpha_1,\alpha_2)Q_\rho^{-1})$$

where

$$\sigma^2(\alpha_1,\alpha_2) = (\alpha_2-\alpha_1)^{-2}(\alpha_1(F^{-1}(\alpha_1))^2 + (1-\alpha_2)(F^{-1}(\alpha_2))^2 + \int_{F^{-1}(\alpha_1)}^{F^{-1}(\alpha_2)} \epsilon^2 dF$$

$$-(\alpha_1 F^{-1}(\alpha_1) + (1-\alpha_2)F^{-1}(\alpha_2) + \lambda)^2).$$

(b) If F is further assumed to be symmetric and we let $\alpha_1 = \alpha = 1 - \alpha_2$, $0 < \alpha < 0.5$, then

$$n^{1/2}(L_G(\alpha, 1-\alpha) - \beta) \rightarrow N(0, \sigma^2(\alpha, 1-\alpha)Q^{-1})$$

where in this situation

$$\sigma^2(\alpha, 1-\alpha) = (1-2\alpha)^{-2}(2\alpha(F^{-1}(1-\alpha))^2 + \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} \epsilon^2 dF).$$

The asymptotic covariance matrix of $L_G(\alpha_1, \alpha_2)$ is also of the form γQ_ρ^{-1} with $\gamma = \sigma^2(\alpha_1, \alpha_2)$ which is the asymptotic variance of the trimmed mean for the location model. How efficient is the GTM compared with the GLSE? Ruppert and Carroll (1980) computed the values of the term $\sigma^2(\alpha, 1-\alpha)$ for e following several contaminated normal distributions. In comparisons of it with σ^2, the variance of e, the GTM is strongly more efficient than the GLSE when the contaminated variance is large.

Theorem 3.3. The PGTM has the same representation as that expressed for the GTM in Theorem 3.2.

References

Aitken, A. C. (1935). On least squares and linear combinations of observations. *Proceedings of the Royal Society*, Edinburgh. **55**, 42-48.

Chen, L-A, Welsh, A. H. and Chan, W. (2001). Linear winsorized means for the linear regression model. *Statistica Sinica.* **11**, 147-172.

Hartley, H. O. and Booker, A. (1965). Nonlinear least squares estimation. *Annals of Mathematical Statistics*, **36**, 638-650.

Huang, J.-Y., Yang, E. K. and Chen, L.-A. (2004). Welsh's trimmed mean for the nonlinear regression model. *Sankhya.* To appear.

Ivanov, A. V. (1976). An asymptotic expansion for the distribution of the least squares estimator of non-linear regression parameters. *Theory of Probability and its Applications*, **21**, 557-570.

Ivanov, A. V. and Zwanzig, S. (1983). An asymptotic expansion of the distribution of least squares estimators in the nonlinear regression model. *Mathematische Operationsforschung und Statistik, Series Statistics*, **14**, 7-27.

Jennrich, R. I. (1969). Asymptotic properties of non-linear least squares estimators. *Annals of Mathematical Statistics*, **40**, 633-643.

Koenker, R. and Bassett, G. J. (1978). Regression quantile. *Econometrica*, **46**, 33-50.

Koenker, R. and Park, B. J. (1992). An interior point algorithm for nonlinear quantile regression. Faculty Working Paper 92-0127. College of Commerce and Business Administration. University of Illinois at Urbana-Champaign.

Lai, Y.-H., Thompson, P. and Chen, L.-A. (2004). Generalized and pseudo-generalized trimmed means for the linear regression with AR(1) error model. *Statistixs and Probability Letters.* **67**, 203-211.

Liese, F. and Vajda, I. (1994). Consistency of M-estimates in general regression models. *Journal of Multivariate Analysis*, **50**, 93-114.

Oberhofer, W. (1982). The consistency of nonlinear regression minimizing the ℓ_1-norm. *Annals of Statistics*, **10**, 316-319.

Procházka, B. (1988). Regression quantiles and trimmed least squares estimator in the nonlinear regression model. *Computational Statistics and Data Analysis*, **6**, 358-391.

Richardson, G. D. and Bhattacharyya, B. B. (1987). Consistent ℓ_1-estimators in nonlinear regression for a noncompact parameter space. *Shankhya A*, **49**, 377-387.

Wang, J. (1995). Asymptotic normality of ℓ_1-estimators in nonlinear regression. *Journal of Multivariate Analysis*, **54**, 227-238.

Welsh, A. H. (1987). The trimmed mean in the linear model. *Annals of Statistics*, **15**, 20-36.

Wu, C. F. (1981). Asymptotic theory of nonlinear least squares estimation. *Annals of Statistics*, **9**, 501-513.