International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:9, No:6, 2015

Two New Relative Efficiencies of Linear Weighted Regression

Shuimiao Wan, Chao Yuan, Baoguang Tian

***Abstract*—In statistics parameter theory, usually the parameter estimations have two kinds,one is the least-square estimation(LSE),and the other is the best linear unbiased estimation (BLUE).Due to the determining theorem of minimum variance unbiased estimator(MVUE),the parameter estimation of BLUE in linear model is most ideal.But since the calculations are complicated or the covariance is not given,people are hardly to get the solution.Therefore people prefer to use LSE rather than BLUE.And this substitution will take some losses.To quantize the losses,many scholars have presented many kinds of different relative efficiencies in different views.For the linear weighted regression model, this paper discusses the relative efficiencies of LSE of β to BLUE of β. It also defines two new relative efficiencies and gives their lower bounds.

***Keywords*—Linear weighted regression, Relative efficiency, Lower bound, Parameter estimation

I. Introduction

THE model we treat here is linear weighted regression, described by

$$\begin{cases} WY = WX\beta + \varepsilon \\ E(\varepsilon) = 0 \\ Cov(\varepsilon) = \sigma^2\Sigma \end{cases} \tag{1}$$

where Y is the $n \times 1$ observation vector, X is a $n \times p$ column full rank design matrix which we are known, β is a $p \times 1$ unknown parameter vector, ε is the $n \times 1$ observation vector, and Σ is the $n \times n$ positive definite covariance matrix and $W = diag(w_1, w_2, \cdots w_n)$, $w_i \neq 0$, $i = 1, 2, \cdots n$.

As Σ is known, the best linear unbiased estimation of β in the model (1) is

$$\beta^* = (X'W\Sigma^{-1}WX)^{-1}X'W\Sigma^{-1}WY$$

And the covariance matrix is

$$Cov(\beta^*) = \sigma^2(X'W\Sigma^{-1}WX)^{-1}$$

Many common questions just like complicated calculations or Σ unknown are given in practical applications. Thus people tend to use the least-square estimation $\hat{\beta}$ of β instead of BLUE β^* where $\hat{\beta} = (X'W^2X)^{-1}X'W^2Y$ and $Cov(\hat{\beta}) = \sigma^2(X'W^2X)^{-1}X'W\Sigma WX(X'W^2X)^{-1}$. By the theorem of Gauss-Markov[1]: $Cov(\beta^*) < Cov(\hat{\beta})$. This kind of conversion will bring some losses to the estimation. To quantize the losses of estimation, relative efficiency is cited. Following are commonly used

$$e_1(\hat{\beta}) = \frac{|Cov(\beta^*)|}{|Cov(\hat{\beta})|},$$

$$e_2(\hat{\beta}) = \frac{tr(Cov\beta^*)}{tr(Cov\hat{\beta})},$$

$$e_3(\hat{\beta}) = \frac{\|tr(Cov\beta^*)\|}{\|tr(Cov\hat{\beta})\|},$$

$$e_4(\hat{\beta}) = [\frac{tr(Cov\beta^*)^q}{tr(Cov\hat{\beta})^q}]^{\frac{1}{q}},$$

$$e_5(\hat{\beta}) = \frac{\lambda_1(Cov\beta^*)}{\lambda_1(Cov\hat{\beta})},$$

$$e_6(\hat{\beta}) = \min_{1\leq i\leq r} \frac{\lambda_i(Cov\beta^*)}{\lambda_i(Cov\hat{\beta})},$$

$$e_7(\hat{\beta}) = [tr(Cov\hat{\mu} - Cov\mu^*)^q]^{\frac{1}{q}},$$

[2]-[8].

It is obvious that the relative efficiencies above-mentioned have transparent defects in some degree as follows[9]-[13]:

(1)The level that $e_1(\hat{\beta})$ depends on the matrix X is so low. And it keeps invariant when the matrix multiplies a reversible matrix on the right.

(2)The dependence that $e_2(\hat{\beta})$ relies on the matrix X is improved but it doesn't take into account the impact of the covariance between each component.

(3)Though $e_3(\hat{\beta})$ measures the size of deviation about the variance and covariance among these components in LSE and BLUE products, the sensitivity is not better than $e_1(\hat{\beta})$.

(4)When $q = 1$, $e_4(\hat{\beta})$ is just $e_2(\hat{\beta})$.When $q = 2$, $e_4(\hat{\beta})$ is just $e_3(\hat{\beta})$.When q becomes bigger and bigger, its lower bound gets smaller and smaller.That is the error produced by $\hat{\beta}$ replacing β^*. And the response of $e_4(\hat{\beta})$ is more sensitive as increasing of q.But the level it depends on the matrix X is still low. Hence people can choose different q according to various cases in practical application.

(5)$e_5(\hat{\beta})$ is used as long as we can take the first principal component with enough information. Since $e_5(\hat{\beta})$ does not use all characteristic roots of $Cov(\beta^*)$ and $Cov(\hat{\beta})$, it causes the loss of information in some degree.

(6)We can find $e_6(\hat{\beta})$ use all nonzero eigenvalues of $Cov\beta$ and it responses the loss sensitively.

(7)$e_7(\hat{\beta})$ overcomes low dependence for the design matrix X. But it is not used widely.

S. Wan,C. Yuan and B. Tian are with the Department of Mathematics, Qingdao University of Science & Technology, Shandong 266061, China, e-mail: (tianbaoguangqd@163.com).

Thanks to the above analysis,this paper studies the relative efficiency of LSE to BLUE. Not only does it define two new relative efficiencies, but also gives their lower bounds

$$e_1(\hat{\beta}/\beta^*) = tr(Cov\hat{\beta} - Cov\beta^*)$$
$$e_2(\hat{\beta}/\beta^*) = \|Cov\hat{\beta} - Cov\beta^*\|_F^2.$$

II. The Lower Bound of Relative Efficiency of $e_1(\hat{\beta})$

At first we give some lemmas which will play key roles in the proof of the following theorems.

Lemma 1. *Assume that $\Delta = diag(\delta_1, \delta_2, \cdots \delta_p)$, $\delta_1 > \cdots > \delta_p > 0$, $A_{nn} > 0$ and U is the $n \times p$ matrix,then*

$$\max_{U'U=\Delta} tr(U'AU) = \sum_{i=1}^{p} \lambda_i(A)\delta_i$$
$$\min_{U'U=\Delta} tr(U'AU) = \sum_{i=1}^{p} \lambda_{n-i+1}(A)\delta_i$$
$$\max_{U'U=\Delta} tr(U'AU)^{-1} = \sum_{i=1}^{p} \lambda_{n-i+1}(A)\delta_{p-i+1}^{-1}$$
$$\min_{U'U=\Delta} tr(U'AU)^{-1} = \sum_{i=1}^{p} \lambda_{p-i+1}^{-1}(A)\delta_i^{-1}.$$

Lemma 2. *Suppose that A and B are n order symmetric matrices, then*

$$\lambda_n(B)\lambda_i(A^2) \leq \lambda_i(ABA) \leq \lambda_1(B)\lambda_i(A^2)$$

where $\lambda_i(A)$ means the i-th characteristic root of matrix A, $i = 1, 2, \cdots n$.

Lemma 3. *Let A be a $n \times n$ Hermite matrix, U is the $n \times k$ positive definite matrix and meets $U'U = I_k$, then we have*

$$\lambda_{n-k+i}(A) \leq \lambda_i(U'AU) \leq \lambda_i(A), i = 1, 2, \cdots k$$

The proofs of the above lemmas can be found in [14]-[16] respectively.

Theorem 1. *In the model of (1),*

$$e_1(\hat{\beta}) \leq \sum_{i=1}^{p} \mu_i^{-1}(\lambda_i - \lambda_{n-p+i})$$

where

$$\lambda_i = \lambda_i(\Sigma)$$
$$\mu_i = \mu_i(\Lambda).$$

Proof. In the model of (1), we assume that the singular value decomposition (SVD) of WX is: $WX = P\Lambda Q$ where $P'P = I_n$, $Q'Q = QQ' = I_p$, and $\Lambda = diag(\sqrt{\mu_1}, \sqrt{\mu_2}, \cdots \sqrt{\mu_p})$. And $\mu_1, \mu_2, \cdots \mu_p$are the sequence characteristic roots of $X'W^2X$.
Then

$$Cov\beta^*/\sigma^2 = (X'W\Sigma^- WX)^{-1} = (Q'\Lambda P'\Sigma^{-1}P\Lambda Q)^{-1} \quad (2)$$

$$Cov(\hat{\beta})/\sigma^2 = (X'W^2X)^{-1}X'W\Sigma WX(X'W^2X)^{-1}$$
$$= (Q'\Lambda P'P\Lambda Q)^{-1}Q'\Lambda P'\Sigma P\Lambda Q(Q'\Lambda P'P\Lambda Q)^{-1}. \quad (3)$$

By the lemma (1), it is readily to get

$$tr(X'W\Sigma^- WX)^{-1} = tr[(Q'\Lambda P'\Sigma^{-1}P\Lambda Q)^{-1}]$$
$$= tr[(\Lambda P'\Sigma^{-1}P\Lambda)^{-1}]$$
$$\geq \Sigma_{i=1}^p \mu_i^{-1}\lambda_{p-i+1}^{-1}(P'\Sigma^{-1}P) \quad (4)$$

and

$$\lambda_{p-i+1}(P'\Sigma^{-1}P) \leq \lambda_{p-i+1}(\Sigma^{-1}) = \lambda_{n-p+i}^{-1}. \quad (5)$$

Insert (5) into (4), we can obtain

$$tr(X'W\Sigma^- WX)^{-1} \geq \Sigma_{i=1}^p \lambda_{n-p+i}\mu_i^{-1}. \quad (6)$$

Note that

$$tr[(X'W^2X)^{-1}X'W\Sigma WX(X'W^2X)^{-1}]$$
$$= tr(\Lambda^{-1}P'\Sigma P\Lambda^{-1})$$
$$\leq \Sigma_{i=1}^p \mu_i^{-1}\lambda_i(P'\Sigma P)$$
$$\leq \Sigma_{i=1}^p \mu_i^{-1}\lambda_i, \quad (7)$$

hence

$$tr(Cov\hat{\beta} - Cov\beta^*) \leq \Sigma_{i=1}^p \mu_i^{-1}(\lambda_i - \lambda_{n-p+i}). \quad (8)$$

III. The Lower Bound of Relative Efficiency of $e_2(\hat{\beta})$

Theorem 2. *In the model of (1),*

$$e_2(\hat{\beta}) \leq \Sigma_{i=1}^p \mu_{p-i+1}^{-1}\left(\frac{\mu_1}{\lambda_i^2} + \frac{\lambda_i^2}{\mu_p} - 2\frac{\lambda_{n-i+1}^2}{\mu_1}\right)$$

where

$$\lambda_i = \lambda_i(\Sigma)$$
$$\mu_i = \mu_i(\Lambda).$$

Proof. According to the proof of theorem 1 and known condition, it is obvious that[14]

$$Cov\beta^*/\sigma^2 = (X'W\Sigma^- WX)^{-1}$$
$$= (Q'\Lambda P'\Sigma^{-1}P\Lambda Q)^{-1}$$
$$Cov(\hat{\beta})/\sigma^2 = (X'W^2X)^{-1}X'W\Sigma WX(X'W^2X)^{-1}$$
$$= (Q'\Lambda P'P\Lambda Q)^{-1}Q'\Lambda P'\Sigma P\Lambda Q(Q'\Lambda P'P\Lambda Q)^{-1}$$

$$tr(X'W\Sigma^{-}WX)^{-2}$$

$$= tr[(Q'\Lambda P'\Sigma^{-1}P\Lambda Q)^{-1}(Q'\Lambda P'\Sigma^{-1}P\Lambda Q)^{-1}]$$

$$= tr(\Lambda P'\Sigma^{-1}P\Lambda^{2}P'\Sigma^{-1}P\Lambda)^{-1}$$

$$\leq \Sigma_{i=1}^{p}\lambda_{n-i+1}(P'\Sigma^{-1}P\Lambda^{2}P'\Sigma^{-1}P)\mu_{p-i+1}^{-1}, \quad (9)$$

thus

$$\lambda_{n-i+1}(P'\Sigma^{-1}P\Lambda^{2}P'\Sigma^{-1}P)$$

$$\leq \lambda_{n-i+1}(P'\Sigma^{-1}P)^{2}\lambda_{1}(\Lambda^{2})$$

$$= \lambda_{n-i+1}(P'\Sigma^{-2}P)\lambda_{1}(\Lambda^{2})$$

$$\leq \lambda_{n-i+1}(\Sigma^{-2})\lambda_{1}(\Lambda^{2})$$

$$= \lambda_{i}^{-2}\mu_{1}. \quad (10)$$

Insert (10) into (9), we have

$$tr(X'W\Sigma^{-}WX)^{-2} \leq \Sigma_{i=1}^{p}\lambda_{i}^{2}\mu_{1}^{-1}\mu_{p-i+1}^{-1} \quad (11)$$

$$tr[(X'W^{2}X)^{-1}X'W\Sigma WX(X'W^{2}X)^{-1}]^{2}$$

$$= tr(\Lambda^{-1}P'\Sigma P\Lambda^{-2}P'\Sigma P\Lambda^{-1})$$

$$\leq \Sigma_{i=1}^{p}\mu_{p-i+1}^{-1}\lambda_{i}(P'\Sigma P\Lambda^{-2}P'\Sigma P)$$

$$\leq \Sigma_{i=1}^{p}\mu_{p-i+1}^{-1}\lambda_{i}(P'\Sigma P)^{2}\lambda_{1}(\Lambda^{-2})$$

$$\leq \Sigma_{i=1}^{p}\mu_{p-i+1}^{-1}\lambda_{i}(\Sigma^{2})\lambda_{1}(\Lambda^{-2})$$

$$\leq \Sigma_{i=1}^{p}\mu_{p-i+1}^{-1}\lambda_{i}^{2}\mu_{p}^{-1}, \quad (12)$$

$$tr[(X'W^{2}X)^{-1}X'W\Sigma WX(X'W^{2}X)^{-1}$$

$$*(X'W\Sigma^{-1}WX)^{-1}]$$

$$= tr[\Lambda^{-1}P'\Sigma P\Lambda^{-1}(\Lambda P'\Sigma^{-1}P\Lambda)^{-1}]$$

$$= tr(\Lambda^{-1}P'\Sigma P\Lambda^{-2}P'\Sigma P\Lambda^{-1})$$

$$\geq \Sigma_{i=1}^{p}\lambda_{n-i+1}(P'\Sigma P\Lambda^{-2}P'\Sigma P)\lambda_{i}(\Lambda^{-2})$$

$$\geq \Sigma_{i=1}^{p}\lambda_{p}(\Lambda^{-2})\lambda_{n-i+1}(P'\Sigma P)^{2}\lambda_{i}(\Lambda^{-2})$$

$$= \Sigma_{i=1}^{p}\lambda_{n-i+1}^{2}\mu_{1}^{-1}\mu_{p-i+1}^{-1}. \quad (13)$$

Insert (11) (12) (13) into the following formula, it is readily verified that

$$\|Cov\hat{\beta} - Cov\beta^{*}\|_{F}^{2}$$

$$= tr[(Cov\hat{\beta} - Cov\beta^{*})'(Cov\hat{\beta} - Cov\beta^{*})]$$

$$= tr[(Cov\hat{\beta})^{2} + (Cov\beta^{*})^{2} - 2Cov\hat{\beta}Cov\beta^{*}]$$

$$\leq \Sigma_{i=1}^{p}\mu_{p-i+1}^{-1}(\frac{\mu_{1}}{\lambda_{i}^{2}} + \frac{\lambda_{i}^{2}}{\mu_{p}} - 2\frac{\lambda_{n-i+1}^{2}}{\mu_{1}}).$$

IV. Conclusions

In statistics parameter theory[17],due to the determining theorem of minimum variance unbiased estimator(MVUE),the parameter estimation of BLUE in linear model is most ideal.But since the calculations are complicated or the covariance is not given,people are hardly to get the solution.Therefore people prefer to use LSE rather than BLUE.And this substitution will take some losses.To quantize the losses,many scholars have presented many different relative efficiencies in different views.That becomes a popular problem recently.This kind of loss will be heavy sometimes.Hence considering this loss is more and more important.

For relative efficiency of linear model,we can make a further research from the following several aspects:

(1)The connections among several defined relative efficiencies should be further investigated in the future;

(2)People can conduct a further study about the contacts between relative efficiency and general relative coefficients;

(3)Try to find another better relative efficiencies such that it can overcome more defects;

(4)An example should be given to illustrate the theoretical results of relative efficiency by computer calculations.

References

[1] J.X. Pan. *Regression parameters in growth curve model of least squares estimate and Gauss-Markov theorem*.Mathematical statistics and applied probability. 1988, 3, pp.169-185.

[2] Q.T. Duan, Y.H. Zhang. *The relative efficiency of least square estimation in linear weighted regression*. Pure mathematics and applied learning, 1998, 4(3), pp.25-29.

[3] J. Wang, W.X. Wang, C.M. Wang. *The relative efficiency of parameter estimation in linear weighted regression*. Guangxi Normal University, 2000, 15(2), pp.37-40.

[4] A.Y. Liu, S.G. Wang. *A new relative efficiency of least square estimation in linear model*. Applied probability and statistics, 1989, 15(2), pp.97-104.

[5] P. Bloomfield, G.S. Watson. *The relative efficiency of least square estimation*. Biometrika, 1975, 62, pp.121-128.

[6] X.X. Chen. *A new relative efficiency of parameter estimation in linear model*. Jiangxi Normal University, 1999, 23(4), pp.313-315.

[7] B.G. Tian, N. Chen. *The relative efficiency of parameter estimation in linear weighted regression*. World Academy of Science, Engineering and Technology, 2014, 8(11), pp.1348-1351.

[8] Y.L. Huan, G.J. Chen. *The relative efficiency of parameter estimation in linear model*. Applied probability and statistics, 1998, 14(2), pp.159-164.

[9] X.M. Liu, J.Zhang. *A new relative efficiency of parameter estimation in linear model*. Mathematical theory and application, 2000, 20(3), pp.26-29.

[10] J.L. Wang, D.D. Gao. *The relative efficiency of Generalized least square estimation under euclidean norm*. Applied probability and statistics, 1991, 7(4), pp.361-366.

[11] H.S. Liu, R.Q. Li. *The Relative Efficiency of the Parameter Estimate in Linear Weighted Regression*. Journal of Taiyuan University of Technology, 2005, 36, pp.612-614.

[12] L. Xia, B.G. Tian, N.N. Wang. *The Relative Efficiency of Linear Weighted Regression*. Science-technology and engineering, 2009, 20(10), pp.6139-6141.

[13] J. Li, L. Xia, N.N. Wang. *A new relative efficiency in generalized linear model*. Proceedings of the Eighth International Conference on Matrix Theory and its Applications. 2009, pp.84-86.

[14] S.G. Wang, M.X. Wu, Z.Z. Jia. *Matrix inequlity*. Beijing Science Press, 2006.

[15] C.G. Khatri, C.R. Rao. *Some generalization of Kantorovich inequality*. Sankhya, 1982, 44, pp.91-102.

[16] S.G. Wang. *The theory and application of linear model*. Anhui Education Press, 1987.

[17] X.R. Chen. *Statistics introduction*. Beijing Science Press, 1981, pp.104-108.