ACADEMIC
World Academic Union

Journal of Uncertain Systems
Vol.1, No.4, pp.315-320, 2007
Online at: www.jus.org.uk

The Admissibility of the Linear Model under Quadratic Loss Function

Wen- Liang Huang[1, 2] and Yan Gao[1]

[1] *School of Management, University of Shanghai for Science and Technology, Shanghai, 200093, China*
[2] *Department of Mathematics, Eeast China University of Science and Technology, Shanghai, 200237, China*
E-mails: hwlsqt@163.com, gaoyan1962@263.net

Received 8 June 2007; Accepted 27 September 2007

Abstract

This paper is devoted to the linear admissible estimate and admissible estimate in the class of homogeneous estimates. For Gauss-Markov model, a necessary and sufficient condition for admissible estimation is proposed when variance is positive.

Keywords: linear admissible estimate, matrix, quadratic loss function

1. Introduction

For the sake of convenience, throughout the paper, we will use the following notations: A is an $m\times n$ matrix, A' is a transpose of A; if an $n\times n$ matrix A is nonsingular, A^{-1} and $tr(A)$ denotes its inversion matrix and trace respectively; I_n is an $n\times n$ unit matrix; $\mu(A)$ is a linear metric space by column vector of A. Let us consider Gauss-Markov model

$$H: Y = X\beta + \varepsilon,\ \ E(\varepsilon)=0, Var(\varepsilon)=\sigma^2 V, \tag{1.1}$$

where Y is an $n-$dimension observable random vector, X is an $n\times p$ design matrix, β is one unknown $p-$dimension parameter vector; ε is a $n-$dimensional random vector, where $V\geq 0$ is known, $\sigma^2>0$ is an unknown parameters. Already homogeneous linear estimates for regression coefficients almost are obtained in a linear model. Rao [1] proposeed a matrix loss function; the loss function with $V>0$ has been obtained in [2]. The linear minimax estimators under quadratic loss function were developed in [3-5]. Although some interesting results have been obtained, they are not satisfactory enough. In this paper, the proposed loss function is different in denominator; it has the minimax admissibility characterization under the linear model $H: Y = X\beta+\varepsilon,\ \ \varepsilon\sim N(0,\sigma^2 V)$ the quadratic loss function:

$$L(S\beta,\sigma^2,\mathbf{d})=\frac{(d-S\beta)'(d-S\beta)}{\sigma^2+\beta' T\beta},\quad R(S\beta,\sigma^2,\mathbf{d})=\frac{E\{(d-S\beta)'(d-S\beta)\}}{\sigma^2+\beta' T\beta},\quad T=X'V^{-1}X\,(V>0). \tag{1.2}$$

When $V>0$, we want to get a necessary and sufficient condition which is given for a linear admissible estimate and admissible estimate in the class of homogeneous estimates.

2 Some Lemmas

The following lemmas are necessary for the proof of our results. In the model (1.2) the quadratic loss function without the denominator has been discussed. So we do not describe it.

Definition 2.1. *AY is said to be identically superior to BY in the linear model H if random variable (β,σ^2) satisfies*

$$R(S\beta,\sigma^2,AY)\le R(S\beta,\sigma^2,BY)$$

and there exists at least (β,σ^2) such that the above inequality happens to be strict one. AY is said to be admissible characterization estimate of $S\beta$ if there are no estimates which are identically superior to AY in the linear model H.

Lemma 2.1 *Let H be a linear model, L a $k\times n$ matrix, and LY an estimate of $S\beta$. For any random parameter vector $\beta\in R^p$ and $\sigma^2>0$, we have*

$$\frac{E(LY-S\beta)'(LY-S\beta)}{\sigma^2+\beta'T\beta}\ge\frac{E(LX\overline{\beta}-S\beta)'(LX\overline{\beta}-S\beta)}{\sigma^2+\beta'T\beta}. \tag{2.1}$$

Moreover, (2.1) happens to be equality if and only if one of the following conditions holds:

(1) $L=LXT^-X'V^{-1}$, (2.2)

(2) $\mu(VL')\subset\mu(X)$. (2.3)

Proof. By deducing, it follows that

$$E(LX\overline{\beta})=ELY=LX\beta,\quad E(Y'Y')=\sigma^2V+X\beta\beta'X\ ,$$

$$E(LY-LX\overline{\beta})'(LX\overline{\beta}-S\beta)=\frac{tr\left\{E(LX\overline{\beta}-S\beta)(LY-LX\overline{\beta})'\right\}}{\sigma^2+\beta'T\beta}=\frac{tr\left\{ELX\overline{\beta}(LY-LX\overline{\beta})'\right\}}{\sigma^2+\beta'T\beta}=0\,.$$

Therefore, we get a quadratic loss function:

$$\begin{aligned}R(S\beta,\sigma^2,LY)&=\frac{E(LY-S\beta)'(LY-S\beta)}{\sigma^2+\beta'T\beta}=\frac{E(LY-S\beta)'(LY-S\beta)+E(LX\overline{\beta}-S\beta)'(LX\overline{\beta}-S\beta)}{\sigma^2+\beta'T\beta}\\&\ge\frac{E(LX\overline{\beta}-S\beta)'(LX\overline{\beta}-S\beta)}{\sigma^2+\beta'T\beta}.\end{aligned}$$

In the above formula, the equality holds if and only if

$$\begin{aligned}\frac{E(LY-LX\overline{\beta})'(LY-LX\overline{\beta})}{\sigma^2+\beta'T\beta}&=\frac{tr\left\{E(LY-LX\overline{\beta})(LY-LX\overline{\beta})'\right\}}{\sigma^2+\beta'T\beta}\\&=\frac{\sigma^2tr\left\{L(I-XT^-X'V^{-1})V(I-XT^-X'V^{-1}L')\right\}}{\sigma^2+\beta'T\beta}=0\,.\end{aligned}$$

It is evident that, the above formula holds if and only if $L(I-XT^-X'V^{-1})V=0$, which is equivalent to $L=LXT^-X'V^{-1}$.

"(1) $\Leftrightarrow$ (2)". By virtue of $VL^{'} = XT^{-}X^{'}L^{'}$, it is easy to obtain $\mu(VL^{'}) \subset \mu(X)$. On the other hand, if $\mu(VL^{'}) \subset \mu(X)$, there exists a matrix M such that $VL^{'} = XM$. Therefore, we have

$$LXT^{-}X^{'}V^{-1} = (LV)V^{-1}XT^{-}X^{'}V^{-1} = M^{'}X^{'}V^{-1}XT^{-}X^{'}V^{-1} = M^{'}X^{'}V^{-1} = LVV^{-1} = L.$$

Thus, the proof of the lemma is complete.

Lemma 2.2 *If the matrix A is not symmetrical, then there exists an orthogonal matrix P such that $tr(PA) > tr(A)$.*

Proof. Suppose A is a 2×2 matrix, denoted by $A = \begin{pmatrix} a & c \\ d & b \end{pmatrix}$. Without loss of generality, we assume $c > d$. Then, we can take an orthogonal matrix $P_\varepsilon = \begin{pmatrix} 1-\varepsilon & -g \\ g & 1-\varepsilon \end{pmatrix}$, where $\varepsilon > 0$ is small enough, and we let $g = \sqrt{2\varepsilon - \varepsilon^2}$. Thus, it is obtained that

$$tr(P_\varepsilon A) > tr(A) = \ \varepsilon(a+b) + g(c - d).$$

Because $c > d$, $g > 0$ and $\varepsilon > 0$ small enough, it follows that

$$tr(P_a A) > tr(A).$$

On the other hand, when $d > c$, we let $g = -\sqrt{2\varepsilon - \varepsilon^2}$. By mathematical induction, we suppose that the conclusion of the lemma holds when matrix A is $(n-1)\times(n-1)$ matrix. Now we suppose A is a $n \times n$ matrix. Because A is asymmetrical, it must have an $(n-1)\times(n-1)$ main submatrix. We may suppose A has the form $A = \begin{pmatrix} A_{n-1} & b \\ d^{'} & c \end{pmatrix}$. As A_{n-1} is asymmetrical, by the assumption, we have $(n-1)\times(n-1)$ orthogonal matrix P_{n-1}, which satisfies $tr(P_{n-1}A_{n-1}) > tr(A_{n-1})$.

Now we assume $P = \begin{pmatrix} P_{nn-1} & 0 \\ 0 & 1 \end{pmatrix}$. It is easy to see that P is an orthogonal matrix. As a consequence,

$$tr(PA) = \ tr(P_{n-1}A_{n-1}) + c > tr(A_{n-1}) + c = tr(A).$$

Lemma 2.3 *Under the model H, if $LY \sim S\beta$ holds, then to any $1\times k$ matrix K, the following relation holds:*

$$KLY \sim KS\beta.$$

The lemma has been proved in [3, 5].

Lemma 2.4 *Supposes $w \in R^p, t \in R^n$ under model H, the $w'\beta$ linearity be estimated, then $t'Y \sim w'\beta$ essential condition is:*

(1) $t^{'} = t^{'}XT^{-}X^{'}V^{-1}$,

(2) $t^{'}XT^{-}X^{'}t \le t^{'}XT^{-}w$.

Proof. (1) It is easy to get the result by applying Lemma 2.1. Now we prove (2).

As $\beta \in R^p$, thus as the assumed the risk of $w'\beta$, $t'Y$ is

$$R(w'\beta, \sigma^2, t^{'}Y) = \frac{E(t^{'}Y - w^{'}\beta)^2}{\sigma^2 + \beta^{'}T\beta} = \frac{\sigma^2 t^{'}Vt + [\beta^{'}(X^{'}t - w)]^2}{\sigma^2 + \beta^{'}T\beta}.$$

Let $t' = bt'XT^{-}X'V^{-1} - (1-b)w'T^{-}V^{-1}$, and $b \in (0,1)$. Thus t' satisfies (2.2) and as the assumed risk of $w'\beta$, $t'Y$ is sign to the following equation:

$$R(w'\beta,\sigma^2,t'Y)^2 = \frac{E(t'Y - w'\beta)^2}{\sigma^2+\beta'T\beta} = \frac{\sigma^2 t'Vt + \left[\beta'(X't-w)\right]^2}{\sigma^2+\beta'T\beta}$$
$$= \frac{\sigma^2(bt'X+(1-b)w')T^{-}(bX't+(1-b)w) + b^2(\beta(X't-w))^2}{\sigma^2+\beta'T\beta}.$$

If there is $t'Y \sim w'\beta$, it is sufficient to show the relations

$$t'Vt \le [bt'X + (1-b)w]T^{-}[bXt + (1-b)w], \text{ for any } b \in (0,1).$$

This means when t satisfies (1), thus we can obtain the inequality like

$$(1-b)^2 t'XT^{-}X't \le (1-b)^2 w'T^{-}w' + 2b(1-b)^2 t'XT^{-}w'.$$

Let both sides of the inequality above divide $(1-b)^2$, at the same time $b \uparrow 1$. Then we get

$$t'XT^{-}X't \le t'XT^{-}w.$$

As a result, (2) is proved.

3 The admissible characteristic in the linear model *H*

Theorem 3.1 *Let L and S be two $k\times n, k\times p$ constant matrixes. We suppose that $S\beta$ is linearly estimable in the model H. LY is the permissible estimate of $S\beta$ in the linear model H if and only if:*

(1) $L = LXT^{-}X'V^{-1}$,

(2) $LXT^{-}X'L' \le LXT^{-}S'$.

Proof. "*Necessity*" From Lemma 2.1, (1) is proved immediately. From Lemma 2.3 and the condition $LY \sim S\beta$, we know that for any k dimensional constant vector t which satisfies $t'LY \sim t'S\beta$. Thus, from Lemma 2.4, for any $t \in R^k$, we obtain

$$t'LXT^{-}X'L't \le t'LXT^{-}S't.$$

In order to prove (2), it is necessary to prove that $LXT^{-}S'$ is symmetrical.

By contraditon, we assume $LXT^{-}S'$ is asymmetric, and then $(S-LX)T^{-}S'$ is asymmetric. From Lemma 2.2, there exists an orthogonal matrix P which satisfies

$$tr(P(S-LX)T^{-}S') > tr((S-LX)T^{-}S').$$

Let $M = (S - P(S-LX))T^{-}X'V^{-1}$. By deducing, it follows that

$$R(S\beta,\sigma^2,MY) = \frac{E(MY-S\beta)'(MY-S\beta)}{\sigma^2+\beta'T\beta} = \frac{\sigma^2 tr(MVM') + \beta'(MX-S)'(MX-S)\beta}{\sigma^2+\beta'T\beta}$$
$$= \frac{\sigma^2 tr\left\{ST^{-}S' + (S-LX)T^{-}(S-LX)' - 2P(S-LX)T^{-}S'\right\}}{\sigma^2+\beta'T\beta} + \frac{\beta'(MX-S)'(MX-S)\beta}{\sigma^2+\beta'T\beta}$$
$$< \frac{\sigma^2 tr\left\{ST^{-}S' + (S-LX)T^{-}(S-LX) \ - 2(S-LX)T^{-}S'\right\} + \beta'(MX-S)'(MX-S)\beta}{\sigma^2+\beta'T\beta}$$

$$=\frac{\sigma^2 tr(LVL') + \beta'(LX-S)'(LX-S)\beta}{\sigma^2+\beta' T\beta} = \frac{E(LY-S\beta)'(LY-S\beta)}{\sigma^2+\beta' T\beta} = R(S\beta,\sigma^2,LY).$$

And it conflicts to $LY \sim S\beta$, so $LXT^{-}S'$ is symmetric.

"*Sufficiency*". M is an arbitrary $k \times n$ constant matrix, and from Lemma 2.1, it is necessary to prove $MX\overline{\beta}$ is no way superior to LY which will be discussed under two different conditions. Above all, let's give two equalities as follows:

$$\frac{E(LY-S\beta)'(LY-S\beta)}{\sigma^2+\beta' T\beta} = \frac{\sigma^2 trLVL' + \beta'(LX-S)'(LX-S)\beta}{\sigma^2+\beta' T\beta} = \frac{\sigma^2 trLXT^{-}X'L' + \beta'(LX-S)'(LX-S)\beta}{\sigma^2+\beta' T\beta}. \quad (3.1)$$

$$\frac{E(MX\overline{\beta}-S\beta)'(MX\overline{\beta}-S\beta)}{\sigma^2+\beta' T\beta} = \frac{\sigma^2 trMXT^{-}X'M' + \beta'(MX-S)'(MX-S)\beta}{\sigma^2+\beta' T\beta}. \quad (3.2)$$

In the first step we consider the case $LX = S$.

(1) When $MX = S$, it is not hard to see that

$$R(S\beta,\sigma^2,LY) = \frac{E(LY-S\beta)'(LX-S\beta)}{\sigma^2+\beta' T\beta} = \frac{E(MX\overline{\beta}-S\beta)'(MX\overline{\beta}-S\beta)}{\sigma^2+\beta' T\beta} = \frac{\sigma^2 trST^{-}S'}{\sigma^2+\beta' T\beta}.$$

Evidently, $MX\overline{\beta}$ is impossibly superior to LY.

(2) If $MX \neq S$, then from (3.1), (3.2) we can obtain the result below by selecting a proper β:

$$\frac{E(LY-S\beta)'(LY-S\beta)}{\sigma^2+\beta' T\beta} < \frac{E(MX\overline{\beta}-S\beta)'(MX\overline{\beta}-S\beta)}{\sigma^2+\beta' T\beta}.$$

So, $MX\overline{\beta}$ is impossibly superior to LY.

Next we discuss the case $LX \neq S$.

(1) In case of $MX = S$. Let $\beta = 0$. Then from (3.1) and (3.2), we obtain:

$$\frac{E(MX\overline{\beta}-S\beta)'(MX\overline{\beta}-S\beta) - E(LY-S\beta)'(LY-S\beta)}{\sigma^2+\beta' T\beta} = \frac{\sigma^2 tr\{MXT^{-}X'M'\} - \sigma^2 tr\{LXT^{-}X'L'\}}{\sigma^2+\beta' T\beta}$$

$$= \frac{\sigma^2 tr\{(LX-S)T^{-}(LX-S)' + 2LXT^{-}S' - 2LXT^{-}X'L'\}}{\sigma^2+\beta' T\beta} \geq \frac{\sigma^2 tr(LX-S)T^{-}(LX-S)'}{\sigma^2+\beta' T\beta} \geq 0.$$

Thus, $MX\overline{\beta}$ is impossibly superior to LY.

(2) Now we check the reverse conclusion $MX \neq S$. When $MX = LX$, from (3.1) and (3.2), it is easy to verify that LY has the same risk as $MX\overline{\beta}$, so $MX\overline{\beta}$ is impossibly superior to LY. On the other hand, $MX \neq LX$. If $MX\overline{\beta}$ superior to LY, then one has that

$$trMXT^{-}X'M' \leq trLXT^{-}X'L'. \quad (3.3)$$

$$(MX-S)'(MX-S) \leq (LX-S)'(LX-S). \quad (3.4)$$

And at least one of (3.3) and (3.4) is not true. From (3.4), it is sufficient to show that $\mu((MX-S)') \subset \mu((LX-S)')$. That is to say,

$$(MX-S)' = (MX-S)'(LX-S)(LX-S)' + (LX-S)(MS-S)'.$$

Taking

$$F=(LX-S)(LX-S)'+(LX-S)(MS-S)',$$

then there exists the equation

$$(MX-S)'=(LX-S)F.$$

Thus from (3.4)，it follows that

$$FF'\le (LX-S)(LX-S)'+(LX-S)(LX-S)'\le I. \tag{3.5}$$

Therefore, from (3.5), we can find it as follows,

$$\begin{aligned}
& tr\{MXT^{-}X'M'\}-tr\{LXT^{-}X'L'\} \\
&= tr\{(F'LX+(I-F')S)T^{-}(X'L'F+S'(I-F))-LXT^{-}X'L'\} \\
&= tr\{(I-F)(I-F')\}ST^{-}S'+2F(I-F)ST^{-}X'L'+FF'LXT^{-}X'L'-LXT^{-}X'L' \\
&= tr\{(I-F)(I-F')ST^{-}S'+2F(I-F)ST^{-}X'L'-(I-FF')LXT^{-}X'L'\} \\
&= tr\{(I-F)(I-F')ST^{-}S'+2F(I-F)ST^{-}X'L'-(I-FF')LXT^{-}X'L'\} \\
&\ge tr\{(I-F)(I-F')ST^{-}S'-2(I-F)ST^{-}X'L'+(I-FF')LXT^{-}X'L'\} \\
&= tr\{(I-F)(I-F')ST^{-}S'-(I-F-F+FF')ST^{-}X'L'\} \\
&= tr\{(I-F)(ST^{-}S'-ST^{-}X'L')(I-F')\} \\
&= tr\{(I-F)(LX-S)T^{-}(LX-S)+(S-LX)T^{-}XL)(I-F)\} \\
&\ge tr\{(I-F)(LX-S)T^{-}(LX-S)'(I-F)\}.
\end{aligned}$$

This implies that

$$(LX-S)'(I-F)=(LX-S)-(MX-S)=LX-MX\ne 0.$$

Therefore, we have

$$tr\{MXT^{-}X'M'\}>tr\{LXT^{-}X'L'\}.$$

This contradicts to (3.3). Thus the sufficiency is proved.

Based on the theorem 3.1, a necessary and sufficient condition for admissible estimation has been accomplished when variance is positive.

Acknowledgements

This work was supported by National Science Foundation of China (under grant:10671126), and Shanghai leading academic discipline project(under grant:T0502)}.

References

[1] Rao, C. R ., Estimation of parameters in a linear model, *Ann, Statistic*, vol.4, pp. 1037-1073, 1997.

[2] Lu, C. Y. and X. H. Zhu, Admissible linear estimation in linear models, *North East Mathematics*, vol.1, pp.34-40, 1994.

[3] Yu, S. H., The linear mini max estimators of estimable function in a general Gauss-Markov model under quadratic loss function, *Acta Math. Appl. Sinica*, vol.26, no.4, pp. 693-701, 2003.

[4] Zhan, J. L. and J. B. Chen, Admissibility of linear estimator of regression coefficients under quadratic loss, *Acta Math. Appl. Sinica*, vol.8, pp. 237-244, 1992.

[5] Lu, C. Y. and X. H. Zhu, Admissible linear estimation in linear models, *North East Mathematics*, vol.1, pp. 34-40, 1994.