第15卷 增刊　　数 学 研 究 与 评 论　　Vol. 15 Supp
1995年11月　　JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION　　Nov. 1995

病态条件下的一个有偏估计*

王 礼 发
(大连水产学院基础部，116023)

摘　要　在线性回归问题中的病态条件下，提出了一个新的有偏估计及容许性证明. [1]中的估计实为此估计的一个特例. 进一步针对两种特殊情况，给出了特殊的有偏估计及其容许性证明.

关键词　最小二乘估计，病态条件，有偏估计，容许性.
分类号　AMS(1991) 62J05/CCL O212.1

在线性回归模型

$$B_{m\times 1} = A_{m\times n}X_{n\times 1} + e_{m\times 1} \tag{1}$$

中 $B_{m\times 1}$是观测向量；$A_{m\times n}$是设计矩阵，且秩 $A_{m\times n}=n$，$n<m$；随机向量 $e_{m\times 1}$的数学期望和协方差阵分别为

$$E(e_{m\times 1}) = O_{m\times 1},$$
$$E(ee^T) = \sigma^2 I,$$

其中 T 表示转置，I 为单位矩阵，$X_{n\times 1}$是需要估计的 n 维向量，它的最小二乘估计是

$$\hat{X}_{n\times 1} = (A^TA)^{-1}A^TB, \tag{2}$$

$$E(\|\hat{X}_{n\times 1} - X_{n\times 1}\|^2) = E[(\hat{X}-X)^T(\hat{X}-X)] = \sigma^2\sum_{i=1}^{n}\frac{1}{\lambda_i},$$

其中 $\|\hat{X}-X\|$ 表示 $\hat{X}$ 到 X 的欧氏距离，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 是(A^TA)的特征值.

在(A^TA)的特征值 λ_i 中至少有一个接近于零时，即所谓在病态条件下 $E[(\hat{X}-X)^T(\hat{X}-X)]$很大，尽管 $E(\hat{X})=X$，但用 $\hat{X}$ 作为 X 的估计值却很不理想. 从 1962 年以来先后出现了几种有偏估计，1979 年林少宫、李楚霖在[1]中对已有的有偏估计进行了总结，指出了有偏估计的构造、并提出了一个新的有偏估计. 本文再提出一个有偏估计

$$\tilde{X}_H = P(H + P^TA^TAP)^{-1}P^TA^TB, \tag{3}$$

其中 P 是 $n\times n$ 阶正交矩阵，满足

$$P^T(A^TA)P = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = \Lambda, \tag{4}$$

* 1994年2月25日收到.

$$H = \mathrm{diag}(h_1^2, h_1^2, \cdots, h_n^2), \tag{5}$$

可以指出[1]中的有偏估计仅是本文提出的有偏估计(3)当 $h_i^2 = k/\lambda_i (i=1,2,\cdots,n)$ 的特殊情形.

一　有偏估计(3)容许性的证明

定理 1　有偏估计(3)中必定存在 H 使相应的 $\widetilde{X}_H$ 有

$$E[(\widetilde{X}_H - X)^T(\widetilde{X}_H - X)] < E[(\hat{X} - X)^T(\hat{X} - X)], \tag{6}$$

且当 $H = \mathrm{diag}(\frac{\sigma^2}{y_1^2}, \frac{\sigma^2}{y_2^2}, \cdots, \frac{\sigma^2}{y_n^2})$ 时(6)左端取最小值.

证明　作正交变换

$$X = PY, \tag{7}$$

又设

$$AP = A^*. \tag{8}$$

于是(1)化为

$$B = A^*Y + e, \tag{9}$$

(3)化为

$$\widetilde{X}_H = P(H + \Lambda)^{-1}A^{*^T}B. \tag{10}$$

由(7)及(10)得

$$\widetilde{Y}_H = (H + \Lambda)^{-1}A^{*^T}B. \tag{11}$$

因为 $X=PY$ 是正交变换,所以

$$\begin{aligned} E[(\widetilde{X}_H - X)^T(\widetilde{X}_H - X)] &= E[(\widetilde{Y}_H - Y)^T(\widetilde{Y}_H - Y)] \\ &= E[(H + \Lambda)^{-1}A^{*^T}B - Y]^T[(H + \Lambda)^{-1}A^{*^T}B - Y]. \end{aligned} \tag{12}$$

再由(9)式,(12)式相当于

$$\begin{aligned} & E[(H + \Lambda)^{-1}A^{*^T}(A^*Y + e) - Y]^T[(H + \Lambda)^{-1}A^{*^T}(A^*Y + e) - Y] \\ &= Y^T[(H + \Lambda)^{-1}\Lambda - I]^2Y + \mathrm{tr}[(H + \Lambda)^{-1}A^{*^T}E(ee^T)A^*(H + \Lambda)^{-1}] \\ &= \sum_{i=1}^{n}\frac{h_i^4 y_i^2}{(\lambda_i + h_i^2)^2} + \sigma^2\sum_{i=1}^{n}\frac{\lambda_i}{(\lambda_i + h_i^2)^2} \\ &= \sum_{i=1}^{n}\frac{\lambda_i\sigma^2 + y_i^2 h_i^4}{(\lambda_i + h_i^2)^2}. \end{aligned} \tag{13}$$

由 $\dfrac{\partial E[(\widetilde{X}_H - X)^T(\widetilde{X}_H - X)]}{\partial h_i} = \sum_{i=1}^{n}\dfrac{4h_i\lambda_i(y_i^2h_i^2 - \sigma^2)}{(h_i^2 + \lambda_i)^3} = 0$,所以 $h_i = 0, h_i^2 = \dfrac{\sigma^2}{y_i^2}(i=1,2,\cdots,n, h_i^2 + \lambda_i \neq 0$ 时为驻点,而 $h_i^2 = -\lambda_i$ 时为导数不存在点.

又由于 $h_i = 0$ 时,(13)式右端为 $\sum_{i=1}^{n}\frac{\sigma^2}{\lambda_i}$,

$h_i^2 = -\lambda_i$ 时,(13)式右端为无穷大,

$$h_i^2=\frac{\sigma^2}{y_i^2}\text{时，(13)式右端为}\sum_{i=1}^{n}\frac{\sigma^2}{\lambda_i+\frac{\sigma^2}{y_i^2}},$$

所以 $h_i^2=\frac{\sigma^2}{y_i^2}$时，即 $H=\mathrm{diag}(\frac{\sigma^2}{y_1^2},\frac{\sigma^2}{y_2^2},\cdots,\frac{\sigma^2}{y_n^2})$时，$E[(\widetilde{X}_H-X)^T(\widetilde{X}_H-X)]$取到最小值

$$\min E[(\widetilde{X}_H-X)^T(\widetilde{X}_H-X)]=\sum_{i=1}^{n}\frac{\sigma^2}{\lambda_i+\frac{\sigma^2}{y_i^2}}, \tag{15}$$

而 $E[(\hat{X}-X)^T(\hat{X}-X)]=\sum_{i=1}^{n}\frac{\sigma^2}{\lambda_i}$，所以存在 $H=\mathrm{diag}(\frac{\sigma^2}{y_1^2},\frac{\sigma^2}{y_2^2},\cdots,\frac{\sigma^2}{y_n^2})$取到最小值，使 $E[(\widetilde{X}_H-X)^T(\widetilde{X}_H-X)]<E[(\widetilde{X}-X)^T(\widetilde{X}-X)]$.

定理 2　对于特殊情况

$$\|X\|^2<\min\frac{\sigma^2}{\lambda_i} \tag{16}$$

时，对任何 $H=\mathrm{giad}(h_1^2,h_2^2,\cdots,h_k^2)\neq 0$ 有偏估计(3)都是容许的，即

$$E[(\widetilde{X}_H-X)^T(\widetilde{X}_H-X)]<E[(\hat{X}-X)^T(\hat{X}-X)]=\sum_{i=1}^{n}\frac{\sigma^2}{\lambda_i}.$$

证明　由(13)知

$$E[(\widetilde{X}_H-X)^T(\widetilde{X}_H-X)]=\sum_{i=1}^{n}\frac{\lambda_i\sigma^2+y_i^2h_i^4}{(\lambda_i+h_i^2)^2} \tag{13$'$}$$

当 $\|X\|<\min\frac{\sigma^2}{\lambda_i}$时由 $y_i^2\leqslant\|Y\|=\|X\|<\min\frac{\sigma^2}{\lambda_i}$.

只要 $h_i\neq 0$ 对一切 i 均有

$$\frac{\lambda_i\sigma^2+y_i^2h_i^4}{(\lambda_i+h_i^2)^2}<\frac{\lambda_i\sigma^2+\frac{\sigma^2}{\lambda_i}h_i^4}{(\lambda_i+h_i^2)^2}=\frac{\sigma^2(\lambda_i+\frac{h_i^4}{\lambda_i})}{(\lambda_i+h_i^2)^2}<\frac{\sigma^2}{\lambda_i},$$

所以 $\sum_{i=1}^{k}\frac{\lambda_i\sigma^2+y_ih_i^4}{(\lambda_i+h_i^2)^2}<\sum_{i=1}^{n}\frac{\sigma^2}{\lambda_i}$.

若(A^TA)的特征值按大小排列

$$\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n,$$

则进一步有

定理 3　若$\frac{\sigma^2}{\lambda_{j-1}}\leqslant\|X\|^2<\frac{\sigma^2}{\lambda_j}$，则 $H=(\underbrace{h_1^2,h_2^2,\cdots,h_j^2}_{j\text{个}},\underbrace{0,0,\cdots,0}_{(n-j)\text{个}})$.

只要 $h_1,h_2,\cdots,h_j$ 不同时为零有偏估计(3)都是容许的，即

$$E[(\widetilde{X}_H-X)^T(\widetilde{X}_H-X)]<E[(\hat{X}-X)^T(\hat{X}-X)].$$

证明　因为

$$E[(\widetilde{X}_H-X)^T(\widetilde{X}_H-X)]=\sum_{i=1}^{n}\frac{\lambda_i\sigma^2+y_i^2h_i^4}{(\lambda_i+h_i^2)^2}=\sum_{i=1}^{j}\frac{\lambda_i\sigma^2+y_i^2h_i^4}{(\lambda_i+h_i^2)^2}+\sum_{i=j+1}^{n}\frac{\sigma^2}{\lambda_i}$$

$$<\sum_{i=1}^{j}\frac{\lambda_i\sigma^2+\frac{\sigma^2}{\lambda_j}h_i^4}{(\lambda_i+h_i^2)^2}+\sum_{i=j+1}^{n}\frac{\sigma^2}{\lambda_i}<\sum_{i=1}^{j}\frac{\sigma^2}{\lambda_i}+\sum_{i=j+1}^{n}\frac{\sigma^2}{\lambda_i}=\sum_{\lambda=1}^{n}\frac{\sigma^2}{\lambda_i}$$

参 考 文 献

[1] 林少宫、李楚霖，关于线性回归的有偏估计构造，华中工学院学报，1979，第二期.

[2] 杨维权、刘兰亭、林鸿洲，多元统计分析，高等教育出版社，1989.

A Biased Estimation Under Ill-Condition

Wang Lifa

(Dalian Fisheries College, 116023)

Abstract

A biased estimation is proposed with admissibility proof for the linear regression problems under ill-conditions. At particular case, it covers an estimations given in [1].