文章编号:1000-582X(2007)12-0098-04

# 广义岭型估计岭参数的确定方法

严羿鹏,杨 虎
(重庆大学 数理学院统计研究所,重庆 400030)

摘 要:在线性回归模型的理论研究中,参数的点估计是一个重要的研究方向。针对设计矩阵的近似复共线性,考虑回归系数的椭球约束,推广了广义岭型估计在均方误差阵意义下优于广义最小二乘估计的有关结论;针对广义岭型估计是一种自适应非线性估计,提出了采用线性 Minimax 估计和平衡损失函数确定广义岭型估计岭参数的具体方法,并应用 R 软件作了算例分析和比较。

关键词:广义岭型估计;椭球约束;岭参数;Minimax 估计;平衡损失函数
中图分类号:O212.4 文献标志码:A

考虑线性模型

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{e},\boldsymbol{e} \sim (0,\sigma^2\boldsymbol{W})。\qquad(1)$$

其中,$\boldsymbol{y}$ 为 $n\times1$ 阶向量,$\boldsymbol{X}$ 为 $n\times p$ 阶设计矩阵,且 $\mathrm{rank}(X)=p$,$\boldsymbol{\beta}$ 为 $p\times1$ 阶未知向量,$\boldsymbol{e}$ 为 $n\times1$ 阶向量,$\boldsymbol{W}$ 为 $n$ 阶正定阵,记为 $\boldsymbol{W}>0$。记号 $\boldsymbol{e}\sim(0,\sigma^2\boldsymbol{W})$ 表示,$e$ 的期望 $E(\boldsymbol{e})=0$,协方差阵 $Cov(\boldsymbol{e})=\sigma^2\boldsymbol{W}$。

对于模型(1)的参数估计问题,在实际应用中可能遇到两种情况。一种是根据以往的经验,参数 $\beta$ 的每一个分量都存在于某个区间。另一种情况是设计阵 $\boldsymbol{X}$ 具有近似的复共线性,$\boldsymbol{\beta}$ 的广义最小二乘估计(GLS)的长度将在 $p$ 维空间中的某些方向上偏大,此时一般的线性约束条件不再适合。为了改进这种情况,考虑对 $\boldsymbol{\beta}$ 的各分量加以限制。针对上述两种情况,考虑约束:$|\boldsymbol{\beta}_i|\leqslant m_i,i=1,2,\cdots,p$,其中 $0<m_1\leqslant\cdots\leqslant m_p$,$\boldsymbol{\beta}_i$ 为 $\boldsymbol{\beta}$ 的分量。

构造如下的约束:

$$\sum_{i=1}^{p}\left(\frac{\boldsymbol{\beta}_i}{m_i}\right)^2 \leqslant k \qquad k\in(0,p],\qquad(2)$$

对任意 $i=1,2,\cdots,p$,不妨令 $n_i=\sqrt{k}m_i$,则上式变为

$$\sum_{i=1}^{p}\left(\frac{\boldsymbol{\beta}_i}{n_i}\right)^2 \leqslant 1,\text{即}\boldsymbol{\beta}'N\boldsymbol{\beta}\leqslant 1,\qquad(3)$$

其中 $N=\mathrm{diag}\left(\frac{1}{n_1^2},\cdots\frac{1}{n_p^2}\right)$,且 $\frac{1}{n_1^2}\geqslant\cdots\geqslant\frac{1}{n_p^2}$。称式(3)为模型(1)的椭球约束。

由文献[1]知,模型(1)中参数向量 $\beta$ 在椭球约束(3)下的估计为

$$\hat{\beta}(\lambda) = (\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{X}+\lambda N)^{-1}\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{y};(\lambda>0),\qquad(4)$$

称为广义岭型估计。其中 $\sigma^2$、$\lambda$、$\boldsymbol{W}$ 都不是容易确定的参数。在实际应用中,可以考虑用 $\hat{\sigma}^2$ 代替 $\sigma^2$,用单位矩阵 $\boldsymbol{I}$ 代替 $\boldsymbol{W}$。对于参数 $\lambda$,文献[1]中提供了一种估计方法:$\hat{\lambda}=\hat{\boldsymbol{\beta}}'_{\mathrm{GLS}}\boldsymbol{X}'\boldsymbol{W}^{-1}(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}_{\mathrm{GLS}})$,其中 $\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$ 表示参数向量 $\boldsymbol{\beta}$ 的广义最小二乘估计。因为 $\hat{\lambda}$ 与样本有关,则广义岭型估计是一种自适应非线性估计。如果用单位矩阵 $\boldsymbol{I}$ 代替上式中的 $\boldsymbol{W}$,易验证 $\hat{\lambda}=0$。因此,需要进一步考虑对参数 $\lambda$ 的估计进行改进。文献[2]提供了估计 $\lambda$ 的另一种方法,即通过式(2)确定 $k$,然后结合式(4)确定 $\lambda$ 的估计值。将对确定广义岭型估计的岭参数 $\lambda$ 提出两种新的方法。

## 1 定义和引理

定义 1 $\hat{\boldsymbol{\beta}}_1$、$\hat{\boldsymbol{\beta}}_2$ 是 $\boldsymbol{\beta}$ 的 2 个估计量,如果满足 $\Delta(\hat{\boldsymbol{\beta}}_1,\hat{\boldsymbol{\beta}}_2)=M(\hat{\boldsymbol{\beta}}_1,\boldsymbol{\beta})-M(\hat{\boldsymbol{\beta}}_1,\boldsymbol{\beta})\geqslant0$,则称估计量 $\hat{\boldsymbol{\beta}}_2$ 关于 MDE-I[2] 有效于估计量 $\hat{\boldsymbol{\beta}}_1$,其中 $M(\hat{\boldsymbol{\beta}},\boldsymbol{\beta})=E(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'$。

其中 $\Delta(\hat{\boldsymbol{\beta}}_1,\hat{\boldsymbol{\beta}}_2)\geqslant0$ 表示半正定矩阵。估计量 $\hat{\boldsymbol{\beta}}_2$ 关于 MDE-I 有效于估计量 $\hat{\boldsymbol{\beta}}_1$,又称为 $\hat{\boldsymbol{\beta}}_2$ 在均方误差阵意义下优于 $\hat{\boldsymbol{\beta}}_1$。

收稿日期:2007-05-07
作者简介:严羿鹏(1971-),男,重庆大学硕士研究生,主要从事数理统计研究。杨虎(联系人),男,教授,(E-mail)yh@cqu.edu.cn。

定义 2　参数 $\boldsymbol{\beta}$ 的估计量 $\hat{\boldsymbol{\beta}}$ 的二次风险函数[2]定义为如下两种形式:

1) $R_1(\hat{\boldsymbol{\beta}},\boldsymbol{\beta},\boldsymbol{A})=E(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{A}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$　($\boldsymbol{A}$ 是一个 $k$ 阶正定阵);

2) $R_2(\hat{\boldsymbol{\beta}},\boldsymbol{\beta},\boldsymbol{A})=E[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{a}]$　($\boldsymbol{a}$ 是一个 $k$ 维向量)。

定义 3　如果

$$\min_{\{\hat{\boldsymbol{\beta}}\}}\sup_{\boldsymbol{\beta}\in B}R_1(\hat{\boldsymbol{\beta}},\boldsymbol{\beta},\boldsymbol{A}) = \sup_{\boldsymbol{\beta}\in B}R_1(\hat{\boldsymbol{\beta}}^*,\boldsymbol{\beta},\boldsymbol{A})$$

则 $\hat{\boldsymbol{\beta}}^*$ 称为 $\boldsymbol{\beta}$ 的 Minmax 估计[2]。其中 $B(\boldsymbol{\beta})\subset R^k$ 是一个关于 $\boldsymbol{\beta}$ 的约束的凸区域。

引理 1　设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 皆为实对称阵,且 $B>0$,则

$$\sup_{x\neq 0}\frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}} = \lambda_{\max}(\boldsymbol{A}\boldsymbol{B}^{-1}),\inf_{x\neq 0}\frac{\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}} = \lambda_{\min}(\boldsymbol{A}\boldsymbol{B}^{-1})。$$

引理 2　设 $\boldsymbol{x}$ 为 $k$ 维向量,$\boldsymbol{A}$ 为 $t$ 阶对称阵,$\boldsymbol{C}$ 为 $t\times k$阶矩阵,则

$$\frac{\partial}{\partial x}\boldsymbol{x}'\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\boldsymbol{x} = 2\boldsymbol{A}\boldsymbol{C}\boldsymbol{x}\boldsymbol{x}'。$$

引理 3　设 $\boldsymbol{A}$ 为 $n\times m$ 矩阵,$\boldsymbol{B}$ 为 $m\times m$ 对称矩阵,且 $\operatorname{rank}(\boldsymbol{A})=m$,则

$$\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}' \geqslant 0\Leftrightarrow\boldsymbol{B}\geqslant 0。$$

引理 4　设 $\boldsymbol{I}$ 为 $k$ 阶单位阵,$\boldsymbol{x}$ 为 $k$ 维向量,则 $\boldsymbol{I}-\boldsymbol{x}\boldsymbol{x}'\geqslant 0\Leftrightarrow\boldsymbol{x}'\boldsymbol{x}\leqslant 1$。

引理 5　设 $\boldsymbol{A}$ 为 $n\times n$ 阶正定阵,$\boldsymbol{B}$ 为 $n\times n$ 阶半正定阵,则 $\boldsymbol{A}^{-1}-(\boldsymbol{A}+\boldsymbol{B})^{-1}\geqslant 0$。

上述引理可以参考文献[3]。

## 2　广义岭型估计的基本性质

性质 1　$E(\hat{\boldsymbol{\beta}}(\lambda))=(\boldsymbol{I}-\lambda\boldsymbol{B}_\lambda^{-1}\boldsymbol{N})\boldsymbol{\beta}$,$Cov(\hat{\boldsymbol{\beta}}(\lambda))=\sigma^2\boldsymbol{B}_\lambda^{-1}\boldsymbol{S}\boldsymbol{B}_\lambda^{-1}$。

其中 $\boldsymbol{B}_\lambda=\boldsymbol{S}+\lambda\boldsymbol{N}$,$\boldsymbol{S}=\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{X}$。

性质 2　$Cov(\hat{\boldsymbol{\beta}}(\lambda))\leqslant Cov(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}})$。

性质 3　$trCov(\hat{\boldsymbol{\beta}}(\lambda))$是关于 $\lambda$ 的单调减少函数,其中 $tr(\cdot)$表示矩阵的迹。

性质 4　在 $\boldsymbol{y}$ 给定的条件下,残差平方和 $S(\lambda)=\|\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}(\lambda)\|^2$ 是关于 $\lambda$ 的单调增加函数。上述性质可以参考文献[1]和文献[2]

定理 1　在椭球约束 $\boldsymbol{\beta}'\boldsymbol{N}\boldsymbol{\beta}\leqslant 1$ 内,估计量 $\hat{\boldsymbol{\beta}}(\lambda)$关于 MDE-I 有效于最小二乘估计量 $\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$的充分条件为下列条件之一成立:

(1) $0<\lambda\leqslant 2\sigma^2$;(2) $\boldsymbol{\beta}'\boldsymbol{S}\boldsymbol{\beta}\leqslant\sigma^2$。

证明:由性质 1,可以得到估计的偏倚为

$\mathrm{Bias}(\hat{\boldsymbol{\beta}}(\lambda),\boldsymbol{\beta}) = E(\hat{\boldsymbol{\beta}}(\lambda)) - \boldsymbol{\beta} = \lambda\boldsymbol{B}_\lambda^{-1}\boldsymbol{N}\boldsymbol{\beta}$。

又因为 $M(\hat{\boldsymbol{\beta}},\boldsymbol{\beta})=E(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'=\mathrm{Cov}(\hat{\boldsymbol{\beta}})+\mathrm{Bias}(\hat{\boldsymbol{\beta}},\boldsymbol{\beta})\mathrm{Bias}(\hat{\boldsymbol{\beta}},\boldsymbol{\beta})'$。

所以 $M(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}},\boldsymbol{\beta})=Cov(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}})$,

$$M(\hat{\boldsymbol{\beta}}(\lambda),\boldsymbol{\beta}) = Cov(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}) + \sigma^2[\boldsymbol{B}_\lambda^{-1}\boldsymbol{S}\boldsymbol{B}_\lambda^{-1} - \boldsymbol{S}^{-1}] + \lambda^2\boldsymbol{B}_\lambda^{-1}\boldsymbol{N}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'\boldsymbol{N}\boldsymbol{B}_\lambda^{-1},\Delta(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}},\hat{\boldsymbol{\beta}}(\lambda)) = \sigma^2[\boldsymbol{S}^{-1}-\boldsymbol{B}_\lambda^{-1}\boldsymbol{S}\boldsymbol{B}_\lambda^{-1}]-\lambda^2\boldsymbol{B}_\lambda^{-1}\boldsymbol{N}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'\boldsymbol{N}\boldsymbol{B}_\lambda^{-1}。$$

根据定义 1,只要证明 $\Delta(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}},\hat{\boldsymbol{\beta}}(\lambda))\geqslant 0$ 即可。由引理 3 可得:

$$\Delta(\hat{\boldsymbol{\beta}}_{\mathrm{GLS}},\hat{\boldsymbol{\beta}}(\lambda))\geqslant 0\Leftrightarrow\sigma^2[\boldsymbol{S}^{-1}-\boldsymbol{B}_\lambda^{-1}\boldsymbol{S}\boldsymbol{B}_\lambda^{-1}]-\lambda^2\boldsymbol{B}_\lambda^{-1}\boldsymbol{N}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'\boldsymbol{N}\boldsymbol{B}_\lambda^{-1}\geqslant 0$$

$$\Leftrightarrow\sigma^2\boldsymbol{B}_\lambda^{-1}(\boldsymbol{B}_\lambda\boldsymbol{S}^{-1}\boldsymbol{B}_\lambda-\boldsymbol{S}-\sigma^{-2}\lambda^2\boldsymbol{N}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'\boldsymbol{N})\boldsymbol{B}_\lambda^{-1}\geqslant 0$$

$$\Leftrightarrow\boldsymbol{B}_\lambda\boldsymbol{S}^{-1}\boldsymbol{B}_\lambda-\boldsymbol{S}-\sigma^{-2}\lambda^2\boldsymbol{N}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'\boldsymbol{N}\geqslant 0$$

$$\Leftrightarrow\lambda^2\boldsymbol{N}\boldsymbol{S}^{-1}\boldsymbol{N}+2\lambda\boldsymbol{N}-\sigma^{-2}\lambda^2\boldsymbol{N}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'\boldsymbol{N}\geqslant 0$$

$$\Leftrightarrow\lambda^2\boldsymbol{N}(\boldsymbol{S}^{-1}+\frac{2}{\lambda}\boldsymbol{N}^{-1}-\sigma^{-2}\boldsymbol{\beta}\cdot\boldsymbol{\beta}')\boldsymbol{N}\geqslant 0$$

$$\Leftrightarrow\boldsymbol{S}^{-1}+\frac{2}{\lambda}\boldsymbol{N}^{-1}-\sigma^{-2}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'\geqslant 0。$$

令 $C=S^{-1}+\frac{2}{\lambda}N^{-1}$,易知 $C>0$,由引理 3、4 可得:

$$S^{-1}+\frac{2}{\lambda}N^{-1}-\sigma^{-2}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'\geqslant 0$$

$$\Leftrightarrow C^{\frac{1}{2}}(\boldsymbol{I}-\sigma^{-2}C^{-\frac{1}{2}}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'C^{-\frac{1}{2}})C^{\frac{1}{2}}\geqslant 0$$

$$\Leftrightarrow\boldsymbol{I}-\sigma^{-2}C^{-\frac{1}{2}}\boldsymbol{\beta}\cdot\boldsymbol{\beta}'C^{-\frac{1}{2}}\geqslant 0$$

$$\Leftrightarrow\sigma^{-2}\boldsymbol{\beta}'(S^{-1}+\frac{2}{\lambda}N^{-1})^{-1}\boldsymbol{\beta}\leqslant 1。\tag{5}$$

根据引理 5,由上式可得:$S-(S^{-1}+\frac{2}{\lambda}N^{-1})^{-1}\geqslant 0$ 和 $\frac{\lambda}{2}N-(S^{-1}+\frac{2}{\lambda}N^{-1})^{-1}\geqslant 0$,

如果 $1\geqslant\sigma^{-2}\boldsymbol{\beta}'S\boldsymbol{\beta}$,此时有 $\boldsymbol{\beta}'S\boldsymbol{\beta}\leqslant\sigma^2$,使得式(5)成立。

如果 $1\geqslant\frac{\lambda\sigma^{-2}}{2}\boldsymbol{\beta}'N\boldsymbol{\beta}\Leftrightarrow\lambda\leqslant\frac{2\sigma^2}{\boldsymbol{\beta}'N\boldsymbol{\beta}}$,又因为 $\boldsymbol{\beta}'N\boldsymbol{\beta}\leqslant 1$,所以 $0<\lambda\leqslant 2\sigma^2$,也使得式(5)成立,证毕。

此定理推广了文献[1]的定理 4。

## 3　广义岭型估计岭参数的确定方法

当广义岭型估计的参数 $\lambda=\frac{\sigma^2}{k}$时,广义岭型估计是一个线性 Minmax 估计。

考虑模型(1),带约束条件 $\boldsymbol{\beta}'\boldsymbol{T}\boldsymbol{\beta}\leqslant k$(其中 $T$ 为 $p$ 阶正定阵)。考虑 $\boldsymbol{\beta}$ 的线性齐次估计类 $\{\hat{\boldsymbol{\beta}}=\boldsymbol{C}\boldsymbol{y}\}$,根据定义 3:

$R_1(C\boldsymbol{y},\beta,A)=\boldsymbol{\beta}'(CX-\boldsymbol{I})'A(CX-\boldsymbol{I})\boldsymbol{\beta}+\sigma^2 tr(ACWC')$

如果允许 $\boldsymbol{A}$ 为非负定阵,设 $\boldsymbol{A}=\boldsymbol{a}\boldsymbol{a}'$($\boldsymbol{a}$ 为任意向量),则 $R_1(C\boldsymbol{y},\boldsymbol{\beta},\boldsymbol{A})=R_2(C\boldsymbol{y},\boldsymbol{\beta},\boldsymbol{A})$

记 $\tilde{A}=\boldsymbol{T}^{-\frac{1}{2}}(CX-\boldsymbol{I})'\boldsymbol{A}(CX-\boldsymbol{I})\boldsymbol{T}^{-\frac{1}{2}}=\boldsymbol{T}^{-\frac{1}{2}}(CX$

$-I)'\boldsymbol{aa}'(\boldsymbol{CX}-\boldsymbol{I})\boldsymbol{T}^{-\frac{1}{2}}$。

则 $R_2(Cy,\boldsymbol{\beta},\boldsymbol{A})=\boldsymbol{\beta}'\boldsymbol{T}^{\frac{1}{2}}\tilde{A}\boldsymbol{T}^{\frac{1}{2}}\boldsymbol{\beta}+\sigma^2a'CWC'a$。

由引理1,$\sup\limits_{\boldsymbol{\beta}}\dfrac{\boldsymbol{\beta}'\boldsymbol{T}^{\frac{1}{2}}\tilde{A}\boldsymbol{T}^{\frac{1}{2}}\boldsymbol{\beta}}{\boldsymbol{\beta}'T\boldsymbol{\beta}}=\lambda_{\max}(\tilde{A})$,

故 $\sup\limits_{\boldsymbol{\beta}'T\boldsymbol{\beta}\leqslant k}R_2(\boldsymbol{Cy},\boldsymbol{\beta},\boldsymbol{A})=\sigma^2\boldsymbol{a}'\boldsymbol{CWC}'a+k\lambda_{\max}(\tilde{A})$。

又因为 $\lambda_{\max}(\tilde{\boldsymbol{A}})=a'(\boldsymbol{CX}-\boldsymbol{I})\boldsymbol{T}^{-1}(\boldsymbol{CX}-\boldsymbol{I})'a$,

故 $\sup\limits_{\boldsymbol{\beta}'T\boldsymbol{\beta}\leqslant k}R_2(Cy,\boldsymbol{\beta},\boldsymbol{A})=\sigma^2a'\boldsymbol{CWC}'a+ka'(\boldsymbol{CX}-\boldsymbol{I})\cdot$
$\boldsymbol{T}^{-1}(CX-\boldsymbol{I})'a$,根据引理2

$\dfrac{1}{2}\dfrac{\partial}{\partial C}\{\sup\limits_{\boldsymbol{\beta}'T\boldsymbol{\beta}\leqslant k}R_2(\boldsymbol{Cy},\boldsymbol{\beta},\boldsymbol{A})\}=(\sigma^2\boldsymbol{W}+k\boldsymbol{XT}^{-1}\boldsymbol{X}')\boldsymbol{C}'$
$aa'-k\boldsymbol{XT}^{-1}aa'$。

因为 $a$ 为任意向量,则由上式为零,可得 $\boldsymbol{C}'=k(\sigma^2\boldsymbol{W}+k\boldsymbol{XT}^{-1}\boldsymbol{X}')^{-1}\boldsymbol{XT}^{-1}$。

则 $\boldsymbol{C}=k\boldsymbol{T}^{-1}\boldsymbol{X}'(\sigma^2\boldsymbol{W}+k\boldsymbol{XT}^{-1}\boldsymbol{X}')^{-1}$,两边同乘以$(\sigma^2\boldsymbol{T}+k\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{X})$,得:

$(\sigma^2T+k\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{X})\boldsymbol{C}=k(\sigma^2\boldsymbol{T}+k\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{X})\boldsymbol{T}^{-1}\boldsymbol{X}'(\sigma^2\boldsymbol{W}+k\boldsymbol{XT}^{-1}\boldsymbol{X}')^{-1}=$
$k\boldsymbol{X}'\boldsymbol{W}^{-1}(\sigma^2\boldsymbol{W}+k\boldsymbol{XT}^{-1}\boldsymbol{X}')(\sigma^2\boldsymbol{W}+k\boldsymbol{XT}^{-1}\boldsymbol{X}')^{-1}=$
$k\boldsymbol{X}'\boldsymbol{W}^{-1}$。

故 $\hat{\boldsymbol{\beta}}=\boldsymbol{Cy}=(\boldsymbol{XW}^{-1}\boldsymbol{X}'+\dfrac{\sigma^2}{k}\boldsymbol{T})^{-1}\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{y}$ 为 $\boldsymbol{\beta}$ 的线性 Minmax 估计[4-5]。

上述线性 Minmax 估计的推导是基于二次损失函数,广义岭型估计的导出是基于椭球约束下的加权最小二乘问题。但是,线性 Minmax 估计可以看成是广义岭型估计$\hat{\boldsymbol{\beta}}(\lambda)$的岭参数$\lambda=\dfrac{\sigma^2}{k}$时的结果,这为确定广义岭型估计的岭参数提供了两种可能:

1)用 $\hat{\sigma}^2$ 代替 $\sigma^2$,则 $\lambda=\dfrac{\hat{\sigma}^2}{k}$;

2)$\boldsymbol{\beta}'\boldsymbol{T\beta}\leqslant\tilde{k}=\dfrac{\sigma^2}{k}$,即可以认为椭球的边界已经包含 $\sigma^2$。

下面分析基于平衡损失的广义岭型估计岭参数的确定方法。

Zellner 于 1994 年提出了一个新的称之为平衡损失函数的标准[6-8]:

$$c\,\|y-X\hat{\boldsymbol{\beta}}\|^2+(1-c)(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{S}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\text{。}\tag{6}$$

其中 $\hat{\boldsymbol{\beta}}$ 为估计量,$\boldsymbol{S}$ 为已知正定阵,$0\leqslant c\leqslant 1$。上式既考虑了估计的精度,又考虑了模型拟合的优良程度,是一个合理的、全面的标准。

根据广义岭型估计的性质3和性质4知道,在确定岭参数的过程中,既要考虑让 $trCov(\hat{\boldsymbol{\beta}}(\lambda))$ 尽量地小,又要考虑不让残差平方和增加太多,这与平衡损失的思想一致。由于平衡损失函数和它的风险中包含未知参数$\boldsymbol{\beta}$,在实际应用中极不方便,因此我们考虑一种近似方法,即选择使

$$J=c\cdot trCov(\hat{\boldsymbol{\beta}}(\lambda))+(1-c)\,\|y-X\hat{\boldsymbol{\beta}}(\lambda)\|^2\text{。}\tag{7}$$

达到最小的$\lambda$,称之为平衡损失估计法,简称 $J$ 方法。参数 $c$ 的选择,在实际应用中有很大的灵活性,如果重视精度,则采用较大的 $c$;更在意模型拟合度,则采用较小的 $c$。当 $c=0.5$ 时,式(7)包含文献[2]的公式作为特殊情况。

参考文献[2]的方法,先固定 $c$,然后通过参数 $k$ 来确定 $\lambda$。

首先,令式(2)中的 $k=1,2,\cdots,p$,$\hat{\boldsymbol{\beta}}_i(\lambda)$ 记为 $\hat{\boldsymbol{\beta}}(\lambda)$ 的第 $i$ 个分量。故$\sum\limits_{i=1}^{p}\dfrac{\hat{\boldsymbol{\beta}}_i(\lambda)}{m_i}=k$。由式(4),可以确定 $\hat{\boldsymbol{\beta}}(\lambda)$ 的分量的表达式,带入上式,解方程可以得到惟一的$\lambda$。当$k$从1取到$p$时,产生$p$个估计:$\hat{\boldsymbol{\beta}}(\lambda_1)$,$\hat{\boldsymbol{\beta}}(\lambda_2)$,$\cdots$,$\hat{\boldsymbol{\beta}}(\lambda_p)$。

其次,验证$|\hat{\boldsymbol{\beta}}_i(\lambda_j)|\leqslant m_i(j=1,\cdots p)$,保留所有分量均满足约束的估计:$\hat{\boldsymbol{\beta}}(\lambda_{(1)})$,$\hat{\boldsymbol{\beta}}(\lambda_{(2)})$,$\cdots$,$\hat{\boldsymbol{\beta}}(\lambda_{(t)})$

最后将上述 $t$ 个估计逐一带入 $J=c\cdot trCov(\hat{\boldsymbol{\beta}}(\lambda))+(1-c)\,\|\boldsymbol{y}-X\hat{\boldsymbol{\beta}}(\lambda)\|^2$,找出使 $J$ 最小的估计。这就是找到的既满足约束的又符合 $J$ 方法的较好的估计。

另一方面,结合1)和2)的方法,在区间$\left[\dfrac{\sigma^2}{p},p\right]$上搜索使 $J$ 达到最小的参数 $\lambda$,也是一种有效的办法。

## 4 算例

笔者的计算基于 R 软件。数据(表1)和模型如下:

表1 数据

| $y$ | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|
| 0.057 8 | 0.762 0 | 0.615 3 | 0.405 7 |
| 0.352 8 | 0.456 4 | 0.791 9 | 0.935 4 |
| 0.813 1 | 0.018 5 | 0.921 8 | 0.916 9 |
| 0.009 8 | 0.821 4 | 0.738 2 | 0.410 2 |
| 0.138 8 | 0.444 7 | 0.176 2 | 0.893 6 |

模型 $y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3$,其中 $\beta_0,\beta_1,\beta_2,\beta_3$ 是待估参数,且已知 $|\beta_1|\leqslant 0.5$,$|\beta_2|\leqslant 0.5$,$|\beta_3|\leqslant 0.2$。假设 $\boldsymbol{e}\sim N(0,\sigma^2\boldsymbol{I})$,式(2)中 $k$ 的最大取值为3,约束矩阵 $\boldsymbol{M}=\mathrm{diag}(4,4,25)$,估计量$\hat{\boldsymbol{\beta}}=(\hat{\beta}_1,\hat{\beta}_2,\hat{\beta}_3)$。

首先考虑无约束条件下参数的最小二乘估计,计算结果如表2:

表 2　无约束条件下参数的最小二乘估计

| $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | 残差平方和 | $\hat{\boldsymbol{\beta}}'M\hat{\boldsymbol{\beta}}$ |
|---|---|---|---|---|---|
| 0. 589 95 | －0. 959 21 | 0. 356 49 | －0. 093 43 | 0. 000 64 | 4. 406 94 |

从表中可以看到,$\hat{\boldsymbol{\beta}}'M\hat{\boldsymbol{\beta}}$ 的值已经超过 $k$ 的最大值3,且参数 $\beta_1$ 的估计值也超过所给范围。因此,进入下一步,即按照第 4 部分所给出的方法得到如表 3 的结果：

表 3　进一步的计算结果

| | | 模型 1 | 模型 2 | 模型 3 | 模型 4 | 模型 5 | 模型 6 |
|---|---|---|---|---|---|---|---|
| $\hat{\lambda}$ | | 0. 000 21 | 1. 131 22 | 0. 414 71 | 2. 999 72 | 0. 010 71 | 0. 78 21 |
| $\hat{\beta}_0$ | | 0. 572 40 | 0. 28191 | 0. 29228 | 0. 27742 | 0. 394 79 | 0. 328 22 |
| $\hat{\beta}_1$ | | －0. 946 96 | －0. 07925 | －0. 18521 | －0. 031 81 | －0. 776 53 | －0. 499 31 |
| $\hat{\beta}_2$ | | 0. 359 56 | 0. 040 98 | 0. 09529 | 0. 016 48 | 0. 365 84 | 0. 251 30 |
| $\hat{\beta}_3$ | | －0. 080 21 | 0. 007 93 | 0. 018 37 | 0. 003 20 | 0. 043 65 | 0. 046 58 |
| $\hat{\boldsymbol{\beta}}'M\hat{\boldsymbol{\beta}}$ | | 4. 264 92 | 0. 033 41 | 0. 181 96 | 0. 005 39 | 2. 994 94 | 1. 304 10 |
| | =0 | 0. 000 66 | 0. 351 98 | 0. 258 45 | 0. 398 68 | 0. 004 98 | 0. 068 01 |
| | $c=0.25$ | 11. 594 63 | 0. 323 23 | 0. 29737 | 0. 350 46 | 1. 894 48 | 0. 587 63 |
| $J$ | $c=0.5$ | 23. 188 61 | 0. 294 49 | 0. 336 30 | 0. 302 25 | 3. 783 99 | 1. 107 25 |
| | $c=0.75$ | 34. 782 58 | 0. 265 74 | 0. 375 22 | 0. 254 03 | 5. 673 49 | 1. 626 87 |
| | $c=1$ | 46. 376 56 | 0. 237 00 | 0. 414 15 | 0. 205 82 | 7. 563 00 | 2. 146 5 |

表中模型 1 数据是按照 $\lambda=\frac{\hat{\sigma}^2}{k}$ 计算得到的,可以看到参数 $\beta_1$ 的估计值超过所给范围且 $\hat{\boldsymbol{\beta}}'M\hat{\boldsymbol{\beta}}$ 的值也大于 3,但是残差平方和(即 $c=0$)最小。模型 4 数据是按照 $\boldsymbol{\beta}'M\boldsymbol{\beta}\leqslant 3=\frac{\sigma^2}{k}$ 计算得到的,可以看到参数 $\beta_1$ 的估计值满足所给范围且 $\hat{\boldsymbol{\beta}}'M\hat{\boldsymbol{\beta}}$ 的值也小于 3,但是残差平方和最大。在 $c=0.5$ 的条件下,所有模型中,模型 2 的 $J$ 值最小。同理,模型 3、模型 4 数据分别在 $c=0.25$, $c=0.75$ 的条件下, $J$ 值取到最小。从表中可以看到,它们的参数估计值均满足所给范围且 $\hat{\boldsymbol{\beta}}'M\hat{\boldsymbol{\beta}}$ 的值都满足约束。

模型 5 的数据是按照 $\hat{\boldsymbol{\beta}}'M\hat{\boldsymbol{\beta}}=3$ 计算得到,但是其中参数 $\beta_1$ 的估计值超过所给范围。模型 6 的数据是按照 $|\hat{\beta}_1|=0.5$ 取得,$J$ 值相应地增大。

综上所述,在实际应用中,可以根据实际问题的需要,采用适当的方法,选择适当的模型。

参考文献：


[1] 杨婷,杨虎. 椭球约束与广义岭型估计[J]. 应用概率统计,2003,19(8):232-236.

[2] 张金槐,线性模型参数估计及其改进[M]. 长沙:国防科技大学出版社,1999.

[3] 王松桂,吴密霞,贾忠贞. 矩阵不等式(二版)[M]. 北京:科学出版社,2006.

[4] MACIEJ WILCZYNSKÍSKÍ. Minimax estimation in linear regression with ellipsoidal constraints[J]. Journal of Statistical Planning and Inference, 2007 (137): 79-86.

[5] SHENG-HYA YU. The linear minimax estimator of stochastic regression coefficients and parameters under quadratic loss function[J]. Statistics & Probability Letters,2007(77):54-62.

[6] ALAN T K WAN. On generalized ridge regression estimators under collinearity and balanced loss[J]. Applied Mathematics and Computation,2002(129):455-467.

[7] 徐兴忠. 平衡损失下回归系数的线性容许估计[J]. 数学物理学报, 2000,20(4):468-473.

[8] 罗汉,柏超. 线性模型中参数的平衡 LS 估计及其性质[J]. 湖南大学学报:自然科学版,2006,33(2):122-124.


(下转第 110 页)

2006,26(B2):307-313

[7] KIM C W, YIM J W. Finsler manifolds with positive constant flag curvature[J]. Geom,Dedicata (to appear).

[8] CHERN S S, SHEN Z. Riemannian-Finsler Geometry[M]. Singapore:World Sci., 2005.

[9] LI B, SHEN Z. On a class of weakly Landsberg Metrics [J]. Science in China Series A: Muthematics, 2007, 50 (4): 573-589.

# Some Properties of Reversible Finsler Metric

*Lu Cong-yin*[1], *Wang Ming-feng*[1] *and Cheng Xin-yue*[2]

1. Department of Mathematics and Physics, Chongqing University, Chongqing 400030, P. R. China;

2. Department of Mathematics and Physics, Chongqing Institute of Technology, Chongqing 400050, P. R. China.

**Abstract**: Let $(M, F)$ be an n-dimensional Finsler manifold ($n \geqslant 3$), using Finsler geometric basic knowledge and methods, it is obtained that reversible Finsler metric $F$ is of many good curvature propertoes. It is proved that reversible $(\alpha,\beta)$-metric $F=\alpha\varphi(s)$ is of relatively isotropic mean Landsberg curvature if and only if the metric $F$ is either Riemannian or Berwaldian., and it develops the result given by Shen Zhong Ming. Finally, it is obtained that if reversible Finsler metric $F$ is of almost isotropic S-curvature, the metric $F$ must be weak Berwaldian. In this case, if the metric $F$ is of scalar flag curvature$K(x,y)$, then$K(x,y)$ must be a constant.

**Key words**: Finsler metric; $(\alpha,\beta)$-metric; Riemann metric; Berwald metric; S-curvature; flag curvature

(编辑 吕建斌)

(上接第 101 页)

# The Method to Determine the Ridge parameter in Generalized Ridge-type Estimator

*YAN Yi-peng, YANG Hu*

(Institute of Statistics, College of Mathematicsl and Physics, Chongqing University, Chongqing 400030, P. R. China)

**Abstract**: Point estimation of parameters is an important research direction in the theory of linear regression model. The conclusion is generalized on the fact that generalized ridge-type estimator is superior to the generalized least-squares estimator with respect to the mean squared error matrix criterion in terms of the approximate multicollinearity of design matrix and the ellipsoidal constraints on regression coefficient. Some methods to determine ridge parameter in generalized ridge-type estimator are given through linear Minimax estimator and balanced loss function aiming at the fact that the generalized ridge-type estimator is an adaptive non-linear estimator. In addition, an example is also presented to analyze and compare above methods using R software.

**Key words**: generalized ridge-type estimator; ellipsoidal restriction; ridge parameter; Minimax estimation; balanced loss function

(编辑 张小强)