第 31 卷　第 3 期　　吉首大学学报(自然科学版)　　Vol. 31　No. 3
2010 年 5 月　　Journal of Jishou University (Natural Science Edition)　　May 2010

文章编号: 1007-2985(2010) 03-0029-03

不等式约束线性模型的参数估计及影响*

王　涛[1], 赵　丹[2]
(1. 山西财经大学统计学院, 山西 太原　030006; 2. 山西财经大学公共管理学院, 山西 太原　030006)

摘　要: 引用不等式约束问题分析处理的思想和对约束线性模型的协方差扰动影响分析的相关结论, 利用分块矩阵求逆和非线性规划的方法, 得出不等式约束线性模型在满足特殊情况时的最小二乘估计 $\hat{\beta}_n$, 接着给出了扰动模型参数的最小二乘估计 $\hat{\beta}_n(\boldsymbol{g}_n)$.

关键词: 不等式约束; 线性模型; 扰动
中图分类号: O221　　文献标识码: A

1　问题的提出

在回归分析中, 对影响点的判断是回归问题重要的一环, 对于无约束问题, 相关结论已经很多, 但对于含不等式约束问题的回归问题, 这方面的结论不多[1].

对于有约束回归问题的主要困难在于不等式约束, 它使得无法得到 β 的解析表达式, 从而难以作进一步的统计分析. 近几年, 数学规划的方法被应用到统计中[2], 得到了很多的重要结论. 文中主要研究有约束的如下线性模型:

$$\begin{cases} \boldsymbol{y}_i = \boldsymbol{x}_i^{\mathrm{T}}\beta_0 + \boldsymbol{e}_i & i = 1, 2, 3, \cdots, n, \\ \boldsymbol{A}_i^{\mathrm{T}}\beta \leqslant \boldsymbol{b}_j & j = 1, 2, \cdots, m. \end{cases} \tag{1}$$

该模型可以转化为非线性规划问题:

$$\begin{cases} \min \sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{x}_i^{\mathrm{T}}\beta)^2, \\ \text{s.t.}\quad \boldsymbol{A}_j^{\mathrm{T}}\beta \leqslant \boldsymbol{b}_j. \end{cases} \tag{2}$$

其中: $j = 1, 2, \cdots, m$; β 为未知参数; $\{(\boldsymbol{x}_n, \boldsymbol{y}_n), i = 1, 2, \cdots, n\}$ 是一组观测数据; β_0 是 β 的未知真实值; $\boldsymbol{e}_i \sim N(0, \sigma^2)(i = 1, 2, 3, \cdots, n)$ 是相互独立的随机误差. 通过非线性规划和矩阵求逆的方法对问题(2) 进行处理[3].

2　记号和假设

问题(2) 是一个数学规划问题, 分别用 $\boldsymbol{F}_n(\beta)$, $\boldsymbol{S}$, $\hat{\beta}_n$ 和 $\lambda_n = (\lambda_{n1}, \lambda_{n2}, \cdots, \lambda_{nm})$ 表示其中的目标函数、可行解集合、最优解以及与之相联系的拉格朗日乘子, 定义 $\hat{\boldsymbol{v}}_n = \begin{pmatrix} \hat{\beta}_n \\ \lambda_n \end{pmatrix}$, $\boldsymbol{Y} = (\boldsymbol{y}_1 \cdots \boldsymbol{y}_n)^{\mathrm{T}}$, $\boldsymbol{e} = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)^{\mathrm{T}}$, $\boldsymbol{X} = (\boldsymbol{x}_1 \cdots \boldsymbol{x}_n)^{\mathrm{T}}$, $\boldsymbol{A} = (\boldsymbol{A}_1 \cdots \boldsymbol{A}_n)^{\mathrm{T}}$, $\boldsymbol{b} = (\boldsymbol{b}_1 \cdots \boldsymbol{b}_n)^{\mathrm{T}}$, 则问题(1) 和(2) 可写成如下的矩阵形式:

$$\boldsymbol{Y} = \boldsymbol{X}\beta_0 + \boldsymbol{e},$$

*　收稿日期: 2010-03-04
作者简介: 王　涛(1981-), 男, 山西阳城人, 山西财经大学统计学院助教, 统计学硕士, 主要从事市场调查和宏观经济等研究.

$$\begin{cases}\min(\boldsymbol{Y}-\boldsymbol{X}\beta)^{\mathrm{T}}(\boldsymbol{Y}-\boldsymbol{X}\beta),\\ \text{s.t.}\quad \boldsymbol{A}\beta\leqslant\boldsymbol{b}.\end{cases}$$

为了保证问题(2) 的解的正则性以及避免复杂的计算, 作出如下假设:

(A1) 当 n 足够大时, $\boldsymbol{k}_n=\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 是一个正定矩阵, 且 $n^{-1}\boldsymbol{k}_n$ 收敛到某一正定矩阵.

(A2) 行向量 $\boldsymbol{A}_j(j=1,\cdots,m)$ 是线性无关的.

(A3) β_0 是如下问题的唯一解, $\lambda=0$ 是唯一的相应的拉格朗日乘子[4], 其中 $j=1,2,\cdots,m$:

$$\begin{cases}\min\sum_{i=1}^{n}(\boldsymbol{x}_i^{\mathrm{T}}\beta-\boldsymbol{x}_i^{\mathrm{T}}\beta_0)^2,\\ \text{s.t.}\quad \boldsymbol{A}_j^{\mathrm{T}}\beta\leqslant\boldsymbol{b}_j.\end{cases}$$

3 $\hat{\beta}_n$ 和 $\hat{\beta}_n(\boldsymbol{g}_n)$ 的估计

首先要求出 $\hat{\beta}_n$ 和 $\hat{\lambda}_n$ 的表达式, 然后在求出对模型施加扰动后的 $\hat{\beta}_n(\boldsymbol{g}_n)$ 和 $\hat{\lambda}_n(\boldsymbol{g}_n)$.

由非线性规划可知, $\hat{\beta}_n$, $\hat{\lambda}_n$ 必须满足如下方程[5]:

$$\begin{cases}\boldsymbol{k}_n\beta+\sum_{l=1}^{k}\lambda_{jl}\boldsymbol{A}_{jl}=\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{y}_i,\\ \boldsymbol{A}_{jl}\beta=\boldsymbol{b}_{jl}\qquad l=1,2,\cdots,k.\end{cases}$$

令 $\boldsymbol{D}=(\boldsymbol{A}_{j1}\cdots\boldsymbol{A}_{jk})$, $\boldsymbol{M}_n=\begin{pmatrix}\boldsymbol{k}_n & \boldsymbol{D}^{\mathrm{T}}\\ \boldsymbol{D} & \boldsymbol{0}\end{pmatrix}$, $\boldsymbol{b}=(b_{j1}\cdots b_{jk})^{\mathrm{T}}$, 则由矩阵求逆公式可得

$$\boldsymbol{M}_{(n)}^{-1}=\begin{pmatrix}\boldsymbol{S}_n & (\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}}\\ \boldsymbol{D}\boldsymbol{k}_n^{-1}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1} & -(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})\end{pmatrix},$$

其中 $\boldsymbol{S}_n=\boldsymbol{k}_n^{-1}[\boldsymbol{I}_P-\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{D}\boldsymbol{k}_n^{-1}]$, $\boldsymbol{k}_n=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$. 由此解得

$$\hat{\boldsymbol{v}}_n=\begin{pmatrix}\hat{\beta}_n\\ \hat{\lambda}_n\end{pmatrix}=\boldsymbol{M}_{(n)}^{-1}\begin{pmatrix}\boldsymbol{X}\boldsymbol{Y}\\ \boldsymbol{b}\end{pmatrix}.$$

再由矩阵逆的性质, 设 $\boldsymbol{M}_{(n)}\boldsymbol{M}_{(n)}^{-1}=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$, 则有

$$a_{11}=\boldsymbol{E}-\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{D}\boldsymbol{k}_n^{-1}+\boldsymbol{D}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{k}_n^{-1}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}=\boldsymbol{E},\tag{3}$$

$$a_{12}=\boldsymbol{k}_n(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{k}_n^{-1}\boldsymbol{D}-\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}=0,\tag{4}$$

$$a_{21}=\boldsymbol{D}\boldsymbol{k}_n^{-1}-\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{D}\boldsymbol{k}_n^{-1}=0,\tag{5}$$

$$a_{22}=\boldsymbol{D}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{k}_n^{-1}\boldsymbol{D}=\boldsymbol{E}.\tag{6}$$

由(3) 至(6) 式可知 $(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{k}_n^{-1}\boldsymbol{D}=\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}$, 则易得到估计 $\hat{\beta}_n$ 以及 $\hat{\lambda}_n$ 的表达式如下:

$$\hat{\beta}_n=\boldsymbol{S}_n\boldsymbol{X}\boldsymbol{Y}+\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{b}=\hat{\beta}+\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}(\boldsymbol{b}-\boldsymbol{D}\hat{\beta}).$$

设

$$\boldsymbol{U}_n=\boldsymbol{Y}-\boldsymbol{X}\hat{\beta}_n=\boldsymbol{Y}-\boldsymbol{X}[\hat{\beta}+\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}(\boldsymbol{b}-\boldsymbol{D}\hat{\beta})]=\boldsymbol{Y}-$$
$$\boldsymbol{X}\{\hat{\beta}-(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{M}\hat{\beta}_n+(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{D}^{\mathrm{T}}[\boldsymbol{D}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{D}^{\mathrm{T}}]^{-1}\boldsymbol{b}\}.$$

其中: $\hat{\beta}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}\boldsymbol{Y}$; $\boldsymbol{M}=\boldsymbol{D}^{\mathrm{T}}[\boldsymbol{D}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{D}^{\mathrm{T}}]^{-1}\boldsymbol{D}$; $\hat{\lambda}_n=(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{D}\boldsymbol{k}_n\boldsymbol{X}\boldsymbol{Y}-(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{b}$. 此时得到 $\hat{\boldsymbol{v}}_n$ 的表达式如下:

$$\hat{\boldsymbol{v}}_n=\begin{pmatrix}\hat{\beta}_n\\ \hat{\lambda}_n\end{pmatrix}=\begin{pmatrix}\hat{\beta}+\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}(\boldsymbol{b}-\boldsymbol{D}\hat{\beta})\\ (\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{D}\boldsymbol{k}_n\boldsymbol{X}\boldsymbol{Y}-(\boldsymbol{D}\boldsymbol{k}_n^{-1}\boldsymbol{D}^{\mathrm{T}})^{-1}\boldsymbol{b}\end{pmatrix}.$$

接着对线性模型进行扰动来求出 $\hat{\beta}_n(\boldsymbol{g}_n)$ 和 $\hat{\lambda}_n(\boldsymbol{g}_n)$ 的表达式.

下面是对线性模型进行随机模型的扰动后的模型:

$$\begin{cases}\boldsymbol{Y}=\boldsymbol{X}\beta+\boldsymbol{e},\boldsymbol{y}_i=\boldsymbol{x}_i^{\mathrm{T}}+\boldsymbol{e}_i\qquad i=1,2,\cdots,n,\\ \boldsymbol{A}_j^{\mathrm{T}}\beta\leqslant\boldsymbol{b}_j\qquad j=1,2,\cdots,m,\\ \boldsymbol{E}(\boldsymbol{e})=0,\boldsymbol{D}(\boldsymbol{e})=\sigma^2\boldsymbol{G}^{-1}.\end{cases}$$

则线性模型的非线性规划的 K-T 方程进一步变为

$$\begin{cases}(X^{T}GX)^{-1}\beta+\sum_{i=1}^{k}\lambda_{jl}A_{jl}=X^{T}GX,\\ A_{jl}\beta=b_{jl}\qquad l=1,2,\cdots,k.\end{cases}$$

此时得 $M_{(n)}^{-1}$ 为

$$M_{(n)}^{-1}(G)=\begin{pmatrix}S_n(G) & (Dk_n^{-1}(G)D^{T})^{-1}k_n^{-1}(G)D^{T}\\ Dk_n^{-1}(G)(Dk_n^{-1}(G)D^{T})^{-1} & -(Dk_n^{-1}(G)D^{T})\end{pmatrix},$$

其中 $S_n(G)=k_n^{-1}(G)[I_P-D^{T}(Dk_n^{-1}(G)D^{T})^{-1}Dk_n^{-1}(G)]$, $k_n^{-1}(G)=(X^{T}GX)^{-1}$. 从而得 $\hat{\beta}_n(G)=\hat{\beta}(G)+k_n^{-1}(G)D^{T}(Dk_n^{-1}(G)D^{T})^{-1}(b-D\hat{\beta}(G))$, 其中 $\hat{\beta}(G)=(X^{T}GX)^{-1}XGY$ 表示线性模型无约束时有扰动的参数估计, 而 $\hat{\beta}_n(G)$ 表示有约束时对线性模型施加扰动的参数估计.

设 $G=G-I$, $M=D^{T}[D(X^{T}X)^{-1}D^{T}]^{-1}D$, $H=X(X^{T}X)^{-1}X^{T}$. 若 $I+HG$ 和 $X(X^{T}X)^{-1}M(X^{T}X)^{-1}XG(I+HG)^{-1}-I$ 都可逆, 则 $\hat{\beta}_n(G)$ 和 $\hat{\beta}_n$ 有如下的关系式:

$$\hat{\beta}_n(G)=\hat{\beta}_n+[(X^{T}X)^{-1}M(X^{T}X)^{-1}-(X^{T}X)^{-1}]X^{T}G(I+HG)^{-1}\{X(X^{T}X)^{-1}\cdot D^{T}[D(X^{T}X)^{-1}D^{T}]^{-1}D(X^{T}X)^{-1}XG(I+HG)^{-1}-I\}^{-1}U_n.$$

证明略.

接下来求施加扰动后的 $\lambda_n(G)$ 的表达式.

因为 $\lambda_n=[D(X^{T}X)^{-1}D^{T}]^{-1}D^{T}(X^{T}X)^{-1}XY-[D(X^{T}X)^{-1}D^{T}]^{-1}b$, 所以对模型进行扰动后, λ_n 变为

$$\lambda_n(G)=\lambda_n-[D(X^{T}X)^{-1}D^{T}]^{-1}D(X^{T}X)^{-1}X^{T}G(I+HG)^{-1}\{X(X^{T}X)^{-1}D^{T}\cdot[D(X^{T}X)^{-1}D^{T}]^{-1}D(X^{T}X)^{-1}XG(I+HG)^{-1}-I\}^{-1}U_n.$$

证明略.

综合以上, 求出了 $\hat{\beta}_n$ 和 λ_n 的表达式, 还有模型施加扰动后的 $\hat{\beta}_n(g_n)$ 和 $\lambda_n(g_n)$ 的表达式.

参考文献:

[1] 陈希孺, 吴启光, 成　平, 等. 线性模型参数的估计理论 [M]. 北京: 科学出版社, 1985.

[2] 方开泰, 王东谦, 吴国富. 一类带约束的回归——配方回归 [J]. 计算数学, 1982, 2(1): 57-69.

[3] 王志忠. 带约束的参数平差函数模型的影响分析 [J]. 中南工业大学学报, 1997, 28(5): 423-426.

[4] WANG J. Asymptotics of Least Square Estimation in Constrained Regression Problem [J]. Ann. Statist., 1996, 24: 1316-1326.

[5] WANG Zhi-zhong. The Restricted Influential Analysis of Observations [J]. 湖南数学年刊, 1997, 17(2): 82-86.

Parameter Estimation of Linear Models Under Inequality Constraint and the Influential Points

WANG Tao[1], ZHAO Dan[2]

(1. College of Statistics, Shanxi University of Finance and Economics, Taiyuan 030006, China; 2. College of Public Administration, Shanxi University of Finance and Economics, Taiyuan 030006, China)

Abstract: By using the correlated conclusion of solving inequality constrained problem and using the method of anti-matrix and nonlinear programming, the least square estimation of $\hat{\beta}_n$ is derived, and then the least square estimation of $\hat{\beta}_n(g_n)$ is also derived by variance disturbance.

Key words: inequality constraint; linear model; disturbance

(责任编辑　向阳洁)