应用概率统计 第二十四卷
第六期 2008年12月

Chinese Journal of Applied Probability
and Statistics Vol.24 No.6 Dec. 2008

最小一乘估计快速算法 *

吕书龙 刘文丽

(福州大学数学与计算机科学学院, 福州, 350108)

摘 要

最小二乘估计容易受奇异点的影响, 最小一乘估计是稳健估计, 可以很好地克服这个缺陷, 但计算困难. 基于非退化模型假设下的稳定极点理论, 本文找到了快速准确求解最小一乘估计的迭代算法, 并给出算法的计算过程及与线性规划求解的比较, 较好地解决了最小一乘估计计算难的问题, 使其成为有效的参数估计方法.

关键词: 最小一乘估计, 最小二乘估计, 迭代算法.

学科分类号: O212.

§1. 引 言

考虑线性回归模型:

$$y_i = \alpha + b' x_i + \varepsilon_i, \qquad i = 1, \cdots, n, \tag{1.1}$$

其中x_i为已知$p-1$维列向量, $x_i = (x_{i1}, \cdots, x_{i(p-1)})'$, $i = 1, \cdots, n$; a, b为待定参数, $a \in R$, $b = (b^{(1)}, \cdots, b^{(p-1)})'$为维回归系数向量; $e_1, \cdots, e_n$独立且各中位数为零. 令$t_i = (x_i', y_i)'$, $i = 1, \cdots, n$为p维欧氏空间R^p中的数据点, 记$T = \{t_i, \cdots, t_n\}$表示模型(1.1)的数据点集合.

最小一乘准则就是寻找回归超平面

$$y = a + b' x. \tag{1.2}$$

其中变量$y \in R$, $x = (x^{(1)}, \cdots, x^{(p-1)})' \in R^p$, 使得$\sum\limits_{i=1}^{n} |y_i - a - b' x_i|$达到最小. 令$c = (a, b')'$, 则目标函数为

$$f(c) = \sum_{i=1}^{n} |y_i - a - b' x_i| = \sum_{i=1}^{n} |y_i - \langle c, (1, x_i')' \rangle|, \tag{1.3}$$

其中$\langle \ \rangle$表示内积.

对最小一乘准则而言, 由于只考虑残差的一次方而非平方, 故所受异常点的影响就小得多, 因此最小一乘准则比最小二乘准则有更好的稳健性[1][6]. Charnds等把最小一乘估计的计算同线性规划联系起来, 实现了求解, 但是在样本量n较大的时候, 转化成线性规划

*福州大学科技发展基金资助(2005-XQ-17).

本文2005年10月27日收到, 2007年10月27日收到修改稿.

时变量、附加变量及约束条件的个数急剧增加, 工作量和计算量都大得惊人[2][3], 在19世纪20年代Fourier提出了类似单纯形法的迭代法[4], 但最小一乘的计算难题一直存在, 本文的快速算法正好解决了上述计算难的问题.

§2. 有关定义和引理

引理 2.1[2] 存在p维向量c_0, 使

1. 由(1.3)式定义的函数$f(c)$, 在$c=c_0$处达到最小.

2. 记$r_i(c)=y_i-c'x_i$, $W_c=\{t_i:1\le i\le n, r_i(c)=0\}$. W_{c_0}中的向量张成的线性子空间就是$t_1,\cdots,t_n$所生成的线性子空间$M(t_1,\cdots,t_n)$.

这就是1809年Gauss得到的定理.

若$t_i\in W_{c_0}$, 则超平面$y=a_0+b_0'x=a_0+b_0^{(1)}x^{(1)}+\cdots+b_0^{(p-1)}x^{(p-1)}$通过样本点$t_i$, 一般地$M(t_1,\cdots,t_n)=$ 全空间R^p. 这时, 由W_{c_0}生成$M(t_1,\cdots,t_n)$可知, W_{c_0}中至少包含p个样本点, 即超平面$y=a_0+b_0'x$至少通过p个样本点. 这个引理同时揭示了最小一乘与最小二乘的一个重大差别, 即最小一乘回归超平面一定通过p个样本点. 曾有人据此认为最小一乘不好, 因为它在全部n个样本点中, 只用了p个. 其实不然, 因为究竟用哪p个样本点, 要由全体n个样本点来决定.

设$c\in R^p$, 记$r_i(c)=y_i-c'x_i$, $Z_c=\{i:r_i(c)=0\}$, $|Z_c|$表示集合Z_c中元素的个数.

定义 2.1 如果$\dim\{t_i, i\in Z_c\}=p$, 则称$c\in R^p$为极点. 显然此时$|Z_c|\ge p$. 若$|Z_c|>p$, 则称c为退化极点; 否则若$|Z_c|=p$, 则称c为非退化极点.

定义 2.2 记$M=\{c\in R^p: f(c)\le f(d), \forall\, d\in R^p\}$, 极点$c\in M$时, 称$c$为最优极点. 显然最优极点就是模型(1.1)的最小一乘估计.

定义 2.3 若模型(1.1)中所有极点都是非退化极点, 则称模型(1.1)是非退化的, 否则称其是退化的.

本文就是要在模型(1.1)非退化的情况下讨论最优极点的求解.

对于欧式空间R^p中的任意p个数据点, 不妨设为$t_1,\cdots,t_p$, 若矩阵

$$((1,x_i')',\cdots,(1,x_p')')'=\begin{bmatrix}1 & x_{11} & \cdots & x_{1(p-1)}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_{p1} & \cdots & x_{p(p-1)}\end{bmatrix}$$

可逆, 将数据点$t_1,\cdots,t_p$坐标代入(1.2)式, 可得通过它们的超平面为$y=a+b'x$, 其中

$$c=(a,b')=\varphi(t_1,\cdots,t_p)=\begin{bmatrix}1 & x_{11} & \cdots & x_{1(p-1)}\\ \vdots & \vdots & \vdots & \vdots\\ 1 & x_{p1} & \cdots & x_{p(p-1)}\end{bmatrix}^{-1}\begin{pmatrix}y_1\\ \vdots\\ y_p\end{pmatrix}, \tag{2.1}$$

其中函数$\varphi(t_1,\cdots,t_p)$表示为覆盖数据点$t_1,\cdots,t_p$的超平面的系数向量，显然它是极点．此时残差绝对值之和$f(c)$也是关于数据点$t_1,\cdots,t_p$的函数，设为$g(t_1,\cdots,t_p)$，即

$$g(t_1,\cdots,t_p)=\sum_{i=1}^{n}|y_i-\langle\varphi(t_1,\cdots,t_p),x_i\rangle|. \tag{2.2}$$

易见作用在数据点集合T上的函数g与$t_1,\cdots,t_p$的排列顺序无关．若$((1,x_i')',\cdots,(1,x_p')')'$不可逆，不妨令$g(t_1,\cdots,t_p)=\infty$.

引理 2.2[5]　对于模型(1.1)，若p个数据点$t_{i_1},\cdots,t_{i_p}\in T$满足

$$g(t_{i_1},\cdots,t_{i_p})=\min_{t_{j_1},\cdots,t_{j_p}\in T}g(t_{j_1},\cdots,t_{j_p}), \tag{2.3}$$

则$c^*=\varphi(t_{i_1},\cdots,t_{i_p})$就是回归系数$c$的最小一乘估计，即最优极点．

显然遍历数据点集合T，可得到满足式(2.3)的p个数据点．但是这样搜寻的计算量很大，需要对n个数据点中的任意p个点计算回归系数及残差绝对值之和，共有C_n^p种组合方式，可想而知当n很大时，计算量将异常之大．

定义 2.4　在线性回归模型(1.1)中若p个数据点$t_{i_1},\cdots,t_{i_p}\in T$满足

$$g(t_{i_1},\cdots,t_{i_p})=\min_{t_k\in T}g(t_{i_1},\cdots,t_{i_{j-1}},t_k,t_{i_{j+1}},\cdots,t_{i_p}),\qquad j=1,\cdots,p, \tag{2.4}$$

则称$c^0=\varphi(t_{i_1},\cdots,t_{i_p})$为稳定极点．

引理 2.3[5]　对于非退化的模型(1.1)，$c^*\in R^p$是最优极点当且仅当c^*是稳定极点．

由引理2.3知，模型(1.1)的稳定极点就是最优极点，即回归系数的最小一乘估计，因此只要找到模型(1.1)的稳定极点，就解决了求解最小一乘估计的问题．下面是求解稳定极点的迭代算法．

§3.　稳定极点的迭代算法描述

若回归模型(1.1)中只含一个自变量的情况，即$p=1$时，我们要寻找a,b使目标函数(1.3)式达到最小．下面先介绍两种特殊情况．

1. 假定$b=0$，求a.

把$y_1,y_2,\cdots,y_n$按由小到大的次序排列为：$y_{(1)},y_{(2)},\cdots,y_{(n)}$，可以证明：若$n$为奇数，则$y_{(n+1)/2}$是使目标函数$\sum\limits_{i=1}^{n}|y_i-a|$达到最小的唯一解；若$n$为偶数，则区间$[y_{n/2},y_{n/2+1}]$内任一数都是该最小一乘问题的解．

2. 假定$a=0$，求b.

此时最小一乘准则就是使表达式$\sum\limits_{i=1}^{n}|y_i-b'x_i|$达到最小．这是一个配过原点的回归直线问题．不妨可设$x_i\neq0$，$i=1,\cdots,n$．计算$y_i/x_i,\cdots,y_n/x_n$，把它们按由小到大的次

序排列为$y_{(1)}/x_{(1)} \le y_{(2)}/x_{(2)} \le \cdots \le y_{(n)}/x_{(n)}$. 令$M=\sum\limits_{i=1}^{n}|x_i|$, 计算$M_1=|x_{(i)}|$, $M_2=|X_{(1)}|+|X_{(2)}|$, 直到$M_i=\sum\limits_{j=1}^{i}|x_{(j)}|$, 满足$M_i \ge M/2$. 可以证明$y_{(i)}/x_{(i)}$就是回归系数$b$的最小一乘估计. 详细证明过程见参考文献[1].

下面再介绍模型(1.1)稳定极点的求解方法.

1. 任意给出初始p个数据点$t_1^{(1)},t_2^{(1)},\cdots,t_p^{(1)}$;

2. 固定$t_2^{(1)},\cdots,t_p^{(1)}$, 在剩余的$n-p+1$个点中寻找$t_1^{(2)}$使

$$g(t_1^{(2)},t_2^{(1)},\cdots,t_p^{(1)})=\min_{t_i\in T} g(t_i,t_2^{(1)},\cdots,t_p^{(1)});$$

固定$t_1^{(2)},t_3^{(1)},\cdots,t_p^{(1)}$, 在剩余的$n-p+1$个点中寻找$t_2^{(2)}$使

$$g(t_1^{(2)},t_2^{(2)},t_3^{(1)},\cdots,t_p^{(1)})=\min_{t_i\in T} g(t_1^{(2)},t_i,t_3^{(1)},\cdots,t_p^{(1)});$$

同理, 依次找到$t_3^{(2)},\cdots,t_p^{(2)}$, 且显然有

$$g(t_1^{(2)},\cdots,t_{i-1}^{(2)},t_i^{(2)},t_{i+1}^{(1)},\cdots,t_p^{(1)}) \le g(t_1^{(2)},\cdots,t_{i-1}^{(2)},t_i^{(1)},\cdots,t_p^{(1)}), \qquad i=1,\cdots,p,$$
$$g(t_1^{(2)},\cdots,t_p^{(2)}) \le g(t_1^{(1)},\cdots,t_p^{(1)});$$

3. 若$\forall\, i=1,\cdots,p$, 有$t_i^{(1)}=t_i^{(2)}$, 则迭代结束. 不然, 转到步骤2继续寻找数据点$t_1^{(3)}$, $t_2^{(3)},\cdots,t_p^{(3)}$, 并产生数据点组合序列$\{(t_1^{(i)},\cdots,t_p^{(i)})|i=1,2,\cdots\}$及相应残差绝对值和序列$\{g(t_1^{(i)},\cdots,t_p^{(i)})|i=1,2,\cdots\}$.

§4. 稳定极点求解迭代算法的证明

定理 4.1 述迭代过程产生的数据点组合序列$\{(t_1^{(i)},\cdots,t_p^{(i)})|i=1,2,\cdots\}$及相应的残差绝对值之和序列$\{g(t_1^{(i)},\cdots,t_p^{(i)})|i=1,2,\cdots\}$均为有限序列, 分别为$\{(t_1^{(i)},\cdots,t_p^{(i)})|i=1,2,\cdots,k\}$, $\{g(t_1^{(i)},\cdots,t_p^{(i)})|i=1,2,\cdots,k\}$, 其中$k\le C_n^p$; 且$(t_1^{(k)},\cdots,t_p^{(k)})$满足式(2.4), 即$\varphi(t_1^{(k)},\cdots,t_p^{(k)})$为稳定极点.

证明: 根据迭代过程, 数据点$t_j^{(i)}$, $j\in 1,\cdots,p$, 满足

$$g(t_1^{(i)},\cdots,t_{j-1}^{(i)},t_j^{(i)},t_{j+1}^{(i-1)},\cdots,t_p^{(i-1)})=\min_{t_k\in T} g(t_1^{(i)},\cdots,t_{j-1}^{(i)},t_k,t_{j+1}^{(i-1)},\cdots,t_p^{(i-1)}),$$

则显然

$$g(t_1^{(i)},\cdots,t_{j-1}^{(i)},t_j^{(i)},t_{j+1}^{(i-1)},\cdots,t_p^{(i-1)}) \le g(t_1^{(i)},\cdots,t_{j-1}^{(i)},t_j^{(i-1)},\cdots,t_p^{(i-1)}),$$

从而

$$g(t_1^{(i)},\cdots,t_p^{(i)}) \le g(t_1^{(i-1)},\cdots,t_p^{(i-1)}).$$

另外存在p个数据点满足(2.3)式, 则序列$g(t_1^{(i)},\cdots,t_p^{(i)})$单调减小且有下界, 则序列$g(t_1^{(i)},\cdots,t_p^{(i)})$收敛. 且数据点集合$T$中仅有$C_n^p$种不同数据点组合方式, 因此存在$1\le k\le C_n^p$, 使

$$t_i^{(k)}=t_i^{(k+1)}, \qquad i=1,2,\cdots,p,$$

即迭代结束, 定理得证. □

事实上, 由算法描述过程及定理4.1易知, 算法产生的数据点组合序列

$$\{(t_1^{(i)},\cdots,t_p^{(i)}),(t_1^{(i+1)},t_2^{(i)},\cdots,t_p^{(i)}),\cdots,(t_1^{(i+1)},\cdots,t_p^{(i+1)})|i=1,2,\cdots\},$$

其中只要连续有p个数据点组合完全相同, 该序列就已经收敛了, 迭代便可结束.

下面说明已知$p-1$个数据点$t_{i_1},\cdots,t_{i_{p-1}}\in T$, 如何在剩余$n-p+1$个数据点中寻找一个样本点$t_{i_p}$使它们组合后回归效果最佳, 即回代残差绝对值和最小的方法, 即该点t_{i_p}满足

$$g(t_{i_1},\cdots,t_{i_{p-1}})=\min_{t_j\in T} g(t_{i_1},\cdots,t_{i_{p-1}}). \tag{4.1}$$

显然遍历数据点集合T就可找到$t_{i_p}\in T$, 事实上并不需要遍历集合T就可找到满足(4.1)式极值条件的数据点t_{i_p}, 其计算过程见定理4.2证明.

定理 4.2 对于回归模型(1.1), 任意不同在$p-1$维欧氏空间的数据点$t_{i_1},\cdots,t_{i_{p-1}}\in T$, $\{c_a\in R^p, a\in A\}$是所有回归系数c的估计类且满足:

$$\{t_{i_1},\cdots,t_{i_{p-1}}\}\subseteq Z_{c_a}, \tag{4.2}$$

则存在估计c_{a^*}, $a^*\in A$满足

$$(1)\quad f(c_{a^*})\le f(c_a), \quad a\in A; \tag{4.3}$$

$$(2)\quad \text{存在数据点 } t_{i_p}\in Z_{c_a^*}, \text{ 且 } t_{i_p}\notin\{t_{i_1},\cdots,t_{i_{p-1}}\}. \tag{4.4}$$

证明: 由于$\{t_{i_1},\cdots,t_{i_{p-1}}\}\subseteq Z_{c_a}$, 即超平面为$y=a_a+b_a'x$通过样本点$t_{i_1},\cdots,t_{i_{p-1}}$, 且因为它们不同在任意$p-1$维欧氏空间, 也就是$(x_{i_1},\cdots,x_{ip-1})'$可逆, 则计算可得

$$b_a=[(x_{i_1},\cdots,x_{i_{p-1}})']^{-1}[(y_{i_1},\cdots,y_{i_{p-1}})'-a_a(1,\cdots,1)'].$$

令$\gamma=[(x_{i_1},\cdots,x_{i_{p-1}})']^{-1}(y_{i_1},\cdots,y_{i_{p-1}})'$, $\beta=[(x_{i_1},\cdots,x_{i_{p-1}})']^{-1}(1,\cdots,1)'$, 则有

$$b_a=\gamma-a_a\beta.$$

从而回归方程为

$$y = a_a + \gamma' x - a_a \beta' x = \gamma' x + a_a(1 - \beta' x).$$

残差绝对值之和为

$$f(c_a) = \sum_{i=1}^{n} |y_i - \gamma' x_i - a_a(1 - \beta' x_i)|. \tag{4.5}$$

上式中$f(c_a)$是仅关于参数a_a的函数, 令

$$Y_i = y_i - \gamma' x_i, \quad X_i = 1 - \beta' x_i, \qquad i = 1, \cdots, n. \tag{4.6}$$

显然a_a是最小一乘估计的解, 即是回归模型(1.1)中$p = 1$, $a = 0$的情况. 定理得证. □

求解过程: 先计算Y_i与X_i的比值

$$l_i = \frac{Y_i}{X_i} = \frac{y_i - \gamma' x_i}{1 - \beta' x_i}, \qquad i = 1, \cdots, n;$$

将它们从小到大排列为: $l_{(1)} \le l_{(2)} \le \cdots \le l_{(n)}$; 计算$M = \sum\limits_{i=1}^{n} |1 - \beta' x_i|$, $M_i = \sum\limits_{j=1}^{i} |X_{(j)}|$, $i = 1, 2, \cdots$, 其中$X_{(j)}$序列为X_i序列随$l_{(i)}$的重排, 直到达到某个i_0使得$M_{i_0} \ge M/2$, 则$a_a = l_{(i_0)}$使(4.5)式达到最小, 即找到$t_{i_p} = t_{i_0} \in T$, 则

$$a_{a^*} = \frac{y_{i_p} - \gamma' x_{i_p}}{1 - \beta' x_{i_p}}, \qquad c_{a^*} = (a_{a^*}, \gamma' - a_{a^*} \beta'), \tag{4.7}$$

使$f(c_{a^*}) \le f(c_a)$, $a \in A$, 且$t_{i_p} \in Z_{c_{a^*}}$.

定理4.2的证明过程说明了已知$p-1$个数据点$t_{i_1}, \cdots, t_{i_{p-1}} \in T$, 如何在剩余$n-p+1$个数据点中寻找一个样本点$t_{i_p}$使回代残差绝对值和达到最小的方法, 使得本算法更加简单可行.

§5. 计算过程演示及与线性规划求解的比较

5.1 计算过程演示

以任意p个数据点作为本算法的初始点组进行迭代, 都能快速准确地得到最小一乘估计解, 下面通过一个随机数据演示本算法的求解过程. 以下数据中$p = 3$, 不妨取(t_1, t_2, t_3)作为初始点组.

表1 随机数据

数据点	t_1	t_2	t_3	t_4	t_5	t_6
x_1	1	2	3	4	5	6
x_2	1	3	2	1	3	2
y	1	4	3	3	5	5

(1) 去掉点组中的第一个点t_1, 得到

$$\widehat{X}=\begin{bmatrix}2&3\\3&2\end{bmatrix},\qquad \widehat{X}^{-1}=\begin{bmatrix}-0.4&0.6\\0.6&-0.4\end{bmatrix},\qquad \widehat{Y}=\begin{bmatrix}4\\3\end{bmatrix},$$

然后在剩下的四个点(t_1,t_4,t_5,t_6)中根据(4.6)式计算:

表2　第一步迭代

数据点	t_1	t_4	t_5	t_6
X_i	0.6	0	-0.6	-0.6
Y_i	-0.4	1	0.4	1.4
l_i	-2/3	∞	-2/3	-7/3

依据定理4.2求解过程可找到样本点t_5或t_1, 可将点t_5替换掉t_1, 得到新的点组(t_5,t_2,t_3).

(2) 去掉点组中的第二个点t_2, 得到

$$\widehat{X}=\begin{bmatrix}5&3\\3&2\end{bmatrix},\qquad \widehat{X}^{-1}=\begin{bmatrix}2&-3\\-3&5\end{bmatrix},\qquad \widehat{Y}=\begin{bmatrix}5\\3\end{bmatrix},$$

再在剩下的四个点(t_1,t_2,t_4,t_6)中根据(4.6)式计算:

表3　第二步迭代

数据点	t_1	t_2	t_4	t_6
X_i	0	-3	3	3
Y_i	0	2	-1	-1
l_i	∞	-2/3	-1/3	-1/3

依据定理4.2求解过程可找到t_4, 将其替换掉点组中的第2点, 得到新的点组(t_5,t_4,t_3). 特别地这里的Y_i/X_i从小到大的排序方式有两种即点(2,4,6,1)和点(2,6,4,1), 任选一种都可以.

(3) 去掉点组中的第三个点t_3, 得到

$$\widehat{X}=\begin{bmatrix}5&3\\4&1\end{bmatrix},\qquad \widehat{X}^{-1}=\begin{bmatrix}-1/7&3/7\\4/7&-5/7\end{bmatrix},\qquad \widehat{Y}=\begin{bmatrix}5\\3\end{bmatrix},$$

然后在剩下的四个点(t_1,t_2,t_3,t_6)中根据(4.6)式计算:

表4　第三步迭代

数据点	t_1	t_2	t_3	t_6
X_i	6/7	6/7	3/7	-3/7
Y_i	-2/7	5/7	-1/7	1/7
l_i	-1/3	5/6	-1/3	-1/3

依据定理4.2可找到t_3, 将其替换点组中的第3点, 得到新的点组(t_5,t_4,t_3). 此处Y_i/X_i从小到大的排序方式有6种点序: (1,3,6,2), (1,6,3,2), (3,1,6,2), (3,6,1,2), (6,1,3,2)和(6,3,1,2).

这样一轮迭代完成, 但这一轮中的点组不完全相同. 所以必须进行下一轮迭代, 继续从第一个点t_5开始. 接下来依次找到$(t_5,t_4,t_3)\to(t_5,t_4,t_3)$, 而该点组在点组序列中连续出现了3次, 由于$p=3$, 故认为迭代结束, 取$(t_5,t_4,t_3)$作为最终的解点组.

表5　算法给出的表1数据的最小一乘估计

有效的解点组	(t_1,t_3,t_4,t_5,t_6)中任取3个都可构成解点组
回归系数	$\beta^*=(-0.333333,0.666667,0.666667)$　　残差绝对值和 $=1.000000$
超平面方程	$y=-0.333333+0.666667x_1+0.666667x_2$　　(保留六位小数)

5.2　利用线性规划进行最小一乘的求解[2]

易知最小一乘估计的目标函数为

$$\min f(\beta)=\sum_{i=1}^n|y_i-a-b'x_i|=\sum_{i=1}^n|y_i-\beta' z_i|,$$

其中$\beta=(a,b)'$, $z_i=(1,x_i)'$, $i=1,\cdots,n$; 向量β可表示成$d-e$的形式, 即$\beta_i=d_i-e_i$, 其中d,e各分量都非负, 即$d_i\geq 0$, $e_i\geq 0$. 引进如下矩阵、向量:

$$A=\begin{pmatrix} z_i' \\ \vdots \\ z_n' \end{pmatrix},\qquad B=\begin{pmatrix} I_n & A & -A \\ I_n & -A & A \end{pmatrix},$$

$$y=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},\qquad r=\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix},\qquad l=\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},\qquad 0=\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

其中I_n为n阶单位矩阵, $y,r,l,0$为n维向量. 可得最小一乘估计的线性规划问题:

$$\text{目标函数: } \min(l',0',0')\begin{pmatrix} r \\ d \\ e \end{pmatrix}=\min\sum_{i=1}^n r_i,\quad \text{约束条件 St: }\begin{cases} B\begin{pmatrix} r \\ d \\ e \end{pmatrix}\geq\begin{pmatrix} y \\ -y \end{pmatrix} \\ \begin{pmatrix} r \\ d \\ e \end{pmatrix}\geq 0 \end{cases}.$$

一旦解得r, d, e的值, 则回归系数$\beta = d - e$就可确定, 它就是使最小一乘估计目标函数$f(\beta)$达到最小的解, 且该目标函数的最小值就是$\sum_{i=1}^{n} r_i$.

表6 采用线性规划软件Lindo6.01的计算程序和结果

Min R1+R2+R3+R4+R5+R6	LP OPTIMUM FOUND AT STEP 4
st	OBJECTIVE FUNCTION VALUE 1.000000
R1+D1+1D2+1D3−E1−1E2−1E3 >=1	VARIABLE VALUE REDUCED COST
R2+D1+2D2+3D3−E1−2E2−3E3 >=4	R1 0.000000 1.000000
R3+D1+3D2+2D3−E1−3E2−2E3 >=3	R2 1.000000 0.000000
R4+D1+4D2+1D3−E1−4E2−1E3 >=3	R3 0.000000 0.000000
R5+D1+5D2+3D3−E1−5E2−3E3 >=5	R4 0.000000 1.000000
R6+D1+6D2+2D3−E1−6E2−2E3 >=5	R5 0.000000 0.000000
R1−D1−1D2−1D3+E1+1E2+1E3 >=−1	R6 0.000000 0.000000
R2−D1−2D2−3D3+E1+2E2+3E3 >=−4	D1 0.000000 0.000000
R3−D1−3D2−2D3+E1+3E2+2E3 >=−3	D2 0.666667 0.000000
R4−D1−4D2−1D3+E1+4E2+1E3 >=−3	D3 0.666667 0.000000
R5−D1−5D2−3D3+E1+5E2+3E3 >=−5	E1 0.333333 0.000000
R6−D1−6D2−2D3+E1+6E2+2E3 >=−5	E2 0.000000 0.000000
End	E3 0.000000 0.000000
回归系数$\beta^* =$ (D1−E1,D2−E2,D3−E3)	$(-0.333333, 0.666667, 0.666667)$
残差绝对值和	1.00000

经比较, 本算法和线性规划得到的结果完全相同. 但线性规划求解最小一乘估计问题需设置$n + 2p$个变量和$3n + 2p$个约束条件, 当$n(p)$较大时变得极其繁琐且有可能难以计算. 若用穷举搜索法, 需针对C_n^p种点组组合求解$\widehat{\beta}$并计算和比较目标函数值$f(\widehat{\beta})$, 同样当$n(p)$较大时, 计算量极其巨大, 几乎无法计算; 而本文算法无变量和约束条件的限制, 在计算量上也大大减少, 且易于编程实现, 具备可行性, 准确性和快速性.

在退化模型前提下, 如何快速找出构成最小一乘超平面的所有点组, 以及这些特殊的点组的应用有待进一步的研究.

作者对审稿专家提出的修改建议表示衷心的感谢.

参 考 文 献

[1] 陈希孺, 最小一乘线性回归(上), 数理统计与管理, **1989(5)**(1989), 48–55.

[2] 陈希孺, 最小一乘线性回归(下), 数理统计与管理, **1989(6)**(1989), 48–55.

[3] 谢开贵, 宋乾坤, 周家启, 最小一乘线性回归模型研究, 系统仿真学报, **14(2)**(2002), 99–102.

[4] Fisher, W.D., A note on curve fitting with minmum deviations by linear programming, *journal of American Statistical Association*, **56**(1961), 359–361.

[5] 刘文丽, 吕书龙, 最小一乘估计的几个性质, 福州大学学报(自然科学版), **32(4)**(2004), 442–444.

[6] 周秀轻, 王金德, 随机删失数据非线性回归模型的最小一乘估计, 中国科学A辑(数学), **35(4)**(2005), 387–403.

[7] Bloomfield, P., Steiger, W.L., *Least Absolute Deviation: Theory, Application and Algorithms*, Birkhäuser, Boston, 1983.

Quick Algorithm for Least Absolute Deviation Estimator

Lu Shulong Liu Wenli

(*Colleage of Mathematics and Computer Science, Fuzhou University, Fuzhou, 350108*)

Least square estimator (LSE) is disturbed easily by singular point; least absolute deviation estimator (LADE) can overcome the influence of singular point, but it is difficult in calculation. A convergent algorithm for LADE based on the stable pole theorem of LADE under non-degenerate model is obtained in this paper. The progress of algorithm and comparison of linear programming are derived. Further this algorithm makes LADE more effective.

Keywords: Least absolute deviation, least square, convergent algorithm.

AMS Subject Classification: 62G35, 65C60.