DOI:10.13203/j.whugis20150218

文章编号:1671-8860(2017)08-1172-07

病态方程基于 Liu 估计的一种迭代估计新方法

姜兆英[1,2]　刘国林[1]　于胜文[1]

1　山东科技大学测绘科学与工程学院,山东 青岛,266590

2　青岛农业大学理学与信息科学学院,山东 青岛,266109

摘　要:当线性回归模型的设计矩阵病态时,最小二乘(least square,LS)估值方差大且不稳定,已不是一种优良估计。为了减弱病态性,许多有偏估计法如岭估计、主成分估计、Liu 估计等被提出。基于 Liu 估计,引入迭代的思想,提出了一种新的有偏估计法—迭代估计法。借助对称正定矩阵的谱分解,将迭代公式转化为便于解算的解析表达式,并证明迭代公式在修正因子 $d\in[-1,1]$是收敛的。基于 Liu 估计中修正因子 d 的确定方法,在均方误差最小的情况下给出最优修正因子 d 的确定公式。最后,分别利用 LS 估计、岭估计、Liu 估计和提出的迭代估计对两个算例进行计算并给出实验结果。在第一个算例中,对观测向量添加不同的扰动,结果表明迭代估计法具有更强的抗干扰能力;第二个算例的结果表明,迭代估计法所得结果更接近于真值,即迭代估计法在均方误差意义下优于 LS 估计、岭估计和 Liu 估计。

关键词:复共线性;条件数;Liu 估计;岭估计;均方误差

中图法分类号:P207　　**文献标志码**:A

最小二乘估计自 Gauss-Markov 定理建立后,便作为一个良好的估计而广泛应用。然而,在应用中人们逐渐发现,法方程系数阵呈病态时,模型设计阵的列向量间存在复共线性关系(multi-collinearity),这往往会引起 LS 估计性质显著变坏,模型的解不稳定。为了解决法矩阵病态的问题,研究学者提出了有偏估计这一理论和方法。目前常用的有偏估计有 Stein 压缩估计[1,2]、主成分估计[3]、岭估计[4-6]、Liu 估计[7-9]等,它们在均方误差(mean squared error, MSE)准则下均局部或一致地改进了参数的 LS 估计。但当存在较严重的复共线性时,岭估计、Liu 估计不能完全克服病态问题。除此之外,谱修正迭代法[10,11]、遗传算法[12]等也是常用的解决病态问题的方法。

考虑线性模型:

$$\boldsymbol{L}=\boldsymbol{A}\boldsymbol{X}+\boldsymbol{\Delta} \tag{1}$$

式中,设计阵 $\boldsymbol{A}$ 为已知的 $m\times n(m>n)$阶列满秩阵;$\boldsymbol{X}$ 为 $n\times 1$ 阶未知参数向量;$\boldsymbol{L}$ 为 $m\times 1$ 阶观测向量;$\boldsymbol{\Delta}$ 为 $m\times 1$ 阶误差向量;且 $E(\boldsymbol{\Delta})=0$, $\mathrm{Cov}(\boldsymbol{\Delta})=\sigma^2\boldsymbol{I}$。其经典最小二乘估值为:

$$\hat{\boldsymbol{X}}_{\mathrm{LS}}=(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{L} \tag{2}$$

当 $\boldsymbol{A}'\boldsymbol{A}$ 病态,即它的特征根中至少有一个近似于零,$(\boldsymbol{A}'\boldsymbol{A})^{-1}$中的元素变得很大,一方面导致最小二乘估值不稳定且偏离真值;另一方面会放大模型观测中干扰的影响。所以,若模型(1)病态,LS 估计已不再适应。判断矩阵 $\boldsymbol{A}'\boldsymbol{A}$ 是否“病态”或“病态”的程度,可以采用条件数来衡量。$\boldsymbol{A}'\boldsymbol{A}$(因 $\boldsymbol{A}$ 是列满秩的,所以 $\boldsymbol{A}'\boldsymbol{A}$ 为对称正定矩阵)的条件数为[13]:

$$\mathrm{cond}(\boldsymbol{A}'\boldsymbol{A})=\frac{r_{\max}}{r_{\min}} \tag{3}$$

式中,$r_{\max}$、$r_{\min}$表示矩阵 $\boldsymbol{A}'\boldsymbol{A}$ 的最大、最小特征值。条件数越大,矩阵 $\boldsymbol{A}'\boldsymbol{A}$ 病态性越严重。为了解决 $\boldsymbol{A}'\boldsymbol{A}$ 病态问题,Hoerl 和 Kennard[4]在 1970 年提出了岭估计,则模型(1)的岭估计为:

$$\hat{\boldsymbol{X}}_{\lambda}=(\boldsymbol{A}'\boldsymbol{A}+\lambda\boldsymbol{I})^{-1}\boldsymbol{A}'\boldsymbol{L} \tag{4}$$

式中,$\lambda>0$ 为岭参数,其确定方法有 L-曲线法[14,15]、岭迹法[16]、广义交叉效验(generalized cross-validation ,GCV)法[17]等。这里 $\boldsymbol{A}'\boldsymbol{A}+\lambda\boldsymbol{I}$ 的条件数为:

收稿日期:2015-09-29

项目资助:国家自然科学基金(41274007,41404003);山东省自然科学基金(ZR2012DM001);高等学校博士学科点专项科研基金(20123718110001);泰山学者建设工程。

第一作者:姜兆英,博士,主要从事 InSAR 数据处理和模型算法研究。15964287012@163.com

通讯作者:刘国林,博士,教授。gliu@sdust.edu.cn

$$\frac{r_{\max}+\lambda}{r_{\min}+\lambda} \triangleq f(\lambda) \tag{5}$$

显然,$\lambda>0$ 时有如下关系:

$$\mathrm{cond}(\boldsymbol{A}'\boldsymbol{A}+\lambda\boldsymbol{I}) < \mathrm{cond}(\boldsymbol{A}'\boldsymbol{A}) \tag{6}$$

且

$$f'(\lambda)=\frac{r_{\min}-r_{\max}}{(r_{\min}+\lambda)^2}<0 \tag{7}$$

故 $\mathrm{cond}(\boldsymbol{A}'\boldsymbol{A}+\lambda I)$是 λ 的递减函数。所以,当 $\boldsymbol{A}'\boldsymbol{A}$ 严重病态时,可选取足够大的 λ 使 $\boldsymbol{A}'\boldsymbol{A}+\lambda\boldsymbol{I}$ 的条件数降到预期值。但参数 λ 太大,虽然可以稳定地求解,但该解与模型(1)的实际解相差甚远,是一个相当糟糕的逼近。但若 λ 太小,则对问题病态性的改善不起作用,解的不稳定性依然存在。1993 年,Liu 结合岭估计和 Stein 估计提出了一种新的有偏估计:

$$\hat{\boldsymbol{X}}_d=(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}(\boldsymbol{A}'\boldsymbol{L}+d\hat{\boldsymbol{X}}_{\mathrm{LS}}) \tag{8}$$

式中,$0<d<1$ 为一修正因子。特别当 $d=1$ 时,Liu 估计为经典的最小二乘估计。同时 Liu 还给出了其推广形式:

$$\hat{\boldsymbol{X}}_d=(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}(\boldsymbol{A}'\boldsymbol{L}+\boldsymbol{D}\hat{\boldsymbol{X}}_{\mathrm{LS}}) \tag{9}$$

式中,$\boldsymbol{D}=\mathrm{diag}(d_i)$,为修正因子 d_i 的对角矩阵,且 $0<d_i<1(i=1,2,\cdots,n)$。文献[7]指出 Liu 估计可得到比 LS 估计更小的 MSE,但相较于岭估计,Liu 估计得到的 MSE 不一定优于岭估计。由于 $\hat{\boldsymbol{X}}_d$ 是参数 d 的线性函数,故 d 的确定比岭参数的确定要容易些。2008 年,Sakallioglu[18] 基于岭估计又提出了另外一种有偏估计:

$$\hat{\boldsymbol{X}}_{\lambda,d}=(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}(\boldsymbol{A}'\boldsymbol{L}+d\hat{\boldsymbol{X}}_{\lambda}) \tag{10}$$

式中,$-\infty<d<+\infty$。该估计就是将 Liu 估计中的 LS 估计用岭估计代替,并将修正因子 d 推广到可取任意实数。在文献[18]中,Sakallioglu 证明了该方法在 MSE 意义下优于最小二乘、岭估计和 Liu 估计。本文在 Liu 估计的基础上引入迭代求解的思想,提出了一种迭代估计方法。通过实验验证了该方法在均方误差衡量准则下优于 LS 估计、岭估计和 Liu 估计。

1　迭代估计公式及其收敛性

1.1　迭代估计公式的推导

虽然 Liu 估计能得到比 LS 更小的 MSE,但相较于岭估计,只有在满足一定条件下 Liu 估计得到的 MSE 才会小于岭估计得到的[19]。为此,本文基于 Liu 估计,引入迭代的思想,提出一种有偏迭代估计法。

对于初值的选取可以采用两种方式,一是 LS 估计,另一种为岭估计。下面以 LS 估计作为初值为例给出迭代公式。

取 $\hat{\boldsymbol{X}}^{(0)}=\hat{\boldsymbol{X}}_{\mathrm{LS}}$,为了方便书写,记 $\boldsymbol{q}=(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}$,则:

$$\begin{aligned}\hat{\boldsymbol{X}}^{(1)}&=(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}(\boldsymbol{A}'\boldsymbol{L}+d\hat{\boldsymbol{X}}_{\mathrm{LS}})=\\&\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+d\boldsymbol{q}\hat{\boldsymbol{X}}^{(0)}\end{aligned} \tag{11}$$

将 $\hat{\boldsymbol{X}}^{(1)}$ 代替式(11)中的 $\hat{\boldsymbol{X}}^{(0)}$,得第二次估值:

$$\begin{aligned}\hat{\boldsymbol{X}}^{(2)}&=\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+\boldsymbol{q}d\hat{\boldsymbol{X}}^{(1)}=\\&\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+\boldsymbol{q}d(\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+\boldsymbol{q}d\hat{\boldsymbol{X}}^{(0)})=\\&\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+d\boldsymbol{q}^2\boldsymbol{A}'\boldsymbol{L}+d^2\boldsymbol{q}^2\hat{\boldsymbol{X}}^{(0)}\end{aligned} \tag{12}$$

同样将 $\hat{\boldsymbol{X}}^{(2)}$ 代替式(12)中的 $\hat{\boldsymbol{X}}^{(1)}$,得第三次估值:

$$\begin{aligned}\hat{\boldsymbol{X}}^{(3)}&=\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+\boldsymbol{q}d\hat{\boldsymbol{X}}^{(2)}=\\&\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+d\boldsymbol{q}(\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+d\boldsymbol{q}^2\boldsymbol{A}'\boldsymbol{L}+\\&d^2\boldsymbol{q}^2\hat{\boldsymbol{X}}^{(0)})=\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+d\boldsymbol{q}^2\boldsymbol{A}'\boldsymbol{L}+\\&d^2\boldsymbol{q}^3\boldsymbol{A}'\boldsymbol{L}+d^3\boldsymbol{q}^3\hat{\boldsymbol{X}}^{(0)}\end{aligned} \tag{13}$$

依次类推,即得迭代公式为:

$$\begin{aligned}\hat{\boldsymbol{X}}^{(k)}&=\boldsymbol{q}\boldsymbol{A}'\boldsymbol{L}+d\boldsymbol{q}\hat{\boldsymbol{X}}^{(k-1)}=(\boldsymbol{q}+d\boldsymbol{q}^2+d^2\boldsymbol{q}^3+\cdots+d^{k-1}\boldsymbol{q}^k)\boldsymbol{A}'\boldsymbol{L}+d^k\boldsymbol{q}^k\hat{\boldsymbol{X}}^{(0)}=\\&\sum_{j=0}^{k-1}d^j\left[(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}\right]^{j+1}\boldsymbol{A}'\boldsymbol{L}+d^k\left[(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}\right]^k\hat{\boldsymbol{X}}_{\mathrm{LS}}=\\&\left\{\sum_{j=0}^{k-1}d^j\left[(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}\right]^{j+1}\boldsymbol{A}'\boldsymbol{A}+d^k\left[(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}\right]^k\right\}\hat{\boldsymbol{X}}_{\mathrm{LS}}\end{aligned} \tag{14}$$

式中,$k=1,2,\cdots$。

1.2　迭代公式的解析表达式

矩阵 $\boldsymbol{A}'\boldsymbol{A}$ 为实对称正定阵,故存在正交矩阵 $\boldsymbol{Q}$(即 $\boldsymbol{Q}'\boldsymbol{Q}=\boldsymbol{I}$)使得 $\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}'$,这里 $\boldsymbol{\Lambda}=\mathrm{diag}\{s_1,s_2,\cdots,s_n\}$是对角矩阵,且 $s_1\geqslant s_2\geqslant\cdots\geqslant s_n>0$为矩阵 $\boldsymbol{A}'\boldsymbol{A}$ 的 n 个特征值。$\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I}=\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}'+\boldsymbol{Q}\boldsymbol{I}\boldsymbol{Q}'=\boldsymbol{Q}(\boldsymbol{\Lambda}+\boldsymbol{I})\boldsymbol{Q}'$,则:

$$\boldsymbol{q}=(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}=\boldsymbol{Q}(\boldsymbol{\Lambda}+\boldsymbol{I})^{-1}\boldsymbol{Q}'=\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}'$$

式中,

$$\boldsymbol{D}=(\boldsymbol{\Lambda}+\boldsymbol{I})^{-1}=\mathrm{diag}\{1/(s_1+1),\cdots,1/(s_n+1)\}$$

进一步有：

$$d^j[(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}]^{j+1}\boldsymbol{A}'\boldsymbol{A}=$$
$$d^j(\boldsymbol{QDQ}'\cdot\boldsymbol{QDQ}'\cdot\cdots\cdot\boldsymbol{QDQ}')\boldsymbol{Q\Lambda Q}'=$$
$$\boldsymbol{Q}(d^j\boldsymbol{D}^{j+1}\boldsymbol{\Lambda})\boldsymbol{Q}' \tag{15}$$

将式(15)代入式(14)，并整理得：

$$\hat{\boldsymbol{X}}^{(k)}=[\boldsymbol{Q}(\sum_{j=0}^{k-1}d^j\boldsymbol{D}^{j+1}\boldsymbol{\Lambda}+d^k\boldsymbol{D}^k)\boldsymbol{Q}']\hat{\boldsymbol{X}}_{\mathrm{LS}}=$$
$$\boldsymbol{QD}_{k,d}\boldsymbol{Q}'\hat{\boldsymbol{X}}_{\mathrm{LS}} \tag{16}$$

式中，$\boldsymbol{D}_{k,d}=\sum_{j=0}^{k-1}d^j\boldsymbol{D}^{j+1}\boldsymbol{\Lambda}+d^k\boldsymbol{D}^k$，为一个 n 阶对角矩阵，第 i 个对角元素为：

$$d_i=\sum_{j=0}^{k-1}(\frac{d}{s_i+1})^j\cdot\frac{s_i}{s_i+1}+(\frac{d}{s_i+1})^k \tag{17}$$

式中，$i=1,\cdots,n$，$k\in N^+$，N^+ 表示正整数。

1.3 迭代公式的收敛性

将式(12)与(11)相减，得：

$$\hat{\boldsymbol{X}}^{(2)}-\hat{\boldsymbol{X}}^{(1)}=d\boldsymbol{q}(\hat{\boldsymbol{X}}^{(1)}-\hat{\boldsymbol{X}}^{(0)}) \tag{18}$$

将式(13)与(12)相减，得：

$$\hat{\boldsymbol{X}}^{(3)}-\hat{\boldsymbol{X}}^{(2)}=d\boldsymbol{q}(\hat{\boldsymbol{X}}^{(2)}-\hat{\boldsymbol{X}}^{(1)}) \tag{19}$$

将式(18)代入式(19)，得：

$$\hat{\boldsymbol{X}}^{(3)}-\hat{\boldsymbol{X}}^{(2)}=d^2\boldsymbol{q}^2(\hat{\boldsymbol{X}}^{(1)}-\hat{\boldsymbol{X}}^{(0)}) \tag{20}$$

以此类推，得：

$$\hat{\boldsymbol{X}}^{(k+1)}-\hat{\boldsymbol{X}}^{(k)}=d^k\boldsymbol{q}^k(\hat{\boldsymbol{X}}^{(1)}-\hat{\boldsymbol{X}}^{(0)})=$$
$$d^k[(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}]^k(\hat{\boldsymbol{X}}^{(1)}-\hat{\boldsymbol{X}}^{(0)}) \tag{21}$$

定理1[20] 设 n 阶任意方阵 $\boldsymbol{M}$，则方阵 $\boldsymbol{M}$ 幂收敛于零矩阵的充要条件为 $\boldsymbol{M}$ 的所有特征值的模都小于1。

设 $s(s>0)$ 为方阵 $\boldsymbol{A}'\boldsymbol{A}$ 的某一特征值，则矩阵 $d(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}$ 相应的特征值为 $\frac{d}{s+1}$。所以，若 $-1\leqslant d\leqslant 1$，则有 $\left|\frac{d}{s+1}\right|<1$。根据定理1可知，当 $k\to+\infty$，$d^k[(\boldsymbol{A}'\boldsymbol{A}+\boldsymbol{I})^{-1}]^k\to 0$，从而 $\hat{\boldsymbol{X}}^{(k+1)}-\hat{\boldsymbol{X}}^{(k)}\to 0$，即迭代公式(14)在 $-1\leqslant d\leqslant 1$ 是收敛的。

2 修正因子 d 的确定

在迭代公式(14)中，若 $k=1$ 则为Liu估计。为此，式(14)中 d 的确定方法依据Liu估计中 d 的确定思路。

将模型(1)改写为典则形式：

$$\boldsymbol{L}=\boldsymbol{BY}+\boldsymbol{\Delta} \tag{22}$$

式中，$\boldsymbol{B}=\boldsymbol{AQ}$；$\boldsymbol{Y}=\boldsymbol{Q}'\boldsymbol{X}$。则有：

$$\boldsymbol{B}'\boldsymbol{B}=\boldsymbol{Q}'\boldsymbol{A}'\boldsymbol{AQ}=\boldsymbol{Q}'\boldsymbol{Q\Lambda Q}'\boldsymbol{Q}=\boldsymbol{\Lambda} \tag{23}$$

则模型(22)的最小二乘估值为 $\hat{\boldsymbol{Y}}_{\mathrm{LS}}=\boldsymbol{\Lambda}^{-1}\boldsymbol{B}'\boldsymbol{L}$，岭估值为 $\hat{\boldsymbol{Y}}_\lambda=(\boldsymbol{\Lambda}+\lambda\boldsymbol{I})^{-1}\boldsymbol{B}'\boldsymbol{L}$，Liu估值为：

$$\hat{\boldsymbol{Y}}_d=(\boldsymbol{\Lambda}+\boldsymbol{I})^{-1}(\boldsymbol{B}'\boldsymbol{L}+d\hat{\boldsymbol{Y}}_{\mathrm{LS}})=$$
$$(\boldsymbol{\Lambda}+\boldsymbol{I})^{-1}(\boldsymbol{\Lambda}+d\boldsymbol{I})\hat{\boldsymbol{Y}}_{\mathrm{LS}} \tag{24}$$

则有如下的线性关系[21]：

$$\hat{\boldsymbol{X}}_d=\boldsymbol{Q}\hat{\boldsymbol{Y}}_d,\ \mathrm{MSE}(\hat{\boldsymbol{X}}_d)=\mathrm{MSE}(\hat{\boldsymbol{Y}}_d)$$

由式(24)知：

$$E(\hat{\boldsymbol{Y}}_d)=(\boldsymbol{\Lambda}+\boldsymbol{I})^{-1}(\boldsymbol{\Lambda}+d\boldsymbol{I})E(\hat{\boldsymbol{Y}}_{\mathrm{LS}})=$$
$$(\boldsymbol{\Lambda}+\boldsymbol{I})^{-1}(\boldsymbol{\Lambda}+d\boldsymbol{I})\boldsymbol{Y} \tag{25}$$

$$\mathrm{cov}(\hat{\boldsymbol{Y}}_d)=\sigma^2(\boldsymbol{\Lambda}+\boldsymbol{I})^{-1}(\boldsymbol{\Lambda}+d\boldsymbol{I})\boldsymbol{\Lambda}^{-1}\cdot$$
$$(\boldsymbol{\Lambda}+d\boldsymbol{I})(\boldsymbol{\Lambda}+\boldsymbol{I})^{-1} \tag{26}$$

所以，Liu估值 $\hat{\boldsymbol{Y}}_d$ 的均方误差为：

$$\mathrm{MSE}(\hat{\boldsymbol{Y}}_d)=\|E(\hat{\boldsymbol{Y}}_d)-\boldsymbol{Y}\|^2+\mathrm{trcov}(\hat{\boldsymbol{Y}}_d)=$$
$$\sum_{i=1}^{n}\frac{(d-1)^2y_i^{\ 2}}{(s_i+1)^2}+\sigma^2\sum_{i=1}^{n}\frac{(s_i+d)^2}{s_i(s_i+1)^2}=f(d) \tag{27}$$

所以，

$$f'(d)=2(d-1)\sum_{i=1}^{n}\frac{y_i^{\ 2}}{(s_i+1)^2}+$$
$$2\sigma^2\sum_{i=1}^{n}\frac{s_i+d}{s_i(s_i+1)^2}$$

且

$$f'(1)=2\sigma^2\sum_{i=1}^{n}\frac{1}{s_i(s_i+1)}>0 \tag{28}$$

因此一定存在 $0<d<1$，满足 $\mathrm{MSE}(\hat{\boldsymbol{Y}}_d)<\mathrm{MSE}(\hat{\boldsymbol{Y}}_{\mathrm{LS}})$。使MSE达到最小的是使 $f'(d)=0$ 的解，解得使MSE最小的修正因子为：

$$d_{\mathrm{opt}}=\frac{\sum_{i=1}^{n}(y_i^{\ 2}-\sigma^2)/(s_i+1)^2}{\sum_{i=1}^{n}(s_iy_i^{\ 2}+\sigma^2)/s_i(s_i+1)^2} \tag{29}$$

式中，σ^2、$y_i(i=1,\cdots,n)$ 未知，可以用LS估计法或者岭估计法得到的 $\hat{\sigma}^2$、$\hat{\boldsymbol{Y}}$ 替代。

3 实验分析

算例1 将 11×9 阶的Hilbert矩阵作为线性模型 $\boldsymbol{AX}=\boldsymbol{L}$ 的观测矩阵 $\boldsymbol{A}$，其元素 $a_{ij}=\frac{1}{i+j-1}$，其中 $i=1,2,\cdots,11$；$j=1,2,\cdots,9$。设未知参数的真值均为1，则没有误差的常数项为：

$\boldsymbol{L}=$[2.829 0 1.929 0 1.519 9 1.269 9 1.096 8 0.968 2 0.868 2 0.787 9 0.721 7

0.666 1　0.618 8]′。

根据式（3），求出法矩阵的条件数为 $8.985\ 9\times10^{16}$，属于严重病态问题。在具体的处理过程中，对 $\boldsymbol{L}$ 的每个分量分别添加不同的扰动误差 ε（ε=0.1，ε=0.01，ε=0.001）。分别用 LS 估计、岭估计（参数的确定选用 L-曲线法）、Liu 估计和本文的迭代估计公式进行计算，比较其估值 $\hat{\boldsymbol{X}}$、$\hat{\boldsymbol{X}}$ 的均方误差以及 $\hat{\boldsymbol{X}}$ 与 $\boldsymbol{X}$ 差的范数 $\|\Delta\hat{\boldsymbol{X}}\|=\|\hat{\boldsymbol{X}}-\boldsymbol{X}\|$ 的大小，其计算结果见表 1～3。

表 1　ε=0.1 时不同方法所得结果

Tab.1　Results of Different Methods when ε=0.1

较值	真值	LS 估计	岭估计	Liu 估计	本文方法
$\hat{\boldsymbol{X}}$	1	1.606 5	1.489 1	1.050 6	1.050 6
	1	−16.160 0	−5.304 3	1.053 7	1.053 9
	1	69.522 0	12.243 6	1.050 9	1.050 2
	1	294.290 0	12.176 4	1.049 1	1.046 1
	1	−2 082.400 0	−6.846 8	1.021 0	1.042 3
	1	3 765.700 0	−19.525 6	1.077 7	1.039 1
	1	−1 717.100 0	−16.402 1	1.018 6	1.036 2
	1	−1 428.500 0	2.268 8	1.019 1	1.033 8
	1	1 128.000 0	32.849 2	1.043 2	1.031 6
$\|\Delta\hat{\boldsymbol{X}}\|$	0	$4.986\ 9\times10^{3}$	45.749 3	0.139 6	0.129 9
MSE($\hat{\boldsymbol{X}}$)	0	$2.486\ 9\times10^{7}$	20.930	0.019 48	0.016 8

表 2　ε=0.01 时不同方法所得结果

Tab.2　Results of Different Methods when ε=0.01

较值	真值	LS 估计	岭估计	Liu 估计	本文方法
$\hat{\boldsymbol{X}}$	1	1.060 7	1.048 9	1.005 1	1.005 1
	1	−7.160 3	−3.695 7	1.005 4	1.005 4
	1	7.852 2	2.124 4	1.005 1	1.005 0
	1	30.329 0	2.117 6	1.004 9	1.004 6
	1	−207.340 0	2.153 2	1.002 1	1.004 2
	1	377.470 0	−1.052 6	1.007 8	1.003 9
	1	−170.810 0	−7.402 1	1.001 9	1.003 6
	1	−141.950 0	1.126 9	1.001 9	1.003 4
	1	113.700 0	4.184 9	1.004 3	1.003 2
$\|\Delta\hat{\boldsymbol{X}}\|$	0	$4.986\ 9\times10^{2}$	4.574 9	0.013 97	0.012 99
MSE($\hat{\boldsymbol{X}}$)	0	$2.486\ 9\times10^{5}$	2.093 0	$1.951\ 1\times10^{-4}$	$1.686\ 8\times10^{-4}$

表 3　ε=0.001 时不同方法所得结果

Tab.3　Results of Different Methods when ε=0.001

较值	真值	LS 估计	岭估计	Liu 估计	本文方法
$\hat{\boldsymbol{X}}$	1	1.006 1	1.004 9	1.000 5	1.000 5
	1	−0.828 4	0.936 9	1.000 5	1.000 5
	1	1.685 2	1.002 4	1.000 5	1.000 5
	1	3.932 9	1.111 8	1.000 5	1.000 5
	1	−1.983 4	0.921 5	1.000 2	1.000 4
	1	3.864 7	0.794 7	1.000 8	1.000 4
	1	−1.618 1	0.825 9	1.000 2	1.000 4
	1	−1.329 5	1.012 7	1.000 2	1.000 3
	1	1.227 0	1.318 5	1.000 4	1.000 3
$\|\Delta\hat{\boldsymbol{X}}\|$	0	4.986 9	0.457 5	$1.406\ 5\times10^{-3}$	$1.298\ 8\times10^{-3}$
MSE($\hat{\boldsymbol{X}}$)	0	$2.486\ 9\times10^{3}$	0.209 30	$1.978\ 2\times10^{-6}$	$1.686\ 8\times10^{-6}$

从表 1～3 可以得出如下结论：

1）在法矩阵严重病态的情况下，LS 的解与真值偏差很大，该估计法已不适应，且偏差随着观测量扰动误差大小的变化而变化，扰动越大，偏差越大。

2）在不同扰动误差的情况下，岭估计有效改善了 LS 估计，但在添加扰动 ε=0.1 时，岭估值与真值的偏差依然很大。所以岭估计虽然能改善法

方程病态时的解算效果，但在扰动误差比较大时其求解效果仍不理想。

3)在加入不同程度的干扰后，Liu估计法和迭代估计法所得结果均与真值偏差不大，都是可靠的。干扰越小，偏差就越小，估值越接近于真值，但本文所提的迭代法在精度上略高于Liu估计法。

4)在该算例中，3种情况的迭代次数均为2次，这是因为Liu估计就是它的第一次迭代，精度已经很高，所以迭代完第二次后，MSE几乎不发生变化，故在这里只取了迭代两次的结果进行对比。

算例2 取自文献[22]中第47页的例子。设 $P_1, P_2, \cdots, P_9$ 为9个已知点，其坐标列于表4。9个已知点到两个未知点 P_{10}, P_{11}（假设模拟真值分别为(0,0,0)和(7,10,−5)）的观测距离也列于表4。两个未知点之间的观测距离为 $d_{10,11}$ = 13.107 85 m。要求根据19个观测距离确定两个未知点的坐标，其中各距离为等精度观测，中误差为±0.000 1 m。

表4 控制点的坐标和观测距离

Tab.4 Coordinate and Observation Distance of Control Points

点号	坐标/m			观测距离/m	
	x	y	z	$d_{i,10}$	$d_{i,11}$
P_1	23.000	10.000	0.010	25.078 69	16.765 17
P_2	10.000	9.990	0.000	14.134 51	17.719 65
P_3	35.000	10.010	−0.010	36.415 88	28.442 94
P_4	100.000	19.990	0.005	101.479 43	93.168 39
P_5	−36.000	10.005	0.000	37.364 22	43.299 05
P_6	0.000	10.010	−0.005	10.010 04	8.600 60
P_7	56.000	9.995	0.010	56.996 06	49.256 18
P_8	−15.000	10.105	−0.010	18.035 90	22.559 66
P_9	−1.7000	10.008	−0.015	10.150 63	10.043 82

该算例中法方程系数阵的条件数为 $\text{cond}(\boldsymbol{A}'\cdot\boldsymbol{A}) = 4.601\ 5\times10^3$，呈病态。分别用LS估计、岭估计、Liu估计和本文的迭代估计进行计算，比较计算的估值 $\hat{\boldsymbol{X}}$、$\hat{\boldsymbol{X}}$ 的均方误差以及 $\hat{\boldsymbol{X}}$ 与 $\boldsymbol{X}$ 差范数 $\|\Delta\hat{\boldsymbol{X}}\| = \|\hat{\boldsymbol{X}}-\boldsymbol{X}\|$ 的大小，计算结果见表5。

表5 LS估计、岭估计、Liu估计以及本文方法的结果比较

Tab.5 Comparison of the Results Among LS Estimator、Ridge Estimator、Liu Estimator and Our Method

较值	真值	LS估计	岭估计	Liu估计	本文方法
$\hat{\boldsymbol{X}}$	0	−0.036 8	−0.041 1	−0.038 0	−0.038 4
	0	0.051 8	0.027 5	0.022 7	0.022 0
	0	9.363 1	0.086 0	0.375 9	0.055 9
	7	7.048 8	6.974 6	6.975 9	6.973 3
	10	5.596 3	10.036 3	9.866 2	10.019 6
	−5	−4.648 2	−4.873 5	−4.904 4	−4.911 7
参数	0	0	0.482 7	0.137 5	0.137 5
$\|\Delta\hat{\boldsymbol{X}}\|$	0	10.353 3	0.166 7	0.214 2	0.118 3
MSE($\hat{\boldsymbol{X}}$)	0	18.281 3	0.219 6	0.170 9	0.014 0

表5中的“参数”是指在不同估计方法中所需要确定的参数，如岭估计法中指的是岭参数 λ，Liu估计法和本文所提的方法中指的是修正因子 d。

从表5可知，岭估计、Liu估计和本文所提出的迭代方法所得结果与真值比较接近，均提高了LS解的精度。进一步对比估值的第三个分量，迭代估计法得到的值比Liu估计法得到的值更接近于真值，说明迭代估计法在病态问题求解中比Liu估计更可靠准确。此次实验中，在采用Liu估计的修正因子 d 后，迭代法所得的MSE和偏差均小于LS估计、岭估计和Liu估计，表明新方法对病态方程的求解具有一定的有效性，所得结果具有可信度。

4 结语

在GPS数据处理、工程控制网平差、形变观测分析、大地测量反演等测量数据处理领域，系统的病态性问题是常见的，并且病态性的危害作用

非常严重。在病态问题的求解中,统计学家们提出了许多有偏估计方法。本文在 Liu 估计或 Sakallioglu 提出的有偏估计法的基础上,引入迭代的思想提出了一种新的有偏迭代估计法。若初值选为 LS 估值,则第一次迭代式就是 Liu 估计;若初值选为岭估计,则第一次迭代就是 Sakallioglu 提出的估计法。针对迭代公式,证明了修正因子在[−1,1]范围内的收敛性。借助矩阵的谱分解给出迭代公式的解析形式,并给出修正因子的选取公式。通过两个不同情况例子的实验结果来看,迭代估计法优于 LS 估计和岭估计,求解精度略高于 Liu 估计。还需要说明的是,本文的两个例子均采用初值是 LS 估值。事实上不管初值选为 LS 估值还是岭估计值,最后迭代公式总是达到同一个精度。

本文所提的有偏迭代估计法受修正因子 $d\in[-1,1]$的限制,不能使该方法所得到的估值达到更高的精度。故如何拓宽修正因子的取值范围以及每次迭代步骤中是否可以采用不同的最优的修正因子都是下一步需要研究的问题。

参　考　文　献

[1] Stein C. Inadmissibility of the Usual Estimator for Mean of Multivariate Normal Distribution[C]. 3th Berkley Symposium on Mathematical and Statistics Probability, Berkley,1956

[2] Malthouse E C. Shrinkage Estimation and Direct Marketing Scoring Model[J]. *Journal of Interactive Marketing*,1999,13(4):10-23

[3] Massy W F. Principal Components Regression in Exploratory Statistical Research[J]. *Journal of the American Statistical Association*, 1965, 60(309): 234-266

[4] Hoerl A E, Kennard R W. Ridge Regression: Biased Estimation for Non-Orthogonal Problems [J]. *Technometrics*,1970, 12 (1):55-67

[5] Hoerl A E, Kennard R W. Ridge Regression: Application for Non-Orthogonal Problems[J]. *Technometrics*,1970,12(1):69-82

[6] Guo Jie, Gui Qingming, Guo Shumei, et al. Truncation Ridge Estimation Based on the Biased Corrected Theory by Multicollinearity Diagnosis[J]. *Geomatics and Information Science of Wuhan University*,2015,40(6):785-789(郭杰,归庆明,郭淑妹,等. 利用复共线性诊断确定偏差矫正项的截断型岭估计[J]. 武汉大学学报·信息科学版,2015,40(6):785-789)

[7] Liu Kejian. A New Class of Biased Estimate in Linear Regression[J]. *Communications in Statistics-Theory and Methods*,1993,22(2):393-402

[8] Liu Kejian. Using Liu-Type Estimator to Combat Collinearity[J]. *Communications in Statistics-Theory and Methods*, 2003,32(5):1 009-1 020

[9] Mansson K. Developing a Liu Estimator for the Negative Binomial Regression Model: Method and Application[J]. *Journal of Statistical Computation and Simulation*, 2013,83(9):1 773-1 780

[10] Wang Xinzhou, Liu Dingyou, Zhang Qianyong, et al. The Iteration by Correcting Characteristic Value and Its Application in Surveying Data Processing [J]. *Journal of Heilongjiang Institute of Technology*, 2001,15(2):3-6(王新洲,刘丁酉,张前勇,等. 谱修正迭代法及其在测量数据处理中的应用[J]. 黑龙江工程学院学报,2001,15(2):3-6)

[11] Liu Bin, Gong Jianya, Jiang Wanshou, et al. Improvement of the Iteration by Correcting Characteristic Value Based on Ridge Estimation and Its Application in RPC Calculating[J]. *Geomatics and Information Science of Wuhan University*, 2012,37(4):399-402(刘斌,龚健雅,江万寿,等. 基于岭参数的谱修正迭代法及其在有理多项式参数求解中的应用[J]. 武汉大学学报·信息科学版,2012,37(4):399-402)

[12] Guo Qiuying, Hu Zhenqi. Application of Genetic Algorithm to Solve Ill-Conditioned Equations for GPS Rapid Positioning[J]. *Geomatics and Information Science of Wuhan University*, 2009, 34 (2): 240-243(郭秋英,胡振琪. 遗传算法在 GPS 快速定位病态方程解算中的应用[J]. 武汉大学学报·信息科学版,2009,34(2):240-243)

[13] Chen Xiru, Wang Guisong. Modern Regression Analysis: Theory and Application [M]. Hefei: Anhui Education Press,1987(陈希孺,王松桂. 近代回归分析:原理方法及应用[M]. 合肥:安徽教育出版社,1987)

[14] Hansen P C, O'Leary D P. The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems [J]. *SIAM Journal on Scientific Computing*, 1993,14(6):1 487-1 503

[15] Wang Zhenjie, Ou Jikun. Determining the Ridge Parameter in a Ridge Estimation Using L-Curve Method [J]. *Geomatics and Information Science of Wuhan University*, 2004, 29 (3):235-238(王振杰,欧吉坤. 用 L-曲线法确定岭估计中的岭参数[J]. 武汉大学学报·信息科学版,2004,29(3):235-238)

[16] Wang Jiaying. Inverse Theory in Geophysics[M]. Beijing: Higher Education Press, 2002(王家映. 地球物理反演理论[M]. 北京:高等教育出版社,2002)

[17] Huang Weibin. Modern Adjustment Theory and Its

Application[M]. Beijing: People's Liberation Army Press,1992(黄维彬. 近代平差理论及其应用[M]. 北京:解放军出版社,1992)

[18] Sakallioglu S, Kaciranlar S. A New Biased Estimator Basedon Ridge Estimation[J]. *Statistics Papers*, 2008,49:669-689

[19] Sakallioglu S, Kaciranlar S, Akdeniz F. Mean Squared Error Comparisons of Some Biased Regression Estimators[J]. *Communications in Statistics: Theory and Methods*, 2001,30(2):347-361

[20] Liu Dingyou. Matrix Analysis [M]. Wuhan: Wuhan University Press,2006(刘丁酉. 矩阵分析[M]. 武汉:武汉大学出版社,2006)

[21] Golam K B M. Applications of Some Improved Estimators in Linear Regression[J]. *Journal of Modern Applied Statistical Methods*, 2006, 5(2): 367-380

[22] Guo Jianfeng. Diagnostics and Processing of Ill-Posed in Surveying Adjustment System[D]. Zhengzhou: Information Engineering University,2002(郭建峰. 测量平差系统病态性的诊断与处理[D]. 郑州:信息工程大学,2002)

A New Iterative Estimator Method Based on Liu Estimator for Ill-Posed Equations

JIANG Zhaoying[1,2] *LIU Guolin*[1] *YU Shengwen*[1]

1 College of Geomatics, Shandong University of Science and Technology, Qingdao 266590, China

2 College of Science and Information, Qingdao Agricultural University, Qingdao 266109, China

Abstract: In the presence of the design matrix's collinearity (which is equivalent to ill-conditioning) in the linear regression model, the least squares (LS) estimator has large variances and its solution is rather unstable, so the LS estimator is not the precise estimation any more. In order to weaken the ill-conditioning, many biased estimator methods are introduced, such as ridge estimator, the principal components estimator, the Liuestimator and so on. In this paper, based on the famous Liuestimator, we present a new biased estimator which is called a biased iterative estimator method. With the aid of spectral decomposition of the symmetric and positive matrix, the iterative formula is converted to a simple analytical expression conveniently for calculating. And the iterative formula is proved to be convergent in the condition of modified parameterd$\in[-1,1]$. Following the deter mination method of modified parameter in the Liuestimator, we give a formula of the optimal modified parameter to minimize the mean squared error (MSE). Finally, we use the proposed biased iterative estimator, LS estimator, ridge estimator and the Liuestimator to calculate two numerical examples and give their experimental results. In the first example, we respectively add different perturbations to the observation vector. The simulation results show that compared with other three methods, the biased iterative estimator is more stable under the perturbation. Comparison results of the second example show that our new biased iterative estimator is more closed to the real value, that is superior, in the mean squared error sense, to the LS estimator, ridge estimator and the Liuestimator.

Key words: multicollinearity; condition number; Liu estimator; ridge estimator; mean squared error

First author: JIANG Zhaoying, PhD, specializes in InSAR data processing and model algorithm. E-mail: 15964287012@163.com

Corresponding author: LIU Guolin, PhD, professor. E-mail: gliu@sdust.edu.cn

Foundation support: The National Natural Science Foundation of China, Nos. 41274007, 41404003; Shandong Province Natural Science Foundation of China, No. ZR2012DM001; Specialized Research Fund for the Doctoral Program of Higher Education, No. 20123718110001; Shandong Taishan Scholar Construction Project Under Special Funding.