第32卷第6期 大地测量与地球动力学 Vol. 32 No. 6
2012年12月 JOURNAL OF GEODESY AND GEODYNAMICS Dec. ,2012

文章编号:1671-5942(2012)06-0111-04

抗差总体最小二乘方法*

陈玮娴[1)] 袁 庆[2)]

(1)武汉市城乡建设委员会,武汉 430023
2)中铁第四勘察设计集团有限公司,武汉 430063)

摘 要 针对变量中的误差模型的观测向量中含有粗差的情况,提出了抗差总体最小二乘方法。根据M估计方法,用增长较缓慢的极小化观测向量残差函数代替观测向量残差平方和项,并推导出基于选权迭代法的抗差总体最小二乘估计计算公式。采用小角度坐标转换算例进行计算分析,证明了该方法的可行性。

关键词 变量中的误差模型;稳健估计;总体最小二乘;抗差总体最小二乘;坐标变换

中图分类号:P207 **文献标识码:**A

A ROBUST TOTAL LEAST-SQUARES METHOD

Chen Weixian[1)] and Yuan Qing[2)]

(1) *Wuhan Urban & Rural Construction Committee*, *Wuhan* 430023
2) *China Railway Siyuan Survey and Design Group Co.*, *Ltd*, *Wuhan* 430063)

Abstract Aiming at the situation that there is single gross error in the observation vector of the Errors-in- variables(EIV) model, we propose the robust total least-squares method. On the basis of the generalized maximum likelihood estimation, we use the observation vector residual function which grows more slowly instead of observation vector squared residuals items, and have deduced robust total least-squares estimation calculation formula with the iteration method with variable weights. Finally, we apply the new method to the coordinate transformation parameters determination, and verify the feasibilities of the new method.

Key words: Errors-in-variables(EIV) model; robust estimation; total least-squares method; robust total least-squares method; coordinate transformation

1 引言

在应用最小二乘进行参数估计时,系数矩阵 $\boldsymbol{A}$ 的数值是不受随机误差影响的,而全部误差都限制在观测向量阵 $\boldsymbol{y}$ 中,此时我们建立高斯-马尔科夫(G-M)模型,采用最小二乘(LS)方法进行参数估计。若观测向量中含有粗差,则采用抗差最小二乘(抗差LS)法[1],通过选权迭代法进行稳健估计,使得含有粗差的奇异观测的权函数接近于零,从而达到剔除粗差的目的[2]。当观测向量 $\boldsymbol{y}$ 和系数矩阵 $\boldsymbol{A}$ 都受到随机误差影响时,通常建立相应的EIV(Errors-In-Variables)模型[3],采用总体最小二乘(TLS)方法进行参数估计。若此时观测向量中含有粗差,若采用TLS方法则会使模型歪曲造成参数估计严重失实。鉴于此,针对在EIV模型中观测向量 $\boldsymbol{y}$ 含有个别粗差的情况,我们提出抗差总体最小二乘的方法(抗差TLS)。根据M估计的思想,用增长较缓慢的极小化观测向量残差函数与系数矩阵残差平方

* **收稿日期:**2012-05-16
基金项目:国家自然科学基金(41074017)
作者简介:陈玮娴,女,1987年生,硕士,研究方向:大地测量与工程测量数据处理. E-mail:15531422@qq.com

和的函数,代替观测向量残差平方和与系数矩阵残差平方和的函数,采用 Lagrange 极值法推导出基于选权迭代法的抗差总体最小二乘估计模型。经小角度坐标转换的算例分析,证明该方法可以有效地剔除粗差。

2 数学模型

许多场合存在着各种因素,如测量误差、人为因素、建模偏差和仪器误差等等,都会降低系数矩阵元素的数值精度,即不仅在观测数据中含有随机误差,而且系数矩阵中的数值也含有随机误差,因此建立

$$\boldsymbol{y}-\boldsymbol{e}_y=(\boldsymbol{A}-\boldsymbol{E}_A)\boldsymbol{\xi} \tag{1}$$

的 EIV 模型。式中,$\boldsymbol{y}\in\boldsymbol{R}^{n\times1}$ 为观测向量,$\boldsymbol{A}\in\boldsymbol{R}^{n\times m}$ 为系数矩阵,$\boldsymbol{\xi}\in\boldsymbol{R}^{m\times1}$ 为待估参数,$\boldsymbol{e}_y\in\boldsymbol{R}^{n\times1}$ 为观测向量的随机误差向量,$\boldsymbol{E}_A\in\boldsymbol{R}^{n\times m}$ 为系数矩阵的随机误差矩阵。相应的随机模型如下:

$$\begin{bmatrix}\boldsymbol{e}_y\\ \boldsymbol{e}_A\end{bmatrix}=\begin{bmatrix}\boldsymbol{e}_y\\ \mathrm{vec}\boldsymbol{E}_A\end{bmatrix}\sim\left(\begin{bmatrix}0\\0\end{bmatrix},\sigma_0^2\begin{bmatrix}P_y^{-1} & 0\\ 0 & P_A^{-1}\end{bmatrix}\right) \tag{2}$$

式中,“vec”为矩阵列向量化算子[4],$\boldsymbol{P}_y\in\boldsymbol{R}^{n\times n}$ 和 $\boldsymbol{P}_A\in\boldsymbol{R}^{nm\times nm}$ 分别为观测向量和列向量化的系数矩阵的权阵,通常为单位阵。根据文献[5]的思想把 EIV 模型看作是非线性的,将式(1)线性化展开,去掉二次 $\boldsymbol{E}_A\boldsymbol{d\xi}$ 小项得:

$$-\boldsymbol{e}_y+\boldsymbol{B}\boldsymbol{e}_A=\boldsymbol{A}\boldsymbol{\delta\xi}-\boldsymbol{\delta y} \tag{3}$$

式中,$\boldsymbol{\xi}=\boldsymbol{\xi}^0+\boldsymbol{\delta\xi}$,$\boldsymbol{\delta y}=\boldsymbol{y}-\boldsymbol{A}\boldsymbol{\xi}^0$,$\boldsymbol{B}=(\boldsymbol{\xi}^{0\mathrm{T}}\otimes\boldsymbol{I}_n)$,$\boldsymbol{I}_n\in\boldsymbol{R}^{n\times n}$ 为单位阵(“$\otimes$”为 kronecker 积:$\boldsymbol{M}\otimes\boldsymbol{N}[m_{ij}\cdot N]$,$\boldsymbol{M}=[m_{ij}]$)[4]。相应的总体最小二乘估计准则为:

$$\boldsymbol{e}_y^{\mathrm{T}}\boldsymbol{P}_y\boldsymbol{e}_y+\boldsymbol{e}_A^{\mathrm{T}}\boldsymbol{P}_A\boldsymbol{e}_A=\min \tag{4}$$

由于观测向量 $\boldsymbol{y}$ 中含有粗差,若采用上述估计准则,个别异常大观测值残差的出现将会导致观测值残差平方和 $\boldsymbol{e}_y^{\mathrm{T}}\boldsymbol{P}_y\boldsymbol{e}_y$ 这一项迅速增大,为了达到平方和极小的目的,估值必然要迁就那些异常值,因此采用个别异常值会对整个估值产生大的影响。为了避免这种情况,根据 $\boldsymbol{M}$ 估计思想,用增长较慢的极小化观测向量残差函数 $\sum_{i=1}^{n}\rho(e_{yi})$ 代替 $\boldsymbol{e}_y^{\mathrm{T}}\boldsymbol{P}_y\boldsymbol{e}_y$,得到抗差总体最小二乘的估计准则如下:

$$\sum_{i=1}^{n}\rho(e_{yi})+\boldsymbol{e}_A^{\mathrm{T}}\boldsymbol{P}_A\boldsymbol{e}_A=\min \tag{5}$$

式中,$\rho(e_{yi})$ 为 M 估计函数,采用 Lagrange 法求稳健 TLS 公式,$\boldsymbol{\lambda}\in\boldsymbol{R}^{n\times1}$ 为 Lagrange 乘子,建立 Lagrange 目标函数:

$$\boldsymbol{\Phi}(e_y,e_A,\lambda,\delta\xi)=\sum_{i=1}^{n}\rho(e_{yi})+\boldsymbol{e}_A^{\mathrm{T}}\boldsymbol{P}_A\boldsymbol{e}_A+2\boldsymbol{\lambda}^{\mathrm{T}}(\mathrm{\delta y}-\mathrm{A\delta\xi}-\mathrm{e_y}+\mathrm{Be_A}) \tag{6}$$

对式(6)式求 $\boldsymbol{e}_y$ 偏导数,并令其为零,以求出其极值点得:

$$\sum_{i=1}^{n}[\rho'(e_{yi})-\boldsymbol{\lambda}_i]=0 \tag{7}$$

写成如下形式:

$$\sum_{i=1}^{n}\left[\frac{\rho'(e_{yi})}{e_{yi}}e_{yi}-\boldsymbol{\lambda}_i\right]=0 \tag{8}$$

令式(8)中的 $\frac{\rho'(e_{yi})}{e_{yi}}$ 为权函数 $p_i(e_{yi})$,即:

$$\frac{\rho'(e_{yi})}{e_{yi}}=p_i(e_{yi}) \tag{9}$$

则式(9)成为:

$$\sum_{i=1}^{n}[\rho_i'(e_{yi})e_{yi}-\boldsymbol{\lambda}_i]=0 \tag{10}$$

并将其以矩阵的形式表示:

$$\boldsymbol{P}(e_y)\boldsymbol{e}_y=\boldsymbol{\lambda} \tag{11}$$

式中,$\boldsymbol{P}(e_y)$ 为权函数矩阵,为观测向量残差的函数,$p_i(e_y)$ 为 $\boldsymbol{P}(e_y)$ 的第 i 个对角元素,$\boldsymbol{P}(e_y)=\mathrm{diag}(p_1(e_y)\quad p_2(e_y)\ \cdots\ p_n(e_y))$。

分别对式(6)求 e_λ、λ、$\delta\xi$ 偏导数,并令其为零,求出其极值点得:

$$\boldsymbol{P}_A\boldsymbol{e}_A+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}=0 \tag{12}$$

$$\boldsymbol{\delta y}-\boldsymbol{A}\boldsymbol{\delta\xi}-\boldsymbol{e}_y+\boldsymbol{B}\boldsymbol{e}_A=0 \tag{13}$$

$$-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}=0 \tag{14}$$

由式(11)、(12)得:

$$\boldsymbol{e}_y=\boldsymbol{P}(e_y)^{-1}\boldsymbol{\lambda} \tag{15}$$

$$\boldsymbol{e}_A=-\boldsymbol{P}_A^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda} \tag{16}$$

将式(15)、(16)代入(13)得:

$$\boldsymbol{\delta y}-\boldsymbol{A}\delta\xi-\boldsymbol{P}(e_y)^{-1}\boldsymbol{\lambda}-\boldsymbol{B}\boldsymbol{P}_A^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}=0 \tag{17}$$

由式(17)得:

$$\boldsymbol{\lambda}=(\boldsymbol{P}(e_y)^{-1}+\boldsymbol{B}\boldsymbol{P}_A^{-1}\boldsymbol{B}^{\mathrm{T}})^{-1}(\boldsymbol{\delta y}-\boldsymbol{A}\delta\xi) \tag{18}$$

将式(17)代入式(14)得法方程为:

$$\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{P}(e_y)^{-1}+\boldsymbol{B}\boldsymbol{P}_A^{-1}\boldsymbol{B}^{\mathrm{T}})^{-1}\boldsymbol{A}\delta\xi=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{P}(e_y)^{-1}+\boldsymbol{B}\boldsymbol{P}_A^{-1}\boldsymbol{B}^{\mathrm{T}})^{-1}\boldsymbol{\delta y} \tag{19}$$

因此抗差总体最小二乘的估计准则为:

$$\boldsymbol{e}_y^{\mathrm{T}}\boldsymbol{P}(e_y)\boldsymbol{e}_y+\boldsymbol{e}_A^{\mathrm{T}}\boldsymbol{P}_A\boldsymbol{e}_A=\min \tag{20}$$

选取相应的权函数[6]为其中 $\sigma=\sqrt{\boldsymbol{e}_y^{\mathrm{T}}\boldsymbol{P}(e_y)\boldsymbol{e}_y/(n-m)}$:

$$p_i(e_{yi})=\begin{cases}1 & |e_{yi}|\leqslant 2\sigma\\ c/|e_{yi}| & |e_{yi}|>2\sigma\end{cases} \tag{21}$$

3 计算步骤与算例分析

3.1 计算步骤

下角标(i)为第 i 步迭代计算,根据文献[5]的迭代方法,每次迭代中需更新系数矩阵 $\boldsymbol{A}$,即 $\boldsymbol{A}_{(i)}=\boldsymbol{A}-\boldsymbol{E}_{A(i)}$,为了表达方便,令 $\boldsymbol{Q}_{1(i)}=\boldsymbol{P}(e_{y(i)})^{-1}+\boldsymbol{B}_{(i)}$

$Q_A B^{T}_{(i)}$。得到相应的计算步骤如下:

1)设置初值 $E_{A(0)}=0, P(e_{y(i)})=I_n, \xi_{(0)}=\xi^0=A^{T}P(e_{y(0)})A)^{-1}(A^{T}P(e_{y(0)})y$;

2)从 $i=0$ 开始,计算 $A_{(i)}=A-E_{A(i)}, B_{(i)}=\xi^{T}_{(i)}\otimes I_n, \delta y_{(i)}=y-A\xi_{(i)}, Q_{1(i)}=P(e_{y(i)})^{-1}+B_{(i)}Q_A B^{T}_{(i)}$,由式(12)得 $\delta\hat{\xi}_{(i+1)}=A^{T}_{(i)}Q^{-1}_{1(i)}A_{(i)})^{-1}(A^{T}_{(i)}(Q^{-1}_{1(i)})(\delta y_{(i)}-E_{A(i)}\xi_{(i)})$,并计算 $\xi_{(i+1)}=\xi_{(i)}+\delta\hat{\xi}_{(i+1)}, e_{y(i+1)}=Q_y Q^{-1}_{1(i)}(\delta y_{(i)}-A_{(i)}\delta\hat{\xi}_{(i+1)}), e_{A(i+1)}=-Q_A Q^{-1}_{(i)}(\delta y_{(i)}-A_{(i)}\delta\hat{\xi}_{(i+1)})$,由式(14)得 $P(e_{y(i+1)})$;

3)重复 2)直到 $\|\delta\hat{\xi}_{(i+1)}\|<\delta_0$,$\delta_0$ 为给定阈值,并计算单位权中误差近似估值:

$$\hat{\sigma}_0=\sqrt{(e_y^{T}P(e_y)e_y+e_A^{T}P_A e_A)/(n-m)} \quad (15)$$

3.2　算例分析

采用基于线性的布尔沙模型的小角度空间直角坐标转换为例进行算法的验证,数据为实测数据(表 1),其中#1 点是加入粗差后的原坐标系 1 号点坐标值。采用 LS、抗差 LS、TLS 和抗差 TLS 方法计算加入粗差后的坐标转换参数,并与未加入粗差时采用 LS 和 TLS 方法计算的转换参数加以比较,结果见表 2,其中 X_0、Y_0、Z_0 为平移参数,μ 为尺度参数,ε_X、ε_Y、ε_Z 为旋转参数。表 3 为加入粗差后 LS、抗差 LS、TLS 和抗差 TLS 4 种方法计算得到的观测向量和系数矩阵的残差值。

由表 2 可以看出,当两套坐标均含有随机误差而未含粗差时,TLS 方法较 LS 方法得到较小的单位权中误差,即在这种情况下 TLS 方法较 LS 方法可

表 1　WGS-84 坐标系和地方坐标系下公共点坐标(单位:m)

Tab. 1　Coordinates under the WGS-84 system and local system(unit:m)

点号	WGS-84 坐标系坐标			点号	地方系坐标		
	X	Y	Z		X	Y	Z
1	−2 802 088.418 2	5 009 123.156 9	2 772 386.569 9	1	−2 802 191.348 2	5 009 064.765 7	2 772 381.176 8
2	−2 810 072.730 9	5 016 144.572 1	2 751 960.500 7	2	−2 810 175.651 5	5 016 086.112 0	2 751 955.053 1
3	−2 820 464.394 2	5 009 963.991 7	2 752 475.921 1	3	−2 820 567.227 2	5 009 905.344 4	2 752 470.385 8
4	−2 817 059.835 1	5 002 512.999 3	2 769 304.628 8	4	−2 817 162.653 8	5 002 454.352 0	2 769 299.110 6
5	−2 841 060.290 2	4 982 041.618 3	2 781 470.180 5	5	−2 841 162.870 7	4 981 982.504 1	2 781 464.448 9
#1	−2 802 070.418 2						

表 2　参数计算结果和精度比较

Tab. 2　Comparison between the estimated parameters and accuracies calculated with different methods

	未加入粗差		加入粗差 LS	
	LS	TLS	LS	抗差 TLS
X_0/m	−3.824 577 9	−3.824 577 9	−3.824 577 9	−3.824 577 9
Y_0/m	7.351 878 4	7.351 879 0	7.351 880 5	7.351 880 5
Z_0/m	3.864 378 3	3.864 380 5	3.864 380 9	3.864 380 9
μ	1.000 000 6	1.000 000 6	1.000 027 9	1.000 004 6
ε_X/rad	−0.000 004 2	−0.000 004 2	−0.000 049 1	−0.000 010 8
ε_Y/rad	0.000 002 4	0.000 002 4	0.000 042 8	0.000 008 3
ε_Z/rad	−0.000 019 3	−0.000 019 3	−0.000 010 5	−0.000 018 0
σ_0/rad	0.084 343 9	0.059 640 1	3.918 798 2	1.499 725 6

表 3　加入粗差的坐标转换后坐标较差比较(单位:m)

Tab. 3　Comparison between the coordinate differences after transformation by adding the gross error(unit:m)

点号	TLS 方法坐标较差			抗差 TLS 方法坐标较差		
	ΔX	ΔY	ΔZ	ΔX	ΔY	ΔZ
1	−109.385 14	−0.411 40	−5.474 04	−7.362 12	−0.489 14	−0.780 95
2	25.495 33	8.258 44	−9.729 60	5.840 47	2.204 10	−1.995 71
3	23.556 31	6.089 49	−4.127 26	5.201 94	1.806 36	−1.008 97
4	30.400 66	−1.609 47	0.970 90	6.269 36	−0.505 92	0.257 89
5	29.989 74	−12.025 83	18.387 01	5.402 68	−2.965 23	3.543 67

(下转第 128 页)