

基于中位参数法相关观测的抗差加权整体最小二乘算法

刘春阳¹ 王 坚¹ 王 彬² 刘 超³ 刘纪平⁴

1 中国矿业大学国土环境与灾害监测国家测绘地理信息局重点实验室,江苏 徐州,221116
2 武汉大学测绘学院,湖北 武汉,430079
3 安徽理工大学测绘学院,安徽 淮南,232001
4 中国测绘科学研究院,北京,100830

摘 要:在抗差加权整体最小二乘算法中,抗差模型的抗差性与初值的好坏关系极大,若以最小二乘或整体最小二乘估值作为初值,必定会受到粗差污染而影响其抗差性。考虑到观测向量和系数矩阵存在相关性,首先推导了部分变量误差(partial errors-in-variables, Partial EIV)模型的加权整体最小二乘算法,在此基础上提出了一种利用中位参数法求解抗差迭代初值的相关观测抗差加权整体最小二乘算法。然后采用中位参数法确定抗差初值,考虑到可能出现的粗差对观测空间与结构空间的综合影响,基于标准化残差构造权因子函数,实现其抗差解法。仿真实验结果表明,此算法具有良好的抗差性能,其参数估计结果比传统算法精度更高,且随着粗差个数的增加,其抗差稳定性较好。

关键词:抗差加权整体最小二乘;中位参数法;Partial EIV 模型;相关观测;标准化残差
中图分类号:P207 **文献标志码:**A

整体最小二乘(total least squares, TLS)估计可以有效解决变量误差(errors in variables, EIV)模型中系数矩阵含有随机误差的问题,国内外学者对此展开了广泛的研究,产生了大量研究成果。文献[1]通过扩展最小二乘(least squares, LS)估计,最早提出了将观测向量和系数矩阵的残差平方和最小化的 TLS 估计准则;文献[2]提出了著名的奇异值分解法,并正式命名了 TLS 估计方法的通用术语,也使得 TLS 估计引起了各领域的广泛关注;文献[3]提出了附加等式约束的 TLS 估计,自此, TLS 估计方法才真正应用到大地测量领域。近年来,国内外学者针对加权整体最小二乘(weighted total least squares, WTLS)算法进行了深入的研究。文献[4]提出了基于拉格朗日函数的 WTLS 迭代算法;文献[5]基于 Gauss-Helmert 模型研究了平面坐标转换的 WTLS 估计;文献[6]提出了非线性条件下 WTLS 的高斯-牛顿迭代法;文献[7-9]研究了一般性权矩阵、非线性、系数矩阵部分元素为固定量

等情形下的 WTLS 算法;文献[10]从非线性 LS 入手,将 WTLS 算法扩展到部分变量误差(Partial EIV)模型中,推导了有限样本下 WTLS 算法的精度评定公式,可以将各种形式下的系数矩阵纳入统一的模型下求解,且有效减少了未知数的数目。

上述讨论仅考虑了观测向量和系数矩阵中观测值包含的随机误差,在实际测量中观测值往往会受到粗差污染,而 LS 和 WTLS 估计并不具备抵御粗差的能力,因此,有必要进一步研究抗差加权整体最小二乘算法(robust weighted total least squares, RWTLS)。在 LS 估计中,粗差的处理方法一般可归纳为两类:(1)基于均值漂移模型,把粗差归入函数模型,研究粗差探测方法^[11],(2)基于方差-协方差膨胀模型,将粗差归入随机模型,研究抗差估计方法^[12-14]。在 TLS 估计中,国内外学者进行了一些关于 RWTLS 的探讨,文献[15]基于 Gauss-Helmert 模型和 Huber 权函数提出了三维坐标转换的稳健 WTLS 方法;文献[16]提出了 WTLS 的选权迭代算法,并用于直线拟合和

坐标转换中;文献[17]将文献[15]中的 Huber 权函数替换成了 IGG 函数,并验证了该方法的优势;文献[18-19]基于 Gauss-Helmert 模型研究了三维坐标转换和 GPS 高程拟合中的 RWTLS 算法;文献[20]认识到在上述关于 RWTLS 的研究中均采用残差构造权因子函数,无法顾及结构空间的抗差性,对于加权情况下的抗差效果不理想,因此基于标准化残差提出了一种 RWTLS 算法,并推导了残差协因数阵的表达式。然而,上述算法均存在共同的缺陷:(1)相关算法未考虑观测向量和系数矩阵的相关性;(2)由于等价权抗差估计的抗差性与初值好坏的关系极大,而抗差迭代前的初值往往采用 LS 和 TLS 估计的结果,受到粗差的影响而容易失真。文献[21]提出采用基于中位数法的抗差总体最小二乘算法,但最终只用到少量观测数据计算参数估值,丢失大量的有效信息,按文中分组选取子样并不十分合理,且抗差权因子函数由残差构造,并不能顾及结构空间抗差。

鉴于上述分析,本文首先结合 Partial EIV 模型和标准化残差各自的优势,基于高斯-牛顿迭代法推导了基于 Partial EIV 模型相关观测的 RWTLS 算法,然后利用中位数法求解抗差迭代的初值,代入 RWTLS 算法中重新定权,进行参数求解。

1 相关观测的 Partial EIV 模型的高斯-牛顿迭代解法

观测向量和系数矩阵完全相关的 Partial EIV 模型的函数模型和随机模型可表示为:

$$\left. \begin{aligned} y - v_y &= (X^T \otimes I_m)(h + B(a - v_a)) \\ v &= \begin{bmatrix} v_y \\ v_a \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_0^2 Q \right) \end{aligned} \right\} \quad (1)$$

式中, y 和 v_y 分别表示 $m \times 1$ 维观测向量及其对应的随机误差向量; X 为 t 维参数向量; h 为非随机元素构成的 mt 维向量; B 为由常数构成的 $mt \times n$ 维系数矩阵; a 为随机元素对应的 n 维列向量; v_a 为 a 的随机误差向量;随机模型中 σ_0 为单位权中误差; Q 为 v 对应的协因数阵,

$$Q = \begin{bmatrix} Q_y & Q_{ya} \\ Q_{ay} & Q_a \end{bmatrix} \quad (2)$$

式中, Q_y 和 Q_a 分别表示 v_y 和 v_a 对应的协因数阵; Q_{ya} 和 Q_{ay} 分别表示互协因数阵。

将式(1)中的函数模型在 $(X_{(i)}, v_{a(i)})$ 处进行一阶泰勒级数展开可得:

$$y - v_y = AX_{(i)} + A_{(i)}\delta X - (X_{(i)}^T \otimes I_m)Bv_a \quad (3)$$

将式(3)整理可得:

$$y = AX_{(i)} + A_{(i)}\delta X + [I_m - (X_{(i)}^T \otimes I_m)B]v \quad (4)$$

式中,

$$A = \text{ivec}(h + Ba) \quad (5)$$

$$A_{(i)} = \text{ivec}(h + B(a - v_{a(i)})) = A - \text{ivec}(Bv_{a(i)}) \quad (6)$$

式中, ivec 表示将列向量恢复为指定维数矩阵的算子。

构造 WTLS 的拉格朗日目标函数:

$$\Phi(v, X, K) = v^T Q^{-1} v + 2K^T (y - AX_{(i)} - A_{(i)}\delta X - [I_m - (X_{(i)}^T \otimes I_m)B]v) \quad (7)$$

对各变量求偏导并令偏导数为零:

$$\frac{1}{2} \frac{\partial \Phi}{\partial v} = Q^{-1} \hat{v} - [I_m - (X_{(i)}^T \otimes I_m)B]^T \hat{K} = 0 \quad (8)$$

$$\frac{1}{2} \frac{\partial \Phi}{\partial \delta X} = -A_{(i)}^T \hat{K} = 0 \quad (9)$$

$$\begin{aligned} \frac{1}{2} \frac{\partial \Phi}{\partial K} &= y - AX_{(i)} - A_{(i)}\delta \hat{X} - \\ &[I_m - (X_{(i)}^T \otimes I_m)B]\hat{v} = 0 \end{aligned} \quad (10)$$

由(8)可得:

$$\hat{v} = Q [I_m - (X_{(i)}^T \otimes I_m)B]^T \hat{K} \quad (11)$$

将式(11)代入式(10)可得:

$$\begin{aligned} \hat{K} &= (MQM^T)^{-1} (L - AX_{(i)} - A_{(i)}\delta \hat{X}) = \\ &Q_{c(i)}^{-1} (L - AX_{(i)} - A_{(i)}\delta \hat{X}) \end{aligned} \quad (12)$$

式中,

$$\begin{aligned} M &= [I_m - (X_{(i)}^T \otimes I_m)B] \\ Q_{c(i)} &= [I_m - (X_{(i)}^T \otimes I_m)B]Q[I_m - \\ &(X_{(i)}^T \otimes I_m)B]^T \end{aligned} \quad (13)$$

将式(12)代入式(9)可以推导出:

$$\delta \hat{X}_{(i+1)} = (A_{(i)}^T Q_{c(i)}^{-1} A_{(i)})^{-1} \cdot A_{(i)}^T Q_{c(i)}^{-1} (y - AX_{(i)}) \quad (14)$$

则在 $(i+1)$ 次迭代后更新的参数估值为:

$$\begin{aligned} \hat{X}_{(i+1)} &= X_{(i)} + \delta \hat{X}_{(i+1)} = (A_{(i)}^T Q_{c(i)}^{-1} A_{(i)})^{-1} A_{(i)}^T Q_{c(i)}^{-1} \cdot \\ &(y - (X_{(i)}^T \otimes I_m)Bv_{a(i)}) \end{aligned} \quad (15)$$

$$\begin{aligned} \hat{v}_{(i+1)} &= Q^{-1} [I_m - (X_{(i)}^T \otimes I_m)B]^T \cdot \\ &Q_{c(i)}^{-1} (L - AX_{(i)} - A_{(i)}\delta \hat{X}_{(i+1)}) \end{aligned} \quad (16)$$

$$\begin{aligned} \hat{v}_{y(i+1)} &= (Q_y + Q_{ya} - (X_{(i)}^T \otimes I_m)B)^T \cdot \\ &Q_{c(i)}^{-1} (L - AX_{(i)} - A_{(i)}\delta \hat{X}_{(i+1)}) \end{aligned} \quad (17)$$

$$\begin{aligned} \hat{v}_{a(i+1)} &= (Q_{ay} + Q_a - (X_{(i)}^T \otimes I_m)B)^T \cdot \\ &Q_{c(i)}^{-1} (L - AX_{(i)} - A_{(i)}\delta \hat{X}_{(i+1)}) \end{aligned} \quad (18)$$

将最小二乘的参数估值作为初值,按照式(14)至式(18)进行迭代计算,当满足 $\|\delta \hat{X}_{(i+1)}\| < \theta$ (θ 为给定小量)时,停止迭代,此时可以获得 Partial EIV 的 WTLS 解。

当观测向量和系数矩阵不相关时, $\boldsymbol{Q}_{\boldsymbol{y}\boldsymbol{a}} = \boldsymbol{0}$, $\boldsymbol{Q}_{\boldsymbol{a}\boldsymbol{y}} = \boldsymbol{0}$, 式(17)和式(18)可变为:

$$\hat{\boldsymbol{v}}_{\boldsymbol{y}(i+1)} = \boldsymbol{Q}_{\boldsymbol{y}} \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)} - \boldsymbol{A}_{(i)} \delta \hat{\boldsymbol{X}}_{(i+1)}) \tag{19}$$
$$\hat{\boldsymbol{v}}_{\boldsymbol{a}(i+1)} = -\boldsymbol{Q}_{\boldsymbol{a}} \boldsymbol{B}^T (\boldsymbol{X}_{(i)} \otimes \boldsymbol{I}_m) \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)} - \boldsymbol{A}_{(i)} \delta \hat{\boldsymbol{X}}_{(i+1)}) \tag{20}$$

2 基于中位参数法的 RWTLS 算法

2.1 基于标准化残差的降权策略

为了同时顾及观测向量和结构空间抗差, 在利用等价权抗差原理进行参数的抗差估计时, 需要构建检验量, 其关键在于残差向量的协因数阵推导。根据式(17)和式(18)得到第 $(i+1)$ 次迭代的预测残差向量:

$$\boldsymbol{v}_{i+1} = \begin{bmatrix} \hat{\boldsymbol{v}}_{\boldsymbol{y}(i+1)} \\ \hat{\boldsymbol{v}}_{\boldsymbol{a}(i+1)} \end{bmatrix} = \begin{bmatrix} (\boldsymbol{Q}_{\boldsymbol{y}} + \boldsymbol{Q}_{\boldsymbol{y}\boldsymbol{a}} - (\boldsymbol{X}_{(i)}^T \otimes \boldsymbol{I}_m) \boldsymbol{B})^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \\ (\boldsymbol{Q}_{\boldsymbol{a}\boldsymbol{y}} + \boldsymbol{Q}_{\boldsymbol{a}} - (\boldsymbol{X}_{(i)}^T \otimes \boldsymbol{I}_m) \boldsymbol{B})^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \end{bmatrix} \cdot (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)} - \boldsymbol{A}_{(i)} \delta \hat{\boldsymbol{X}}_{(i+1)}) = \begin{bmatrix} \boldsymbol{\omega}_i \\ \boldsymbol{\psi}_i \end{bmatrix} \boldsymbol{\varphi}_{i+1} \tag{21}$$

式中, \boldsymbol{v}_{i+1} 为第 $(i+1)$ 次迭代的残差向量; $\hat{\boldsymbol{v}}_{\boldsymbol{y}(i+1)}$ 和 $\hat{\boldsymbol{v}}_{\boldsymbol{a}(i+1)}$ 分别表示观测向量和随机元素向量对应的第 $(i+1)$ 次迭代的残差向量; $\boldsymbol{\omega}_i = (\boldsymbol{Q}_{\boldsymbol{y}} + \boldsymbol{Q}_{\boldsymbol{y}\boldsymbol{a}} - (\boldsymbol{X}_{(i)}^T \otimes \boldsymbol{I}_m) \boldsymbol{B})^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1}$; $\boldsymbol{\psi}_i = (\boldsymbol{Q}_{\boldsymbol{a}\boldsymbol{y}} + \boldsymbol{Q}_{\boldsymbol{a}} - (\boldsymbol{X}_{(i)}^T \otimes \boldsymbol{I}_m) \boldsymbol{B})^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1}$; $\boldsymbol{\varphi}_{i+1} = \boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)} - \boldsymbol{A}_{(i)} \delta \hat{\boldsymbol{X}}_{(i+1)}$ 。

则第 $(i+1)$ 次迭代的预测残差向量的协因数阵为:

$$\boldsymbol{Q}_{\boldsymbol{v}_{i+1}} = \begin{bmatrix} \boldsymbol{\omega}_i \\ \boldsymbol{\psi}_i \end{bmatrix} \cdot \boldsymbol{Q}_{\boldsymbol{\varphi}_{i+1}} \cdot [\boldsymbol{\omega}_i^T \quad \boldsymbol{\psi}_i^T] \tag{22}$$

将式(14)代入 $\boldsymbol{\varphi}_{i+1}$ 得:

$$\boldsymbol{\varphi}_{i+1} = \boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)} - \boldsymbol{A}_{(i)} \delta \hat{\boldsymbol{X}}_{(i+1)} = \boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)} - \boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \cdot \boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)}) = (\boldsymbol{I}_m - \boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1}) \cdot (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)}) \tag{23}$$

根据协因数传播率可以求得^[20]:

$$\boldsymbol{Q}_{\boldsymbol{\varphi}_{i+1}} = (\boldsymbol{I}_m - \boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1}) \cdot \boldsymbol{Q}_{(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)})} (\boldsymbol{I}_m - \boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1})^T = (\boldsymbol{I}_m - \boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1}) \cdot \boldsymbol{Q}_{\boldsymbol{c}(i)} (\boldsymbol{I}_m - \boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1})^T = \boldsymbol{Q}_{\boldsymbol{c}(i)} - \boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T \tag{24}$$

将式(24)代入式(22)可得:

$$\boldsymbol{Q}_{\boldsymbol{v}_{i+1}} = \begin{bmatrix} \boldsymbol{\omega}_i \\ \boldsymbol{\psi}_i \end{bmatrix} \cdot \boldsymbol{Q}_{\boldsymbol{\varphi}_{i+1}} \cdot [\boldsymbol{\omega}_i^T \quad \boldsymbol{\psi}_i^T] = \begin{bmatrix} \boldsymbol{\omega}_i \\ \boldsymbol{\psi}_i \end{bmatrix} (\boldsymbol{Q}_{\boldsymbol{c}(i)} - \boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1}) \begin{bmatrix} \boldsymbol{\omega}_i^T \\ \boldsymbol{\psi}_i^T \end{bmatrix}$$

$$\boldsymbol{A}_{(i)} (\boldsymbol{A}_{(i)}^T \boldsymbol{Q}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T) \cdot [\boldsymbol{\omega}_i^T \quad \boldsymbol{\psi}_i^T] \tag{25}$$

由于上述推导获得的残差协因数阵与迭代过程中的参数和系数阵残差有关, 故残差协因数阵需要在迭代中不断更新。

在此基础上, 对式(21)的残差进行标准化:

$$\hat{v}_i = \frac{v_i}{\sigma_0 (\sqrt{q_{v_i}})} \quad (q_{v_i} \neq 0) \tag{26}$$

式中, \hat{v}_i 为标准化残差向量 $\hat{\boldsymbol{v}}$ 中的第 i 个元素; v_i 表示残差向量 \boldsymbol{v} 中的第 i 个元素; σ_0 为单位权中误差; q_{v_i} 为残差协因数阵 $\boldsymbol{Q}_{\boldsymbol{v}}$ 中对角线上第 i 个元素。 σ_0 由中位函数计算, 可表示为:

$$\sigma_0 = \underset{i=1}{\overset{t}{\text{med}}} (|v_i| / \sqrt{q_{v_i}}) \cdot 1.4826 \quad (q_{v_i} \neq 0) \tag{27}$$

基于标准化残差进行抗差估计, 可选的权因子函数众多, 本文采用 IGG 3 权因子函数, 为了计算方便, 采用其对应的协因数因子函数:

$$R_{\hat{v}_i} = \begin{cases} 1.0, & |\hat{v}_i| \leq k_0 \\ \frac{|\hat{v}_i|}{k_0} \left(\frac{k_1 - k_0}{k_1 - |\hat{v}_i|} \right)^2, & k_0 < |\hat{v}_i| \leq k_1 \\ 10^{10}, & k_1 < |\hat{v}_i| \end{cases} \tag{28}$$

式中, 第 3 段理论上应该为无穷大, 为了方便计算, 可采用极大值替代 (如 1×10^{10}), 满足实际需求; k_0 取为 $2.0 \sim 3.0$, k_1 取为 $4.0 \sim 8.0$ ^[20]。

采用双因子模型, 由式(28)可得观测向量和系数矩阵中观测值的等价协因数:

$$\bar{q}_{ij} = q_{ij} \sqrt{R_{\hat{v}_i}} \sqrt{R_{\hat{v}_j}} \tag{29}$$

式中, \bar{q}_{ij} 和 q_{ij} 分别为观测值的等价协因数阵 $\bar{\boldsymbol{Q}}$ 和协因数阵 \boldsymbol{Q} 中的对应位置元素。采用 $\bar{\boldsymbol{Q}}$ 替代 \boldsymbol{Q} , 构建类似于式(7)的加权整体最小二乘问题的拉格朗日目标函数, 对各变量求导并令导数为零, 类似于式(8)至式(10), 可得第 $(i+1)$ 次迭代的待求参数抗差估计的改正数为:

$$\delta \hat{\boldsymbol{X}}_{(i+1)} = (\boldsymbol{A}_{(i)}^T \bar{\boldsymbol{Q}}_{\boldsymbol{c}(i)}^{-1} \boldsymbol{A}_{(i)})^{-1} \boldsymbol{A}_{(i)}^T \bar{\boldsymbol{Q}}_{\boldsymbol{c}(i)}^{-1} \cdot (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{X}_{(i)}) \tag{30}$$

式中,

$$\bar{\boldsymbol{Q}}_{\boldsymbol{c}(i)} = [\boldsymbol{I}_m - (\boldsymbol{X}_{(i)}^T \otimes \boldsymbol{I}_m) \boldsymbol{B} \bar{\boldsymbol{Q}} \cdot [\boldsymbol{I}_m - (\boldsymbol{X}_{(i)}^T \otimes \boldsymbol{I}_m) \boldsymbol{B}]^T$$

采用类似于式(14)至式(18)的迭代过程, 在给定的阈值条件下, 可获得参数的最终估值及其精度评定指标。

2.2 基于中位参数法确定抗差初值

文献[21]中基于中位数法的思路为:(1)采用中位数法求解分组后的中位数参数估值, 并用其计算观测向量与系数矩阵中含有的误差向量的估值;(2)依据绝对误差中位数与中误差的关系, 分

类确定观测向量与系数矩阵中的观测元素对应的中误差,并应用抗差权函数对观测元素进行重新分类定权;(3)根据重新定义的权值分组求解参数估值,并求取其参数的中位数值,计算各组模型参数的估值与中位参数的差值向量,以各组中差值向量的最小范数对应的参数估值作为参数的最终估值。

然而,文献[21]方法在解算最终参数估值的过程中使用的是分组的观测数据,即最终只用到少量观测数据计算参数估值,损失了大量的有效信息。而本文只采用中位数法确定抗差迭代前的初始值,在抗差迭代过程中使用了全部的观测数据,理论上更加合理。

式(1)函数模型中有 $m \times 1$ 维观测向量,待估参数 $\hat{\mathbf{X}}$ 为 t 维向量($t < m$),因此,采用相关观测的 Partial EIV 模型可以获得 C_m 组参数估值,可表示为:

$$\hat{\mathbf{X}} = [\hat{\mathbf{X}}_1 \ \hat{\mathbf{X}}_2 \ \cdots \ \hat{\mathbf{X}}_{C_m^t}] = \begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_{C_m^t}^1 \\ x_1^2 & x_2^2 & \cdots & x_{C_m^t}^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^t & x_2^t & \cdots & x_{C_m^t}^t \end{bmatrix}$$

(31)

式中, x 表示待估参数元素。

对式(31)中获得的 C_m^t 组参数估值(行向量)求中位数,可得到待估参数 $\hat{\mathbf{X}}$ 的 t 个参数的中位数值,表示为:

$$\hat{\mathbf{X}}_{\text{med}} = [x_{\text{med}}^1 \ x_{\text{med}}^2 \ \cdots \ x_{\text{med}}^t]^T$$

(32)

式中,下标 med 表示中位数值。

计算 C_m^t 组估值与中位参数向量的差值,并求二次范数 $\|d_{\hat{\mathbf{X}}_i}\|$,可表示为:

$$\|d_{\hat{\mathbf{X}}_i}\| = \|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}_{\text{med}}\|, i \in [1, 2 \cdots C_m^t]$$

(33)

式中, $\|d_{\hat{\mathbf{X}}_i}\|$ 的最小值对应的参数估值 $\hat{\mathbf{X}}_i$ 为所求的中位参数解,即抗差迭代前的初值,根据式(17)和式(18)可以分别求出观测向量和系数矩阵对应的残差向量,再根据式(21)至式(29)进行重新定权,重新定权后根据标准化残差的降权策略(见 § 2.1)进行抗差估计即可。

应用中位参数法进行参数估值的求取,可以有效降低受观测值含有的粗差污染的影响。众所周知,单参数中位数法的崩溃污染率接近 50%,当一半以上的观测值正常时,其中位数也是正常的。因此,当观测值中含有 k 个粗差时,则参数的 C_m^t 组参数估值中,不含粗差的解向量为 $C_m^t - k$ 个。为保证中位参数估值的抗差性,需要满足 $C_m^t - k \geq C_m^t/2$,此时,可确保中位参数解由不含粗差的观

测值求得。

3 算例与分析

假设一直线方程为 $y = 5x + 9$,采用均匀分布函数在 $(0, 18)$ 之间选择 18 个 x 值,并计算对应 y 值,将对应的 18 个点位坐标 (x, y) 视为真值,并加入非等精度的随机误差,其数学期望为 0,标准差在 $0 \sim 0.05$ 间随机产生。考虑坐标点存在的相关性,设同一点位内 x 和 y 的相关系数 $\rho_{x_i y_i}$ 均为 0.6,点间的相关系数 $\rho_{x_i x_j}$ 和 $\rho_{y_i y_j}$ 均为 0.3,在实际计算中,通过 $\sigma_{x_i y_i} = \rho_{x_i y_i} \sigma_{x_i} \sigma_{y_i}$, $\sigma_{x_i x_j} = \rho_{x_i x_j} \sigma_{x_i} \sigma_{x_j}$, $\sigma_{y_i y_j} = \rho_{y_i y_j} \sigma_{y_i} \sigma_{y_j}$ 计算观测数据的协方差,完全相关观测情况下的协方差阵为(其中 $\sigma_0^2 = 1$):

$$\mathbf{D} = \begin{bmatrix} \frac{\sigma_0^2}{\sigma_{x_1}^2} & \cdots & \frac{\sigma_0^2}{\sigma_{x_1 x_{18}}} & \frac{\sigma_0^2}{\sigma_{x_1 y_1}} & \cdots & \frac{\sigma_0^2}{\sigma_{x_1 y_{18}}} \\ \vdots & & & & & \vdots \\ \frac{\sigma_0^2}{\sigma_{x_{18} x_1}} & \cdots & \frac{\sigma_0^2}{\sigma_{x_{18}}^2} & \frac{\sigma_0^2}{\sigma_{x_{18} y_1}} & \cdots & \frac{\sigma_0^2}{\sigma_{x_{18} y_{18}}} \\ \frac{\sigma_0^2}{\sigma_{y_1 x_1}} & \cdots & \frac{\sigma_0^2}{\sigma_{y_1 x_{18}}} & \frac{\sigma_0^2}{\sigma_{y_1}^2} & \cdots & \frac{\sigma_0^2}{\sigma_{y_1 y_{18}}} \\ \vdots & & & & & \vdots \\ \frac{\sigma_0^2}{\sigma_{y_{18} x_1}} & \cdots & \frac{\sigma_0^2}{\sigma_{y_{18} x_{18}}} & \frac{\sigma_0^2}{\sigma_{y_{18} y_1}} & \cdots & \frac{\sigma_0^2}{\sigma_{y_{18}}^2} \end{bmatrix}$$

(34)

加入的粗差个数分别为 1、2、3,粗差大小介于标准差绝对值的 5~20 倍之间,位置随机产生。反复进行 500 次模拟,方案设计如下。

- 1)方案 1:加入粗差前,使用 WTLS 估计。
- 2)方案 2:加入粗差后,使用 WTLS 估计。
- 3)方案 3:加入粗差后,使用基于标准化残差的 RWTLS 方法(文献[20])。
- 4)方案 4:加入粗差后,使用基于中位参数法的 RWTLS 方法。

上述方案中,方案 1 和方案 2 分别是在未加入粗差前和加入粗差后使用相关观测的 Partial EIV 模型进行求解;方案 3 是基于文献[20]的算法;方案 4 的算法是将文献[20]的算法推广到相关观测的情形,且在此基础上利用了中位参数法相关观测的 RWTLS 算法。

现有的 RWTLS 方法在进行抗差迭代前使用的初值一般采用 WTLS 的结果,而受到粗差污染的 WTLS 方法所求得的抗差迭代初值会严重失真。当粗差个数为 3 时,方案 3 和方案 4 抗差迭代初值的偏差序列分别如图 1 和图 2 所示,可以看出方案 4 中使用中位参数法求出的抗差迭代

初值明显优于方案 3 中使用 WTLS 方法的结果。

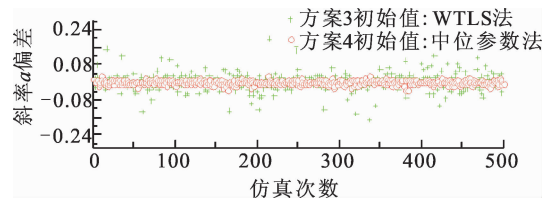


图 1 方案 3 和方案 4 抗差迭代初值(斜率 a) 偏差序列对比

Fig. 1 Deviation Sequence Comparison of Resistance Values (Slope a) Between Schemes 3 and 4

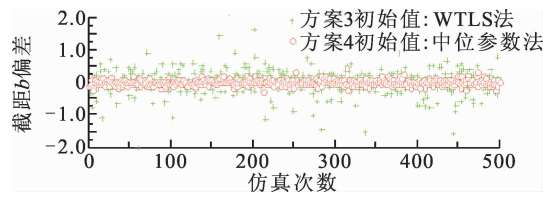


图 2 方案 3 和方案 4 抗差迭代初值(截距 b) 偏差序列对比

Fig. 2 Deviation Sequence Comparison of Resistance Values (Intercept b) Between Schemes 3 and 4

当粗差个数分别为 1、2、3 时,采用 4 种方案进行参数估计,分析不同方案的效果如下。

表 1 至表 3 分别给出了在不同粗差个数时各方案直线拟合斜率 a 与截距 b 的均方根误差 (root mean square error, RMSE) 和最大绝对偏差 (max) 的对比情况。

分析表 1 至表 3 得出以下结论:(1)由于方案 1 中没有混入粗差,所求得的均方根误差和最大绝对偏差均最小;(2)由于方案 2 中混入了粗差,均方根误差和最大绝对偏差均明显大于方案 1,说明混入粗差的观测值使得 WTLS 方法求得的参数估值严重失真,WTLS 方法缺乏抵御粗差的能力;(3)当粗差个数为 1 时,方案 3 采用了基于标准化残差的 RWTLS 方法,相较方案 2,其均方根误差和最大绝对偏差均显著减小,说明方案 3 在粗差个数较少时,由于同时顾及了观测空间和结构空间抗差,其模型具有较好的抵御粗差的能力;但随着粗差个数增加到 2 和 3 时,方案 3 中最大绝对偏差相较方案 2 并未发生显著变化,说明在考虑相关观测的情况下,方案 3 对于个别仿真数据,抗差会失效,整体抗差能力在减弱;(4)当粗差个数分别为 1、2、3 时,方案 4 的均方根误差相较于方案 3 均进一步减小,且随着粗差个数的增加,效果更加明显。当粗差个数为 3 时,方案 4 相比于方案 3,斜率 a 和截距 b 分别减小了约 32.70% 和 47.90%,抗差效果改善明显,尤其是最大绝对偏差显著减小,分别减小了约 67.16%

和 70.83%,说明方案 4 的抗差方案稳定性较好,相较其他方案,效果总体最佳。

图 3 和图 4 给出了当粗差个数为 3 时,方案 3 和方案 4 直线拟合斜率 a 与截距 b 的解算结果与真值的偏差序列。从图 3、图 4 中可以看出,方案 3 中个别偏差值较大,部分仿真数据抗差效果不佳,而方案 4 的拟合结果与真值偏差序列总体比方案 3 更稳定,即对于方案 3 抗差失效的数据,方案 4 仍能实现有效抗差估计,抗差性能稳定。

表 1 1 个粗差时各方案的统计结果

Tab. 1 Statistical Results of Each Scheme When There Exists One Gross Error

方案	RMSE (a)	RMSE (b)	max ($ a $)	max ($ b $)
1	0.009 6	0.104 0	0.008 1	0.512 1
2	0.022 8	0.213 9	0.155 8	1.444 7
3	0.010 1	0.111 1	0.073 3	0.635 2
4	0.009 8	0.107 2	0.072 4	0.629 9

表 2 2 个粗差时各方案的统计结果

Tab. 2 Statistical Results of Each Scheme When There Exists Two Gross Errors

方案	RMSE (a)	RMSE (b)	max ($ a $)	max ($ b $)
1	0.008 8	0.123 2	0.031 2	0.480 5
2	0.029 9	0.426 0	0.183 7	2.819 0
3	0.013 4	0.292 7	0.105 1	2.819 0
4	0.010 5	0.128 6	0.054 8	1.342 5

表 3 3 个粗差时各方案的统计结果

Tab. 3 Statistical Results of Each Scheme When There Exists Three Gross Errors

方案	RMSE (a)	RMSE (b)	max ($ a $)	max ($ b $)
1	0.008 9	0.118 3	0.027 8	0.349 7
2	0.037 9	0.428 4	0.210 4	2.749 9
3	0.021 1	0.309 8	0.197 6	2.729 5
4	0.014 2	0.161 4	0.064 9	0.796 2

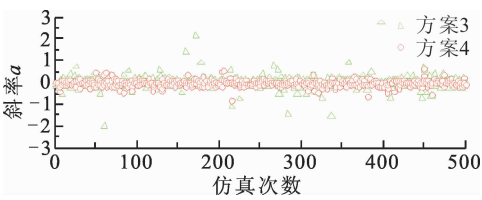


图 3 方案 3 和方案 4 拟合结果(斜率 a)对比
Fig. 3 Fitting Results (Slope a) Comparison Between Schemes 3 and 4

4 结 语

本文在推导相关观测 Partial-EIV 模型的 WTLS 算法的基础上,结合标准化残差提出了一种基于中位参数法求解抗差迭代初值的相关观测

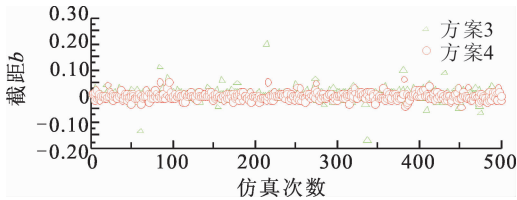


图 4 方案 3 和方案 4 拟合结果(截距 b)对比

Fig. 4 Fitting Results (Intercept b) Comparison Between Schemes 3 and 4

RWTLS 迭代算法。本文算法解决了抗差迭代初值受粗差污染会严重失真的问题,并顾及了观测空间和结构空间的抗差性,同时考虑了坐标点位的相关性。算例结果表明:(1)与基于标准化残差的 RWTLS 算法相比,本文算法得到的斜率与截距的均方根误差显著减小,算法的精度较现有的算法有一定的提高;(2)考虑到点位之间的相关性,通过比较基于标准化残差方案的 RWTLS 算法和本文算法,可以看出基于标准化残差的 RWTLS 算法虽然总体表现良好,但随着粗差个数的增加,对于个别仿真数据抗差失效,而本文算法仍表现出了良好的抗差性能,总体抗差效果更稳定。

参考文献

[1] Adcock R J. Note on the Method of Least Squares [J]. *The Analyst*, 1877, 4(6): 183-184

[2] Golub H G, Vanloan F C. An Analysis of the Total Least Squares Problem[J]. *SIAM Journal on Numerical Analysis*, 1980, 17(6): 883-893

[3] Schaffrin B, Felus Y A. On Total Least-Squares Adjustment with Constraints [J]. *IAG-Symp*, 2005, 128: 417-421

[4] Schaffrin B, Felus Y A. On the Multivariate Total Least-Squares Approach to Empirical Coordinate Transformations[J]. *Journal of Geodesy*, 2008, 82(6): 373-383

[5] Neitzel F. Generalization of Total Least-Squares on Example of Unweighted and Weighted 2D Similarity Transformation[J]. *Journal of Geodesy*, 2010, 84(12): 751-762

[6] Shen Y Z, Li B F, Chen Y. An Iterative Solution of Weighted Total Least-Squares Adjustment [J]. *Journal of Geodesy*, 2011, 85(4): 229-238

[7] Fang X. Weighted Total Least Squares Solutions for Applications in Geodesy[D]. Germany: Leibniz University Hannover, 2011

[8] Fang X. Weighted Total Least Squares; Necessary and Sufficient Conditions, Fixed and Random Pa-

rameters[J]. *Journal of Geodesy*, 2013, 87(8): 733-749

[9] Deng Caihua, Zhou Yongjun, Zhu Jianjun, et al. General Weighted Total Least Squares Method for a Type of New Functional Model [J]. *Journal of China University of Mining and Technology*, 2015, 44(5): 952-958 (邓才华, 周拥军, 朱建军, 等. 一类新函数模型及通用加权总体最小二乘平差方法 [J]. *中国矿业大学学报*, 2015, 44(5): 952-958)

[10] Xu P L, Liu J N, Shi C. Total Least Squares Adjustment in Partial Errors-in-Variables Models: Algorithm and Statistical Analysis[J]. *Journal of Geodesy*, 2012, 86(8): 661-675

[11] Ou Jikun. Quasi-Accurate Detection of Gross Errors (QUAD) [J]. *Acta Geodaetica et Cartographica Sinica*, 1999, 28(1): 15-20 (欧吉坤. 粗差的拟准检定法 (QUAD 法) [J]. *测绘学报*, 1999, 28(1): 15-20)

[12] Yang Y X. Robust Estimation of Geodetic Datum Transformation[J]. *Journal of Geodesy*, 1999, 73(5): 268-274

[13] Zhou Jiangwen, Huang Youcai, Yang Yuanxi, et al. Robust Least Squares Method [M]. Wuhan: Huazhong University of Science and Technology Press, 1997 (周江文, 黄幼才, 杨元喜, 等. 抗差估计最小二乘法 [M]. 武汉: 华中理工大学出版社, 1997)

[14] Gao Jingxiang, Zhang Huahai, Yu Xuexiang. Study on Robust Estimation Model for Coordinate Transformation of GPS Network in Mining Areas [J]. *Journal of China University of Mining and Technology*, 1999, 28(2): 99-103 (高井祥, 张华海, 余学祥. 矿区 GPS 网坐标转换的抗差模型 [J]. *中国矿业大学学报*, 1999, 28(2): 99-103)

[15] Chen Yi, Lu Jue. Performing 3D Similarity Transformation by Robust Total Least Squares[J]. *Acta Geodaetica et Cartographica Sinica*, 2012, 41(5): 715-722 (陈义, 陆珏. 以三维坐标转换为例解算稳健总体最小二乘方法 [J]. *测绘学报*, 2012, 41(5): 715-722)

[16] Mahboub V, Amiri-Simkooei A R, Sharifi M A. Iteratively Reweighted Total Least Squares: A Robust Estimation in Errors-in-Variables Models[J]. *Surv Rev*, 2013, 45(329): 92-99

[17] Gong Xunqiang, Li Zhilin. A Robust Weighted Total Least Squares Method [J]. *Acta Geodaetica et Cartographica Sinica*, 2014, 43(9): 888-894 (龚循强, 李志林. 稳健加权总体最小二乘法 [J]. *测绘学报*, 2014, 43(9): 888-894)

[18] Lu J, Chen Y, Li B, et al. Robust Total Least Squares with Reweighting Iteration for Three-Di-

mensional Similarity Transformation[J]. *Surv Rev*, 2014, 46(334): 28-36

[19] Tao Y Q, Gao J X, Yao Y F. TLS Algorithm for GPS Height Fitting Based on Robust Estimation [J]. *Surv Rev*, 2014, 46(336): 184-188

[20] Wang B, Li J C, Liu C. A Robust Weighted Total Least Squares Algorithm and Its Geodetic Applications[J]. *Studia Geophysica et Geodaetica*, 2016, 60(2): 177-194

[21] Tao Yeqing, Gao Jingxiang, Yao Yifei. Solution for Robust Total Least Squares Estimation Based on Median Method [J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(3): 297-301 (陶叶青, 高井祥, 姚一飞. 基于中位数法的抗差总体最小二乘估计[J]. *测绘学报*, 2016, 45(3): 297-301)

Robust Weight Total Least Squares Algorithm of Correlated Observation
Based on Median Parameter Method

LIU Chunyang¹ WANG Jian¹ WANG Bin² LIU Chao³ LIU Jiping⁴

- 1 Key Laboratory of Land Environment and Disaster Monitoring, NASG, China University of Mining and Technology, Xuzhou, 221116, China
- 2 School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
- 3 School of Geodesy and Geomatics, Anhui University of Science and Technology, Huainan 232001, China
- 4 Chinese Academy of Surveying and Mapping, Beijing 100830, China

Abstract: In the robust weighted total least squares(RWTLS) algorithm, its robustness of the robust model is highly related to the initial values. If the least squares or total least squares estimates is used as the initial value, it will be affected by gross error, and certainly impacted the robust characteristics of RWTLS estimates. Considering the correlation between the observed vector and the coefficient matrix, we first deduce the weighted least-squares solution of Partial-EIV model, and a new RWTLS algorithm of correlated observation is proposed to solve the initial values of robust iterations by using the median parameter method. Then the median parameter method is used to determine the initial value, and on this basis we propose a new robust estimated method, which is based on the standardized residual error and considered the influence of gross error both on observation and structure spaces. The experiment results show that the proposed estimated method has a good performance to resist gross error, and the presented solution is more accurate than the traditional method for line fitting, and with the increase of the number of gross errors, the stability of the algorithm is superior to the traditional method.

Key words: robust weighted total least squares; median parameter method; partial errors-in-variables model; correlated observation; standardized residual error

First author: LIU Chunyang, PhD candidate, specializes in surveying data processing. E-mail: cylu6666@163.com

Corresponding author: WANG Jian, PhD, professor. E-mail: wjian@cumt.edu.cn

Foundation support: The National Key Research and Development Program of China, No. 2016YFC0803103.