

一种加权整体最小二乘估计的高效算法

王建民^{1,2}, 倪福泽³, 赵建军²

(1. 太原理工大学 矿业工程学院, 山西 太原 030024; 2. 成都理工大学 地质灾害防治与地质环境保护国家重点实验室, 四川 成都 610059; 3. 中煤(西安)航测遥感研究院有限公司, 陕西 西安 710100)

摘要: 加权整体最小二乘法(WTLS)是估计 errors-in-variables (EIV)模型参数严密的方法, 当面临大数据集时, 其计算效率有限。针对EIV模型中设计矩阵呈现出的结构性特征, 在最小二乘准则的约束条件下, 通过仅给设计矩阵的随机列赋予权重, 推证了适用于EIV模型参数估计的部分加权整体最小二乘法(PWTLS)。PWTLS无需借助拉格朗日辅助法, 能够精确估计EIV模型参数; 另外, 该算法缩减了矩阵的维数, 同时在迭代过程中避免了估计设计矩阵的随机误差, 从而减小了矩阵运算量, 提升了计算效率。最后以真实数据和模拟数据为例与其他7种同类算法进行对比, 结果表明, PWTLS取得了与同类算法相同的精度, 但计算效率显著提高, 验证了算法的可行性。

关键词: 整体最小二乘; 变量误差模型; 计算效率; 坐标转换
中图分类号: P207 **文献标志码:** A

The PWTLS can obtain an exact solution of the EIV model without applying Lagrange multipliers in a straightforward manner. In addition, the PWTLS reduces the dimensions of the cofactor matrix and does not estimate the random error of the design matrix, as this would greatly improve the computational efficiency. Finally, real and simulated examples are used to demonstrate the accuracy and computational performance of the proposed algorithms. The results show that the PWTLS can obtain the same accuracy as the existing seven improved algorithms, but the computational efficiency is significantly improved.

Key words: total least-squares; errors-in-variables; computational efficiency; coordinate transformations

An Efficient Algorithm for Weighted Total Least Squares Method

WANG Jianmin^{1,2}, NI Fuze³, ZHAO Jianjun²

(1. College of Mining Engineering, Taiyuan University of Technology, Taiyuan 030024; 2. State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, China; 3. Aerial Photogrammetry and Remote Sensing Research Institute Co., Ltd., Xi'an 710100, China)

Abstract: The weighted total least-squares (WTLS) adjustment is a rigorous method for estimating parameters in the errors-in-variables (EIV) model. However, the WTLS are not proper for larger data problem in terms of computational efficiency. Aimed at the structural characteristics of the design matrix in the EIV model, a partially weighted total least-squares (PWTLS) algorithm is proposed based on weighted least-squares (WLS) adjustment by weighting the random column of the design

高斯-马尔可夫模型在许多工程实践中得到成功应用, 通常认为模型中的设计矩阵是无误差的。然而, 在许多情景下, 设计矩阵中的元素并不能保证是常数, 其本身也是含有随机误差的观测值, 设计矩阵和观测向量中含有随机误差的模型通常被定义为变量误差模型(errors-in-variables, EIV)^[1-2]。

最初, 因为求解EIV模型参数的方法不成熟, 通常忽略设计矩阵中的随机误差, 而采用加权最小二乘法(weighted least-squares, WLS)近似估计EIV参数, 这种方法是不严密的。1980年, Golub等^[3]提出采用整体最小二乘法(total least-squares, TLS)估计EIV模型的未知参数。当设计矩阵和观测向量均含有随机误差时, TLS就是最小二乘(least-squares, LS)的扩展且更具一般性^[4]。TLS将设计矩阵中的所有元素按等权对待, 但实际上设计矩阵中的元素并非都含有随机误差, 且精度也不尽相同, 因此TLS被进一步扩展为加权整体最小二乘法(weighted

收稿日期: 2020-06-23

基金项目: 地质灾害防治与地质环境保护国家重点实验室开放基金(SKLGP2020K027); 山西省自然科学基金(201901D111048)

第一作者: 王建民(1976—), 男, 副教授, 工学博士, 主要研究方向为测绘数据处理理论与方法。
E-mail: 8844.4321@163.com



论文
拓展
介绍

TLS, WTLS), WTLS分别给观测向量和设计矩阵附加了不同的权重,比TLS更具一般性。

一般来说,主要有两类方法实现TLS和WTLS的求解,一类方法是基于奇异值分解(singular value decomposition, SVD)的直接解算方法^[4-8], Akyilmaz^[6]在SVD分解中给设计矩阵和观测向量赋予了相应的权矩阵,但引入的权阵并没有真正起到调配误差的作用,其本质上是对每个参数和观测值进行了缩放^[9]。另一类常用的方法是基于拉格朗日辅助法构建极值函数分别对未知量进行求偏导数,许多文献据此提出了WTLS的迭代算法^[2,5,8,10-12],建立在统计框架下的这些迭代算法估计EIV模型参数是严密的^[13]。因为WTLS同时顾及了设计矩阵和观测向量的权重,在工程实践中应用迭代法估计EIV模型参数更为普遍。

从现有的研究成果来看,虽然国内外学者就EIV模型提出了许多卓有成效的解算方法,但是与传统的WLS相比,算法比较复杂,不具有操作上的优势^[14],而且大多数迭代算法仅考虑了参数估计精度和迭代收敛速率,当面临大数据集时计算效率有限。

为了提高WTLS的效率,一些学者从不同侧面改进了WTLS算法。Shen等^[10]提出基于Newton-Gauss的非线性WLS平差方法, Jazaeri等^[15]导出了基于WLS准则下的迭代算法,这些方法都明显提升了迭代收敛速率。但仅通过增强迭代收敛速率,计算效率提升有限。事实上,绝大多数EIV模型的设计矩阵具有结构性特征,仅有部分元素是随机的,其通常的解法是构造特定的权矩阵,使系数矩阵中固定元素位置的改正数为零,但这样会导致部分矩阵维数增大,从而加大了矩阵的运算量,不利于计算效率的提高。为此, Xu等^[9]提出了partial EIV (PEIV)模型,并给出了相应的算法。PEIV模型一个显著的优点是未知数的总量明显减小,然而多次迭代收敛拖慢了其计算效率^[9,16-17]。Zhao^[17]应用拉格朗日辅助法对参量求导增强了PEIV算法的效率,王乐洋等^[16]应用间接平差的原理进行改进,使算法更加简洁、紧凑。这两种方法改进的结果是迭代快速收敛,效率明显提高,但不足之处是需要根据原设计矩阵的结构解析出一个固定矩阵和固定向量,而且在迭代过程仍需重构设计矩阵,不利于效率的提升。

WTLS在大地测量等众多科学研究和工程领域中得到应用^[18-19], WTLS算法基本成熟,渐成理论体系。算法之间的差异主要体现在计算效率和算法原

理的复杂程度上。在大数据的背景下,计算效率是决定算法能有效应用的一个重要因素。例如应用WTLS处理点云数据,数据量会显著增加,算法性能就会显得尤为重要。

本文针对EIV模型设计矩阵呈现出的结构性特征,将设计矩阵分成随机和固定部分,仅给随机部分赋予权重,并且不涉及拉格朗日辅助法对参量求偏导,在WLS平差约束准则下,推证了一种估计EIV模型参数的迭代算法,该算法能够显著降低矩阵的运算量,极大提升计算效率。

1 基于WLS准则的WTLS高效算法原理

1.1 EIV模型及WTLS算法

EIV模型的函数模型和随机模型通常表示为^[8-9]

$$L = (A - E)\theta + e \quad (1)$$

$$\begin{bmatrix} e \\ \text{vec}(E) \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} Q_e & 0 \\ 0 & Q_E \end{bmatrix} \right) \quad (2)$$

式(1)、(2)中: E 为 $(n \times t)$ 阶设计矩阵 A 的随机误差; e 为 n 维观测向量 L 的随机误差; $(n \times n)$ 阶矩阵 Q_e 和 $(nt \times nt)$ 阶矩阵 Q_E 分别为 e 和 $\text{vec}(E)$ 的协因数矩阵; $\text{vec}(\cdot)$ 表示矩阵的拉直变换; θ 为待估计的 t 个未知参数; σ^2 为单位权方差。

对于EIV模型,可直接应用最小二乘准则 $e^T Q_e^{-1} e = \min$,其参数估值为^[15]

$$\hat{\theta} = \left[(A - \hat{E})^T Q_e^{-1} (A - \hat{E}) \right]^{-1} (A - \hat{E})^T Q_e^{-1} L \quad (3)$$

式中: \hat{E} 和 $\hat{\theta}$ 分别表示 E 和 θ 的估计值。

仅对随机误差 e 加以约束,难以估计 E ,一般的WTLS的平差准则为^[10-11,15]

$$e^T Q_e^{-1} e + \text{vec}(E)^T Q_E^{-1} \text{vec}(E) = \min \quad (4)$$

在式(4)的约束条件下,利用拉格朗日辅助法可以精确估计未知参数 θ ^[8],即

$$\hat{\theta} = \left[(A - \hat{E})^T Q^{-1} (A - \hat{E}) \right]^{-1} \times (A - \hat{E})^T Q^{-1} (L - \hat{E}\hat{\theta}) \quad (5)$$

$$\text{vec}(\hat{E}) = -Q_E X Q^{-1} (L - A\hat{\theta}) \quad (6)$$

式(5)、(6)中: $Q = Q_e + \hat{X}^T Q_E \hat{X}$, $\hat{X} = (\hat{\theta} \otimes I_n)$, I_n 为单位阵, \otimes 表示两个矩阵的直积运算。

式(5)、(6)是WTLS常用的迭代算式,一般的WTLS方法在迭代过程中需要更新矩阵 \hat{E} ,不利于算法效率的提升。接下来,根据设计矩阵 A 呈现出

的结构性特征,推导一种基于WLS准则的EIV模型参数估计高效算法。

1.2 顾及设计矩阵结构的WTLS算法

理论上EIV模型中的设计矩阵的元素均可包含随机误差,但在实际应用中,绝大多数EIV模型的设计矩阵都会存在部分固定列并且不含有随机误差^[9,17,20]。对于含有固定列的设计矩阵,通常是构建一个特定的协因数矩阵 Q_E ,在迭代的过程中使得固定元素的改正数为零,但这样会导致 Q_E 的维数被固定列“扩大”,制约了计算效率。

设 A 中有 t_1 列固定列和 t_2 列随机列,将 A 中的固定列和随机列进行矩阵分块,对应的参数 θ 也分成两部分,即

$$A=[A_1, A_2] \quad (7)$$

$$\theta=[\theta_1, \theta_2]^T \quad (8)$$

式(7)、(8)中: A_1 为由固定列组成的 $(n \times t_1)$ 阶矩阵; A_2 为由随机列组成的 $(n \times t_2)$ 阶矩阵;与 A_1 和 A_2 对应的参数分别为 θ_1 和 θ_2 。

根据式(7)和式(8),EIV模型(式(1)和式(2))重新改写为

$$L=(A_1, A_2 - E_2) \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + e \quad (9)$$

$$\begin{bmatrix} e \\ \text{vec}(E_2) \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} Q_e & 0 \\ 0 & Q_{E_2} \end{bmatrix} \right) \quad (10)$$

式(9)、(10)中: E_2 为 A_2 的随机误差矩阵,其阶数为 $(n \times t_2)$; Q_{E_2} 为 $\text{vec}(E_2)$ 的协因数矩阵,其阶数为 $(nt_2 \times nt_2)$ 。

将式(9)改写为以下误差方程:

$$L - A_1 \hat{\theta}_1 = \hat{X}_2^T \text{vec}(\hat{A}_2) + \hat{e} \quad (11)$$

式中: $\hat{X}_2^T = (\hat{\theta}_2^T \otimes I_n)$ 。

顾及式(11), $\text{vec}(A_2)$ 表示为

$$\text{vec}(A_2) = \text{vec}(\hat{A}_2) + \text{vec}(\hat{E}_2) \quad (12)$$

将式(11)和式(12)组合形成如下的非线性方程,即

$$H = B \text{vec}(\hat{A}_2) + \hat{\Omega} \quad (13)$$

$$\text{式中: } H = \begin{bmatrix} L - A_1 \hat{\theta}_1 \\ \text{vec}(A_2) \end{bmatrix}; B = \begin{bmatrix} \hat{X}_2^T \\ I_{nt_2} \end{bmatrix}; \hat{\Omega} = \begin{bmatrix} \hat{e} \\ \text{vec}(\hat{E}_2) \end{bmatrix}。$$

在 $\hat{\Omega}^T P_H \hat{\Omega} = \min$ 的约束下,应用WLS原理,式(13)中的 $\text{vec}(\hat{A}_2)$ 的估值为

$$\begin{aligned} \text{vec}(\hat{A}_2) &= [B^T P_H B]^{-1} B^T P_H H = \\ & (Q_{E_2}^{-1} + \hat{X}_2 Q_e^{-1} \hat{X}_2^T)^{-1} \times \\ & [\hat{X}_2 Q_e^{-1} (L - A_1 \hat{\theta}_1) + Q_{E_2}^{-1} \text{vec}(A_2)] \end{aligned} \quad (14)$$

$$\text{式中: } P_H = \begin{bmatrix} Q_e^{-1} & 0 \\ 0 & Q_{E_2}^{-1} \end{bmatrix}。$$

为进一步化简式(14),需利用式(15)的矩阵求逆引理^[21],即

$$(J^{-1} - K D^{-1} Z)^{-1} = J + J K (D - Z J K)^{-1} Z J \quad (15)$$

令 $J^{-1} = Q_{E_2}^{-1}$, $K = -\hat{X}_2$, $D = Q_e$ 且 $Z = \hat{X}_2^T$,式(14)化简为

$$\begin{aligned} \text{vec}(\hat{A}_2) &= [Q_{E_2}^{-1} - Q_{E_2}^{-1} \hat{X}_2 (Q_e + \hat{X}_2^T Q_{E_2}^{-1} \hat{X}_2)^{-1} \times \\ & \hat{X}_2^T Q_{E_2}^{-1}] [\hat{X}_2 Q_e^{-1} (L - A_1 \hat{\theta}_1) + Q_{E_2}^{-1} \text{vec}(A_2)] = \\ & \text{vec}(A_2) + Q_{E_2}^{-1} \hat{X}_2 (Q_e + \hat{X}_2^T Q_{E_2}^{-1} \hat{X}_2)^{-1} \times \\ & (L - A_1 \hat{\theta}_1 - A_2 \hat{\theta}_2) = \\ & \text{vec}(A_2) + Q_{E_2}^{-1} \hat{X}_2 Q_2^{-1} (L - A \hat{\theta}) \end{aligned} \quad (16)$$

式中: $Q_2 = (Q_e + \hat{X}_2^T Q_{E_2}^{-1} \hat{X}_2)$ 。

由式(16)可知,设计矩阵 A_2 的随机误差的估值为

$$\text{vec}(\hat{E}_2) = -Q_{E_2} \hat{X}_2 Q_2^{-1} (L - A \hat{\theta}) \quad (17)$$

由式(7)可知,观测向量 L 的残差可表示为

$$\begin{aligned} \hat{e} &= L - \hat{A} \hat{\theta} = \\ & L - [A_1, A_2 - \hat{E}_2] [\hat{\theta}_1, \hat{\theta}_2]^T = \\ & L - A \hat{\theta} + \hat{X}_2^T \text{vec}(\hat{E}_2) = \\ & L - A \hat{\theta} - \hat{X}_2^T Q_{E_2} \hat{X}_2 Q_2^{-1} (L - A \hat{\theta}) = \\ & (I - \hat{X}_2^T Q_{E_2} \hat{X}_2 Q_2^{-1}) (L - A \hat{\theta}) = \\ & Q_e Q_2^{-1} (L - A \hat{\theta}) \end{aligned} \quad (18)$$

由式(18)可知

$$L = \hat{A} \hat{\theta} + Q_e Q_2^{-1} (L - A \hat{\theta}) \quad (19)$$

将式(19)代入式(3),得

$$\begin{aligned} \hat{\theta} &= (\hat{A}^T Q_e^{-1} \hat{A})^{-1} \hat{A}^T Q_e^{-1} [\hat{A} \hat{\theta} + Q_e Q_2^{-1} (L - A \hat{\theta})] = \\ & \hat{\theta} + (\hat{A}^T Q_e^{-1} \hat{A})^{-1} \hat{A}^T Q_2^{-1} (L - A \hat{\theta}) \end{aligned} \quad (20)$$

则

$$\hat{A}^T \lambda = (A - \hat{E})^T \lambda = 0 \quad (21)$$

式中: $\lambda = Q_2^{-1} (L - A \hat{\theta})$, λ 为 $n \times 1$ 的向量。

由式(21)可知

$$A^T \lambda = \hat{E}^T \lambda \quad (22)$$

根据 A 的矩阵分块(式(7)),式(22)可分解为两部分,即

$$A_1^T \lambda = \hat{E}_1^T \lambda = O \lambda \quad (23)$$

$$A_2^T \lambda = \hat{E}_2^T \lambda \quad (24)$$

式中: $\hat{E}_1^T = O$, O 为 $(t_1 \times n)$ 零矩阵。

式(25)和式(26)是两个克罗内克乘积法则,用于式(24)。

$$\text{vec}(\mathbf{S}_{n \times 1}) = \text{vec}(\mathbf{S}_n^T \mathbf{1}) \quad (25)$$

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}) \quad (26)$$

式(24)可表示为

$$\mathbf{A}_2^T \lambda = \hat{E}_2^T \lambda = \mathbf{I}_{t_2} \hat{E}_2^T \lambda = \text{vec}(\lambda^T \hat{E}_2 \mathbf{I}_{t_2}) \quad (27)$$

则

$$\mathbf{A}_2^T \lambda = \mathbf{A}_2^T \mathbf{Q}_2^{-1} (\mathbf{L} - \mathbf{A} \hat{\theta}) = (\mathbf{I}_{t_2} \otimes \lambda^T) \text{vec}(\hat{E}_2) \quad (28)$$

将式(17)代入式(28),得

$$\mathbf{A}_2^T \mathbf{Q}_2^{-1} (\mathbf{L} - \mathbf{A} \hat{\theta}) = -(\mathbf{I}_{t_2} \otimes \lambda^T) \mathbf{Q}_{E_2} \hat{X}_2 \mathbf{Q}_2^{-1} (\mathbf{L} - \mathbf{A} \hat{\theta}) = \mathbf{G} (\mathbf{L} - \mathbf{A} \hat{\theta}) \quad (29)$$

式中: $\mathbf{G} = -(\mathbf{I}_{t_2} \otimes \lambda^T) \mathbf{Q}_{E_2} \hat{X}_2 \mathbf{Q}_2^{-1}$, \mathbf{G} 为 $(t_2 \times n)$ 的矩阵。

组合式(23)和式(29),得

$$\begin{bmatrix} \mathbf{A}_1^T \\ \mathbf{A}_2^T \end{bmatrix} \mathbf{Q}_2^{-1} (\mathbf{L} - \mathbf{A} \hat{\theta}) = \begin{bmatrix} \mathbf{O} \\ \mathbf{G} \end{bmatrix} (\mathbf{L} - \mathbf{A} \hat{\theta}) \quad (30)$$

式(30)可简化为

$$(\mathbf{A}^T \mathbf{Q}_2^{-1} \mathbf{A} - \mathbf{U} \mathbf{A}) \hat{\theta} = (\mathbf{A}^T \mathbf{Q}_2^{-1} - \mathbf{U}) \mathbf{L} \quad (31)$$

式中: $\mathbf{U} = [\mathbf{O} \quad \mathbf{G}]^T$ 。

式(31)中的 $(\mathbf{A}^T \mathbf{Q}_2^{-1} \mathbf{A} - \mathbf{U} \mathbf{A})$ 是列满秩矩阵,则未知参数 $\hat{\theta}$ 最终的估计结果为

$$\hat{\theta} = (\mathbf{A}^T \mathbf{Q}_2^{-1} \mathbf{A} - \mathbf{U} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{Q}_2^{-1} - \mathbf{U}) \mathbf{L} \quad (32)$$

由式(17)和式(18)可估计 σ^2 , 即

$$\hat{\sigma}^2 = \frac{\hat{e}^T \mathbf{Q}_e^{-1} \hat{e} + \text{vec}(\hat{E}_2)^T \mathbf{Q}_{E_2}^{-1} \text{vec}(\hat{E}_2)}{n - t} = \frac{\lambda^T (\mathbf{A} \hat{\theta} - \mathbf{L})}{n - t} \quad (33)$$

上述推导是基于 WLS 最小二乘准则,并且式(32)在形式上与间接平差的法方程类似,但方程两侧都存在未知参数 $\hat{\theta}$, 仍需迭代计算。为了加以区别和简写,上述算法称为部分加权整体最小二乘法 (partially weighted TLS, PWTLS), PWTLS 的算法流程步骤如下:

(1) 初始化初值, 即

$$\hat{\theta}_{(0)} = (\mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Q}_L^{-1} \mathbf{L}$$

(2) 开始迭代, 即

$$\hat{\theta}_{(i)} = [\hat{\theta}_{1(i)} \quad \hat{\theta}_{2(i)}]^T$$

$$\mathbf{X}_{2(i)}^T = \hat{\theta}_{2(i)}^T \otimes \mathbf{I}_n$$

$$\mathbf{Q}_{2(i)} = (\mathbf{Q}_e + \hat{X}_{2(i)}^T \mathbf{Q}_{E_2} \hat{X}_{2(i)})$$

$$\lambda_{(i)} = \mathbf{Q}_{2(i)}^{-1} (\mathbf{L} - \mathbf{A} \hat{\theta}_{(i)})$$

$$\mathbf{G}_{(i)} = -(\mathbf{I}_{t_2} \otimes \lambda_{(i)}^T) \mathbf{Q}_{E_2} \hat{X}_{2(i)} \mathbf{Q}_{2(i)}^{-1}$$

$$\mathbf{U}_{(i)} = [\mathbf{O} \quad \mathbf{G}_{(i)}]^T$$

$$\hat{\theta}_{(i+1)} = [\mathbf{A}^T \mathbf{Q}_{2(i)}^{-1} \mathbf{A} - \mathbf{U}_{(i)} \mathbf{A}]^{-1} (\mathbf{A}^T \mathbf{Q}_{2(i)}^{-1} - \mathbf{U}_{(i)}) \mathbf{L}$$

(3) 重复步骤(2)直到 $|\hat{\theta}_{(i+1)} - \hat{\theta}_{(i)}| < \epsilon$, 迭代结束(下标 i 表示迭代次数, ϵ 为设定的收敛阈值)。

由上述算法可知,在迭代过程中,并不涉及对设计矩阵 \mathbf{A} 拆分成 \mathbf{A}_1 和 \mathbf{A}_2 , 仅需把随机列 \mathbf{A}_2 对应的参数 θ_2 分解出来参与迭代。只要 EIV 模型的设计矩阵存在固定列, 根据列与参数的对应关系, 很容易将 \mathbf{A} 组合成固定和随机部分。通常情况下, 习惯上总是将 EIV 模型的固定列排列在一起, 而且绝大多数 EIV 模型至少有一列是固定列, 因而 PWTLS 估计 EIV 参数具有普适性。另外, 在实际应用中, 设计矩阵的随机列可能会含有少量固定元素, 例如三维七参数坐标转换模型中, 有 3 列固定列组成 \mathbf{A}_1 , 在随机列中有一个固定零元素, 对于随机列中的个别固定元素赋予值为零的协因数后, 仍可组成随机列 \mathbf{A}_2 。

从算法流程可以看出, PWTLS 算法不同于一般的 WTLS 算法, 在每次迭代的过程中并不需要估计随机误差 E ; 通过给设计矩阵的随机列赋予权重, 使得相关矩阵运算的时间复杂度由 $O(nt)$ 降低到 $O(nt_2)$, 占用的内存空间由 $O(nt^2)$ 减小到 $O(nt_2^2)$, 例如原参数矩阵 \hat{X} 的维数 $(nt \times nt)$ 降为 $\hat{X}_2 (nt_2 \times nt_2)$, 基于这两点, 使 PWTLS 算法的效率显著提高。

2 算例与分析

为了验证 PWTLS 的性能, 本文选择了 7 种具有代表性的 WTLS 算法^[10-11, 15-17, 20, 22]与本文算法进行对比分析, 这 7 种算法从不同角度对 WTLS 进行效率增强, 具体原理见参考文献, 有的算法在相关文献中也与其他算法进行过对比研究。

2.1 真实算例

直线拟合简单而有效, 经常用于证明 WTLS 算法的有效性。设有直线方程为 $y = ax + b$, 其中 x 和 y 是直线的坐标分量, a 和 b 是直线的斜率和截距。实验数据来自于文献[23], 见表 1, 其中 (x_i, y_i) 和 (w_{x_i}, w_{y_i}) 表示坐标和相应的权值。

根据直线方程和坐标点数量, 设计矩阵 \mathbf{A} , 参数向量 θ 及观测向量 \mathbf{L} 具体形式为

表1 坐标观测值和相应的权重

Tab. 1 Coordinate observations and weights

编号	x_i	W_{xi}	y_i	W_{yi}
1	0	1 000.0	5.9	1.0
2	0.9	1 000.0	5.4	1.8
3	1.8	500.0	4.4	4.0
4	2.6	800.0	4.6	8.0
5	3.3	200.0	3.5	20.0
6	4.4	80.0	3.7	20.0
7	5.2	60.0	2.8	70.0
8	6.1	20.0	2.8	70.0
9	6.5	1.8	2.4	100.0
10	7.4	1.0	1.5	500.0

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{10} \end{bmatrix}, L = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{10} \end{bmatrix}, \theta = \begin{bmatrix} a \\ b \end{bmatrix} \quad (34)$$

协因数矩阵 Q_e 和 Q_E 分别为

$$\begin{aligned} Q_e &= \text{diag}(W_{y_1}^{-1}, W_{y_2}^{-1}, \dots, W_{y_{10}}^{-1}), \\ Q_E &= \text{diag}(\underbrace{0, 0, \dots, 0}_{10}, W_{x_1}^{-1}, W_{x_2}^{-1}, \dots, W_{x_{10}}^{-1}) \end{aligned} \quad (35)$$

在式(35)中,设计矩阵 A 可分成 A_1 和 A_2 两部分,即

$$A_1 = [\underbrace{1, 1, \dots, 1}_{10}]^T, A_2 = [x_1, x_2, \dots, x_{10}]^T \quad (36)$$

A_2 的协因数矩阵 Q_{E_2} 的结构取自 Q_E 的一部分,即

$$Q_{E_2} = \text{diag}(W_{x_1}^{-1}, W_{x_2}^{-1}, \dots, W_{x_{10}}^{-1}) \quad (37)$$

设迭代收敛阈值 $\epsilon=10^{-10}$,分别应用8种算法估计参数和单位权中误差,实验结果见表2。从表2中可明显看出,8种严密算法估计的参数相差在 10^{-11} 量级且中误差相同,表明PWTLs能够获得和另外7种严密算法相同的解。文献[22]对PEIV模型进行改进,其目的也是提升效率,从本例的结果来看,收敛速度高达454次,其余算法的收敛速度都较快。文献[16]同样是对PEIV模型算法进行优化,取得了理想的效果。比较而言,文献[20]的算法收敛最快,收敛速度仅是影响计算效率一方面的因素。本例中直线拟合的设计矩阵结构简单而且仅有10个点,主要用于PWTLs参数估计精度的验证及收敛速度的初步分析,接下来通过大数据量应用仿真模拟实验来对比分析各个算法的效率。

表2 8种不同算法的参数和中误差的估计结果

Tab. 2 Estimated parameters and standard deviations obtained using eight algorithms

算法	参数 a	参数 b	中误差 σ	迭代次数
PWTLs	5.479 910 224 033	-0.4805 334 074 462	1.217 91	8
文献[11]	5.479 910 224 033	-0.4805 334 074 462	1.217 91	8
文献[17]	5.479 910 224 033	-0.4805 334 074 462	1.217 91	8
文献[16]	5.479 910 22 4032	-0.4805 334 074 461	1.217 91	7
文献[20]	5.479 910 224 032	-0.4805 334 074 462	1.217 91	5
文献[15]	5.479 910 224 032	-0.4805 334 074 461	1.217 91	7
文献[10]	5.479 910 224 030	-0.4805 334 074 456	1.217 91	13
文献[22]	5.479 910 226 342	-0.4805 334 079 855	1.217 91	454

2.2 仿真算例

平面相似四参数转换模型是常用的坐标基准转换模型之一,通常由式(38)表示^[24]

$$\begin{bmatrix} x \\ y \end{bmatrix}_T = \begin{bmatrix} \xi \\ \eta \end{bmatrix} + k \begin{bmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_S \quad (38)$$

式中: ξ 和 η 为平移参数; k 为尺度参数; a 为旋转角;下标S和T分别表示原坐标系和目标坐标系。

假设有 d 个重合点用于求解转换参数,式(38)经过变换可由EIV模型描述,设计矩阵 A ,参数 θ 及观测向量 L 表示为

$$A_{2d,4} = \begin{bmatrix} 1 & 0 & x_1 & y_1 \\ 0 & 1 & y_1 & -x_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_d & y_d \\ 0 & 1 & y_d & -x_d \end{bmatrix}_S, \theta = \begin{bmatrix} \xi \\ \eta \\ u \\ w \end{bmatrix}, L_{2d,1} = \begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ x_d \\ y_d \end{bmatrix}_T \quad (39)$$

式中: $u=k \cos a$; $w=k \sin a$ 为辅助参数。

参数 ξ 和 η 对应的系数是1和0的常量,设计矩阵 A 的第1列和第2列是固定列,组成 A_1 ;其他两列组成 A_2 形成随机列, A_1 和 A_2 表示为

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} x_1 & y_1 \\ y_1 & -x_1 \\ \vdots & \vdots \\ x_d & y_d \\ y_d & -x_d \end{bmatrix} \quad (40)$$

对应于 \mathbf{A}_1 和 \mathbf{A}_2 ,相应的参数可分成两部分,即

$$\boldsymbol{\theta}_1 = [\xi, \eta]^T, \boldsymbol{\theta}_2 = [u, w]^T \quad (41)$$

原坐标系中的坐标由随机函数产生 d 个坐标点,目标坐标由给定的参数真值 $\xi = -27.366$, $\eta = -71.185$, $u = 1.000\ 001\ 092$ 和 $w = 6.400\ 15 \times 10^{-7}$ 通过转换模型反算得到。原坐标和目标坐标中的随机误差由随机函数(式(42)、(43))随机生成,并将随机误差附加到相应的坐标分量中。

$$\mathbf{Q}_{E_2} = \begin{bmatrix} q_{x_1}^{1,1} & 0 & 0 & 0 & 0 & 0 & q_{x_1}^{1,2d+2} & 0 & 0 & 0 \\ 0 & q_{y_1}^{2,2} & 0 & 0 & 0 & -q_{y_1}^{2,2d+1} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & q_{x_d}^{2d-1,2d-1} & 0 & 0 & 0 & 0 & 0 & q_{x_d}^{2d-1,4d} \\ 0 & 0 & 0 & 0 & q_{y_d}^{2d,2d} & 0 & 0 & 0 & -q_{y_d}^{2d,4d-1} & 0 \\ 0 & -q_{y_1}^{2d+1,2} & 0 & 0 & 0 & q_{y_1}^{2d+1,2d+1} & 0 & 0 & 0 & 0 \\ q_{x_1}^{2d+2,1} & 0 & 0 & 0 & 0 & 0 & q_{x_1}^{2d+2,2d+2} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -q_{y_d}^{4d-1,2d} & 0 & 0 & 0 & q_{y_d}^{4d-1,4d-1} & 0 \\ 0 & 0 & 0 & q_{x_d}^{4d,2d-1} & 0 & 0 & 0 & 0 & 0 & q_{x_d}^{4d,4d} \end{bmatrix} \quad (45)$$

式中: $q_{x_i} = \frac{\Delta_{x_i}^2}{\sigma^2}$; $q_{y_i} = \frac{\Delta_{y_i}^2}{\sigma^2}$ ($i=1, 2, \dots, d$);上标表示元素在矩阵中的位置,其中 Δ_{x_i} 和 Δ_{y_i} 来自于向量 \mathbf{A}_s 。

基于上述的模型和数据,随机模拟1000次实验,固定重合点数 $d=200$,应用均方根误差(RMSE)来衡量参数估计的精度,参数 ξ 、 η 、 u 和 w 的RMSE由式(46)、(47)估计

$$\hat{\sigma}_\xi = \sqrt{\frac{1}{1000} \sum_{i=1}^{1000} (\xi - \hat{\xi}_i)^2}, \hat{\sigma}_\eta = \sqrt{\frac{1}{1000} \sum_{i=1}^{1000} (\eta - \hat{\eta}_i)^2} \quad (46)$$

$$\hat{\sigma}_u = \sqrt{\frac{1}{1000} \sum_{i=1}^{1000} (u - \hat{u}_i)^2}, \hat{\sigma}_w = \sqrt{\frac{1}{1000} \sum_{i=1}^{1000} (w - \hat{w}_i)^2} \quad (47)$$

式(46)、(47)中: $\hat{\xi}_i$ 、 $\hat{\eta}_i$ 、 \hat{u}_i 和 \hat{w}_i 为经第 i 次模拟的对应参数的估计值。

在相同数据和计算平台的实验条件下(收敛阈值 $\epsilon = 10^{-10}$),分别应用7种具有较高收敛速度的算

$$\mathbf{A}_s = [\Delta_{x_1} \ \Delta_{y_1} \ \cdots \ \Delta_{x_d} \ \Delta_{y_d}]^T = \text{mvnrnd}(0, 0.05, 2d) \quad (42)$$

$$\mathbf{A}_T = [\Delta_{x_1} \ \Delta_{y_1} \ \cdots \ \Delta_{x_d} \ \Delta_{y_d}]^T = \text{mvnrnd}(0, 0.05, 2d) \quad (43)$$

式中: \mathbf{A}_s 和 \mathbf{A}_T 分别表示原始坐标和目标坐标中的真误差,mvnrnd($0, \sigma, 2d$)生成期望为0、标准差为 $\sigma = 0.05$ m、维数为 $2d \times 1$ 的随机向量。

由随机误差和标准差定义观测量向量 L 的协因数阵 \mathbf{Q}_e

$$\mathbf{Q}_e = \text{diag} \left(\frac{\Delta_{x_1}^2}{\sigma^2} \ \frac{\Delta_{y_1}^2}{\sigma^2} \ \cdots \ \frac{\Delta_{x_d}^2}{\sigma^2} \ \frac{\Delta_{y_d}^2}{\sigma^2} \right)_T \quad (44)$$

根据设计矩阵 \mathbf{A}_2 的结构,vec(\mathbf{A}_2)的协因数阵 \mathbf{Q}_{E_2} 定义为

法进行模拟,参数的RMSE的统计结果和1000次模拟的累计耗时见表3,结果表明,严密算法中的RMSE差异为 10×10^{-5} mm,进一步证明了PWTLs与其他严密算法的参数估计精度相同。

从计算的总耗时可以看出,PWTLs是所有严密算法中计算效率最高的,与其他算法计算效率相比,最少提高了161%,最多提高了近416%。其主要是因为PWTLs使用 \mathbf{Q}_{E_2} 代替了 \mathbf{Q}_E ,维数由 800×800 降到 400×400 ;另外,在每次迭代过程中,PWTLs算法并不需要估计随机误差矩阵 E 。

为了突出PWTLs计算效率的优越性,从另一个角度执行了10次模拟实验,要求重合点数据以间隔为100个点,逐渐从 $d=100$ 增加到1000。计算耗时统计结果如图1所示,结果表明,PWTLs的耗时曲线变化相对平缓,意味着随着数据量的增大,PWTLs的优势越来越明显。

表3 1 000次模拟的RMSE和累计耗时

Tab. 3 Results of parameter estimation and details of the time consumed for different algorithms

算法	$\hat{\sigma}_z/\text{mm}$	$\hat{\sigma}_y/\text{mm}$	$\hat{\sigma}_u/10^{-6}$	$\hat{\sigma}_w/10^{-6}$	平均迭代次数	总耗时/s
PWTLS	6.600 843 647 7	6.769 967 561 0	7.825 463 890 4	8.310 172 114 1	3.63	7.61
文献[11]	6.600 843 647 7	6.769 967 561 0	7.825 463 890 4	8.310 172 114 1	3.63	19.86
文献[17]	6.600 857 165 2	6.769 954 304 2	7.825 463 137 6	8.310 192 209 9	3.19	20.71
文献[16]	6.600 835 421 5	6.769 953 801 2	7.825 473 558 9	8.310 158 454 4	3.08	22.97
文献[20]	6.600 843 647 6	6.769 967 560 9	7.825 463 889 6	8.310 172 114 3	3.01	24.14
文献[15]	6.600 843 647 1	6.769 967 560 9	7.825 463 888 3	8.310 172 114 5	3.08	38.95
文献[10]	6.600 827 049 8	6.769 950 721 3	7.825 406 436 6	8.310 176 011 5	3.73	39.28

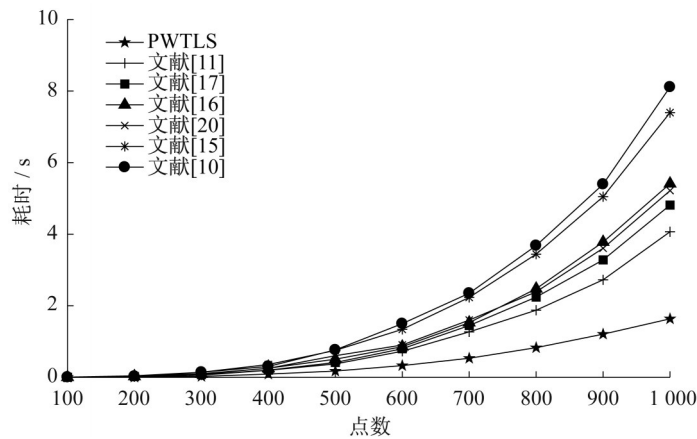


图1 7种算法的计算效率比较

Fig. 1 Computational efficiency of seven algorithms

3 结语

WTLS已经在众多科学和工程技术中得到应用,当面临处理大数据集时,其计算效率有限,降低算法的复杂度和提高计算效率对WTLS算法将显得越来越重要。

本文针对EIV模型的设计矩阵呈现出的结构性特征,提出了基于WLS准则约束条件下的PWTLS算法,无需引入辅助量,迭代方程与经典的间接平差法方程类似,便于理解,易于实现,估计的参数是严密的。PWTLS通过给设计矩阵中的随机列赋予权重,使得设计矩阵的协因数阵维数降低,从而在迭代过程中缩减了所有参与运算矩阵的维数,极大地减少了计算量,促进了效率的提升;另外,PWTLS算法避免了重复估计设计矩阵随机列的误差,这两方面的因素使矩阵的运算量降低,从而显著提升了计算效率。最后,通过算例验证了所提出的PWTLS算法与同类算法参数估计精度相当,但计算效率显著提高。

作者贡献说明:

王建民:提出了论文中PWTLS算法原理,完成了论文的整体写作。

倪福泽:编写了计算程序,完成了论文中的实验计算和分析,绘制了图表。

赵建军:提供研究经费的资助,修改了中英文摘要。

参考文献:

- [1] 刘经南,曾文宪,徐培亮.整体最小二乘估计的研究进展[J].武汉大学学报(信息科学版),2013,38(5):505.
LIU Jingnan, ZENG Wenxian, XU Peiliang. Overview of total least squares methods[J]. Geomatics and Information Science of Whuhan University, 2013, 38(5): 505.
- [2] FANG X. Weighted total least squares solutions for applications in geodesy [D]. Hannover: Gottfried Wilhelm Leibniz Universität Hannover, 2011.
- [3] GOLUB G H, VAN LOAN C F. An analysis of the total least-squares problem [J]. SIAM Journal on Numerical Analysis, 1980, 17(6): 883.
- [4] MARKOVSKY I, VAN HUFFEL S. Overview of total least-squares methods[J]. Signal Processing, 2007, 87(10): 2283.
- [5] TONG X, JIN Y, LI L. An improved weighted total least

- squares method with applications in linear fitting and coordinate transformation [J]. *Journal of Surveying Engineering*, 2011, 137(4): 120.
- [6] AKYILMAZ O. Total least squares solution of coordinate transformation[J]. *Survey Review*, 2007, 39(303): 68.
- [7] FELUS Y A. Application of total least squares for spatial point process analysis [J]. *Journal of Surveying Engineering*, 2004, 130(3): 126.
- [8] SCHAFFRIN B, WIESER A. On weighted total least-squares adjustment for linear regression[J]. *Journal of Geodesy*, 2008, 82(7): 415.
- [9] XU P, LIU J, SHI C. Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis [J]. *Journal of Geodesy*, 2012, 86(8): 661.
- [10] SHEN Y, LI B, CHEN Y. An iterative solution of weighted total least-squares adjustment [J]. *Journal of Geodesy*, 2011, 85(4): 229.
- [11] MAHBOUB V. On weighted total least-squares for geodetic transformations[J]. *Journal of Geodesy*, 2012, 86(5): 59.
- [12] MAHBOUB V, SHARIFI M A. On weighted total least-squares with linear and quadratic constraints [J]. *Journal of Geodesy*, 2013, 87(3): 279.
- [13] XU P, LIU J, ZENG W, *et al.* Effects of errors-in-variables on weighted least squares estimation [J]. *Journal of Geodesy*, 2014, 88(7): 705.
- [14] 姚宜斌, 孔建. 顾及设计矩阵随机误差的最小二乘组合新解法[J]. *武汉大学学报(信息科学版)*, 2014, 39(9): 1028.
YAO Yibin, KONG Jian. A new combined LS method considering random errors of design matrix [J]. *Geomatics and Information Science of Whuhan University*, 2014, 39(9): 1028.
- [15] JAZAERI S, AMIRI-SIMKOOEI A R, SHARIFI M A. Iterative algorithm for weighted total least squares adjustment [J]. *Survey Review*, 2014, 46(334): 19.
- [16] 王乐洋, 余航, 陈晓勇. Partial EIV模型的解法[J]. *测绘学报*, 2016, 45(1): 22.
- WANG Leyang, YU Hang, CHEN Xiaoyong. An algorithm for partial EIV model [J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(1): 22.
- [17] ZHAO J. Efficient weighted total least-squares solution for partial errors-in-variables model [J]. *Survey Review*, 2017, 49(356): 346.
- [18] 王乐洋. 地壳应变参数反演的总体最小二乘方法[J]. *大地测量与地球动力学*, 2013, 33(3): 106.
WANG Leyang. Inversion of crustal strain parameters based on total least squares [J]. *Journal of Geodesy and Geodynamics*, 2013, 33(3): 106.
- [19] 方兴, 曾文宪, 刘经南, 等. 三维坐标转换的通用整体最小二乘算法[J]. *测绘学报*, 2014, 43(11): 1139.
FANG Xing, ZENG Wenxian, LIU Jingnan, *et al.* A general total least squares algorithm for three-dimensional coordinate transformations [J]. *Acta Geodaetica et Cartographica Sinica*, 2014, 43(11): 1139.
- [20] ZHOU Y, KOU X, LI J, *et al.* Comparison of structured and weighted total least-squares adjustment methods for linearly structured errors-in-variables models [J]. *Journal of Surveying Engineering*, 2017, 143(1): 04016019.
- [21] HENDERSON H V, SEARLE S R. On deriving the inverse of a sum of matrices [J]. *SIAM Review*, 1981, 23(1): 53.
- [22] SHI Y, XU P, LIU J, *et al.* Alternative formulae for parameter estimation in partial errors-in-variables models [J]. *Journal of Geodesy*, 2015, 89(1): 13.
- [23] NERI F, SAITTA G, CHIOFALO S. An accurate and straightforward approach to line regression analysis of error-affected experimental data [J]. *Journal of Physics E: Scientific Instruments*, 1989, 22(4): 215.
- [24] WANG J. Locating and estimating multiple gross errors during coordinate transformation [J]. *Survey Review*, 2015, 47(345): 458.