文章编号:1671-8860(2010)06-0711-04　　**文献标志码**:A

整体最小二乘的迭代解法

孔　建[1]　姚宜斌[1]　吴　寒[1]

(1　武汉大学测绘学院,武汉市珞喻路 129 号,430079)

摘　要:在引入整体最小二乘平差准则的基础上,推导了整体最小二乘的迭代解法;同时,引入多元函数隐函数求导的方法以确定未知参数对观测数据的线性信息,解决了整体最小二乘下的精度评定问题。给出了运用新的解法在拟合函数确定以及坐标转换参数确定等方面的应用实例,验证了新算法的可行性。

关键词:整体最小二乘;非线性方程求解;迭代计算;拟合函数;坐标转换

中图法分类号:P207.2

在测量数据处理过程中,经常会遇到平差模型中系数矩阵也有误差的情况,传统最小二乘往往忽略掉这项误差,这样做显然是不合理的,估计出来的结果,从统计上来看是有偏的,而不是最优的。整体最小二乘的提出正是为了解决这个问题,但是整体最小二乘的解法却制约了其自身的推广应用[1]。本文提出的迭代解法较好地解决了算法复杂度的问题。

1　算法推导

整体最小二乘的观测方程为:

$$\underset{n\times 1}{\boldsymbol{L}}+\underset{n\times 1}{\boldsymbol{V}}=(\underset{n\times 1}{\boldsymbol{B}}+\Delta\underset{n\times 1}{\boldsymbol{B}})\underset{n\times 1}{\boldsymbol{X}}-\underset{n\times 1}{\boldsymbol{d}} \tag{1}$$

此时,系数矩阵是有误差的。设观测值的平差值为 $\hat{\boldsymbol{L}}$,系数矩阵的平差值为 $\hat{\boldsymbol{B}}$,未知参数平差值为 $\hat{\boldsymbol{X}}$,引入平差准则:

$$\sum_{i=1}^{n}(\hat{L}_i-L_i)^2+\sum_{j=1,i=1}^{j=t,i=n}(\hat{B}_{ij}-B_{ij})^2=\min \tag{2}$$

$\hat{\boldsymbol{L}}$、$\hat{\boldsymbol{B}}$、$\hat{\boldsymbol{X}}$ 是在满足式(1)的基础上,使得式(2)取得最小值的最优解。将式(1)代入式(2),有:

$$\sum_{i=1}^{n}(\sum_{j=1}^{n}\hat{B}_{ij}\hat{X}_{ij}-d_i-L_i)+\sum_{j=1,i=1}^{j=t,i=n}(\hat{B}_{ij}-B_{ij})^2=\min \tag{3}$$

将式(3)设为 F,F 分别对待求参数求导:

$$\begin{cases}\dfrac{\partial F}{B_{11}}=(\hat{B}_{11}-B_{11})+(\sum\limits_{j=1}^{t}\hat{B}_{ij}\hat{X}_j-d_1-L_1)\hat{X}_1=0\\ \dfrac{\partial F}{B_{12}}=(\hat{B}_{12}-B_{12})+(\sum\limits_{j=1}^{t}\hat{B}_{ij}\hat{X}_j-d_1-L_1)\hat{X}_2=0\\ \qquad\vdots\\ \dfrac{\partial F}{B_{nt}}=(\hat{B}_{nt}-B_{nt})+(\sum\limits_{j=1}^{t}\hat{B}_{nj}\hat{X}_j-d_n-L_n)\hat{X}_t=0\\ \dfrac{\partial F}{x_1}=(\sum\limits_{j=1}^{t}\hat{B}_j\hat{X}_j-d_1-L_1)\hat{B}_{11}+(\sum\limits_{j=1}^{t}\hat{B}_{2j}\hat{X}_j-d_2-L_2)\hat{B}_{21}+\cdots+(\sum\limits_{j=1}^{t}\hat{B}_{nj}\hat{X}_j-d_n-L_n)\hat{B}_{n1}=0\\ \qquad\vdots\\ \dfrac{\partial F}{x_t}=(\sum\limits_{j=1}^{t}\hat{B}_{ij}\hat{X}_j-d_1-L_1)\hat{B}_{1n}+(\sum\limits_{j=1}^{t}\hat{B}_{2j}\hat{X}_j-d_2-L_2)\hat{B}_{2n}+\cdots+(\sum\limits_{j=1}^{t}\hat{B}_{nj}\hat{X}_j-d_n-L_n)\hat{B}_{nt}=0\end{cases} \tag{4}$$

式(4)给出了 $t(n+1)$ 个方程,其中 $t(n+1)$ 个未知数。但是,方程是非线性的,而且方程

收稿日期:2010-04-28。

项目来源:国家自然科学基金资助项目(40774008,40721001);国家 973 计划资助项目(2006CB701301)。

的形式非常复杂,如果通过线性化求解,则会面临方程不收敛的问题。常规的解法是利用矩阵的奇异值分解,但是这种方法不仅实现复杂,而且不利于编程[2,3]。本文采用了一种新的迭代方法求解。首先,对式(4)中的式子分类,分为对 B_{ij} 求导的式子以及对未知参数 $\hat{X}_i$ 求导的式子,对于第一类式子,以第一式为例化简:

$$(\hat{B}_{11}-B_{11})+(\sum_{j=1}^{t}\hat{B}_j\hat{X}_j-d_1-L_1)\hat{X}_1=0$$
$$\Rightarrow\hat{B}_{11}+\hat{B}_{11}\hat{X}_1^2+\hat{B}_{12}\hat{X}_1\hat{X}_2+\cdots+\hat{B}_{1t}\hat{X}_1\hat{X}_t=B_{11}+(L_1+d_1)\hat{X}_1$$

对第二类式子,仍以第一式为例化简:

$$(\sum_{j=1}^{t}\hat{B}_{1j}\hat{X}_j-d_1-L_1)\hat{B}_{11}+(\sum_{j=1}^{t}\hat{B}_{2j}\hat{X}_j-d_2-L_2)\hat{B}_{21}+\cdots+(\sum_{j=1}^{t}\hat{B}_{nj}\hat{X}_j-d_n-L_n)\hat{B}_{n1}=0$$
$$\Rightarrow\sum_{i=1}^{n}\hat{B}_{i1}^2\hat{X}_1+\sum_{i=1}^{n}\hat{B}_{i1}\hat{B}_{i2}\hat{X}_2+\cdots+\sum_{i=1}^{n}\hat{B}_{it}\hat{B}_{it}\hat{X}_t=\sum_{i=1}^{n}(d_i+L_i)\hat{B}_{i1}$$

将化简之后的式子整理为:

$$\left\{\begin{array}{l}(1+\hat{X}_1^2)\hat{B}_{11}+\cdots+\hat{X}_1\hat{X}_i\hat{B}_{1i}+\cdots+\hat{X}_1\hat{X}_t\hat{B}_{1t}=B_{11}+(L_1+d_1)\hat{X}_1\\ \hat{X}_1\hat{X}_2\hat{B}_{11}+(1+\hat{X}_2^2)\hat{B}_{12}+\cdots+\hat{X}_2\hat{X}_i\hat{B}_{1i}+\cdots+\hat{X}_2\hat{X}_t\hat{B}_{1t}=B_{12}+(L_1+d_1)\hat{X}_2\\ \vdots\\ \hat{X}_1\hat{X}_t\hat{B}_{n1}+\cdots+\hat{X}_t\hat{X}_i\hat{B}_{ni}+\cdots+(1+\hat{X}_t^2)\hat{B}_{nt}=B_{nt}+(L_n+d_n)\hat{X}_t\\ \sum_{i=1}^{n}\hat{B}_{i1}^2\hat{X}_1+\sum_{i=1}^{n}\hat{B}_{i1}\hat{B}_{i2}\hat{X}_2+\cdots+\sum_{i=1}^{n}\hat{B}_{i1}\hat{B}_{it}\hat{X}_t=\sum_{i=1}^{n}(d_i+L_i)\hat{B}_{i1}\\ \vdots\\ \sum_{i=1}^{n}(\hat{B}_{it}\hat{B}_{i1})\hat{X}_1+\sum_{i=1}^{n}(\hat{B}_{it}\hat{B}_{i2})\hat{X}_2+\cdots+\sum_{i=1}^{n}\hat{B}_{it}^2\hat{X}_t=\sum_{i=1}^{n}(d_i+L_i)\hat{B}_{it}\end{array}\right. \tag{5}$$

由整理之后的前 $n\times t$ 个式子可以看出,获取了观测值和未知参数的平差值就可以得到系数矩阵的平差值;由后 t 个式子可以看出,获取了观测值和系数矩阵的平差值就可以得到未知参数的平差值。这是一个迭代的过程,具体操作为:① 获取未知参数的初值 $\boldsymbol{X}^0$;② 根据观测值信息以及未知参数初值 $\boldsymbol{X}^0$,由前 $n\times t$ 个式子求取系数矩阵平差值;③ 根据求得的系数矩阵和观测值信息,由后 t 个式子求取未知参数的平差值;④ 重复步骤②、③,直到两次计算的参数值之差小于一定的域值则停止迭代,输出结果。

对于估计公式中的后 t 个式子,可以写成矩阵形式:

$$\hat{\boldsymbol{B}}^{\mathrm{T}}\hat{\boldsymbol{B}}\hat{\boldsymbol{X}}=\hat{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{L} \tag{6}$$

估计公式中前 $n\times t$ 个式子可以表示为:

$$\boldsymbol{N}_b\hat{\boldsymbol{B}}^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}+\hat{\boldsymbol{X}}\boldsymbol{L}^{\mathrm{T}} \tag{7}$$

式中,$\boldsymbol{N}_b=\boldsymbol{E}+\hat{\boldsymbol{X}}\hat{\boldsymbol{X}}^{\mathrm{T}}$;$\boldsymbol{L}=[L_1-d_1\quad L_2-d_2\quad\cdots\quad L_n-d_n]^{\mathrm{T}}$。

2 精度评定

2.1 单位权中误差

单位权中误差的计算公式为 $\hat{\sigma}_0=\sqrt{\dfrac{\boldsymbol{V}^{\mathrm{T}}\boldsymbol{PV}}{f}}$,其中,$f$ 为自由度。平差后得到 $\boldsymbol{B}$ 与 $\boldsymbol{L}$ 的平差值,进而得到残差平方和为 $\sum\boldsymbol{V}^2=\sum(\hat{B}_{ij}-B_{ij})^2+\sum(\hat{L}_i-L_i)^2$,自由度 f 仍然等于观测值的个数减去必要观测的个数。代入到单位权中误差的计算公式中,即可求得 $\hat{\sigma}_0$。

2.2 未知参数精度评定

在得到参数的估值之后,还应给出所求参数的精度信息。但是,最后的参数估计方程是形式非常复杂的非线性方程。为了得到参数准确的误差信息,可以采用 Wolf 所给出的顾及函数二次项的误差传播公式[4]。下文以仅考虑一次项为例说明如何解决精度评定的问题。对于仅顾及一次项误差传播而言,首先应把待评定参数整理成观测信息的一次函数。上文推导得到了 $\boldsymbol{N}_b\hat{\boldsymbol{B}}^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}+\hat{\boldsymbol{X}}\boldsymbol{L}^{\mathrm{T}}$,$\hat{\boldsymbol{B}}^{\mathrm{T}}\hat{\boldsymbol{B}}\hat{\boldsymbol{X}}=\hat{\boldsymbol{B}}^{\mathrm{T}}\boldsymbol{L}$ 两个方程,这是一个隐函数方程组,直接线性化很困难,但它内在地确定了两个函数关系:

$$\begin{aligned}\hat{\boldsymbol{B}}&=F(\boldsymbol{B},\boldsymbol{L})\\ \hat{\boldsymbol{X}}&=G(\boldsymbol{B},\boldsymbol{L})\end{aligned} \tag{8}$$

$\hat{\boldsymbol{B}}$、$\hat{\boldsymbol{X}}$ 对观测信息 $\boldsymbol{L}$、$\boldsymbol{B}$ 的线性信息可以用 $\hat{\boldsymbol{B}}$、$\hat{\boldsymbol{X}}$ 和对 $\boldsymbol{L}$、$\boldsymbol{B}$ 的一次偏导数来表示,即

$$\begin{aligned}\mathrm{d}\hat{\boldsymbol{B}}&=\frac{\partial\hat{\boldsymbol{B}}}{\partial\boldsymbol{L}}\mathrm{d}\boldsymbol{L}+\frac{\partial\hat{\boldsymbol{B}}}{\partial\boldsymbol{B}}\mathrm{d}B\\ \mathrm{d}\hat{\boldsymbol{X}}&=\frac{\partial\hat{\boldsymbol{X}}}{\partial\boldsymbol{L}}\mathrm{d}\boldsymbol{L}+\frac{\partial\hat{\boldsymbol{X}}}{\partial\boldsymbol{B}}\mathrm{d}\boldsymbol{B}\end{aligned} \tag{9}$$

可以运用多元函数的隐函数求导法则来确定上文的 4 项偏导数。但如果要直接获取 4 个偏导数,根据隐函数求导的知识,需要对上文推导得到的 $t(n+1)$ 个式子同时求导,这时计算量很大。

为了简化运算,根据全微分不变定理,可以仍然把上面的式子按原有的分类分开求导,从式(6)中可以得到:

$$d\hat{\boldsymbol{B}} = \frac{\partial \hat{\boldsymbol{B}}}{\partial \boldsymbol{L}} d\boldsymbol{L} + \frac{\partial \hat{\boldsymbol{B}}}{\partial \boldsymbol{B}} d\boldsymbol{B} + \frac{\partial \hat{\boldsymbol{B}}}{\partial \boldsymbol{X}} d\hat{\boldsymbol{X}} \tag{10}$$

从式(7)中可以得到:

$$d\hat{\boldsymbol{X}} = \frac{\partial \hat{\boldsymbol{X}}}{\partial \boldsymbol{L}} d\boldsymbol{L} + \frac{\partial \hat{\boldsymbol{X}}}{\partial \boldsymbol{B}} d\boldsymbol{B} + \frac{\partial \hat{\boldsymbol{X}}}{\partial \hat{\boldsymbol{B}}} d\hat{\boldsymbol{B}} \tag{11}$$

这样求导可以减少解算各项偏导数的复杂度。根据全微分的不变性,式(10)与式(11)的表示与式(9)是等价的。在得到式(10)与式(11)之后,由于这两个式子都是线性的,$\boldsymbol{B}$、$\boldsymbol{L}$ 是测量数据,已知它们的误差信息,从而可以根据误差传播定律确定 $\hat{\boldsymbol{B}}$、$\hat{\boldsymbol{X}}$ 的误差信息。

以式(7)为例来说明求取偏导数的具体方法。对式(7)两端分别对 $\hat{X}_1$ 求导,有:

$$\begin{cases} 2\hat{X}_1\hat{B}_{11} + (1+\hat{X}_1^2)\dfrac{\partial \hat{B}_{11}}{\partial \hat{X}_1} + \hat{X}_2\hat{B}_{12} + \hat{X}_1\hat{X}_2\dfrac{\partial \hat{B}_{12}}{\partial \hat{X}_1} + \cdots + \hat{X}_t\hat{B}_{1t} + \hat{X}_1\hat{X}_t\dfrac{\partial \hat{B}_{1t}}{\partial \hat{X}_1} = L_1 + d_1 \\ \hat{X}_2\hat{B}_{11} + \hat{X}_1\hat{X}_2\dfrac{\partial \hat{B}_{21}}{\partial \hat{X}_1} + (1+\hat{X}_2^2)\dfrac{\partial \hat{B}_{22}}{\partial \hat{X}_1} + \cdots + \hat{X}_2\hat{X}_t\dfrac{\partial \hat{B}_{2t}}{\partial \hat{X}_1} = 0 \\ \vdots \\ \hat{X}_t\hat{B}_{n1} + \hat{X}_1\hat{X}_t\dfrac{\partial \hat{B}_{n1}}{\partial \hat{X}_1} + \hat{X}_2\hat{X}_t\dfrac{\partial \hat{B}_{n2}}{\partial \hat{X}_1} + \cdots + (1+\hat{X}_t^2)\dfrac{\partial \hat{B}_{nt}}{\partial \hat{X}_1} = 0 \end{cases}$$

上面的线性方程的系数矩阵是 $\boldsymbol{N}_b = \boldsymbol{E} + \hat{\boldsymbol{X}}\hat{\boldsymbol{X}}^{\mathrm{T}}$,$|\boldsymbol{N}_b| \neq 0$(可以先求出 $\boldsymbol{N}_b$ 的特征向量,用反证法求证 $\boldsymbol{N}_b$ 的可逆性),所以这个方程存在唯一解,从中可以解出 $\hat{B}_{ij}$ 对 $\hat{X}_1$ 的偏导数,同理可得 $\hat{B}_{ij}$ 对其他 $\hat{X}_i$ 的偏导。对于式(6)、式(7),两边同时对 B_{ij} 和 L_i 求偏导,从中可以解得所需的偏导数。

如果对式(9)继续求偏导可以得到相应的二阶偏导数,代入到参考文献[4]给出的顾及二次项的误差传播公式中,可以得到更准确的误差信息。

3　实例分析

3.1　拟合函数的确定

在实际的测量工作中,为了确定一条直线、一个平面或者曲面等几何图形,往往采用拟合的办法,即对待拟合图形采用一定的观测手段获取描述这个图形的观测数据,给出既定的函数模型,用最小二乘来确定函数模型中的参数。以直线拟合为例,观测数据为直线上的点坐标,传统的最小二乘往往是把其中一个方向上的坐标当作是没有误差的,这样处理显然是不合理的,求出的参数在统计角度来看是有偏的,此时应该采用整体最小二乘的方法进行拟合函数的确定。参考文献[1]给出的正交最小二乘实质就是整体最小二乘。下面以空间直线拟合为例,实例分析整体最小二乘在拟合函数确定方面的应用。根据参考文献[5]中给出的一组空间直线的实测数据,基于本文介绍的方法,编写了相应的计算程序,计算的方向矢量结果为 $\alpha = 0.267\ 179\ 8$,$\beta = 0.534\ 622$,$\gamma = 0.801\ 744\ 5$,与直线方向向量的真值 $\alpha = 0.267\ 261$,$\beta = 0.534\ 622$,$\gamma = 0.801\ 744$ 非常接近。

测点到直线的误差 Δ 很容易由计算结果得到,从结果中可以看出,$\sum\Delta = 0.023\ 9$,直线度为 0.007 6,结果优于参考文献[5]中的结果。

3.2　坐标转换参数的确定

以二维坐标系转换为例,分析整体最小二乘在坐标转换参数确定方面的应用。选取参考文献[6]中一组误差较大的公共点信息。根据文献中的结论,在该地区采用仿射变换模型较好。基于下文介绍的方法,编程进行了计算,计算过程中取 A、B、C、D、E、F、G、J、L、M、N、O 12 个点的公共坐标进行求参,取 H、I、K 点坐标作为检核。计算得到的结果见表 1、2。

表 1　整体最小二乘求得的转换参数/m

Tab.1　Transformation Parameters Cdculated by TLS/m

参数	a	b	c	d	e	f
参数值	1.000 000 967	−4.555 3E-06	−70.380 9	4.886E−06	1.000 001 016	−119.714 3

表 2　检核点转换坐标与已知坐标差值/mm

Tab.2　Cooridinate Difference of Check List Point/mm

检核点点号	坐标较差	
	X 方向	Y 方向
H	−6.1	8.3
I	−0.6	0.07
K	−6.8	6.8

将计算的结果与参考文献[4]中的结果进行比较可以看出,在仅考虑一组坐标有误差的情况下,点坐标的最大改正为 12 mm;而整体最小二乘的计算结果中,最大改正为 6.2 mm。从两种方法计算得到的单位权方差因子中也可以看出整体最

小二乘比经典最小二乘的统计优势(见表3)。

表3 不同方法的单位权方差因子/mm

Tab. 3 Unit Weight Variance Factor of Two Methods/mm

	整体最小二乘	常规最小二乘
单位权方差因子	3.095	6.177

4 结 语

传统整体最小二乘的解法十分复杂,限制了它的推广应用。本文给出的整体最小二乘的迭代解法计算简便,易于编程。同时,本文还给出了相应的精度评定方法,对整体最小二乘的推广是有意义的。

整体最小二乘解算后的精度评定还未得到广泛的研究,本文中提出的精度评定方法,首先通过对隐函数求导提取估计量对观测量的线性信息,然后通过误差传播定律估计误差。这从数据处理的角度完善了整体最小二乘的理论。

参 考 文 献

[1] 丁克良,欧吉坤,赵春梅. 正交最小二乘曲线拟合法[J]. 测绘科学, 2007,32(3):17-19

[2] 康正九,石毅明,胡保生. 整体最小二乘参数估计的并行算法[J]. 西安交通大学学报, 1998,32(8):1-4

[3] 黄开斌,俞锦成. 整体最小二乘问题的解集与极小范数解[J]. 南京师范大学学报(自然科学版)1997,20(4):1-5

[4] 胡圣武,陶本藻. 非线性模型的误差传播及其在GIS中的应用[J]. 武汉测绘科技大学学报,1997, 22(2):129-131

[5] 陈基伟. 工业测量数据拟合研究[D]. 上海:同济大学,2005

[6] 褚永海,田福娟,马晶. 局部坐标转换方法与应用[J]. 湖北水利水电职业技术学院学报, 2006,2(1):42-47

第一作者简介:孔建,硕士生,现从事测量数据处理与理论方面的研究。

E-mail:LiuHuKJ@163.com

Iterative Method for Total Least-Squares

KONG Jian[1] *YAO Yibin*[1] *WU Han*[1]

(1 School of Geodesy and Geomatics, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: The conventional solution method for TLS is based on matrix singular value decomposition. This method is rigorous in theory, but is complex and not easy to be programed. The complexity of the method is an important reason restricting TLS application in the field of Geomatics. By introducing the total least-squares adjustment standard, we derive the total least-squares iterative method, which is simple and easy to be programed. Through the introduction of multi-function derivative knowledge of the implicit function to determine the linear information of the parameters to observational data, we solve the problem of assessing the accuracy based on TLS. Finally, we apply the new method to the fitting function determination and coordinate transformation parameters determination based on measured data, and verify the feasibilities of the new method. The new method has a great significance to the popularization and application of TLS.

Key words: total least-squares; solving non-linear equation; iteration; fitting function; coordinate transformation

About the first author: KONG Jian, postgraduate, majors in measurement data processing and theory.

E-mail: LiuHuKJ@163.com