DOI:10.13203/j.whugis20130204　　文章编号:1671-8860(2014)12-1463-04

求解自回归模型参数的整体最小二乘新方法

姚宜斌[1]　黄书华[1]　陈家君[1]

1　武汉大学测绘学院, 湖北 武汉,430079

摘　要:在应用整体最小二乘法求解自回归模型的参数时,针对传统的 SVD 方法和迭代法并没有顾及到系数矩阵和观测向量构成的增广矩阵中不同位置上相同元素的改正数却不相同这一不足,推导了一种新的迭代解法,有效地解决了传统方法的不足,使得增广矩阵中不同位置的同一元素具有相同的改正数,更加符合实际情况且平差精度也有所提高。最后通过具体的算例,验证了本文方法的可行性和有效性。

关键词:整体最小二乘;自回归 AR 模型;迭代法

中图法分类号:P207.2　　**文献标志码**:A

整体最小二乘(TLS)法最早是由 GOLUB 和 Van Loan[1] 引入到超定线性方程组参数的求解中,其目的是为了解决观测方程中的系数矩阵也是观测值,即系数矩阵和观测值同时含有误差的问题。设有超定线性方程为:

$$(\boldsymbol{A}+\boldsymbol{E}_A)\boldsymbol{x}=\boldsymbol{b}+\boldsymbol{e}_b \tag{1}$$

式中,$\boldsymbol{A}\in\boldsymbol{R}^{n\times t}(n>t)$,$\boldsymbol{b}\in\boldsymbol{R}^{n\times 1}$ 都是含有误差的观测值;$\boldsymbol{x}\in\boldsymbol{R}^{t\times 1}$ 是待估参数向量;$\boldsymbol{E}_A\in\boldsymbol{R}^{n\times t}$ 是系数矩阵 A 的误差阵;$e_b\in\boldsymbol{R}^{n\times 1}$ 是观测值 b 的误差阵。为了求得参数 x 的最佳估值,使用的较为广泛的一种方法是基于目标方程:

$$\boldsymbol{e}_b^{\mathrm{T}}\boldsymbol{e}_b+\boldsymbol{e}_A^{\mathrm{T}}\boldsymbol{e}_A= \min \quad \text{s.t. } \boldsymbol{b}+\boldsymbol{e}_b-(\boldsymbol{A}+\boldsymbol{E}_A)\boldsymbol{x}=0 \tag{2}$$

式中,$\boldsymbol{e}_A=\mathrm{vec}(\boldsymbol{E}_A)$,$\mathrm{vec}(\cdot)$ 表示矩阵的拉直运算。建立拉格朗日方程式,然后对方程中的变量使用迭代法进行求导[2-3]。但是在使用上述方法时,要求每一个观测值的误差和系数矩阵中每一个元素的误差是相互独立的[4],这也就是要求观测值的误差和系数矩阵的误差不同源。但在有些情况下无法满足。例如在求解自回归(AR)模型的参数时,系数矩阵 A 和观测值 b 中的误差都是来自于同一类观测值,并且在由 A 和 b 组成的增广矩阵的不同位置有规律性地重复出现了同样的观测值。这时若仍然采用传统的 SVD 分解法[5] 或文献[2-3]中的迭代法进行解算,在不同位置的同一观测值将会有不同的改正数,这和实际情况是不相符的,在理论上也是不严密的。为解决这一问题,本文推导了一种适合自回归(autoregressive,AR)模型参数求解的新迭代算法。

1　AR 模型参数估计新方法的算法推导

在变形分析与预报中,常采用 AR 模型来建立变形模型[6],其实质是根据过往自身的变化规律来建立可以预测其未来发展趋势的函数模型。设有一个 t 阶自回归模型,其模型可表示为:

$$\begin{cases} z_{t+1}=\varphi_1 z_t+\varphi_2 z_{t-1}+\cdots+\varphi_t z_1 \\ z_{t+2}=\varphi_1 z_{t+1}+\varphi_2 z_t+\cdots+\varphi_t z_2 \\ \vdots \\ z_{t+n}=\varphi_1 z_{t+n-1}+\varphi_2 z_{t+n-2}+\cdots+\varphi_t z_n \end{cases} \tag{3}$$

将其表示成矩阵形式可写成:

$$\boldsymbol{b}=\boldsymbol{A}\boldsymbol{x} \tag{4}$$

式中,$\boldsymbol{b}=\begin{bmatrix} z_{t+1} \\ z_{t+2} \\ \vdots \\ z_{t+n} \end{bmatrix}$;$\boldsymbol{A}=\begin{bmatrix} z_1 & z_2 & \cdots & z_t \\ z_2 & z_3 & \cdots & z_{t+1} \\ \vdots & \vdots & \ddots & \vdots \\ z_n & z_{n+1} & \cdots & z_{t+n-1} \end{bmatrix}$;

$\boldsymbol{x}=[\varphi_t \quad \varphi_{t-1} \quad \cdots \quad \varphi_1]^{\mathrm{T}}$。考虑系数矩阵 A 和观测值向量 b 的误差,可以将式(4)改写为式(1),所以误差矩阵的增广矩阵形式可以写成:

收稿日期:2013-06-06

项目来源:国家自然科学基金资助项目(41174012、41274022);国家 863 计划资助项目(2013AA122502);教育部新世纪优秀人才支持计划资助项目(NCET-12-0428)。

第一作者:姚宜斌,博士,教授,现主要从事测量数据处理基础理论与方法研究。E-mail:ybyao@whu.edu.cn

$$[\boldsymbol{E}_A \quad \boldsymbol{e}_b] = \begin{bmatrix} \Delta z_1 & \Delta z_2 & \cdots & \Delta z_t & \Delta z_{t+1} \\ \Delta z_2 & \Delta z_3 & \cdots & \Delta z_{t+1} & \Delta z_{t+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Delta z_n & \Delta z_{n+1} & \cdots & \Delta z_{t+n-1} & \Delta z_{t+n} \end{bmatrix}$$

其中 $\Delta z_i (i=1,2,\cdots,t+n)$ 表示增广矩阵中每个元素对应的误差项。由此可见，采用 TLS 方法处理 AR 模型是一种非常特殊的情况，其观测向量误差和矩阵误差同源，而且误差增广矩阵中不同位置有规律性地重复相同元素。

在一般整体最小二乘算法中，不特别考虑误差增广矩阵中不同位置具有相同元素这一事实，直接使用 SVD 分解法进行参数求解或者使用式(2)构建拉格朗日方程[2-3]：

$$\Phi(\boldsymbol{e}_b, \boldsymbol{e}_A, \boldsymbol{\lambda}, \boldsymbol{x}) = \boldsymbol{e}_b^{\mathrm{T}} \boldsymbol{e}_b + \boldsymbol{e}_A^{\mathrm{T}} \boldsymbol{e}_A + 2\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b} + \boldsymbol{e}_b - \boldsymbol{A}\boldsymbol{x} - \boldsymbol{E}_A \boldsymbol{x}) \tag{5}$$

这种处理将导致同一个观测值求得的误差改正数不同，这与实际情况是不相符的。为了使误差矩阵的增广矩阵 $[\boldsymbol{E}_A \quad \boldsymbol{e}_b]$ 中的相同元素具有相同的改正数，可将其中不同的元素单独取出，构成新的向量 $\Delta z = [\Delta z_1 \quad \Delta z_2 \quad \cdots \quad \Delta z_n \quad \cdots \quad \Delta z_{t+n}]^{\mathrm{T}}$，此时目标函数可以写成：

$$(\Delta z)^{\mathrm{T}} \boldsymbol{W} \Delta z = \min \tag{6}$$

式中，对角阵 $W \in \boldsymbol{R}^{(n+t)\times(n+t)}$ 类似于权阵，其主对角线上的元素依次为 $\Delta z_i (i=1,2\cdots n+t)$ 在增广矩阵中出现的次数。之所以考虑 Δz_i 出现的次数，是因为在利用时间序列数据建立 AR 模型时，同一个观测值可能会依据 AR 模型的阶数按不同的次数出现在多个不同的观测方程中，也即把同一个观测值当做了多次相同观测参与建模。为了将向量 Δz 引入到观测方程中，可以将式(1)改写成如下形式：

$$\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} = \boldsymbol{E}_A \boldsymbol{x} - \boldsymbol{e}_b = [\boldsymbol{E}_A \quad \boldsymbol{e}_b] \begin{bmatrix} \boldsymbol{x} \\ -\boldsymbol{1} \end{bmatrix} = \boldsymbol{C} \times \Delta z \tag{7}$$

式中，$\boldsymbol{C} = \begin{bmatrix} \varphi_t & \varphi_{t-1} & \cdots & \varphi_1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & \varphi_t & \varphi_{t-1} & \cdots & \varphi_1 & -1 & 0 & \cdots & 0 \\ 0 & 0 & \varphi_t & \varphi_{t-1} & \cdots & \varphi_1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \varphi_t & \varphi_{t-1} & \cdots & \varphi_1 & -1 \end{bmatrix} \in \boldsymbol{R}^{n\times(n+t)}$。

构造出和式(2)相类似的目标方程：

$$(\Delta z)^{\mathrm{T}} \boldsymbol{W} \Delta z = \min \quad \mathrm{s.t.} \quad \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} - \boldsymbol{C} \times \Delta z = 0 \tag{8}$$

利用式(8)可以构成拉格朗日方程为：

$$\Phi(\Delta z, \boldsymbol{x}, \boldsymbol{\lambda}) = (\Delta z)^{\mathrm{T}} \boldsymbol{W} \Delta z + 2\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} - \boldsymbol{C} \times \Delta z)$$

对三个参数分别求导后写成矩阵形式：

$$\begin{bmatrix} \boldsymbol{W} & 0 & \boldsymbol{C}^{\mathrm{T}} \\ 0 & 0 & \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{C} & \boldsymbol{A} & 0 \end{bmatrix} \begin{bmatrix} \Delta z \\ \boldsymbol{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{b} \end{bmatrix} \tag{9}$$

实际上式(8)是和式(2)等价的，只是表示的形式不同，通过权阵的形式避免了目标方程中重复元素的出现。但正是因为这两者表示的形式的不同，可以解决增广矩阵中同一元素含有不同改正数的问题。从上面的推导过程可以看出，Δz 中包含了系数矩阵 A 和观测值 b 的改正数 $\boldsymbol{E}_A$ 和 $\boldsymbol{e}_b$，矩阵 C 是由最小二乘法求得参数 x 中对应元素构成的。所以要求得式(9)的稳定解必须采用迭代法，具体解算步骤可总结如下为：① 采用最小二乘法求得参数的初始解 $x_0 = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}} b$，利用求得的 x 可以构造出矩阵 C 的初始值；② 将 C 代入到式(9)中，可以求得参数 Δz 和 x 的值；③ 求得参数 Δz 之后，矩阵 A 和向量 b 加上 Δz 中对应的改正项，重复步骤②可以求得新的 Δz 和 x 的值；④ 重复步骤②、③，直至相邻两次求得的 Δz 和 x 的值之差小于一个阈值 ε 为止。

使用上述迭代步骤可以求得参数 x 的平差值，同时可以得到增广矩阵中各不同元素的改正数 Δz。本文方法平差模型的自由度为 $(n+n\times t)-(t+n\times t)=n-t$，和经典最小二乘是一致的，因此单位权中误差计算公式为：

$$\delta_0 = \sqrt{\frac{(\Delta z)^{\mathrm{T}} \boldsymbol{W} \Delta z}{n-t}} \tag{10}$$

2 实例与分析

现以某建筑物某个时段定期进行的 36 次沉降观测的高程观测值[7]为例，具体观测值如表 1 所示。

由文献[7]中的分析可知对上表中的数据应建立一个阶数 $t=3$ 的自回归模型，本例中的 W 阵为：

$$\boldsymbol{W} = \mathrm{diag}(1,2,3,4,\underbrace{4\cdots\cdots4}_{(m+n)-2(n+1)},4,3,2,1)$$

表 1　沉降观测数据/mm

Tab. 1　Observation Data of Settlement/mm

序数	高程	序数	高程	序数	高程	序数	高程	序数	高程	序数	高程
1	26.33	7	25.93	13	26.67	19	28.09	25	26.81	31	26.81
2	26.27	8	26.43	14	27.95	20	26.78	26	28.50	32	28.50
3	26.43	9	26.52	15	26.74	21	28.66	27	27.68	33	27.68
4	25.56	10	25.46	16	27.53	22	26.75	28	26.57	34	26.57
5	26.82	11	26.12	17	25.31	23	27.24	29	28.36	35	28.36
6	26.56	12	27.28	18	26.90	24	28.02	30	27.94	36	27.94

利用表 1 中的数据可以建立一个形如式(3)的三阶自回归方程，现分别采用传统最小二乘法、整体最小二乘的 SVD 分解法和本文的方法(阈值 $\varepsilon=0.000\ 000\ 1$)对建立的模型进行参数估计，求得的结果如表 2 所示。为了比较三种结果的精度，求得各种方法的单位权中误差也列于表 2 中。

从表 2 中的结果可以发现，在应用本文方法求解 AR 模型参数时，单位权中误差明显要小于 LS 法和 SVD 法。其原因主要是本文所采用的方法不仅考虑了观测方程中系数矩阵的误差，并且顾及到了增广矩阵中同一元素应该有相同的改正数这一事实，更加符合实际情况。为了更加直观地比较三种方法求得的参数值，现将三种方法求得的观测向量的平差值和原始数据 b 进行比较，如表 3 所示。

表 2　三种方法求得的参数值及单位权中误差/mm

Tab. 2　The Parameter Value and the Error of Unit Weight of Three Methods/mm

使用的方法	LS 法	SVD 法	本文方法
φ_3	0.635 059	1.090 148	0.634 924
φ_2	0.327 809	0.289 835	0.328 491
φ_1	0.041 087	−0.374 782	0.041 870
单位权中误差	0.804 749	0.604 558	0.147 438

表 3　三种方法求得的观测量平差值与原始数据/mm

Tab. 3　Adjustment of Observation Values and the Original Data of Three Methods/mm

序数	原始值	LS 法	TLS 法	本文方法	序数	原始值	LS 法	TLS 法	本文方法
4	25.56	26.419	26.412	26.454	21	28.66	27.392	27.430	27.428
5	26.82	26.397	26.719	26.432	22	26.75	27.795	27.643	27.832
6	26.56	26.265	26.169	26.300	23	27.24	27.501	27.475	27.538
7	25.93	26.115	25.683	26.151	24	28.02	28.089	28.788	28.125
8	26.43	26.804	27.218	26.839	25	26.81	27.069	26.555	27.106
9	26.52	26.453	26.564	26.488	26	28.50	27.586	27.769	27.622
10	25.46	26.221	25.989	26.256	27	27.68	27.754	27.635	27.791
11	26.12	26.524	26.957	26.559	28	26.57	27.506	27.113	27.543
12	27.28	26.261	26.501	26.295	29	28.36	28.265	29.134	28.300
13	26.67	25.852	25.102	25.888	30	27.94	27.454	27.247	27.490
14	27.95	26.626	26.386	26.662	31	26.81	27.318	26.714	27.356
15	26.74	27.215	26.994	27.252	32	28.50	28.271	28.967	28.307
16	27.53	27.198	27.153	27.234	33	27.68	27.703	27.548	27.740
17	25.31	27.647	27.902	27.683	34	26.57	27.506	27.113	27.543
18	26.90	27.046	27.644	27.081	35	28.36	28.265	29.134	28.300
19	28.09	26.885	27.266	26.920	36	27.94	27.454	27.247	27.490
20	26.78	26.046	24.861	26.082					

从表 3 中可以发现，使用本文方法求得的观测量的平差值和原始数据的变化趋势最为相符，而整体最小二乘的结果却是最差的。由于本文方法中，目标方程中只取出其中不同的元素，相同的元素通过类似于权阵的形式表示出来，这就使得增广矩阵中相同的元素有了同样的改正数。所以本文方法不仅完善了 TLS 法求解 AR 模型参数的理论，也使得该方法在实际中变得更为实用。

为了使上述结论具有一般性，采用 MATLAB 软件仿真了一个三阶 AR 模型：

$$z(i)=0.1\cdot z(i-1)+0.2\cdot z(i-2)+0.6\cdot z(i-3)+a(i)$$

其中，$a(i)$ 表示正态白噪声。限于篇幅就不将仿真的原始数据列出，现分别采用三种不同的方法对仿真数据进行参数的求解，将求得的结果与真值对比于表 4 中。

从表 4 中可以看到，本文方法求得的参数值较其他方法更接近于真值，其他结论和上述变形监测实例是一致的。进一步实验可以发现，当不加入白噪声 $a(i)$ 时，使用本文方法求得的参数值

表4 三种方法求得的参数值与真值比较

Tab.4 The Parameters of Three Methods Compared with the True Value

使用的方法	LS法	SVD法	本文方法	真值
φ_3	0.515 652	1.139 847	0.577 932	0.6
φ_2	0.232 248	0.060 997	0.273 974	0.2
φ_1	0.133 252	−0.171 805	0.107 639	0.1
单位权中误差	1.125 063	0.895 325	0.796 322	

和真值是一样的。这表明本文方法可以较好地收敛于真值点，也进一步说明使用该方法进行自回归模型参数估计的有效性和可行性。

3 结 语

在对AR模型进行参数估计时，传统的最小二乘法因没有考虑系数矩阵的误差而存在着明显的不足。当采用TLS法时，虽然解决了传统最小二乘的缺陷，但是又有新的问题出现，即增广矩阵中不同位置的同一元素的改正数不相同，这和实际情况是不相符的。本文给出的算法不仅可以同时克服上述两个问题，而且求得的参数也具有更高的精度。这使得应用TLS法求解AR模型参数的理论变得更加完善。

参 考 文 献

[1] Markovsky I, Van Huffel S. Overview of Total Least-squares Methods[J]. *Signal Processing*, 2007, 87(10):2 283-2 302

[2] Schaffrin B, Felus Y A. An Algorithmic Approach to the Total Least-Squares Problem with Linear and Quadratic Constraints[J]. *Studia Geophysica et Geodaetica*, 2009,53(1):1-16

[3] Schaffrin B. A Note on Constrained Total Least-Squares Estimation[J]. *Linear Algebra and its Application*, 2006,417(1):245-258

[4] Van H S, Zha Hongyuan. The Total Least Squares Problem[J]. *Handbook of Statistics*, 1993:377-408

[5] Golub G H, Van Loan C F. An Analysis of the Total Least Squares Problem[J]. *SIAM J Numer Anal*, 1980,17:883-893

[6] Zhao Xinxiu, Wang Jiexian. Application of Kalman Filter Method Based on AR(P) Model in GPS Deformation Prediction[J]. *GNSS World of China*, 2009,5:42-44(赵新秀，王解先. 基于AR(P)的卡尔曼滤波在GPS变形监测中的应用[J]. 全球定位系统，2009,5:42-44)

[7] Wang Xinzhou, Tao Benzao, Qiu Weining, et al. Higher Surveying Adjustment[M]. Beijing: Surveying and Mapping Press, 2006(王新洲，陶本藻，邱卫宁，等. 高等测量平差[M]. 北京：测绘出版社，2006)

A New Method of TLS to Solving the Autoregressive Model Parameter

YAO Yibin[1] *HUANG Shuhua*[1] *CHEN Jiajun*[1]

1 School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China

Abstract: When using total least squares to solve the regression model parameter, both the traditional SVD and iteration methods do not consider that corrections of the same elements in different positions of the augmented matrix formed by the coefficient matrix and the observation vector are different. This paper proposes a new iteration method which can effectively solve the shortage problem in traditional methods, puts the same element in different positions of an augmented matrix into the same correction. This method is more compatible to the actual situation. Adjustment precision can also be improved. Finally, a concrete example was conducted to verify the feasibility and effectiveness of this method.

Key words: total least squares; AR model; iterative method

First author: YAO Yibin, PhD, professor, specializes in measurement data processing theory and methods. E-mail: ybyao@sgg.whu.edu.cn

Foundation support: The National Natural Science Foundation of China, Nos. 41174012, 41274022; the National High Technology Research and Development Program of China(863 Program), No. 2013AA122502; the New Century Excellent Talents in University, No. NCET-12-0428.