

DOI :10.3876/j.issn.1000-1980.2011.02.009

G-M 模型病态问题的多步解法

蒋敏卫

(河海大学地球科学与工程学院 江苏 南京 210098)

摘要:针对 G-M 模型的病态问题,分析了导致模型病态的原因和各种解决办法,重点研究了两步解法的计算过程.通过对两步解法的借鉴和引申,提出了克服模型病态性的新方法——多步解法.模拟了 1 个病态问题算例,应用 maple 数值计算软件,采用不同的方法计算该模型的参数估值,比较它们计算结果的优劣.结果表明,多步解法比两步解法计算精度更高,并且其计算步数与阈值的大小有关.

关键词:多步解法;病态问题;G-M 模型;两步解法;正则化理论

中图分类号:P228 文献标志码:A 文章编号:1000-1980(2011)02-0165-04

病态问题是测量数据处理中长期关注和研究的问题.病态是相对于良态而言的.对于一个参数估计模型,如果其数据的微小变化会引起解的巨大变化,则称该模型是病态的,否则就是良态的.测量平差系统病态问题的产生,既涉及现代数理统计、最优化理论、计算方法等理论基础,又与各种具体平差问题的内在因素,如控制点分布、几何结构、观测精度、参数选择等有关.而参数模型的病态效应主要体现在计算和统计 2 个方面^[1],如估值远远偏离真值,或者估值的均方误差很大.

如果参数模型是病态的,那么由此计算得到的参数估值将变得很不可靠.因此,很多学者对这一问题给予了长期的关注,在解算方法方面,有一批重要成果.为了消除或减弱 G-M 模型的病态性,本文分析了 G-M 病态问题的机理,并在原有解法的基础上,提出了一种新的解法——多步解法.

1 G-M 模型的病态性

测量平差中最基本的数学模型为 Gauss-Markov 模型(简称 G-M 模型)^[2],该模型也是病态问题的基本数学模型.G-M 模型为

$$\begin{cases} L = AX + \Delta \\ E(\Delta) = 0 \quad D(\Delta) = \sigma_0^2 P^{-1} \end{cases} \quad (1)$$

式中: L —— $n \times 1$ 观测向量; A —— $n \times m$ 阶设计矩阵, $\text{rank}(A) = m$; X —— $m \times 1$ 未知参数向量; Δ —— $n \times 1$ 观测误差向量, $\Delta \sim N(0, \sigma_0^2 I)$; σ_0^2 ——单位权方差; P ——权矩阵; $E(\Delta)$ ——误差期望值.

由式(1)可得 G-M 模型的最小二乘估计(LS 估计)解为

$$\begin{aligned} \hat{X} &= N^{-1}W = (A^T P A)^{-1} A^T P L \\ N &= A^T P A \quad W = A^T P L \end{aligned} \quad (2)$$

式中 N 为法矩阵.

根据 Gauss-Markov 定理,在观测值服从正态分布的前提下,LS 估计量 $\hat{X}$ 是 X 的最优线性无偏估计量,因此,LS 估计一直是被广泛采用的重要的估计方法.但是,在测量实际应用中,由于参数选择、观测和计算方法等原因,LS 估计会变得很不稳定,精度较差.

在式(2)中,当 N 非奇异时,存在唯一解.假设 N, W 有微小的扰动 $\delta N, \delta W$,引起估计量 $\hat{X}$ 有偏差 $\delta \hat{X}$,可表达为

$$\hat{X} + \delta \hat{X} = (N + \delta N)^{-1} (W + \delta W) \quad (3)$$

根据 Banach 引理的推论可知,只要 $\|N^{-1}\| \|\delta N\| < 1$,则矩阵 $N + \delta N$ 总是非奇异的.显然 $\delta N, \delta W$ 对 $\delta \hat{X}$ 的影响与 N^{-1} 密切相关.因此,判断参数估计模型是否病态,需要考虑法矩阵 N .克服模型病态性的方法之一就是要在一定的准则条件下通过改变法矩阵 N 的状态而使得模型病态性减弱或变成良性.

2 病态问题的两步解法

克服病态问题的方法有多种,其中,Stein 估计^[3]、岭估计^[4]、广义岭估计和截断奇异值法^[5]是常用的几种方法,Householder 正交变换法、约束条件法和谱修正迭代法^[6]等是目前较新的几种方法.针对模型的病态问题,当前研究的重点之一就是以上这些解算方法的深入分析和改进,或者将以上几种方法相结合而提出一些能更好地克服模型病态性的方法,其基本思路是通过正则化方法将病态问题转化为病态性较轻或良性问题^[7-8],然后再进行解算.

王振杰等^[8]对测量中最基本的 G-M 模型的病态问题进行了深入研究,提出了一种新的克服模型病态性的方法——两步解法^[9].两步解法是由 Tikhonov 正则化理论^[8]推导而来的.对于 G-M 模型,Tikhonov 正则化理论的估计准则为

$$\|A\hat{X} - L\|^2 + \alpha\chi(\hat{X}) = \|A\hat{X} - L\|^2 + \alpha\hat{X}^T R \hat{X} = \min \tag{4}$$

式中: α ——正则化参数; R ——正则化矩阵; $\chi(\hat{X})$ ——稳定泛函; $\|\cdot\|$ ——2-范数.

两步解法的第 1 步计算实质上是岭估计,正则化矩阵 R_1 是单位阵.根据 Tikhonov 正则化理论,由式(4)可得解及其均方误差矩阵^[10]为

$$\hat{X}_1 = (A^T A + \alpha_1 R_1)^{-1} A^T L \tag{5}$$

$$\text{MSEM}(\hat{X}_1) = \sigma_0^2 (A^T A + \alpha_1 I)^{-1} \tag{6}$$

$$R_1 = I$$

式中: α_1 ——正则化参数,由 L 曲线法^[11]求得; σ_0^2 ——单位权方差.

两步解法的第 2 步计算实质上是广义岭估计,此时的正则化矩阵 R_2 不是单位阵,而是由第 1 步计算的结果确定的.根据 Tikhonov 正则化理论,由式(4)可得解及其均方误差矩阵^[10]为

$$\hat{X}_2 = (A^T A + \alpha_2 R_2)^{-1} A^T L \tag{7}$$

$$\text{MSEM}(\hat{X}_2) = \sigma_0^2 (A^T A + \alpha_2 R_2)^{-1} \tag{8}$$

式中 α_2 由 L 曲线法^[11]求得,令

$$R_2 = \text{diag}(\text{MSEM}(\hat{X}_1)^{-1}) = \text{diag}\left(\frac{1}{\sigma_0^2}(A^T A + \alpha_1 I)\right) \tag{9}$$

需要说明的是,以上计算结果是在权矩阵 P 为单位阵时求得的,当权矩阵 P 不是单位阵时,可将其转换为单位阵^[12].在第 2 步计算中,由于正则化矩阵 R_2 不再是单位阵,此时用 L 曲线法确定 α_2 时,需要进行一个转换^[13].

通过分析两步解法的计算过程可以看到,两步解法的第 2 步是在第 1 步的基础上,对法矩阵进行进一步的优化,利用第 1 步得到的均方误差阵的逆的主对角元素作为第 2 步的正则化矩阵的元素,从而再次改变法矩阵的状态,使得参数估计精度更高.由此,可以深入下去,通过更多相似步骤的计算,而改进法矩阵的状态.

3 病态问题的多步解法

多步解法是在两步解法的基础上,通过多次对法矩阵的优化而达到某种精度要求的病态问题的参数估计方法,这里的计算次数取决于要达到的精度,可能是 2 次,也可能是 2 次以上.多步解法的计算流程见图 1.图 1 中: $\|\cdot\|$ 为 ∞ 范数; b 是相邻两次计算所得估计参数差值的绝对值的最大值; $\hat{X}_k$ 为最终参数

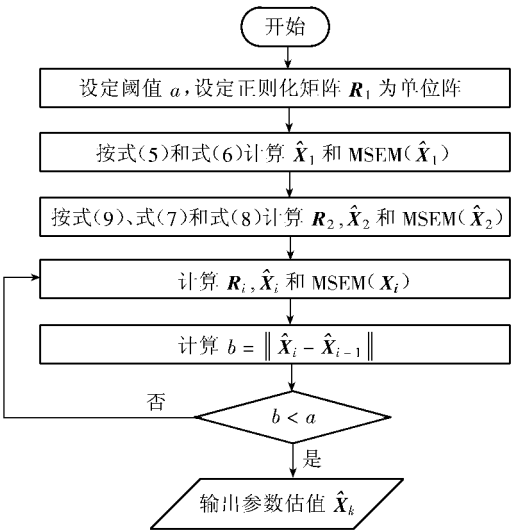


图 1 多步解法的计算流程
Fig. 1 Flow chart of multi-step solution

估值.在此,令

$$R_i = \text{diag}(\text{MSEM}(\hat{X}_{i-1})^{-1}) \tag{10}$$

多步解法和两步解法有很多相似之处:两者都是根据正则化理论通过改变法矩阵的状态来进行参数估计的,两者第 1 步计算时的正则化矩阵都是单位矩阵,其余所有的正则化矩阵都是利用上一步计算结果得到的,两者每一步的正则化参数都是由 L 曲线法求得.多步解法与两步解法的不同之处是:前者比后者的计算步骤更多,精度更高,前者要设定一个阈值,以控制计算过程,而后者不要.需要注意的是:阈值要设定适当.若阈值太小,计算步骤会很多,甚至无法结束;若阈值太大,计算会很快结束但精度太低.一般情况下,可以设阈值为参数估计所要求达到的精度值,如参数估计值的精度要求达到 0.002 m,则可以设阈值 a 为 0.002.

多步解法还具有如下优良性质:(a)多步解法是 LS 估计的一个线性变换,并且是一种压缩型有偏估计;(b)多步解法的解比 LS 估计的解具有更强的抗干扰性能,观测数据的扰动对它的影响比后者小;(c)在一定条件下,多步解法的解在均方误差意义下优于 LS 估计.

这些性质的证明与两步解法的证明^[8]相似.

4 算例分析

设计一个 G-M 模型的模拟病态问题,未知参数有 3 个,其真值为 $X=(10,15,6)^T$,设计矩阵为

$$A = \begin{bmatrix} 1 & 3.6 & 2.4 & 1 & 3.5 & -1 & 5 & 1 & 4 & 3 \\ 2 & 1 & 1.5 & 2 & 1 & 3 & 0.5 & 2 & 1 & 1 \\ 4 & 2.1 & 3 & 3.9 & 2 & 6 & 1.1 & 4.1 & 1.9 & 2 \end{bmatrix}^T$$

观测值权矩阵 P 为单位阵,观测向量为 $L=(64.5,64,64,63,62.7,71.1,64,64,66.9,57.4)^T$,观测误差 $\Delta \sim N(0,\sigma^2 I)$, $\sigma=1$.为了分析多步解法的一些特性,需要观察尽可能多步的计算结果,所以,在此设定一个很小的阈值 $a=0.00001$,并取前 9 步计算结果列于表 1 中.

表 1 多步解法的计算结果

Table 1 Calculated results of multi-step solution

步数	$\hat{X}_1$	$\hat{X}_2$	$\hat{X}_3$	正则化参数	$\ \hat{X}_i - \hat{X}_{i-1}\ $	$\ \hat{X}_i - X\ _2$
1	10.0260	6.0511	10.4504	0.16470		9.99445
2	10.0379	13.7100	6.6165	0.00305	7.65886	1.43025
3	10.0383	13.7419	6.6007	0.00301	0.03194	1.39463
4	10.0377	13.7356	6.6036	0.00308	0.00636	1.40163
5	10.0380	13.7392	6.6019	0.00304	0.00364	1.39763
6	10.0378	13.7365	6.6032	0.00307	0.00274	1.40064
7	10.0380	13.7392	6.6019	0.00304	0.00274	1.39763
8	10.0379	13.7374	6.6028	0.00306	0.00183	1.39964
9	10.0377	13.7356	6.6036	0.00308	0.00182	1.40164

分析表 1 中的计算结果,可以发现以下几个特性:(a) $\|\hat{X}_i - X\|_2$ 值第 3 步明显小于第 2 步,从第 3 步开始, $\|\hat{X}_i - X\|_2$ 值基本保持稳定;(b)从第 2 步开始,正则化参数的大小变化不大;(c) $b = \|\hat{X}_i - \hat{X}_{i-1}\|$ 值随着计算步骤的增加呈逐渐减小的趋势,虽然并不具有递减的性质,但其变化幅度趋于平缓;(d)若设定阈值 $a=0.01$,则第 4 步的 b 值小于 a ,计算过程将到第 4 步结束;若设定阈值 $a=0.002$,则计算过程将到第 8 步结束.

由此,可以得到以下几点结论:(a)多步计算过后,变化相对稳定的参数估值为该解法下的最优近似解;(b)多步解法比两步解法的参数估计精度更高;(c)多步解法的计算步数与阈值 a 的大小有关,阈值越小则计算步数就越多.

5 结 语

克服病态问题的解法有很多种,通过改变法矩阵的状态来减弱模型的病态性是重要的途径之一,而两步解法便是其中较新的一种方法.两步解法是根据正则化理论通过两步计算而达到比岭估计方法精度较高的解决病态问题的参数估计方法.本文提出的多步解法则是在两步解法的基础上,借鉴两步解法的思想,通过更多步计算,而使得参数估计精度更高的一种减弱或消除模型病态性的方法,它是对两步解法的进一步的延

伸和提高. 本文仅是在大量模拟计算的基础上对多步解法的一些特性做了初步的研究. 通过算例分析可知, 多步解法具有较好的参数估计特性, 值得更深入的探讨和理论研究.

参考文献:

[1] 戴伯新. 共线性的诊断与评价[J]. 数学的实践与认识, 1989, 19(4) : 53-61. (DAI Bo-xin. Diagnosis and evaluation of collinearity [J]. Mathematics in Practice and Theory, 1989, 19(4) : 53-61. (in Chinese))

[2] 陈希孺. 线性模型参数的估计理论[M]. 北京: 科学出版社, 2010.

[3] 谢宇. 回归分析[M]. 北京: 社会科学文献出版社, 2010.

[4] HOERL A E, KENNARD R W. Ridge regression biased estimation for non-orthogonal problem[J]. Technometrics, 1970, 12(1) : 55-67.

[5] WIGGINS R A. The generalized linear inverse problem: implication of surface waves and free oscillations for earth structure[J]. Reviews of Geophysics and Space Sciences, 1972, 10(10) : 251-285.

[6] 王新洲. 在无偏估计类中改进最小二乘估计的方法[J]. 武汉测绘科技大学学报, 1995, 20(1) : 46-50. (WANG Xin-zhou. A method of improved least-squares estimation in class of unbiased estimator[J]. Journal of Wuhan University of Science and Technology, 1995, 20(1) : 46-50. (in Chinese))

[7] 刘继军. 不适定问题的正则化方法及应用[M]. 北京: 科学出版社, 2008.

[8] TIKHONOV A N, ARSENIN V Y. Solutions of ill-posed problems[M]. New York: Wiley, 1997.

[9] 王振杰, 欧吉坤, 柳林涛. 一种解算病态问题的方法: 两步解法[J]. 武汉大学学报: 信息科学版, 2005, 30(9) : 821-823. (WANG Zhen-jie, OU Ji-kun, LIU Lin-tao. A method of resolving ill conditioning problem: two-step solution[J]. Journal of Wuhan University: Information Science, 2005, 30(9) : 821-823. (in Chinese))

[10] 杨文采. 地球物理反演和地震层析成像[M]. 北京: 地质出版社, 1989.

[11] HANSEN P C. Analysis of discrete ill-posed problems by means of the I-curve[J]. SIAM Review, 1992, 34(4) : 561-580.

[12] 王穗辉. 误差理论与测量平差[M]. 上海: 同济大学出版社, 2010.

[13] HANSEN P C. Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion[M]. Philadelphia: SIAM, 1998.

Multi-step solution to ill-posed problems of Gauss-Markov model

JIANG Min-wei

(College of Geoscience and Engineering, Hohai University, Nanjing 210098, China)

Abstract: With regard to the ill-posed problem of the Gauss-Markov model, the causes and various methods were analyzed, and the computational process of the two-step solution was highlighted. Based on the extension of the two-step solution, a new method for solving the ill-posed problems was proposed, that is, multi-step solution. An example of the ill-posed model was simulated. Various methods were employed to calculate the parametric values of this model by use of numerical software maple, and the calculated results were compared. It is concluded that the multi-step solution has higher precision than the two-step solution, and the calculation steps are related to the magnitude of the threshold.

Key words: multi-step solution; ill-posed problem; Gauss-Markov model; two-step solution; regularization theory