

一种基于丹麦法的改进型双步 M 估计

林国标^{1,2,3} 刘立龙^{1,2,3} 蔡成辉^{1,2,3} 黎峻宇^{1,2,3} 黄良珂^{1,2,3}

1 广西矿冶与环境科学实验中心,桂林市建干路 12 号, 541004
2 桂林理工大学测绘地理信息学院,桂林市建干路 12 号, 541004
3 广西空间信息与测绘重点实验室,桂林市建干路 12 号, 541004

摘 要: 针对抗差估计方法初值与临界值选取不当的问题,提出一种以 L_1 -范估计的平差值作为初值、以改进的丹麦法权函数作为第一步抗差阶段权函数的双步 M 估计,该方法既具有较强的抗差性,又具有最小二乘的最优性。试验表明,该方法估计结果与无粗差时的 LS 估计结果基本一致,抗差效果明显优于传统抗差估计方法与其他两步抗差估计。

关键词: 丹麦法;双步 M 估计;临界值;均方差
中图分类号: P207 **文献标识码:** A

抗差估计的抗差性及有效性主要取决于初值的准确性与权函数的合理性^[1]。目前,一般选取最小二乘估计的结果作为抗差估计的初值^[2],但最小二乘估计对粗差具有均衡性和不敏感性,致使粗差观测值的残差并非最大,导致选权迭代时出现错误的判断。同时,合理的权函数应该分 3 段:正常段、可疑段、淘汰段,其临界值和均方差因子应根据实际情况选取,具有可变性^[1,3]。在传统的抗差估计方法中,Andrews 法、Tukey 法不含正常段,Huber 法、丹麦法没有淘汰段,IGG 法、Hampel 法虽然具有正常段、可疑段及淘汰段,但其临界值往往取某一固定的先验值。这些方法仅采用一步抗差估计^[4],具有一定的局限性。

针对以上问题,本文提出一种基于丹麦法的改进型双步 M 估计方法。采用抗差性较强的 L_1 -范估计的平差值作为抗差估计的初值,在丹麦法等价权函数的基础上引入淘汰域,将改进的丹麦法权函数作为抗差权,选取可变动的临界值,通过双步 M 估计,提高抗差估计的精度。

1 改进型双步 M 估计

1.1 初值选取

由于 LS 估计不具抗差性,LS 估计的残差不能有效反映粗差的位置。而 L_1 -范估计的平差值具有较好的抗差性,能较为准确地体现真实残

差^[5],本文将其作为双步 M 估计的初值。

首先,列误差方程:

$$\mathbf{V} = \mathbf{A}\hat{\mathbf{X}} - \mathbf{l} \tag{1}$$

解得残差:

$$\mathbf{V} = \mathbf{A}(\mathbf{A}^T\mathbf{P}_i\mathbf{A})^{-1}\mathbf{A}^T\mathbf{P}_i\mathbf{l} - \mathbf{l} \tag{2}$$

然后,以 L_1 -范权函数($\bar{\mathbf{P}} = \frac{\mathbf{P}_i}{|\mathbf{V}|+c}$)重新定权,再进行迭代计算,收敛时得残差初值:

$$\mathbf{V}_0 = \mathbf{A}(\mathbf{A}^T\bar{\mathbf{P}}\mathbf{A})^{-1}\mathbf{A}^T\bar{\mathbf{P}}\mathbf{l} - \mathbf{l} \tag{3}$$

式中, \mathbf{P}_i 为观测权, $\bar{\mathbf{P}}$ 为等价权, \mathbf{V}_0 为初始残差, \mathbf{l} 为常数项, c 为一个很小的正常数。

1.2 等价权函数选择

丹麦法的权函数如下^[3]:

$$\bar{P}_i = \begin{cases} p_i, & |V_i| \leq k \\ p_i \exp(1 - \frac{|V_i|}{k}), & |V_i| > k \end{cases} \tag{4}$$

式中, $|V_i| = |v_i/\sigma_{v_i}|$, 其中 $\sigma_{v_i} = \sigma_0\sqrt{q_{v_i}} = \sigma_0\sqrt{1/p_i - \mathbf{a}_i^T\mathbf{N}^{-1}\mathbf{a}_i}$, V_i 即标准化残差, σ_{v_i} 为 V_i 的均方差, σ_0 为单位权中误差, $\mathbf{N} = \mathbf{A}^T\mathbf{P}\mathbf{A}$, \mathbf{a}_i 为设计矩阵 \mathbf{A} 的第 i 行元素组成的向量。

丹麦法权函数缺乏淘汰段,导致抗差效率降低。加入淘汰段,对丹麦法进行改进:

$$\bar{P}_i = \begin{cases} p_i, & |V_i| \leq k_0 \\ p_i \exp(1 - \frac{|V_i|}{k_0}), & k_0 < |V_i| \leq k_1 \\ 0, & |V_i| > k_1 \end{cases} \tag{5}$$

收稿日期:2014-03-19
项目来源:国家自然科学基金(41064001);广西自然科学基金(2012GXNSFAA053183,2012GXNSFGA060001);广西空间信息与测绘重点实验室基金(桂科能 1207115-07);广西研究生教育创新计划(YCSZ2013077);广西“八桂学者”岗位专项。
第一作者简介:林国标,硕士生,主要研究方向为 GNSS 技术与数据处理,E-mail: gblin168@163.com。
通讯作者:刘立龙,博士,教授,研究方向为 GNSS 技术及应用,E-mail: hu_liulilong@163.com。

式中, $|V_i| = |v_i/\sigma_{v_i}|$, 其中 $\sigma_{v_i} = \sigma_0 \sqrt{q_{v_i}} = \sigma_0 \sqrt{1/p_i - a_i N^{-1} a_i^T}$ 。

1.3 第一步——综合抗差阶段

该阶段有两个关键问题需要解决:一是确定临界值 k_0 、 k_1 ,二是选取等价权。 k_0 、 k_1 一般分别取某一固定值^[3], $k_0=1\sim 2.5$, $k_1=3\sim 6$ 。临界值在 k_0 、 k_1 基础上再乘以一个可变因子 $1/\sqrt{r}$ (r 为多余度),可解决临界值取值不当的问题,实现系数阵空间与观测空间的同时抗差^[6]。 r 的平均值为 $(n-m)/n$ (n 、 m 分别为观测值个数与参数个数),故可变因子为 $d = 1/\sqrt{r} = \sqrt{n/(n-m)}$ 。此时临界值会随具体问题而变化,增强了抗差估计的灵活性。本文取 $k_0=2d$, $k_1=4d$ 。

选取改进后的丹麦法等价权作为抗差权,单位权中误差 σ_0 采用第一次或第二次抗差估计所求得的观测残差来计算,能够获得比较准确的均方差因子^[7]:

$$\sigma_0 = \pm \text{med}(|\sqrt{p_i} \cdot v_i|)/\alpha, \alpha=0.674\ 5 \tag{6}$$

1.4 第二步——优化阶段

将第一步所得的估计值作为该阶段的初值,利用截尾 LS 法重新平差计算。取权函数^[8]:

$$\bar{P}_i = \begin{cases} p_i, & |v_i| \leq k\sigma_{v_i} \\ 0, & |v_i| > k\sigma_{v_i} \end{cases} \tag{7}$$

临界值取 $k=3$ 。根据此时的 v 与 σ^2 ,将大于 3 倍均方差的观测值删去,然后利用剩余的观测值进行加权 LS 平差,得到最终的参数解。精度评定采用:

$$\hat{\sigma}_0^2 = \frac{v^T \bar{P} v}{n-m-t} \tag{8}$$

式中, n 、 m 、 t 分别为总观测数、必要观测数和删去的粗差观测数; \bar{P} 为最终权。

2 算例分析

本文结合一个测边网算例和一个水准网算例^[9],对比本文提出的改进型双步 M 估计与丹麦法、Huber 法、 L_1 法、IGG I 法及两步抗差估计 5 种方法的平差结果。将无粗差时的 LS 法平差值当作真值,传统抗差估计方法和两步抗差估计的

临界值取先验值,均方差因子通过初次平差的残差由中位数法计算,即 $\sigma_0 = \pm \text{med}(|\sqrt{p_i} \cdot v_i|)/0.674\ 5$ 。

算例 1 如图 1,测边网中 A 、 B 、 C 、 D 为已知点, P_1 、 P_2 、 P_3 、 P_4 为未知点,观测 13 条边长,测距精度 $\sigma_S=3\text{ mm}+1\times 10^{-6}S$ 。起算数据及观测边长见表 1。在第 3、12 条边上分别附上一 30 dm、50 dm 的粗差,各种抗差估计方法的计算结果见表 2。

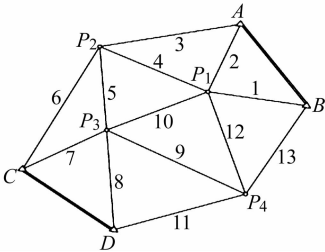


图 1 测边网
Fig. 1 Trilateration network

算例中,测边网的污染率为 15.4%,属于较严重污染。由表 2 可知:

1) 在参数估计方面,几种传统抗差估计方法的估计结果与真值之差较大:丹麦法、Huber 法、 L_1 法和 IGG I 法最大的差值分别为 11.2 cm、30.4 cm、20.3 cm、43.1 cm;两步抗差估计的结果与真值之差较小,最大为一 3.2 cm;本文方法与真值之差较小,最大差值为一 2.5 cm,整体比两步抗差估计小一些。

2) 在最终权方面, P_3 、 P_{12} 为粗差权,其余为正常观测值的最终权,传统抗差估计方法的 P_3 、 P_{12} 未能同时降到最小或 0,如丹麦法、Huber 法、 L_1 法和 IGG I 法的 P_3 分别为 0.03、0.11、0.03、0.13, L_1 法正常观测值的最终权受粗差影响波动很大;两步抗差估计的 P_3 、 P_{12} 均降为 0,而 P_6 、 P_8 、 P_{13} 分别降为 0.42、0.40、0.73;本文方法的 P_3 、 P_{12} 均降为 0,其余权值均为观测权。

若在上述粗差基础上再加上一个粗差,即在第 3、10、12 条边上分别附上一 30 dm、-20 dm、50 dm 的 3 个粗差,再一次进行平差计算,则各种抗差估计方法的计算结果见表 3。此时测边网的污染率为 23.1%。由表 3 可知:

表 1 测边网观测数据及起算数据
Tab. 1 Observed value and known data

序号	1	2	3	4	5	6	7	8	9	10	11	12	13
边观测值 S/m	5 760.706	5 187.342	7 838.880	5 483.158	5 731.788	8 720.162	5 598.570	7 494.881	7 493.323	5 438.382	5 487.073	8 884.587	7 228.367
观测权	1.29	1.49	0.86	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.71	0.96
	A			B			C			D			
起算坐标/m	X			47 943.002			40 049.229			36 924.728			
	Y			66 225.854			53 782.790			61 027.086			

表 2 各种抗差估计方法的计算结果
Tab. 2 Results by various robust estimation methods

与真值之差/cm	ΔX_1	ΔY_1	ΔX_2	ΔY_2	ΔX_3	ΔY_3	ΔX_4	ΔY_4
丹麦法	-1.9	6.0	9.0	11.2	1.3	-1.4	2.3	-2.6
Huber 法	7.8	10.7	20.8	30.4	2.5	11.3	-17.6	18.6
L_1 法	0.1	0.8	6.4	20.3	-0.2	4.7	-0.9	2.8
IGG I 法	-4.4	21.7	33.7	43.1	4.8	-2.1	3.3	-4.2
两步抗差估计	-1.1	0.6	0.5	0.7	0.8	-2.2	3.0	-3.2
本文方法	-1.3	0.8	0.8	0.5	0.2	-1.5	2.3	-2.5

最终权	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}
丹麦法	1.29	1.49	0.03	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.00	0.96
Huber 法	1.29	1.49	0.11	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.05	0.96
L_1 法	854.19	725.61	0.03	0.66	43.61	12.73	2.18	67.94	3.64	55.49	1 351.44	0.01	9.32
IGG I 法	1.29	1.49	0.13	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.00	0.96
两步抗差估计	1.29	1.49	0.00	1.38	1.32	0.42	1.35	0.40	0.91	1.42	1.38	0.00	0.73
本文方法	1.29	1.49	0.00	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.00	0.96

表 3 各种抗差估计方法的计算结果
Tab. 3 Results by various robust estimation methods

与真值之差/cm	ΔX_1	ΔY_1	ΔX_2	ΔY_2	ΔX_3	ΔY_3	ΔX_4	ΔY_4
丹麦法	-38.1	21.9	108.3	109.0	65.6	67.7	-57.3	62.5
Huber 法	19.6	-6.4	85.4	100.0	59.6	134.9	-162.1	175.6
L_1 法	-0.1	0.7	9.8	42.9	52.0	115.1	-61.5	64.6
IGG I 法	-7.9	7.1	96.3	104.3	62.5	102.9	-112.2	121.8
两步抗差估计	-7.9	7.1	96.3	104.3	62.5	102.9	-112.2	121.8
本文方法	0.1	1.7	0.7	1.5	-1.1	-2.3	2.4	-2.7

最终权	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}
丹麦法	1.29	1.49	0.86	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.14	0.96
Huber 法	1.29	1.49	0.86	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.71	0.96
L_1 法	991.77	48.98	0.03	0.29	3.35	10.77	0.10	5.80	3.13	0.15	593.06	0.02	0.18
IGG I 法	1.29	1.49	0.86	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.35	0.96
两步抗差估计	1.29	1.49	0.86	1.38	1.32	0.73	1.35	0.91	0.91	1.42	1.38	0.35	0.96
本文方法	1.29	1.49	0.00	1.38	1.32	0.73	1.35	0.91	0.91	0.00	1.38	0.00	0.96

1) 在参数估计方面,传统抗差估计方法和两步抗差估计的结果与真值之差比较大,如丹麦法、Huber 法、 L_1 法、IGG I 法和两步抗差估计的最大差值分别为 109.0、175.6、115.1、121.8、121.8 cm;本文方法的结果与真值之差较小,最大仅为 -2.7 cm。

2) 在最终权方面, P_3 、 P_{10} 、 P_{12} 为粗差权,其余为正常观测值的最终权。传统抗差估计方法和两步抗差估计的 P_3 、 P_{10} 仍为观测权,分别为 1.86、14.2, L_1 法的粗差权虽然最后有所降低,但是正常观测权之比波动很大;本文方法的 P_3 、 P_{10} 、 P_{12} 均降为 0,其余权仍为观测权。

综上所述,当测边网含有 2 个粗差时,传统抗差估计方法的抗差效果较差,两步抗差估计和本文方法的抗差效果较好,而本文方法的整体抗差效果优于两步抗差估计;当测边网含有 3 个粗差时,传统抗差估计方法和两步抗差估计基本丧失了抗差估计的能力,而本文方法的抗差效果依然

比较显著。这是由于传统抗差估计方法和两步抗差估计往往采用对粗差具有均衡性的 LS 法的平差结果作为抗差初值,不能准确反映粗差的位置。而前者临界值是固定的,不能有效适应实际问题,因此它们抵御多维粗差的效果并不显著。

算例 2 如图 2,水准网中 A、B 为已知点, $H_A=8.016$ m, $H_B=9.016$ m,点 P_1 、 P_2 、 P_3 为待定点。进行 7 条水准路线观测,见表 4。在第 1、4 路线上分别附加 -20 mm、40 mm 的粗差,各种方法的计算结果见表 5。

图 2 水准网图
Fig. 2 Leveling network

表 4 水准网观测高差值及路线长
Tab. 4 Observed value and length

观测序号	1	2	3	4	5	6	7
观测高差 h/m	1.359	2.009	0.363	1.012	0.657	0.238	-0.595
路线长/km	1.1	1.7	2.3	2.7	2.4	1.4	2.6
观测权	2.73	1.76	1.30	1.11	1.25	2.14	1.15

表 5 各种抗差估计方法的计算结果
Tab. 5 Results by various robust estimation methods

与真值 之差/mm	丹麦法	Huber 法	L_1 法	IGG I 法	两步抗 差估计	本文 方法	
ΔX_1	-8.7	-8.0	-6.7	-8.3	-8.8	-0.3	
ΔX_2	1.9	5.4	-2.8	3.8	1.4	-0.2	
ΔX_3	-5.6	-5.1	-6.0	-5.3	-5.7	-0.1	
最终权	P_1	P_2	P_3	P_4	P_5	P_6	P_7
丹麦法	2.73	1.76	1.30	0.44	1.25	2.14	1.15
Huber 法	2.73	1.76	1.30	0.76	1.25	2.14	1.15
L_1 法	0.21	16.95	0.12	0.03	986.30	60.51	0.24
IGG I 法	2.73	1.76	1.30	0.61	1.25	2.14	1.15
两步抗差估计	2.73	1.76	1.30	0.40	1.25	2.14	1.15
本文方法	0.00	1.76	1.30	0.00	1.25	2.14	1.15

算例中,多余观测数为 4,污染率为 28.6%,属于比较严重的污染,因此本算例只进行包含 2 个粗差的比较。由表 5 可知:

1) 在参数估计方面,传统抗差估计方法和两步抗差估计的估计结果与真值之差较大,丹麦法、Huber 法、 L_1 法、IGG I 法和两步抗差估计的最大差值分别为 -8.7 mm、-8.0 mm、-6.7 mm、-8.3 mm、-8.8 mm;本文方法的结果与真值之差较小,最大仅为 -0.3 mm。

2) 在最终权方面, P_1 、 P_4 为粗差权,其余为正常观测值的最终权,传统抗差估计和两步抗差估计的 P_1 、 P_4 未能同时降为 0,如丹麦法、Huber 法、IGG I 法及两步抗差估计的 P_1 观测权均为 2.73, L_1 法的 P_1 为 0.21, L_1 法的 P_3 、 P_7 有所降低;本文方法的 P_1 、 P_4 均降为 0,其余最终权维持不变。

和算例 1 同理,因为传统抗差估计和两步抗差估计常常采用不能准确反映粗差位置的 LS 法平差结果作为抗差初值,前者临界值是固定的,所以往往不能抵抗多维粗差的干扰。而本方法以 L_1 -范估计结果作为初值,第一步的抗差权采用改进的丹麦法等价权,临界值可变;第二步采用截尾 LS 法,取得了较好的估计结果。

3 结 语

基于丹麦法的改进型双步 M 估计是一种以 L_1 -范估计的平差值作为初值、以改进后的丹麦等

价权作为综合抗差阶段权因子的抗差估计方法,比较有效地解决了抗差估计初值与临界值选取不当的问题,提高了抗差估计的精度。算例表明,无论是测边网平差,还是水准网平差,该方法都能有效抵御多维粗差的干扰,获得优于传统抗差估计和两步抗差估计的平差结果,具有较强的抗差性。

参考文献

[1] 杨元喜,吴富梅. 临界值可变的抗差估计等价权函数[J]. 测绘科学技术学报, 2006(5): 320-324 (Yang Yuanxi, Wu Fumei. Modified Equivalent Weight Function with Variable Criterion for Robust Estimation[J]. Journal of Geomatics Science and Technology, 2006,(5): 320-324)

[2] 邱卫宁. 具有稳健初值的选权迭代法[J]. 武汉大学学报: 信息科学版, 2003(4): 452-454 (Qiu Weining. Method for Selecting Weight Iteration with Robust Initial Value[J]. Geomatics and Information Science of Wuhan University, 2003,(4): 452-454)

[3] 李浩军,唐诗华,黄杰. 经典选权迭代法研究与两步抗差估计的提出[J]. 海洋测绘, 2007(1): 17-20 (Li Haojun, Tang Shihua, Huang Jie. The Discussion about Some Kinds of Selecting Weight Iteration Method[J]. Hydrographic Surveying and Charting, 2007 (1): 17-20)

[4] 靳奉祥. 抗差估计理论与方法研究[J]. 山东科技大学学报: 自然科学版, 2003(4): 1-6 (Jin Fengxiang. Study on Robust Estimation Theory and Method[J]. Journal of Shandong University of Science and Technology: Natural Science, 2003 (4): 1-6)

[5] 王振杰,欧吉坤,曲国庆. 抗差估计的初值选择[J]. 地壳形变与地震, 2001(3): 32-35 (Wang Zhenjie, Ou Jikun, Qu Guoqing. Choice of Initial Value for Robust Estimation[J]. Crustal Deformation and Earthquake, 2001,(3): 32-35)

[6] 欧吉坤. 一种三步抗差方案的设计[J]. 测绘学报, 1996(3): 173-179 (Ou Jikun. Design of a New Scheme of Robust Estimation by Three Steps[J]. Acta Geodaetica et Cartographica Sinica, 1996(3): 173-179)

[7] 高晓,戴吾蛟. 抗差 Helmert 方差分量估计在 GPS/BDS 组合定位中的应用[J]. 大地测量与地球动力学, 2014, 34(1): 173-176 (Gao Xiao, Dai Wujiao. Application of Robust-helmert Variance Component Estimation to Position in Combination of GPS and BDS[J]. Journal of Geodesy and Geodynamics, 2014, 34(1): 173-176)

[8] 王福丽,成英燕,韦钺,等. 利用抗差多项式拟合法探测修复 GNSS 周跳[J]. 大地测量与地球动力学, 2013, 33(3): 129-132 (Wang Fuli, Cheng Yingyan, Wei Cheng, et al. GNSS Cycle Slip Detection and Correction Using Robust Polynomial Fitting [J]. Journal of Geodesy and Geodynamics, 2013, 33(3): 129-132)

[9] 武汉大学测绘学院测量平差学科组. 误差理论与测量平差基础 [M]. 武汉: 武汉大学出版社, 2009 (Research Group of Surveying Adjustment in School of Geodesy and Geomatics of Wuhan University. Error Theory and Foundation of Surveying Adjustment [M]. Wuhan: Wuhan University Press, 2009)