

文章编号:1671-8860(2005)03-0242-04

文献标志码:A

应用抗差估计理论分析 2000 国家重力基本网

郭春喜^{1,2} 李斐¹ 王斌²

(1 武汉大学测绘学院,武汉市珞喻路 129 号,430079)

(2 国家测绘局大地测量数据处理中心,西安市友谊东路 334 号,710054)

摘要:应用抗差估计理论,结合 2000 国家重力基本网的实际数据,通过大量的试算对 2000 国家重力基本网的数据进行分析比较,提出了符合 2000 国家重力基本网平差计算的抗差估计方法。

关键词:重力基本网;抗差估计;等价权

中图法分类号:P223.0;P207

粗差探测是测量数据处理、测量质量控制的重要研究主题之一。如何有效地消除或减弱观测数据中的粗差对参数估值的不良影响,一直是人们关注的问题。而传统的排除粗差方法比直接用抗差估计方法要困难得多,抗差估计能在拒绝和接受一个观测值之间起一种平滑作用,这样就保留了一些不是很好但仍然可以利用的观测值^[1]。

在国家 1985 重力基本网的基础上重建的 2000 国家重力基本网(以下简称 2000 网),是国家重要的基础测绘工程项目,将为我国测绘、地质、石油、城建等部门提供更精确、更现势的重力基准成果,及时、有效地为我国经济建设、空间技术、地质勘探、海洋开发、地震预报、国防建设等领域的基础测绘服务。在重力网的施测时,虽然通过相对中误差限差、闭合环闭合差限差检验,大大降低了平差中的粗差观测量,但难免会存在少量粗差观测量,因此,对 2000 网来说,在平差中应该通过抗差估计对粗差予以剔除。

1 抗差估计

本文着眼测量实际,重点研究观测参数的估计,要求估值在减免粗差影响的同时,又具有小方差。其基本方案是抗差最小二乘法,即在构造参数解式时,兼取抗差估计和最小二乘法的理论和算法。基于误差有界分布理论,将数据分成三个

基本段:小误差取原观测权;大误差的观测予以淘汰;中间段观测采用等价权,使抗差估值方程形式上和最小二乘法相同^[2]。

抗差估计与粗差探测不同,它通过一定的统计量构造权函数,并基于权函数对异常观测进行控制。权函数一般是连续函数,当观测误差达到临界值时,权函数值为零;当观测误差近于临界值时,权函数值也近于零,于是可有效地控制观测异常对参数估值的影响。

由于抗差估计方法很多,并且每种抗差估计方法可以根据参数等的不同选取分很多种。针对这一情况,需结合抗差估计有关理论和 2000 国家重力基本网的数据特点,通过大量的研究分析与试算,确定出适合于 2000 国家重力基本网平差计算的抗差估计方法及其参数,确定最终的平差计算方案。

在 2000 网数据处理中设计了如下等价权模型:

$$\bar{p}_i = \begin{cases} p_i, & |v'_i| = \left| \frac{v_i}{\sigma_{v_i}} \right| \leq k_0 \\ p_i \cdot \frac{k_0}{|v'_i|} \left(\frac{k_1 - |v'_i|}{k_1 - k_0} \right)^2, & k_0 < |v'_i| \leq k_1 \\ 0, & |v'_i| > k_1 \end{cases} \quad (1)$$

式中, $k_0 = 1.0 \sim 1.5$, $k_1 = 3.0 \sim 4.5$; $\sigma_{v_i} = \sqrt{q_{v_i}} \sigma_0$, q_{v_i} 为残差协因数矩阵 Q_v 的第 i 个对角线元素,对

收稿日期:2004-11-15。

项目来源:国家自然科学基金资助项目(40374005);国家测绘局测绘科技发展基金资助项目(20010103, WKL03-0201);地理空间信息工程国家测绘局重点实验室基础测绘若干重大关键技术课题资助项目(1460130424210)。

于法方程的解式, $\mathbf{Q}_v = \mathbf{P}^{-1} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T, \mathbf{N} = \mathbf{A}^T\mathbf{P}\mathbf{A}$ 。

解算采用直接解法或迭代法, 第 $k+1$ 次解式为:

$$\hat{\mathbf{X}}^{(k+1)} = (\mathbf{A}^T \bar{\mathbf{P}}^{(k+1)} \mathbf{A})^{-1} \mathbf{A}^T \bar{\mathbf{P}}^{(k+1)} \mathbf{L}$$

1) 由初始解求得 $\hat{\mathbf{X}}^{(0)}, \mathbf{V}^{(0)}, \hat{\sigma}_{v_i}$;

2) 由 $v_i^{(0)}$ 经式(1)求得 $\bar{P}_i^{(1)}$;

3) 求解参数估值 $\hat{\mathbf{X}}^{(1)}$:

$$\hat{\mathbf{X}}^{(1)} = (\mathbf{A}^T \bar{\mathbf{P}}^{(1)} \mathbf{A})^{-1} \mathbf{A}^T \bar{\mathbf{P}}^{(1)} \mathbf{L}, \mathbf{V}^{(1)} = \mathbf{A} \hat{\mathbf{X}}^{(1)} - \mathbf{L}$$

直至 $\max|\hat{x}_j^{k+1} - \hat{x}_j^k| < \epsilon$ (小数) 为止, 否则返回步骤 2)。

2 各类重力观测量的等价权分析

2000 网参加平差的观测量种类较多, 不但精度不同, 而且对于基本网的贡献也各不相同。如绝对重力观测量精度高, 对于整网具有基准意义; 相对重力观测量, 不但采用了飞机、汽车等运载工具, 而且长基线联测与一般联测在整网中的作用也不相同。因此, 在利用抗差估计重新定义权值时, 要根据试算情况进行分析处理。

2.1 绝对重力观测量的等价权分析

在 2000 网平差计算时采用弱基准, 各基准点采用最近观测的几期成果作为独立观测量参与平差计算。在抗差估计试算过程中发现, 若采用抗差估计等价权模型重新定义绝对重力观测成果的权值, 一部分绝对重力观测成果的权将会改变, 甚至有的观测量的权值变为 0 (即该观测量被剔除), 主要原因与各基准点的绝对重力观测成果的稳定性有关。从选用的 21 个基准点的各期绝对重力观测成果来看, 21 个基准点可分为三类: ① 喀什和狮泉河只有一期绝对重力成果, 且这两个基准点均处于地理条件较差的西部地区; ② 漠河、武汉、昆明、南宁、广州、乌鲁木齐、拉萨、北京白家疃、北京房山(新)、海口和九峰等 11 个基准点的几期成果互差较小 ($10 \times 10^{-8} \text{ m} \cdot \text{s}^{-2}$ 左右), 属于比较稳定的点; ③ 海拉尔、哈尔滨、西安、成都、上海、香港、格尔木和西宁等 8 个基准点的几期成果相差较大, 最大可以达到 $30 \times 10^{-8} \text{ m} \cdot \text{s}^{-2}$ 。

由于重力场变化、地下水位、观测时间等因素的影响, 不同时间观测的基准点绝对重力成果存在差异, 在平差计算时, 采用相距相对重力测量施测年代较近的几期成果的目的是为了综合反映各期成果信息, 确定更为可靠的基准点绝对重力成果和其他重力点的重力值。而抗差估计试算结果显示, 几期成果相差较大的 8 个基准点的绝对重力观测成果均有一期被剔除 (权降为 0), 绝对重

力观测量权值的改变不但改变了基准点的平差重力值, 而且削弱了绝对重力成果的基准意义。为了保证绝对重力成果的基准意义, 在抗差估计过程中, 绝对重力观测成果不进行降权处理, 始终采用先验权。

2.2 长基线标定与飞机联测成果的等价权分析

长基线标定联测 (飞机联测和短距离地面联测) 成果本身精度较高, 同时又对整网起一定的控制作用 (长基线标定的比例因子用于外业数据处理, 并控制外业成果的质量)。在基本网相对重力联测数据中, 虽然飞机联测成果在数量远低于其他联测成果数量, 但考虑到采用飞机联测时具有运载平稳、闭合时间短的优点, 而且从精度统计来看, 飞机联测成果的精度高于汽车等运载工具联测的成果, 由于精度较高, 在平差中的贡献应大于其他联测成果。但试算结果显示, 由于地面联测成果数远大于联测成果数, 有很大一部分飞机联测成果及长基线标定联测成果被降权甚至剔除, 这就削弱了飞机联测成果与长基线标定结果在平差中的作用, 降低了整网平差结果的可靠性。因此, 为了保证平差结果的可靠性, 在抗差估计过程中, 所有飞机联测成果与长基线标定联测成果不进行降权处理。

2.3 其他相对重力联测成果的等价权分析

其他相对重力联测成果主要是指汽车、轮船等运载工具联测的成果, 因为汽车、轮船联测闭合时间较长, 且在自然条件较差的西部地区主要是用汽车联测的。同时, 在外业联测时, 每个测段都用 6 台仪器进行了联测, 这样采用抗差估计既能削弱较大误差的数据对平差结果的影响, 又能充分保证平差采用的联测成果数满足规范要求。

综上所述, 绝对重力观测成果、所有飞机联测成果与长基线标定联测成果始终采用先验权; 其他相对重力观测量采用抗差估计的等价权进行重新定权。

3 抗差估计等价权模型参数选取

在 2000 网平差计算模型中使用的抗差估计等价权模型的主要思想是: 对于标准化残差小于 k_0 的测量成果保持其原有的权; 对于标准化残差大于 k_0 但小于 k_1 的测量成果依据等价权模型计算相应的权, 即降权; 而对于标准化残差大于 k_1 的测量成果予以剔除。但是在 2000 网数据处理方案中, 对 k_0, k_1 的值只给出了一个范围, 即 k_0 取 $1.0 \sim 1.5, k_1$ 取 $3.0 \sim 4.5$ 。因此从该模型中

可以看出,参数 k_0 、 k_1 的选取对于测量成果在平差中的影响至关重要,参数选取合理与否将直接影响到观测成果的降权与取舍,进而影响平差成果的可靠性与精度。

在试算时,采用变换模型参数的方法进行反复试算,结合试算结果进行综合分析,确定出切实符合 2000 网实际的模型参数。主要试算结果见表 1。

表 1 等价权模型参数试算结果

Tab. 1 Calculation Result Parameter of Equivalence Weight Model

k_0	k_1	平差结果比较				
		m_0	m_1	解算次数	剔除成果数	降权成果数
1.0	3.0	15.1	6.8	6	62	723
1.0	3.5	15.3	6.9	4	31	754
1.0	4.0	15.5	6.9	4	13	776
1.0	4.5	15.6	7.0	4	4	783
1.5	3.0	16.5	7.2	4	56	280
1.5	3.5	16.7	7.2	4	32	303
1.5	4.0	16.9	7.3	4	14	323
1.5	4.5	17.0	7.3	4	4	336

注: m_0 表示单位权中误差, m_1 表示重力点平均中误差, 单位均为 $10^{-8} \text{ m} \cdot \text{s}^{-2}$ 。

从表 1 可以看出以下两个方面。

1) 当 k_0 取为 1.0 时,无论 k_1 取 3.0~4.5 的任何值,均有大量的观测成果被降权使用(700 以上)。虽然平差结果的内符合精度较高,但从平差结果来看,许多平差后残差在 $20 \times 10^{-8} \text{ m} \cdot \text{s}^{-2}$ 左右的观测成果都被降权使用(200 多个),这样会削弱外业成果的作用,同时有违平差使用数据尽可能采用原始成果的原则,显然会影响平差结果的可靠性,因此, k_0 不能取 1.0 或接近 1.0 的值。

2) 当 k_1 取 3.0, k_0 取 1.0 或 1.5 时,分别有 62 个和 56 个观测成果被剔除,因为在本次平差试算前,已经通过初步试算剔除了一部分含粗差的观测量,因此,通过抗差估计剔除如此多的观测量显然是不合理的。通过对被剔除的观测量进行分析发现,很多平差后残差在 $50 \times 10^{-8} \text{ m} \cdot \text{s}^{-2}$ ~ $60 \times 10^{-8} \text{ m} \cdot \text{s}^{-2}$ 且对平差有用的观测成果被剔除,这些观测量主要集中在少数地区和几台仪器,说明这些观测量与观测条件和仪器性能有关,这些信息对平差是非常有用的,不能随便剔除,因此 k_1 不能取 3.0。

综合表 1 中的试算结果,本着既保证平差结果的可靠性,又尽量采用原始成果的原则, k_0 取 1.5、 k_1 取 4.5 则显得更有效。既最大限度地采用了原始测量成果,又削弱了含有较大误差的测量成果对平差结果的影响,提高了平差计算结果

的精度,保证了成果的质量与可靠性。

4 抗差估计的作用分析

原来的重力网平差均采用最小二乘估计求解,对粗差数据的控制主要是在平差计算前完成,即在数据预处理过程中通过制定一个较为合理的取舍原则(主要是限差),符合限差要求的就参加平差计算,超过限差的就舍去。这样往往会造成数据的不合理取舍,尤其对于那些处于限差边缘的数据,其误差对参数估值的影响是相当的,剔除或保留都有可能降低估值效率。而对于在这一范围的数据,采用常规的在预处理中剔除粗差的方法是无法解决的,因此,往往含有粗差的数据被纳入平差采用数据中。

最小二乘估计是通过求解余差的平方和最小来求参数的估值的,其估值的解是唯一的,但是这种估计不具备抗差的能力,而且由这种方法得到的参数估值的可靠性往往会被少数几个粗差破坏。而抗差估计是通过求解余差的另外某种函数的极小,来得到既具有较好统计特性又具有较好的抗差能力的估值的。

通过抗差估计对数据质量进行控制,主要体现在以下几个方面。

1) 改善联测成果对平差结果的影响。

在采用最小二乘估计时,对于满足限差要求的观测量一律采用先验权,即较好的成果与较差的成果都等权对待,这样可能会降低参数估值的可靠性,而抗差估计通过余差构造权函数很好地削弱了这种影响。从平差计算结果来看,余差绝对值大的(主要体现为标准化残差大)观测量对应的权小,余差绝对值小的观测量对应的权大,这表明等价权可以增强未知量估值的可靠性。在 2000 国家重力网平差计算中,总共剔除了 4 个观测量,其余差(单位为 $10^{-8} \text{ m} \cdot \text{s}^{-2}$)分别是 -89.271 、 -104.957 、 -88.894 、 -89.287 ,其标准化残差均大于 4.5。总共有 336 个相对重力观测量的权被降低,其具体分布见表 2。

从表 2 可以看出,通过余差构造的权函数是一个连续的函数,这样对于较好与较差的成果在平差计算时都能给予合适的权,从而有效地控制各种观测量对平差结果的影响。

2) 提高了平差结果的可信度。

为了检验 2000 网平差结果的正确性与可靠性,比较客观地评价 2000 网的实际精度,在方案设计与数据处理时,在全国均匀地选择了 8 个未

表2 权被降低的相对重力观测量的分布
Tab. 2 Distribution of Comparative Gravity Observation Whose Influence is Reduced

最终权与原始权之比	观测量个数
<10%	48
10%~20%	22
20%~30%	33
30%~40%	30
40%~50%	40
50%~60%	29
60%~70%	26
70%~80%	46
80%~90%	24
>90%	38

被选为2000网基准点的绝对重力点,作为2000网的外部检核点。在此利用外部检核点来评价最小二乘估计与抗差估计的计算结果的可信度,比较结果见表3。

表3 外部检核点绝对观测值与最小二乘估计和抗差估计计算结果比较/ $10^{-8} \text{ m} \cdot \text{s}^{-2}$

Tab. 3 Difference Between Absolute Observation and Least Squares, and Difference Between Absolute Observation and Robust Estimation

点名	最小二乘估计结果 与绝对观测值差值	抗差估计结果 与绝对观测值差值
地壳网长春基准站	-6.50	-7.93
武汉地震所绝对点	-3.42	-6.21
地壳网昆明基准站	-4.98	-5.87
地壳网德令哈基准站	-15.61	-10.18
拉萨国家测绘局绝对点	-6.36	-6.49
地壳网北京十三陵基准点	8.38	5.39
乌鲁木齐国家测绘局绝对点	-1.05	0.89
香山地震台绝对点	11.20	10.59
不符值中误差	8.4	7.3

从表3可以看出,最小二乘估计与抗差估计计算结果中大部分点的不符值都比较接近,不符值差异最大的是地壳网德令哈基准站,该点处于地理环境较差的西部地区,外业联测受地理、交通条件的限制,精度较差,而通过抗差估计将不符值由 $-15.61 \times 10^{-8} \text{ m} \cdot \text{s}^{-2}$ 提高为 $-10.18 \times 10^{-8} \text{ m} \cdot \text{s}^{-2}$,提高了平差结果的可信度。

2000网平差计算是一个复杂的过程,该部分试算与分析工作的完成为平差计算的顺利实施提供了技术支持。从最终平差结果来看,采用抗差估计剔除粗差和削弱部分成果对平差结果的影响,减小了采用常规人工干预方法对成果质量的影响,提高了数据处理方法的科学性以及平差结果的精度与可靠性。抗差估计理论在2000网平差计算中的成功应用,为该理论在以后的大型国家控制网平差计算中的应用积累了一定的经验。

参 考 文 献

- 1 黄幼才. 数据探测与抗差估计. 北京: 测绘出版社, 1990
- 2 周江文, 杨元喜, 欧吉坤, 等. 见: 抗差估计论文集. 北京: 测绘出版社, 1992

第一作者简介: 郭春喜, 博士生。主要从事物理大地测量与大地测量数据处理研究。

E-mail: sacgdp@bylink.com.cn

Application of Robust Estimation Theory to Analyzing 2000 China Gravity Base Net

GUO Chunxi^{1,2} LI Fei¹ WANG Bin²

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

(2) Geodetic Data Processing Center, State Bureau of Surveying and Mapping,
334 East Youyi Road, Xi'an 710054, China)

Abstract: This paper presents a method for robust estimation that corresponds to the adjustment of 2000 China gravity base net. It mainly adopts the theory of robust estimation, combined with the actual data of 2000 China gravity base net, compares and analyzes the data of 2000 China gravity base net through a great deal of computation.

Key words: gravity base net; robust estimation; equivalence weight