

# 隶属度函数改进的IGGIII抗差方案及其在测量数据处理中的应用

雷新强<sup>1</sup>, 李屹旭<sup>2\*</sup>

<sup>1</sup>贵州大学矿业学院, 贵州 贵阳

<sup>2</sup>贵州大学农学院, 贵州 贵阳

收稿日期: 2024年11月25日; 录用日期: 2024年12月19日; 发布日期: 2024年12月27日

## 摘要

针对IGGIII抗差估计方法根据残差大小是否满足设定的界限阈值进行分段定权不能很好处理粗差模糊边界的问题, 提出一种利用标准化残差属于粗差的模糊子集隶属度函数改进的IGGIII抗差估计方法(MF-IGGIII), 具体为: 首先通过模糊统计的方法确定标准化残差最大程度受粗差污染的模糊集合的隶属函数, 然后根据计算的隶属度函数值构建等价权函数并进行选权迭代过程。通一个实际测角网平差算例验证本文方法, 即在角度观测值中分别模拟4个、5个和6个三种不同数目粗差后再分别进行平差处理。选取参数估值误差的范数作为精度指标, 结果表明, 与传统IGGIII抗差估计方法相比, 新方法精度分别提升了35.9%、32.5%和32.7%。说明同等条件下, 由于新方法考虑了粗差的模糊性质, 故取得了更高的参数估计精度, 且在粗差个数增加到观测值总数的1/3情况下, 仍然获得了较好的结果, 其抗差性能更稳定, 结果更可靠。

## 关键词

粗差, 模糊子集, 隶属函数, 标准化残差, IGGIII

# Improved IGGIII Robust Scheme with Membership Function and Its Application in Measurement Data Processing

Xinqiang Lei<sup>1</sup>, Yixu Li<sup>2\*</sup>

<sup>1</sup>Mining College of Guizhou University, Guiyang Guizhou

<sup>2</sup>College of Agriculture, Guizhou University, Guiyang Guizhou

Received: Nov. 25<sup>th</sup>, 2024; accepted: Dec. 19<sup>th</sup>, 2024; published: Dec. 27<sup>th</sup>, 2024

\*通讯作者。

文章引用: 雷新强, 李屹旭. 隶属度函数改进的IGGIII抗差方案及其在测量数据处理中的应用[J]. 应用数学进展, 2024, 13(12): 5315-5321. DOI: 10.12677/aam.2024.1312512

## Abstract

In response to the problem that the IGGIII robust estimation method cannot effectively handle rough fuzzy boundaries by segmenting and weighting based on whether the residual size meets the set threshold, a modified IGGIII robust estimation method (MF-IGGIII) is proposed, which utilizes the membership function of the fuzzy subset of standardized residuals belonging to rough errors. Specifically, the membership function of the fuzzy set that is most heavily contaminated by rough errors in the standardized residuals is first determined through fuzzy statistics, and then an equivalent weight function is constructed based on the calculated membership function values and subjected to a weight selection iteration process. Verify the method proposed in this paper through an actual angle measurement network adjustment example, which simulates four, five, and six different numbers of gross errors in the angle observation values, and then performs adjustment processing separately. The analysis results of the case study show that under the same conditions, even when the number of gross errors increased to one-third of the total number of observations, good results were still obtained, with more stable robustness and more reliable results.

## Keywords

Gross Error, Fuzzy Subset, Membership Function, Standardized Residuals, IGGIII

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

当平差系统不包含粗差且性态良好(非病态系统)时,经典最小二乘估计可以获得参数的最优估计值[1][2],但是,当平差系统观测值中包含粗差时,观测值服从某种污染分布[3],此时,经典最小二乘估计解不再是最优的。关于粗差处理,大致可分为两种途径,即将粗差影响归入函数模型的均值漂移模型和将粗差影响归入随机模型的方差膨胀模型[4][5]。前者通过假设检验法检测定位和删除粗差[5]-[7],后者通过残差“表现”采用分段定权的方式抑制粗差对数据处理结果的影响[8]-[10]。上述两种粗差处理方式,无论采取哪一种方式进行数据处理,其本质上都需要经过对粗差进行识别和定位的过程[4],在这个过程中,粗差识别总是以残差绝对值是否满足事先设定的阈值范围为依据判定粗差。然而,这样的粗差识别过程忽略了粗差边界的模糊性。一般认为,粗差是一种比正常条件下最大误差还要大的误差[11],其显著特征是“离群”,然而,无论是“大得多”还是“离群”都是一种相对模糊的概念,应当采用模糊数学的思想对粗差处理才比较合理。例如,若事先设定以2倍后验残差中误差为粗差判定的阈值标准,则当某一残差绝对值等于1.99999倍其后验残差中误差时被认为是正常值予以保留,而当另一残差绝对值等于2.00001倍其后验中误差时即被判定为粗差加以剔除,二者只相差0.00002,一般这样小的差异对于实际工程测量并不会造成任何影响,但是,二者在数据处理中却得到截然不同的处理方式,这显然是不合理的。本文旨在将模糊数学中的隶属度概念及隶属函数引入粗差处理中,以最大程度受粗差污染的标准残差的模糊子集的隶属度值代替IGGIII抗差方案中直接以标准化残差绝对值为依据的等价权确定过程,理论及数值结果都表明,利用隶属度函数改进的IGGIII粗差处理方案(为方便,本文将改进的方法记为MF-IGGIII,意即Membership Function IGGIII)抗差性能更优,参数精度更高,结果更可靠。

## 2. 最大程度受粗差污染的标准化残差的模糊子集及其隶属函数

前文已述及粗差是一种比正常条件下最大误差还要大的误差, 这种所谓“比最大误差还要大的误差”是一种模糊概念, 因为对“最大误差”本身的判定并没有严格分明的界线, 使得当给出 1 个误差时, 能借以明确无误地得出该误差属于或不属于“最大误差”的结论。然而, 此时, 若引入隶属度概念, 并根据某种规则给出其隶属度值, 以此判定该误差多大程度属于“比正常误差还要大的误差”则是比较合适的, 这就是模糊集合的基本思想。模糊集合可以用隶属函数来表示, 根据模糊数学中关于模糊子集的相关理论[12], 最大程度受粗差污染的标准化残差的模糊子集及其隶属函数定义如下:

设  $A$  为最大程度受粗差污染的标准化残差  $W$  的模糊子集, 标准化残差绝对值  $|W|$  为论域  $U$ , 则称映射

$$\begin{aligned} \mu_A: U &\rightarrow [0,1] \\ w &\mapsto \mu_A(w) \in [0,1] \end{aligned} \quad (1)$$

确定了一个  $U$  上的模糊子集  $A$ , 映射  $\mu_A$  称为  $A$  的隶属函数,  $\mu_A(w_i)$  称为  $w_i$  对  $A$  的隶属度。由上述定义可知, 当  $\mu_A(w_i)=1$  时,  $w_i$  完全属于  $A$ , 即观测值  $L_i$  最大程度上受粗差污染, 当  $\mu_A(w_i)=0$  时,  $w_i$  完全不属于  $A$ , 粗差对观测值  $L_i$  完全没有影响。随着  $\mu_A(w_i)$  趋近于 1,  $w_i$  属于  $A$  的程度也增大。大量实践及经验都表明, 不同的模糊模型, 其隶属函数曲线形状各不相同[12], 但就测量平差中最大程度受粗差污染的标准化残差模糊子集的隶属函数而言, 正常模式下, 由系列  $\mu_A(w_i)(i=1,2,\dots,n)$  所描绘的曲线应当是一条以 0 和 1 为渐近线的二维平面曲线, 该曲线表示标准化残差属于粗差的程度由小变大的过程, 后文的算例分析将给予展示。

## 3. MF-IGGIII 抗差估计原理

### 3.1. MF-IGGIII 等价权函数

IGGIII 抗差估计是我国学者周江文教授 1989 年提出的一种粗差处理方法[13], 其本质上与其他稳健 M 估计一样, 通过构建等价权函数利用选权迭代过程降低粗差对整个平差系统的不良影响。IGGIII 估计将权函数分为三段, 即正常段、可疑段和淘汰段, 正常段保持原权不变, 仍采用最小二乘估计, 可疑段采用降权方式对大误差对参数估计的影响加以抑制, 淘汰段则对超大误差给予剔除处理, 具体 IGGIII 等价权函数表达式如式(2)所示[14]:

$$\bar{P}_{i+1} = \begin{cases} 1 & |w_i| \leq k_1 \\ \frac{k_1}{|w_i|} \left( \frac{k_1 - |w_i|}{k_2 - k_1} \right) & k_1 < |w_i| \leq k_2 \\ 0 & |w_i| > k_2 \end{cases} \quad (2)$$

式中,  $|w_i| = |V_i|/\hat{\sigma}_{V_i}$  为观测值  $L_i$  的残差  $V_i$  的标准化残差的绝对值,  $i$  表示迭代次数,  $k_1$  和  $k_2$  通常取经验值 1.5 和 2.5, 其含义是迭代过程中, 各残差最小二乘估计验后中误差的 1.5 倍和 2.5 倍。

显然式(2)给出的 IGGIII 抗差方案是根据最小二乘残差在数值上是否位于或超过事先设定的残差范围定义等价权函数的, 其依据是若观测误差无异常, 即近似服从正态分布时, 误差落在 1.5 倍中误差之外的概率仅为 0.13, 而落在 2.5 倍中误差之外的概率则降为 0.01, 据此, 正常模式下, 观测误差应具有一个明显合理的范围, 若超出这个范围, 则应予以降权或剔除。事实证明, IGGIII 抗差方案是一种比较符合测量实际的粗差处理方法, 能很大程度抑制粗差干扰带来的不良影响。但是, IGGIII 抗差估计忽略了粗差本身是一个模糊变量的客观事实, 模糊变量必然具有模糊性质, 采用硬阈值对粗差进行分类处理可能

会受最小二乘估计“均摊”效应(残差本身未超限但因粗差污染而超限)影响而出现“不公正”降权现象发生。针对此问题, 本文受文献[15]启发, 提出采用模糊数学中模糊子集隶属函数值代替式(2)中的标准化残差进行选权迭代过程。由于隶属度从概念上是一个弹性界线, 不同于硬阈值, 它可以将属于粗差程度大致相同的标准化残差给予相同的降权幅度, 所以利用最大程度属于粗差的模糊子集的标准化残差隶属函数值代替标准化残差本身进行分段定权, 可以极大改善受粗差污染的其它观测值残差被不合理赋权的问题。此时, 相应的 MF-IGGIII 等价权函数定义变为:

$$\bar{P}_{i+1} = \begin{cases} 1 & \mu(w_i) \leq \mu(k_1) \\ \frac{\mu(k_1)}{\mu(w_i)} \left( \frac{\mu(k_2) - \mu(w_i)}{\mu(k_2) - \mu(k_1)} \right) & \mu(k_1) < \mu(w_i) \leq \mu(k_2) \\ 10^{-10} & \mu(w_i) > \mu(k_2) \end{cases} \quad (3)$$

式中,  $\mu(\bullet)$  表示由式(1)定义的标准化残差最大程度属于粗差的模糊子集的隶属函数, 为避免迭代过程中权阵发生奇异, 在淘汰段将权值用  $10^{-10}$  代替。

### 3.2. MF-IGGIII 中的隶属度函数

不同的模糊问题中其模糊子集的隶属函数是不同的, 隶属函数确定方法也多种多样[12]。测量中的误差具有鲜明的特性, 那就是观测误差一般都服从某种有界分布[4], 文献[15]据此通过模糊统计的方法经过反复试验, 最终给出一种符合测量误差实际的隶属函数, 其具体形式如下:

$$\mu_A(w_i) = \begin{cases} 0 & w_i < 1 \\ \frac{1}{1 + \frac{20 + w_i}{w_i^2 c^2}} & w_i \geq 1 \end{cases} \quad (4)$$

式中

$$c = \begin{cases} w_i & w_i < 3 \\ 3 & w_i \geq 3 \end{cases} \quad (5)$$

根据式(3)给出的等价权函数定义, MF-IGGIII 抗差估计步骤如下:

**步骤 1** 根据具体平差问题确定观测值初始权阵, 若为等精度, 则直接用单位阵代替初始权阵, 然后利用最小二乘估计获取初始残差  $V$  并计算获取标准化残差  $w_i$ ;

**步骤 2** 根据(4)式计算各标准化残差的隶属函数值, 带入式(3)计算等价权;

**步骤 3** 利用步骤 2 获取的等价权, 根据选权迭代停止条件  $\|X^{i+1} - X^i\| < \varepsilon$  (本文取  $\varepsilon = 10^{-8}$ ), 重复步骤 1 工作, 更新参数和残差估值;

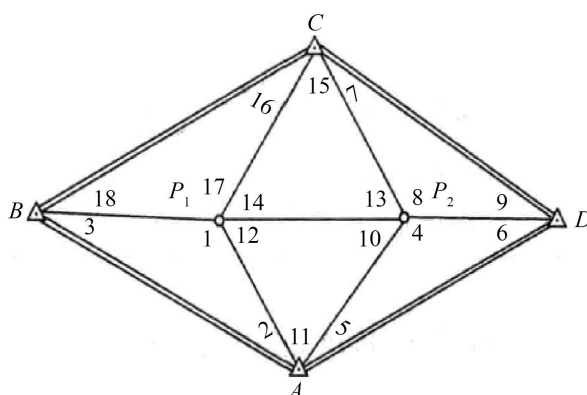
**步骤 4** 结束迭代过程, 输出参数估值。

## 4. 算例及分析

### 4.1. 算例描述

采用文献[11]例 7~9 观测数据。有一测角网, 如图 1 所示, 网中 A、B、C、D 是已知点,  $P_1$ 、 $P_2$  是待定点, 同精度观测获取了 18 个角度观测值, 起算数据及观测值详见文献[11]。为分析本文方法的抗差性能, 在数目不同的观测值上附加不小于  $4.3\hat{\sigma}_0$  (其中  $\hat{\sigma}_0$  为单位权中误差)的粗差, 进行如下粗差处理实验:

**实验 1:** 在观测序号为 2、11、15、17 的 4 个观测值上附加大小分别为  $-7.0''$ 、 $7.0''$ 、 $-5.6''$ 、 $5.6''$  的粗差;



**Figure 1.** Reference [11] Angle measurement network  
**图 1.** 文献[11]测角网

**实验 2:** 在观测序号为 2、7、11、15、17 的 5 个观测值上附加大小分别为 7.0"、6.8"、7.0"、-5.6"、5.6"的粗差;

**实验 3:** 在观测序号为 2、7、10、11、15、17 的 6 个观测值上附加大小分别为 7.0"、6.8"、-6.8"、7.0"、-5.6"、5.6"的粗差;

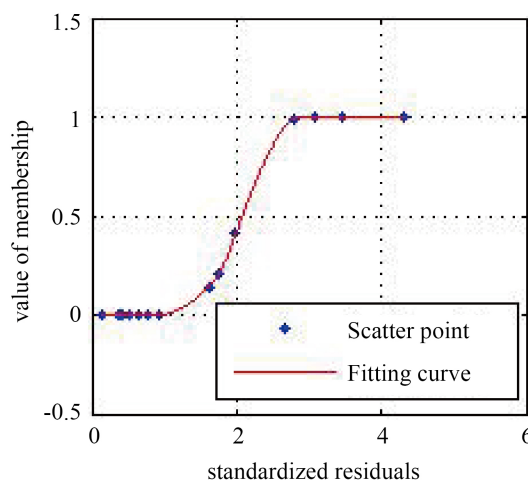
针对以上 3 个数值实验, 分别采用如下 3 种方案进行平差计算:

**方案 1:** 最小二乘估计;

**方案 2:** IGGIII 抗差估计;

**方案 3:** MF-IGGIII 抗差估计。

以上 3 种方案参数估计结果列于表 1, 为方便比较, 可将未附加粗差的参数最小二乘估计结果  $X_{\text{真}} = [-0.1030 \ 2.3208 \ -1.2069 \ -0.5348] \text{dm}$  作为参数真值, 并以各方案参数估值与真值之差的范数  $\|\hat{X} - X_{\text{真}}\|$  作为各方案精度优劣的衡量指标, 一并列入表 1 中。另外, 为展示正常模式下标准化残差属于最大程度受粗差污染的标准化残差模糊子集的隶属函数曲线特性, 本算例还绘制了不附加粗差情况下的最小二乘估计标准化残差的隶属函数曲线, 如图 2 所示。



**Figure 2.** Membership curve of normal standardized residuals

**图 2.** 正常模式下标准化残差的隶属度曲线

**Table 1.** Parameter estimation results  
**表 1.** 参数估计结果

实验	方案	参数估值 $\hat{X}/dm$	$\ \hat{X} - X_{真}\ $
实验 1	方案 1	[-0.3652 1.4230 -1.0294 -0.5342]	0.9688
	方案 2	[-0.1717 1.5084 -0.7993 -0.3643]	0.9286
	方案 3	[-0.068 1.8802 -0.8420 -0.3805]	0.5952
实验 2	方案 1	[-0.4248 1.3558 -1.2606 -0.5807]	1.0213
	方案 2	[-0.1207 1.4940 -0.7448 -0.3227]	0.9742
	方案 3	[-0.0652 1.8297 -0.8127 -0.3465]	0.6654
实验 3	方案 1	[-0.6906 1.6319 -0.7076 0.2329]	1.2887
	方案 2	[-0.5272 1.1558 -0.8589 -0.4233]	1.2942
	方案 3	[-0.0872 1.6886 -0.6921 -0.2333]	0.8704

## 4.2. 算例分析

从图 2 可以看出, 当观测值中不包含粗差时, 最大程度属于粗差的模糊子集  $A$  的标准化残差的隶属函数  $\mu_A(w)$  是以 0 和 1 为渐近线的二维平面曲线, 且当  $w \in [0, 1.5]$  时,  $\mu_A(w)$  缓慢增加, 从  $\mu_A(0) = 0$  增加到  $\mu_A(1.5) = 0.104$ ;  $w \in (1.5, 2.7]$  时,  $\mu_A(w)$  急剧增加, 从  $\mu_A(1.5) = 0.104$  增加到  $\mu_A(2.7) = 0.9840$ 。特别地, 若以 2 倍中误差作为淘汰界线, 此时, 计算  $\mu_A(1.9999) = 0.421$ ,  $\mu_A(2.0001) = 0.421$ , 说明 MF-IGGII 利用标准化残差的隶属函数进行分段定权, 能很好地处理以 2 倍中误差为阈值的模糊界线问题;  $w \in (2.7, +\infty)$  时,  $\mu_A(w)$  缓慢增长, 例如  $\mu_A(2.7) = 0.984$ ,  $\mu_A(3) = 0.998$ ,  $\mu_A(3.5) = 0.999$ ,  $\mu_A(3.7) = 1.000$ ,  $\mu_A(4.0) = 1.000$ , 直至以 0 为渐近线。可见, 正常模式下, 本文采用的标准化残差隶属函数与测量实际以及现行测量规范关于测量限差的规定都是非常吻合的。

从表 1 给出的数值实验结果来看, 在实验 1 和实验 2 中, 方案 2 和方案 3 的参数估计结果都一致地优于方案 1, 表明最小二乘估计不具备抗差能力; 就方案 2 和方案 3 估计结果来看, 当观测粗差从实验 1 的 4 个增加到实验 2 的 5 个时, 方案 3 都比方案 2 的参数估计结果更优, 这说明当量级相同且粗差个数相同时, MF-IGGIII 抗差方法优于 IGGIII 抗差方法; 而在实验 3 中, 当观测粗差增加至 6 个时, 方案 1 和方案 2 估计结果的精度相当, 且参数估值与真值相比都发生扭曲, 说明当粗差数目较多(达到总数的 1/3)时, IGGIII 抗差估计结果变得很不可靠, 而方案 3 此时仍保持了与实验 1 中只有 4 个粗差时 IGGIII 抗差的精度水平, 说明方案 3 具有更稳定的抗差性能。

## 5. 结论

1) 经典 IGGIII 抗差估计将残差分为正常、可疑和淘汰 3 段进行选权迭代过程, 忽略了最小二乘残差是属于粗差这一概念本身具有模糊性的特点, 应当采用模糊数学的方法处理粗差更为合理;

2) 本文引入最大程度受粗差污染的标准化残差的模糊子集的概念, 并提出利用标准化残差属于该模糊子集的隶属函数值代替 IGGIII 抗差估计中直接以标准化残差为依据定义等价权函数的 MF-IGGIII 改进抗差估计方法, 新方法充分顾及了残差属于粗差这一概念本身的模糊性, 有效解决了以往硬阈值条件下属于粗差的程度近似相同但降权幅度却差异较大甚至被错误淘汰的不合理情况发生;

3) 观测值中包含粗差的个数、量级大小对于抗差估计的影响显著, 粗差量级越大, 数目越多, 抗差估计结果会随之变差, 但当粗差数目较少时, MF-IGGII 参数估计结果及精度都显著地优于 IGGIII; 当粗差数目达到上限(本文实验 3 粗差个数达到 6 个, 占观测值总量的 1/3)时, IGGIII 抗差性能急剧下降, 仅与最小二乘估计结果相当。MF-IGGII 抗差参数估计结果及精度较实验 1 和实验 2 也相对变差, 但仍然达到了与实验 1 中仅包含 4 个粗差的 IGGIII 抗差结果的精度水平。总体来看, 3 种不同数目的粗差处理实验结果表明, 与传统 IGGIII 抗差估计方法相比, 参数估计值与真值差值范数指标在三个实验中, 其精度分别提升了 35.9%、32.5%和 32.7%, 表明其抗差性能相对比较稳定, 结果更加可靠。

## 基金项目

1. 贵州省科学技术基础研究计划项目(黔科合基础[2017]1054);
2. 贵州省科技厅科技支撑计划项目: 高分四号卫星和 CYGNSS 协同的贵州多云山区土壤湿度智能反演方法研究(黔科合支撑[2022]一般 204);
3. 资源与环境测绘工程研究生工作站(2024GZZ17)。

## 参考文献

- [1] 黄维彬. 测量平差的当代进展—近代测量平差[J]. 测绘通报, 1994(2): 3-9.
- [2] 黄维彬. 近代平差理论及其应用[M]. 北京: 解放军出版社, 1992.
- [3] 沈云中, 陶本藻. 实用测量数据处理方法[M]. 第 2 版. 北京: 测绘出版社, 2012.
- [4] 杨元喜. 抗差估计理论及其应用[M]. 北京: 八一出版社, 1993.
- [5] 张勤, 张菊清, 岳东杰. 近代测量数据处理与应用[M]. 北京: 测绘出版社, 2011.
- [6] Karl-Rudolf, K. (1988) Parameter Estimation and Hypothesis Testing in Linear Models. Springer.
- [7] Baarda, W. (1968) A Testing Procedure for Use in Geodetic Networks. <https://ncgeo.nl/wp-content/uploads/2024/06/09Baarda.pdf>
- [8] Huber, P.J. (1972) The 1972 Wald Lecture Robust Statistics: A Review. *The Annals of Mathematical Statistics*, **43**, 1041-1067. <https://doi.org/10.1214/aoms/1177692459>
- [9] Hampel, F.R. (1973) Robust Estimation: A Condensed Partial Survey. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **27**, 87-104. <https://doi.org/10.1007/bf00536619>
- [10] Huber, P.J. (1992) Robust Estimation of a Location Parameter. In: *Springer Series in Statistics*, Springer, 492-518. [https://doi.org/10.1007/978-1-4612-4380-9\\_35](https://doi.org/10.1007/978-1-4612-4380-9_35)
- [11] 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]. 第 3 版. 武汉: 武汉大学出版社, 2014.
- [12] 谢季坚, 刘承平. 模糊数学方法及其应用[M]. 武汉: 华中科技大学出版社, 2013.
- [13] 周江文. 经典误差理论与抗差估计[J]. 测绘学报, 1989(2): 115-120.
- [14] 徐爱功, 唐龙江, 徐宗秋, 等. 一种拟合 GPS 高程异常的新方法[J]. 辽宁工程技术大学学报(自然科学版), 2018, 37(3): 591-594.
- [15] 王新洲. 顾及模糊逻辑关系的稳健估计[J]. 武汉测绘科技大学学报, 1996, 21(4): 338-343.