# 再生权最小二乘法稳健估计

葛永慧　著

科学出版社

# 再生权最小二乘法稳健估计

葛永慧　著

科学出版社
北京

## 内容简介

本书根据作者多年从事测量数据处理的教学与研究工作成果撰写而成。首先，介绍了再生权最小二乘法的基本原理和计算方法，讨论了再生权最小二乘法稳健估计在测量控制网平差、多元线性回归和坐标系统转换中的具体应用。然后，介绍了一种确定稳健估计方法的稳健特性的仿真实验方法，并通过仿真实验，讨论了再生权最小二乘法和 13 种常用稳健估计方法的稳健特性以及它们中相对更为有效的稳健估计方法。最后，对测量控制网平差、一元线性和非线性回归的有关问题进行了讨论。

本书可作为高等院校测绘类相关专业研究生用书，也可供涉及参数估计的相关专业人员和广大工程技术人员参考。

图书在版编目(CIP)数据

再生权最小二乘法稳健估计/葛永慧著. —北京：科学出版社，2015
ISBN 978-7-03-043487-6

Ⅰ.①再… Ⅱ.①葛… Ⅲ.①测绘学-最小二乘法-研究 Ⅳ.①P2-0

中国版本图书馆 CIP 数据核字(2015)第 038204 号

责任编辑：裴 育 / 责任校对：桂伟利
责任印制：张 倩 / 封面设计：蓝正设计

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

三河市骏杰印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2015 年 2 月第 一 版 开本：720×1000 1/16
2015 年 2 月第一次印刷 印张：18
字数：412 000

定价：98.00 元

(如有印装质量问题，我社负责调换)

# 前　　言

参数估计在生产实践与科学实验中具有十分重要的作用。统计学家根据大量数据分析指出，粗差出现的概率约占观测值总数的1%～10%，对含粗差观测值的处理主要采用两种方法，其一是将含粗差观测值视为期望异常，用统计检验法剔除含粗差观测值后再进行参数估计；其二是将含粗差观测值视为方差异常，采用稳健估计方法处理。最早引起重视的是统计检验法，归纳起来是一个辨别、定位和调节改正的过程，其实质是假设观测误差服从均值漂移模型，将粗差归于函数模型处理。当存在多个粗差，且系统结构不佳时，仅仅依靠最小二乘法的残差检测来定位粗差的办法具有很大的局限性。稳健估计方法则能有效地消除或减弱粗差对参数估计的影响，稳健估计理论是建立在符合于观测数据的实际分布模式上，而不是建立在某种理想的分布模式上，主要包括M估计(极大似然估计)、L估计(排序线性组合估计)和R估计(秩估计)等。其中，基于M估计的稳健估计方法是应用最为广泛的方法，其基本原理是用观测值的改正数(残差)构造观测值的权(等价权)，用最小二乘法通过迭代消除或减弱粗差对参数估计的影响。权函数的选取是基于M估计的稳健估计方法的核心，很多学者进行了大量的研究，提出了各种形式的权函数并构造了相应的稳健估计方法，如Huber法、L1法、Danish法和IGG方案等。

再生权最小二乘法是一种稳健估计方法，它充分利用观测值之间的条件方程所提供的有效信息构造观测值的权，而不像传统基于M估计的稳健估计方法那样用最小二乘法得到的观测值改正数(残差)构造观测值的权，因此能更有效地消除或减弱粗差对参数估计的影响。本书介绍了再生权最小二乘法的基本原理和计算方法，讨论了再生权最小二乘法在测量控制网平差(不同规模的水准网和测边网)、多元线性回归(一元至五元)和坐标系统转换(四参数和七参数模型坐标系统转换、相似变换和仿射变换图像纠正等)中的具体应用；提供了再生权最小二乘法、最小二乘法和不同稳健估计方法在测量控制网平差、多元线性回归、坐标系统转换中的多个算例，对测量控制网平差、一元线性和非线性回归的有关问题进行了讨论。

稳健估计方法消除或减弱粗差对参数估计影响的能力，即稳健估计方法的稳健特性，取决于稳健估计方法自身、具体的参数估计问题及其观测值的数量等。本书介绍了一种确定稳健估计方法的稳健特性的仿真实验方法，包括等权或不等权的包含粗差的随机误差的生成、任意两种稳健估计方法比较的绝对指标(观测值估值的残余真误差均方误差)和相对指标(稳健估计方法之间的相对增益)。对于测

量控制网平差、多元线性回归和坐标系统转换，通过所述仿真实验方法，讨论了当观测值中包含粗差时再生权最小二乘法和13种常用稳健估计方法的稳健特性以及它们中相对更为有效的稳健估计方法；当观测值中不包含粗差时它们相对于最小二乘法的精度损失。对于测量控制网平差，通过进一步的仿真实验，讨论了在不同的误差分布模式（广义高斯分布的形状参数分别是1.5、2.0和2.5）下和观测值不等权时，再生权最小二乘法和常用稳健估计方法的稳健特性以及它们中相对更为有效的稳健估计方法。

本书根据作者多年从事测量数据处理的教学与研究工作成果撰写而成，基础理论是测量平差与稳健估计，而应用不限于测绘类专业。本书既可作为高等院校测绘类相关专业研究生用书，还可供涉及参数估计的相关专业人员和广大工程技术人员参考。

在本书的撰写过程中，得到了许多同行和同事的热心帮助，并提出了许多宝贵意见，姜佃高、赵晓囡、魏世丽和王晓辉对本书的插图和表格进行了整理与校对，作者在此深表感谢。对本书参阅和引用的有关文献资料的作者同样表示真诚的感谢。

限于作者水平，书中难免有不妥之处，敬请读者批评指正。

# 目　　录

# 第 1 章　稳健估计

## 1.1　稳健估计概述

测量都具有观测误差。观测误差分为三类：一类是具有随机性的偶然误差；一类是带有规律性的系统误差；此外还有粗差(outlier 或 gross error)，泛指离群的误差[1]。统计学家根据大量观测数据分析指出，在生产实践和科学实验中，粗差出现的概率约占观测总数的 1%～10%[2]。这些少量的粗差会对参数估计结果造成严重的干扰。随着科学技术的进步，人们对测量结果的精度要求越来越高。因此，寻求有效的方法消除或减弱粗差显得越来越重要。

目前，对含粗差观测值的处理主要采用两种方法：其一是将含粗差观测值视为期望异常，用统计检验方法剔除含粗差的观测值后再用最小二乘法进行处理；其二是将含粗差观测值视为方差异常，采用稳健估计方法处理[3]。最早引起重视的是统计检验法，归纳起来是一个辨别、定位和调节改正的过程。其实质是假设观测误差服从均值漂移模型，将粗差归于函数模型处理。当存在多个粗差，且系统结构不佳时，仅仅依靠最小二乘法的残差检测来定位粗差的办法具有很大的局限性[1]。有鉴于此，稳健估计的理论和方法应运而生。

稳健估计(robust estimation)也称抗差估计，是指在粗差不可避免的情况下，选择适当的估计方法，使参数估值尽可能地减免其影响，得出正常模式下的最优或接近最优的参数估值[4]。

早在 19 世纪初，已有学者提出了减免粗差干扰的估计方法。但直到 20 世纪五六十年代，随着电子计算机的发展，稳健估计理论和方法的研究才得以深入。Box 于 1953 年首次提出“稳健性”(robustness)的概念。随后，Tukey 于 1960 年提出了污染分布模式。Huber[5] 于 1964 年发表“定位参数的稳健估计”一文，提出了 M 估计理论。Hampel 于 1968 年提出了影响函数和崩溃点的概念。Holland 和 Welsch[6] 于 1977 年提出了选权迭代法。Stiger[7] 于同年提出了中位数估计法。之后，Stiger 与 Bloomfield 又提出了 $L_1$ 估计法(本书中记为 L1 法)。Huber[8]、Hampel[9]、Rousseeuw 和 Leroy[10] 等均对稳健估计进行了卓有成效的研究，并先后发表了有影响力的论著，为稳健估计奠定了理论基础。经过众多数理统计学家不断地开拓和研究，稳健估计理论深入发展，成为应用到众多学科的分支科学。

丹麦的 Krarup 和 Kubik 等于 1980 年将稳健估计理论引入测量界，提出了著

名的“丹麦法”。由于稳健估计方法能够较好地处理测量数据中含有粗差的问题，大地测量界掀起了稳健估计的研究热潮，产生了大量有价值的研究成果。

1983 年，Rousseeuw 等提出了最小剪切二乘法（LTS 法）。LTS 法对杠杆点具有很好的抵抗性，但是计算效率比较低。随后，Rousseeuw[11]等又提出了最小中位数二乘法（LMS 法）、S 估计法和 $\tau$ 估计法。Yohai 于 1987 年提出 MM 估计法，在保证 M 估计稳健性的前提下，提高了 M 估计的计算效率。1989 年，周江文[12,13]提出等价权的概念，将 M 估计最小二乘化，使传统最小二乘法具备了抗差能力，并提出两种有效的估计方案——IGGI 方案和 IGGII 方案。杨元喜[14,15]对等价权原理进行了扩充，提出了 IGGIII 方案，并且针对相关等价权不对称的问题，构造了双因子方差膨胀模型和双因子等价权模型，导出了各种平差模型的参数抗差估计公式。徐培亮[16]也给出了相关观测的稳健估计方法。刘经南和姚宜斌等[17]提出了基于等价方差-协方差阵的稳健最小二乘估计理论，这种方法不仅可以控制观测异常的影响，而且保持了原有观测的相关性不变。欧吉坤[18]提出了一种三步抗差方案，用分步变常数法提高了参数估值的计算效率。为了控制设计空间误差的影响，提出了杠杆点评估和设计空间抗差的 IGGIV 方案。徐培亮[19]提出了符号约束的抗差估计。王志忠和朱建军[20]等研究了适合污染误差模型估计的最优性准则，提出了均方差极小原则下的参数抗差估计。杨元喜[21,22]提出了依据误差分布实际情形的自适应抗差估计和抗差方差分量估计，导出了抗差拟合推估解法。针对病态性与粗差同时存在的问题，Nyquist 和 Silvapulle 提出了基于 M 估计的抗差岭估计。隋立芬[23,24]对其原理和性质进行了研究，提出了抗差组合主成分估计和抗差单参数主成分估计。归庆明等[25]运用有偏估计的压缩变换方法，提出了压缩型抗差估计。彭军还[26]证明了基于误差方差膨胀模型与基于误差均值漂移模型所得到的无偏估计公式的等价性。$L_p$ 估计作为一类重要的抗差估计，也得到了广泛而深入的研究。孙海燕[27]和周世健[28,29]等研究了 $p$-范分布的密度函数，$L_p$ 估计的抗差性和效率，误差分布和估计方法之间的关系。周秋生[30]提出了利用线性规划求解 $L_\infty$ 估计问题的方法，并且依据线性规划的对偶原理给出了求解问题的实用方法。在动态数据处理方面，杨元喜[31~33]提出了抗差 Kalman 滤波，分析了多种抗差滤波的理论基础，讨论了抗差自适应滤波解的性质，构建了抗差自适应滤波理论体系。

## 1.2 最小 $N$ 乘法原理

### 1. 最小二乘法

最小二乘法，又称最小平方法，是一种数学优化技术。它是通过最小化误差的

平方和寻找数据的最佳函数匹配。自 Gauss 于 1809 年提出以来,最小二乘法广泛应用于测量及其他科学工程领域。

在测量数据处理中,Gauss-Markov 模型是最常见的模型之一。其基本模型是

$$\mathrm{E}(\boldsymbol{L}) = \boldsymbol{B}\widetilde{\boldsymbol{X}}, \quad \boldsymbol{D}_L = \sigma_0^2 \boldsymbol{P}^{-1} \tag{1-1}$$

模型(1-1)还可表示为

$$\boldsymbol{L} = \boldsymbol{B}\widetilde{\boldsymbol{X}} - \boldsymbol{\Delta}, \quad \mathrm{E}(\boldsymbol{\Delta}) = 0, \quad \boldsymbol{D}_\Delta = \boldsymbol{D}_L = \sigma_0^2 \boldsymbol{P}^{-1} \tag{1-2}$$

测量平差中一般将式(1-2)表示为

$$\boldsymbol{V} = \boldsymbol{B}\hat{\boldsymbol{X}} - \boldsymbol{L}, \quad \boldsymbol{D}_L = \sigma_0^2 \boldsymbol{P}^{-1} \tag{1-3}$$

式(1-3)称为误差方程。上式中,$\boldsymbol{L}$ 表示 $n\times 1$ 阶观测值矩阵;$\boldsymbol{P}(n\times n)$ 表示观测值 $\boldsymbol{L}$ 的权阵;$\boldsymbol{D}_L$ 表示观测值 $\boldsymbol{L}$ 的协方差阵;$\sigma_0^2$ 表示单位权方差;$\boldsymbol{\Delta}$ 表示观测值 $\boldsymbol{L}$ 的真误差;$\boldsymbol{V}$ 表示观测值 $\boldsymbol{L}$ 的改正数,是真误差 $\boldsymbol{\Delta}$ 的估值;$\boldsymbol{B}$ 表示 $n\times t$ 阶系数矩阵;$\widetilde{\boldsymbol{X}}$ 表示 $t\times 1$ 阶未知数真值矩阵;$\hat{\boldsymbol{X}}$ 表示 $t\times 1$ 阶未知数估值矩阵。

按最小二乘法求解 Gauss-Markov 模型中的参数估值 $\hat{\boldsymbol{X}}$,即是要求准则函数

$$\Phi = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{V} = (\boldsymbol{B}\hat{\boldsymbol{X}} - \boldsymbol{L})^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{B}\hat{\boldsymbol{X}} - \boldsymbol{L}) = \min(\text{最小}) \tag{1-4}$$

将 $\Phi$ 对 $\hat{\boldsymbol{X}}$ 求导并令其为零,得

$$\frac{\mathrm{d}\Phi}{\mathrm{d}\hat{\boldsymbol{X}}} = 2\boldsymbol{V}^{\mathrm{T}}\boldsymbol{P}\frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}\hat{\boldsymbol{X}}} = \boldsymbol{0} \tag{1-5}$$

考虑到式(1-3),有$\frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}\hat{\boldsymbol{X}}}=\boldsymbol{B}$,则

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{V} = \boldsymbol{0} \tag{1-6}$$

将式(1-3)代入得

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{B}\hat{\boldsymbol{X}} - \boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{L} = \boldsymbol{0} \tag{1-7}$$

令

$$\boldsymbol{N} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{B}, \quad \boldsymbol{W} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{L}$$

则式(1-7)写成

$$\boldsymbol{N}\hat{\boldsymbol{X}} = \boldsymbol{W} \tag{1-8}$$

式(1-8)称为法方程(normal equations),其解为

$$\hat{\boldsymbol{X}} = \boldsymbol{N}^{-1}\boldsymbol{W} = (\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{L} \tag{1-9}$$

由式(1-9)求得的参数估值确保了 $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{V}=\min$。

将式(1-9)代入式(1-3)得观测值的改正数 $\boldsymbol{V}$ 和观测值的估值 $\hat{\boldsymbol{L}}$:

$$\boldsymbol{V} = \boldsymbol{B}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{L} - \boldsymbol{L} \tag{1-10}$$

$$\hat{\boldsymbol{L}} = \boldsymbol{L} + \boldsymbol{V} \tag{1-11}$$

单位权中误差的估值：

$$\hat{\sigma}_0 = \sqrt{\frac{\boldsymbol{V}^{\mathrm{T}}\boldsymbol{PV}}{n-t}} \tag{1-12}$$

未知数的协因数阵：

$$\boldsymbol{Q}_{\hat{X}\hat{X}} = \boldsymbol{N}^{-1} = \begin{bmatrix} q_{\hat{X}_1\hat{X}_1} & q_{\hat{X}_1\hat{X}_2} & \cdots & q_{\hat{X}_1\hat{X}_t} \\ q_{\hat{X}_2\hat{X}_1} & q_{\hat{X}_2\hat{X}_2} & \cdots & q_{\hat{X}_2\hat{X}_t} \\ \vdots & \vdots & & \vdots \\ q_{\hat{X}_t\hat{X}_1} & q_{\hat{X}_t\hat{X}_2} & \cdots & q_{\hat{X}_t\hat{X}_t} \end{bmatrix} \tag{1-13}$$

未知数 $\hat{X}_i$ 的中误差：

$$\hat{\sigma}_{\hat{X}_i} = \hat{\sigma}_0 \sqrt{q_{\hat{X}_i\hat{X}_i}}, \quad i = 1,2,\cdots,t \tag{1-14}$$

应用最小二乘准则时，并不需要知道观测向量服从何种概率分布，而只需知道它的先验权阵即可。

当 $\boldsymbol{P}$ 为非对角阵时，表示观测值相关，按 $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{PV}=\min$ 进行的平差称为相关观测平差。

当 $\boldsymbol{P}$ 为对角阵时，表示观测值不相关，此时最小二乘准则可表示为纯量形式，即

$$\Phi = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{PV} = p_1 v_1^2 + p_2 v_2^2 + \cdots + p_n v_n^2 = \min$$

特别地，当观测值不相关且等精度时，权阵 $\boldsymbol{P}$ 为单位阵，此时最小二乘准则可表示为

$$\Phi = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{PV} = v_1^2 + v_2^2 + \cdots + v_n^2 = \min$$

### 2. 最小 $N$ 乘法

当观测值不相关且等精度时，最小 $N$ 乘准则函数为

$$\sum_{i=1}^{n} |v_i|^N = |v_1|^N + |v_2|^N + \cdots + |v_n|^N = \min \tag{1-15}$$

当 $N=1$ 时，即为最小一乘法的准则函数：

$$\sum_{i=1}^{n} |v_i| = |v_1| + |v_2| + \cdots + |v_n| = \min \tag{1-16}$$

当 $N=2$ 时，即为最小二乘法的准则函数：

$$\sum_{i=1}^{n} v_i^2 = v_1^2 + v_2^2 + \cdots + v_n^2 = \min \tag{1-17}$$

当 $N=3$ 时，即为最小三乘法的准则函数：

$$\sum_{i=1}^{n} |v_i|^3 = |v_1|^3 + |v_2|^3 + \cdots + |v_n|^3 = \min \tag{1-18}$$

当 $N=4$ 时，即为最小四乘法的准则函数：

$$\sum_{i=1}^{n} v_i^4 = v_1^4 + v_2^4 + \cdots + v_n^4 = \min \tag{1-19}$$

…

当 $N\to\infty$ 时，即为最小无穷乘法的准则函数：

$$\max |v| = \min \tag{1-20}$$

3. 算例

设有线性方程组

$$\begin{aligned} v_1 - v_2 + v_4 + 9.0 &= 0 \\ v_2 - v_3 + v_5 - 11.0 &= 0 \\ v_1 - v_3 - v_6 + 5.0 &= 0 \end{aligned} \tag{1-21}$$

式(1-21)中，方程的数量 $r=3$，未知数的数量 $n=6$。未知数的数量大于方程的数量，所以未知数具有无穷多组解。表 1.1 列出了 $N$ 为 1、2、3、4 和无穷大时式(1-21)未知数(精确到 0.1)的部分解。由表 1.1 可知，当未知数的解精确到 0.1 时，最小一乘法有 31 组解(第 1～31 行)，约束条件式(1-16)的值为 14.0；最小二乘法有 1 组解(第 32 行)，约束条件式(1-17)的值为 59.0；最小三乘法有 1 组解(第 33 行)，约束条件式(1-18)的值为 214.3；最小四乘法有 1 组解(第 34 行)，约束条件式(1-19)的值为 765.7；最小无穷乘法有 25 组解(第 35～59 行)，约束条件式(1-20)的值为 3.7。

(1) 不同的约束条件下未知数的解是不尽相同的。

(2) 最小一乘法没有唯一解，但它是测量数据处理特别是粗差探测的有效方法之一。可用选权迭代法或单纯形法进行解算。

(3) 最小二乘法是应用十分广泛的参数估计方法，也是测量数据处理的基础方法。最小二乘法的实质是在未知数无穷多组解中确定满足约束条件的唯一的一组解。可用求条件极值或自由极值的方法进行解算。式(1-21)的最小二乘法解是

$$[v_1 \quad v_2 \quad v_3 \quad v_4 \quad v_5 \quad v_6]^{\mathrm{T}} = [-1.0 \quad 5.0 \quad -4.0 \quad -3.0 \quad 2.0 \quad -2.0]^{\mathrm{T}} \tag{1-22}$$

(4) 最小三乘法、最小四乘法等通常不用。式(1-21)的最小三乘法和最小四乘法的解分别是

$$[v_1 \quad v_2 \quad v_3 \quad v_4 \quad v_5 \quad v_6]^{\mathrm{T}} = [-1.6 \quad 4.2 \quad -3.9 \quad -3.2 \quad 2.9 \quad -2.7]^{\mathrm{T}} \tag{1-23}$$

$$[v_1 \quad v_2 \quad v_3 \quad v_4 \quad v_5 \quad v_6]^{\mathrm{T}} = [-1.8 \quad 4.0 \quad -3.8 \quad -3.2 \quad 3.2 \quad -3.0]^{\mathrm{T}} \tag{1-24}$$

(5) 最小无穷乘法没有唯一解，通常不用。

**表 1.1 方程组(1-21)的部分解**

| 序号 | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ | $v_6$ | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_\infty$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.0 | 6.0 | −5.0 | −3.0 | 0.0 | 0.0 | **14.0** | 70.0 | 368.0 | 2002.0 | 6.0 |
| 2 | 0.0 | 6.1 | −4.9 | −2.9 | 0.0 | −0.1 | **14.0** | 69.6 | 369.0 | 2031.8 | 6.1 |
| 3 | 0.0 | 6.2 | −4.8 | −2.8 | 0.0 | −0.2 | **14.0** | 69.4 | 370.9 | 2069.9 | 6.2 |
| 4 | 0.0 | 6.3 | −4.7 | −2.7 | 0.0 | −0.3 | **14.0** | 69.2 | 373.6 | 2116.4 | 6.3 |
| 5 | 0.0 | 6.4 | −4.6 | −2.6 | 0.0 | −0.4 | **14.0** | 69.0 | 377.1 | 2171.2 | 6.4 |
| 6 | 0.0 | 6.5 | −4.5 | −2.5 | 0.0 | −0.5 | **14.0** | 69.0 | 381.5 | 2234.3 | 6.5 |
| 7 | 0.0 | 6.6 | −4.4 | −2.4 | 0.0 | −0.6 | **14.0** | 69.0 | 386.7 | 2305.6 | 6.6 |
| 8 | 0.0 | 6.7 | −4.3 | −2.3 | 0.0 | −0.7 | **14.0** | 69.2 | 392.8 | 2385.2 | 6.7 |
| 9 | 0.0 | 6.8 | −4.2 | −2.2 | 0.0 | −0.8 | **14.0** | 69.4 | 399.7 | 2473.1 | 6.8 |
| 10 | 0.0 | 6.9 | −4.1 | −2.1 | 0.0 | −0.9 | **14.0** | 69.6 | 407.4 | 2569.4 | 6.9 |
| 11 | 0.0 | 7.0 | −4.0 | −2.0 | 0.0 | −1.0 | **14.0** | 70.0 | 416.0 | 2674.0 | 7.0 |
| 12 | 0.0 | 7.1 | −3.9 | −1.9 | 0.0 | −1.1 | **14.0** | 70.4 | 425.4 | 2787.0 | 7.1 |
| 13 | 0.0 | 7.2 | −3.8 | −1.8 | 0.0 | −1.2 | **14.0** | 71.0 | 435.7 | 2908.5 | 7.2 |
| 14 | 0.0 | 7.3 | −3.7 | −1.7 | 0.0 | −1.3 | **14.0** | 71.6 | 446.8 | 3038.5 | 7.3 |
| 15 | 0.0 | 7.4 | −3.6 | −1.6 | 0.0 | −1.4 | **14.0** | 72.2 | 458.7 | 3177.0 | 7.4 |
| 16 | 0.0 | 7.5 | −3.5 | −1.5 | 0.0 | −1.5 | **14.0** | 73.0 | 471.5 | 3324.3 | 7.5 |
| 17 | 0.0 | 7.6 | −3.4 | −1.4 | 0.0 | −1.6 | **14.0** | 73.8 | 485.1 | 3480.3 | 7.6 |
| 18 | 0.0 | 7.7 | −3.3 | −1.3 | 0.0 | −1.7 | **14.0** | 74.8 | 499.6 | 3645.1 | 7.7 |
| 19 | 0.0 | 7.8 | −3.2 | −1.2 | 0.0 | −1.8 | **14.0** | 75.8 | 514.9 | 3818.9 | 7.8 |
| 20 | 0.0 | 7.9 | −3.1 | −1.1 | 0.0 | −1.9 | **14.0** | 76.8 | 531.0 | 4001.9 | 7.9 |
| 21 | 0.0 | 8.0 | −3.0 | −1.0 | 0.0 | −2.0 | **14.0** | 78.0 | 548.0 | 4194.0 | 8.0 |
| 22 | 0.0 | 8.1 | −2.9 | −0.9 | 0.0 | −2.1 | **14.0** | 79.2 | 565.8 | 4395.5 | 8.1 |
| 23 | 0.0 | 8.2 | −2.8 | −0.8 | 0.0 | −2.2 | **14.0** | 80.6 | 584.5 | 4606.5 | 8.2 |
| 24 | 0.0 | 8.3 | −2.7 | −0.7 | 0.0 | −2.3 | **14.0** | 82.0 | 604.0 | 4827.2 | 8.3 |
| 25 | 0.0 | 8.4 | −2.6 | −0.6 | 0.0 | −2.4 | **14.0** | 83.4 | 624.3 | 5057.7 | 8.4 |
| 26 | 0.0 | 8.5 | −2.5 | −0.5 | 0.0 | −2.5 | **14.0** | 85.0 | 645.5 | 5298.3 | 8.5 |
| 27 | 0.0 | 8.6 | −2.4 | −0.4 | 0.0 | −2.6 | **14.0** | 86.6 | 667.5 | 5549.0 | 8.6 |
| 28 | 0.0 | 8.7 | −2.3 | −0.3 | 0.0 | −2.7 | **14.0** | 88.4 | 690.4 | 5810.1 | 8.7 |
| 29 | 0.0 | 8.8 | −2.2 | −0.2 | 0.0 | −2.8 | **14.0** | 90.2 | 714.1 | 6081.9 | 8.8 |
| 30 | 0.0 | 8.9 | −2.1 | −0.1 | 0.0 | −2.9 | **14.0** | 92.0 | 738.6 | 6364.4 | 8.9 |
| 31 | 0.0 | 9.0 | −2.0 | 0.0 | 0.0 | −3.0 | **14.0** | 94.0 | 764.0 | 6658.0 | 9.0 |
| 32 | −1.0 | 5.0 | −4.0 | −3.0 | 2.0 | −2.0 | 17.0 | **59.0** | 233.0 | 995.0 | 5.0 |
| 33 | −1.6 | 4.2 | −3.9 | −3.2 | 2.9 | −2.7 | 18.5 | 61.4 | **214.3** | 777.8 | 4.2 |
| 34 | −1.8 | 4.0 | −3.8 | −3.2 | 3.2 | −3.0 | 19.0 | 63.2 | 217.2 | **765.7** | 4.0 |
| 35 | −2.4 | 3.6 | −3.7 | −3.0 | 3.7 | −3.7 | 20.1 | 68.8 | 239.4 | 844.4 | **3.7** |
| 36 | −2.4 | 3.7 | −3.7 | −2.9 | 3.6 | −3.7 | 20.0 | 68.2 | 236.8 | 834.1 | **3.7** |
| 37 | −2.3 | 3.6 | −3.7 | −3.1 | 3.7 | −3.6 | 20.0 | 68.2 | 236.6 | 831.1 | **3.7** |

续表

| 序号 | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ | $v_6$ | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_\infty$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 38 | −2.3 | 3.7 | −3.7 | −3.0 | 3.6 | −3.6 | 19.9 | 67.6 | 233.8 | 819.7 | **3.7** |
| 39 | −2.3 | 3.7 | −3.6 | −3.0 | 3.7 | −3.7 | 20.0 | 68.3 | 237.8 | 839.2 | **3.7** |
| 40 | −2.2 | 3.6 | −3.7 | −3.2 | 3.7 | −3.5 | 19.9 | 67.7 | 234.3 | 821.1 | **3.7** |
| 41 | −2.2 | 3.7 | −3.7 | −3.1 | 3.6 | −3.5 | 19.8 | 67.0 | 231.3 | 808.6 | **3.7** |
| 42 | −2.2 | 3.7 | −3.6 | −3.1 | 3.7 | −3.6 | 19.9 | 67.8 | 235.1 | 826.5 | **3.7** |
| 43 | −2.1 | 3.6 | −3.7 | −3.3 | 3.7 | −3.4 | 19.8 | 67.2 | 232.5 | 814.5 | **3.7** |
| 44 | −2.1 | 3.7 | −3.7 | −3.2 | 3.6 | −3.4 | 19.7 | 66.6 | 229.3 | 800.7 | **3.7** |
| 45 | −2.1 | 3.7 | −3.6 | −3.2 | 3.7 | −3.5 | 19.8 | 67.2 | 232.9 | 817.2 | **3.7** |
| 46 | −2.0 | 3.6 | −3.7 | −3.4 | 3.7 | −3.3 | 19.7 | 66.8 | 231.2 | 811.0 | **3.7** |
| 47 | −2.0 | 3.7 | −3.7 | −3.3 | 3.6 | −3.3 | 19.6 | 66.1 | 227.8 | 796.0 | **3.7** |
| 48 | −2.0 | 3.7 | −3.6 | −3.3 | 3.7 | −3.4 | 19.7 | 66.8 | 231.2 | 811.0 | **3.7** |
| 49 | −1.9 | 3.6 | −3.7 | −3.5 | 3.7 | −3.2 | 19.6 | 66.4 | 230.5 | 810.8 | **3.7** |
| 50 | −1.9 | 3.7 | −3.7 | −3.4 | 3.6 | −3.2 | 19.5 | 65.8 | 226.9 | 794.3 | **3.7** |
| 51 | −1.9 | 3.7 | −3.6 | −3.4 | 3.7 | −3.3 | 19.6 | 66.4 | 230.1 | 808.1 | **3.7** |
| 52 | −1.8 | 3.6 | −3.7 | −3.6 | 3.7 | −3.1 | 19.5 | 66.2 | 230.2 | 813.6 | **3.7** |
| 53 | −1.8 | 3.7 | −3.7 | −3.5 | 3.6 | −3.1 | 19.4 | 65.4 | 226.5 | 795.7 | **3.7** |
| 54 | −1.8 | 3.7 | −3.6 | −3.5 | 3.7 | −3.2 | 19.5 | 66.1 | 229.4 | 808.2 | **3.7** |
| 55 | −1.7 | 3.6 | −3.7 | −3.7 | 3.7 | −3.0 | 19.4 | 65.9 | 230.5 | 819.6 | **3.7** |
| 56 | −1.7 | 3.7 | −3.7 | −3.6 | 3.6 | −3.0 | 19.3 | 65.2 | 226.5 | 800.1 | **3.7** |
| 57 | −1.7 | 3.7 | −3.6 | −3.6 | 3.7 | −3.1 | 19.4 | 65.8 | 229.3 | 811.5 | **3.7** |
| 58 | −1.6 | 3.7 | −3.7 | −3.7 | 3.6 | −2.9 | 19.2 | 65.0 | 227.1 | 807.5 | **3.7** |
| 59 | −1.6 | 3.7 | −3.6 | −3.7 | 3.7 | −3.0 | 19.3 | 65.6 | 229.7 | 817.8 | **3.7** |

注：$v_1$～$v_6$ 表示未知数；$S_1$、$S_2$、$S_3$、$S_4$ 和 $S_\infty$ 分别表示最小一乘法、最小二乘法、最小三乘法、最小四乘法和最小无穷乘法的约束条件的值。表中每一行是未知数的一组解。

## 1.3　稳健估计原理、方法和计算

### 1.3.1　稳健估计原理[34]

稳健估计讨论问题的方式是：对于实际问题有一个假定模型，同时又认为这个模型并不准确，而只是实际问题理论模型的一个近似。它要求稳健估计方法应达到的三个目标是：

(1) 在假定的观测分布模型下，估值应是最优的或接近最优的。

(2) 当假定的分布模型与实际的理论分布模型有较小差异时，估值受到粗差的影响较小。

(3) 当假设的分布模型与实际的理论分布模型有较大偏离时，估值不至于受

到破坏性影响。

稳健估计是建立在观测数据的实际分布上，而不是理论分布上，这是稳健估计理论与经典估计理论的根本区别。稳健估计的原则是充分利用观测数据中的有效信息，限制利用可用信息，排除有害信息。

1. 极大似然估计准则

设独立观测样本为 $L_1, L_2, \cdots, L_n$，$\boldsymbol{X}$ 为待估参数，$L_i$ 的分布密度为 $f(l_i, \hat{\boldsymbol{X}})$，其极大似然估计准则为

$$f(l_1, l_2, \cdots, l_n, \hat{\boldsymbol{X}}) = f(l_1, \hat{\boldsymbol{X}}) \times f(l_2, \hat{\boldsymbol{X}}) \times \cdots \times f(l_n, \hat{\boldsymbol{X}}) = \max \quad (1\text{-}25)$$

或

$$\sum_{i=1}^{n} \ln f(l_i, \hat{\boldsymbol{X}}) = \max \quad (1\text{-}26)$$

2. 正态分布密度下的极大似然估计准则

设独立观测样本 $L_i \sim N(\mu_i, \sigma^2)$，其密度函数为

$$f(l_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(l_i - \hat{\mu}_i)^2}{2\sigma^2}\right\} \quad (1\text{-}27)$$

参数 $\boldsymbol{X}$ 的极大似然估计准则为

$$f(l_1, l_2, \cdots, l_n, \hat{\boldsymbol{X}}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{\sum_{i=1}^{n}(l_i - \hat{\mu}_i)^2}{2\sigma^2}\right\} = \max \quad (1\text{-}28)$$

或

$$\sum_{i=1}^{n}(l_i - \hat{\mu}_i)^2 = \sum_{i=1}^{n} v_i^2 = \min \quad (1\text{-}29)$$

即正态分布密度下的极大似然估计准则就是最小二乘估计准则。

3. 稳健估计的极大似然估计准则

稳健估计基本可以分为三大类型，即

(1) M 估计：又称为极大似然估计，基于 1964 年 Huber 所提出的 M 估计理论。

(2) L 估计：又称为排序线性组合估计，在测绘界有一定范围应用。

(3) R 估计：又称为秩估计，目前在测绘界应用还很少。

其中，M 估计是测量平差中最主要的稳健估计(抗差估计)准则。

设独立观测样本为 $L_1, L_2, \cdots, L_n$，$\boldsymbol{X}$ 为待估参数，$L_i$ 的分布密度为 $f(l_i, \hat{\boldsymbol{X}})$，

其极大似然估计准则为

$$\sum_{i=1}^{n} \ln f(l_i, \hat{\boldsymbol{X}}) = \max \tag{1-30}$$

若以 $\rho(\cdot)$代替$-\ln f(\cdot)$,则极大似然估计准则可改写为

$$\sum_{i=1}^{n} \rho(l_i, \hat{\boldsymbol{X}}) = \min \tag{1-31}$$

对式(1-31)求导并令其为0,得

$$\sum_{i=1}^{n} \varphi(l_i, \hat{\boldsymbol{X}}) = 0 \tag{1-32}$$

式中,$\varphi(l_i, \hat{\boldsymbol{X}}) = \dfrac{\partial \rho(l_i, \hat{\boldsymbol{X}})}{\partial \hat{\boldsymbol{X}}}$。

因此,M 估计就是指由式(1-30)或式(1-31)定义的一类估计。一个 $\rho$(或 $\varphi$)函数,定义一个 M 估计。通常,取对称、连续、严凸或者在正半轴上非降的函数作为 $\rho$ 函数,取 $\rho$ 函数的导函数作为 $\varphi$ 函数。

确定 $\rho$(或 $\varphi$)函数是 M 估计的关键。作为一种稳健估计方法,在选取 $\rho$ 函数时,必须要求满足参数稳健估计的三个目标。

### 1.3.2 基于选权迭代法的稳健估计方法[35]

M 估计的估计方法有很多种,而选权迭代法在测量平差中应用最广泛,因为其计算简单,算法类似于最小二乘平差,且易于程序实现。

设独立观测值为 $\underset{n\times1}{\boldsymbol{L}}$,未知参数向量为 $\underset{t\times1}{\hat{\boldsymbol{X}}}$,误差方程及权阵为

$$\boldsymbol{V} = \boldsymbol{B}\hat{\boldsymbol{X}} - \boldsymbol{L} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix} \hat{\boldsymbol{X}} - \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} p_1 & & & \\ & p_2 & & \\ & & \ddots & \\ & & & p_n \end{bmatrix} \tag{1-33}$$

式中,$\boldsymbol{b}_i$ 为 $1\times t$ 阶系数向量。

考虑误差方程,M 估计的函数 $\rho(l_i, \hat{\boldsymbol{X}})$可表述为

$$\rho(l_i, \hat{\boldsymbol{X}}) = \rho(v_i) \tag{1-34}$$

1. 独立等权观测的选权迭代法

设式(1-33)中的权阵 $\boldsymbol{P}=\boldsymbol{I}$,即 $p_1 = p_2 = \cdots = p_n = 1$,按 M 估计准则并取 $\rho$ 函数为式(1-34),则有

$$\sum_{i=1}^{n} \rho(v_i) = \min \tag{1-35}$$