

测量数据不确定性度量的最小模糊熵算法

魏冠军¹ 党亚民² 章传银²

1 兰州交通大学测绘与地理信息学院,甘肃 兰州,730070

2 中国测绘科学研究院,北京,100830

摘要: 测量数据的质量及可靠性取决于测量数据不确定性的大小。从如何评价测量数据的不确定性入手,以测量不确定度理论与模糊数学为基础,构建以测量不确定度为未知参数的测量数据不确定性评价的函数模型,提出“模糊熵测度”作为函数模型求解的最优准则并建立相应的算法,应用高程监测网数据进行解算并与最小二乘估计结果进行比较,结果证明了该方法的可行性。

关键词: 数据不确定性; 测量不确定度; 模糊熵准则; 函数模型; 最小模糊熵估计

中图法分类号:P207

文献标志码:A

测量数据质量及可靠性取决于测量数据不确定性的大小。经典的测量平差理论是基于“观测值的不确定性就是随机性”的假设。Helmert、Tienstra、Baarda等学者将测量数据的不确定性分为偶然误差、系统误差和粗差三种类型^[1]。标准差(σ)在一定程度上反映了随机不确定性,也是以 2σ 或 3σ 作为观测限差的理论基础^[2]。随着测量数据获取手段的更新,异常值、有色噪声、不完备信息等新的误差不断产生,测量数据的不确定性并非仅由随机误差组成,而是多种不确定性因素的综合影响,既有可以数值化的不确定性因素(误差、噪声),也有无法数值化、参数化的不确定性因素^[3]。1993年国际标准化组织(International organization for Standardization, ISO)等7个组织引入“测量不确定度”概念,公布了《Guide to the Expression of Uncertainty in Measurement》^[4]。测量不确定度是对测量结果不确定性的定量表征,决定了测量结果质量及可靠性。文献[5,6]基于模糊数学,建立以不确定度作为参数的测量数据质量评价模型;文献[7-11]从不同的角度对测量不确定度理论进行研究。随着测量数据不确定性研究的深入,文献[12-16]将熵理论应用到空间数据不确定性研究之中,信息熵作为信源平均不确定性的度量指标,具有不受置信概率

选取时的主观因素影响的优点,还可以简化某些不确定性问题的数学处理方法。但已有研究大多关注测量数据的随机性,而测量数据的不确定性是随机性、模糊性及不完备性等多种因素的综合。针对此,本文在前人研究的基础上,运用信息熵、测量不确定度及现代优化算法等理论和方法来构建测量数据不确定性度量的模糊熵算法,并开展应用实例计算及对比分析。

1 测量数据不确定性的模糊数学描述

观测量的真值(设为 Y_m)总是客观、唯一存在的,设测量值为 y_i ($i=1,2,\dots,n$),若定义精确集合: $A=\{Y_m\}$,则测量值 y_i 和集合 A 可由特征函数 $G_A(y)$ 来刻画:

$$G_A(y)=\begin{cases} 1, & y_i \in A \\ 0, & y_i \notin A \end{cases} \quad (1)$$

式(1)表明,观测值要么属于集合 A ,要么不属于集合 A 。

由于误差、认识水平等原因,把任何一个测量值判断为真值或者非真值都非常不现实。根据模糊理论,可以把观测值 y_i ($i=1,2,\dots,n$)处理成其真值的一个模糊集 \tilde{A} ,模糊集 \tilde{A} 可根据隶属函数

收稿日期:2015-09-07

项目资助:国家自然科学基金(41201004,41364001);甘肃省自然科学基金(1508RJZA065);兰州交通大学科技支撑计划(ZC2014002);2014甘肃省高等学校科研项目(2014B-040)。

第一作者:魏冠军,博士,副教授。主要从事测量数据处理的理论与算法研究。wchampion@sina.com

通讯作者:党亚民,博士,研究员。dangym@casm.ac.cn

唯一定义:

$$\tilde{A} = \{(y, \mu_{\tilde{A}}(y)) \mid y \in I\mathbb{R}\}$$

$$\mu_{\tilde{A}}(y) : I\mathbb{R} \rightarrow [0, 1] \quad (2)$$

测量值 $y_i (i=1, 2, \dots, n)$ 与模糊集 \tilde{A} 的关系可通过隶属函数 $\mu_{\tilde{A}}(y)$ 来描述^[17], $\mu_{\tilde{A}}(y)$ 作为测量值 $y_i (i=1, 2, \dots, n)$ 与真值的隶属度, 把“观测值 $y_i (i=1, 2, \dots, n)$ 是其真值的可能性”用模糊分布 μ_y 来表示。如图 1 所示, 模糊集的隶属函数可用左 $L(\cdot)$ 、右 $R(\cdot)$ 参照函数表示。

$$\mu_{\tilde{A}}(y) = \begin{cases} L\left(\frac{Y_m - y - r}{c_l}\right), & y < Y_m < -r \\ 1, & Y_m - r \leq y \leq Y_m + r \\ R\left(\frac{y - Y_m - r}{c_r}\right), & y > Y_m + r \end{cases} \quad (3)$$

式中, Y_m 表示真值; r 为核半径; c_l, c_r 表示单调参照函数的参数; $c_l + c_r$ 的值定量描述了真值的不确定性区间。

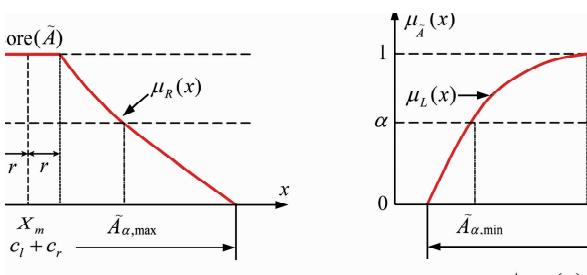


图 1 隶属函数与 α -截集

Fig. 1 Membership Function and Its α -cut

若取 $\alpha \in [0, 1]$, α 为置信水平, 用 \tilde{A}_α 表示 α -截集下的模糊集, 如图 1 所示。若 $\tilde{A}_{\alpha,r}$ 表示模糊区间半径(即一定置信水平下不确定区间的半宽), 则有:

$$\tilde{A}_\alpha := \{y \in I\mathbb{R} \mid \mu_{\tilde{A}}(y) \geq \alpha\} \quad (4)$$

$$\begin{cases} \tilde{A}_{\alpha,\min} = \min(\tilde{A}_\alpha) \\ \tilde{A}_{\alpha,\max} = \max(\tilde{A}_\alpha) \\ \tilde{A}_{\alpha,r} = \frac{\tilde{A}_{\alpha,\max} - \tilde{A}_{\alpha,\min}}{2} \end{cases} \quad (5)$$

考虑到观测值 $y_i (i=1, 2, \dots, n)$ 在其真值 Y_m 两侧出现的可能性相等, 以真值为中心, 则观测值是对称模糊数, 若用 $I = (Y_m, \delta)_L$ 表示对称模糊数 I , 其隶属函数可表示为:

$$\mu_I(y) = L\left(\frac{y - Y_m}{\delta_y}\right), \delta_y > 0 \quad (6)$$

式中, $L(z)$ 为参照函数; $z = \frac{y - Y_m}{\delta_y}$; Y_m 为观测值的真值; δ_y 为模糊幅度, 描述观测值波动范围。

测量不确定度用于表征合理地赋予被测量值

的分散性, 在测量结果中用来说明测量值变化范围的参数; 当观测值用式(6)的对称模糊数表示时, 模糊幅度 δ_y 是描述观测值变化范围的参数。因此, 文献[18]提出用模糊幅度来衡量测量结果的不确定度, 测量数据的不确定性越大, 其模糊幅度也就越大。

2 最小模糊熵估计原理

2.1 基于模糊数的高斯-马尔柯夫模型

在经典平差或近代平差中, 一般均采用高斯-马尔柯夫模型(G-M 模型)来进行测量数据处理, 对于该模型的解算, 一般采用 MA、MV、LS、Bayes 或 L 估计, 这些方法主要根据含有随机误差的观测值来确定未知参数的最佳估值^[19]。由于观测量子样容量限制, 不完全符合概率统计理论, 而是有一定的模糊性(式(6))。当观测向量 \mathbf{Y} 和未知参数向量 \mathbf{X} 均为模糊数时, G-M 模型为:

$$\mathbf{Y} = \mathbf{A} \cdot \mathbf{X} + \mathbf{A}_0 \quad (7)$$

式中, \mathbf{A} 为 $n \times t$ 设计矩阵,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1t} \\ a_{21} & a_{22} & \cdots & a_{2t} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nt} \end{bmatrix}; \mathbf{A}_0 = \begin{bmatrix} a_{01} \\ a_{02} \\ \vdots \\ a_{0n} \end{bmatrix};$$

$\mathbf{X} = (x_1 \ x_2 \ \cdots \ x_t)^T = ((\hat{x}_1, c_1)_L \ (\hat{x}_2, c_2)_L \ \cdots \ (\hat{x}_t, c_t)_L)^T$; $\mathbf{Y} = (y_1 \ y_2 \ \cdots \ y_t)^T = ((\hat{y}_1, \delta_{y_1})_L \ (\hat{y}_2, \delta_{y_2})_L \ \cdots \ (\hat{y}_t, \delta_{y_t})_L)^T$, 其中 \hat{y}_i, δ_{y_i} 分别表示观测值的真值及其测量不确定度; \hat{x}_i, c_i 分别表示参数的真值及其测量不确定度。

2.2 模型求解的模糊熵准则

为了估计式(7)中的未知参数向量 \mathbf{X} , 必须事先确定最优估计准则。从信息论的角度分析, 信息熵是信源不确定性的唯一度量, 模糊熵测度是模糊集合理论中度量模糊子集不确定性的测度之一。第一个与概率无关, 能正确地描述模糊集模糊程度的信息量是文献[20]提出的模糊熵 $e_{LT}(\mathbf{A})$, 即

$$e_{LT}(\mathbf{A}) = -k \sum_{i=1}^n [\mu_A(y_i) \ln \mu_A(y_i) + (1 - \mu_A(y_i)) \ln (1 - \mu_A(y_i))] \quad (8)$$

对每一个观测值而言, 观测值 y_i 越接近真值 \hat{y}_i , “ y_i 是 \hat{y}_i ”的可能性越大, 其模糊隶属度 $\mu_A(y_i)$ 增大(图 2), 模糊性减少, 相应的模糊熵减少。当所有观测值最接于真值时, 即 n 个“ y_i 是 \hat{y}_i ”的可能性都达到最大时, 模糊隶属度达到最

大, 模糊不确定程度就最小, 所对应的模糊熵最小, 即模糊熵最小。

若取 $k=1$, 由式(8)可得到最小模糊熵准则:

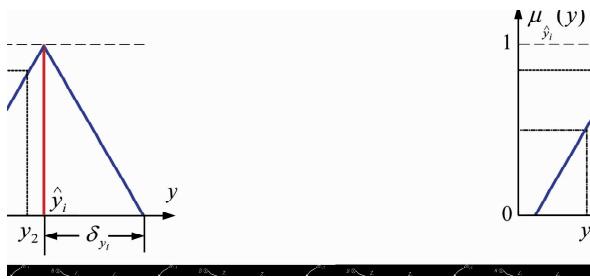


图 2 观测值的三角模糊数

Fig. 2 Triangular Fuzzy Number Observed Value

$$\left\{ \begin{array}{l} - \sum_{i=1}^n [\mu_{y_i}(y_i) \ln \mu_{y_i}(y_i) + (1 - \mu_{y_i}(y_i)) \cdot \\ \ln(1 - \mu_{y_i}(y_i))] = \min \end{array} \right. \quad (9)$$

对于 \hat{y}_i 、 $\mu_{y_i}(y_i)$ 及 δ_{y_i} 的计算, 根据式(7), 由对称模糊数的运算性质可得:

$$\hat{y}_i = a_{0i} + \sum_{j=1}^t a_{ij} \hat{x}_j, \delta_{y_i} = \sum_{j=1}^t c_j |a_{ij}|, \quad i = 1, 2, \dots, n \quad (10)$$

根据式(6)与式(10), 每一个观测值 y_i 对于其对称中心(真值) \hat{y}_i 的隶属度为:

$$\mu_{y_i}(y_i) = L\left(\frac{y_i - \hat{y}_i}{\delta_{y_i}}\right) = L\left(\frac{y_i - a_{0i} - \sum_{j=1}^t a_{ij} \hat{x}_j}{\sum_{j=1}^t c_j |a_{ij}|}\right) \quad (11)$$

2.3 模型的解算

根据式(9)的模糊熵准则, 对以模糊数表示的 G-M 模型式(7)进行求解时, 还应满足两个条件: ① $\delta_{y_i} > 0, i = 1, 2, \dots, n$, 即测量不确定度恒为正值; ② $|v_i| \leq \delta_{y_i}$, 即观测值的改正数 v 的绝对值应小于等于该观测值的不确定度 δ_{y_i} 。在实际计算中, 结合式(10), 观测数改正数的计算采用:

$$v_i = \hat{y}_i - y_i = a_i + \sum_{j=1}^t a_{ij} \hat{x}_j - y_i \quad (12)$$

综上所述, 测量数据不确定性评估的最小模糊熵估计(least fuzzy entropy estimation, LFEE)如式(13)所示:

$$\left\{ \begin{array}{l} - \sum_{i=1}^n [\mu_{y_i}(y_i) \ln \mu_{y_i}(y_i) + (1 - \mu_{y_i}(y_i)) \cdot \\ \ln(1 - \mu_{y_i}(y_i))] = \min \\ \text{st } \delta_{y_i} > 0 \text{ and } |v_i| \leq \delta_{y_i} \end{array} \right. \quad (13)$$

在式(13)中, 将未知参数 \mathbf{X} 和观测值的测量

不确定度 δ_{y_i} 均视为参数一并求解, 该模型的参数估计可以转换为非线性规划问题进行求解。

从式(13)可以看出, 最小模糊熵估计随观测值隶属函数 $\mu_{y_i}(y_i)$ 的不同而不同。因此, 最小模糊熵估计不是一个估计, 而是一类估计。对于不同的参照函数, 根据式(11), 可以计算出观测值的隶属函数不同, 观测值的三角模糊隶属数为:

$$\mu_{y_i}(y_i)_{\text{trg}} = 1 - \frac{|v_i|}{\delta_{y_i}} \quad (14)$$

观测值的抛物线模糊隶属数为:

$$\mu_{y_i}(y_i)_{\text{para}} = 1 - \frac{v_i^2}{\delta_{y_i}^2} \quad (15)$$

观测值的高斯模糊隶属数为:

$$\mu_{y_i}(y_i)_{\text{gauss}} = \exp\left[-\left(\frac{v_i}{\delta_{y_i}}\right)^2\right] \quad (16)$$

如果将观测值的不同隶属函数(式(14)、式(15)、式(16))代入最小模糊熵估计(式(13)), 得到基于三角模糊数的最小模糊熵估计:

$$\left\{ \begin{array}{l} - \sum_{i=1}^n \left[\frac{|v_i|}{\delta_{y_i}} \ln \frac{|v_i|}{\delta_{y_i}} + (1 - \frac{|v_i|}{\delta_{y_i}}) \cdot \\ \ln(1 - \frac{|v_i|}{\delta_{y_i}}) \right] = \min \\ \text{st } \delta_{y_i} > 0 \text{ and } |v_i| \leq \delta_{y_i} \end{array} \right. \quad (17)$$

基于抛物线模糊数的最小模糊熵估计:

$$\left\{ \begin{array}{l} - \sum_{i=1}^n \left[\frac{v_i^2}{\delta_{y_i}^2} \ln \frac{v_i^2}{\delta_{y_i}^2} + (1 - \frac{v_i^2}{\delta_{y_i}^2}) \ln(1 - \frac{v_i^2}{\delta_{y_i}^2}) \right] = \min \\ \text{st } \delta_{y_i} > 0 \text{ and } |v_i| \leq \delta_{y_i} \end{array} \right. \quad (18)$$

基于高斯模糊数的最小模糊熵估计:

$$\left\{ \begin{array}{l} - \sum_{i=1}^n \left[\exp\left[-\left(\frac{v_i}{\delta_{y_i}}\right)^2\right] \cdot \ln\left(\exp\left[-\left(\frac{v_i}{\delta_{y_i}}\right)^2\right]\right) + \right. \\ \left. (1 - \exp\left[-\left(\frac{v_i}{\delta_{y_i}}\right)^2\right]) \cdot \ln(1 - \exp\left[-\left(\frac{v_i}{\delta_{y_i}}\right)^2\right]) \right] = \min \\ \text{st } \delta_{y_i} > 0 \text{ and } |v_i| \leq \delta_{y_i} \end{array} \right. \quad (19)$$

3 应用实例计算

为验证本文提出的最小模糊熵估计应用于测量数据不确定性评估的可行性, 以文献[21]中的某高程监测网数据为例, A、B、C、D 为基准点(图 3), 其高程分别为 41.431 m、38.744 m、58.732 m、62.160 m, BM1、BM2、BM3 为工作基

点,各测段水准路线的长度及某期的高差观测值均列于表1。

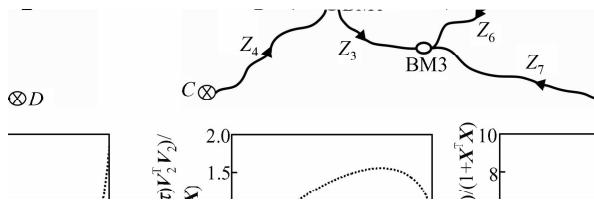


图3 高程监测网

Fig. 3 Elevation Monitoring Network

表1 高差观测值

Tab. 1 Elevation Difference Observations

水准路线	长度 L/km	高差观测值 h/m
Z_1	2.04	+13.019
Z_2	1.88	+3.092
Z_3	2.30	+2.175
Z_4	1.95	-6.949
Z_5	1.61	+7.251
Z_6	2.32	+5.255
Z_7	1.71	-8.238

设待定点 $H_{\text{BM}1}, H_{\text{BM}2}, H_{\text{BM}3}$ 的高程为未知参数 $\hat{x}_1, \hat{x}_2, \hat{x}_3$, ,用 \hat{y}_i 表示高差观测值 h_i 的真值($i=1, 2, \dots, 7$),将未知参数和观测值均看成对称的模糊数,根据水准测量原理建立如 G-M 模型:

$$\left\{ \begin{array}{l} (\hat{y}_1, \delta_{y_1}) = (\hat{x}_1, \delta_{x_1}) - H_B \\ (\hat{y}_2, \delta_{y_2}) = (\hat{x}_1, \delta_{x_1}) - (\hat{x}_2, \delta_{x_2}) \\ (\hat{y}_3, \delta_{y_3}) = (\hat{x}_3, \delta_{x_3}) - (\hat{x}_1, \delta_{x_1}) \\ (\hat{y}_4, \delta_{y_4}) = (\hat{x}_1, \delta_{x_1}) - H_C \\ (\hat{y}_5, \delta_{y_5}) = (\hat{x}_2, \delta_{x_2}) - H_A \\ (\hat{y}_6, \delta_{y_6}) = (\hat{x}_3, \delta_{x_3}) - (\hat{x}_2, \delta_{x_2}) \\ (\hat{y}_7, \delta_{y_7}) = (\hat{x}_3, \delta_{x_3}) - H_D \end{array} \right. \quad (20)$$

根据式(10),有式(21)、式(22)成立,即

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \\ \hat{y}_5 \\ \hat{y}_6 \\ \hat{y}_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} +$$

$$\begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \times \begin{bmatrix} H_A \\ H_B \\ H_C \\ H_D \end{bmatrix} \quad (21)$$

$$\begin{bmatrix} \delta_{y_1} \\ \delta_{y_2} \\ \delta_{y_3} \\ \delta_{y_4} \\ \delta_{y_5} \\ \delta_{y_6} \\ \delta_{y_7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \delta_{x_1} \\ \delta_{x_2} \\ \delta_{x_3} \end{bmatrix} \quad (22)$$

对式(20)建立的 G-M 模型,待定点的高程 \hat{x}_i 和高差观测值的不确定度 δ_{y_i} 均视作未知参数,总共 10 个参数,分别按基于不同模糊数的最小模糊熵估计公式(式(17)~(19)),利用遗传算法进行优化计算,计算结果分别见表 2、表 3。

通过比较,从表 2 可以看出,参数的最小模糊熵估计结果与最小二乘估计结果基本一致,特别是基于高斯模糊数的最小模糊熵估计与最小二乘估计结果最为接近;从表 3 计算的观测值不确定性度量方面来看,最小二乘估计采用 3σ 计算的结果均高于 30 mm,而最小模糊熵估计计算的结果有大有小,更符合观测值的随机波动规律。

表2 不同参数估计方法的结果比较

Tab. 2 Comparison Results of Different Estimation Methods

参数	最小二乘估			最小模糊熵估计(LFEE)/m			LFEE 与 LSE 参数估计之差/mm
	计(LSE)/m	三角	抛物线	高斯	三角	抛物线	
X_1	51.769 1	51.773 4	51.767 9	51.765 3	4.3	-1.2	-3.8
X_2	48.679 1	48.682 1	48.680 3	48.679 1	3.0	1.2	0
X_3	53.932 2	53.926 3	53.935 7	53.932 2	-5.9	3.5	0

口,以测量不确定度理论和模糊数学为基础,提出了最小模糊熵估计理论。该估计是以“测量不确定度”为指标对测量数据不确定性进行直接评估的新的估计类,未知参数和观测值的测量不确定

4 结语

本文以如何评价测量数据的不确定性为突破

表 3 两种估计方法计算的观测量平差值结果比较

Tab. 3 Comparison of Different Methods of Estimation Results

观 测 量	观测值的平差值/m				观测值的不确定性/mm		
	最小二乘估计	最小模糊熵估计			中误差 (3σ)	测量不确定度/mm	
		三角	抛物线	高斯		三角	抛物线
h_1	13.025 1	13.029 4	13.023 9	13.021 3	33.7	48.8	42.0
h_2	3.090 0	3.091 3	3.087 6	3.084 1	32.4	23.3	39.1
h_3	2.163 1	2.152 9	2.167 8	2.166 9	36.0	22.1	38.8
h_3	-6.963 0	-6.958 6	-6.964 1	-6.966 7	33.1	47.4	43.8
h_5	7.248 1	7.251 1	7.249 3	7.250 2	30.1	32.2	38.2
h_6	5.253 1	5.244 2	5.255 4	5.251 0	36.0	46.2	39.1
h_7	-8.227 8	-8.233 7	-8.224 3	-8.227 8	31.0	37.9	43.3
							30.7

度作为参数同时进行估计, 较为完整地实现了观测数据不确定性建模及数据质量评定的问题。本文研究工作仅仅是初步, 关于模糊数的选择、不同精度的观测数据定权以及参数的不确定性等问题还需要进一步的深入研究。

参 考 文 献

- [1] Kutterer H. Uncertainty Assessment in Geodetic Data Analysis [C]. IAG First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS, Zurich, Switzerland, 2001
- [2] Sun Haiyan. Entropy and the Uncertainty Interval [J]. *Journal of Wuhan Technical University of Surveying and Mapping*, 1994, 19(1):63-71(孙海燕. 熵与不确定度区间[J]. 武汉测绘科技大学学报, 1994, 19(1):63-71)
- [3] Song Yingchun, Jin Hao, Cui Xianqiang. Adjustment Algorithm About Observation Data with Uncertain[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(7):788-792(宋迎春, 金昊, 崔先强. 带有不确定性的观测数据平差解算方法[J]. 武汉大学学报·信息科学版, 2014, 39(7):788-792)
- [4] Ye Depei. Understanding Evaluation and Application of Measurement Uncertainty [M]. Beijing: China Metrology Publishing House, 2013(叶德培. 测量不确定度理解评定与应用[M]. 北京: 中国计量出版社, 2013)
- [5] Chen Wei, Wang Xinzhou. Least Uncertainty Estimation Theory and Its Application to Resolving Morbid Problems[J]. *Geomatics and Information Science of Wuhan University*, 2008, 33(7):752-754(陈伟, 王新洲. 最小不确定度估计原理及其病态问题解法研究[J]. 武汉大学学报·信息科学版, 2008, 33(7):752-754)
- [6] Fu Jiangque, Gao Jingxiang, Cheng Zhengfeng, et al. Weighted Solution Theory for Maximum Possibility Estimation and Its Data Quality Evaluation [J]. *Journal of Geomatics Science and Technology*,
- 2010, 27(4):239-242(付江缺, 高井祥, 程正逢, 等. 极大可能性估计带权解算理论及数据质量评定[J]. 测绘科学技术学报, 2010, 27(4):239-242)
- [7] Yang Yuanxi. Some Notes on Uncertainty. [J]. *Acta Geodaetica et Cartographica Sinica*, 2012, 41(5): 646-650(杨元喜. 卫星导航的不确定性、不确定度与精度若干注记[J]. 测绘学报, 2012, 41(5): 646-650)
- [8] Tao Benzao. Theory of Uncertainty of Quality Control in GIS [J]. *Journal of Institute of Surveying and Mapping*, 2000, 17(4):235-238(陶本藻. GIS 质量控制中不确定度理论[J]. 测绘学院学报, 2000, 17(4):235-238)
- [9] Zou Yonggang, Zhai Jingsheng, Liu Yanchun, et al. Seabed DEM Construction Based on Uncertainty [J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(8):964-968 (邹永刚, 翟京生, 刘雁春, 等. 利用不确定度的海底数字高程模型构建[J]. 武汉大学学报·信息科学版, 2011, 36(8): 964-968)
- [10] Tao Benzao, Lan Yueming. The Statistical Estimation Method of GIS Overlay Uncertainty[J]. *Geomatics and Information Science of Wuhan University*, 2001, 26(2):101-104(陶本藻, 蓝悦明. GIS 叠置位置不确定度的统计估计方法[J]. 武汉大学学报·信息科学版, 2001, 26(2):101-104)
- [11] Zhang Zhenglu, Fan Guoqing, Zhang Songlin, et al. General Reliability of Measurement [J]. *Geomatics and Information Science of Wuhan University*, 2012, 37(5):577-581(张正禄, 范国庆, 张松林, 等. 测量的广义可靠性研究[J]. 武汉大学学报·信息科学版, 2012, 37(5):577-581)
- [12] Li Dajun, Gong Jianya, Xie Gangsheng, et al. Error Entropy Band for Linear Segment in GIS[J]. *Geomatics and Information Science of Wuhan University*, 2002, 27(5):462-466(李大军, 龚健雅, 谢刚生, 等. GIS 中线元的误差熵带研究[J]. 武汉大学学报·信息科学版, 2002, 27(5):462-466)
- [13] Shi Yufeng, Shi Wenzhong, Jin Fengxiang. Hybrid

- Entropy Model of Spatial Data Uncertainty in GIS [J]. *Geomatics and Information Science of Wuhan University*, 2006, 31(1): 82-85(史玉峰, 史文中, 靳奉祥. GIS 中空间数据不确定性的混合熵模型研究[J]. 武汉大学学报·信息科学版, 2006, 31(1): 82-85)
- [14] Zhu Changqing, Zhang Guoqin, Wang Guangxia. An Error Entropy Model for GIS Spatial Linear [J]. *Geomatics and Information Science of Wuhan University*, 2005, 30(5): 405-407(朱长青, 张国芹, 王光霞. GIS 中三维空间直线的误差熵模型[J]. 武汉大学学报·信息科学版, 2005, 30(5): 405-407)
- [15] Shi Yufeng, Jin Fengxiang, Wang Jian. Information Entropy Based Gross Error Discrimination Method for Surveying Data [J]. *Bulletin of Surveying and Mapping*, 2002, 2: 9-13(史玉峰, 靳奉祥, 王健. 基于信息熵的测量数据粗差识别法[J]. 测绘通报, 2002, 2: 9-13)
- [16] You Yangsheng, Ma Li, Liu Xing. On Estimation of Error Entropy [J]. *Geomatics and Information Science of Wuhan University*, 2008, 33(7): 748-751(游扬声, 马力, 刘星. 误差熵的估计问题研究[J]. 武汉大学学报·信息科学版, 2008, 33(7): 748-751)
- [17] Alkhatib H, Neumann I, Kutterer H. Uncertainty Modeling of Random and Systematic Errors of Monte Carlo and Fuzzy Techniques [J]. *Journal of Applied Geodesy*, 2009, 3(2): 67-80
- [18] Chen Wei, Zhang Jian. Least Uncertainty Estimation Theory and Its Applications in Survey Data Processing [J]. *Bulletin of Surveying and Mapping*, 2013, 1: 16-18(陈伟, 张践. 最小不确定度估计及其在测量数据处理中的应用[J]. 测绘通报, 2013, 1: 16-18)
- [19] Guo Jinyun, Li Chengyao, Cui Xianguo, et al. The Gauss-Markov Model Solved by Fuzzy Mathematics Theory [J]. *Site Investigation and Surveying Practice*, 1996, 2: 21-26(郭金运, 李成尧, 崔先国, 等. 高斯-马尔柯夫模型的模糊解算[J]. 军工勘察, 1996, 2: 21-26)
- [20] Deluca A, Termini S. A Definition of Non-probabilist Entropy in Setting of Fuzzy Sets Theory [J]. *Information Control*, 1972, 10(20): 301-312
- [21] Bai Dimou. Deformation Observation and Analysis of Engineering Building [M]. Chengdu: Southwest Jiaotong University Press, 2002(白迪谋. 工程建筑变形观测与变形分析[M]. 成都: 西南交通大学出版社, 2002)

Using Minimum Fuzzy Entropy Algorithm to Measure Uncertainty of Geodetic Data

WEI Guanjun¹ DANG Yamin² ZHANG Chuanyin²

¹ Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730070, China

² Chinese Academy of Surveying and Mapping, Beijing 100830, China

Abstract: Quality and reliability of geodetic data depend on the size of its uncertainty. In this paper, uncertainty in measurement data is effectively evaluated based on the theory of measurement uncertainty and fuzzy mathematics, A Function model using measurement uncertainty as unknown parameter is established to directly evaluate the uncertainty of survey data. The Fuzzy Entropy Measure is proposed as the optimal criterion to solve the function model. A corresponding algorithm is established; Experimental results and comparisons with the least squares estimation show that the proposed method using a elevation monitoring network data solver is feasible.

Key words: data uncertainty; measurement uncertainty; fuzzy entropy criterion; function model; least fuzzy entropy estimation

First author: WEI Guanjun, PhD, associate professor, specializes in geodetic data processing. E-mail: whampion@sina.com

Corresponding author: DANG Yamin, PhD, professor. E-mail: dangym@casm.ac.cn

Foundation support: The National Natural Science Foundation of China, Nos. 41201004, 41364001; Natural Science Foundation of Gansu Province, No. 1508RJZA065; Science and Technology Support Program of Lanzhou Jiaotong University, No. ZC2014002; Higher Education Research Project of Gansu Province, No. 2014B-040.