

文章编号: 1001-1595(2010)02-0141-05

P-范分布混合整数模型极大似然估计

李博峰¹, 沈云中^{1, 2}

1. 同济大学 测量与国土信息工程系, 上海 200092; 2. 现代工程测量国家测绘局重点实验室, 上海 200092

Maximum Likelihood Estimation in Mixed Integer Linear Model with P-norm Distribution

Li Bofeng¹, SHEN Yunzhong^{1, 2}

1. Department of Surveying and Geo-informatics Engineering, Tongji University, Shanghai 200092, China; 2. Key Laboratory of Advanced Surveying Engineering of SBSM, Shanghai 200092, China

Abstract: The integer parameters must be primarily estimated in the GNSS and InSAR application, which essentially introduces a special mixed integer model with both real- and integer-valued parameters from mathematical point of view. Up to now, all methods for mixed integer model are based on the least squares criterion. In this paper, the parameter estimation will be investigated in mixed integer model with the P-norm distributed observation noises. First of all, we will employ the maximum likelihood estimation theory to derive the criterion for integer searching, considering the fact that only real parameters can be differentiated but not the integer parameters due to their discrete property, and further verify that least squares based integer searching criterion is just a case with normally distributed noises. Secondly, the approach and iterative procedure are given for estimating p , searching integers and solving real-valued parameters. Finally, the simulated experiments are implemented to verify the correctness of the derived formulae and proposed algorithm.

Key words: GNSS; mixed integer model; P-norm distribution; maximum likelihood estimation; ambiguity resolution

摘 要: 研究 P-范分布混合整数模型参数估计理论, 首先采用极大似然理论, 顾及实参数可导而整参数不可导, 导出 P-范分布整数搜索准则, 并证明最小二乘整数搜索准则是它的一个特例; 然后给出 P-范分布混合整数模型的 p 参数估计、整参数搜索和实参数求解的方法和计算流程; 最后通过模拟算例分析和验证本文模型和计算方法的正确性。

关键词: 全球导航卫星系统; 混合整数模型; P-范分布; 极大似然估计; 模糊度解算

中图分类号: P229 文献标识码: A

基金项目: 国家自然科学基金(40674003, 40874016); 现代工程测量国家测绘局重点实验室开放课题(TJES0809)

1 引言

从早期的相位电磁波测距到目前的高精度 GNSS 定位以及 InSAR 解缠, 都需要求解整数参数, 它们的观测模型都含有实参数和整参数。徐培亮引入“混合整数模型”的概念来描述这种特殊的数学模型, 并指出所有的纯实数和纯整数模型都只是它的退化形式, 但其对混合整数模型的研究都是基于观测噪声服从正态分布的^[1]。

只有当观测误差服从正态分布时, 采用最小二乘准则才能获得参数的最优无偏估值。但在测量数据的采集、录入及处理过程中, 不可避免地受到外界条件、人为因素等的影响, 导致观测误差不是正态分布, 而是根据观测值合理选取 p 值的 P-范分布^[2], 此时采用最小二乘估计不能得到参数的最优估值。P-范分布的常用退化分布有: 拉普拉斯分布、正态分布、均匀分布等^[3]。对于实参数

观测模型的 P-范分布估计理论已有深入研究。孙海燕等对 P-范分布理论进行了较系统的研究^[3-8], 并提出采用拉普拉斯分布与正态分布或正态分布与均匀分布的线性组合去逼近 P-范分布, 回避了 P-范分布概率密度函数的复杂性^[4]。於宗涛等从附有约束条件的条件平差模型(又称概括模型)出发, 导出了多元和多母体 P-范极大似然平差公式^[5-6]。文献[7]构造了 P-范分布方差估值的 χ^2 统计量, 并研究了它的假设检验方法。此外, 文献[2]和[8]分别讨论了一元有界 P-范分布的参数估计和一元 P-范分布的半参数模型, 试验表明有界 P-范估计能够更合理地确定 P-范分布的各项参数。然而, 目前对混合整数模型的参数估计只是直接采用最小二乘准则, 当观测噪声服从 P-范分布时, 混合整数模型的估计理论研究尚处于空白。

研究 P-范分布混合整数模型估计理论, 首先

采用极大似然估计理论, 顾及实参数可导而整参数不可导, 导出了 P-范分布整数搜索准则, 并证明最小二乘整数搜索准则是它的一个特例; 然后给出了 P-范分布混合整数模型的 p 参数估计、整参数搜索和实参数求解的方法和计算流程。理论上, 合理的 p 值体现观测误差最符合的分布, 即使观测噪声服从或者近似地服从正态分布, 基于最小二乘的混合整数模型也仅为基于 P-范分布极大似然估计的特例。

2 P-范分布混合整数模型的等价整参数估计

2.1 P-范分布整数搜索准则

混合整数线性(或线性化)观测模型^[1]

$$\varepsilon = Ax + Bz - l \quad (1)$$

其中, l 与 ε 分别为 n 维观测值向量与观测噪声向量, 其数学期望和方差分别为

$$\left. \begin{aligned} E(l) = 0, \quad E(l) = Ax + Bz = \mu_l \\ D(l) = D(\varepsilon) = D_l = \alpha_0^2 Q_\varepsilon \end{aligned} \right\} \quad (2)$$

其中, α₀ 和 Q_ε 分别为先验方差因子和先验协因数阵; x 为 m 维实参数向量, A 为对应设计矩阵; 混合整数模型式(1)与实参数模型的区别在于包含了 t 维整参数向量 z, B 为它的设计矩阵。不失一般性, 假设观测噪声 ε 服从多元 P-范分布, 则它的概率密度函数为^[10]

$$f(\varepsilon, p, \alpha_0) = \frac{(p\lambda)^n}{2^n \Gamma^n \left[\frac{1}{p} \right] \alpha_0^n \prod_{i=1}^n |Q_\varepsilon|^{1/2}} \exp \left\{ - \left[\frac{\lambda}{\alpha_0} \right]^p \|\mathbf{Q}_\varepsilon^{-1/2} \varepsilon\|_p^p \right\} \quad (3)$$

其中, λ = √[Γ(3/p)/Γ(1/p)], (x) 为伽玛函数;

$\|x\|_p = \sqrt[p]{\sum_{i=1}^n |x_i|^p}$ 表示向量 x 的 p 范值, |·| 为取绝对值; p 是 P-范分布参数, 通常 Q_ε 为对称正定矩阵, 可分解 Q_ε = Q_ε^{1/2}(Q_ε^{1/2})^T, 从而 Q_ε⁻¹ = Q_ε^{-1/2}(Q_ε^{-1/2})^T。给式(1) 两边同时左乘矩阵 Q_ε^{-1/2}, 得线性变换后的混合整数观测模型

$$l = Ax + Bz + \varepsilon \quad (4)$$

其中, ε = Q_ε^{-1/2}ε, A = Q_ε^{-1/2}A, B = Q_ε^{-1/2}B, l = Q_ε^{-1/2}l。变换后的误差向量 ε 的协因数阵 Q_ε 为单位阵。P-范分布概率密度函数式(3)变为^[10]

$$f(\varepsilon, p, \alpha_0) = \frac{(p\lambda)^n}{2^n \Gamma^n (1/p) \alpha_0^n} \exp \left\{ - \left[\frac{\lambda}{\alpha_0} \right]^p \|\varepsilon\|_p^p \right\} \quad (5)$$

两边同取自然对数

$$\ln f(\varepsilon, p, \alpha_0) = n \left[\ln p + \ln \lambda - \ln 2 - \ln \Gamma \left[\frac{1}{p} \right] - \ln \alpha_0 \right] - \left[\frac{\lambda}{\alpha_0} \right]^p \|\varepsilon\|_p^p \quad (6)$$

根据极大似然估计理论, 在极大化式(6)的情况下求解参数 x、z、p 和 α₀, 等价于最小化目标函数

$$g(\varepsilon, p, \alpha_0) = \left[\frac{\lambda}{\alpha_0} \right]^p \sum_{i=1}^n |\varepsilon_i|^p - n \left[\ln p + \ln \lambda - \ln \Gamma \left[\frac{1}{p} \right] - \ln \alpha_0 \right] = \min \quad (7)$$

如果将 ε 也当作未知参数, 则需要引入约束方程式(4), 构造拉格朗日极值函数

$$\Phi = g(\varepsilon, p, \alpha_0) + k^T (Ax + Bz - \varepsilon - l) = \min \quad (8)$$

其中, k = (k₁ ... k_n)^T 为未知联系数向量。先对 ε 求导

$$\frac{\partial \Phi}{\partial \varepsilon} = \left(\frac{\partial \Phi}{\partial \varepsilon_1} \quad \dots \quad \frac{\partial \Phi}{\partial \varepsilon_n} \right) = \left[\frac{\lambda}{\alpha_0} \right]^p \begin{pmatrix} |v_1|^{p-2} v_1 \\ \vdots \\ |v_n|^{p-2} v_n \end{pmatrix}^T - k^T = 0 \quad (9)$$

其中

$$v = Ax + Bz - l \quad (10)$$

是 ε 的估值。令 W = diag (|v₁|^{p-2} ... |v_n|^{p-2}), v = (v₁ ... v_n)^T, 同时考虑 (Nα₀)^p ≠ 0, 则

$$k = Wv \quad (11)$$

在纯实数模型中, 只需要对所有的参数求导, 但在混合整数模型中, 由于整参数的非连续性导致不能对其求导^[9], 所以求实参数 x 的导数

$$\frac{\partial \Phi}{\partial x} = \left(\frac{\partial \Phi}{\partial x_1} \quad \dots \quad \frac{\partial \Phi}{\partial x_n} \right) = k^T A = 0 \quad (12)$$

将式(10)、式(11)代入式(12), 得

$$A^T W A \hat{x} + A^T W B \hat{z} = A^T W l \quad (13)$$

得到实参数 x 的估值

$$\hat{x} = (A^T W A)^{-1} (A^T W l - A^T W B \hat{z}) \quad (14)$$

为了求解参数 p 和 α₀, 还需要计算目标函数式(8)分别对 p 和 α₀ 的导数, 为此引入普西函数 φ(x) = Γ'(x)/Γ(x), 经过类似文献[10]的推导得

$$p + \phi \left[\frac{1}{p} \right] - \ln \left[\frac{n}{p} \right] + \ln \left[\sum_{i=1}^n |v_i|^p \right] - \frac{p \sum_{i=1}^n (|v_i|^p \ln |v_i|)}{\sum_{i=1}^n |v_i|^p} = 0 \quad (15)$$

$$\alpha_0 = \frac{p}{n} \left[\frac{\Gamma(3/p)}{\Gamma(1/p)} \right]^{\frac{1}{p}} \sum_{i=1}^n |v_i|^p \quad (16)$$

如果参数 p 和 σ_0 事先已知或者在迭代求解过程中固定, 则式(6)等价于

$$h(z) = \|\varepsilon\|_p = \|Ax + Bz - I\|_p = \min \quad (17)$$

将式(13)代入式(17),

$$h(z) = \|A(A^T W A)^{-1} A^T W(I - Bz) + Bz - I\|_p = \|RBz - RI\|_p = \min \quad (18)$$

其中, $R = I - A(A^T W A)^{-1} A^T W = I - J$ 是幂等阵, 即 $RR = R$, 且 $R^T W = (I - J)^T W$ 是对称阵, I 表示单位阵。将幂等阵 R 左乘于观测方程式(4), 得整参数观测方程

$$R\varepsilon = RI - RBz \quad (19)$$

由于整参数为离散型参数, 无法直接求解, 只能通过搜索得到, 因此式(19)、(18)为 P-范分布混合整数模型的等价观测方程和整参数解算准则。确定整参数 z 后, 回代到式(14)求解实参数。

2.2 正态分布整参数搜索准则

当观测误差服从正态分布时, 即 $p = 2$, $W = I$, $R = I - A(A^T A)^{-1} A^T$, 则整参数搜索准则式(18)可简化为

$$h(z) = \|RBz - RI\|_2^2 = (RBz - RI)^T (RBz - RI) = I^T R^T RI - I^T R^T RBz - z^T B^T R^T RI + z^T B^T R^T RBz \quad (20)$$

考虑到 R 为对称幂等阵, 满足 $RR = R^T R = R$ 和 $R^T = R$, 则

$$h(z) = I^T RI - I^T RBz - z^T B^T RI + z^T B^T RBz = z^T Q_f^{-1} z - \hat{z}_f^T Q_f^{-1} z - z^T Q_f^{-1} \hat{z}_f + \hat{z}_f^T Q_f^{-1} \hat{z}_f - \hat{z}_f^T Q_f^{-1} \hat{z}_f + I^T RI = (z - \hat{z}_f)^T Q_f^{-1} (z - \hat{z}_f) + I^T RI - \hat{z}_f^T Q_f^{-1} \hat{z}_f = \min \quad (21)$$

其中, $\hat{z}_f = (B^T RB)^{-1} B^T RI = Q_f^{-1} B^T RI$ 和 $Q_f^{-1} = B^T RB$ 分别为整参数浮点解及其协方差矩阵。由于此时 R 为常数矩阵, 因此准则式(21)严格等价于

$$(z - \hat{z}_f)^T Q_f^{-1} (z - \hat{z}_f) = \min \quad (22)$$

这也就是解算 GPS 模糊度通常采用的最小二乘整参数搜索准则^[11], 同时也说明当观测误差服从正态分布时, 采用极大似然估计与最小二乘估计是等价的。

3 P-范分布混合整数模型参数估计算法

P-范混合整数模型参数估计算法分三个部分(如图1的虚线框所示): 部分 iv 为观测方程预处理, 即观测方程权矩阵的单位化; 部分 ② 求解整参数, 先求浮点解, 然后搜索固定; 部分 ③ 求解实参数的固定解, 即利用固定的整参数将混合整数模

型转化为纯实参数模型, 按照常规的 P-范实参数模型求解。下面详细讨论算法步骤:

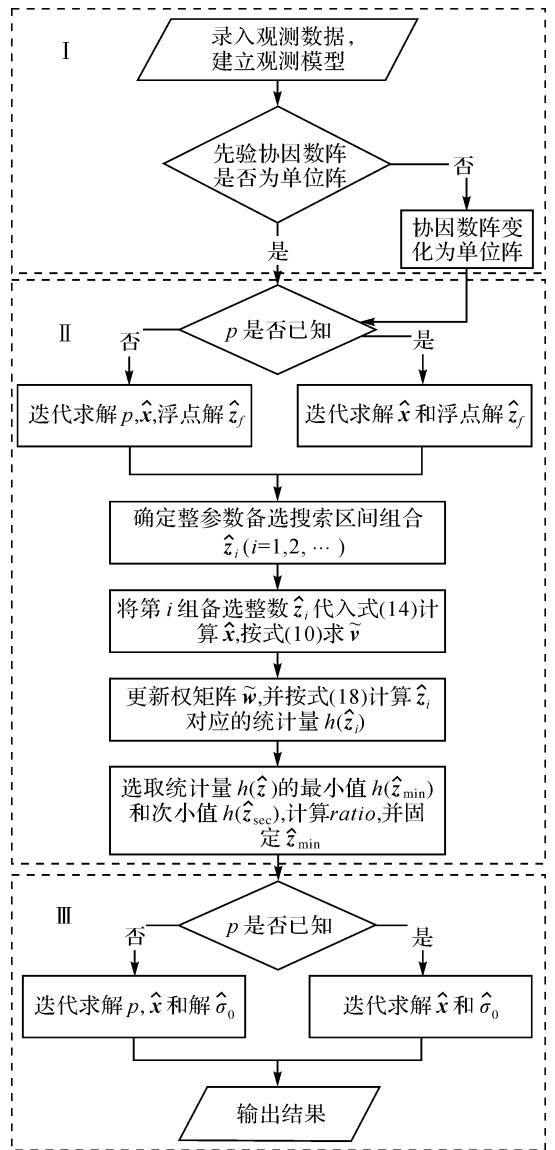


图 1 P-范混合整数模型参数估计算法流程

Fig. 1 Flowchart for parameter estimation in mixed integer model with P-norm distribution

1. 若观测方程的协因数阵不是单位阵, 对观测方程按照式(4)变换, 使变换后的协因数阵为单位阵;

2. 对整参数作实数近似, 对其也求导, 经过类似 2.1 的推导, 得到法方程

$$\begin{bmatrix} A^T W A & A^T W B \\ B^T W A & B^T W B \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} A^T W I \\ B^T W I \end{bmatrix} \quad (23)$$

采用式(10)、(23)、(15)和(16)迭代计算参数浮点解 \hat{x}_f 、 \hat{z}_f 、 p 和 σ_{0f} ;

3. 根据整参数浮点解 \hat{z}_f 与其精度 $D_z = \sigma_0^2(B^T WB)^{-1}$, 确定整参数搜索备选组合 $\hat{z}_i \in [\hat{z}_f - k \text{diag}(D_z^{1/2})\hat{z}_f + k \text{diag}(D_z^{1/2})]$, $i = 1, 2, \dots$, 其中 $\text{diag}(\cdot)$ 为取矩阵对角元素的函数, 常数 k 一般取 3;

4. 将第 i 个备选整数组合 \hat{z}_i 代入式(14)求解实参数估值 \hat{x}_i , 然后根据式(10)求解残差 v_i ;

5. 根据残差 v 计算 W , 然后计算备选组合 \hat{z}_i 对应的 P-范统计量 $h(\hat{z}_i) = \|RB\hat{z}_i - RI\|_p$;

6. 选取对应统计量 $h(\hat{z})$ 最小和次小的备选组合 \hat{z}_{\min} 和 \hat{z}_{sec} , 计算指标 $ratio = \frac{h(\hat{z}_{\text{sec}})}{h(\hat{z}_{\min})}$, 并固定 \hat{z}_{\min} ;

7. 整参数固定后, 混合整数模型转化为纯实数模型, 再采用式(10)、(14)、(15)和(16)迭代计算参数 \hat{x} 、 p 和 σ_0 的固定解。

4 算例与分析

算例旨在验证导出的 P-范混合整数模型参数估计模型和算法的正确性。采用模型 $l = Ax + z + \varepsilon$ 模拟观测数据。其中, 取实参数 $x = 0.5$, 整参数 $z = 5$; $(n \times 1)$ 设计矩阵 A 的元素 $a_i = \sin(2\pi \times (i-1)/99)$, $i = 1, \dots, n$; 取先验精度 $\sigma_0 = 0.1$, 模拟均值为零的正态分布观测噪声 ε 。模拟意味着当观测数据个数 n 趋于无穷时, p 和 σ_0 的估值趋于真值 2 和 0.1, 而在有限采样下, 估值只能反映实际数据本身的最佳估值, 即反映观测数据最符

合的分布。

取 $n = 200, 300, \dots, 600$ 做 5 次试验, 每次试验先模拟观测噪声 ε 然后生成观测向量 l , 并计算观测噪声的先验精度。表 1 给出了 5 次试验的计算过程与结果, 第 1 列为每次试验采用的观测值数目, 第 2 列为每次试验数据的先验精度; 第 3 到第 5 列给出浮点解(即算法部分 ③), 第 6 列为浮点解的迭代次数, 第 7 列为固定整参数的 $ratio$ 值; 整参数固定后, 混合整数模型转化为纯实数模型, 第 8 到第 10 列给出纯实数模型的参数估值(即算法部分 ④的固定解), 第 11 列为固定解的迭代次数。其中验后精度指观测值一次采样的验后精度。应说明的是, 所有试验整参数都正确固定。

从观测噪声的先验精度得出, 随着观测值数目的增多, 观测数据所体现的先验精度更加接近于实际模拟所采用的精度, 说明模拟试验一定要有足够的采样。总体而言, 无论是浮点解还是固定解, 都可在 4 次迭代收敛, 算法较稳定; 随着观测值数量的增大, p 估值更加接近于 2(即模拟采用的 p 值), 固定整参数的 $ratio$ 值也随之增大, 验后精度也更加接近于先验精度。值得注意的是, 尽管模拟采用正态分布, 但实际估值并非严格的正态分布, 尤其当 $n = 200$ 时, 这些观测数据所体现的是 $p \approx 1.54$ 的 P-范分布, 这也说明在实际应用中采用 P-范分布极大似然估计是必要且合理的。

表 1 P-范混合整数模型解算过程

Tab. 1 The process of parameter estimation in the mixed integer model with P-norm distribution

观测值 数目	先验 精度	浮点解					固定解			
		x	p	验后 精度	迭代 次数	$ratio$	x	p	验后 精度	迭代 次数
200	0.090 9	0.498 8	1.542 4	0.093 1	4	45.69	0.498 8	1.554 5	0.093 0	3
300	0.092 4	0.505 8	1.819 9	0.092 5	2	80.69	0.505 8	1.829 2	0.092 7	2
400	0.094 2	0.503 0	1.867 7	0.094 3	4	86.14	0.50295	1.860 4	0.094 4	4
500	0.094 7	0.506 9	1.944 2	0.094 5	3	99.36	0.50688	1.955 0	0.094 7	3
600	0.095 1	0.512 8	1.998 1	0.094 6	2	110.37	0.51273	2.007 8	0.094 8	2

5 结论

本文研究了 P-范分布混合整数模型极大似然估计理论, 导出了该模型等价的整参数搜索准则, 并给出了该模型的 p 参数估计、整参数搜索和实参数求解的计算方法, 通过模拟算例验证了

本文公式和方法正确, 得出以下结论: ① 观测数据的分布受外界因素和样本量的影响, 往往不能严格地服从它的真实分布, 采用 P-范估计能够确定观测数据最符合的分布; ④ 随着样本量的增加, 观测数据所体现的分布更加接近它的真实分布, 验后精度更加接近观测数据的实际精度;

(四)即使观测噪声服从或者近似地服从正态分布, 基于最小二乘的混合整数模型也仅为基于 P-范分布极大似然估计的特例; $1/4$ 与基于最小二乘的混合整数模型相同, 整参数必须通过搜索来确定, 因此, 研究提高 P-范分布混合整数模型整参数搜索效率的算法具有重要的意义。

参考文献:

- [1] XU Peiliang, CANNON E, LACHAPPELLE G. Mixed Integer Programming for the Resolution of GPS Carrier Phase Ambiguities [C] // IUGCG95 Assembly. Boulder: [s. n.], 1995.
- [2] PAN Xiong, FU Zongtang. Parameter Adaptive Estimation of Bounded P-norm Distribution[J]. Geomatics and Information Science of Wuhan University, 2007, 32(4): 323-325. (潘雄, 付宗堂. 一元有界 P-范分布的参数自适应估计[J]. 武汉大学学报: 信息科学版, 2007, 32(4): 323-325.)
- [3] SUN Haiyan, PAN Xiong. Parameter Estimation of the Monadic P-norm Distribution[J]. Geomatics and Information Science of Wuhan University, 2003, 28(5): 49-52. (孙海燕, 潘雄. 一元 P-范分布的参数估计[J]. 武汉大学学报: 信息科学版, 2003, 28(5): 49-52.)
- [4] SUN Haiyan. Approximate Representation of the P-norm Distribution[J]. Geomatics and Information Science of Wuhan University, 2001, 26(3): 222-225. (孙海燕. P-范分布的近似表示[J]. 武汉大学学报: 信息科学版, 2001, 26(3): 222-225.)
- [5] YU Zongchou, SUN Haiyan, CHEN Zhizhong. Maximum Likelihood Adjustment of Multiple-family P-norm Distribution Samples[J]. Acta Geodaetica et Cartographica Sinica, 1997, 26(1): 42-46. (於宗侑, 孙海燕, 陈之中. 多母体 P-范极大似然平差[J]. 测绘学报, 1997, 26(1): 42-46.)
- [6] YU Zongchou, SUN Haiyan, CHEN Zhizhong. The Maximum Likelihood Adjustment of the P-norm Distribution

[J]. Acta Geodaetica et Cartographica Sinica, 1996, 25(4): 241-246. (於宗侑, 孙海燕, 陈之中. 多元 P-范极大似然平差[J]. 测绘学报, 1996, 25(4): 241-246.)

- [7] HU Hongchang, SUN Haiyan. Parameter σ Estimation of P-norm Distribution[J]. Geomatics and Information Science of Wuhan University, 2002, 27(5): 483-485. (胡宏昌, 孙海燕. P-范分布参数 σ 的估计[J]. 武汉大学学报: 信息科学版, 2002, 27(5): 483-485.)
- [8] PAN Xiong, SUN Haiyan. Semiparametric P-norm Maximum Likelihood Regression[J]. Acta Geodaetica et Cartographica Sinica, 2005, 34(1): 30-34. (潘雄, 孙海燕. 半参数 P-范极大似然回归[J]. 测绘学报, 2005, 34(1): 30-34.)
- [9] XU Peiliang. Voronoi Cells, Probabilistic Bounds and Hypothesis Testing in Mixed Integer Linear Models[J]. IEEE Transaction on Information Theory, 2006, 52(2): 3122-3138.
- [10] CUI Xizhang, YU Zongchou, TAO Benzao, et al. Generalized Surveying Adjustment (new edition)[M]. Wuhan: Wuhan University Press, 2001. (崔希璋, 於宗侑, 陶本藻, 等. 广义测量平差(新版)[M]. 武汉: 武汉大学出版社, 2001.)
- [11] TEUNISSEN P. A New Method for Fast Carrier Phase Ambiguity Estimation[C] // Proceedings of IEEE PLANS' 94. Las Vegas: [s. n.], 1994: 562-573.

(责任编辑: 雷秀丽)

收稿日期: 2008-11-26

修回日期: 2009-06-03

第一作者简介: 李博峰(1983—), 男, 博士生, 主要从事测量数据处理, GPS 理论与算法方面的研究工作。

First author: LI Bofeng (1983—), male, PhD candidate, majors in surveying data processing, theory and algorithms of GPS.

E-mail: bofeng_li@163.com

声 明

《测绘学报》2009 年第 5 期文章《基于 X 波段雷达图像序列反演海洋表面流的算法研究》(作者: 王福友, 何宜军, 卢志忠, 袁赣南, 唐艳红——编者注)一文, 文章研究内容主要由第一作者王福友负责, 而第二、三作者何宜军, 卢志忠对研究内容不承担责任。

特此声明!

王福友