

基于自适应稳健最小 p 范数估计的轨道 确定方法研究*

潘晓刚^{1,2†} 李 强¹ 周海银¹

(1 国防科学技术大学九院 长沙 410073)

(2 北京航天飞行控制中心载人航天与深空探测飞行力学实验室 北京 100094)

摘要 当测轨数据误差不服从正态分布时, 传统的最小二乘 (LSE) 轨道确定方法将不是最优的. 为了获得高精度的定轨结果, 一种可行的策略是采用基于最小 p 范数 (L_p) 的轨道确定方法. 通过分析 L_p 估计的相关性质, 得出普通 L_p 估计不具有良好的抗差性的结论. 为抑制模型误差和异常值的影响, 提出了基于数据深度加权的稳健最小 p 范数估计方法, 并证明了相关性质, 得出了其崩溃点可以达到 $1/2$ 的结论. 最后, 通过残差分析和矩估计法自适应估计相关参数, 使得估计达到最大效率. 以天基空间目标监视系统为背景进行了仿真试验. 结果表明, 当观测数据存在系统误差或异常值时, 或者当目标动力学模型存在误差或者天基观测平台存在系统误差时, 即使观测数据服从正态分布, LSE 也不是最优的, 在这种意义下自适应稳健 L_p 估计轨道确定方法比传统轨道确定方法更加稳健, 定轨精度也更高.

关键词 天体力学; 轨道计算和定轨, 方法: 数值

中图分类号: P 135; **文献标识码:** A

1 引言

轨道确定是根据包含卫星轨道信息的测量数据, 在某一最优准则下, 从符合卫星动力学模型的轨道空间中, 寻找最优轨道的过程. 最优准则确定了轨道的选择标准, 不同的准则下, 将产生不同的轨道结果. 通常的轨道确定算法中, 最优准则是使得观测残差平方和最小, 即最小二乘准则. 在高斯 - 马尔科夫条件下, 最小二乘准则选择的轨道是最优的, 中心极限定理也说明当观测量无限多、影响因素相互独立时, 观测数据的随机误差服从正态分布. 然而在卫星轨道测控中, 首先很难保证影响因素的相互独立性, 尤其在天基测控过程中, 太空环境的复杂性和设备的精密性使得测量影响因素繁多且彼此间难以保证独立; 其次天基测控环境下的轨道确定往往需要每天或更短时间进行处理, 数据量少, 分布特性难以确定; 再次, 天基测控的模式下, 天基卫星的轨道作为已知条件直接

2009-02-25 收到原稿, 2010-01-18 收到修改稿

* 国家自然科学基金 (60974124) 及载人航天与深空探测飞行力学实验室开放基金资助项目

† panxiaogang@tom.com

参与轨道确定, 而受太空环境影响, 天基设备 (挠性太阳帆板等) 会发生高频抖动和低频震动等, 天基星历误差和平台误差将传播到定轨过程中, 在这种意义下, 测量数据将偏离正态分布, LSE 估计将不再是最优的. 同时, 当观测数据中包含有粗差时, 由于 LSE 不具有抗差性, 将会导致估计失效甚至发散. 虽然可通过残差剔除异常值, 但很难剔除那些临近边缘的粗差. 导致这些情况发生的原因是观测数据量的高斯 - 马尔科夫假设与实际情况发生了偏颇, 因此我们将误差分布的假设扩大, 以 p 范分布作为基本假设.

p 范分布是一族分布, 通常的退化分布、拉普拉斯分布、正态分布、均匀分布等都是 p 范分布的特例. 国内外许多学者已经对 L_p 估计进行了详细的研究^[1-3]. 本文主要研究 L_p 估计的稳健性和相应的改进措施. 通过残差分析自适应确定 p 值, 使估计达到最大效率, 同时应用数据深度加权技术, 使得 L_p 估计的崩溃点达到 1/2, 增强了估计的稳健性, 使之适合于天基环境下的轨道确定.

2 基于 LSE 的轨道确定算法

2.1 带不确定性误差的测量模型推导

设 $f_0(\mathbf{X}, t)$ 为精确建模的目标卫星受力的总表达式, $f_\varepsilon(\mathbf{X}, t)$ 为未知的摄动力或者目前无法认识的摄动力, 则卫星的状态模型可写为

$$\ddot{\mathbf{r}} = f_0(\mathbf{r}, \dot{\mathbf{r}}) + f_\varepsilon(\mathbf{r}, \dot{\mathbf{r}}), \quad (1)$$

其中 $\mathbf{r} = (x, y, z)^\top$, $\dot{\mathbf{r}} = (\dot{x}, \dot{y}, \dot{z})^\top$.

相应的卫星的状态转移矩阵为

$$\mathbf{X} = [\Phi(t, t_0) \cdot \Phi_\varepsilon(t, t_0)] \mathbf{X}_0, \quad (2)$$

其中 Φ_ε 为未知摄动力引起的状态转移矩阵, $\mathbf{X} = (\mathbf{r}, \dot{\mathbf{r}})$ 为目标卫星的状态.

设天基测控方程为

$$\mathbf{H} = \tilde{\mathbf{G}}(\mathbf{X}_T + \Delta\mathbf{X}_T, \mathbf{X}, t_0, t) + \varepsilon, \quad (3)$$

其中 \mathbf{H} 为天基观测数据, \mathbf{X}_T 为天基卫星星历, $\Delta\mathbf{X}_T$ 为星历误差和平台误差构成的总误差, ε 为观测误差.

根据 (2) 式, 观测方程可写作

$$\mathbf{H} = \tilde{\mathbf{G}}(\mathbf{X}_T + \Delta\mathbf{X}_T, (\Phi \cdot \Phi_\varepsilon) \mathbf{X}_0, t_0, t) + \varepsilon = \mathbf{G}(\mathbf{X}_T, \mathbf{X}_0, t_0, t) + \Delta\mathbf{G}_\varepsilon + \varepsilon, \quad (4)$$

其中 $\Delta\mathbf{G}_\varepsilon = \tilde{\mathbf{G}}(\mathbf{X}_T + \Delta\mathbf{X}_T, (\Phi \cdot \Phi_\varepsilon) \mathbf{X}_0, t_0, t) - \mathbf{G}(\mathbf{X}_T, \mathbf{X}_0, t_0, t)$, 表示动力学模型误差、平台误差所引起的观测误差.

在传统的轨道确定方法中, $\Delta\mathbf{G}_\varepsilon$ 将归结到 ε , 导致测量数据的分布特性发生改变.

2.2 轨道改进模型

设 t 时刻状态矢量 \mathbf{X}^* 与实际轨道足够接近, 则实际轨道就可在 \mathbf{X}^* 处进行泰勒展

开. 令 $\beta = X_0 - X_0^*$, 对观测模型进行泰勒展开可得

$$H = G(X_T, X_0^*, t_0, t) + \left. \frac{\partial G}{\partial X_0} \right|_{X_0=X_0^*} \Delta X_0 + \cdots + \Delta G_\varepsilon + \varepsilon. \quad (5)$$

令 $Y = H - G(X_T, X_0^*, t_0, t)$, $B = \left. \frac{\partial G}{\partial X_0} \right|_{X_0=X_0^*}$, $\beta = \Delta X_0$, (5) 式可表示为

$$Y = B \cdot \beta + \Delta G_\varepsilon + \varepsilon = B \cdot \beta + v, \quad (6)$$

(6) 式为轨道改进的基本方程, 其中 $v = \Delta G_\varepsilon + \varepsilon$ 表示观测总误差.

3 基于 L_p 的轨道确定算法

3.1 M 估计

选定一个定义于 R^1 的函数 ρ, β 的 M 估计定义为 $H_\rho(\beta)$ 的一个最小值点

$$H_\rho(\hat{\beta}) = \min \left(H_\rho(\beta) : H_\rho(\beta) = \sum_{i=1}^n \rho(Y_i - B_i \beta) \right). \quad (7)$$

常用 M 估计的函数有 3 类^[4], 分别满足以下条件:

(B ρ) $\rho(x)$ 在 $x = 0$ 处达到极小值 -1; 当 $x < 0$ 时, ρ 非增; 当 $x > 0$ 时, ρ 非减; 且当 $|x| \rightarrow \infty$ 时, $\rho(x) \rightarrow 0$;

(U ρ) $\rho(x)$ 在 $x = 0$ 处达到极小值 0; $\rho(x)$ 关于 0 对称; 当 $x > 0$ 时, ρ 非减, 当 $|x| \rightarrow \infty$ 时, $\rho(x) \rightarrow \infty$; $\psi = \rho'$ 在 R^1 上连续, 且存在 $x_0 > 0$, 使 ψ 在 $(0, x_0]$ 上非减, 在 (x_0, ∞) 上非增;

(C ρ) $\rho(x)$ 是凸函数, $\psi = \rho'$ 处处存在且 $\psi(-\infty) < 0 < \psi(+\infty)$.

设 $\rho(u)$ 的导数 $\rho' = \psi$ 在 R^1 上处处存在, 则 $\hat{\beta}$ 满足方程

$$\sum_{i=1}^n \psi(Y_i - B_i \beta) B_i^T = 0. \quad (8)$$

(8) 式可以看作 M 估计的另一种定义. 令 $w_i = \frac{\psi(v_i)}{v_i}$, 则 $\sum_{i=1}^n w_i (Y_i - B_i \beta) B_i^T = 0$, 即 β 的 M 估计解为

$$\hat{\beta} = (B^T w B)^{-1} B^T w Y. \quad (9)$$

L_p 估计是 M 估计的一种, 因此 M 估计的优良性质 (强相合性、渐进正态) 都可以在 L_p 估计中得以体现, 具体请参考文献 [5].

3.2 p 范分布

设 v 服从 m 维 p 范分布, 密度函数可表示为

$$f(v) = \frac{p^m}{2^m \Gamma^m(\frac{1}{p}) |D_v|^{\frac{1}{2}}} (\lambda)^m \exp \left(- \left(\lambda \left\| D_v^{-\frac{1}{2}} (v - \mu) \right\|_p \right)^m \right), \quad (10)$$

$$\lambda = \sqrt{\frac{\Gamma(\frac{3}{p})}{\Gamma(\frac{1}{p})}},$$

其中 μ, λ, D_v 为特征参数.

在轨道确定中, 残差序列可以看作是一维数据, 即上式 $m=1$, 一维 p 范分布可化为

$$f(v) = \frac{p\lambda}{2\Gamma\left(\frac{1}{p}\right)\sigma} \exp\left(-\left(\frac{\lambda}{\sigma}|(v-\mu)|\right)^p\right).$$

易得 $Ev = \mu$. 因此, 对于轨道确定中的残差可以假设 $\mu = 0$, 否则进行残差均值化处理, 即令 $v = v - E\mu$.

令 ρ 函数为 $\rho = \ln(f(v))$, 在确定 D_v 和 p 后, $\min(\ln(f(v)))$ 等价于 $\min\left(\left\|D_v^{-\frac{1}{2}}v\right\|_p\right)$, 即 L_p 估计和极大似然估计是等价的, 则 $\psi(v) = |v|^{p-1} \text{sign}(v)$, $\psi'(v) = (p-1)|v|^{p-2}$.

根据 M 估计求解算法可以获得 L_p 准则下的估计量.

3.3 p 范分布的稳健性分析

评价估计稳健性可以通过影响函数和崩溃污染率来描述. 影响函数描述估计的局部稳健性, 崩溃污染率描述了整体稳健性, 即使得估计崩溃的粗差个数的下确界.

(1) 影响函数

定义 1: 设样本观测值 Y_1, Y_2, \dots, Y_n 独立同分布, 其模型分布为 F , $X(F)$ 为 F 的泛函, 对任何实数 l , 以 δl 记退化于点 l 的概率分布, 并记污染分布为 $F_\varepsilon = (1-\varepsilon)F + \varepsilon\delta l$ ($0 \leq \varepsilon \leq 1$). 当极限

$$IF(l; F, X) = \lim_{\varepsilon \rightarrow 0} \frac{X(F_\varepsilon) - X(F)}{\varepsilon} = \left. \frac{dX(F_\varepsilon)}{d\varepsilon} \right|_{\varepsilon=0}$$

存在时, 称之为泛函 $X(F)$ 的影响函数.

对于 M 估计而言, 由于

$$\int \psi(\theta, x) f(\theta, x) dx = \int \frac{\partial \ln f(\theta, x)}{\partial x} f(\theta, x) dx = \int f'(\theta, x) dx = 0,$$

所以对于污染分布有

$$\int \psi(\theta_\varepsilon, x) f_\varepsilon(\theta, x) dx = 0,$$

$$\frac{d\theta_\varepsilon}{d\varepsilon} \int \psi'(\theta, x) f(\theta, x) dx - \int \psi(\theta, x) f(\theta, x) dx + \int \psi(\theta, x) \delta(x - x_0) dx = 0,$$

则影响函数可描述为

$$IF(\theta, x_0) = \frac{d\theta_\varepsilon}{d\varepsilon} = \frac{-\psi(\theta, x_0)}{\int \psi'(\theta, x) f(\theta, x) dx}.$$

由于 x_0 的任意性, 可以得到影响函数的另一种形式

$$IF(v) = \frac{d\theta_\varepsilon}{d\varepsilon} = \frac{-\psi(v)}{\int \psi'(v) f(v) dv} = \frac{-\psi(v)}{E(\psi'(v))}. \quad (11)$$

将 L_p 估计的 ψ 函数代入,

$$\begin{aligned} E(\psi'(v)) &= \int_{-\infty}^{\infty} \left((p-1)|v|^{p-2} \right) \frac{p}{2\Gamma\left(\frac{1}{p}\right)\sigma} \lambda \exp\left(-\left(\left|\frac{v}{\sigma}\right|\right)^p\right) dv \\ &= 2 \int_0^{\infty} \left((p-1) \left(\frac{\sigma}{\lambda} y^{\frac{1}{p}}\right)^{p-2} \right) \frac{p_0}{2\Gamma\left(\frac{1}{p}\right)\sigma} \lambda \exp(-y) \frac{\sigma}{p\lambda} y^{\frac{1}{p}-1} dy \\ &= \frac{p-1}{\Gamma\left(\frac{1}{p}\right)} \left(\frac{\sigma}{\lambda}\right)^{p-2} \Gamma\left(\frac{p-1}{p}\right). \end{aligned}$$

由此可得到 L_p 估计的影响函数为

$$IF(p) = \frac{-|v|^{p-1} \text{sign}(v)}{\frac{p-1}{\Gamma\left(\frac{1}{p}\right)} \left(\frac{\sigma}{\lambda}\right)^{p-2} \Gamma\left(\frac{p-1}{p}\right)}. \quad (12)$$

对不同的 p 值, 可以得到影响函数的具体形式为

$$\begin{aligned} IF(p=1) &= \frac{\sigma}{\sqrt{2}} \text{sign}(v) \\ IF(1 < p < 2) &= (k \cdot v^{\text{mod}(p)} \sigma^{1-\text{mod}(p)}) \text{sign}(v) \\ IF(p=2) &= -v \text{sign}(v) \\ IF(p > 2) &= -\frac{k \cdot v^{p-1} \text{sign}(v)}{\sigma^{p-2}} \end{aligned},$$

其中 mod 为取余计算, p 的取值影响系数 k 的数值.

从 L_p 估计的影响函数可以看出, 只有当 $p=1$ 时, 影响函数才有界, 具有很强的抗差能力, 即所谓的最小一乘估计. 当 $1 < p < 2$ 时, L_p 估计具有较强的抗差能力, 但仍然会受到粗差的污染, 需要新的加权方式, 以改善估计的抗差性.

(2) 崩溃点

崩溃点有渐进崩溃点和样本崩溃点两类. Hampel 的渐进崩溃点体现了大范围程度上的统计量所能承受模型偏差的程度. 样本崩溃点从大范围上描述统计量能承受多大比例的离群值.

$$\varepsilon^* \triangleq \varepsilon(F_0, \mathbf{Y}) \triangleq \inf\{\varepsilon > 0 : \xi(\varepsilon, F_0, \mathbf{T}) = \infty\}, \quad (13)$$

其中 $\xi(n, \beta, \mathbf{Y})$ 表示 $\hat{\beta}(\tilde{\mathbf{Y}})$ 和 $\hat{\beta}(\mathbf{Y})$ 的最大偏差.

$$\xi(n, \beta, \mathbf{Y}) = \sup \left\| \hat{\beta}(\tilde{\mathbf{Y}}) - \hat{\beta}(\mathbf{Y}) \right\|, \quad (14)$$

其中 $\tilde{\mathbf{Y}}$ 表示用 r 个粗差 \mathbf{Y}_r 代替原样本中 r 个测量值所得到的新样本, 对应的崩溃污染率称为替换型 $\varepsilon_R^*(\beta, \mathbf{Y})$. 如果在样本 \mathbf{Y} 后面添加 r 个污染数据, 对应的崩溃污染率称为添加型 $\varepsilon_A^*(\beta, \mathbf{Y})$.

定义 2: 替换崩溃污染率

$$\varepsilon_R^*(\beta, \mathbf{Y}) = \min\left(\frac{r}{n} : \xi(n, \beta, \mathbf{Y}) = \infty\right). \quad (15)$$

定义 3: 添加崩溃污染率

$$\varepsilon_A^*(\beta, \mathbf{Y}) = \min\left(\frac{r}{n} : \xi(n, \beta, \mathbf{Y}) = \infty\right). \quad (16)$$

对于 M 估计, 文献 [6] 证明了下列定理:

定理 1: 如果满足条件 $(U\rho)$, 则对任何 n , 3 种样本崩溃点分别是 $\varepsilon_A = 1/2$, $\varepsilon_{SR} = \varepsilon_R = \frac{\lfloor (n+1)/2 \rfloor}{n}$, 其中 $\lfloor \cdot \rfloor$ 是取整算子.

定理 2: 假设满足条件 $(B\rho)$. 记 $N = \{1, 2, \dots, n\}$, $I_k = \{i_1, i_2, \dots, i_k\}$, 其中 $i_s \in N, 1 \leq s \leq k, 1 \leq k \leq n$. 对给定的样本 $X = (x_1, x_2, \dots, x_n)$, 令

$$A_0 = 0, B_0 = 0,$$

$$A_k = \sup_t \sum_{i=1}^k -\rho(x_i - t), B_k = \min_{I_k \subset N} \sup_t \sum_{i \in I_k} -\rho(x_i - t),$$

$$m = \min \{k : k \geq A_{n-k}\}, r = \min \{k : k \geq B_{n-k}\}.$$

设 a 是满足 $\lceil A_n \rceil \leq a \leq \lfloor A_n \rfloor + 1$ 的整数, 则

$$\varepsilon_A = \frac{a}{n+a}, \frac{m}{n} \leq \varepsilon_{SR} \leq \frac{m+1}{n}, \frac{r}{n} \leq \varepsilon_R \leq \frac{r+1}{n}.$$

如果存在 $0 < c < +\infty$, 使 $|x| > c$ 时 $\rho(x) = 0$, 则

$$a = \lceil A_n \rceil, \varepsilon_{SR} = \frac{m}{n}, \varepsilon_R = \frac{r}{n}.$$

定理 3: 设 ρ 满足 $(C\rho)$, 则对任何 n , 当 ψ 有界时,

$$\varepsilon_{SR} = \varepsilon_R = \varepsilon_A = \min \left\{ \frac{-\psi(-\infty)}{\psi(\infty) - \psi(-\infty)}, \frac{\psi(\infty)}{\psi(\infty) - \psi(-\infty)} \right\};$$

当 ψ 无界时,

$$\varepsilon_{SR} = \varepsilon_R = \frac{1}{n}, \varepsilon_A = \frac{1}{n+1}.$$

从上述定理可以看出, 对于 p 范估计的 $\psi(v) = |v|^{p-1} \text{sign}(v)$, 其崩溃点为 $\varepsilon_{SR} = \varepsilon_R = \frac{1}{n} \rightarrow 0, \varepsilon_A = \frac{1}{n+1} \rightarrow 0$. 因此, 基于 p 范估计的轨道确定算法不是稳健的估计算法.

4 稳健的 p 范估计

从前面的讨论可以看出, 普通的 p 范估计并不是稳健估计, 虽然通过 p 值的合理确定可以获得良好的局部稳健性, 但对于估计算法的整体抗差性能而言, 不管 p 值如何选择, 对于高维数据的 p 范估计不具有抗差性. 从崩溃点的定义可以看出, 如果对 p 范估计的函数加以约束, 则可以有效提高崩溃点. 我们采用数据深度理论对函数进行加权.

4.1 数据深度的定义及性质

数据深度是高维数据的次序统计量, 描述了给定样本数据集下, 数据点距离数据中心的靠近程度. 越靠近中心的高维数据深度越大, 越远离中心的高维数据深度越小. 在这样的意义下, 隐藏在低维数据中的信息可以通过数据深度进行量化. 本文将数据深度应用在估计理论中, 利用数据深度作为量化高维粗差的工具. 数据深度作为一种新的统计工具, 具有广阔的应用前景.

Zuo 等 [7] 在 2000 年提出一般数据深度的定义, 并给出构造数据深度的 4 种方法.

定义 [7]: 设 $D(\cdot; \cdot)$ 是 $R^d \times \mathcal{F} \rightarrow R$ 的有界非负函数, 其中 \mathcal{F} 表示 R^d 上所有分布组成的类, 且满足以下 4 条性质:

C1 (仿射不变性) 对任意 $q \times q$ 阶的可逆矩阵 \mathbf{A} , 任意的 q 维向量 \mathbf{b} 和 \mathbf{x} , 以及 R^d 上的任意随机向量 \mathbf{Y} , 有 $D(\mathbf{Ax} + \mathbf{b}; F_{\mathbf{AY} + \mathbf{b}}) = D(\mathbf{x}; F_{\mathbf{Y}})$, 其中 $F_{\mathbf{Y}}$ 表示 \mathbf{Y} 的分布;

C2 (对称中心最深) $D(\theta; F) = \sup_{\mathbf{x} \in R^d} D(\mathbf{x}; F_{\mathbf{Y}})$, 对 \mathcal{F} 中任意以 θ (相对于某种多元分布的定义) 为中心的分佈 F 都成立;

C3 (沿最深点向外的射线深度递减) 对任意给定的 $F \in \mathcal{F}$, 若 D 在 θ 处达到最大, 则对所有 $\lambda \in [0, 1]$, 都有 $D[\mathbf{x}; F] \leq D[\theta + \lambda(\mathbf{x} - \theta); F]$ 成立;

C4 (无穷远处深度为零) 对任意给定的 $F \in \mathcal{F}$, 当 $\|\mathbf{x}\| \rightarrow \infty$ 时, 则称 $D(\mathbf{x}; F) \rightarrow 0$ 是一个数据深度 (统计深度函数).

目前比较常见的数据深度主要包括: 半空间深度 [8]、单纯型深度 [9]、单形体积深度 [10]、投影深度 [11]. 由于投影深度的意义直观、稳健性好, 本文采用投影深度作为数据深度的加权方法.

投影深度是基于“离群测度” $O(\mathbf{x}; F)$ 提出来的一种数据深度, R^d 中 \mathbf{x} 关于分布 F 的投影深度为 $PD(\mathbf{x}; F) = (1 + O(\mathbf{x}; F))^{-1}$, 其中 $O(\mathbf{x}; F) = \sup_{\|\mathbf{u}\|=1} \frac{|\mathbf{u}'\mathbf{x} - \text{Med}(\mathbf{u}'\mathbf{Y})|}{\text{MAD}(\mathbf{u}'\mathbf{Y})}$, $\text{MAD}(\mathbf{Y}) = \text{Med}(|\mathbf{Y} - \text{Med}(\mathbf{Y})|)$, Med 表示一元中位数. $O(\mathbf{x}; F)$ 是点 \mathbf{x} 在分布 F 所有方向的一维投影中相对偏离最大的值, 度量了点 \mathbf{x} 偏离分布 F 的程度.

4.2 数据深度加权的 p 范估计性质

定义 4: 选定一个定义于 R^1 的函数 ρ, β 的深度加权 M 估计定义为 $H_\rho(\beta)$ 的一个最小值点

$$H_\rho(\tilde{\beta}) = \min \left(H_\rho(\beta) : H_\rho(\beta) = \sum_{i=1}^n \rho(\omega(PD(z_i; F_n))(Y_i - \mathbf{B}_i\beta)) \right), \quad (17)$$

其中 $\mathbf{Z}^n = (z_1, \dots, z_i, \dots, z_n)$, $z_i = (B_i, y_i)^T \in R^{d+1}$.

当 ρ 函数取 $\rho = \ln(f(v))$ 时, 上式 $\tilde{\beta}$ 即为基于数据深度加权的 p 范估计.

定理 4: 若权函数的一阶导数有界, 则深度加权函数的崩溃点为 $1/2$, 即

$$\varepsilon^*(\omega(PD(\mathbf{z}; \mathbf{Z}^n)); \mathbf{Z}) = 1/2. \quad (18)$$

证明: 当 $|z_i| \rightarrow \infty$ 时, 有

$$\sup_{\|\mathbf{u}\|=1} \text{Med}(\mathbf{u}'\mathbf{Z}^n) < \infty, \quad \sup_{\|\mathbf{u}\|=1} \text{MAD}(\mathbf{u}'\mathbf{Z}^n) < \infty,$$

则

$$\begin{aligned} PD(z_i; \mathbf{Z}^n) &= \left(1 + \sup_{\|\mathbf{u}\|=1} \frac{|\mathbf{u}'z_i - \text{Med}(\mathbf{u}'\mathbf{Z}^n)|}{\text{MAD}(\mathbf{u}'\mathbf{Z}^n)} \right)^{-1} \\ &\leq \left(1 + \frac{\sup_{\|\mathbf{u}\|=1} |\mathbf{u}'z_i - \text{Med}(\mathbf{u}'\mathbf{Z}^n)|}{\sup_{\|\mathbf{u}\|=1} \text{MAD}(\mathbf{u}'\mathbf{Z}^n)} \right)^{-1} \end{aligned}$$

$$\begin{aligned}
& \leq \frac{\sup_{\|\mathbf{u}\|=1} \text{MAD}(\mathbf{u}'\mathbf{Z}^n)}{\sup_{\|\mathbf{u}\|=1} |\mathbf{u}'z_i - \text{Med}(\mathbf{u}'\mathbf{Z}^n)| + \sup_{\|\mathbf{u}\|=1} \text{MAD}(\mathbf{u}'\mathbf{Z}^n)} \\
& \leq \frac{\sup_{\|\mathbf{u}\|=1} \text{MAD}(\mathbf{u}'\mathbf{Z}^n)}{\|z_i\| - \sup_{\|\mathbf{u}\|=1} |\text{Med}(\mathbf{u}'\mathbf{Z}^n)| + \sup_{\|\mathbf{u}\|=1} \text{MAD}(\mathbf{u}'\mathbf{Z}^n)} \\
& \leq \frac{2 \sup_{\|\mathbf{u}\|=1} \text{MAD}(\mathbf{u}'\mathbf{Z}^n)}{\|z_i\|}
\end{aligned}$$

由权函数的性质可以得到 $\omega(PD(z; \mathbf{Z}^n)) \leq M \cdot PD(z; \mathbf{Z}^n)$, 由于 $\varepsilon^*(\text{Med}(\mathbf{u}'\mathbf{Z}^n)) = \varepsilon^*(\text{MAD}(\mathbf{u}'\mathbf{Z}^n)) = 1/2$, 则有

$$\|z_i \omega(PD(z_i; \mathbf{Z}^n))\| \leq 2M \sup_{\|\mathbf{u}\|=1} \text{MAD}(\mathbf{u}'\mathbf{Z}^n) \leq 1/2.$$

定理 5: 若 $\mathbf{B}^n = \{B_1, \dots, B_n\} \subset R^d$ 为在一般位置的数据集, $0 < \omega(\cdot) \leq 1$, 则当 $n > d$ 时, 统计量的崩溃点和权函数崩溃点一致^[12].

4.3 数据深度加权的 p 范估计求解方法

根据 L_p 估计的求解方程

$$\sum_{i=1}^n \left\| \omega(D(z_i; F_n)) (Y_i - \mathbf{B}_i \tilde{\beta}) \right\|_p = \min,$$

令等价权 $w'_i = \omega^p |v_i|^{p-2}$, 则上述方程转化为最小二乘的估计形式

$$\sum_{i=1}^n w'_i (Y_i - \mathbf{B}_i \tilde{\beta})^2 = \min.$$

根据最小二乘估计方法可以估计上式.

4.4 p 值的自适应估计方法

(1) 矩估计法

根据该方法, 可以得到 $Ev^2 = \mu^2 + \sigma^2$. 同理,

$$Ev^4 = \frac{\Gamma\left(\frac{1}{p}\right) \Gamma\left(\frac{5}{p}\right)}{\Gamma^2\left(\frac{3}{p}\right)} \sigma_0^4 + 6\sigma^2 \mu^2 + \mu^4.$$

由此, 可以得到

$$\begin{aligned}
\hat{\mu} &= \frac{1}{n} \sum_{i=1}^n v_i, \\
\hat{\sigma}_0^2 &= \frac{1}{n} \sum_{i=1}^n v_i^2 - \hat{\mu}^2, \\
\frac{\Gamma\left(\frac{1}{p}\right) \Gamma\left(\frac{5}{p}\right)}{\Gamma^2\left(\frac{3}{p}\right)} &= \frac{1}{\hat{\sigma}_0^4} \left(\frac{1}{n} \sum_{i=1}^n x_i^4 - 6\hat{\sigma}_0^2 \hat{\mu}^2 - \hat{\mu}^4 \right).
\end{aligned}$$

(2) 极大似然估计法

L_p 估计的对数似然函数为

$$\Phi = \ln(L|\mu, \sigma_0, p) = n \{ \ln p - \ln 2 - \ln \sigma_0 - \ln \Gamma(1/p) \} \\ + n \{ \ln \Gamma(3/p) - \ln \Gamma(1/p) \} / 2 - \sigma_0^p (\Gamma(3/p) / \Gamma(1/p))^{p/2} \sum_{i=1}^n |L_i - \mu|^p.$$

对似然函数求导, 组成非线性方程组, 即可以得到残差 v 的 μ, p, σ_0 的自适应估计值. 具体求解方法可以见文献 [13].

4.5 基于自适应稳健 L_p 估计的轨道确定算法

设初始收敛因子为 λ , 允许误差为 δ , $p_0 = 2$; 观测值和计算值的残差均方根为

$$\text{RMS} = \left(\sum_{i=1}^n \frac{1}{n-m} [v_i^T v_i] \right)^{1/2}.$$

设计算法如下:

Step1 令 $p_i = p_0$, 计算轨道初值 $\mathbf{X}_0 = \mathbf{X}_0^* + \lambda \Delta \mathbf{X}_0$;

Step2 计算 OC 残差, 得到估计值 $\hat{\mu}, \hat{p}_i, \hat{\sigma}_0$; 若 $|p_i - p_{i-1}| < 0.001$, 转向 Step4; 否则转向 Step3;

Step3 根据 OC 残差计算数据深度权, 令 $p = \hat{p}_i$ 并构建等价权矩阵; 按照 L_p 估计求解公式计算轨道初值 $\mathbf{X}_0 = \mathbf{X}_0^* + \lambda \Delta \mathbf{X}_0$, 转向 Step2;

Step4 以 \mathbf{X}_0 为初值, 采用 Adams-Cowell 数值积分法积分轨道, 计算 RMS. 若 $\text{RMS}^j < \text{RMS}^{j-1}$, 转向 Step5; 若 $\text{RMS}^j > \text{RMS}^{j-1}$, 令 $\lambda = \frac{\lambda}{2}$, 转向 Step1;

Step5 若 $\|\text{RMS}^j - \text{RMS}^{j-1}\| < \delta$, $\mathbf{X}_0^{**} = \mathbf{X}_0$, 转向 Step6; 否则, 令 $\lambda = 1$, $\mathbf{X}_0^* = \mathbf{X}_0$, 转向 Step1;

Step6 \mathbf{X}_0^{**} 为初轨, 计算卫星轨道, 并与标准轨道作差, 统计误差特性.

值得说明的是, 当目标函数为非严格凸时, 会有多个极值点, 为了避免收敛到局部最优值点, 需要有较高精度的初始轨道作为初值. 事实上, 当 $p \geq 2$ 时, 目标函数严格凸; 当 $1 < p \leq 2$ 时, 在 $|v| > 0$ 处为凸函数; 只有当 $p = 1$ 时为非凸函数, 此时需要用线性规划方法求解. 同时为避免当 $p < 2$, 残差 $v_i \rightarrow 0$ 时, 等价权 $w'_i \rightarrow \infty$, 令 $w'_i = w_i^{p/2} |v_i + e|^{p-2}$, 其中 e 为任意小的正数 [1].

5 数据验证

5.1 仿真条件

以天基空间目标监视为仿真环境. 设天基卫星为轨道高度 800 km 的极地卫星, 目标卫星选择为 Cosmos 2227 Rocket 卫星. 天基卫星通过星载光学传感器对目标卫星常规扫描, 当捕获到目标后采用闭环跟踪方式获得测量数据, 并进行轨道确定, 为尽可能模拟太空真实环境, 具体仿真条件如表 1.

卫星平台的高频抖动和震动等价于传感器中心点不稳定, 仿真中在平台每个方向加入系统误差, 如图 1.

表 1 仿真试验条件

Table 1 Conditions of simulation experiments

Models and conditions	Description
standard dynamic model	perturbation of nonspheric gravitational field (JGM3, 70×70), atmospheric drag (DTM94), lunisolar gravitational perturbation, solar radiation pressure
calculating dynamic model	perturbation of nonspheric gravitational field (JGM3, 20×20), atmospheric drag (DTM94, to be estimated) lunisolar gravitational perturbation, solar radiation pressure (to be estimated)
platform error of space based satellite	jitter error with high frequency and vibration error with low frequency $a \sin\left(\frac{2\pi}{T}t + \varepsilon\right)$
measuring accuracy	random error (10 arcsec, 3σ)
time and frequency	86 400 s, 0.016 Hz
ephemeris of target satellite	Cosmos 2227 Rocket 1 22285U 92093B 08240.66512461 -.00000062 00000-0 -72814-5 0 7026 2 22285 071.0211 115.9845 0005051 323.5284 036.5482 14.14727218807734

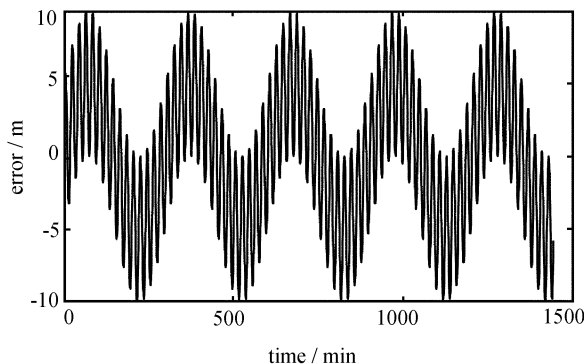


图 1 天基卫星星历误差

Fig.1 Ephemeris error of space based satellite

对天基光学传感器而言, 目标的光照强度和阳光幅角是最重要的观测条件, 传感器的作用距离也是重要的条件之一. 在这种假设下, 得到目标的阳光幅角如图 2, 可观测性如图 3.

5.2 仿真结果及分析

为验证本文提出的算法的可靠性, 采用如下几种策略进行仿真试验.

试验 1: 目标动力学特性完全已知, 即计算用动力学模型和仿真一致, 天基平台没有误差, 测量数据中不包含异常值.

试验 2: 测量数据中没有异常值, 其它条件见表 1.

试验 3: 测量数据中包含有较小异常值, 其它条件见表 1.

试验 4: 测量数据中包含有粗大异常值, 其它条件见表 1.

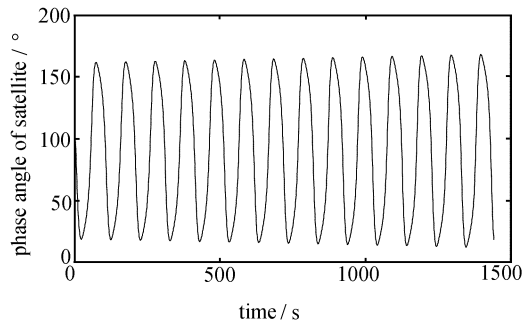


图 2 目标卫星的阳光幅角 (90° 以下为可见区域)
 Fig.2 Phase angle determined by the relative orientation of the sun, the space based satellite and the target (visibility area within 90°)

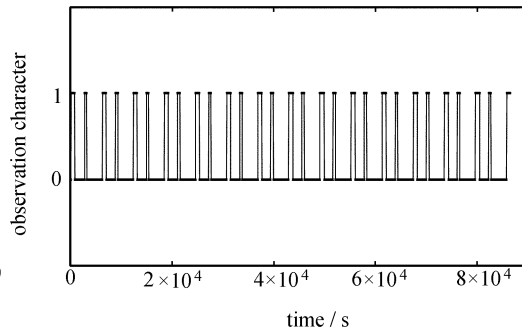


图 3 目标卫星的可观测性 (1 为可见弧段; 0 为不可见弧段)
 Fig.3 Observability of target satellite (1 is visible; 0 is invisible)

在没有其它误差源的情况下, 测量数据误差是正态分布, LSE 估计和 L_p 估计基本一致, 见图 4 和图 5. 轨道确定结果图中, 实线、虚线和点线分别表示径向、法向和切向的定轨误差. 此时 p 值的估计值在 2 附近, 轨道确定的结果也基本一致.

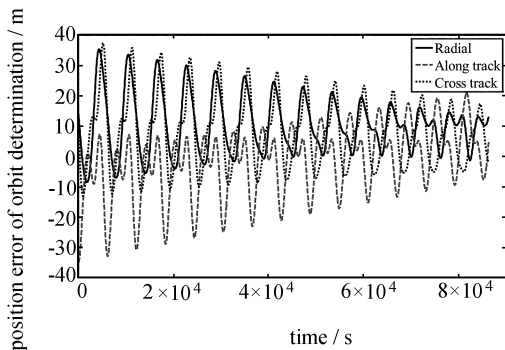


图 4 试验 1 LSE 的轨道确定结果
 Fig.4 Orbit determination results of LSE method in experiment 1

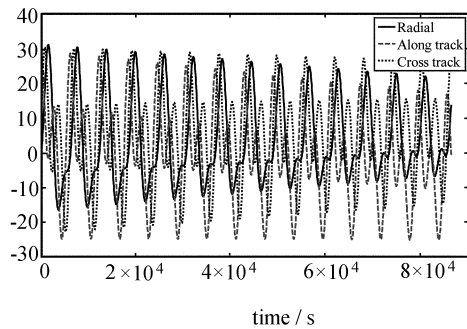


图 5 试验 1 自适应稳健 L_p 估计的轨道确定结果
 Fig.5 Orbit determination results of the self-adaptive stable L_p method in experiment 1

虽然测量数据没有其它误差, 当存在动力学模型误差和传感器平台误差时, 将直接传播到轨道确定中, 测量误差不再服从正态分布. 因此, L_p 估计的优越性得以体现, 估计精度明显高于 LSE 估计, 见图 6 和图 7.

存在异常值时, 由于 LSE 没有抗差性, 定轨误差将变大, 见图 8. 但稳健 L_p 估计能

够有效抑制异常值的影响, 定轨结果优于 LSE 估计, 见图 9.

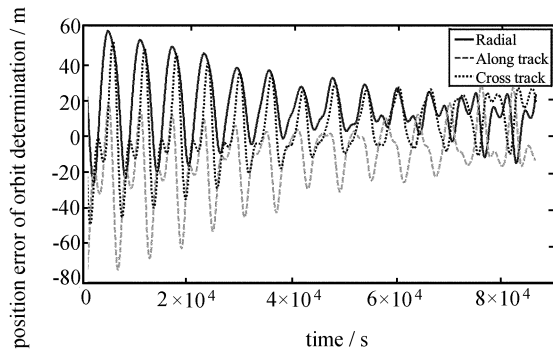


图 6 试验 2 LSE 的轨道确定结果
 Fig.6 Orbit determination results of LSE method in experiment 2

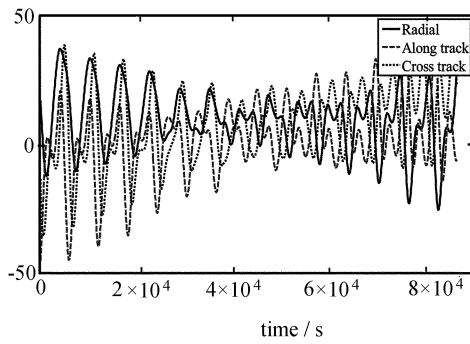


图 7 试验 2 自适应稳健 Lp 估计的轨道确定
 Fig.7 Orbit determination results of the self-adaptive stable Lp method in experiment 2

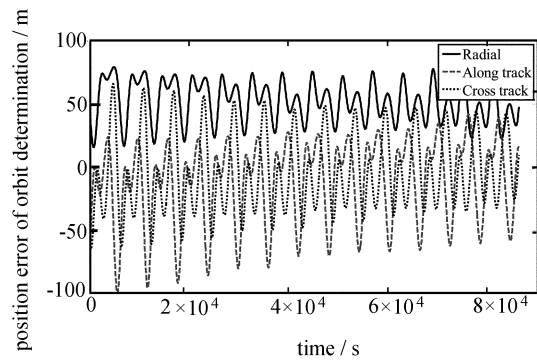


图 8 试验 3 LSE 的轨道确定结果
 Fig.8 Orbit determination results of LSE method in experiment 3

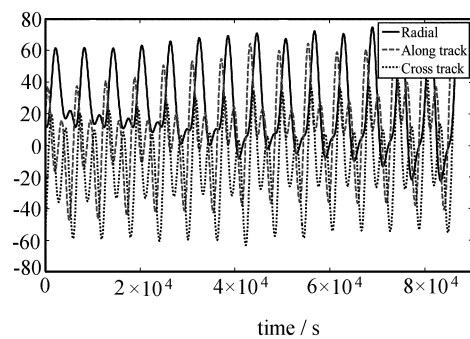


图 9 试验 3 自适应稳健 Lp 估计的轨道确定
 Fig.9 Orbit determination results of the self-adaptive stable Lp method in experiment 3

当存在较大的异常值时, LSE 的估计结果直接受到污染, 造成较大轨道确定误差, 见图 10. 从残差图 (见图 11) 可以看出, 存在明显的系统误差. 而此时稳健 Lp 估计的优越性得以最大的体现, 定轨精度受异常值影响不大, 见图 12. 从稳健 Lp 估计的残差 (图 13) 来看, 呈明显的白噪声趋势. 因此系统误差可以通过 Lp 估计有效抑制, 从而提高轨道确定精度.

6 结论

天基测控是未来测控系统的主要发展方向, 在不了解天基环境的情况下, 研究稳健的轨道确定方法对于提高整个系统的稳定性具有重要意义. 由于自适应稳健 Lp 估计轨道确定算法具有大崩溃点的优良性质, 在观测环境受到非正常干扰时, 仍然可以完成轨道

确定. 常规状态下, 当目标动力学模型存在不确定性误差以及天基卫星存在星历误差和平台误差时, 观测方程的误差分布不再是正态分布, 传统的轨道确定算法将不再最优, 在这种意义下自适应稳健 L_p 估计将明显优于传统算法, 而自适应的 p 值估计方法可以使得估计达到最大效率. 由于数据深度具有良好的理论背景和应用价值, 数据深度加权方法为稳健估计理论提供了一种新的加权思路, 但一个明显的缺点是计算复杂度高, 如何降低计算复杂度是实际应用的关键问题.

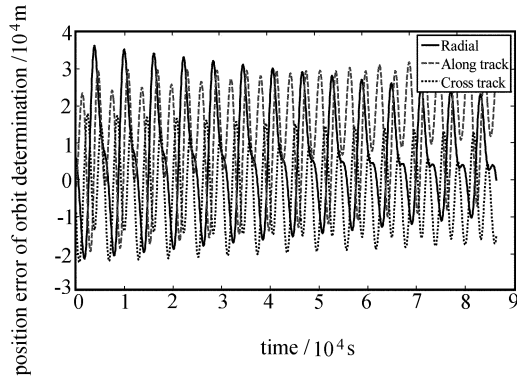


图 10 试验 4 LSE 的轨道确定结果

Fig. 10 Orbit determination results of LSE method in experiment 4

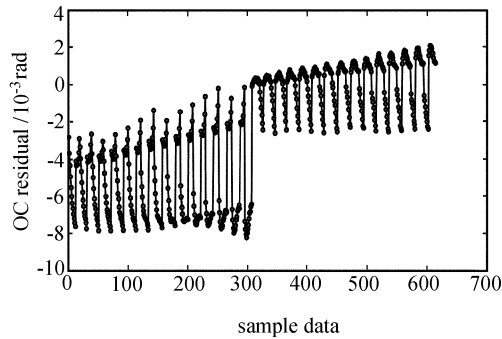


图 11 试验 4 LSE 的轨道确定残差

Fig. 11 Residuals of LSE method in experiment 4

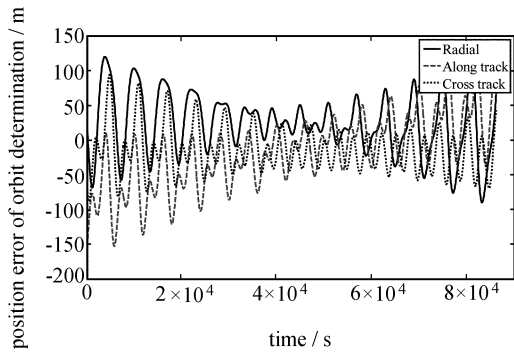


图 12 试验 4 自适应稳健 L_p 估计的轨道确定结果

Fig. 12 Orbit determination results of the self-adaptive stable L_p method in experiment 4

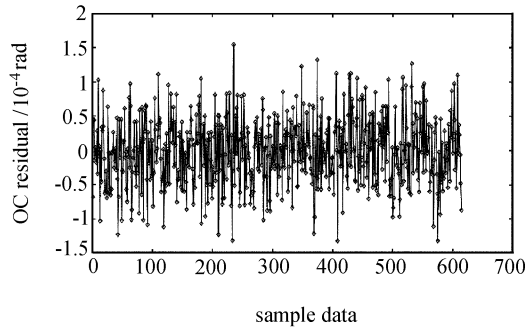


图 13 试验 4 自适应稳健 L_p 估计的轨道确定残差

Fig. 13 Residuals of the self-adaptive stable L_p method in experiment 4

致谢 非常感谢审稿专家的宝贵意见.

参 考 文 献

- [1] 郑作亚, 卢秀山, 彭军还. 天文学报, 2007, 48: 211
- [2] 孙海燕. p 范分布理论及其在现代测量数据处理中的应用. 武汉: 武汉测绘科技大学, 1995

- [3] 潘雄. 半参数模型的估计理论及其应用. 武汉: 武汉大学, 2005
- [4] 张建, 李国英. 数学进展, 1998, 27: 403
- [5] 陈希孺, 赵林城. 线性模型中的 M 方法. 上海: 上海科学技术出版社, 1996: 59-117
- [6] Zhang J, Li G Y. The Annals of Statistics, 1998, 26: 1170
- [7] Zuo Y J, Serfling R. The Annals of Statistics, 2000, 28: 461
- [8] Tukey J W. Proceedings of the International Congress of Mathematicians. 1974: 523-531
- [9] Liu R Y. The Annals of Statistics, 1990, 18: 405
- [10] Oja H. Statistics and Probability Letters, 1983, 1: 327
- [11] Donoho D L. Breakdown Properties of Multivariate Location Estimators. Cambridge: Harvard University, 1982
- [12] 李强. 两类数据深度及深度加权 M 估计. 长沙: 国防科技大学, 2008
- [13] 孙海燕, 潘雄. 武汉大学学报, 2003, 28: 551

Research on the Method of Orbit Determination Based on the Self-adaptive Stable Least p -norm Estimate

PAN Xiao-gang^{1,2} LI Qiang¹ ZHOU Hai-yin¹

(1 College 9, National University of Defense Technology, Changsha 410073)

(2 Manned Space Flight and Deep Space Exploration Flight Mechanics Lab, Beijing Aerospace Control Center, Beijing 100094)

ABSTRACT The traditional LSE (least square estimation) method for orbit determination will not be optimal if the error of observational data of orbit determination does not obey Gaussian distribution. In order to solve this problem, the least p -norm (Lp) estimation method for orbit determination is presented to deal with the non-Gaussian distribution cases because of the adaptability of p -norm distribution. Through estimating the parameter p , the results can be more reasonable. The corresponding character of Lp estimation is analyzed. It is proved that the traditional Lp estimation method is not a robust method. So a stable Lp estimation based on data depth weighting is put forward to deal with the model error and outlier. In orbit determination process, the outlier of observational data and coarse model error can be quantitatively described. The farther is the data from the data center, the smaller is the value of data depth and the smaller is the weighted value accordingly. The result of the new Lp method is stabler than that of the common Lp estimation and the break down point could be up to $1/2$. In addition, the orbit parameter is adaptively estimated by residual analysis and matrix estimation method, and the estimation efficiency is more perfect. Finally, taking the Space Based Space Surveillance System as an example to make simulation experiments, the results show that if there are system error or abnormal value in the observational data or system error in satellite dynamical model and space based observation platform, LSE will not be optimal even though the observational data obeys Gaussian distribution and the orbit determination precision by the self-adaptive stable Lp estimation method is much higher than that by the common LSE method.

Key words celestial mechanics: orbit calculation and determination, method: numerical