总体最小二乘法在初轨计算中的应用

陈务深 [1,2] 掌 静 [1] 马静远 [1] 陆本魁 [1]

(1 中国科学院紫金山天文台 南京 210093)
(2 南京工业大学理学院 南京 210009)

摘要 在分析了 UVM2 方法条件方程特点的基础上，将总体最小二乘法 (TLS) 引入轨道计算中，避免了原算法中对活力公式的线性化处理. 仿真和实测计算结果表明，将 TLS 应用于 UVM2 是有效的.

关键词 天体力学：初轨计算，单位矢量法，总体最小二乘法
中图分类号： P 135; **文献标识码**： A

1 引言

初始轨道计算，是人造地球卫星轨道测定的重要内容之一 [1]. 提高初始轨道测定精度，特别是轨道半长轴的测定精度，不仅有利于提高轨道改进的收敛性和收敛速度，而且对保证卫星下次过境预报精度起着关键性作用.

改进的无摄初轨的单位矢量法 (UVM2) 是一种既适用于长弧段 [2,3], 也适用于短弧段的有效方法. 其最大特点是将半长轴 a 同其他 5 个根数区别开来单独处理. 为此，需要将位置向量 $\boldsymbol{r}$ 和速度矢量 $\dot{\boldsymbol{r}}$ 中所包含的 a 分离出来，根据这种思想，就可以得到如下形式的条件方程 [2,3]:

$$\begin{cases} (f_1\cdot\boldsymbol{A}^*)\cdot\boldsymbol{r}'_0+(g_1\cdot\boldsymbol{A}^*)\cdot\dot{\boldsymbol{r}}'_0+\left(\boldsymbol{A}^*\cdot\dfrac{\partial\boldsymbol{r}_c}{\partial n}\right)\cdot\Delta n=\boldsymbol{A}^*\cdot\boldsymbol{R} \\ (f_1\cdot\boldsymbol{h}^*)\cdot\boldsymbol{r}'_0+(g_1\cdot\boldsymbol{h}^*)\cdot\dot{\boldsymbol{r}}'_0+\left(\boldsymbol{h}^*\cdot\dfrac{\partial\boldsymbol{r}_c}{\partial n}\right)\cdot\Delta n=\boldsymbol{h}^*\cdot\boldsymbol{R} \\ (f_1\cdot\boldsymbol{\rho}^*)\cdot\boldsymbol{r}'_0+(g_1\cdot\boldsymbol{\rho}^*)\cdot\dot{\boldsymbol{r}}'_0+\left(\boldsymbol{\rho}^*\cdot\dfrac{\partial\boldsymbol{r}_c}{\partial n}\right)\cdot\Delta n=\boldsymbol{\rho}^*\cdot\boldsymbol{R}+\rho \end{cases}, \tag{1}$$

$$(\dot{f}_1\cdot\boldsymbol{\rho}^*)\cdot\boldsymbol{r}'_0+(\dot{g}_1\cdot\boldsymbol{\rho}^*)\cdot\dot{\boldsymbol{r}}'_0+\left(\boldsymbol{\rho}^*\cdot\frac{\partial\dot{\boldsymbol{r}}_c}{\partial n}\right)\cdot\Delta n=\boldsymbol{\rho}^*\cdot\dot{\boldsymbol{R}}+\dot{\rho}\,. \tag{2}$$

解方程时，需要考虑活力积分约束：

$$\frac{2}{r'_0}-\|\dot{\boldsymbol{r}}'_0\|^2=1\,, \tag{3}$$

利用条件最小二乘法求解未知量 $\boldsymbol{r}'_0$, $\dot{\boldsymbol{r}}'_0$ 和 Δn, 其中： $\boldsymbol{r}'_0=\boldsymbol{r}_0/a$, $\dot{\boldsymbol{r}}'_0=\sqrt{a}\dot{\boldsymbol{r}}_0$.

2005-04-01 收到原稿， 2005-05-23 收到修改稿

在公式 (1) 、公式 (2) 的系数和等式的右端常数项中，单位矢量 $\boldsymbol{\rho}^*$ 、 $\boldsymbol{A}^*$ 和 $\boldsymbol{h}^*$ 是用测角资料计算得到的， ρ 和 $\dot{\rho}$ 分别对应于测距和测速资料.

由 UVM2 的原理可知，条件方程的前 6 个未知量 $\boldsymbol{r}_0'$ 和 $\dot{\boldsymbol{r}}_0'$ 是相互关联的，从本质上讲，只相当于 5 个相互独立的未知量. 不仅如此，在力学模型和观测资料都没有误差的理想情况下，所有的条件方程都严格成立. 也就是说，若用 M_c 表示由条件方程的系数和右端常数项所组成的列数为 8 的矩阵，则有 $\mathrm{rank}(M_c) = 6$. 而事实上，观测误差和力学模型误差是一定存在的，这使得 M_c 为列满秩矩阵，即 $\mathrm{rank}(M_c) = 8$, 也就是由于存在误差而掩盖了未知量的相关性. 在文 [2, 3] 的算法中，采用活力公式约束求解条件方程，并对活力公式进行了线性化处理. 线性化处理又带来了新的误差，在此情况下，本文引入了总体最小二乘法.

2 原理

对于任意一个复矩阵 $A_{m\times n}$, 若存在矩阵 $Y_{n\times m}$, 满足

$$\begin{aligned} AYA &= A \\ YAY &= Y \\ (AY)^H &= AY \\ (YA)^H &= YA \end{aligned}$$

则称矩阵 Y 为矩阵 A 的一个 Moore-Penrose 广义逆矩阵，简称 M-P 逆. 可以证明，同时满足以上 4 个条件的 M-P 逆唯一存在，通常记为 A^+.

M-P 逆实际上是逆矩阵概念的推广，它在与线性最小二乘方法的相关理论中有着非常广泛的应用. 假设 $A \in R^{m\times n}$, $B \in R^m$, 那么，对于如下的最小二乘问题

$$\min \|Ax - B\|_2^2$$

其通解就为：

$$X = A^+B + (I - A^+A)Y$$

其中： $X = A^+B$ 为最小 $2-$ 范数解.

总体最小二乘问题是 1980 年由 Golub 等 [4] 首先提出的，近 20 年来，人们已经对总体最小二乘解的算法以及解的形式、总体最小二乘 (TLS) 解与最小二乘 (LS) 解之间的关系、总体最小二乘解的扰动理论以及数值实验做了许多工作.

假设 M 是 $m \times p$ 的矩阵， $\|\cdot\|$ 为酉不变范数，我们希望给出一个 $m \times p$ 的矩阵 $\hat{M}$, 使得

$$\mathrm{rank}(\hat{M}) \le r$$

$$\| \hat{M} - M \| = \inf_{\mathrm{rank}[\tilde{M}]\le r} \| \tilde{M} - M \|$$

在 1936 年，Eckart 等 [5] 就给出了该问题在 Frobenius 范数意义下的最优解. 后来 Mirsky[6] 给出了在任意酉不变范数意义下的最优解，它可以根据矩阵 M 的奇异值分解被简单地

构造出来. 假设:

$$M = U\,\Psi\,V^T\,,$$

其中: $UU^T = V^TV = I$, $\Psi = \mathrm{diag}(\psi_1, \psi_2, \cdots, \psi_n)$ 且 $\psi_1 \geq \psi_2 \geq \cdots \geq \psi_n \geq 0$.

若记 $\hat{\Psi} = \mathrm{diag}(\psi_1, \psi_2, \cdots, \psi_r, 0, \cdots 0)$, 则

$$\hat{M} = U\ \hat{\Psi} V^T$$

就为矩阵 M 的秩为 r 的最佳近似矩阵. 如果 $\mathrm{rank}(M) \geq r$, 那么, 在 Frobenius 范数意义下, 当且仅当 $\psi_r > \psi_{r+1}$ 时, 矩阵 $\hat{M}$ 唯一存在. 这就是 Eckart-Young-Mirsky 定理.

设 $A \in R^{m\times n}$, $B \in R^{m\times d}$, 考虑线性方程组

$$A\,X \simeq B\,, \tag{4}$$

其中: $X \in R^{n\times d}$ 为未知量. 利用以上方法, 转换处理成总体最小二乘 (TLS) 方法, 其通常定义如下 [4,7]: 对于某个在 1 和 n 间的正整数 p, 求修正矩阵 $(\Delta A, \Delta B)$, 使得

$$\min \|(\Delta A, \Delta B)\|_F\,,$$
$$\mathrm{s.t.}\,\mathrm{rank}[\bar{A}] = \mathrm{rank}[(\bar{A}, \bar{B})] = p\,,$$

其中: $(\bar{A}, \bar{B}) = (A, B) - (\Delta A, \Delta B)$. 此时, 线性方程组

$$\bar{A}X = \bar{B} \tag{5}$$

是相容方程组, 其解就为方程组 (4) 的总体最小二乘解.

如果矩阵 (A, B) 的奇异值分解为 $(A, B) = U\Sigma V^T$, 其中: $U \in R^{m\times m}$, $V \in R^{(n+d)\times(n+d)}$ 为正交矩阵, $\Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_p, \sigma_{p+1}, \cdots, \sigma_{n+d})$, 则

$$(\bar{A}, \bar{B}) = U\bar{\Sigma}V^T\,,$$

其中: $\bar{\Sigma} = \mathrm{diag}(\sigma_1, \cdots, \sigma_p, 0, \cdots, 0)$, 且有

$$\min \|(\Delta A, \Delta B)\|_F = \sqrt{\sum_{i>p} \sigma_i^2}\,.$$

对矩阵 V 作如下划分

$$V = \begin{array}{c} \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \begin{array}{c} n \\ d \end{array} \\ \begin{array}{cc} p & n+d-p \end{array} \end{array},$$

则线性方程组 (5) 的通解为

$$X = \bar{A}^+\bar{B} + (I - \bar{A}^+\bar{A})Y = -V_{22}^+V_{12} + (I - (V_{11}V_{11}^+)^T)Y\,,$$

其中: $I \in R^{n\times n}$ 为单位矩阵, $Y \in R^{n\times d}$ 为任意矩阵.

3　应用

在实际计算的迭代过程中，利用各组观测资料，根据 (1) 式和 (2) 式形成许多条件方程，假设其对应的增广矩阵为 M_c, 为了节省存储空间，就不直接形成矩阵 M_c, 而是利用 Givens 旋转将矩阵 M_c 的各行压缩至 8×8 的矩阵 C 之中 [7], 这样就得到了压缩条件方程组

$$C\cdot \boldsymbol{x}\simeq \boldsymbol{0}\,,$$

其中：　$\boldsymbol{x}=[(\boldsymbol{r}_0')^T,(\dot{\boldsymbol{r}}_0')^T,\Delta n,-1]^T$.

对矩阵 C 进行奇异值分解：

$$C=U\Sigma V^T\,,$$

其中：　$\Sigma=\mathrm{diag}(\sigma_1,\cdots,\sigma_8)$ 且 $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_7\geq\sigma_8$, 对矩阵 V 作如下划分

$$V=\begin{pmatrix} V_{11} & V_{12}\\ V_{21} & V_{22}\end{pmatrix}\begin{matrix}7\\1\end{matrix}\,,\qquad \begin{matrix}6 & 2\end{matrix}$$

当 $V_{22}\neq\boldsymbol{0}$ 时，条件方程的总体最小二乘问题的通解为：

$$\boldsymbol{x}=-V_{22}^{+}V_{12}+(I-(V_{11}V_{11}^{+})^T)\boldsymbol{y}\,,$$

其中：　$\boldsymbol{y}$ 为任意矢量.

在实际的处理过程中，可以求旋转矩阵 $G(\theta)$, 使得

$$\begin{pmatrix} V_{12}\\ V_{22}\end{pmatrix}G(\theta)=\begin{pmatrix} \boldsymbol{s}_0 & \boldsymbol{s}_1\\ u & 0\end{pmatrix}\begin{matrix}7\\1\end{matrix}\,,\qquad \begin{matrix}1 & 1\end{matrix}$$

则总体最小二乘问题的通解具有以下形式

$$\boldsymbol{x}=-\frac{1}{u}\boldsymbol{s}_0+\lambda\boldsymbol{s}_1\,,\tag{6}$$

其中：　λ 为待定参数.

将 (6) 式代入活力积分约束方程 (3), 就形成了以 λ 为参数的一个非线性方程，选择合适的初始值，使用牛顿迭代法定出 λ, 就可以得到满足条件方程的总体最小二乘解 $\boldsymbol{x}$.

4　仿真

根据上面的思想，我们基于文 [2,3] 方法和 TLS 方法，对 UVM2 进行了仿真计算. 采用了每秒一组共 10 分钟的四元素仿真数据 (考虑了一阶摄动), 用两种不同的 UVM2 仿真计算 150 次，在仿真过程中，测角资料误差为 0.01°, 测距误差为 5 m, 测速误差为每秒 1 m. 表 1, 表 2 列出了半长轴 a 的均方差 σ_a.

表 1 小偏心率 ($e < 0.1$) 仿真结果: σ_a(km)

Table 1 σ_a(km) of Simulation results with $e < 0.1$

Method	orbit I	orbit II	orbit III
Paper [2,3]	0.5516E+01	0.3845E+01	0.2715E+01
TLS	0.5323E+01	0.3471E+01	0.2358E+01

表 2 大偏心率 ($e > 0.7$) 仿真结果: σ_a(km)

Table 2 σ_a(km) of Simulation results with $e > 0.7$

Method	orbit I	orbit II	orbit III
Paper [2,3]	0.502249E+01	0.485396E+01	0.485018E+01
TLS	0.457024E+01	0.425251E+01	0.424465E+01

5 实例

我们采用中国卫星海上测控部实测的四元素资料进行了同样的计算. 测距精度为 10m, 测角精度为 $43''.2$, 测速精度为 0.02m/s, 对比标准采用多站多圈的轨道改进结果.

表 3 小偏心率 ($e < 0.1$) 实例, 弧段为 3^m20^s

Table 3 results of real data with $e < 0.1$, arc length 3^m20^s

Method	Δa(km)	Δe	Δi(deg.)	$\Delta\omega$(deg.)	$\Delta\Omega$(deg.)	$\Delta\lambda$(deg.)
Paper[2, 3]	-1.161	7.740E-5	-3.023E-2	8.855	-6.396E-2	4.568E-2
TLS	-1.134	5.491E-5	-3.017E-2	8.852	-6.390E-2	4.827E-2

表 4 大偏心率 ($e > 0.7$) 实例, 弧段为 7^h35^m

Table 4 results of real data with $e > 0.7$, arc length 7^h35^m

Method	Δa(km)	Δe	Δi(deg.)	$\Delta\omega$(deg.)	$\Delta\Omega$(deg.)	$\Delta\lambda$(deg.)
Paper[2, 3]	-8.60868	1.65179E-4	-6.498E-2	-3.46339E-2	-6.14054E-2	0.714361
TLS	-8.60585	1.65121E-4	-6.498E-2	-3.46353E-2	-6.14504E-2	0.714364

6 分析

由上述计算结果可见, 将总体最小二乘法 (TLS) 应用到 UVM2 的初轨计算中是有效的. 其方法的精度与文 [2,3] 大致相当, 且半长轴的测定精度有所提高. 它的主要特点是避免了文 [2,3] 解方程时必须对约束方程进行线性化处理, 从而改善了方法的收敛性, 放宽了对初值的要求.

参 考 文 献

1 易照华. 天体力学引论. 上海：科学出版社， 1978

2 LU Ben-Kui, MA Jing-Yuan, XIA Yi, et.al. ChA&A, 2004, 28: 88

3 LU Ben-Kui, MA Jing-Yuan, XIA Yi, et.al. AcASn, 2003, 44(4): 369 (陆本魁，马静远，夏益等. 天文学报， 2003, 44(4): 369)

4 Golub G H, Van Loan C F. SIAM J Numer Anal, 1980, 17: 883

5 Eckart G, Young G. The approximation of one matrix by another of lower rank, Psychometrika, 1936, 1: 211

6 Mirsky L. Quart J Math Oxford, 1960, 11: 50

7 Golub G H, Van Loan C F. Matrix Computations[M], The Johns Hopkings University Press, 1989

On the Application of Total Least Squares Method in Preliminary Orbit Determination

CHEN Wu-shen[1,2] ZHANG Jing[1] MA Jing-yuan[1] LU Ben-kui[1]

(1 *Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing* 210008)

(2 *School of Sciences, Nanjing University of Technology, Nanjing* 210009)

Abstract Based on the mathematical model of a nonlinear least squares problem with an equality constrain, the UVM2 method is introduced in preliminary orbit determination. It can be inferred from the principle of UVM2 that the 6 variables of $\boldsymbol{r}'_0$ and $\dot{\boldsymbol{r}}'_0$ in the condition equations are essentially 5 dependent ones. To counter this characteristic of UVM2, the Total Least Squares method is introduced in preliminary orbit determination. It is shown that this method is effective from the simulation results and the real ones.

Key words celestial mechanics: orbit determination, Unit Vector Method, Total Least Squares