第 48 卷 第 3 期 天 文 学 报 Vol.48 No.3
2007 年 7 月 ACTA ASTRONOMICA SINICA Jul., 2007

单位矢量法的数学模型 (MMUVM) 及其简化形式 (PUVM1) 的迭代法的收敛性

陈务深 [1,2†] 陆本魁 [1] 马静远 [1]

(1 中国科学院紫金山天文台 南京 210008)
(2 南京工业大学理学院 南京 210009)

摘要 利用人造地球卫星观测资料测定有摄初轨的单位矢量法 (PUVM1), 已得到了非常广泛地实际应用. 为了对单位矢量法作进一步地完善和改进, 首先在考虑测量误差模型的基础上, 建立单位矢量法所对应的数学模型 MMUVM. 它本质上就是一个非线性最优化问题. 针对 MMUVM, 先分别使用多圈仿真数据和实测数据, 形成了与之相对应的目标函数, 再利用求解最优化问题的一种三对角二次插值模型的直接搜索方法, 分别对其进行了数值处理. 计算结果表明, 所建立的优化模型 MMUVM 正确合理, 所采用的直接搜索方法实用有效. 其次, 进一步指明了 PUVM1 和 MMUVM 之间关系, 即: 从本质上讲, PUVM1 就是 MMUVM 的一种简化形式. 从数学原理上, 清楚地解释了利用 PUVM1 的准法化方程, 只能使用单圈短弧段数据进行初始轨道确定, 而不能使用长弧段多圈资料进行轨道确定或轨道改进的根本原因. 最后, 对 PUVM1 的迭代算法的收敛性问题进行了初步的理论分析, 并给出了相应的数值验证实例, 指出了 PUVM1 的迭代格式是条件收敛的, 即: 只有在满足一定条件后, 才能收敛. 这也就意味着: 有的时候, 尽管准法化方程是合情合理的, 但是, 此时该迭代法却是发散的, 无法迭代求出所要的解.

关键词 天体力学: 轨道计算和定轨, 天体力学: 数值方法
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. 努力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计算速度, 而且对保证卫星下次过境预报精度起着关键性的作用.

文 [1] 提出了有摄初轨计算的单位矢量法 (PUVM1). 就计算收敛率和稳定性而言, PUVM1 既高于常用的现有各种初轨计算方法, 也不低于微分轨道改进. 又由于它避免了微分轨道改进要对所有轨道根数求有摄偏导数的繁复计算, 故 PUVM1 与微分轨道改进相比, 可以提高计算速度, 特别是摄动计算采用数值积分方法时, 计算速度的提高尤其明显. 实际使用结果也表明, 该方法具有定轨精度高、适用范围广、稳定性好、收敛速度快等显著优点.

2006-02-20 收到原稿, 2006-06-02 收到修改稿
†chenwushen@163.com

但是，该方法还存在一些缺陷．它缺少精确的数学模型，所给出的迭代法仅仅适用于单圈短弧段情况，而无法处理多圈资料的情况．本文针对上述问题，建立了单位矢量法新的数学模型，对原有方法作了进一步完善和改进.

2 单位矢量法的非线性条件方程

在不考虑力学模型和观测误差的理想情况下，可以得到严格成立的单位矢量法的条件方程，其一般形式如下 [1]:

$$\left\{\begin{array}{lcl}(f\,A^*)\cdot r_0+(g\,A^*)\cdot\dot r_0+(w\,A^*)\cdot W_0 &=& A^*\cdot R\\ (f\,h^*)\cdot r_0+(g\,h^*)\cdot\dot r_0+(w\,h^*)\cdot W_0 &=& h^*\cdot R\\ (f\,\rho^*)\cdot r_0+(g\,\rho^*)\cdot\dot r_0+(w\,\rho^*)\cdot W_0 &=& \rho^*\cdot R+\rho\\ (\dot f\,\rho^*)\cdot r_0+(\dot g\,\rho^*)\cdot\dot r_0+(\dot w\,\rho^*)\cdot W_0 &=& \rho^*\cdot\dot R+\dot\rho\end{array}\right.,\tag{1}$$

或

$$\left\{\begin{array}{l}\delta_A=(f\,\boldsymbol{r}_0+g\,\dot{\boldsymbol{r}}_0+w\,\boldsymbol{W}-\boldsymbol{R})\cdot\boldsymbol{A}^*=0\\ \delta_h=(f\,\boldsymbol{r}_0+g\,\dot{\boldsymbol{r}}_0+w\,\boldsymbol{W}-\boldsymbol{R})\cdot\boldsymbol{h}^*=0\\ \delta_\rho=(f\,\boldsymbol{r}_0+g\,\dot{\boldsymbol{r}}_0+w\,\boldsymbol{W}-\boldsymbol{R})\cdot\boldsymbol{\rho}^*-\rho=0\\ \delta_{\dot\rho}=\left(\dot f\,\boldsymbol{r}_0+\dot g\,\dot{\boldsymbol{r}}_0+\dot w\,\boldsymbol{W}-\dot{\boldsymbol{R}}\right)\cdot\boldsymbol{\rho}^*-\dot\rho=0\end{array}\right..\tag{2}$$

其中：

$$\left\{\begin{array}{lcl}f &=& \boldsymbol{F_0}\cdot\boldsymbol{r(t,r_0,\dot r_0)}\\ g &=& \boldsymbol{G_0}\cdot\boldsymbol{r(t,r_0,\dot r_0)}\\ w &=& \boldsymbol{W_0}\cdot\boldsymbol{r(t,r_0,\dot r_0)}\end{array}\right.,$$

$$\left\{\begin{array}{lcl}\dot f &=& F_0\cdot\dot r(t,r_0,\dot r_0)\\ \dot g &=& G_0\cdot\dot r(t,r_0,\dot r_0)\\ \dot w &=& W_0\cdot\dot r(t,r_0,\dot r_0)\end{array}\right.,$$

$$\left\{\begin{array}{lcl}F_0 &=& -p_0^{-1/2}(W_0\times\dot r_0)\\ G_0 &=& p_0^{-1/2}(W_0\times r_0)\\ W_0 &=& p_0^{-1/2}(r_0\times\dot r_0)\\ p_0 &=& \|r_0\times\dot r_0\|_2^2\end{array}\right.,$$

$$\left\{\begin{array}{lcccccc}\rho^* &=& \cos h\cos A\,\varphi^* &+& \cos h\sin A\,S^* &+& \sin h\,R^*\\ A^* &=& -\sin A\,\varphi^* &+& \cos A\,S^* & & \\ h^* &=& -\sin h\cos A\,\varphi^* &-& \sin h\sin A\,S^* &+& \cos h\,R^*\end{array}\right.,$$

$$\left\{\begin{array}{lcccccc}R^* &=& \cos\varphi\cos S\,i &+& \cos\varphi\sin S\,j &+& \sin\varphi\,k\\ S^* &=& -\sin S\,i &+& \cos S\,j & & \\ \varphi^* &=& -\sin\varphi\cos S\,i &-& \sin\varphi\sin S\,j &+& \cos\varphi\,k\end{array}\right..$$

符号 $\|\cdot\|_2$ 在下文中将继续出现，它表示一个矢量的 2 范数，即对于一个 n 维矢量 $\boldsymbol{s}$, 有：

$$\|\boldsymbol{s}\|_2 = \sqrt{\boldsymbol{s}\cdot\boldsymbol{s}}\,.$$

从上面的式子可以看出，四类条件方程是关于 $\boldsymbol{r}_0$ 和 $\dot{\boldsymbol{r}}_0$ 的非线性方程．其中与测量数据有关的量为：测距资料 ρ, 测速资料 $\dot{\rho}$ 和由测角资料形成的单位矢量 $\boldsymbol{\rho}^*$ 、 $\boldsymbol{A}^*$ 和 $\boldsymbol{h}^*$.

3　单位矢量法的条件方程加权方案及其相应的数学模型

3.1　考虑观测资料误差情况下单位矢量法条件方程的加权方案

在实际应用中，观测数据必然存在观测误差．下面，在不考虑力学模型误差而仅考虑观测资料误差的前提下，给出了单位矢量法条件方程的加权方案.

为了便于表述，记：

$$\boldsymbol{c} = (A,\, h,\, \rho,\, \dot{\rho})^T\,,$$

假设 $\boldsymbol{c}$ 的测量误差

$$\Delta\boldsymbol{c} = (\Delta A,\, \Delta h,\, \Delta\rho,\, \Delta\dot{\rho})^T$$

为服从正态分布

$$\Delta\boldsymbol{c} \sim N(\mu_{\boldsymbol{c}}, \Sigma)$$

的随机变量．其中 $\boldsymbol{\mu_c} = (\mu_A, \mu_h, \mu_\rho, \mu_{\dot{\rho}})^T$ 为期望值， Σ 为协方差矩阵．因此，条件方程的残差也为随机变量

$$\begin{cases} \varepsilon_A = \boldsymbol{A}^*\cdot\boldsymbol{\eta}_1 \\ \varepsilon_h = \boldsymbol{h}^*\cdot\boldsymbol{\eta}_1 \\ \varepsilon_\rho = \boldsymbol{\rho}^*\cdot\boldsymbol{\eta}_1 - \rho \\ \varepsilon_{\dot{\rho}} = \boldsymbol{\rho}^*\cdot\boldsymbol{\eta}_2 - \dot{\rho} \end{cases}, \tag{3}$$

其中：

$$\begin{cases} \boldsymbol{\eta}_1 = f\boldsymbol{r}_0 + g\dot{\boldsymbol{r}}_0 + w\boldsymbol{W}_0 - \boldsymbol{R} \\ \boldsymbol{\eta}_2 = \dot{f}\boldsymbol{r}_0 + \dot{g}\dot{\boldsymbol{r}}_0 + \dot{w}\boldsymbol{W}_0 - \dot{\boldsymbol{R}} \end{cases}.$$

容易看出，残差 (3) 是四类观测资料 ρ 、 $\dot{\rho}$ 、 A 和 h 的非线性函数，如果将其在精确值点处线性展开，则可得期望值和方差的近似关系式 [2].

因此，残差 (3) 的期望值为：

$$\begin{cases} \mu_{\varepsilon_A} = \left(\dfrac{\partial\varepsilon_A}{\partial\rho}\right)\mu_\rho + \left(\dfrac{\partial\varepsilon_A}{\partial h}\right)\mu_h + \left(\dfrac{\partial\varepsilon_A}{\partial A}\right)\mu_A \\ \mu_{\varepsilon_h} = \left(\dfrac{\partial\varepsilon_h}{\partial\rho}\right)\mu_\rho + \left(\dfrac{\partial\varepsilon_h}{\partial h}\right)\mu_h + \left(\dfrac{\partial\varepsilon_h}{\partial A}\right)\mu_A \\ \mu_{\varepsilon_\rho} = \left(\dfrac{\partial\varepsilon_\rho}{\partial\rho}\right)\mu_\rho + \left(\dfrac{\partial\varepsilon_\rho}{\partial h}\right)\mu_h + \left(\dfrac{\partial\varepsilon_\rho}{\partial A}\right)\mu_A \\ \mu_{\varepsilon_{\dot{\rho}}} = \left(\dfrac{\partial\varepsilon_{\dot{\rho}}}{\partial\rho}\right)\mu_\rho + \left(\dfrac{\partial\varepsilon_{\dot{\rho}}}{\partial h}\right)\mu_h + \left(\dfrac{\partial\varepsilon_{\dot{\rho}}}{\partial A}\right)\mu_A + \left(\dfrac{\partial\varepsilon_{\dot{\rho}}}{\partial\dot{\rho}}\right)\mu_{\dot{\rho}} \end{cases}, \tag{4}$$

其中：

$$\begin{cases} \left(\dfrac{\partial \varepsilon_A}{\partial A}\right) = \left(\dfrac{\partial \boldsymbol{A}^*}{\partial A}\right)\cdot\boldsymbol{\eta}_1 \\ \left(\dfrac{\partial \varepsilon_A}{\partial h}\right) = 0 \\ \left(\dfrac{\partial \varepsilon_A}{\partial \rho}\right) = 0 \\ \left(\dfrac{\partial \varepsilon_A}{\partial \dot{\rho}}\right) = 0 \end{cases},$$

$$\begin{cases} \left(\dfrac{\partial \varepsilon_h}{\partial A}\right) = \left(\dfrac{\partial \boldsymbol{h}^*}{\partial A}\right)\cdot\boldsymbol{\eta}_1 \\ \left(\dfrac{\partial \varepsilon_h}{\partial h}\right) = \left(\dfrac{\partial \boldsymbol{h}^*}{\partial h}\right)\cdot\boldsymbol{\eta}_1 \\ \left(\dfrac{\partial \varepsilon_h}{\partial \rho}\right) = 0 \\ \left(\dfrac{\partial \varepsilon_h}{\partial \dot{\rho}}\right) = 0 \end{cases},$$

$$\begin{cases} \left(\dfrac{\partial \varepsilon_\rho}{\partial A}\right) = \left(\dfrac{\partial \boldsymbol{\rho}^*}{\partial A}\right)\cdot\boldsymbol{\eta}_1 \\ \left(\dfrac{\partial \varepsilon_\rho}{\partial h}\right) = \left(\dfrac{\partial \boldsymbol{\rho}^*}{\partial h}\right)\cdot\boldsymbol{\eta}_1 \\ \left(\dfrac{\partial \varepsilon_\rho}{\partial \rho}\right) = -1 \\ \left(\dfrac{\partial \varepsilon_\rho}{\partial \dot{\rho}}\right) = 0 \end{cases},$$

$$\begin{cases} \left(\dfrac{\partial \varepsilon_{\dot{\rho}}}{\partial A}\right) = \left(\dfrac{\partial \boldsymbol{\rho}^*}{\partial A}\right)\cdot\boldsymbol{\eta}_2 \\ \left(\dfrac{\partial \varepsilon_{\dot{\rho}}}{\partial h}\right) = \left(\dfrac{\partial \boldsymbol{\rho}^*}{\partial h}\right)\cdot\boldsymbol{\eta}_2 \\ \left(\dfrac{\partial \varepsilon_{\dot{\rho}}}{\partial \rho}\right) = 0 \\ \left(\dfrac{\partial \varepsilon_{\dot{\rho}}}{\partial \dot{\rho}}\right) = -1 \end{cases},$$

$$\begin{cases} \dfrac{\partial \boldsymbol{\rho}^*}{\partial A} = -\sin A\cos h\,\boldsymbol{\varphi}^* + \cos A\cos h\,\boldsymbol{S}^* = \cos h\,\boldsymbol{A}^* \\ \dfrac{\partial \boldsymbol{\rho}^*}{\partial h} = -\cos A\sin h\,\boldsymbol{\varphi}^* - \sin A\sin h\,\boldsymbol{S}^* + \cos h\,\boldsymbol{R}^* = \boldsymbol{h}^* \end{cases},$$

$$\begin{cases} \dfrac{\partial \boldsymbol{h}^*}{\partial A} = \sin A\sin h\,\boldsymbol{\varphi}^* - \cos A\sin h\,\boldsymbol{S}^* = -\sin h\,\boldsymbol{A}^* \\ \dfrac{\partial \boldsymbol{h}^*}{\partial h} = -\cos A\cos h\,\boldsymbol{\varphi}^* - \sin A\cos h\,\boldsymbol{S}^* - \sin h\,\boldsymbol{R}^* = -\boldsymbol{\rho}^* \end{cases},$$

$$\begin{cases} \dfrac{\partial \boldsymbol{A}^*}{\partial A} = -\cos A\,\boldsymbol{\varphi}^* - \sin A\,\boldsymbol{S}^* = -\cos h\,\boldsymbol{\rho}^* + \sin h\,\boldsymbol{h}^* \\ \dfrac{\partial \boldsymbol{A}^*}{\partial h} = \boldsymbol{0} \end{cases}.$$

残差 (3) 的方差为

$$\begin{cases} \sigma_{\varepsilon_A}^2 = \xi_{\varepsilon_A}^T \cdot \Sigma \cdot \xi_{\varepsilon_A} \\ \sigma_{\varepsilon_h}^2 = \xi_{\varepsilon_h}^T \cdot \Sigma \cdot \xi_{\varepsilon_h} \\ \sigma_{\varepsilon_\rho}^2 = \xi_{\varepsilon_\rho}^T \cdot \Sigma \cdot \xi_{\varepsilon_\rho} \\ \sigma_{\varepsilon_{\dot\rho}}^2 = \xi_{\varepsilon_{\dot\rho}}^T \cdot \Sigma \cdot \xi_{\varepsilon_{\dot\rho}} \end{cases}, \tag{5}$$

其中：

$$\begin{cases} \xi_{\varepsilon_A} = \left(\dfrac{\partial \varepsilon_A}{\partial A}, \dfrac{\partial \varepsilon_A}{\partial h}, \dfrac{\partial \varepsilon_A}{\partial \rho}, \dfrac{\partial \varepsilon_A}{\partial \dot\rho} \right)^T \\ \xi_{\varepsilon_h} = \left(\dfrac{\partial \varepsilon_h}{\partial A}, \dfrac{\partial \varepsilon_h}{\partial h}, \dfrac{\partial \varepsilon_h}{\partial \rho}, \dfrac{\partial \varepsilon_h}{\partial \dot\rho} \right)^T \\ \xi_{\varepsilon_\rho} = \left(\dfrac{\partial \varepsilon_\rho}{\partial A}, \dfrac{\partial \varepsilon_\rho}{\partial h}, \dfrac{\partial \varepsilon_\rho}{\partial \rho}, \dfrac{\partial \varepsilon_\rho}{\partial \dot\rho} \right)^T \\ \xi_{\varepsilon_{\dot\rho}} = \left(\dfrac{\partial \varepsilon_{\dot\rho}}{\partial A}, \dfrac{\partial \varepsilon_{\dot\rho}}{\partial h}, \dfrac{\partial \varepsilon_{\dot\rho}}{\partial \rho}, \dfrac{\partial \varepsilon_{\dot\rho}}{\partial \dot\rho} \right)^T \end{cases}.$$

将条件方程的残差标准化处理，得到服从标准正态分布的随机变量

$$\begin{cases} \epsilon_A = \dfrac{\varepsilon_A - \mu_{\varepsilon_A}}{\sigma_{\varepsilon_A}} \\ \epsilon_h = \dfrac{\varepsilon_h - \mu_{\varepsilon_h}}{\sigma_{\varepsilon_h}} \\ \epsilon_\rho = \dfrac{\varepsilon_\rho - \mu_{\varepsilon_\rho}}{\sigma_{\varepsilon_\rho}} \\ \epsilon_{\dot\rho} = \dfrac{\varepsilon_{\dot\rho} - \mu_{\varepsilon_{\dot\rho}}}{\sigma_{\varepsilon_{\dot\rho}}} \end{cases}. \tag{6}$$

利用观测资料，根据标准化的条件方程残差，使用最小二乘法，就可以得到以 $\boldsymbol{r}_0$ 、$\dot{\boldsymbol{r}}_0$ 为自变量的函数

$$G(\boldsymbol{r}_0, \dot{\boldsymbol{r}}_0) = \sum_t \left(\epsilon_A^2 + \epsilon_h^2 + \epsilon_\rho^2 + \epsilon_{\dot\rho}^2 \right)_t . \tag{7}$$

它是一个服从卡方分布的随机变量.

3.2 单位矢量法的数学模型

综合以上可知，单位矢量法的数学模型为：

$$\min_{\boldsymbol{r}_0, \dot{\boldsymbol{r}}_0} G(\boldsymbol{r}_0, \dot{\boldsymbol{r}}_0) . \tag{8}$$

该最优化问题的最优解就为轨道预报或轨道改进的结果. 为了方便表述，在下文中，用 MMUVM 表示该模型.

4 求解最优化问题的三对角二次插值模型的直接搜索算法

无约束最优化问题的标准数学模型为

$$\min_{x \in R^n} f(\boldsymbol{x}) ,$$

其中 $f: R^n \to R$ 是光滑非线性函数.

针对此类最优化问题，常见的处理方法有两大类 [3]: 一类是用到偏导数的梯度法，另一类是不用偏导数的直接方法. 直接搜索方法是只使用目标函数值，而不使用目标函数导数值的一种最优化方法. 3 对角二次插值模型算法是一种比较有效的直接搜索方法 [4,5], 它在存储、运行速度、函数值计算次数、迭代次数、运行时间等方面都具有明显的优越性.

众所周知，若将目标函数在极小值点展开，则目标函数可以用一个二次型来进行近似表示. 而如果知道二次型的一些函数值, 则在一定的条件下, 可利用它们来确定该二次型. 二次插值模型就是基于这种事实构造出来的直接算法，该算法通过构造并求解一系列的二次型，最终得到问题的最优解. 为了减少计算量，简化所构造的二次型的形式，即限定二次项系数矩阵为 3 对角对称矩阵. 此矩阵使用海森矩阵的主要部分，忽略非主要部分. 基于此方法，就可以构造出 3 对角二次插值模型的直接搜索算法.

在下文的数值计算中，对于优化模型 MMUVM, 就使用了 3 对角二次插值模型的直接搜索算法进行处理.

5 仿真资料和实测资料的计算结果

利用仿真数据和实测数据，分别形成了相应的 MMUVM, 然后，使用求解最优化问题的直接方法进行求解，可得如下的数值仿真结果.

5.1 仿真资料的计算结果

仿真次数为 100, 仿真资料的精度：测角资料精度为 $43''.2$, 测距资料的精度为 10 m, 测速资料的精度为每秒 0.02 m. 详细结果见表 1 和表 2:

表 1 偏心率为 $e = 0.000500409$, 每秒 1 组, 共 3 圈 900 组仿真资料

Table 1 With $e = 0.000500409$, a set of measurement per second, 3 arcs in 3 circles, 900 seconds totally

Type	σ_a(m)	σ_i(deg.)	σ_Ω(deg.)	σ_ξ	σ_η	σ_λ(deg.)
A,h,ρ,$\dot\rho$	7.889	0.001329	0.000735	0.000007	0.000016	0.001570
A,h,ρ	12.213	0.002525	0.001221	0.000014	0.000031	0.002653
A,h,$\dot\rho$	12.385	0.003346	0.001354	0.000015	0.000034	0.002604
A,h	23.345	0.005668	0.001855	0.000018	0.000063	0.004803

表 2 偏心率为 $e = 0.646001$, 每 10 秒 1 组, 共 3 圈 3756 组仿真数据

Table 2 With $e = 0.646001$, 6 set of measurement per minute, 3 arcs in 3 circles, 18780 seconds 3756 sets totally

Type	σ_a(m)	σ_e	σ_i(deg.)	σ_Ω(deg.)	σ_ω(deg.)	σ_M(deg.)
A,h,ρ,$\dot\rho$	28.841	0.000019	0.003836	0.010150	0.007806	0.000700
A,h,ρ	39.662	0.000014	0.003966	0.009092	0.007755	0.000609
A,h,$\dot\rho$	44.311	0.000007	0.000568	0.000706	0.000655	0.000526
A,h	44.339	0.000007	0.000565	0.000697	0.000642	0.000517

5.2 实测资料的计算结果 (见表 3 和表 4)

表 3　偏心率为 $e = 0.000255$. 每秒 1 组，共 5 圈 2197 组四元素实测资料

Table 3　with $e = 0.000255$, a set of measurement per second, 5 arcs in 5 circles, 2197 sets totally

Type	Δ_a(m)	Δ_i(deg.)	Δ_Ω(deg.)	Δ_ξ	Δ_η	Δ_λ(deg.)
A,h,ρ,$\dot{\rho}$	106.138	0.00265	-0.05721	0.0001692	-0.000989	0.0939
A,h,ρ	106.292	0.00287	-0.06279	0.0001940	-0.000992	0.0991
A,h,$\dot{\rho}$	-176.098	0.07973	-0.02077	0.0000293	-0.001485	0.0953
A,h	-203.838	0.08674	-0.02016	0.0000305	-0.001532	0.0981

表 4　偏心率为 $e = 0.725$, 共 8 小时 15 分钟，1900 组四元素实测资料

Table 4　with $e = 0.725$, 1900 set of measurement, 8 hours 15 minutes totally

Type	Δ_a(m)	Δ_e	Δ_i(deg.)	Δ_Ω(deg.)	Δ_ω(deg.)	Δ_M(deg.)
A,h,ρ,$\dot{\rho}$	-108.83	0.0000703	-0.00164	0.00095	0.00016	-0.0136
A,h,ρ	-303.91	0.0000389	0.00081	-0.00064	-0.00012	-0.0035
A,h,$\dot{\rho}$	119.64	-0.0000122	0.00226	-0.00132	0.00138	0.0010
A,h	387.90	-0.0000263	0.00022	-0.00035	0.00134	0.0001

5.3 结果分析

以上的数值计算结果表明，使用 MMUVM 方法可以解决原先 PUVM1 方法无法处理的多圈数据资料，并可达到满意的精度.

6　MMUVM 与 PUVM1 方法之间的关系

由 (1) 式可见，利用观测资料所形成的条件方程具有如下形式：

$$\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} \simeq \boldsymbol{b}(\boldsymbol{x})\,,$$

相应地， MMUVM 的目标函数 (7) 式也可表示为：

$$\boldsymbol{G}(\boldsymbol{x}) = [\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} - \boldsymbol{b}(\boldsymbol{x})]^T[\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} - \boldsymbol{b}(\boldsymbol{x})]\,.$$

对 $\boldsymbol{x}$ 求偏导数，可以得到：

$$\frac{\partial \boldsymbol{G}(\boldsymbol{x})}{\partial \boldsymbol{x}} = 2\frac{\partial[\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} - \boldsymbol{b}(\boldsymbol{x})]^T}{\partial \boldsymbol{x}}[\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} - \boldsymbol{b}(\boldsymbol{x})]\,,$$

$$\frac{\partial \boldsymbol{G}(\boldsymbol{x})}{\partial \boldsymbol{x}} = 2[\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} - \boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x})] + 2[\frac{\partial \boldsymbol{A}(\boldsymbol{x})}{\partial \boldsymbol{x}}\boldsymbol{x} - \frac{\partial \boldsymbol{b}(\boldsymbol{x})}{\partial \boldsymbol{x}}]^T\,[\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} - \boldsymbol{b}(\boldsymbol{x})]\,,$$

所以，只当且仅当：

$$\boldsymbol{\tau} = [\frac{\partial \boldsymbol{A}(\boldsymbol{x})}{\partial \boldsymbol{x}}\boldsymbol{x} - \frac{\partial \boldsymbol{b}(\boldsymbol{x})}{\partial \boldsymbol{x}}]^T\,[\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} - \boldsymbol{b}(\boldsymbol{x})] \tag{9}$$

为可以忽略的小量时，才能有：

$$\frac{\partial \boldsymbol{G}(\boldsymbol{x})}{\partial \boldsymbol{x}} \simeq 2[\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} - \boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x})]\,.$$

这样就有准法化方程：

$$\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{A}(\boldsymbol{x})\,\boldsymbol{x} = \boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{b}(\boldsymbol{x})\,. \tag{10}$$

此时，可以求解该准法化方程，给出目标函数 $\boldsymbol{G}\,(\boldsymbol{x})$ 驻点的近似值.

显然，对于短时段的观测资料，所忽略的量 $\boldsymbol{\tau}$ 可以为小量，在一定精度的情况下，求解准法化方程 (10) 和求解优化问题 MMUVM 是等价的. 而对于多圈长弧段的观测资料，由于所忽略的量 $\boldsymbol{\tau}$ 不为小量，则两种求解方法也就不可能等价.

从本质上而言， PUVM1 是利用准法化方程 (10) 进行数据处理的，因此， PUVM1 实际上是 MMUVM 的简化形式. 这样，我们就可以知道， PUVM1 只适用于利用短弧段的观测资料进行初始轨道确定，而不适用于利用多圈长弧段观测资料进行轨道确定或轨道改进.

7 PUVM1 准法化方程的迭代算法及其收敛性

如前所言，在只有短弧段的观测资料的情况下，求解 MMUVM 所对应的最优化问题，经过合理地近似和简化，可以转换为求非线性的准法化方程组 (10). 由实际背景可知，此方程组一定存在精确解的比较理想的近似解. 众所周知，有多种算法可以用来求解这类问题. 在本文中，我们将着重讨论文 [1] 中所提出的迭代算法及其有关问题.

7.1 PUVM1 准法化方程的迭代算法

在实际应用中，求解非线性准法化方程 (10) 时，采用迭代法进行求解. 根据单位矢量法的条件方程 (1), 可构造出如下的迭代方程 [1] ：

$$\left\{\begin{array}{l}
(f^{(k)}\,\boldsymbol{A}^*)\cdot\boldsymbol{r}_0 + (g^{(k)}\,\boldsymbol{A}^*)\cdot\dot{\boldsymbol{r}}_0 = -(w^{(k)}\,\boldsymbol{A}^*)\cdot\boldsymbol{W}_0 + \boldsymbol{A}^*\cdot\boldsymbol{R} \\
(f^{(k)}\,\boldsymbol{h}^*)\cdot\boldsymbol{r}_0 + (g^{(k)}\,\boldsymbol{h}^*)\cdot\dot{\boldsymbol{r}}_0 = -(w^{(k)}\,\boldsymbol{h}^*)\cdot\boldsymbol{W}_0 + \boldsymbol{h}^*\cdot\boldsymbol{R} \\
(f^{(k)}\,\boldsymbol{\rho}^*)\cdot\boldsymbol{r}_0 + (g^{(k)}\,\boldsymbol{\rho}^*)\cdot\dot{\boldsymbol{r}}_0 = -(w^{(k)}\,\boldsymbol{\rho}^*)\cdot\boldsymbol{W}_0 + \boldsymbol{\rho}^*\cdot\boldsymbol{R} + \rho \\
(\dot{f}^{(k)}\,\boldsymbol{\rho}^*)\cdot\boldsymbol{r}_0 + (\dot{g}^{(k)}\,\boldsymbol{\rho}^*)\cdot\dot{\boldsymbol{r}}_0 = -(\dot{w}^{(k)}\,\boldsymbol{\rho}^*)\cdot\boldsymbol{W}_0 + \boldsymbol{\rho}^*\cdot\dot{\boldsymbol{R}} + \dot{\rho}
\end{array}\right. , \tag{11}$$

其中：

$$\begin{aligned}
f^{(k)} &= f(t,\boldsymbol{x}_0^{(k)}), \quad \dot{f}^{(k)} = \dot{f}(t,\boldsymbol{x}_0^{(k)})\,, \\
g^{(k)} &= g(t,\boldsymbol{x}_0^{(k)}), \quad \dot{g}^{(k)} = \dot{g}(t,\boldsymbol{x}_0^{(k)})\,, \\
w^{(k)} &= w(t,\boldsymbol{x}_0^{(k)}), \quad \dot{w}^{(k)} = \dot{w}(t,\boldsymbol{x}_0^{(k)})\,,
\end{aligned}$$

并且 $\boldsymbol{x}_0^{(k)} = \left(\begin{array}{c} \boldsymbol{r}_0^{(k)} \\ \dot{\boldsymbol{r}}_0^{(k)} \end{array}\right)$ 为优化问题精确解 $\boldsymbol{x}_0^* = \left(\begin{array}{c} \boldsymbol{r}_0^* \\ \dot{\boldsymbol{r}}_0^* \end{array}\right)$ 的第 k 次近似结果.

对每组观测资料形成的如上形式的条件方程，进行加权处理，就可以得到如下的最小二乘问题

$$\min \|A(\boldsymbol{x}_0^{(k)})\,\boldsymbol{x}_0 - b(\boldsymbol{x}_0^{(k)})\|_2^2\,, \tag{12}$$

其中：$\boldsymbol{x}_0^{(k)} = \begin{pmatrix} \boldsymbol{r}_0^{(k)} \\ \dot{\boldsymbol{r}}_0^{(k)} \end{pmatrix}$ 为已知量，$\boldsymbol{x}_0 = \begin{pmatrix} \boldsymbol{r}_0 \\ \dot{\boldsymbol{r}}_0 \end{pmatrix}$ 为未知量.

由实际背景容易知道，$A(\boldsymbol{x}_0^{(k)})$ 和 $A(\boldsymbol{x}_0^*)$ 都为列满秩矩阵. 通过对此线性最小二乘问题进行求解，就可得第 $k+1$ 次的迭代结果 $\boldsymbol{x}_0^{(k+1)}$:

$$\boldsymbol{x}_0^{(k+1)} = [A(\boldsymbol{x}_0^{(k)})]^+ \, b(\boldsymbol{x}_0^{(k)}) \, . \tag{13}$$

若此迭代格式收敛，则可计算出优化问题的精确结果.

7.2 准法化方程迭代算法的收敛性

针对上面的迭代格式，我们给出如下定理：

定理：　设 $[A(\boldsymbol{x}_0)]^+ \in C^{n\times m}$ 和 $b(\boldsymbol{x}_0) \in C^m$ 在 $\boldsymbol{x}_0^*$ 的某个邻域 $\boldsymbol{\Omega}$ 内是 Lipschitz 连续的，且对任意的 $\boldsymbol{x}_0 \in \boldsymbol{\Omega}$, 都有 $rank[A(\boldsymbol{x}_0)] = n$, 则有：

$$\|\boldsymbol{x}_0^{(k+1)} - \boldsymbol{x}_0^*\| \le (\alpha_1 \, (1+\epsilon)\|b(\boldsymbol{x}_0^*)\| + \alpha_2 \, \|[A(\boldsymbol{x}_0^*)]^+\|) \, \|\boldsymbol{x}_0^{(k)} - \boldsymbol{x}_0^*\| \, . \tag{14}$$

证明：　由于

$$\|A(\boldsymbol{x}_0^*) \cdot \boldsymbol{x}_0^* - b(\boldsymbol{x}_0^*)\|_F^2 = \min \, ,$$

即

$$\boldsymbol{x}_0^* = [A(\boldsymbol{x}_0^*)]^+ \, b(\boldsymbol{x}_0^*) \, ,$$

又由 (13) 式，可知：

$$\begin{aligned} \|\boldsymbol{x}_0^{(k+1)} - \boldsymbol{x}_0^*\| &= \|[A(\boldsymbol{x}_0^{(k)})]^+ \, b(\boldsymbol{x}_0^{(k)}) - [A(\boldsymbol{x}_0^*)]^+ \, b(\boldsymbol{x}_0^*)\| \\ &\le \|[A(\boldsymbol{x}_0^{(k)})]^+ - [A(\boldsymbol{x}_0^*)]^+\| \cdot \|b(\boldsymbol{x}_0^{(k)})\| + \|[A(\boldsymbol{x}_0^*)]^+\| \cdot \|b(\boldsymbol{x}_0^{(k)}) - b(\boldsymbol{x}_0^*\| \, . \end{aligned}$$

又因为 $[A(\boldsymbol{x}_0)]^+$ 和 $b(\boldsymbol{x}_0)$ 在 $\boldsymbol{\Omega}$ 内是连续的，所以存在参数 α_1 和 α_2, 使得：

$$\|[A(\boldsymbol{x}_0^{(k)})]^+ - [A(\boldsymbol{x}_0^*)]^+\| \le \alpha_1 \, \|\boldsymbol{x}_0^{(k)} - \boldsymbol{x}_0^*\|$$

和

$$\|b(\boldsymbol{x}_0^{(k)}) - b(\boldsymbol{x}_0^*)\| \le \alpha_2 \, \|\boldsymbol{x}_0^{(k)} - \boldsymbol{x}_0^*\| \, .$$

所以，可有：

$$\|\boldsymbol{x}_0^{(k+1)} - \boldsymbol{x}_0^*\| \le (\alpha_1 \, \|b(\boldsymbol{x}_0^{(k)})\| + \alpha_2 \, \|[A(\boldsymbol{x}_0^*)]^+\|) \, \|\boldsymbol{x}_0^{(k)} - \boldsymbol{x}_0^*\| \, .$$

由于 $\|b(\boldsymbol{x}_0)\|$ 是 $\boldsymbol{x}_0$ 的连续函数，所以，对于任意的 $\epsilon > 0$ 存在一个以 $\boldsymbol{x}_0^*$ 为中心的某个邻域，使得

$$\|b(\boldsymbol{x}_0)\| \le (1+\epsilon)\|b(\boldsymbol{x}_0^*)\| \, .$$

因而，可得：

$$\|\boldsymbol{x}_0^{(k+1)} - \boldsymbol{x}_0^*\| \le (\alpha_1 \, (1+\epsilon)\|b(\boldsymbol{x}_0^*)\| + \alpha_2 \, \|[A(\boldsymbol{x}_0^*)]^+\|) \, \|\boldsymbol{x}_0^{(k)} - \boldsymbol{x}_0^*\| .$$

证毕.

从以上的结果可见，定理给出了判别迭代方程是否收敛的充分条件. 即当：

$$L = \alpha_1\,(1+\epsilon)\|b(\boldsymbol{x}_0^*)\| + \alpha_2\,\|[A(\boldsymbol{x}_0^*)]^+\| < 1$$

时，迭代方程是收敛的. 否则，迭代可能发散.

7.3 算例

由于 $A(\boldsymbol{x})$ 和 $b(\boldsymbol{x})$ 的复杂性，很难通过公式推导给出分析结果. 为了说明问题，我们给出了初始误差在一次迭代后所发生变化情况的数值结果. 下面我们基于根数为：

$$a = 7031.407(\mathrm{km}),\ \mathrm{i} = 67.994^\circ,\ \Omega = 95.481^\circ,\ \xi = -0.00197,\ \eta = 0.00046,\ \lambda = 0.0$$

的轨道，给出了收敛和不收敛情况下的数值例子.

例 1: 小偏心率收敛的情况

采用每 2 秒一组共 708 秒的单圈仿真数据，迭代初始值和收敛结果的差别为 0.01, 相应地，有：

$$\begin{aligned}
\|[A(\boldsymbol{x}_0^*)]^+\|_2 &= 0.206492\,,\\
\|b(\boldsymbol{x}_0^{(0)})\|_2 \simeq \|b(\boldsymbol{x}_0^*)\|_2 &= 23.5883\,,\\
\|\boldsymbol{x}_0^{(0)} - \boldsymbol{x}_0^*\|_2 &= 0.0244949\,,\\
\|b(\boldsymbol{x}_0^{(0)}) - b(\boldsymbol{x}_0^*)\|_2 &= 0.0000861984\,,\\
\|[A(\boldsymbol{x}_0^{(0)})]^+ - [A(\boldsymbol{x}_0^*)]^+\|_2 &= 0.000389619\,.
\end{aligned}$$

所以有：

$$\frac{\|b(\boldsymbol{x}_0^{(0)}) - b(\boldsymbol{x}_0^*)\|_2\,\|[A(\boldsymbol{x}_0^*)]^+\|_2}{\|\boldsymbol{x}_0^{(0)} - \boldsymbol{x}_0^*\|} = 0.000726654\,,$$

$$\frac{\|[A(\boldsymbol{x}_0^{(0)})]^+ - [A(\boldsymbol{x}_0^*)]^+\|_2 \cdot \|b(\boldsymbol{x}_0^*)\|_2}{\|\boldsymbol{x}_0^{(0)} - \boldsymbol{x}_0^*\|} = 0.375198\,.$$

以上数据可见 $L < 0.38$, 即初始值的误差，经过一次迭代之后，新数据的误差减少了不到原来误差的 0.38 倍，因此，这种迭代法应该有着比较快收敛速度.

例 2: 小偏心率不收敛的情况

使用同样的例子， 2 圈的仿真数据，迭代初始值和收敛结果的差别为 10^{-4}.

$$\begin{aligned}
\|[A(\boldsymbol{x}_0^*)]^+\|_2 &= 0.0491942,\\
\|b(\boldsymbol{x}_0^{(0)})\|_2 \simeq \|b(\boldsymbol{x}_0^*)\|_2 &= 40.8155,\\
\|\boldsymbol{x}_0^{(0)} - \boldsymbol{x}_0^*\|_2 &= 0.0000244949,\\
\|b(\boldsymbol{x}_0^{(0)}) - b(\boldsymbol{x}_0^*)\|_2 &= 0.000226505,\\
\|[A(\boldsymbol{x}_0^{(0)})]^+ - [A(\boldsymbol{x}_0^*)]^+\|_2 &= 0.000058034\,.
\end{aligned}$$

所以有：

$$\frac{\|b(\boldsymbol{x}_0^{(0)}) - b(\boldsymbol{x}_0^*)\|_2 \, \|[A(\boldsymbol{x}_0^*)]^+\|_2}{\|\boldsymbol{x}_0^{(0)} - \boldsymbol{x}_0^*\|} = 0.454899\,,$$

$$\frac{\|[A(\boldsymbol{x}_0^{(0)})]^+ - [A(\boldsymbol{x}_0^*)]^+\|_2 \cdot \|b(\boldsymbol{x}_0^*)\|_2}{\|\boldsymbol{x}_0^{(0)} - \boldsymbol{x}_0^*\|} = 96.7013\,.$$

从上面数据可见 $L > 96$, 对应于发散的情况，迭代一次，误差可能被放大约 100 倍.

由上可知，利用 MMUVM 目标函数梯度方程得到 PUVM1 准法化方程时，所忽略的量一定要是小量，使得 MMUVM 的最优解和 PUVM1 准法化方程的解的差别量在所容许的精度之内. 由此可知，即使所忽略的量为足够小的量，此时非线性的 PUVM1 准条件方程的解也是非常理想的，由于 PUVM1 的迭代方程 (11) 是条件收敛的，因此，它有时是发散的. 该部分的数值计算结果，都很好地解释了实际计算中所发生的现象，即：PUVM1 的迭代方法为什么只能处理短弧段的观测资料进行初始轨道确定，却不能使用长弧段的多圈资料进行轨道确定或轨道改进.

8　结论

本文首先在只考虑观测资料误差，而不考虑动力学模型误差的前提下，对条件方程给出了合理的加权策略，然后，再基于最小二乘的思想，建立了非线性的单位矢量法数学模型 MMUVM. 该模型的目标函数是服从 χ^2 分布的随机变量，使得进一步的研究工作，如：在一定的可信度的要求下进行区间预报，从数学原理上是可行的. 在实际应用上，MMUVM 解决了 PUVM1 方法只能处理单圈短弧段数据进行初始轨道确定，不能使用长弧段多圈资料进行轨道确定或轨道改进的这一问题.

本文还分别利用仿真数据和实测数据，建立了相应的优化模型 MMUVM. 然后，选取了直接方法对这些最优化问题进行了数值求解，并得到了满意的数值结果. 在使用直接方法求解优化问题的时候，仅仅需要构造相应的目标函数，不需要计算梯度方法中所要涉及到的导数问题，因此，求解最优化问题的直接方法具有相当的实用价值.

通过分析 MMUVM 的目标函数的梯度方程和 PUVM1 准法化方程之间的关系，进一步指出了 PUVM1 只是 MMUVM 的一种简化形式. 从根本上，解释了利用 PUVM1 准法化方程，只能使用短弧段观测资料进行初始轨道确定，不能用于处理多圈长弧段数据进行轨道确定或轨道改进.

本文对文 [1] 中所提及的 PUVM1 迭代法的收敛性问题进行了初步的讨论和研究，并给出与之相应的数值验证例子，指出求解 PUVM1 非线性准法化方程的迭代法 (11) 式是条件收敛的. 这就意味着，即使在构造准法化方程时忽略的量 $\boldsymbol{\tau}$(见 (9) 式) 比较小，所构造的法化方程也是合情合理的. 但是，用如上的迭代法进行求解时，迭代过程可能会是发散的，也就得不到准法化方程的解.

致谢　在南京航空航天大学理学院的倪勤教授的关心和帮助下, 本文才能得以顺利完成. 学生谨向倪勤老师和其他各位对本文提供帮助和建议的老师和同事，表示由衷地感谢.

参 考 文 献

1 陆本魁，李剑峰，马静远，等．宇航学报． 1999, 20(1): 14
2 朱国钦，王天波．误差理论及其在石油测井中的应用．北京：石油工业出版社， 2000
3 袁亚湘．非线性规划数值方法．上海：科学技术出版社， 1992
4 Ni Q, Hu S. ACTA mathematica Scientia Ser B, 2004, 24: 281
5 Ni Q. SIAM J Optimization, 2005, 15: 826

The Mathematical Model of Unit Vector Method and the Iterative Method for its Simplification

CHEN Wu-shen[1,2] LU Ben-kui[1] MA Jing-yuan[1]
(1 *Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing* 210008)
(2 *School of Sciences, Nanjing University of Technology, Nanjing* 210009)

Abstract In practical application, the perturbed unit vector method (PUVM1) is widely used to determine the orbit by the observation data related to artificial satellites. Some improvement efforts have been done to this method in this paper. Firstly, based on the model of measurement error and the idea of unit vector method, a new mathematical model MMUVM which is essentially a nonlinear optimization problem is introduced. Using simulation data and observation data, the objective function of MMUVM can be formed and solved respectively. Numerical results show that the mathematical model MMUVM is correct and the direct search method, i.e. the quadratic tri-diagonal interpolation DFO method is practical and effective. Secondly, based on the relationship of the semi-normal equation of PUVM1 and the gradient of MMUVM, it points out that MMUVM is a reasonable generalization of PUVM1 and PUVM1 is essentially a simplification to MMUVM. It is a big residual problem when MMUVM is formed by multi-circle long arc data, on the contrary, it is a little residual one when MMUVM is obtained by short arc data. The optimization problem of MMUVM can be solved by a suitable algorithm. The original PUVM1 algorithm is only suitable for the little residual problem. It can be obtained that PUVM1 may be used to determine the preliminary orbit by short arc data, but can not be used to deal with the job of orbit improvement by multi-circle long arc data. Finally, a discussion is made on the convergence of the iterative scheme which is used to solve the semi-normal equation of PUVM1. The corresponding numerical examples are also presented to point out that the iterative scheme of PUVM1 is convergent conditionally. It means that even if there is a reasonable solution to the semi-normal equation of PUVM1, it may be also divergent using the iterative scheme of PUVM1 in some cases.

Key words Celestial Mechanics: Determination of Orbit, Celestial Mechanics: Numerical Method