

·信号与信息处理·

样条拟合在光测数据处理中的应用

郭金虎, 赵贤超

(92853 部队, 辽宁 兴城 125106)

摘要:针对目前光测数据处理中,应用多项式拟合原理进行参数估计与插值计算带来较大截断误差的问题,提出了递推样条最小二乘拟合方法。该方法根据飞行目标航迹数据特点,将样条函数和递推思想相结合。经仿真计算证明,该方法能够提高拟合的精度与计算速度,具有很实用的价值。

关键词:样条函数;最小二乘法;截断误差

中图分类号: TP391.4

文献标识码: A

文章编号: 1673-1255(2015)-01-0032-03

Application of Spline Fitting in Optical Measurement Data Processing

GUO Jin-hu, ZHAO Xian-chao

(92853 Army Unit, Xingcheng 125106, China)

Abstract: During optical measurement data processing, multinomial fitting principle is used to estimate parameters and calculate interpolation, which brings higher block error. According to the problem, recursion spline least square fitting method is proposed. Based on the characteristics of flight target track datum, and with the spline function and recursion idea, the simulation calculations show that the accuracy and calculation speed of the fitting can be improved by the method. It has higher practical value.

Key words: spline function; least square method; block error

在光测数据的拟合与补点、野值辨识与剔除的合理性检验中,目前多采用多项式去逼近飞行目标航路轨迹,从而带来较大的截断误差^[1],若用分段多项式逼近,则严重影响了航迹的光滑性与完整性。而采用样条拟合在解决数据逼近问题将比用多项式逼近取得更好的结果,对于提高数据处理精度也起到了不错的效果。

构造一个函数的关键是要估计出该函数的各项系数,也就是参数的估计。工程中应用较多的仍然是最小二乘估计,但实际工程中,往往因为光测数据的相关性,导致估计系数矩阵存在一定的列相关,即系数矩阵存在一定病态,从而影响了最小二乘估计的收敛性和精度^[2]。另外,新获得一个测量数据时,若想重新确定拟合多项式,所有工作都要

从头开始,使得本次估计和前面的估计结果无任何关系,即估计之间缺乏连续性和继承性。目前,追求估计方法的递推性,是理论研究和工程实际刻意解决的一个问题^[3-4],也是提高估计速度、减少计算机内存的需要。

文中将从 Kalman 滤波递推思想出发,构造样条函数拟合的递推计算方法,以期减少截断误差,提高计算精度和速度。

1 样条函数的递推最小二乘计算

如前所述,样条函数能有效解决截断误差问题,但在拟合过程中,目前仍然使用的是逐点求解的方法,极大地影响了计算速度和应用效果。如果

建立具有递推性质的样条拟合方法,是目前工程实际中急需解决的问题之一。

1.1 最小二乘的递推改进

设观测模型为

$$y_i = \beta_{1x_{i,1}} + \cdots + \beta_{mx_{i,m}} + e_i \quad i = 1, 2, \cdots, n \quad (1)$$

且 $E[e_i] = 0, D[e_i] = \sigma_i^2, E[e_i e_j] = 0, i \neq j$; 现记

$$Y_n = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}, H_n = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

$$E_n = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (2)$$

于是有 $Y_n = H_n \beta + E_n$, 且

$$\text{Var } E_n = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \sigma_n^2 \end{bmatrix} = R_n \quad (3)$$

现在问题是获得 n 次观测后得到了系数 β 的估计 $\hat{\beta}_n$, 假若又获得了一个新的观测 y_{n+1} 后, 如何重新估计系数 $\hat{\beta}_{n+1}$? 不妨设

$$y_{n+1} = \beta_1 x_{n+1,1} + \cdots + \beta_m x_{n+1,m} + e_{n+1} \quad (4)$$

记 $M_{n+1} = [x_{n+1,1}, x_{n+1,2}, \cdots, x_{n+1,m}]$, 则式(4)可写为

$$y_{n+1} = M_{n+1} \beta + e_{n+1} \quad (5)$$

即在进行了 $n+1$ 次观测之后, β 的估计为 $\hat{\beta}_{n+1}$

$$\hat{\beta}_{n+1} = [H_{n+1}^T R_{n+1}^{-1} H_{n+1}]^{-1} H_{n+1}^T R_{n+1}^{-1} Y_{n+1} \quad (6)$$

现在考虑式(6)的递推形式。利用矩阵和向量的分块表示, 式(6)可为

$$\hat{\beta}_{n+1} = \left(\begin{bmatrix} H_n^T M_{n+1}^T \\ 0 \end{bmatrix} \begin{bmatrix} R_n & 0 \\ 0 & \sigma_{n+1}^2 \end{bmatrix}^{-1} \begin{bmatrix} H_n \\ M_{n+1} \end{bmatrix} \right)^{-1} \cdot$$

$$\begin{bmatrix} H_n^T M_{n+1}^T \\ 0 \end{bmatrix} \begin{bmatrix} R_n & 0 \\ 0 & \sigma_{n+1}^2 \end{bmatrix}^{-1} \begin{bmatrix} Y_n \\ y_{n+1} \end{bmatrix} =$$

$$\left[H_n^T R_n^{-1} H_n + \frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2} \right]^{-1} [H_n^T R_n^{-1} Y_n + M_{n+1}^T \sigma_{n+1}^{-2} y_{n+1}] \quad (7)$$

$$\text{或} \left[H_n^T R_n^{-1} H_n + \frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2} \right] \hat{\beta}_{n+1} =$$

$$[H_n^T R_n^{-1} Y_n + M_{n+1}^T \sigma_{n+1}^{-2} y_{n+1}] \quad (8)$$

即 $\hat{\beta}_{n+1} = \hat{\beta}_n + (\hat{\beta}_{n+1} - \hat{\beta}_n)$, 于是

$$\left[H_n^T R_n^{-1} H_n + \frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2} \right] [\hat{\beta}_{n+1} - \hat{\beta}_n] = \left[H_n^T R_n^{-1} Y_n + M_{n+1}^T \sigma_{n+1}^{-2} y_{n+1} \right] - [H_n^T R_n^{-1} Y_n] \quad (9)$$

$$[H_n^T R_n^{-1} Y_n + M_{n+1}^T \sigma_{n+1}^{-2} y_{n+1}]$$

由于 $(H_n^T R_n^{-1} H_n) \hat{\beta}_n = H_n^T R_n^{-1} Y_n$, 于是

$$\left[H_n^T R_n^{-1} H_n + \frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2} \right] (\hat{\beta}_{n+1} - \hat{\beta}_n) = - \frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2} \hat{\beta}_n +$$

$$\frac{M_{n+1}^T y_{n+1}}{\sigma_{n+1}^2} = \frac{M_{n+1}^T}{\sigma_{n+1}^2} (y_{n+1} - M_{n+1} \hat{\beta}_n) \quad (10)$$

由此可解得

$$\hat{\beta}_{n+1} = \hat{\beta}_n + \left[H_n^T R_n^{-1} H_n + \frac{M_{n+1}^T M_{n+1}}{\sigma_{n+1}^2} \right]^{-1} \frac{M_{n+1}^T}{\sigma_{n+1}^2} (y_{n+1} - M_{n+1} \hat{\beta}_n) \quad (11)$$

该式右端需要对 $m \times m$ 阶矩阵求逆, 在未知参数向量维数增大时, 计算是不方便的。现根据分块矩阵求逆引理^[5], 设:

$$(1) P_n = [H_n^T R_n^{-1} H_n]^{-1}, \text{ 则 } P_{n+1} = [H_{n+1}^T R_{n+1}^{-1} H_{n+1}]^{-1},$$

可得

$$P_{n+1} = P_n - P_n M_{n+1}^T [M_{n+1} P_n M_{n+1}^T + \sigma_{n+1}^2]^{-1} M_{n+1} P_n \quad (12)$$

$$(2) \hat{\beta}_{n+1} = \hat{\beta}_n + P_{n+1} \frac{M_{n+1}^T}{\sigma_{n+1}^2} (y_{n+1} - M_{n+1} \hat{\beta}_n)$$

$$\text{由此 } P_{n+1} \frac{M_{n+1}^T}{\sigma_{n+1}^2} = P_n M_{n+1}^T [M_{n+1} P_n M_{n+1}^T + \sigma_{n+1}^2]^{-1} \quad (13)$$

于是得最小二乘的递推式如下

$$\hat{\beta}_{n+1} = \hat{\beta}_n + K_n (y_{n+1} - M_{n+1} \hat{\beta}_n) \quad (14)$$

$$K_n P_n M_{n+1}^T [M_{n+1} P_n M_{n+1}^T + \sigma_{n+1}^2]^{-1} \quad (15)$$

$$P_{n+1} = [H_{n+1}^T R_{n+1}^{-1} H_{n+1}]^{-1} =$$

$$P_n - K_n M_{n+1} P_n = (I - K_n M_{n+1}) P_n \quad (16)$$

和传统的二乘方法比较, 显然递推最小二乘估计方法简便易行, 且节省内存。当获得新的观测数据后, 可在前一步基础上递推得到未知参数的新的估计, 计算速度得到明显提高。

1.2 样条拟合的递推计算

光电经纬仪测量数据是等间隔采样的, 因此利用样条函数特别是 B 样条函数拟合是方便的, 现在给出具体的构造方法。

设有一组观测数据: $(t_1, y_1), (t_2, y_2), \cdots, (t_m, y_m)$

$$y_i = f(t_i) \quad (i = 1, 2, \cdots, m)$$

记 $a = \min\{t_i\}, b = \max\{t_i\}$

样条函数最小二乘拟合问题是在空间 $S_n(\Delta)$ 中, 找对于 $f(x)$ 关于范数 $\|\cdot\|$ 的最佳逼近, 即找

$S(x) \in S_n(\Delta)$ 使

$$\|f(x) - S(x)\| = \min_{S \in S_n(\Delta)} \|f(x) - S(x)\| \quad (17)$$

$$S_n(\Delta) = \text{Span}\{B_0(x), B_1(x), \dots, B_n(x)\} \quad (18)$$

即对任何 $S(x) \in S_n(\Delta)$, 都能表示为

$$S(x) = \sum_{i=0}^n \beta_i B_i(x) \quad (19)$$

理论证明已知^[5], $S(x)$ 是 $S_n(\Delta)$ 中对 $Y = \{y_1, y_2, \dots, y_m\}^T$ 的离散最佳逼近的充要条件是 $S(x) - y_k (k=0, 1, \dots, m)$ 与所有的 $B_i(x) (i=0, 1, \dots, n)$ 正交, 即

$$\langle S(x) - y_k, B_i \rangle = 0; \quad k=0, 1, \dots, m; \quad i=0, 1, \dots, n \quad (20)$$

因此, 求 $S(x)$ 的问题, 实际上就是求 B 样条总数的系数 β_i , 即求式(19)使

$$\sum_{j=0}^n \langle B_i, B_j \rangle \beta_j = \langle B_i, y_k \rangle \quad i=0, 1, \dots, n \quad (21)$$

式(21)解的唯一性有定理保证^[6]。

现设 $y_i = \beta_0 B_0(t_i) + \dots + \beta_m B_m(t_i) + e_i \quad i=0, 1, \dots, n$

其中, e_i 满足 $E[e_i] = 0, D[e_i] = \sigma_i^2, E[e_i e_j] = 0, i \neq j$ 的假设。

在拟合区间内很容易给出 B 样条函数^[7]的节点和递推公式^[8], 以求出 $B_i(x) (i=0, 1, \dots, n)$, 此部分不再赘述。由此, 可利用式(14), 式(15)递推求得拟合多项式系数的估计值 $\hat{\beta}$ 。

2 应用结果说明

现用某次任务中光电经纬仪测得的数据双站交会处理后, 对所提出的方法和传统的算法做比较分析。测量数据的时间段为 30 s (不含特征段), 采样频率每秒 10 点, 其间含有丢点和野值。图 1 表示一条三阶多项式拟合处理后的估值与测量值之差, 显然误差较大。图 2 是分三段用不同的三阶多项式拟合后的差值图。可见误差得到了明显改善, 但三个多项式接头出现拐点, 影响了目标飞行航迹的光滑性。图 3 是用三阶 B 样条函数对整段数据采用递推拟合计算的差值图, 效果是显而易见的。

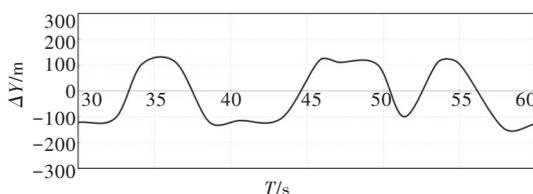


图 1 三阶多项式拟合处理后 Y 方向估值与测量值差图

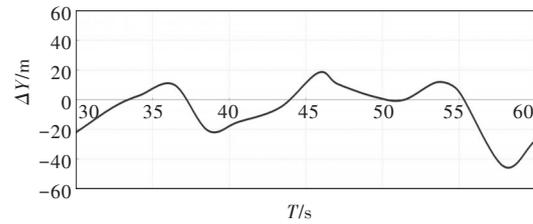


图 2 三段不同三阶多项式拟合后 Y 方向估值与测量值差图

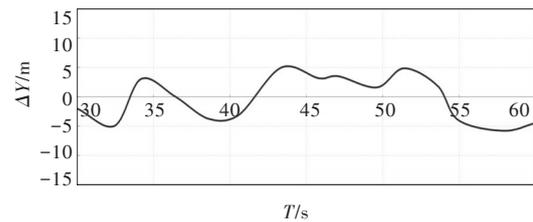


图 3 三阶 B 样条函数递推拟合后 Y 方向估值与测量值差图

在利用 B 样条递推计算时, 应注意如下三个问题: (1) 递推初值 $\hat{\beta}_0$ 和 P_0 的确定可利用开始几个观测数据, 例如在 t_1, t_2, \dots, t_n (n 不一定太大, 只需保证 $H_n^T H_n$ 可逆即可^[9-10]) 时刻已得到观测值, 则可得到下式

$$\hat{\beta}_n = P_n H_n^T R_n^{-1} y_n \quad \text{和} \quad P_n = [H_n^T R_n^{-1} H_n]^{-1} \quad (22)$$

以 $\hat{\beta}_n, P_n$ 作为递推运算的起始条件, 即令

$$\hat{\beta}_0 = \hat{\beta}_n, \quad P_0 = P_n$$

往后的计算就可利用式(14), 式(15)进行。(2) 为了改善拟合效果, 对原始数据最好先进行野值的剔除处理^[11-17], 当然, 文中方法曾经亦作为野值处理的一种方法在任务中加以应用, 效果亦可。该方法的极大价值在于快速求出较高精度的拟合观测值, 用以确定目标飞行航迹和精度分析。(3) 关于拟合段落的确。经实际任务检验, 在用样条函数拟合计算时, 时间段落不能无限长, 在平稳段可适当长一些, 但一般取 20~40 s 为宜。在特征段则以不超过 20 s 为好。

另外, 在样条函数的三类边值条件中^[2], 取一类即可满足精度要求。

表 1 是对某特征段附近数据的两种方法截断误差比较结果。多项式的推断误差均用理论航迹做了补偿处理, 但极大影响了计算速度, 而且使实测数据在特征段失去了真实性。而用样条函数拟合计算, 则不会出现此类现象。

(下转第 38 页)

和外磨工件,由于不同加工方式同一表面粗糙度对比度、平均值、方差数值不同,平磨工件数值比外磨工件数值大。同时偏度和峰值也存在一些差异,且平磨工件比外磨工件波动大。这为远距离探测金属物体表面粗糙度和加工工艺提供一种途径。

参考文献

- [1] 武勇军,李向.超精加工表面粗糙度的激光测量仪[J].应用激光,1996,12(2):60-62.
- [2] 李田泽,张静华.一种无触点快速测量表面粗糙度的方法[J].激光技术,1998,22(4).
- [3] 张蕴东,浦军.光探针式表面形貌测量技术的研究[J].计量与测试技术,2005,32(7):4-5.
- [4] 丰明坤,隋成华.用于表面粗糙度检测的双波长光纤传感器研究[J].量子电子学报,2004,21(6):873-878.
- [5] 于昊,王劲松.表面粗糙度二维信息同步测量方法[J].科学技术与工程,2005,5(2):110-112.

- [6] 周烜,刘秀玲.工件表面粗糙度微机测量系统[J].控制工程,2003,10(3):285-288.
- [7] 郑小兵.一维随机粗糙表面散射特征的数值分析[J].光子·激光,1997,8(5):398-402.
- [8] 齐文宗,李志刚.平面电磁波理想金属表面散射场分布及粗糙度参数关系的理论分析[J].光学·精密工程,1998,6(5):122-126.
- [9] 刘映栋,王亚伟.散斑法测量表面粗糙度 R_q 的研究[J].江苏工学院学报,1994,15(2):89-93.
- [10] 赵学增,周莉莉.入射角对激光散斑测量表面粗糙度的影响[J].中国机械工程,2003,14(7):563-564.
- [11] Yang L, LU R S, Lei L Q. Grinding surface roughness measurement based on Gauss-Markov random field model of laser speckle pattern texture[C]. SPIE 2010, 7997(1X):1-7.
- [12] Toh S I, Shang H M, Jay C. Surface-roughness study using laser speckle method[J]. Optics and Lasers in Engineering, 1998, 29(3): 217-225.

(上接第34页)

表1 多项式拟合与样条拟合在特征段附近的截断误差

时间/s	117.5	118.5	119.5	120.5	121.5	122.5	123.5	124.5	125.5
多项式之差/m	1.223	4.451	8.536	13.347	21.490	14.341	9.418	4.131	1.756
样条函数之差/m	0.024	0.309	0.768	1.156	1.345	1.075	0.835	0.578	0.116

3 结论

实际应用结果证明,利用传统方法仅计算位置和速度6个参数时, m 个采样点就需计算 $6m$ 个参数。而用文中方法,则可以把较长时间段的位置和速度参数集中处理,建立估计航迹参数的拟合模型,极大地减少了计算量,同时也提高了其速度和精度。同时运用该方法,可以精确地估计拟合出飞行目标的航迹数据,为经纬仪提供引导,确保了经纬仪在丢失目标的情况下,通过自引导和互引导方式及时准确地捕获目标,故可在光电经纬仪的跟踪测量中得到推广应用。

参考文献

- [1] 刘利生.外弹道测量数据处理[M].北京:国防工业出版社,2000.

- [2] 祝转民.外测数据的自适应样条滤波与精度估计[D].长沙:国防科技大学,1997.
- [3] 雍L,索德斯图姆T.递推辨识的理论与实践[M].田立生,译.北京:科学出版社,1989.
- [4] 祝转民.外弹道测量数据野值的递推辨识[J].飞行器测控学报,2001,20(3).
- [5] A kaik H1. Filtering auto regressive model for prediction[J] INST Statist Moth, 1969: 21.
- [6] Parzen E1 some recent advance in time series modeling[J]. IEEE Trans1 Auto-control, 1974(19): 723-730.
- [7] 王省富.样条函数及其应用[M].西安:西北工业大学出版社,1998:3-12.
- [8] 王峰.一种新型插补算法的研究与开发[D].太原:华北工学院机,2001.
- [9] Boehm W. Inserting new knots into B-spline curves[J]. Computer Aided Design, 1980, 12(4): 199-201.
- [10] 王朝瑞,史荣昌.矩阵分析[M].北京:北京理工大学出版社,1987.