

带插值条件的移动最小二乘曲线拟合

倪 慧, 李 重, 宋红星, 李静芳

(浙江理工大学理学院, 杭州 310018)

摘 要: 提出一种带插值条件的移动最小二乘曲线拟合方法。首先考虑带插值条件的最小二乘拟合,构造新的带插值条件最小二乘拟合方法,它具有拟合函数次数低、计算方便等优点。然后,将该方法推广到移动最小二乘拟合中,得到带插值条件的移动最小二乘曲线,实验结果表明该方法拟合效果更好。

关键词: 插值; 曲线拟合; 最小二乘法; 移动最小二乘法

中图分类号: O241.2 **文献标识码:** A

0 引 言

数据拟合是一个重要的研究课题,经常用于处理科学实验中涉及到的大量数据,在生物、化工、信号处理、计算机图形学、统计等方面有着重要应用^[1]。对于实验中用到的一些数据,需要找到它们隐含的内在规律,通常会对这些离散的数据进行曲线拟合,做出逼近曲线,这样要求的曲线不必经过所有的数据点。最常用的方法就是最小二乘法,要求误差的平方和达到最小,可通过解线性方程组求出拟合曲线。后来出现了一些优于最小二乘法的拟合方法,有一种方法称为移动最小二乘法^[2-3],相对于最小二乘法有着许多优点^[4-5]。Lancaster 和 Salkauskas 最早在曲面生成中使用了移动最小二乘法。近年来,移动最小二乘法在无网格方法等领域中有着广泛的应用。

当进行数据拟合的同时,有时还需要拟合曲线经过一些点,这就需要对通常的最小二乘法 and 移动最小二乘法进行插值处理。文献[6]中针对带插值条件的最小二乘法进行了讨论,并给出了构造公式。本文提出一种新的带插值条件最小二乘方法,与文献[6]相比,构造的拟合函数具有次数低、计算方便等优点,并可将该方法应用于移动最小二乘拟合中,得到的拟合效果更好。

1 带插值条件的最小二乘曲线拟合

1.1 拟合方法介绍

将带插值条件的最小二乘法应用于曲线拟合时,即对给定离散数据点,要求拟合曲线同时经过其中的某些节点。文献[6]给出了下面方法:

方法一:假设给定离散点为 $(x_i, y_i), i=1, 2, \dots, n$,插值条件为 $(x_s, y_s), s=1, 2, \dots, t, t \leq n$ 。若 $P(x) = a_{m-1}x^{m-1} + a_{m-2}x^{m-2} + \dots + a_1x + a_0$ 为拟合函数,则带插值条件的最小二乘拟合曲线为

$$y = (a_{m-1}x^{m-1} + a_{m-2}x^{m-2} + \dots + a_1x + a_0) \prod_{s=1}^t (x - x_s) + \sum_{s=1}^t l_s(x)y_s, \quad (1)$$

$$\text{其中 } l_s(x) = \prod_{\substack{j=1 \\ s \neq j}}^t \frac{(x - x_j)}{(x_s - x_j)}, \quad s = 1, 2, \dots, t.$$

式(1)可通过解线性方程组求出系数 a_i , 具体可参见文献[6]。

若 $P(x)$ 的次数为 $m-1$, 插值点个数为 t , 则插值条件下的最小二乘拟合曲线次数为 $(m-1) \times t$ 。

本文给出新的构造公式, 即

方法二: 假设给定离散点为 $(x_i, y_i), i=1, 2, \dots, n$, 插值条件为 $(x_s, y_s), s=1, 2, \dots, t, t \leq n$ 。若 $P(x) = a_{m-1}x^{m-1} + a_{m-2}x^{m-2} + \dots + a_1x + a_0$ 为拟合函数, 则带插值条件的最小二乘拟合曲线为

$$y = P(x) - \sum_{s=1}^t l_s(x) \delta_s, \quad (2)$$

其中 $\delta_s = P(x_s) - y_s = a_{m-1}x_s^{m-1} + a_{m-2}x_s^{m-2} + \dots + a_1x_s + a_0 - y_s$ 。

令

$$\lambda = \sum_{j=1}^n \left[y_j - P(x_j) + \sum_{s=1}^t l_s(x_j) \delta_s \right]^2,$$

λ 是以 $P(x)$ 系数为自变量的函数, 将问题转化为求 λ 在哪些点处取最小值, 可通过求解方程组

$$\lambda_{a_i}(a_0, a_1, \dots, a_{m-1}) = 0, \quad i=1, 2, \dots, m-1$$

来求出 a_i 。该方程组即为

$$\begin{cases} \sum_{j=1}^n \left[y_j - (a_{m-1}x_j^{m-1} + a_{m-2}x_j^{m-2} + \dots + a_1x_j + a_0) + \sum_{s=1}^t l_s(x_j) \delta_s \right] \cdot \\ \quad \left[-1 + \sum_{s=1}^t l_s(x_j) \right] = 0 \\ \sum_{j=1}^n \left[y_j - (a_{m-1}x_j^{m-1} + a_{m-2}x_j^{m-2} + \dots + a_1x_j + a_0) + \sum_{s=1}^t l_s(x_j) \delta_s \right] \cdot \\ \quad \left[-x_j + \sum_{s=1}^t l_s(x_j) x_s \right] = 0 \\ \dots\dots \\ \sum_{j=1}^n \left[y_j - (a_{m-1}x_j^{m-1} + a_{m-2}x_j^{m-2} + \dots + a_1x_j + a_0) + \sum_{s=1}^t l_s(x_j) \delta_s \right] \cdot \\ \quad \left[-x_j^{m-1} + \sum_{s=1}^t l_s(x_j) x_s^{m-1} \right] = 0 \end{cases}$$

将 a_0, a_1, \dots, a_{m-1} 计算出来代入式(2)中, 可得到插值条件下的最小二乘曲线拟合。我们注意到, 若 $P(x)$ 的次数为 $m-1$, 插值点个数为 t , 则插值条件下的最小二乘拟合曲线次数为 $\text{Max}\{m-1, t-1\}$, 比方法1的带插值条件最小二乘拟合函数次数低。另外, 我们也可将该构造公式看成是一般最小二乘拟合公式的修正, 其修正项为 $\sum_{s=1}^t l_s(x) \delta_s$ 。因此可先通过一般最小二乘方法得到拟合函数 $P(x)$, 再计算修正项

$\sum_{s=1}^t l_s(x) \delta_s$, 得到的带插值条件最小二乘拟合曲线可看成是方法二的有效近似, 但计算更加简洁方便。

1.2 算例

给定一组离散数据点 $x=[1, 2, 5, 4, 5, 6, 7, 8, 9, 10]; y=[1, 2, 2.5, 3, 4, 5, 5.5, 7]$, 要求拟合这组数据的同时曲线要经过 $(x_1, y_1), (x_8, y_8)$, 即 $(1, 1), (10, 7)$ 两点, 分别采用一般最小二乘法以及上面几种带插值条件的最小二乘方法进行拟合比较。图1是一般最小二乘拟合效果(经验函数为线性函数), 图2是方法一得到的插值条件最小二乘拟合结果(经验函数为3次多项式)。图3中用实线表示的曲线是方法二得到的插值条件最小二乘拟合效果(经验函数为2次多项式), 虚线表示的曲线是通过对应修正项计算得到的插值条件最小二乘拟合效果。图4中实线表示的曲线是使用方法二并取经验函数为3次多项式得到的拟合效果, 虚线表示的曲线是通过修正项计算得到的拟合效果。实验结果发现, 本文方法可以使用次数较低的经验函数构造带插值条件最小二乘拟合, 且拟合效果较好。另外, 通过修正项计算得到的带插值条件最小二乘拟合曲线跟方法二相比, 拟合效果相差不大, 但计算更加方便。

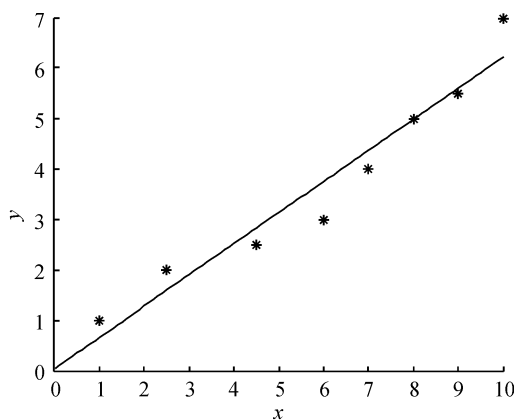


图 1 通常情况下最小二乘拟合

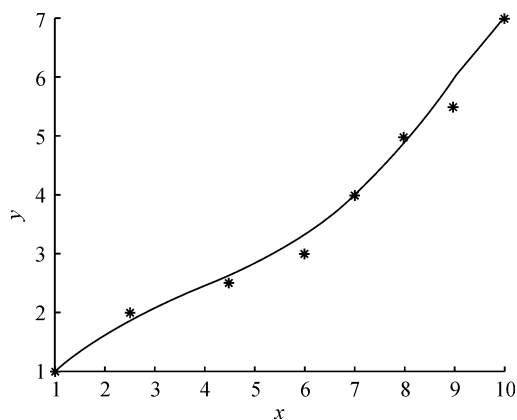


图 2 插值条件下方法一拟合(5次多项式)

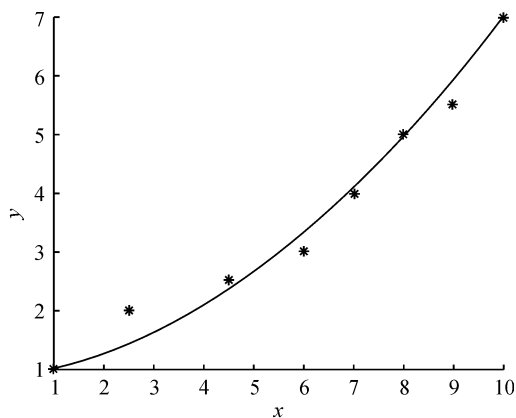


图 3 插值条件下方法二拟合(2次多项式)

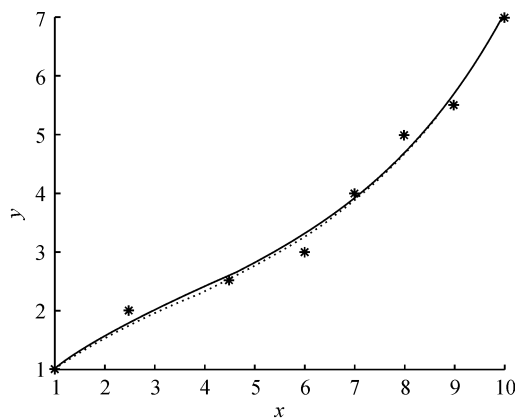


图 4 插值条件下方法二拟合(3次多项式)

2 移动最小二乘法

2.1 移动最小二乘曲线拟合

将拟合函数表述为如下形式:

$$f(x) = \sum_{i=1}^m p_i(x) a_i(x) = p^T(x) a(x), \quad (3)$$

其中 $a(x) = (a_1(x), a_2(x), \dots, a_m(x))^T$ 为待定系数, $p(x) = (p_1(x), p_2(x), \dots, p_m(x))^T$ 为基函数向量, 通常选择完备多项式基, 例如二维情况

$$\text{线性基} \quad p(x) = (1, x, y)^T \quad (m=3)$$

$$\text{二次基} \quad p(x) = (1, x, y, x^2, xy, y^2)^T \quad (m=6)$$

为了得到较为精确的局部近似值, 需使局部近似值 $f(x_i)$ 和节点值 y_i 之差平方带权最小, 因此残差的离散加权 L_2 范式为:

$$J = \sum_{i=1}^n w(x - x_i) [f(x) - y_i]^2 = \sum_{i=1}^n w(x - x_i) [p^T(x_i) a(x) - y_i]^2, \quad (4)$$

其中 n 是求解区域内的节点数, $f(x)$ 是拟合函数, $w(x - x_i)$ 是节点 x_i 的权函数。

权函数应该是非负的, 且随着 $\|x - x_i\|_2$ 的增加单调递减, 权函数还应该具有紧支性, 即在支持域 (x 的影响区域) 内不等于 0, 在支持域之外全为 0, 一般选用圆形作为权函数的支持域, 半径记为 r 。常用的权函数是样条函数, 记 $s' = x - x_i$, $s = \frac{s'}{r}$, 则三次样条函数形式如下:

$$\omega(s) = \begin{cases} \frac{2}{3} - 4s^2 + 4s^3 & s \leq \frac{1}{2} \\ \frac{4}{3} - 4s + 4s^2 - \frac{4}{3}s^3 & \frac{1}{2} < s \leq 1 \\ 0 & s > 1 \end{cases} \quad (5)$$

要求出待定系数 $a(x)$, 先要使 J 取得最小值, 先将式(4)写成矩阵形式:

$$J = (Pa(x) - Y)^T W(x) (Pa(x) - Y)$$

其中

$$Y = (y_1, y_2, \dots, y_n)^T,$$

$$W(x) = \text{diag}(w_1(x), w_2(x), \dots, w_n(x)), w_i(x) = w(x - x_i)$$

$$P = \begin{bmatrix} p_1(x_1) & p_2(x_1) & \cdots & p_m(x_1) \\ p_1(x_2) & p_2(x_2) & \cdots & p_m(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_1(x_n) & p_2(x_n) & \cdots & p_m(x_n) \end{bmatrix}$$

根据最小二乘原理求得待定系数为:

$$a(x) = A^{-1}(x)B(x)Y$$

其中 $A(x) = P^T W(x) P$, $B(x) = P^T W(x)$ 。

代入式(3), 得拟合函数为

$$f(x) = \sum_{i=1}^n \phi_i^k(x) y_i = \psi^k(x) Y,$$

其中 $\psi^k(x)$ 为形函数, k 表示基函数的阶数,

$$\psi^k(x) = [\phi_1^k, \phi_2^k, \dots, \phi_n^k] = p^T(x) A^{-1}(x) B(x)$$

2.2 带插值条件的移动最小二乘曲线拟合

当拟合曲线要求经过某些节点的时候, 就需要将通常的移动最小二乘法增加插值条件, 可利用上面带插值条件的最小二乘曲线拟合方法二进行构造。

假设若给出的离散点为 (x_i, y_i) , $i=1, 2, \dots, n$, 插值条件为 (x_s, y_s) , $s=1, 2, \dots, t$, $t \leq n$, 采用一般的移动最小二乘法得到的拟合曲线为 $f(x)$, 则带插值条件的移动最小二乘拟合曲线可表示为

$$y = f(x) - \sum_{s=1}^t l_s(x) \delta_s \quad (6)$$

$$\text{其中 } l_s(x) = \prod_{\substack{j=1 \\ s \neq j}}^t \frac{(x - x_j)}{(x_s - x_j)}, s = 1, 2, \dots, t, \delta_s = f(x_s) - y_s.$$

该构造公式也可看成是对移动最小二乘拟合公式的修正, 修正项为 $\sum_{s=1}^t l_s(x) \delta_s$, 其具体计算过程为:

- 先求出不带插值条件的移动最小二乘拟合曲线;
- 接着计算插值节点下的偏差 $\delta_s = f(x_s) - y_s$;
- 最后计算 $l_s(x)$ 并利用公式(6)得到插值条件下的移动最小二乘拟合曲线。

我们通过实验对一般移动最小二乘和带插值条件移动最小二乘方法进行比较。给定离散数据点为 $x = [1, 2.5, 4.5, 6, 7, 8, 9, 10]$; $y = [1.5, 2, 2.2, 3, 4, 5.5, 6.5, 7]$ 。采用移动最小二乘方法进行拟合的时候, 采用线性基 $p(x) = (1, x, y)^T$, 并采用公式(5)的三次样条函数作为权函数。图5是一般移动最小二乘方法的拟合结果; 图6是通过两节点 (x_3, y_3) , (x_7, y_7) , 即 $(4.5, 2.2)$, $(9, 6.5)$ 的移动最小二乘拟合结果。比较两种方法拟合的误差, 见表1, 其中 x 与 y 分别表示给定离散点的横纵坐标, y_1 与 y_2 分别表示通常情况与插值条件下的移动最小二乘拟合得到的函数值, δ 取的是拟合产生的离差平方和。实验和误差结果发现, 本文给出的带插值条件的移动最小二乘方法整体拟合效果更好。

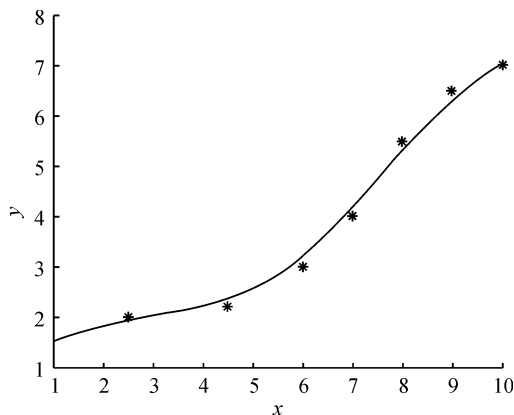


图5 通常情况下移动最小二乘拟合

表 1 拟合误差分析表

x	y	y_1	y_2
1	1.5	1.500 0	1.297 9
2.5	2	1.923 8	1.833 9
4.5	2.2	2.363 6	2.200 0
6	3	3.202 7	3.135 3
7	4	4.173 2	4.177 9
8	5.5	5.301 5	5.457 6
9	6.5	6.310 1	6.500 0
10	7	7.053 8	7.198 0
δ		0.182 0	0.159 4

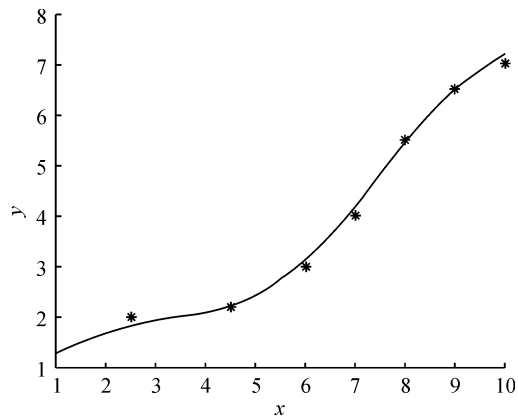


图 6 插值条件下移动最小二乘拟合

3 结 语

本文给出了一种新的带插值条件下最小二乘曲线拟合方法,得到的拟合函数具有次数低、计算方便等优点,并将该方法推广到移动最小二乘方法中,得到了带插值条件的移动最小二乘拟合方法,其数据拟合效果更好。该方法也可推广到带插值条件的最小二乘及移动最小二乘曲面拟合,这些内容将是今后的一个研究方向。

参考文献:

[1] Lancaster P, Salksuskas K. Surfaces generated by moving least squares methods[J]. Mathematics of Computation, 1981, 37(155): 141-158.

[2] Belytschko T, Krongauz Y, Organ D, et al. Meshless method: An overview and recent developments[J]. Computer Methods in Applied Mechanics and Engineering, 1996, 139: 3-47.

[3] Liu H, Shi P. Discontinuity-preserving moving least squares method[C]//Computational and Information Science, Shanghai, 2004: 562-569.

[4] 左传伟, 聂玉峰, 赵美玲. 移动最小二乘法中影响半径的选取[J]. 工程数学学报, 2005, 22(5): 833-838.

[5] 曾清红, 卢德堂. 基于移动最小二乘法的曲线曲面拟合[J]. 工程图学学报, 2004, 25(1): 84-89.

[6] 颜宁生. 带插值条件的最小二乘法[J]. 北京服装学院学报, 2007, 27(2): 42-48.

Moving Least Square Curve Fitting with Interpolation Conditions

NI Hui, LI Zhong, SONG Hong-xing, LI Jing-fang

(School of Sciences, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: This paper presents a method for moving least square curve fitting with interpolation conditions. The problem of least square fitting is first considered. The authors construct a new method for least square fitting with interpolation conditions. It has more advantages : the degree of fitting function is reduced and the computation is convenient. Then, this method is extended to solve the problem of the moving least square fitting with interpolation conditions. The experimental result shows it obtains better fitting effect.

Key words: interpolation; fitting of curves; least square approximation; moving least square approximation

(责任编辑: 马春晓)