

最小二乘分段多项式拟合 及其计算机程序

杨成海 黄振声

(西北农业大学农机系)

摘 要

本文运用数理统计的原理, 阐述了对平面离散数据进行最小二乘分段多项式拟合的数学模型和方法, 估计了拟合曲线的误差, 并编制了相应的通用FORTRAN 计算程序, 能对平面离散数据构成的单值函数曲线和多值函数曲线(在自变量 x 单调变化的转折点, 函数连续而不光滑)进行分段多项式拟合, 并能通过适当地选择曲线分界节点、各段多项式次数和各分界节点处连续的导数的阶数, 满足很多实际问题的需要和所提出的精度要求。

关键词: 分段多项式拟合; 最小二乘; 分界节点; FORTRAN计算程序

在自然科学和生产实践中常常需要从一组观测数据 (x_i, y_i) ($i=1, 2, \dots, n$) 中找到描述自变量 x 和因变量 y 之间关系的数学表达式 $y=f(x)$ 。通常是采用最小二乘法, 以一个普通的函数或样条函数去拟合给定的数据, 从而解决了许多平面数据的拟合问题。尤其是近些年来发展起来的样条函数的最小二乘法已成为曲线拟合中一个十分重要而又非常有效的方法^[1-3]。然而, 这种方法对于分段多项式拟合中各段多项式次数和各段联接处的光滑性有不同要求的一类问题却有一定的局限性。本文的目的就是从数理统计的原理出发, 提出解决这一问题的方法, 并编制相应的通用FORTRAN计算程序。

1 最小二乘分段多项式拟合的数学模型和方法

设有一组观测数据 (x_i, y_i) ($i=1, 2, \dots, n$), 总可以在 $x_1, x_2, \dots, x_n$ 之间适当选取 m_0 个分界节点, 把其分为 m 段, 并使每段之间的 x_i 单调变化。

现把 m_0 个分界节点记为

$$z_1, z_2, \dots, z_{m_0}$$

这些分界节点可以是原来的点, 也可以是新增加的点, 且有

$$m_0 = \begin{cases} m-1, & \text{对开形曲线} \\ m, & \text{对闭形曲线} \end{cases}$$

本文于1986年3月8日收到。

此时, m 段数据可记为

开形曲线:

$$\begin{array}{ccccccc} x_{11}, x_{12}, \dots, x_{1n_1}, z_1 & & & & & & \\ z_1, x_{21}, x_{22}, \dots, x_{2n_2}, z_2 & & & & & & \\ & \dots & & \dots & & \dots & \\ & z_{m-1}, x_{m1}, x_{m2}, \dots, x_{mn_m} & & & & & \end{array}$$

闭形曲线:

$$\begin{array}{ccccccc} z_m, x_{11}, x_{12}, \dots, x_{1n_1}, z_1 & & & & & & \\ z_1, x_{21}, x_{22}, \dots, x_{2n_2}, z_2 & & & & & & \\ & \dots & & \dots & & \dots & \\ & z_{m-1}, x_{m1}, x_{m2}, \dots, x_{mn_m}, z_m & & & & & \end{array}$$

其中 n_j ($j=1, 2, \dots, m$) 为第 j 段中观测数据的个数, 且满足

$$\sum_{j=1}^m n_j = n$$

设第 j 段需要拟合一个 p_j ($p_j \leq n_j - 1$) 阶多项式, 则拟合曲线的数学模型为

$$y_j = b_{j0} + b_{j1}x_j + b_{j2}x_j^2 + \dots + b_{jp_j}x_j^{p_j} + e_j$$

简写成

$$y_j = \sum_{l=0}^{p_j} b_{jl}x_j^l + e_j, \quad j=1, 2, \dots, m \quad (1)$$

其中 b_{jl} ($l=0, 1, 2, \dots, p_j$) 是 p_j+1 个待估的第 j 段多项式的系数, e_j 为服从正态分布 $N(0, \sigma)$ 的随机变量, σ 为未知参数。

将第 j 段中 n_j 个数据代入 (1) 得

$$\begin{array}{l} y_{ji} = \sum_{l=0}^{p_j} b_{jl}x_{ji}^l + e_{ji}, \quad i=1, 2, \dots, n_j \\ j=1, 2, \dots, m \end{array} \quad (2)$$

令

$$Y_j = \begin{pmatrix} y_{j1} \\ y_{j2} \\ \vdots \\ y_{jn_j} \end{pmatrix}, \quad X_j = \begin{pmatrix} 1 & x_{j1} & x_{j1}^2 & \dots & x_{j1}^{p_j} \\ 1 & x_{j2} & x_{j2}^2 & \dots & x_{j2}^{p_j} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{jn_j} & x_{jn_j}^2 & \dots & x_{jn_j}^{p_j} \end{pmatrix}$$

$$b = \begin{pmatrix} b_{j0} \\ b_{j1} \\ \vdots \\ b_{jp_j} \end{pmatrix}_{(p_j+1) \times 1}, \quad \varepsilon = \begin{pmatrix} \varepsilon_{j1} \\ \varepsilon_{j2} \\ \vdots \\ \varepsilon_{jn_j} \end{pmatrix}_{n_j \times 1}$$

则(2)可写成

$$Y_j = X_j b_j + \varepsilon_j, \quad j=1, 2, \dots, m \quad (3)$$

再令

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_m \end{pmatrix}_{n \times 1}, \quad X_j = \begin{pmatrix} X_1 & & 0 \\ & X_2 & \\ & & \ddots \\ 0 & & X_m \end{pmatrix}_{n \times k}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}_{k \times 1}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{pmatrix}_{n \times 1}$$

则(3)式可表示为

$$Y = XB + \varepsilon \quad (4)$$

其中, $k = \sum_{j=1}^m (p_j + 1)$ 为未知参数的个数, ε 为随机向量, 且满足条件

$$E(\varepsilon) = 0, \quad COV(\varepsilon, \varepsilon) = \sigma^2 I_n \quad (5)$$

I_n 是 $n \times n$ 阶单位阵。

由(4), (5)式不难求得

$$E(Y) = XB, \quad COV(Y, Y) = \sigma^2 I_n \quad (6)$$

为使拟合函数及 t 阶导数在中间各分界点连续, 相邻两个多项式在分界点上必须满足以下的约束条件。

在第 j 个分界点上:

$$d \frac{(t_j)}{dx} \left(\sum_{l=0}^{p_j} b_{jl} x_j^l \right) = d \frac{(t_j)}{dx} \left(\sum_{l=0}^{p_{j+1}} b_{(j+1)l} x_{j+1}^l \right) \\ t_j = 0, 1, 2, \dots, q_j; \quad j=1, 2, \dots, m_0 \quad (7)$$

其中

$$q_j = \begin{cases} \leq p_j, & \text{当 } p_j < p_{j+1} \\ < p_j, & \text{当 } p_j = p_{j+1} \\ \leq p_{j+1}, & \text{当 } p_j > p_{j+1} \end{cases}$$

对于闭形曲线, 当 $j = m_0 = m$ 时, $j+1$ 可理解为 1, 以下类同。

若第 j 个分界节点 x_j 为 x 单调变化的转折点, 即 x 由递增(减)变为递减(增)的分界点时, 如果要保证曲线在此点光滑, 相邻两个多项式在该点的导数均为无穷大, 即有 $q_j = 0$ 。

将(7)式展开, 并把 x_j 和 x_{j+1} 用 z_j 代替有

$$\sum_{l=t_j}^{p_j} b_{(j+1)l} \cdot l(l-1)\cdots(l-t_j+1)z_j^{l-t_j} = \sum_{l=t_j}^{p_{j+1}} b_{(j+1)l} \cdot l(l-1)\cdots(l-t_j+1)z_j^{l-t_j}$$

$$t_j = 0, 1, 2, \dots, q_j, \quad j = 1, 2, \dots, m. \quad (8)$$

令

$$Z_{(q_i+1) \times (p_i+1)} = \begin{pmatrix} 1 & z_j & z_j^2 & \cdots & z_j^{q_j} & z_j^{q_j+1} & \cdots & z_j^{p_i-1} & z_j^{p_i} \\ 1 & 2z_j & \cdots & q_j z_j^{q_j-1} & (q_j+1)z_j^{q_j} & \cdots & (p_i-1)z_j^{p_i-2} & p_i z_j^{p_i-1} & p_i z_j^{p_i-1} \\ 2! & \cdots & \frac{q_j!}{(q_j-2)!} z_j^{q_j-2} & \frac{(q_j+1)!}{(q_j-1)!} z_j^{q_j-1} & \cdots & \frac{(p_i-1)!}{(p_i-3)!} z_j^{p_i-3} & \frac{p_i!}{(p_i-2)!} z_j^{p_i-2} & \cdots & \cdots \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ q_j! & \cdots & \frac{(q_j+1)!}{1!} z_j & \cdots & \frac{(p_i-1)!}{(p_i-1-q_j)!} z_j^{p_i-1-q_j} & \frac{p_i!}{(p_i-q_j)!} z_j^{p_i-q_j} \end{pmatrix}$$

$$i = j, j+1$$

则(8)式可表示为

$$Z_{j,j} b_j = Z_{j,j+1} b_{j+1}, \quad j = 1, 2, \dots, m,$$

或写成

$$\begin{bmatrix} Z_{j,j} & -Z_{j,j+1} \end{bmatrix} \begin{bmatrix} b_j \\ b_{j+1} \end{bmatrix} = 0, \quad (9)$$

$$j = 1, 2, \dots, m,$$

再令

对开形曲线:

$$A = \begin{pmatrix} Z_{1,1} & -Z_{1,2} & & & & 0 \\ & Z_{2,1} & -Z_{2,2} & & & \\ & & \ddots & \ddots & & \\ & & & Z_{j,j} & -Z_{j,j+1} & \\ & & & & \ddots & \\ 0 & & & & & Z_{(m-1)(m-1)} & -Z_{(m-1)m} \end{pmatrix}$$

对闭形曲线

$$A = \begin{pmatrix} Z_{11} & -Z_{12} & & & & & 0 \\ & Z_{22} & -Z_{23} & & & & \\ & & \ddots & \ddots & & & \\ & & & Z_{jj} & -Z_{j(j+1)} & & \\ & & & & \ddots & \ddots & \\ 0 & & & & & Z_{(m-1)(m-1)} - Z_{(m-1)m} & \\ Z_{m1} & & & & & & Z_{mm} \end{pmatrix}$$

其中 $r = \sum_{j=1}^{m_0} (q_j + 1)$ 为约束条件数。

则(9)可写成

$$AB = 0 \quad (10)$$

现在, 这个问题的数学模型可表示为

$$\begin{cases} Y = XB + \varepsilon \\ E(Y) = XB, \text{ COV}(Y, Y) = \sigma^2 I_n \\ AB = 0 \end{cases} \quad (11)$$

这是一个带有线性约束条件的线性模型, 可以采用带有约束条件的最小二乘法求得未知参数向量 B 的估计值 $\hat{B}$ 及 σ^2 的估计值 s^2 。

现在的问题就是在约束条件(10)下, 使偏差平方和

$$\|Y - X\hat{B}\|^2 = (Y - X\hat{B})^T (Y - X\hat{B}) \quad (12)$$

为最小。用拉格朗日乘数法组成以下目标函数

$$\Phi(\hat{B}, K) = (Y - X\hat{B})^T (Y - X\hat{B}) + 2K^T (A\hat{B}) \quad (13)$$

这里 K 为 $r \times 1$ 阶拉格朗日乘数向量。

将 Φ 对 $\hat{B}$ 求导, 并令其等于零有

$$\frac{\partial \Phi}{\partial \hat{B}} = 2(Y - X\hat{B})^T (-X) + 2K^T A = 0$$

整理得

$$X^T X \hat{B} + A^T K = X^T Y \quad (14)$$

联合约束条件(10), 即得 $\hat{B}$ 和 K 的法方程为

$$\begin{cases} X^T X \hat{B} + A^T K = X^T Y \\ A \hat{B} = 0 \end{cases} \quad (15)$$

在一般情形下, $|X^T X| \neq 0$, 故 $(X^T X)^{-1}$ 存在, 则由(14)式得

$$\hat{B} = (X^T X)^{-1} X^T Y - (X^T X)^{-1} A^T K \quad (16)$$

在无约束条件时, 上式第二项不存在, 即

$$\hat{B}_0 = (X^T X)^{-1} X^T Y \quad (17)$$

把(16)式求得的 $\hat{B}$ 代入(15)的第二式, 并整理得

$$K = [A(X^T X)^{-1} A^T]^{-1} A(X^T X)^{-1} X^T Y \quad (18)$$

把(18)式代入(16)式得

$$\hat{B} = (X^T X)^{-1} \{X^T Y - A^T [A(X^T X)^{-1} A^T]^{-1} A(X^T X)^{-1} X^T Y\} \quad (19)$$

这就是所求分段多项式的系数向量。

为了确定拟合曲线的精度, 可采用剩余均方误差来衡量, 为此需求得 σ^2 的无偏估计。

求出 $\hat{B}$ 之后, 即可求得 Y 的估计值为

$$\hat{Y}_{n \times 1} = X \hat{B} \quad (20)$$

实测值与估计值的剩余偏差向量为

$$V_{n \times 1} = Y - X \hat{B} \quad (21)$$

由此也可看出拟合模型与观测值拟合的好坏。

记 R^2 为剩余平方和, 可证得

$$R^2 = V^T V = \|Y - X \hat{B}\|^2 = \|Y - X \hat{B}_0\|^2 + \|X(\hat{B}_0 - \hat{B})\|^2 \quad (22)$$

由(17)式有

$$\begin{aligned} \|Y - X \hat{B}_0\|^2 &= [Y - X(X^T X)^{-1} X^T Y]^T [Y - X(X^T X)^{-1} X^T Y] \\ &= Y^T [I_n - X(X^T X)^{-1} X^T] [I_n - X(X^T X)^{-1} X^T] Y \\ &= Y^T [I_n - X(X^T X)^{-1} X^T] Y \end{aligned} \quad (23)$$

由(17)、(19)式, 并注意到 $A \hat{B} = 0$, 有

$$\begin{aligned} \|X(\hat{B}_0 - \hat{B})\|^2 &= [X(\hat{B}_0 - \hat{B})]^T [X(\hat{B}_0 - \hat{B})] \\ &= [A \hat{B}_0]^T [A(X^T X)^{-1} A^T]^{-1} A \hat{B}_0 \end{aligned}$$

$$\begin{aligned}
&= [\hat{A}\hat{B}_0 - AB + AB]^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} [\hat{A}\hat{B}_0 - AB + AB] \\
&= [A(\hat{B}_0 - B)]^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} [A(\hat{B}_0 - B)] \\
&\quad + (AB)^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} AB + 2[A(\hat{B}_0 - B)]^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} AB \\
&= [A(\hat{B}_0 - B)]^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} [A(\hat{B}_0 - B)] \quad (24)
\end{aligned}$$

下面计算 $E(R^2)$, 并求得 σ^2 的无偏估计 s^2 。

$$\begin{aligned}
E(\|Y - X\hat{B}_0\|^2) &= E\{Y^\tau [I_n - X(X^\tau X)^{-1}X^\tau]Y\} \\
&= E(Y^\tau) [I_n - X(X^\tau X)^{-1}X^\tau] E(Y) + \sigma^2 \text{tr}[I_n - X(X^\tau X)^{-1}X^\tau] \\
&= (XB)^\tau [I_n - X(X^\tau X)^{-1}X^\tau] XB + \sigma^2 \{n - \text{tr}[X(X^\tau X)^{-1}X^\tau]\} \\
&= \sigma^2(n - \text{tr}I_k) \\
&= \sigma^2(n - k) \quad (25)
\end{aligned}$$

这里记号 $\text{tr}A = \sum_{i=1}^k a_{ii}$ 为 $k \times k$ 阶矩阵 A 的迹。

$$\begin{aligned}
E[\|X(\hat{B}_0 - \hat{B})\|^2] &= E\{[A(\hat{B}_0 - B)]^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} [A(\hat{B}_0 - B)]\} \\
&= E\{(\hat{B}_0 - B)^\tau A^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} A(\hat{B}_0 - B)\} \\
&= \text{tr}\{A^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} A \cdot \text{COV}(\hat{B}_0, \hat{B}_0)\} \\
&= \text{tr}\{A^\tau [A(X^\tau X)^{-1}A^\tau]^{-1} A(X^\tau X)^{-1}X^\tau \cdot \text{COV}(Y, Y) \\
&\quad X(X^\tau X)^{-1}\} \\
&= \text{tr}\{[A(X^\tau X)^{-1}A^\tau]^{-1} A(X^\tau X)^{-1}A^\tau \sigma^2\} \\
&= \sigma^2 \text{tr}I_r \\
&= \sigma^2 r \quad (26)
\end{aligned}$$

所以有

$$\begin{aligned}
E(R^2) &= E(\|Y - X\hat{B}_0\|^2) + E[\|X(\hat{B}_0 - \hat{B})\|^2] \\
&= \sigma^2(n - k) + \sigma^2 r \\
&= \sigma^2(n - k + r) \quad (27)
\end{aligned}$$

因此

$$s^2 = \frac{R^2}{n - k + r} = \frac{V^\tau V}{n - k + r} \quad (28)$$

是 σ^2 的无偏估计。

由此可得拟合曲线的剩余均方误差为

$$s^2 = \frac{V^T V}{n - k + r} \quad (29)$$

s 越小, 表明拟合曲线精度越高。

2 计算机程序

根据上述算法, 可绘制图 1 所示的计算机程序框图。程序中所用符号与本文计算公式所用符号对照见表 1。

表 1

程序符号	N	M	M ₀	M ₁ (M)	P(M)	Q(M ₀)	N ₁	MA	X ₁ (N)
公式符号	n	m	m ₀	n _j	p _j	q _j	k	r	x _i
程序符号	X(N, N ₁)	Y(N)	Z(M ₀)	A(MA, N ₁)	K(MA)	B(N ₁)	XB(N)	V(N)	S
公式符号	X	y _j	z _j	A	K	B	X _B	V	s

上表中有角标的公式符号表示一组数。

另外, 还有以下三个程序符号:

ISW 控制参数, 如果矩阵求逆成功, **ISW** = 1, 否则, **ISW** = 0, 计算停止。

Y I 函数名, 表示第 I ($I = 1, 2, \dots, M$) 段曲线的函数表达式。

S O 控制参数, 为给定的均方误差, 如果 $S \leq S O$, 计算停止, 否则, 可通过适当地选择曲线分界节点, 各段多项式次数和各段联接点的光滑性条件重新计算, 直到满足所提出的精度要求为止。

根据前述算法及以上程序框图, 我们编制了相应的通用 FORTRAN 计算程序。下面通过一个实例来说明该方法和程序的应用。

表 2 给出了一组观测数据, 要求对其用分段多项式拟合成封闭曲线, 且使均方误差 $s \leq 0.03$ 。

该组数据可用图 2 所示的曲线表示。由图 2 可以看出, 该曲线可看成由曲线段 $\widehat{ADB}$ 、 $\widehat{BC}$ 和直线段 CA 组成的。按照这种分划, 表 3 给出了一些拟合结果。

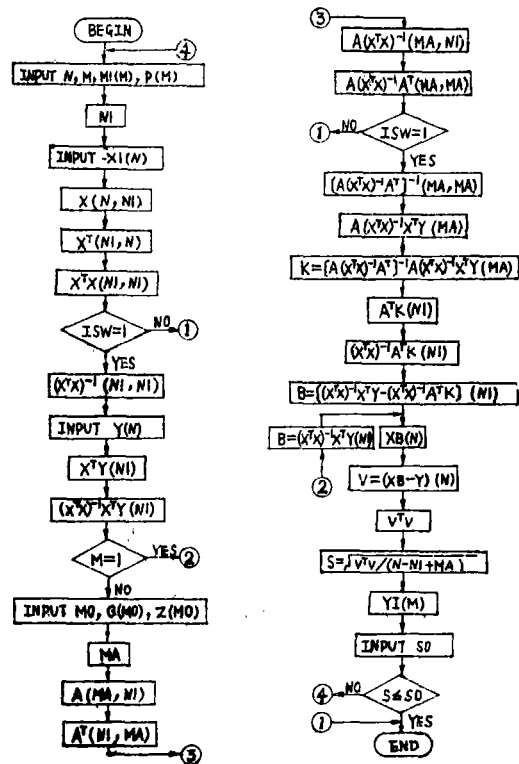


图 1 最小二乘分段多项式拟合计算机程序框图

表2

x	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	9.5	9.0	8.5	7.5	6.5	5.0	3.5	2.0
y	2.9	4.2	4.3	3.6	2.7	1.8	1.2	1.1	1.7	3.1	4.0	4.5	4.8	4.9	4.6	4.1	3.6	3.2

表3

分段数 m	各段数据个数			各段多项式次数			各段联接点连续的 导数的阶数			分界节点			均方误差 s
	n ₁	n ₂	n ₃	p ₁	p ₂	p ₃	q ₁	q ₂	q ₃	z ₁	z ₂	z ₃	
3	10	5	3	3	2	1							0.1907427
				3	3	1							0.1065254
				3	4	1							0.1102145
				4	2	1							0.1800744
				4	3	1							0.02907914
				4	4	1	0	1	0	10.0	6.0	1.0	0.03031500
				5	3	1							0.02608386
				5	4	1							0.02734415
				6	3	1							0.02746335
				7	3	1							0.02764714
				8	3	1							0.02807973
				9	3	1							0.05379601

由表3可以看出,在分界节点和这些点处的光滑性不变的条件下,各段多项式的次数有一个最佳组合,并非多项式次数越高,拟合曲线误差越小。这主要是由于各段多项式相互约束的缘故。对应于表3的分界节点及各联接点的光滑性条件,各段多项式次数的最佳组合为5、3、1,相应的均方误差为 $s=0.02608386$,各段曲线的方程为

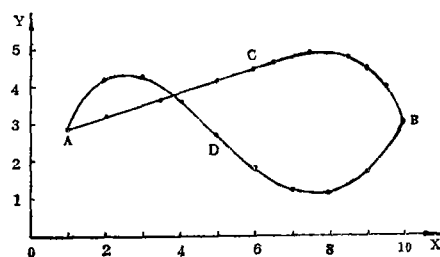


图2

$$y = \begin{cases} -0.3830668 + 4.502626x - 1.382985x^2 + 0.1542102x^3 \\ \quad - 0.7510064 \times 10^{-2}x^4 + 0.1766172 \times 10^{-3}x^5 & (1 \leq x \leq 10) \\ 21.12444 - 8.296527x + 1.321611x^2 - 0.06720433x^3 & (6 \leq x \leq 10) \\ 2.578711 + 0.3047398x & (1 \leq x \leq 6) \end{cases}$$

一般来说,若固定分界节点,各段多项式次数和各联接点光滑性三者中任意两个,而让其中之一在可能的范围内变化,总有一个使拟合曲线误差最小的最佳组合。实际上,在具体计算时,也不一定要找到最佳的组合条件,只要满足所提出的要求就行了。

由于高次多项式缺乏灵合适应数据的能力,虽然在观测点上可以压低平差,但在这些点之间或其外,则往往出现不合理的波动。而且随着多项式次数的提高,在计算时有效数位丢失增多,计算误差增大。因此,在满足要求精度的前提下,可采用低次多项式进行拟合。如上面的问题也可采用各段多项式次数为4、3、1的组合进行拟合。若精度不能满足要求或为了避免由于采用低次多项式而把数据本身的一些起伏波动滤掉,可采用增加分段数的办法来解决。例如上面的问题如果要求 $s \leq 0.026$,且多项式次不能超过四次。我们可以在曲线 $\widehat{ADB}$ 之间再取一分界节点 $z_1 = 5$,把其分为两段,每段用一个四次多项式去拟合,在联接点D保证三阶导数连续,其它节点和联接点的光滑性同前。用各段多项式次数为4, 4, 3, 1的组合可得拟合曲线的均方误差 $s = 0.02566935$ 。

因此,在对某一组数据进行拟合时,在满足实际问题的一些特定要求的条件下,如某段为直线,某点为角点等等,可根据观测数据所构成的近似曲线的变化规律,适当选取分界节点,各段多项式的次数和各段联接点的光滑性条件进行计算,以满足所提出的精度要求。

3 结 论

3.1 本文提出的最小二乘分段多项式拟合的方法,弥补了样条函数最小二乘拟合的一些不足,可以适用于更一般的情形。

3.2 本文所述的曲线拟合方法是建立数理统计基础上的数学方法,拟合曲线的数学模型是一个附有线性约束条件的线性模型,拟合曲线的精度可用均方误差来评价。

3.3 按本文所述的曲线拟合方法编制的计算机程序,能对平面离散数据构成的曲线进行分段多项式拟合,并能通过适当地选择曲线分界节点,各段多项式次数和各分界点的光滑性条件满足很多实际问题的要求和所提出的精度要求。

本文承蒙我校计算机中心王乃信副教授审阅,特此致谢。

参 考 文 献

- 1 李岳生、齐东旭著.样条函数方法.科学出版社,1979
- 2 冯康等著.数值计算方法.国防工业出版社,1978
- 3 李庆海、陶本藻.概率统计原理和在测量中的应用.测绘出版社,1984
- 4 Kendall and Stuart, The advanced theory of statistics, volume 2; Inference and relationship, 4th ed, 1978
- 4 复旦大学编.概率论(第二册), 数理统计(第一分册).人民教育出版社,1979
- 6 陈希孺著.数理统计引论.科学出版社,1981
- 7 李木、杨家全著.FORTRAN 程序设计导引.辽宁科学出版社,1984
- 8 杨振山、金瑞椿编著.FoRTAN 77程序设计语言.同济大学出版社,1985

LEAST SQUARE PIECEWISE POLYNOMIAL FITTING AND ITS PROGRAM

Yang Chenghai Huang zhensheng

(*Department of Farm Machinery, Northwestern Agricultural
University*)

Absrtact

By means of the principles of mathematical statistics, this paper describes a mathematical model and a method of discrete data of plane by least square piecewise polynomial fitting, estimates the errors of the fitting curve, and develops a corresponding current FORTRAN program. This can be used to fit both the single valued function curve and the multiple valued function curve (at the turning points of monotone change of independent x , the function is not smooth but continuous) formed by discrete data of plane. Also, this can meet the needs of many practical problems and suggested precisions through properly selecting nodal points of demarcation of the curve and the degree of polynomial of every piece as well as the order of continuous derivative at every nodal points of demarcation.

Key words: piecewise polynomial fitting, least square, nodal points of demarcation, FORTRAN program