计算方法

第一章 插值方法

胡 敏
合肥工业大学 计算机与信息学院
jsjxhumin@hfut.edu.cn
uhnim@163.com

2015年2月

第 1 章　插值方法

1.6　埃尔米特插值

1.7 分段插值法

1.8 样条函数

1.9 曲线拟合的最小二乘法

1.6 埃尔米特（Hermite）插值

在某些问题中，为了保证插值函数能更好地密合原来的函数，不但要求“过点”，即两者在节点上具有相同的函数值，而且要求“相切”，即在节点上还具有相同的导数值，这类插值称之为切触插值，或称为埃尔米特（Hermite）插值，这是泰勒插值和拉格朗日插值的综合和推广。

一、回忆

1、泰勒插值，特点：多项式；插值条件：

$$p_n^{(i)}(x_i) = y_i^{(i)}$$

2、拉格朗日插值，特点：多项式；插值条件：

$$p_n(x_i) = y_i$$

二、Hermite插值

1、二次插值

可能有的插值条件

条件1:

$$p_2(x_0) = y_0, p_2'(x_0) = y_0', p_2(x_1) = y_1$$

条件2:

$$p_2(x_0) = y_0, p_2'(x_1) = y_1', p_2(x_1) = y_1$$

2、问题5：求作二次多项式，满足

$$p_2(x_0)=y_0,\ p_2'(x_0)=y_0',\ p_2(x_1)=y_1$$

设用这一插值函数 $p_2(x)$ 逼近某个取值为 $f(x_0)=y_0$ $f'(x_0)=y_0', f(x_1)=y_1$ 的函数 $f(x)$，那么，从图形上看，曲线 $y=p_2(x)$ 与 $y=f(x)$ 不但有两个交点 (x_0,y_0)，(x_1,y_1)，而且在点(x_0,y_0)处两者还相切。

3、问题5的求解

方法1：基于承袭性方法

条件为

$$p_2(x_0)=y_0,\ p_2'(x_0)=y_0',\ p_2(x_1)=y_1$$

如何确定**c**？

$$p_2(x)=p_1(x)+c(x-x_0)(x-x_1)$$

线性**Lagrange**插值多项式

$$=y_0+\frac{y_1-y_0}{x_1-x_0}(x-x_0)$$

$$+\frac{1}{x_1-x_0}\left(\frac{y_1-y_0}{x_1-x_0}-y_0'\right)(x-x_0)(x-x_1)$$

方法2：用基函数方法，取下列插值函数

$$p_2(x) = y_0\varphi_0(x) + y_1\varphi_1(x) + y'_0\psi_0(x)$$

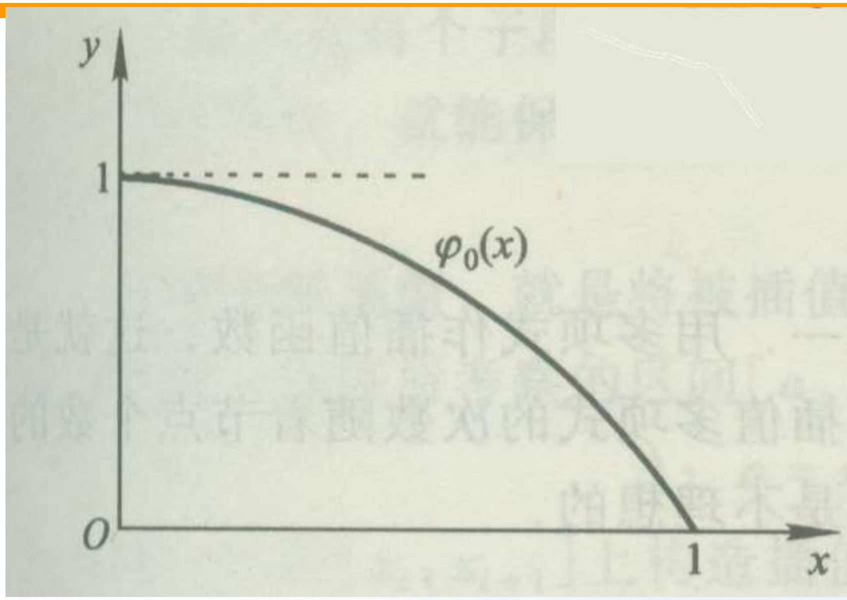
其中$\varphi_0(x), \varphi_1(x), \phi_0(x)$是二次函数，满足下列条件：

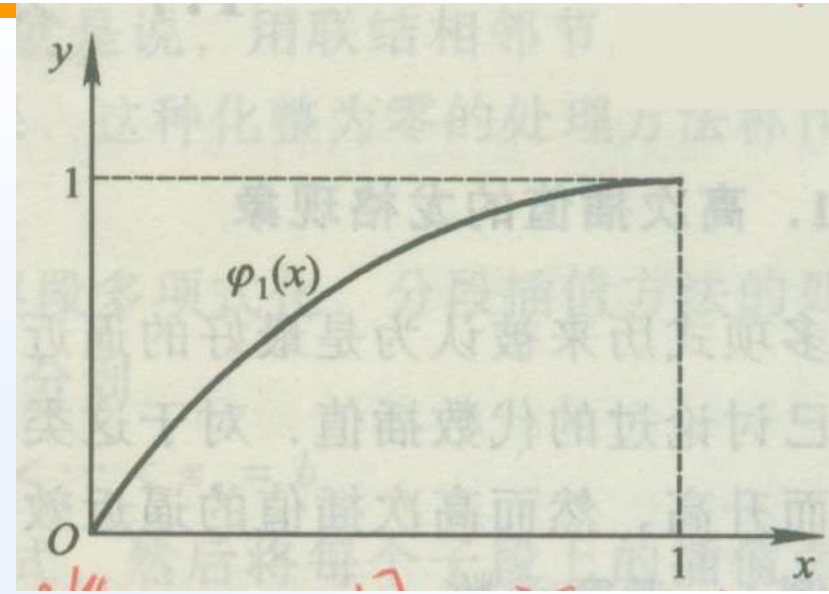
$$\begin{cases} \varphi_0(0) = 1, \varphi_0(1) = \varphi_0'(0) = 0 \\ \varphi_1(1) = 1, \varphi_1(0) = \varphi_1'(1) = 0 \\ \psi_0^1(0) = 1, \psi_0(0) = \psi_0(1) = 0 \end{cases}$$

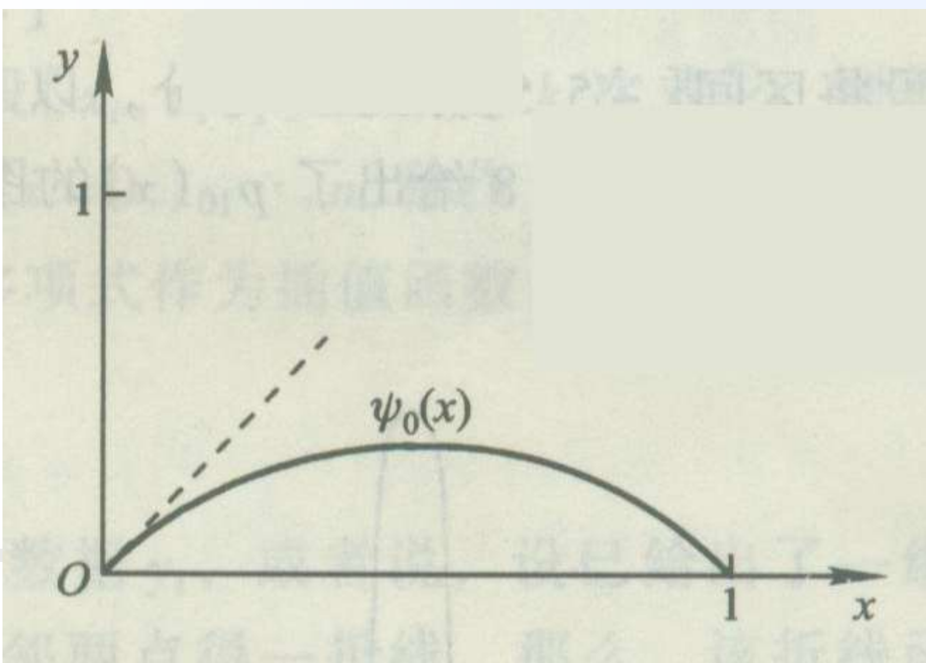
两点三次Hermite插值

如果插值节点为0，1的基函数为：

$$\begin{cases} \varphi_0(x) = 1 - x^2 \\ \varphi_1(x) = 2x - x^2 \\ \psi_0(x) = x(1-x) \end{cases}$$


y
1
$\varphi_0(x)$
O
1
x


y
1
$\varphi_1(x)$
O
1
x


y
1
$\psi_0(x)$
O
1
x

类似地，对于插值节点为 x_0, x_1 的基函数为：

$$p_2(x) = y_0\varphi_0\left(\frac{x-x_0}{h}\right) + y_1\varphi_1\left(\frac{x-x_0}{h}\right) + hy'_0\,\phi_0\left(\frac{x-x_0}{h}\right)$$

注意：通过如下变换可得到上述插值公式。

$$[x_0, x_1] \xrightarrow{\text{变换}} [0,1]$$

$$t = \frac{x-x_0}{h}, h = x_1 - x_0$$

4、高次插值

仿照上述类似地求其插值多项式！

问题6：求作三次式 $p_3(x)$满足插值条件

$$p_3(x_0)=y_0 \qquad p_3(x_1)=y_1$$

$$p_3'(x_0)=y_0' \qquad p_3'(x_1)=y_1'$$

$p_3(x)$应用四个插值基函数表示　记$h=x_1-x_0$

设$p_3(x)$的插值基函数为$\varphi_i(x),\psi_i(x),i=0,1$

$$p_3(x)=y_0\varphi_0\left(\frac{x-x_0}{h}\right)+y_1\varphi_1\left(\frac{x-x_0}{h}\right)+hy_0'\psi_0\left(\frac{x-x_0}{h}\right)+hy_1'\psi_1\left(\frac{x-x_0}{h}\right)$$

(28)

希望插值系数与Lagrange插值一样简单

其中 $\varphi_0(x), \varphi_1(x), \psi_0(x), \psi_1(x)$ 均为代数多项式，且

$$\begin{cases}\varphi_0(x_0)=1\\ \varphi_0(x_1)=0\\ \varphi_0'(x_0)=0\\ \varphi_0'(x_1)=0\end{cases},\quad \begin{cases}\varphi_1(x_0)=0\\ \varphi_1(x_1)=1\\ \varphi_1'(x_0)=0\\ \varphi_1'(x_1)=0\end{cases},\quad \begin{cases}\psi_0(x_0)=0\\ \psi_0(x_1)=0\\ \psi_0'(x_0)=1\\ \psi_0'(x_1)=0\end{cases},\quad \begin{cases}\psi_1(x_0)=0\\ \psi_1(x_1)=0\\ \psi_1'(x_0)=0\\ \psi_1'(x_1)=1\end{cases}$$

可知 x_1是$\varphi_0(x)$的二重零点,即可假设

$$\varphi_0(x)=(x-x_1)^2(ax+b)$$

由 $\varphi_0(x_0)=1$ $\varphi_0'(x_0)=0$

可得 $a = -\dfrac{2}{(x_0 - x_1)^3}$ $b = \dfrac{1}{(x_0 - x_1)^2} + \dfrac{2x_0}{(x_0 - x_1)^3}$

$$\varphi_0(x) = (x - x_1)^2 (ax + b)$$

$$= (x - x_1)^2 \left(-\frac{2x}{(x_0 - x_1)^3} + \frac{1}{(x_0 - x_1)^2} + \frac{2x_0}{(x_0 - x_1)^3} \right)$$

$$= \frac{(x - x_1)^2}{(x_0 - x_1)^2} \left(1 + \frac{2x_0}{x_0 - x_1} - \frac{2x}{x_0 - x_1} \right)$$

Lagrange
插值基函数

$$= \left(1 + 2\frac{x - x_0}{x_1 - x_0} \right) \left(\frac{x - x_1}{x_0 - x_1} \right)^2 = (1 + 2l_1(x)) \cdot l_0^2(x)$$

即 $$\varphi_0(x) = (1+2l_1(x))\cdot l_0^2(x) = \left(1+2\frac{x-x_0}{x_1-x_0}\right)\left(\frac{x-x_1}{x_0-x_1}\right)^2$$

类似可得

$$\varphi_1(x) = (1+2l_0(x))\cdot l_1^2(x) = \left(1+2\frac{x-x_1}{x_0-x_1}\right)\left(\frac{x-x_0}{x_1-x_0}\right)^2$$

$$\psi_0(x) = (x-x_0)\cdot l_0^2(x) = (x-x_0)\left(\frac{x-x_1}{x_0-x_1}\right)^2$$

$$\psi_1(x) = (x-x_1)\cdot l_1^2(x) = (x-x_1)\left(\frac{x-x_0}{x_1-x_0}\right)^2$$

如果插值节点为0，1的基函数为：

$$\varphi_0(x)=\left(1+2\frac{x-x_0}{x_1-x_0}\right)\left(\frac{x-x_1}{x_0-x_1}\right)^2=(x-1)^2(2x+1)$$

$$\varphi_1(x)=\left(1+2\frac{x-x_1}{x_0-x_1}\right)\left(\frac{x-x_0}{x_1-x_0}\right)^2=x^2(-2x+3)$$

$$\psi_0(x)=(x-x_0)\left(\frac{x-x_1}{x_0-x_1}\right)^2=x(x-1)^2$$

$$\psi_1(x)=(x-x_1)\left(\frac{x-x_0}{x_1-x_0}\right)^2=x^2(x-x_1) \qquad (29)$$

5、Hermite插值余项

$$f(x)-p_2(x)=\frac{f'''(\xi_1)}{3!}(x-x_0)^2(x-x_1)$$

$$f(x)-p_3(x)=\frac{f^{(4)}(\xi_2)}{4!}(x-x_0)^2(x-x_1)^2$$

(30)

小 结

- 在节点一定的条件下，可以构造多种插值条件；
- 埃尔米特插值具有少节点得到高次插值多项式的特点；
- 插值多项式灵活多样；
- 构造插值多项式的过程：注意算法的承袭性，并使用Lagrange插值多项式和待定系数。

1.7 分段插值法

一、为何要进行分段低次插值

1、是否是插值多项式的次数越高，精度就越高？

2、龙格(**Runge**)现象：在端点发生激烈的震荡现象！

3、解决问题的方法:分段插值.

1、是否是插值多项式的次数越高，精度就越高？

来讨论一下这个问题。

我们已经知道：f(x)在n+1个节点xi(i=0，1，2，…，n) 上的n次插值多项式Pn (x) 的余项

$$R(x)=f(x)-P_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}\prod_{i=0}^{n}(x-x_i)$$

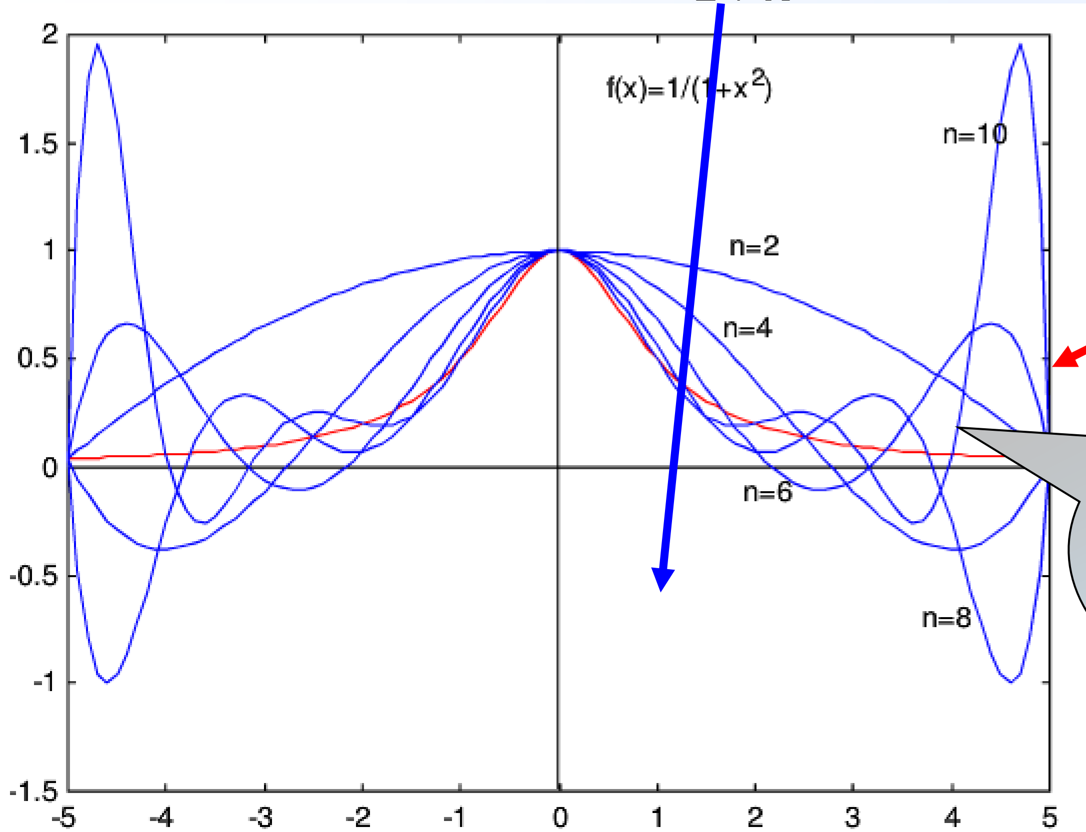
设想当节点数增多时会出现什么情况。由插值余项可知，当f(x)充分光滑时，余项随n增大而趋于0的，这说明可用增加节点的方法达到这个目的，那么实际是这样吗？

2、龙格Runge现象

随着插值多项式次数的增大以及逼近区间的增大，使得在逼近区间发生振荡现象。从而使得逼近效果不理想（**龙格Runge现象**）。从如图 1 – 8 可以看出这一点。

龙格Runge现象

例：在[−5, 5]上考察 $f(x)=\dfrac{1}{1+x^2}$ 的$L_n(x)$。取 $x_i=-5+\dfrac{10}{n}i \quad (i=0,\dots,n)$



$L_n(x) \not\to f(x)$

$L_{10}(x)$

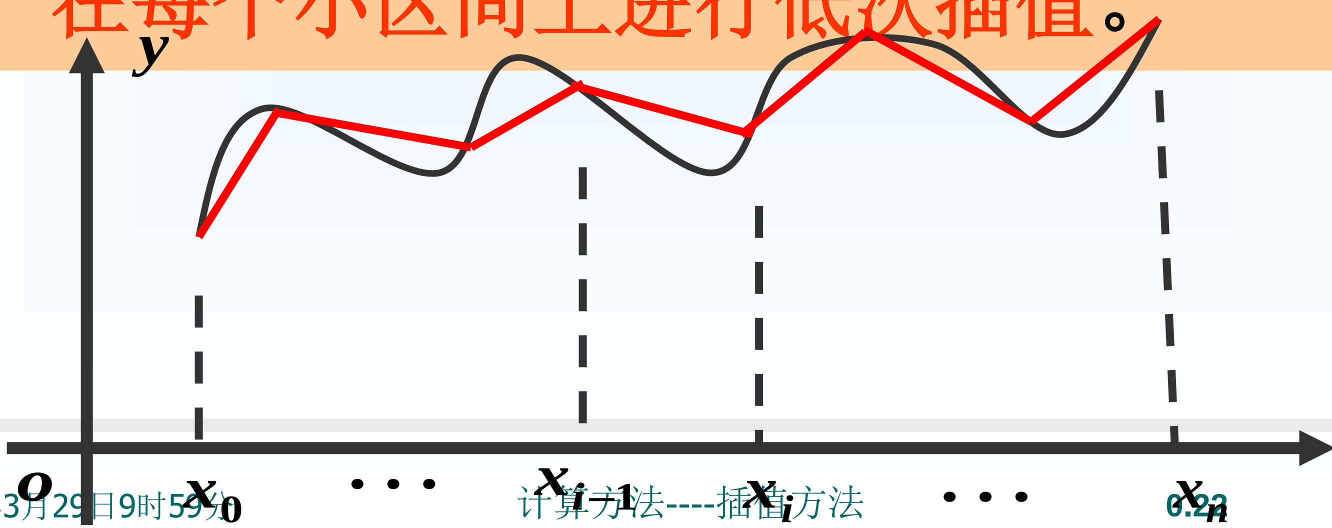
n 越大，
端点附近抖动
越大，称为
Runge 现象

Runge现象

这种插值多项式当节点增加时反而不能更好地接近被插之数的现象，称为龙格现象。

插值误差除来自截断误差外，还来自初始数据的误差和计算过程中的舍入误差。插值次数越高，计算工作量越大，积累误差也可能越大。

因此，在实际操作过程中，常常用**分段低次插值**进行计算，**即把整个插值区间分成若干个小区间，在每个小区间上进行低次插值。**



二、分段插值的基本概念

1、分段插值

就是将被插值函数逐段多项式化。

2、基本方法

•分划: $\Delta: a = x_0 < x_1 < \cdots < x_n = b$

并在每个子段 $[x_i, x_{i+1}]$上构造插值多项式

•将每个子段上的插值多项式组合在一起，作为整个区间 $[a,b]$上的插值函数。这样构造的插值多项式就是分段插值多项式。

如果函数 $S_k(x)$ 在分划 Δ 的每个子段上 $[x_i, x_{i+1}]$ 都是 k 次式，则称 $S_k(x)$ 为具有分划 Δ 的分段 k 次式。点 $x_i(i=0,1,\cdots,n)$ 称作 $S_k(x)$ 的节点。

三、分段线性插值

1、问题7　求作具有分划 Δ 的分段一次 $S_1(x)$，使成立

$$S_1(x_i)=y_i,\ (i=0,1,\cdots,n)$$

解： 注意到每个子段 $[x_i,x_{i+1}]$上 $S_1(x)$ 都是一次式，且成立 $S_1(x_i)=y_i$，$S_1(x_{i+1})=y_{i+1}$，知

$$S_1(x)=\varphi_0\left(\frac{x-x_i}{h_i}\right)y_i+\varphi_1\left(\frac{x-x_i}{h_i}\right)y_{i+1}$$

式中 $h_i=x_{i+1}-x_i$,而

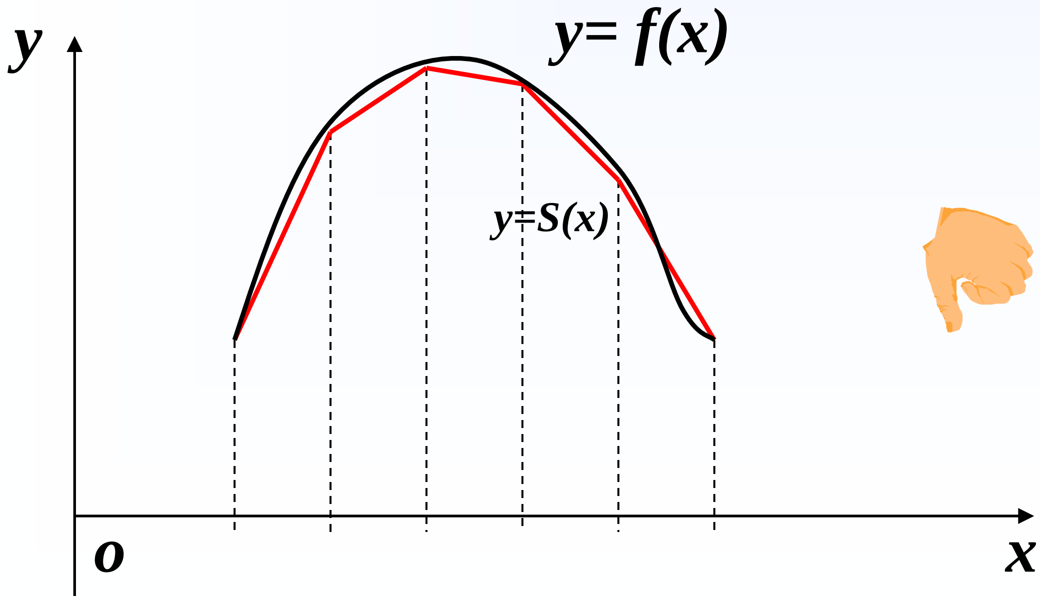
$$\varphi_0(x)=1-x,\varphi_1(x)=x$$

满足
$$\begin{cases}\varphi_0(0)=1,\varphi_0(1)=0\\ \varphi_1(0)=0,\varphi_1(1)=1\end{cases}$$

分段线性插值

在每个区间$[x_i, x_{i+1}]$上，用**1阶多项式** (直线) 逼近 $f(x)$:

$$f(x) \approx S_1(x) = \frac{x - x_{i+1}}{x_i - x_{i+1}} y_i + \frac{x - x_i}{x_{i+1} - x_i} y_{i+1} \qquad x \in [x_i, x_{i+1}]$$



失去了原函数的光滑性。

2、误差估计

定理5

设 $f(x) \in C^2[a,b], f(x_i)=y_i \quad (i=0,1,\cdots,n)$ 已给出，则当 $x \in [a,b]$ 时，对于问题7的解 $S_1(x)$ 成立

$$|f(x)-S_1(x)| \le \frac{h^2}{8}\max_{a\le x\le b}|f''(x)|$$

式中 $h=\max\limits_i h_i$ ，由此得知，$S_1(x)$ 在 $[a,b]$ 上一致收敛到 $f(x)$。

分析：$$|f(x)-S_1(x)| = \left|\frac{f''(\xi)}{2}(x-x_i)(x-x_{i+1})\right| \le \frac{h^2}{8}\max_{a\le x\le b}|f''(x)|$$

分段线性插值的余项

$$|f(x)-s_1(x)| \le \max_{x_i \le x \le x_{i+1}} |f(x)-s_1(x)|$$

$$\le \left| \max_{x_i \le x \le x_{i+1}} \frac{f''(\xi_i)}{2!}(x-x_i)(x-x_{i+1}) \right|$$

$$\le \frac{\max_{x_i \le x \le x_{i+1}} |f''(x)|}{2!} \left(\frac{x_j - x_i}{2} \right)^2 \le \frac{h_i^2}{8} \max_{x_i \le x \le x_{i+1}} |f''(x)|$$

例：已知函数$f(x)=\dfrac{1}{1+x^2}$在区间[0,5]上取等距插值节点（如下表），求区间[0,5]上分段线性插值函数，并利用它求出$f(4.5)$的近似值。

x_i	0	1	2	3	4	5
y_i	1	0.5	0.2	0.1	0.05882	0.03846

解：在每个小区间$[i,i+1]$上：

$$S_1(x)=y_i\varphi\left(\frac{x-x_i}{h_i}\right)+y_{i+1}\varphi\left(\frac{x-x_i}{h_i}\right)$$

$$=y_i\left(1-\frac{x-x_i}{h_i}\right)+y_{i+1}\left(\frac{x-x_i}{h_i}\right)=y_i(1-x+i)+y_{i+1}(x-i)$$

$$
s(x)=\begin{cases}(1-x)+0.5x,\ x\in[0,1]\\ 0.5(2-x)+0.2(x-1),\ x\in[1,2]\\ 0.2(3-x)+0.1(x-2),\ x\in[2,3]\\ 0.1(4-x)+0.05882(x-2),\ x\in[3,4]\\ 0.05882(5-x)+0.03846(x-4),\ x\in[4,5]\end{cases}
$$

$$
P(4.5)\approx 0.05882(5-4.5)+0.03846(4.5-4)
$$

$$
\approx 0.04864
$$

四、分段三次Hermite插值

1、问题8 求作具有分划 Δ 的分段三次式 $S_3(x)$，使成立

$$S_3(x_i)=y_i,\ S_3'(x_i)=y_i',\quad (i=0,1,\cdots,n)$$

解： 注意到每个子段 $[x_i,x_{i+1}]$ 上 $S_3(x)$ 都是三次式，且成立 $S_3(x_i)=y_i$， $S_3'(x_i)=y_i'$， $S_3(x_{i+1})=y_{i+1}$， $S_3'(x_{i+1})=y_{i+1}'$，根据Hermite插值方法，由（28）式知

$$S_3(x)=\varphi_0\left(\frac{x-x_i}{h_i}\right)y_i+\varphi_1\left(\frac{x-x_i}{h_i}\right)y_{i+1}+h_i\psi_0\left(\frac{x-x_i}{h_i}\right)y_i'+h_i\psi_1\left(\frac{x-x_i}{h_i}\right)y_{i+1}'$$

(31)

式中$x_i \le x \le x_{i+1}$，根据Hermite插值基函数（29）式，有

$$\varphi_0(x)=(x-1)^2(2x+1)$$

$$\varphi_1(x)=x^2(-2x+3)$$

$$\psi_0(x)=x(x-1)^2 \qquad (32)$$

$$\psi_1(x)=x^2(x-1)$$

2、误差估计

定理 6

设 $f(x)\in C^4[a,b],\ f(x_i)=y_i,\ f'(x_i)=y_i'\ \ (i=0,1,\cdots,n)$ 已给出，则当 $x\in[a,b]$时，对于问题8的解 $S_3(x)$成立

$$\left|f(x)-S_3(x)\right|\le\frac{h^4}{384}\max_{a\le x\le b}\left|f^{(4)}(x)\right|$$

式中 $h=\max\limits_i h_i$ ，由此得知，$S_3(x)$在 $[a,b]$ 上一致收敛到 $f(x)$。

关于整体误差，若 $f(x) \in C^4[a,b]$, 则可按如下方式考虑：

$$|R(x)| = |f(x) - H(x)| \le \max_{1\le i\le n} |R_i(x)|$$

$$\le \max_{1\le i\le n}\left(\frac{M_i}{384} h_i^4\right) \le \frac{1}{384} \max_{1\le i\le n} M_i \max_{1\le i\le n} h_i^4$$

记 $M = \max_{1\le i\le n} M_i$ $\quad h = \max_{1\le i\le n} h_i$ $\quad M_i = f^{(i)}(x)$

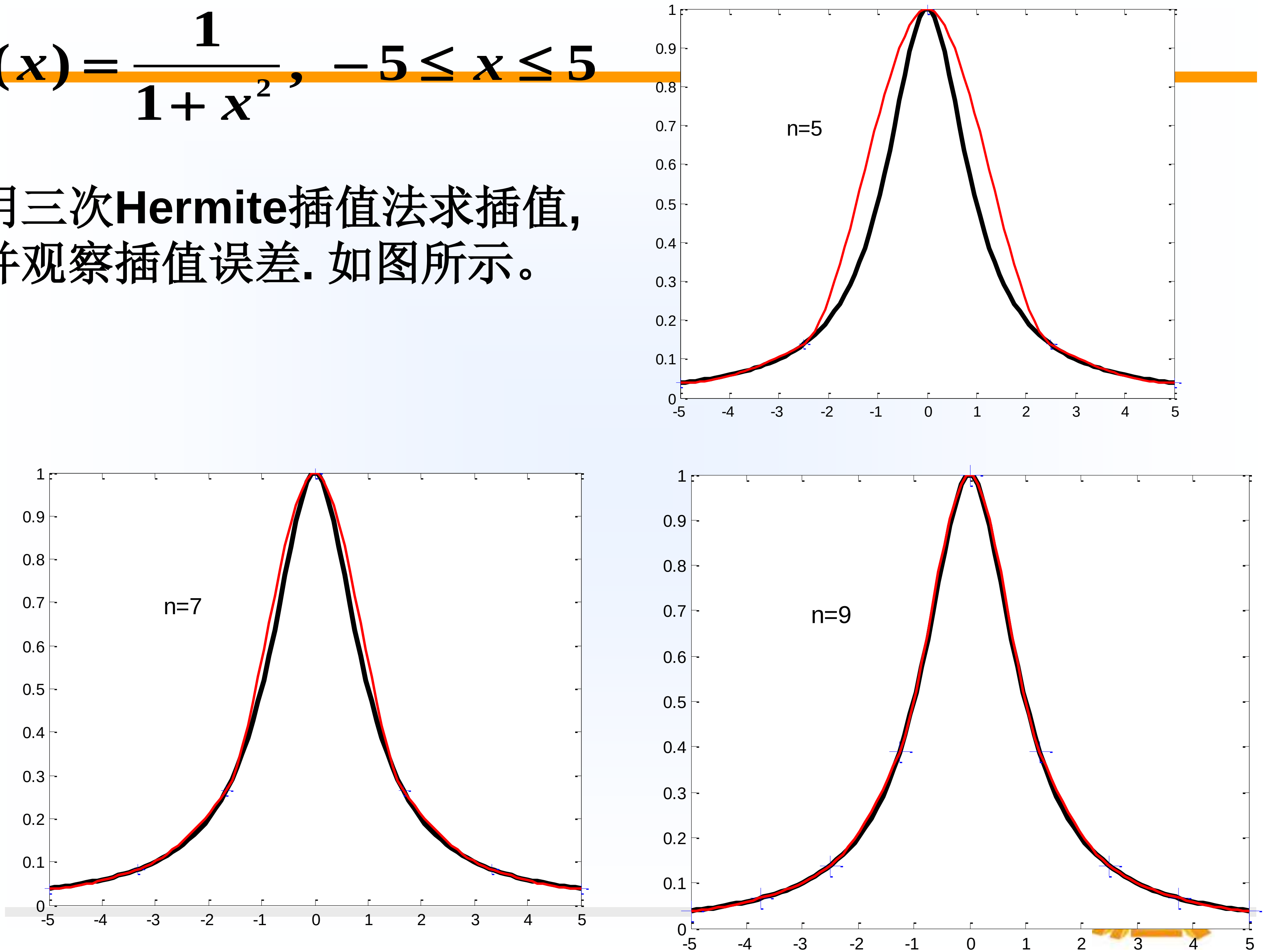
则有： $|R(x)| \le \dfrac{M}{384} h^4$

于是,当$h\to 0$ 时，$R(x)\to 0$.说明分段三次Hermite插值 $H(x)$ 收敛于$f(x)$ 。

例：已知函数

$$g(x)=\frac{1}{1+x^2},\quad -5\le x\le 5$$

用三次Hermite插值法求插值，并观察插值误差. 如图所示。



本节问题

1. 什么是高次插值的Runge 现象，应如何避免？

2. 分段线性插值有何优缺点？如何估计误差？

3. 分段三次Hermite插值有何优缺点, 如何估计误差

$$R_i(x)=f(x)-H_i(x)=\frac{f^{(4)}(\xi_i)}{4!}(x-x_{i-1})^2(x-x_i)^2$$

4. 分段线性插值算法的程序如何设计？

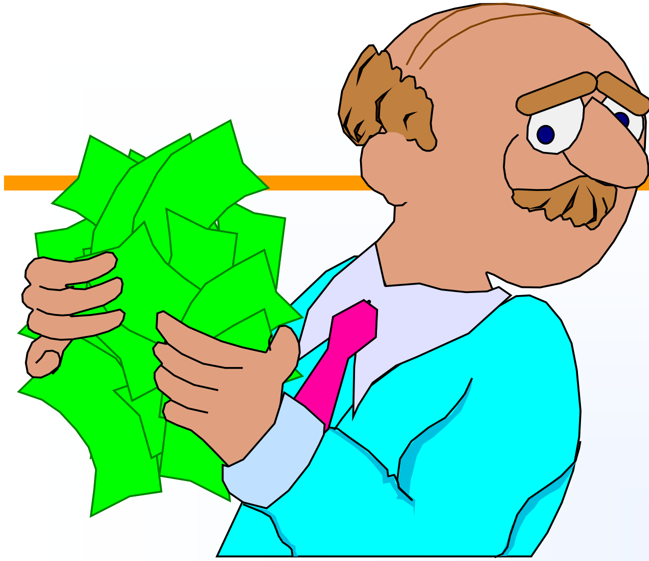
5. 如何构造满足以下条件的插值多项式并估计误差？

x_i	x_0	x_1	x_2
$f(x_i)$	y_0	y_1	y_2
$f'(x_i)$	y_0'		

分段插值的优缺点

1、优点：显示算法，方法简单，收敛性好，只要节点距离充分小，分段插值总能达到所要的精度要求，而不会象高次插值那样发生龙格现象。另一个重要特点就是局部性质。如果修改某个数据，那么插值曲线仅仅在某个局部范围受到影响，而代数插值却会影响到整个插值区间。

2、缺点：分段线性插值与分段三次埃尔米特插值（问题8）虽然改善了精度，但是这种插值要求给出各个节点上的导数值，所要提供的信息太多，同时它的光滑性也不高（只有连续的一阶导数）。改进这种插值以克服其缺点，这就是下一节介绍的三次样条插值方法（问题）。



实际上，上面介绍的分段低次插值，虽然具有计算简便，收敛性有保证，数值稳定性又好且易在计算机上实现等优点，但它却不能保证整条曲线的光滑性，从而不能满足某些工程技术上的要求，从六十年代开始，首先由于航空、造船等工程设计的需要而发展起来的样条插值（spline)方法，既保留了分段低次插值的各种优点，又提高了插值函数的光滑性，在许多领域显得越来越广泛的应用。

要求：插值曲线既要简单，又要在曲线的连接处比较光滑。

这样的分段插值函数在分段上要求多项式次数低，而在节点上不仅连续，还存在连续的低阶导数，

把满足这样条件的插值函数，称为样条插值函数，

它所对应的曲线称为样条曲线，其节点称为样点，

这种插值方法称为——样条插值。

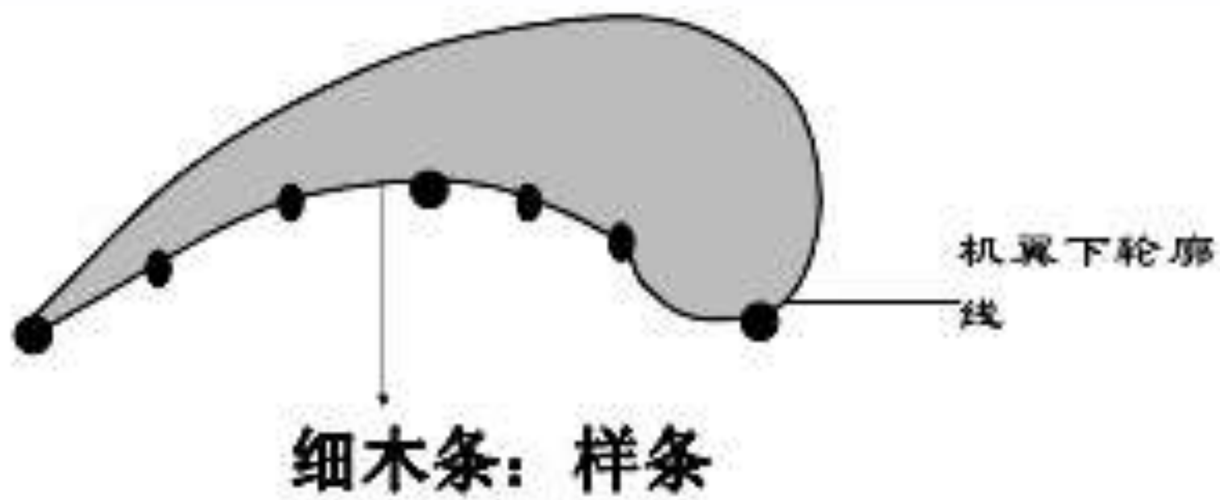


图1.1 早期机翼下轮廓的放样

如图2.1所示，在早期的板材曲线切割时，常把富有弹性的细长木条（样条）固定在样点上，其它地方让其自由弯曲，然后画出长条的曲线称为样条曲线，由此启发设计整体连续光滑的样条插值函数。

1.8 样条函数插值

一、 样条函数的概念

1.定义 具有分划Δ：$a = x_0 < x_1 < \cdots < x_n = b$ 的分段k次式 $S_k(x)$为k次样条，如果它在每个结点 $x_i(i = 1, 2, \cdots n-1)$ 上具有直到k-1阶连续导数.

点 $x_i(1 \le i \le n-1)$则称作样条函数 $S_k(x)$的节点。

特点：光滑性即外形美观，间断性则使它能转折自如，即灵活。

$k-1$

2、三次样条插值

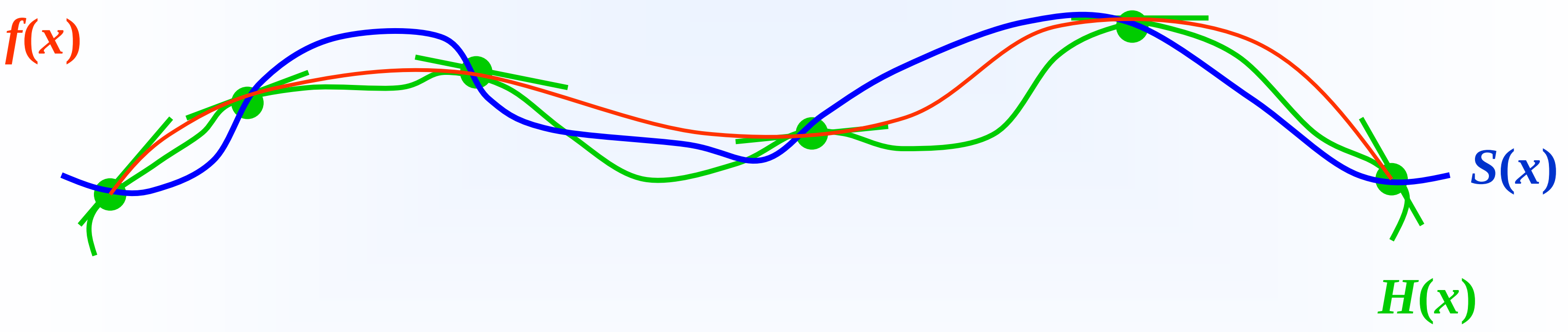
特别地，称 $S_3(x)$为具有分划 Δ 的三次样条，如果它在分划 Δ 的每个小区间 $[x_i, x_{i+1}]$上都是三次式，且在每个内结点 $x_i\ , i=1\sim n-1$ 上具有连续的二阶导数，即成立

$$\begin{cases} S_3(x_i-0)=S_3(x_i+0) \\ S_3'(x_i-0)=S_3'(x_i+0) \\ S_3''(x_i-0)=S_3''(x_i+0) \end{cases} ,\quad i=1,2,\cdots,n-1$$

如果 $S_3(x_i) = f(x_i)\ (i = 0, 1, 2,\ldots, n)$，

则称$S_3(x)$为$y = f(x)$的三次样条插值函数。

注：三次样条与分段 Hermite 插值的根本区别在于 $S(x)$自身光滑，不需要知道 f 的导数值（除了在2个端点可能需要）；而Hermite插值依赖于f 在所有插值点的导数值。



三次样条插值的实质与特点

实质：分段插值。

特点：插值函数具有二阶连续导数。

三次样条插值函数的求解

分析：

1、区间 $[x_k, x_{k+1}]$ 上的三次多项式，共需待定系数 $4n$ 个。

2、已知条件有

$S(x_j) = f(x_j), \quad j = 0,1,2,\cdots,n,$ $\quad n+1$ 个

$S(x_j - 0) = S(x_j + 0), \quad j = 1,2,\cdots,n-1,$ $\quad n-1$ 个

$S'(x_j - 0) = S'(x_j + 0), \quad j = 1,2,\cdots,n-1,$ $\quad n-1$ 个

$S''(x_j - 0) = S''(x_j + 0), \quad j = 1,2,\cdots,n-1,$ $\quad n-1$ 个

共计 $4n-2$ 个

解决的办法：引入边界条件。

边界条件：在确定三次样条插值函数时，所缺少的两个条件由插值区间$[a, b]$的边界点a、b处给出，这个条件通常被称为边界条件。

边界条件的类型

- (1) 已知一阶导数值；
- (2) 已知二阶导数值；
- (3) 被逼近函数是周期函数。

1、 $S'(x_0)=f'(x_0),\quad S'(x_n)=f'(x_n)$

边界条件

2、 $S''(x_0)=f''(x_0),\quad S''(x_n)=f''(x_n)$

特别 $S''(x_0)=S''(x_n)=0$

称自然边界条件

3、 周期性条件

$$S'(x_0+0)=S'(x_n-0),$$
$$S''(x_0+0)=S''(x_n-0)$$

问题9 求作具有分划Δ 的三次样条 $S_3(x)$，使满足

$$S_3(x_i) = y_i \ , \quad i = 0,1,\cdots,n$$

$$S_3'(x_0) = y_0' \ , \quad S_3'(x_n) = y_n'$$

求三次样条插值函数的基本思想：先利用一阶（或二阶）导数 $S'(x)$ ($S''(x)$) 在内节点 $x_i\,(i=1,2,\cdots,n-1)$ 上的连续性以及边界条件，列出确定二阶（一阶）导数(例如：$m_i = S'(x_i)(i=0,1,2,\cdots,n-1)$) 的线性方程组，并由此解出 m_i，然后用 m_i 来表达 $S_3(x)$.

1. 用一阶导数表示的三次样条

设 $S_3'(x_i)=m_i=f'(x_i),\ i=0\sim n$ ，则在 $[x_i,x_{i+1}]$ 上插值条件为

$$S_3(x_i)=y_i,\ S_3'(x_i)=m_i\ ,S_3(x_{i+1})=y_{i+1},\ S_3'(x_{i+1})=m_{i+1}$$

问题：关键在于参数导数值的选择。

方法：样条函数的构造用待定系数法。

三次样条插值函数为:

$$s_3(x)=\varphi_0\left(\frac{x-x_i}{h_i}\right)y_i+\varphi_1\left(\frac{x-x_i}{h_i}\right)y_{i+1}$$

$$+h_i\psi_0\left(\frac{x-x_i}{h_i}\right)m_i+h_i\psi_1\left(\frac{x-x_i}{h_i}\right)m_{i+1} \tag{35}$$

其中 $h_i=x_{i+1}-x_i,x\in[x_i,x_{i+1}]$，而

$$\varphi_0(x)=(x-1)^2(2x+1),\varphi_1(x)=x^2(-2x+3)$$

$$\psi_0(x)=x(x-1)^2,\qquad \psi_1(x)=x^2(x-1)$$

不论如何确定参数 m_i ，这样构造出的三次样条插值函数在每个节点 x_i 上均连续且有连续的一阶导数，现在的问题是如何确定参数 m_i 使其二阶导数也连续。对 $s_3(x)$ 求两次导数，并计算在子区间 $[x_i, x_{i+1}]$ 的端点上的导数值有

$$s_3''(x_i) = 6\frac{y_{i+1} - y_i}{h_i^2} - \frac{4m_i + 2m_{i+1}}{h_i} \qquad (36)$$

$$s_3''(x_{i+1}) = -6\frac{y_{i+1} - y_i}{h_i^2} - \frac{2m_i + 4m_{i+1}}{h_i} \qquad (37)$$

为了保证二阶导数的连续性，要求成立

$$s_3''(x_i-0)=s_3''(x_i+0), \qquad (38) \quad (i=1,2,\cdots,n-1)$$

即要求（36）与（37）相容，即把（37）式中的 i+1 改写为 i，i 改写为 i-1，因而有

$$s_3''(x_i)=-6\frac{y_i-y_{i-1}}{h_{i-1}^2}-\frac{2m_{i-1}+4m_i}{h_{i-1}} \qquad (37')$$

把（37'）和（36）式代入（38），有

$$\frac{m_{i-1}+2m_i}{h_{i-1}}+\frac{2m_i+m_{i+1}}{h_i}=3\left(\frac{y_i-y_{i-1}}{h_{i-1}^2}+\frac{y_{i+1}-y_i}{h_i^2}\right)$$

令

$$\alpha_i=\frac{h_{i-1}}{h_{i-1}+h_i},\beta_i=3\left((1-\alpha_i)\frac{y_i-y_{i-1}}{h_{i-1}}+\alpha_i\frac{y_{i+1}-y_i}{h_i}\right)$$

（39）

$f[x_{i-1},x_i]$　　$f[x_i,x_{i+1}]$

则式（38）可表示为

$$(1-\alpha_i)m_{i-1}+2m_i+\alpha_i m_{i+1}=\beta_i$$

差商

（40）

并注意到　$m_0=y_0', m_n=y_n'$

因而有三对角方程组（基本方程组）

$$\begin{cases} 2m_1+\alpha_1 m_2=\beta_1-(1-\alpha_1)y'_0 \\ (1-\alpha_2)m_1+2m_2+\alpha_2 m_3=\beta_2 \\ \qquad\cdots\cdots \\ (1-\alpha_{n-2})m_{n-3}+2m_{n-2}+\alpha_{n-2}m_{n-1}=\beta_{n-2} \\ (1-\alpha_{n-1})m_{n-2}+2m_{n-1}=\beta_{n-1}-\alpha_{n-1}y'_n \end{cases}$$

其系数行列式是一个三对角行列式，在用追赶方法求其解，于是得到分段插值多项式，即三次样条函数。

或者

$$\begin{pmatrix} 2 & 0 & & & & \\ 1-\alpha_1 & 2 & \alpha_1 & & & \\ & 1-\alpha_2 & 2 & \alpha_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1-\alpha_{n-1} & 2 & \alpha_{n-1} \\ & & & & 0 & 2 \end{pmatrix} \begin{pmatrix} m_0 \\ m_1 \\ m_2 \\ \vdots \\ m_{n-1} \\ m_n \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{n-1} \\ \beta_n \end{pmatrix}$$

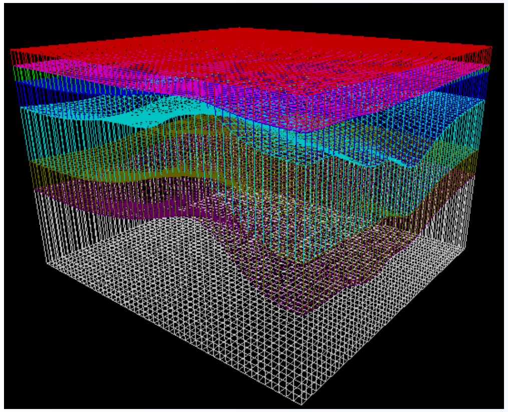
其中

$$\alpha_i = \frac{h_{i-1}}{h_i + h_{i-1}} \quad , \qquad 1-\alpha_i = \frac{h_i}{h_i + h_{i-1}}$$

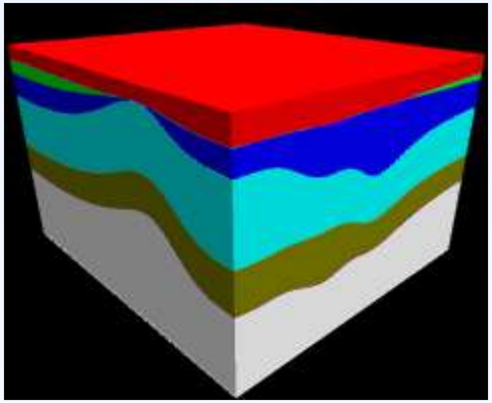
$$h_i = x_{i+1} - x_i$$

地质三维建模与可视化决策支持系统

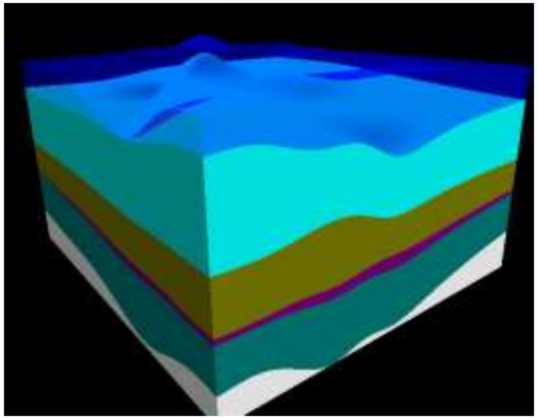
三维地质建模与地质体真实感图形显示



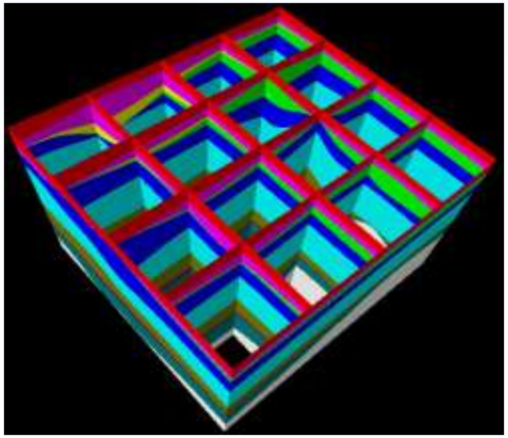
地质体网格模型



地质体真实感图形



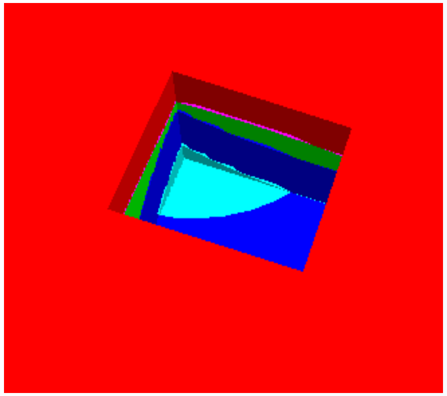
地质体选层显示

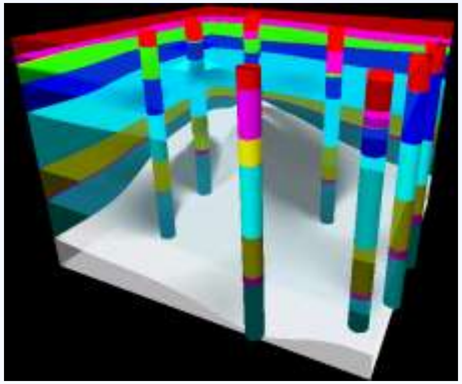


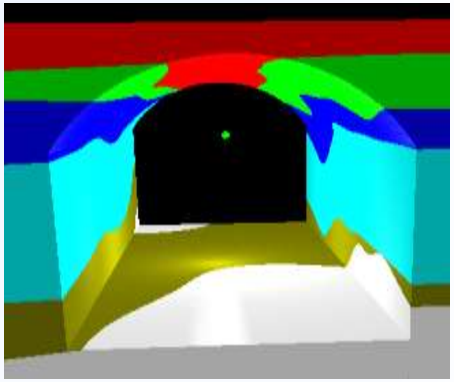
地质体栅格显示

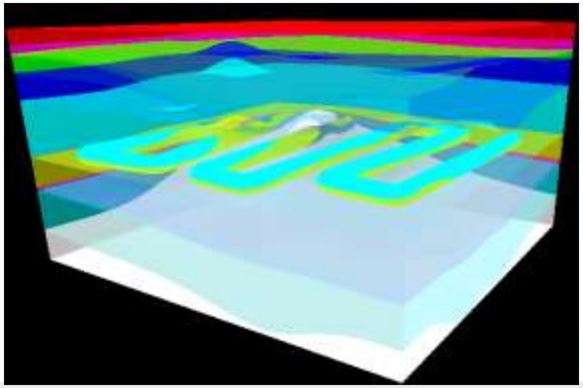
采用空间插值和NURBS(Non-Uniform Rational B-Splines)曲线曲面造型技术实现复杂地质体的建模

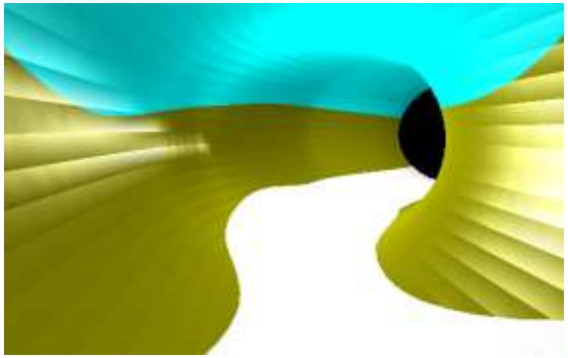
地质可视化分析


虚拟挖方


虚拟钻孔


虚拟直隧道


虚拟管道


虚拟管道漫游

1.9 曲线拟合的最小二乘法

1．教学内容：

曲线拟合的概念、直线拟合、多项式拟合、正则方程组。

2．重点难点：

拟合曲线的类型、正则方程组的建立、拟合多项式的求解。

3．教学目标：

了解曲线拟合的概念、对给出的一组数据点，能判断其拟合曲线的类型、建立相应的正则方程组、求得拟合多项式

引　言

一、问题的提法

在科学实验和生产实践中，经常要从一组实验数据出发，寻找函数$y=f(x)$的一个近似公式（称为经验公式）。已有的多项式插值法解决这类问题有明显的缺陷：实验数据有误差；实验数据量大等。

二、目　的

实际应用中并不刻意要求曲线经过所有的观测点，而是在符合数据分布特征的某类曲线中，在某个函数类中寻找一个“最好”的函数来拟合这组数据。

三、方 法

曲线拟合方法. 数据拟合最常用的近似标准是最小二乘法则.

最小二乘法

一、基本概念：残差

$$e_i = y_i - \hat{y}_i \quad (i = 1, 2, \cdots, N)$$

拟合的目的：使得残差最小，其中 $\hat{y} = \varphi(x)$ 为所要找的函数。

二、残差的选取方法（原则）

1、选取 $\varphi(x)$ 使残差绝对值之和最小，即

$$\sum_{i=1}^{N} |e_i| = \sum_{i=1}^{N} |y_i - \hat{y}_i| = \min \qquad (1)$$

2、选取 $\varphi(x)$，使残差最大绝对值最小，即

$$\max_i |e_i| = \max_i |y_i - \hat{y}_i| = \min \qquad (2)$$

3、选取 $\varphi(x)$，使残差平方之和最小，即

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} [y_i - \hat{y}_i]^2 = \min \qquad (4)$$

三、最小二乘原则（方法）

1、定义：使“残差平方和最小”的原则称为最小二乘原则。

2、定义：按照最小二乘原则选取拟合曲线的方法，称为最小二乘法。

1、直线拟合

假设给定的数据点 (x_i, y_i)，$i=1, 2, \ldots, n$ 的分布大致成一直线，虽然我们不能要求所做的拟合直线

$$y = a + bx$$

严格地通过所有的数据点 (x_i, y_i)，但总希望它尽可能地从所给数据点附近通过，即要求近似成立

$$y_i \approx a + bx_i, \qquad i = 1,2, \ldots ,n$$

由于数据点数目通常远远大于待定系数的个数，因此，拟合直线的构造实际上是求解超定方程（矛盾）方程组的代数问题。

（1）、（2）两种由于含有绝对值运算不便于实际应用。

基于准则（3）来选取拟合曲线的方法称为曲线拟合的**最小二乘法**

直线拟合问题可用数学语言描述如下：

问题10 对于给定的数据点 (x_i, y_i)， $i = 1, 2, \ldots, n$

求作一次式 y=a+bx，使总误差

$$Q = \sum_{i=1}^{N} [y_i - (a + bx_i)]^2$$

为最小。

要使 Q 达到极值，参数a，b 应满足

$$\frac{\partial Q}{\partial a}=2\sum_{i=1}^{n}[y_i-(a+bx_i)]\frac{\partial[y_i-(a+bx_i)]}{\partial a}=-2\sum_{i=1}^{n}[y_i-(a+bx_i)]$$

即

$$\sum_{i=1}^{N}[y_i-(a+bx_i)]=0$$

$$\sum_{i=1}^{N}[y_i-(a+bx_i)]x_i=0$$

由此可得：

$$\begin{cases} aN+b\sum_{i=1}^{N}x_i=\sum_{i=1}^{N}y_i \\ a\sum_{i=1}^{N}x_i+b\sum_{i=1}^{N}x_i^2=\sum_{i=1}^{N}x_iy_i \end{cases} \qquad (42)$$

解线性方程组（42）既可得到a，b

例： 炼钢是个氧化脱碳的过程，钢液含碳量的多少直接影响冶炼时间的长短，下表是某平炉的生产记录，表中 i 是次数，x_i 为全部炉料熔化完毕时的钢液的含碳量，y_i 为熔毕至出钢所许的冶炼时间。

I	1	2	3	4	5
x_i	165	123	150	123	141
y_i	187	126	172	125	148

解：

设所求的拟合直线为 y=a+bx

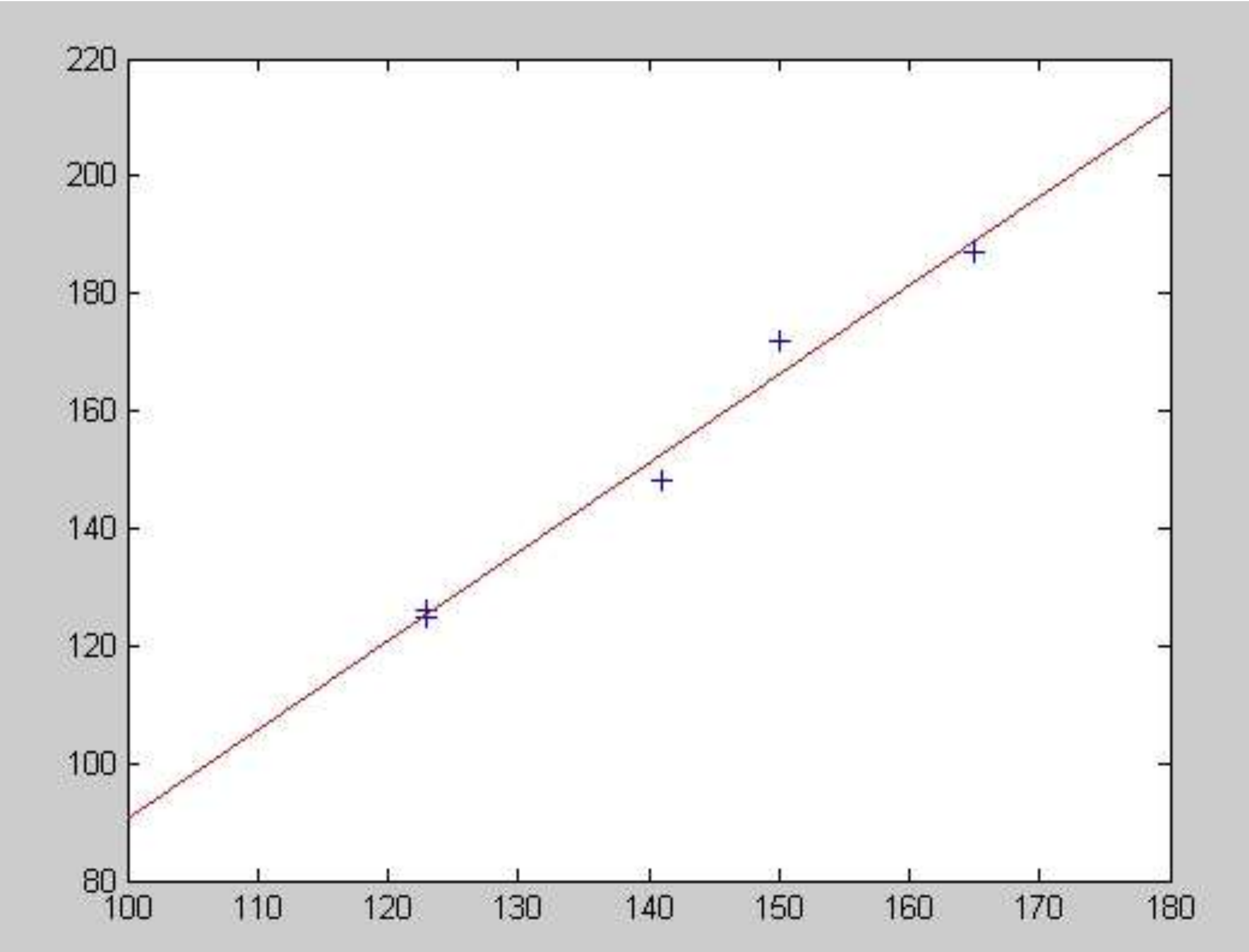
由（42）式可得关于a，b的线性方程组

$$\begin{cases} 5a + 702b = 758 \\ 702a + 99864b = 108396 \end{cases}$$

解此线性方程组得：a=-60.9392，b=1.5138

故拟合直线为：

$$y = -60.9392 + 1.5138x$$


220
200
180
160
140
120
100
80
100
110
120
130
140
150
160
170
180

2、多项式拟合

多项式拟合，是最流行的数据处理方法之一．它常用于把观测数据（离散的数据）归纳总结为经验公式（连续的函数），以利于进一步的推演分析或应用．

问题11 对于给定的数据点（x_i，y_i），$i=1，2，\ldots，n$ 求作m（m<<N）次多项式

$$y=\sum_{j=0}^{m}a_jx^j$$

使总误差

$$Q=\sum_{i=1}^{N}[y_i-\sum_{j=0}^{m}a_jx_i^j]^2$$

为最小。

由于Q可以看成是关于 $a_j(j=0, 1, \ldots, m)$ 的多元函数，故上述拟合多项式的构造问题可归结为多元函数的极值问题。

令
$$\frac{\partial Q}{\partial a_k}=0, \quad k=0, 1, \ldots, m$$

得：

$$\sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}a_j x_i^j)x_i^k=0, \quad k=0, 1, \ldots, m$$

既

$$
\begin{cases}
a_0 N + a_1 \sum_{i=1}^{N} x_i + \ldots + a_m \sum_{i=1}^{N} x_i^m = \sum_{i=1}^{N} y_i \\
a_0 \sum_{i=1}^{N} x_i + a_1 \sum_{i=1}^{N} x_i^2 + \ldots + a_m \sum_{i=1}^{N} x_i^{m+1} = \sum_{i=1}^{N} x_i y_i \\
\cdots\cdots\cdots \\
a_0 \sum_{i=1}^{N} x_i^m + a_1 \sum_{i=1}^{N} x_i^{m+1} + \ldots + a_m \sum_{i=1}^{N} x_i^{2m} = \sum_{i=1}^{N} x_i^m y_i
\end{cases}
\qquad (43)
$$

这个关于系数 a_j 的线性方程组称为**正则方程组**

定理7　正则方程组（43）有唯一解

证： 用反证法，若不然，则对应的齐次方程组

$$
\begin{cases}
a_0 N + a_1 \sum_{i=1}^{N} x_i + \cdots + a_m \sum_{i=1}^{N} x_i^m = 0 \\
a_0 \sum_{i=1}^{N} x_i + a_1 \sum_{i=1}^{N} x_i^2 + \cdots + a_m \sum_{i=1}^{N} x_i^{m+1} = 0 \\
\cdots\cdots \\
a_0 \sum_{i=1}^{N} x_i^m + a_1 \sum_{i=1}^{N} x_i^{m+1} + \cdots + a_m \sum_{i=1}^{N} x_i^{2m} = 0
\end{cases}
$$

有非零解

而

$$a_0\sum_{i=1}^{N}x_i^k+a_1\sum_{i=1}^{N}x_i^{k+1}+\cdots+a_m\sum_{i=1}^{N}x_i^{k+m}=\sum_{j=0}^{m}a_j\sum_{i=0}^{N}x_i^{k+j}$$

从而有

$$\sum_{j=0}^{m}a_j\sum_{i=1}^{N}x_i^{k+j}=0,\quad k=0,\ 1,\quad \ldots,\quad m$$

所以

$$\sum_{k=0}^{m}a_k\left(\sum_{j=0}^{m}a_j\sum_{i=1}^{N}x_i^{k+j}\right)=0$$

而

$$\sum_{k=0}^{m} a_k (\sum_{j=0}^{m} a_j \sum_{i=1}^{N} x_i^{k+j}) = \sum_{i=1}^{N}\sum_{k=0}^{m}\sum_{j=0}^{m} a_k a_j x_i^{k+j}$$

$$= \sum_{i=1}^{N} (\sum_{k=0}^{m} a_k x_i^k)(\sum_{j=0}^{m} a_j x_i^j) = \sum_{i=1}^{N} (\sum_{j=0}^{m} a_j x_i^j)^2$$

因此有：

$$\sum_{j=0}^{m} a_j x_i^j = 0 \quad (i = 1,2, \dots ,N)$$

即拟合多项式 $y=\sum_{j=0}^{m} a_j x^j$

有N个零点 $x_i (i=1,2,\dots,N)$

当N>m时，由代数学基本定理知必有 $\sum_{j=0}^{m} a_j x^j \equiv 0$

从而

$$a_j = 0 \qquad (j=1,2,\dots,m)$$

故与正则方程组的题设矛盾，定理得证。

定理8 设 $a_j\,(j=0,1,\cdots,m)$ 为正则方程（43）的解，则 $y=\sum_{j=0}^{m}a_jx^j$ 必为问题11的解。

证： 任给一组值 $b_j \quad (j=0,1,\cdots,\ m)$ 有

$$\sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}b_jx_i^j)^2-\sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}a_jx_i^j)^2$$

$$=\sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}a_jx_i^j+\sum_{j=0}^{m}a_jx_i^j-\sum_{j=0}^{m}b_jx_i^j)^2-\sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}a_jx_i^j)^2$$

$$=\sum_{i=1}^{N}(y_i-\cdots x_i^j)(\sum_{j=0}^{m}a_jx_i^j-\sum_{j=0}^{m}b_jx_i^j)+$$

$$(\sum_{j=0}^{m}a_jx_i^j-\cdots x_i^j)^2$$

利用正则方程组（43）可以知道，该项应该为零

$$=2\sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}\cdots x_i^j)(\sum_{j=0}^{m}a_jx_i^j-\sum_{j=0}^{m}b_jx_i^j)+(\sum_{j=0}^{m}a_jx_i^j-\sum_{j=0}^{m}b_jx_i^j)^2$$

$$\geq 2\sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}a_jx_i^j)(\sum_{j=0}^{m}a_jx_i^j-\sum_{j=0}^{m}b_jx_i^j)=0$$

因而有

$$\sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}a_jx_i^j)^2 \le \sum_{i=1}^{N}(y_i-\sum_{j=0}^{m}b_jx_i^j)^2$$

所以，只有 $a_j\,(j=0,1,\cdots,m)$ 使得残差的平方和最小

故

$$y=\sum_{j=0}^{m}a_jx^j$$

必为问题11的解。

多项式拟合的一般方法可归纳为：

(1)根据具体问题，确定拟合多项式的次数*n*；（描点）

(2)计算正则方程组的系数和右端项

$$S_k=\sum_{i=0}^{m}x_i^k, \qquad t_k=\sum_{i=0}^{m}x_i^k\, y_i.$$

(3)写出正则方程组

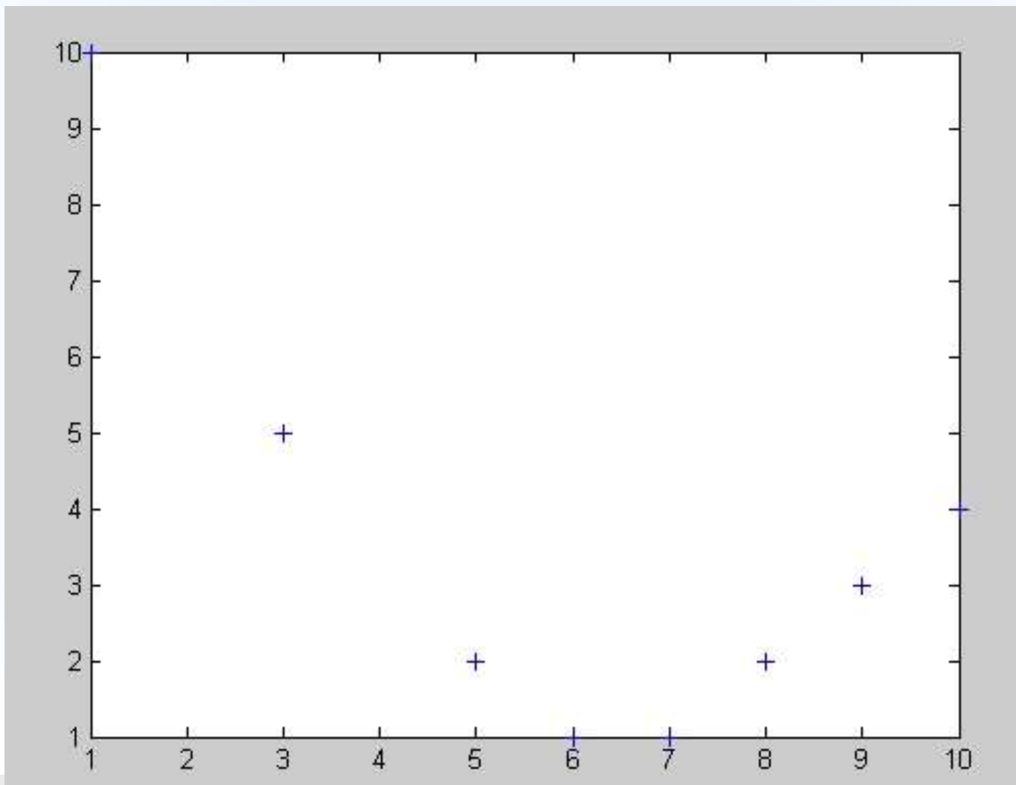
(4)解正则方程组,求出$a_0,a_1,\ldots,a_n$；

(5)写出拟合多项式$P_n(x)$

例：试求一个多项式拟合下列数据。

x	1	3	5	6	7	8	9	10		

如图所示，它们大体分布在一条抛物线附近，故考虑用二次多项式函数拟合这些数据。



解：

设所求的拟合多项式为 $y = a_0 + a_1 x + a_2 x^2$

由（43）式

$$
\begin{cases}
a_0 N + a_1 \sum_{i=1}^{N} x_i + \ldots + a_m \sum_{i=1}^{N} x_i^m = \sum_{i=1}^{N} y_i \\
a_0 \sum_{i=1}^{N} x_i + a_1 \sum_{i=1}^{N} x_i^2 + \ldots + a_m \sum_{i=1}^{N} x_i^{m+1} = \sum_{i=1}^{N} x_i y_i \\
\cdots\cdots\cdots\cdots \\
a_0 \sum_{i=1}^{N} x_i^m + a_1 \sum_{i=1}^{N} x_i^{m+1} + \ldots + a_m \sum_{i=1}^{N} x_i^{2m} = \sum_{i=1}^{N} x_i^m y_i
\end{cases}
\qquad (43)
$$

得其正则方程组为：

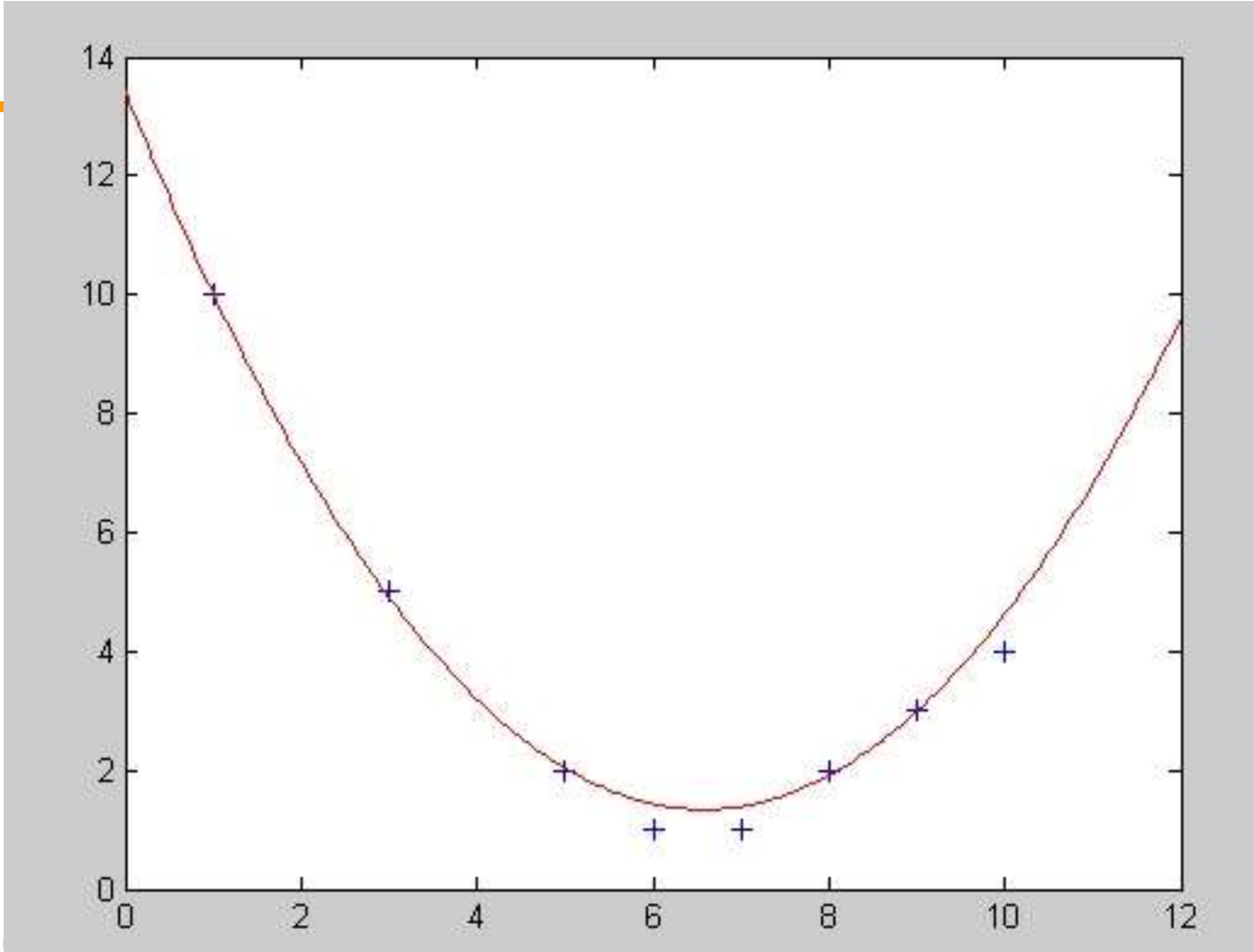
$$\begin{cases} 8a_0 + 49a_1 + 365a_2 = 28 \\ 49a_0 + 365a_1 + 2953a_2 = 131 \\ 365a_0 + 2953a_1 + 25061a_2 = 961 \end{cases}$$

解此方程组得：

$$a_0 = 13.43,\quad a_1 = -3.68,\quad a_2 = 0.28$$

所以，所求拟合多项式为：

$$y = 13.43 - 3.68x + 0.28x^2$$


14
12
10
8
6
4
2
0
0
2
4
6
8
10
12

本 章 小 结

插值问题

设函数f(x)在区间[a ，b]上有定义，且已知在一组互异点 $a \le x_0 < x_1 < \ldots < x_n \le b$ 上的函数值 y_0，y_1， …，y_n，寻求一个简单的函数p(x)，使满足

$$p(x_i) = y_i, \quad i = 0, 1, 2, \ldots, n \qquad (1.1)$$

并用p(x)近似代替f(x)，上述问题称为**插值问题**。

拉格朗日插值基函数

$$l_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{(x - x_j)}{(x_k - x_j)}$$

拉格朗日插值公式：

$$P_n(x) = \sum_{k=0}^{n} l_k(x) y_k = \sum_{k=0}^{n} \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x - x_j}{x_k - x_j} y_k$$

拉格朗日插值多项式存在并且唯一，并有估计式

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{k=0}^{n} (x - x_k)$$

n阶差商可以递推定义为：

$$\boxed{f(x_0, x_1, \ldots, x_n) = \frac{f(x_1, x_2, \ldots, x_n) - f(x_0, x_1, \ldots, x_{n-1})}{x_n - x_0}}$$

n阶差商的性质：

$$\boxed{f(x_0, x_1, \ldots, x_n) = \sum_{k=0}^{n} \frac{f(x_k)}{\prod\limits_{\substack{j=0 \\ j \neq k}}^{n} (x_k - x_j)} \qquad (24)}$$

n阶差商关于节点是对称的

差商与导数的关系

$$f(x_0, x_1, \ldots, x_n) = \frac{f^{(n)}(\xi)}{n!}$$

差商表的建立与使用

x	$f(x)$	一阶差商	二阶差商	三阶差商
x_0	$f(x_0)$			
x_1	$f(x_1)$	$f(x_0,x_1)$		
x_2	$f(x_2)$	$f(x_1,x_2)$	$f(x_0,x_1,x_2)$	
x_3	$f(x_3)$	$f(x_2,x_3)$	$f(x_1,x_2,x_3)$	$f(x_0,x_1,x_2,x_3)$

牛顿插值公式

$$
\begin{aligned}
p_n(x) = f(x_0) + f(x_0, x_1)(x - x_0) + \\
f(x_0, x_1, x_2)(x - x_0)(x - x_1) + \ldots + \\
f(x_0, x_1, \ldots, x_n)(x - x_0)(x - x_1)\ldots(x - x_{n-1})
\end{aligned}
$$

差分的定义

$$\Delta^n y_i = \Delta^{n-1} y_{i+1} - \Delta^{n-1} y_i$$

有限差分公式

$$p_n(x_0 + th) = y_0 + \sum_{k=1}^{n} \frac{\Delta^k y_0}{k!} \prod_{j=0}^{k-1} (t - j)$$

差分表

x_i	y_i	一	阶	[illegible]	[illegible]
x_0	y_0				
x_1	y_1	Δy_0			
x_2	y_2	Δy_1	$\Delta^2 y_0$		
x_3	y_3	Δy_2	$\Delta^2 y_1$	$\Delta^3 y_0$	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

正则方程组

$$
\begin{cases}
a_0 N + a_1 \sum_{i=1}^{N} x_i + \ldots + a_m \sum_{i=1}^{N} x_i^m = \sum_{i=1}^{N} y_i \\
a_0 \sum_{i=1}^{N} x_i + a_1 \sum_{i=1}^{N} x_i^2 + \ldots + a_m \sum_{i=1}^{N} x_i^{m+1} = \sum_{i=1}^{N} x_i y_i \\
\cdots\cdots\cdots\cdot \\
a_0 \sum_{i=1}^{N} x_i^m + a_1 \sum_{i=1}^{N} x_i^{m+1} + \ldots + a_m \sum_{i=1}^{N} x_i^{2m} = \sum_{i=1}^{N} x_i^m y_i
\end{cases}
\tag{43}
$$

作业

P54　6、11、12、13、16、17、31、36、37