第二章 插值方法和曲线拟合的最小二乘法

实践中常有这样的问题：由实验测得某一解析表达式未知的函数 $f(x)$ 在一系列点 x_0，x_1，$\cdots,x_n$ 处的值 $f(x_0),f(x_1),\cdots,f(x_n)$，根据所测数据构造一个简单函数 $p(x)$，作为 $f(x)$ 的近似表达式且要求 $p(x_i)=f(x_i),i=1,2,\cdots,n$；或者 $f(x)$ 虽有解析式，但表达式复杂，不便于计算，需要构造一个简单函数 $p(x)$ 近似地代替 $f(x)$ 且要求 $p(x_i)=f(x_i),i=1,2,\cdots,n$；这就是函数插值问题.

2.1 Lagrange 插值

2.1.1 Lagrange（拉格朗日）插值

设函数 $y=f(x)$ 在区间 $[a,b]$ 上有定义，已知在点 $a\le x_0<x_1<\cdots<x_n\le b$ 上的值 $y_0,y_1,\cdots,y_n$，若存在一个简单函数 $P(x)$，使

$$P(x_i)=y_i \quad (i=0,1,\cdots,n), \tag{1.1}$$

成立，就称 $P(x)$ 为 $f(x)$ 的**插值函数**，点 $x_0,x_1,\cdots,x_n$ 称为**插值节点**，包含插值节点的区间 $[a,b]$ 称为**插值区间**，求插值函数 $P(x)$ 的方法称为**插值法**. 若 $P(x)$ 是次数不超过 n 的代数多项式，即

$$P(x)=a_0+a_1x+\cdots+a_nx^n \quad , \tag{1.2}$$

其中 a_i 为实数，就称 $P(x)$ 为**插值多项式**，相应的插值法称为**多项式插值**.

首先讨论线性插值，即 $n=1$ 的情形。给定区间 $[x_k,x_{k+1}]$ 及端点函数值 $y_k=f(x_k)$，$y_{k+1}=f(x_{k+1})$，要求线性插值多项式 $L_1(x)$，使它满足

$$L_1(x_k)=y_k，\ L_1(x_{k+1})=y_{k+1}.$$

$y=L_1(x)$ 的几何意义就是通过两点 (x_k,y_k) 与 (x_{k+1},y_{k+1}) 的直线，$L_1(x)$ 的表达式可由几何意义直接给出

$$L_1(x)=\frac{x_{k+1}-x}{x_{k+1}-x_k}y_k+\frac{x-x_k}{x_{k+1}-x_k}y_{k+1},$$

由上式看出，$L_1(x)$ 是由两个线性函数

$$l_k(x)=\frac{x-x_{k+1}}{x_k-x_{k+1}},\qquad l_{k+1}(x)=\frac{x-x_k}{x_{k+1}-x_k},$$

的线性组合得到，其系数分别为 y_k 及 y_{k+1}，即

$$L_1(x)=y_k l_k(x)+y_{k+1}l_{k+1}(x).\tag{1.3}$$

显然，$l_k(x)$ 及 $l_{k+1}(x)$ 也是线性插值多项式，在节点 x_k 及 x_{k+1} 上满足条件

$$l_k(x_k)=1,\qquad l_k(x_{k+1})=0;$$

$$l_{k+1}(x_k)=0,\qquad l_{k+1}(x_{k+1})=1.$$

我们称函数 $l_k(x)$ 及 $l_{k+1}(x)$ 为**线性插值基函数**.

其次讨论二次插值，即 $n=2$ 的情形。假定节点为 x_{k-1},x_k,x_{k+1}，求二次插值多项式 $L_2(x)$，使它满足

$$L_2(x_j)=y_j,\ (j=k-1,k,k+1).$$

类似线性插值，需要构造插值基函数 $l_{k-1}(x)$，$l_k(x)$，$l_{k+1}(x)$，且插值基函数满足

$$l_{k-1}(x_{k-1})=1,\ l_{k-1}(x_k)=0,\ l_{k-1}(x_{k+1})=0$$

$$l_k(x_{k-1})=0,\ l_k(x_k)=1,\ l_k(x_{k+1})=0$$

$$l_{k+1}(x_{k-1})=0,\ l_{k+1}(x_k)=0,\ l_{k+1}(x_{k+1})=1$$

下面构造 $l_{k-1}(x)$。由于 $l_{k-1}(x)$ 有 x_k，x_{k+1} 两个零点，因此，$l_{k-1}(x)$ 含因子 $(x-x_k)(x-x_{k+1})$，又由于 $l_{k-1}(x)$ 是一个次数不超过二的多项式，所以，$l_{k-1}(x)$ 还可能含有一个常数因子。于是，设 $l_{k-1}(x)=A(x-x_k)(x-x_{k+1})$。利用 $l_{k-1}(x_{k-1})=1$ 得

$$l_{k-1}(x)=\frac{(x-x_k)(x-x_{k+1})}{(x_{k-1}-x_k)(x_{k-1}-x_{k+1})},$$

类似的方法可得，

$$l_k(x)=\frac{(x-x_{k-1})(x-x_{k+1})}{(x_k-x_{k-1})(x_k-x_{k+1})},$$

$$l_{k+1}(x)=\frac{(x-x_{k-1})(x-x_k)}{(x_{k+1}-x_{k-1})(x_{k+1}-x_k)}.$$

利用 $l_{k-1}(x), l_k(x), l_{k+1}(x)$，立即得到二次插值多项式

$$L_2(x)=y_{k-1}l_{k-1}(x)+y_kl_k(x)+y_{k+1}l_{k+1}(x). \tag{1.4}$$

最后构造通过 $n+1$ 个节点 $x_0<x_1<\cdots<x_n$ 的 n 次插值多项式 $L_n(x)$，假定它满足条件

$$L_n(x_j)=y_j,\ (j=0,1,\cdots,n). \tag{1.5}$$

为了构造 $L_n(x)$，我们先定义 n 次插值基函数.

定义 1 若 n 次多项式 $l_j(x)(j=0,1,\cdots n)$ 在 $n+1$ 个节点 $x_0<x_1<\cdots<x_n$ 上满足条件

$$l_j(x_k)=\begin{cases}1, & k=j;\\ 0, & k\neq j.\end{cases}\quad (j,k=0,1,\cdots,n), \tag{1.6}$$

就称这 $n+1$ 个 n 次多项式 $l_0(x), l_1(x),\cdots,l_n(x)$ 为节点 $x_0, x_1, \cdots, x_n$ 上的 **n 次插值基函数.** 由条件（1.6）和类似 $n=2$ 的方法可得

$$l_k(x)=\frac{(x-x_0)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n)}{(x_k-x_0)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)}\quad (k=0,1,\cdots,n). \tag{1.7}$$

于是，满足条件（1.5）的插值多项式 $L_n(x)$ 可表示为

$$L_n(x)=\sum_{k=0}^{n}y_kl_k(x), \tag{1.8}$$

由插值基函数的定义，知

$$L_n(x_j)=\sum_{k=0}^{n}y_kl_k(x_j)=y_j,\ (j=0,1,\cdots,n)\ .$$

形如（1.8）的插值多项式 $L_n(x)$ 称为 **Lagrange 插值多项式.**

若引人记号

$$\omega_{n+1}(x)=(x-x_0)(x-x_1)\cdots(x-x_n), \tag{1.9}$$

容易求得

$$\omega'_{n+1}(x_k)=(x_k-x_0)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n).$$

于是公式（1.8）可改写成

$$L_n(x)=\sum_{k=0}^{n}y_k\frac{\omega_{n+1}(x)}{(x-x_k)\omega'_{n+1}(x_k)}. \tag{1.10}$$

定理 1 在次数不超过 n 的多项式集合 H_n 中，满足条件（1.5）的插值多项式 $L_n(x)\in H_n$ 是存在唯一的.

证明 $L_n(x)$ 的构造证明了插值多项式的存在性，下面用反证法证明唯一性. 假定还有 $p(x)\in H_n$ 使 $p(x_i)=f(x_i),\quad i=0,1,\cdots,n$ 成立. 于是有 $L_n(x_i)-p(x_i)=0$ 对 $i=0,1,\cdots,n$ 成立，它表明多项式 $L_n(x)-p(x)\in H_n$ 有 $n+1$ 个零点 $x_0,x_1,\cdots,x_n$，这与 n 次多项式只有 n 个零点的代数基本定理矛盾，故只能 $p(x)\equiv L_n(x)$.

根据存在唯一性定理，令

$$f(x)=x^m,\quad m=0,1,\cdots,n. \tag{1.11}$$

取 $m=0$，则

$$\sum_{k=0}^{n}l_k(x)=1, \tag{1.12}$$

它可用来检验数组 $\{l_k(x),\ k=0,1,\cdots,n\}$ 的正确性.

例 1. 求经过点 $A(0,\ 1)$, $B(1,\ 2)$, $C(2,\ 3)$ 三点的 Lagrange 插值多项式.

解 设 $x_0=0,\ x_1=1,\ x_2=2,\ y_0=1,\ y_1=2,\ y_2=3$，则

$$l_0(x)=\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}=\frac{(x-1)(x-2)}{(0-1)(0-2)}=\frac{1}{2}(x^2-3x+2),$$

$$l_1(x)=\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}=\frac{(x-0)(x-2)}{(1-0)(1-2)}=-(x^2-2x),$$

$$l_2(x)=\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}=\frac{(x-0)(x-1)}{(2-0)(2-1)}=\frac{1}{2}(x^2-x).$$

故所求插值多项式为

$$p_2(x)=y_0l_0(x)+y_1l_1(x)+y_2l_2(x)=x+1.$$

2.1.2 插值余项与误差估计

若在$[a\,,\,b]$上用$L_n(x)$近似$f(x)$，则其截断误差为

$$R_n(x)=f(x)-L_n(x),$$

也称为插值多项式的余项. 关于插值余项估计有以下定理.

定理 2 设$f^{(n)}(x)$在$[a\,,\,b]$上连续，$f^{(n+1)}(x)$在$(a\,,\,b)$内存在，节点

$$a\le x_0<x_1<\cdots<x_n\le b,$$

$L_n(x)$是满足条件（1.5）的插值多项式，则对任何$x\in[a\,,\,b]$，插值余项

$$R_n(x)=f(x)-L_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}\omega_{n+1}(x), \tag{1.13}$$

这里$\xi\in(a,b)$且依赖于$x,\omega_{n+1}(x)$是（1.9）所定义的.

证明 由给定条件知$R_n(x)$在节点$x_k\,(k=0,1,\cdots,n)$上为零，即

$$R_n(x_k)=0\quad(k=0,1,\cdots,n),$$

于是

$$R_n(x)=K(x)(x-x_0)(x-x_1)\cdots(x-x_n)=K(x)\omega_{n+1}(x) \tag{1.14}$$

其中$K(x)$是与x有关的待定函数. 现在把x看成$[a,b]$上的一个固定点，作函数

$$\varphi(t)=f(t)-L_n(t)-K(x)(t-x_0)(t-x_1)\cdots(t-x_n),$$

根据插值条件及余项定义，可知$\varphi(t)$在点$x_{0,}x_1,\cdots x_n$及x处均为零，故$\varphi(t)$在$[a,b]$上有$n+2$个零点，根据罗尔（Rolle）定理，$\varphi'(x)$在$\varphi(t)$的两个零点间至少有一个零点，故$\varphi'(x)$在$[a,b]$内至少有$n+1$个零点. 对$\varphi'(x)$再应用罗尔定理，可知$\varphi''(t)$在$[a,b]$内至少有n个零点. 依此类推，$\varphi^{(n+1)}(t)$在(a,b)内至少有一个零点，记为$\xi\in(a,b)$，使

$$\varphi^{(n+1)}(\xi)=f^{(n+1)}(\xi)-(n+1)!K(x)=0,$$

于是

$$K(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!},\xi\in(a,b),$$

得到余项表达式（1.13）.

推论 若 $f(x)$ 是次数不超过 n 次的多项式，$f(x)$ 的 n 次 Lagrange 插值多项式为 $p_n(x)$，则 $p_n(x)\equiv f(x)$.

余项表达式只有在 $f(x)$ 的高阶导数存在时才能应用. ξ 在 (a,b) 内的具体位置通常不可能给出，如果我们可以求出 $\max\limits_{a\le x\le b}\left|f^{(n+1)}(x)\right|=M_{n+1}$ 那么插值多项式 $L_n(x)$ 逼近 $f(x)$ 的截断误差限是

$$\left|R_n(x)\right|\le\frac{M_{n+1}}{(n+1)!}\left|\omega_{n+1}(x)\right| \tag{1.15}$$

当 $n=1$ 时，一次插值（又称线性插值）余项为

$$R_1(x)=\frac{1}{2}f''(\xi)\omega_2(x)=\frac{1}{2}f''(\xi)(x-x_0)(x-x_1),\ x\in[x_0,x_1]; \tag{1.16}$$

当 $n=2$ 时，二次插值（又称抛物插值）余项为

$$R_2(x)=\frac{1}{6}f'''(\xi)(x-x_0)(x-x_1)(x-x_2),\ x\in[x_0,x_2]. \tag{1.17}$$

例 2 已给 $\sin 0.32=0.314567, \sin 0.34=0.333487, \sin 0.36=0.352274$，用线性插值及抛物插值计算 $\sin 0.3367$ 的值并估计截断误差.

解 由题意取

$$x_0=0.32, y_0=0.314567, x_1=0.34, y_1=0.333487, x_2=0.36, y_2=0.352274.$$

用线性插值计算，取 $x_0=0.32$ 及 $x_1=0.34$，由公式（1.3）得：

$$\sin 0.3367\approx L_1(0.3367)=y_0+\frac{y_1-y_0}{x_1-x_0}(0.3367-x_0)=0.314567+\frac{0.01892}{0.02}\times0.0167=0.330365$$

其截断误差由（1.16）得

$$\left|R_1(x)\right|\le\frac{M_2}{2}\left|(x-x_0)(x-x_1)\right|,$$

线性插值

其中 $M_2=\max\limits_{x_0\le x\le x_1}\left|f''(x)\right|$，

因 $f(x)=\sin x, f''(x)=-\sin x$，

可取 $M_2=\max\limits_{x_0\le x\le x_1}\left|\sin x\right|=\sin x_1\le 0.3335$，

于是

$$\left|R_1(0.3367)\right|=\left|\sin 0.3367-L_1(0.3367)\right|\le\frac{1}{2}\times0.3335\times0.0167\times0.0033\le0.92\times10^{-5};$$

用抛物插值计算 $\sin 0.3367$ 时，由公式（1.4）得

$$\sin 0.3367 \approx y_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + y_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2 \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

$$= L_2(0.03367)$$

$$= 0.314567 \times \frac{0.7689\times 10^{-4}}{0.0008} + 0.333487 \times \frac{3.89\times 10^{-4}}{0.0004} + 0.352274 \times \frac{-0.5511\times 10^{-4}}{0.0008}$$

$$= 0.330374$$

这个结果与 6 位有效数字的正函数表完全一样，这说明查表时用二次插值精度已相当高了.
其截断误差限由（1.17）得

$$|R_2(x)| \le \frac{M_3}{6}|(x-x_0)(x-x_1)(x-x_2)|,$$

二次插值

其中 $M_3 = \max\limits_{x_0\le x\le x_2} |f'''(x)| = \cos x_0 < 0.828,$

于是

$$|R_2(0.3367)| = |\sin 0.3367 - L_2(0.3367)| \le \frac{1}{6}\times 0.828\times 0.0167\times 0.033\times 0.0233 < 0.178\times 10^{-6}$$

2.2 Newton 插值

利用插值基函数很容易得到 Lagrange 插值多项式，公式结构紧凑，在理论分析中甚为方便，但当插值节点增减时全部插值基函数 $l_k(x)$ $(k=0,1,\cdots,n)$ 均要随之变化，整个公式也将发生变化，这在实际计算中是很不方便的，为了克服这一缺点，引入 Newton（牛顿）插值公式.
为了得到 n 次 Newton 插值公式，先介绍差商及其性质。

2.2.1 差商及其性质

定义 2 称 $f[x_i,x_j] = \dfrac{f(x_j)-f(x_i)}{x_j-x_i}$ 为函数 $f(x)$ 关于点 x_i,x_j 的**一阶差商.**

$f[x_i,x_j,x_k] = \dfrac{f[x_j,x_k]-f[x_i,x_j]}{x_k-x_i}$ 称为 $f(x)$ 的**二阶差商.** 一般地，称

$$f[x_0,x_1,\cdots,x_k] = \frac{f[x_1,\cdots,x_{k-1},x_k]-f[x_0,x_1,\cdots x_{k-1}]}{x_k-x_0} \tag{2.1}$$

为 $f(x)$ 的 k **阶差商.** 差商有如下的基本性质：

(1) 若 $F(x)=Cf(x)$， C 是常数，则 $F[x_0,x_1,\cdots,x_k]=Cf[x_0,x_1,\cdots,x_k]$.

(2) 若 $F(x)=f(x)+g(x)$，则 $F[x_0,x_1,\cdots,x_k]=f[x_0,x_1,\cdots,x_k]+g[x_0,x_1,\cdots,x_k]$.

(3) k 阶差商可表为函数值 $f(x_0),\cdots,f(x_k)$ 的线性组合：

$$f[x_0,\cdots,x_k]=\sum_{j=0}^{k}\frac{f(x_j)}{(x_j-x_0)\cdots(x_j-x_{j-1})(x_j-x_{j+1})\cdots(x_j-x_k)} \tag{2.2}$$

这个性质也表明差商与节点的排列次序无关，称为差商的对称性. 即

$$f[x_0,\cdots,x_k]=f[x_1,x_0,x_2,\cdots,x_k]=\cdots=f[x_1,\cdots x_k,x_0].$$

(4) 若 $f(x)$ 在 $[a,b]$ 上存在 n 阶导数，且节点 $x_0,\cdots,x_n\in[a,b]$，则 n 阶差商与导数关系如下：

$$f[x_0,\cdots,x_n]=\frac{f^{(n)}(\xi)}{n!},\quad \xi\in[a,b] \tag{2.3}$$

性质（1-2）由定义可直接推出；性质(3)可用归纳法证明；性质(4)可直接用罗尔定理证明.

差商计算可列差商表如下（表 2—1）

表 2—1 差商计算表

x_k	$f(x_k)$	一阶差商	二阶差商	三阶差商	四阶差商
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	$f(x_3)$	$f[x_2,x_3]$	$f[x_1,x_2,x_3]$	$f[x_0,x_1,x_2,x_3]$	
x_4	$f(x_4)$	$f[x_3,x_4]$	$f[x_2,x_3,x_4]$	$f[x_1,x_2,x_3,x_4]$	$f[x_0,x_1,x_2,x_3,x_4]$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

2.2.2 Newton 插值公式

定理 3 设 $x_0,x_1,\cdots,x_n$ 是一组互异的点，$y_i=f(x_i)$ $(i=0,1,2,\cdots,n)$，则 n 次多项式

$$N_n(x)=f(x_0)+f[x_0,x_1](x-x_0)+f[x_0,x_1,x_2](x-x_0)(x-x_1)$$

$$+\cdots+f[x_0,x_1,\cdots,x_n](x-x_0)(x-x_1)\cdots(x-x_{n-1}) \tag{2.4}$$

满足插值条件 $p_n(x_i)=y_i$ $(i=0,1,2,\cdots,n)$，并称(3. 13)为 **Newton 插值多项式**. 且余项为

$$R_n(x)=f(x)-N_n(x)=f[x_0,x_1,\cdots,x_n,x](x-x_0)(x-x_1)\cdots(x-x_n) \tag{2.5}$$

证 明 因为 $f[x_0,x]=\dfrac{f(x)-f(x_0)}{x-x_0}$，所以

$$f(x)=f(x_0)+f[x_0,x](x-x_0) \tag{2.6}$$

令

\ $$N_0(x)=f(x_0),R_0(x)=f(x)-N_0(x)=f[x_0,x](x-x_0);$$

因为 $f[x_0,x_1,x]=f[x_1,x_0,x]=\dfrac{f[x_0,x]-f[x_1,x_0]}{x-x_1}=\dfrac{f[x_0,x]-f[x_0,x_1]}{x-x_1}$，所以

$$f[x_0,x]=f[x_0,x_1]+f[x_0,x_1,x](x-x_1),$$

代入(2.6)，得

$$f(x)=f(x_0)+f[x_0,x_1](x-x_0)+f[x_0,x_1,x](x-x_0)(x-x_1) \tag{2.7}$$

令

$$N_1(x)=f(x_0)+f[x_0,x_1](x-x_0),$$

$$R_1(x)=f(x)-N_1(x)=f[x_0,x_1,x](x-x_0)(x-x_1),$$

依次类推，得

$$f(x)=f(x_0)+f[x_0,x_1](x-x_0)+f[x_0,x_1,x](x-x_0)(x-x_1)$$
$$+\cdots+f[x_0,x_1,\cdots x_n,x](x-x_0)(x-x_1)\cdots(x-x_n)$$

其中

$$N_n(x)=f(x_0)+f[x_0,x_1](x-x_0)+f[x_0,x_1,x_2](x-x_0)(x-x_1)$$
$$+\cdots+f[x_0,x_1,\cdots,x_n](x-x_0)(x-x_1)\cdots(x-x_{n-1})$$

$$R_n(x)=f(x)-N_n(x)=f[x_0,x_1,\cdots,x_n,x](x-x_0)(x-x_1)\cdots(x-x_n).$$

例 3 给出 $f(x)$ 的函数表（见表 2-3），求 4 次 Newton 插值多项式，并由此计算 $f(0.596)$ 的近似值. $N_4(x)$

首先根据给定函数表构造差商表.

表 2—2 差商表

0.4	0.41075			
0.55	0.57815	1.11600		
0.65	0.69675	1.18600	0.28000	
0.80	0.88811	1.27573	0.35893	0.19733

0.90	1.02652	1.38410	0.43348	0.21300	0.03134	
1.05	1.25382	1.51533	0.52493	0.22863	0.03126	−0.00012

从差商表看到 4 阶差商已近似常数．故取 4 次插值多项式 $p_4(x)$ 做近似即可.

$$N_4(x)=0.41075+1.116(x-0.4)+0.28(x-0.4)(x-0.55)$$

$$+0.19733(x-0.4)(x-0.55)(x-0.65)$$

$$+0.03134(x-0.4)(x-0.55)(x-0.65)(x-0.8)\text{，}$$

于是

$$f(0.596)\approx N_4(0.596)=0.63192\text{，}$$

截断误差

$$|R_4(x)|\approx|f[x_0,\cdots,x_5]\omega_5(0.596)|\le 3.63\times10^{-9}\text{ .}$$

这说明截断误差很小，可以忽略不计.

此例说明截断估计中，5 阶差商 $f[x,x_0,\cdots,x_4]$ 用 $f[x_0,x_1\cdots,x_5]=-0.00012$ 近似.另一种方法是取 $x=0.596$，由 $f(0.596)\approx0.63192$，可得 $f[x,x_0,\cdots,x_4]$ 的近似值，从而可求得 $|R_4(x)|$ 的近似.

2. 3 Hermite 插值

不少实际的插值问题不但要求在节点上函数值相等，而且还要求对应节点的导数值也相等，甚至要求节点处高阶导数值也相等，满足这种要求的插值多项式就是 **Hermite (埃尔米特)插值公式**.下面只讨论函数值与导数值个数相等的情况。为了简化问题，我们只研究两个节点情形。设在节点 $x_i<x_{i+1}$ 上，$y_j=f(x_j),m_j=f'(x_j)(j=i,i+1)$，要求插值多项式 $H(x)$，满足条件

$$H(x_j)=y_j\text{，}\quad H'(x_j)=m_j\quad(j=i,i+1)\qquad(3.1)$$

这里给出了 4 个条件，可惟一确定一个次数不超过 3 的多项式

$$H_3(x)=a_0+a_1x+a_2x^2+a_3x^3\text{ .}$$

我们采用 Lagrange 插值多项式的基本方法．先求插值基函数 $\alpha_j(x)$ 及 $\beta_j(x)(j=i,i+1)$，共有 4 个，每一个基函数都是 3 次多项式，且满足条件

$$\begin{cases} \alpha_j(x_k)=\delta_{jk}=\begin{cases}0, j\neq k,\\ 1, j=k,\end{cases} \quad \alpha_j'(x_k)=0; \\ \beta_j(x_k)=0;\ \beta_j'(x_k)=\delta_{jk}(j,k=i,i+1) \end{cases}, \tag{3.2}$$

于是满足条件（3.1）的插值多项式 $H(x)=H_3(x)$ 可写成插值基函数表示的形式

$$H_3(x)=y_i\alpha_i(x)+y_{i+1}\alpha_{i+1}(x)+m_i\beta_i(x)+m_{i+1}\beta_{i+1}(x) \tag{3.3}$$

由条件（3.2），显然 $H_3(x_j)=y_j$，$H_3'(x_j)=m_j,(j=i,i+1)$．下面问题是求满足条件（3.2）的基函数 $\alpha_j(x)$ 及 $\beta_j(x), j=i,i+1$．为此，可利用 Lagrange 插值基函数 $l_j(x)$．令 $\alpha_j(x)=(ax+b)\,l_j^2(x)$，其中 $l_j(x)$ 是（1.7）所表示的基函数．由条件（3.2）有

$$\alpha_j(x_j)=(ax_j+b)\,l_j^2(x_j)=1,$$

$$\alpha_j'(x_j)=l_j(x_j)[al_j(x_j)+2(ax_j+b)l_j'(x_j)]=0,$$

整理得

$$a=-2l_j'(x_j),\quad b=1+2x_jl_j'(x_j),$$

由于

$$l_i(x)=\frac{(x-x_{i+1})}{(x_i-x_{i+1})},\quad l_{i+1}(x)=\frac{(x-x_i)}{(x_{i+1}-x_i)}.$$

利用两端求导，得

$$l_i'(x_i)=\frac{1}{x_i-x_{i+1}},\quad l_{i+1}'(x_{i+1})=\frac{1}{x_{i+1}-x_i},$$

于是

$$\alpha_i(x)=(1+2\frac{x-x_i}{x_{i+1}-x_i})(\frac{x-x_{i+1}}{x_i-x_{i+1}})^2,\quad \alpha_{i+1}(x)=(1+2\frac{x-x_{i+1}}{x_i-x_{i+1}})(\frac{x-x_i}{x_{i+1}-x_i})^2$$

同理整理可得

$$\beta_i(x)=(x-x_i)(\frac{x-x_{i+1}}{x_i-x_{i+1}})^2,\quad \beta_{i+1}(x)=(x-x_{i+1})(\frac{x-x_i}{x_{i+1}-x_i})^2$$

还可以证明满足条件（3.1）的插值多项式是惟一的．用反证法，假设 $H_3(x)$ 及 $\overline{H}_3(x)$ 均满足条件（3.1），于是

$$\varphi(x)=H_3(x)-\overline{H}_3(x)$$

在每个节点 x_k 上均有二重根，即 $\varphi(x)$ 有 4 重根．但 $\varphi(x)$ 是不高于 3 次多项式，故 $\varphi(x)\equiv 0$，

惟一性得证.

仿照 Lagrange 插值余项的证明方法，若 $f(x)$ 在 (a,b) 内的 4 阶导数存在，则其插值余项

$$R(x)=f(x)-H_3(x)=\tfrac{f^{(4)}(\xi)}{4!}\omega_2^2(x), \tag{3.4}$$

其中 $\xi\in(a,b)$ 且与 x 有关.

于是满足条件(4.1)的插值多项式的余项为 $R_3(x)=f(x)-H_3(x)$. 由（3.4）得

$$R_3(x)=\frac{1}{4!}f^{(4)}(\xi)(x-x_k)^2(x-x_{k+1})^2,\xi\in(x_k,x_{k+1}).$$

2.4 分段低次插值

由后面 2.7.2 节的 Runge 现象的数值实验知，加密插值节点不一定能保证插值函数很好的逼近被插值函数，因此，引进分段插值的概念。

2.4.1 分段线性插值

所谓分段线性插值就是通过插值点用折线段连接起来逼近 $f(x)$.

定义 3 设已知节点 $a=x_0<x_1<\cdots<x_n=b$ 上的函数值 $f_0,f_1,\cdots,f_n$，记 $h_k=x_{k+1}-x_k$, $h=\max\limits_k h_k$，求一折线函数 $I_h(x)$ 满足：

1） 记 $I_h(x)\in C[a,b]$，

2） $I_h(x_k)=f_k$， （k=0，1，2$\cdots n$)

3）$I_h(x)$ 在每一个小区间 $[x_k,x_{k+1}]$ 上是线性函数.

则称 $I_h(x)$ 为**分段线性插值函数**.

由定义可知 $I_h(x)$ 在每个小区间 $[x_k,x_{k+1}]$ 上可表示为

$$I_h(x)=\frac{x-x_{k+1}}{x_k-x_{k+1}}f_k+\frac{x-x_k}{x_k-x_{k+1}}f_{k+1}\quad x_k\le x\le x_{k+1} \tag{4.1}$$

若用插值基函数表示，则在整个区间 $[a,b]$ 上 $I_h(x)$ 为

$$I_h(x)=\sum_{j=0}^{n} f_j l_j(x), \qquad (4.2)$$

基函数$l_j(x)$满足条件$l_j(x_k)=\delta_{jk}(j,k=0,1,\cdots,n)$，其形式是

$$l_j(x)=\begin{cases}\dfrac{x-x_{j-1}}{x_j-x_{j-1}}, & x_{j-1}\le x\le x_j \quad (j=0\text{略去})；\\ \dfrac{x-x_{j+1}}{x_j-x_{j+1}}, & x_j\le x\le x_{j+1} \quad (j=n\text{略去})；\\ 0, & x\in[a,b],\ x\notin[x_{j-1},x_{j+1}].\end{cases} \qquad (4.3)$$

2.4.2 分段三次 Hermite 插值

分段线性插值函数$I_h(x)$的导数是间断的，若在节点$x_k(k=0,1,\cdots,n)$上除已知函数值f_k外还给出了导数值$f_k'=m_k(k=0,1,\cdots,n)$，这样就构造一个导数连续的分段插值函数$I_h(x)$，它满足条件：

1）. $I_h(x)\in C^1[a,b]$

2）. $I_h(x_k)=f_k$， $I_h'(x_k)=f_k'$, $k=0,1,\cdots,n$；

3）. $I_h(x)$在每个区间$[x_k,x_{k+1}]$上是三次多项式，其表达式为

$$I_h(x)=\left(\frac{x-x_{k+1}}{x_k-x_{k+1}}\right)^2\left(1+2\frac{x-x_k}{x_{k+1}-x_k}\right)f_k+\left(\frac{x-x_k}{x_{k+1}-x_k}\right)^2\left(1+2\frac{x-x_{k+1}}{x_k-x_{k+1}}\right)f_{k+1}$$
$$+\left(\frac{x-x_{k+1}}{x_k-x_{k+1}}\right)^2(x-x_k)f_k'+\left(\frac{x-x_k}{x_{k+1}-x_k}\right)^2(x-x_{k+1})f_{k+1}' \qquad (4.4)$$

若在整个区间$[a,b]$上定义一组分段三次插值基函数$\alpha_j(x),\beta_j(x)$ ，$j=0,1,\cdots,n$，则$I_h(x)$可表示为

$$I_h(x)=\sum_{j=0}^{n}[f_j\alpha_j(x)+f_j'\beta_j(x)] \qquad (4.5)$$

其中

$$\alpha_j(x)=\begin{cases}\left(\dfrac{x-x_{j-1}}{x_j-x_{j-1}}\right)^2\left(1+2\dfrac{x-x_j}{x_{j-1}-x_j}\right) & x_{j-1}\le x\le x_j \\ & (j=0\text{略去}) \\ \left(\dfrac{x-x_{j+1}}{x_j-x_{j+1}}\right)^2\left(1+2\dfrac{x-x_j}{x_{j+1}-x_j}\right) & x_j\le x\le x_{j+1} \\ & (j=n\text{略去}) \\ 0 & \text{其他}\end{cases} \tag{4.6}$$

$$\beta_j(x)=\begin{cases}\left(\dfrac{x-x_{j-1}}{x_j-x_{j-1}}\right)^2(x-x_j) & x_{j-1}\le x\le x_j \\ & (j=0\text{略去}) \\ \left(\dfrac{x-x_{j+1}}{x_j-x_{j+1}}\right)^2(x-x_j) & x_j\le x\le x_{j+1} \\ & (j=n\text{略去}) \\ 0 & \text{其他}\end{cases} \tag{4.7}$$

2.5 三次样条插值

上面讨论的分段低次插值函数都有一致收敛性，但光滑性较差，只能构造一个整体上具有一阶连续微商的插值函数，且对于实际问题，要知道在节点上的微商值（特别是高阶微商值）是比较困难的．本节介绍在只给出节点上函数值的情况下构造一个整体上充分光滑的函数的方法——样条函数插值．

2.5.1 三次样条函数

定义 4 设在区间$[a,b]$上给定一个划分：

$\Delta: a=x_0<x_1<x_2<\cdots<x_{n-1}<x_n=b$, $s(x)$为实值函数，如果它满足条件：

（1） $s(x)$在子区间$[x_i,x_{i+1}](i=0,1,\cdots,n-1)$上是三次多项式；

（2） $s(x)$在$[a,b]$上具有 2 阶连续导数，即$s(x)\in C^2[a,b]$；

则称$y=s(x)$为**三次样条函数**．若在节点x_j上给定函数值

$$y_j=f(x_j)(j=0,1,\cdots,n),$$

并成立

$$S(x_j)=y_j \qquad (j=0,1,\cdots,n), \tag{5.1}$$

则称$S(x)$为**三次样条插值函数**.

从定义知道要求出$S(x)$，在每个小区间$[x_j,x_{j+1}]$上要确定 4 个待定系数，共有 n 个小

区间，故应确 4n 个参数．根据 $S(x)$ 在$[a,b]$上二阶导数连续，在节点$x_j(j=1,2,\cdots,n-1)$处应满足连续条件

$$\begin{aligned}&S(x_j-0)=S(x_j+0), S'(x_j-0)=S'(x_j+0),\\&S''(x_j-0)=S''(x_j+0).\end{aligned}\qquad(5.2)$$

共有 3n-3 个条件，再加上$S(x)$满足（5.1），共有 4n-2 个条件，因此还需要有 2 个条件才能确定$S(x)$．通常可在区间$[a,b]$端点$a=x_0,b=x_n$上各加一个条件（称为**边界条件**），可根据实际问题的要求给定．常见的有以下 3 种：

1° 已知两端的一阶导数值，即

$$S'(x_0)=f_0', S'(x_n)=f_n'.$$

2° 两端的二阶导数已知，即

$$S''(x_0)=f_0'', S''(x_n)=f_n'',$$

其特殊情况为

$$S''(x_0)=S''(x_n)=0.\qquad(5.3)$$

称为**自然边界条件**.

3° 当$f(x)$是以x_n-x_0为周期的周期函数时，则要求$S(x)$也是周期函数．这是边界条件应满足

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

而此时（5.1）中$y_0=y_n$．这样确定的样条函数$S(x)$称为**周期样条函数.**

2.5.2 样条插值函数的建立

构造满足差值条件（5.1）及相应边界条件的三次样条插值函数$S(x)$的表达式可以有多种方法，例如，可以直接利用分段三次 Hermite 插值，只要假定$S'(x_j)=m_j$，$(j=0,1,\cdots,n)$，再由（5.1）可得

$$S(x)=\sum_{j=0}^{n}[y_j a_j(x)+m_j\beta_j(x)]\qquad(5.4)$$

其中$\alpha_j(x),\beta_j(x)$是由（2.7），（2.8）表示的插值基函数，利用条件（5.2）及相应边界条件可得到关于$m_j(j=0,1,\cdots,n)$的三对角方程组，求出m_j，则得到所求的三次样条函数

$S(x)$.

下面我们利用 $S(x)$ 的二阶导数值 $S''(x_j)=M_j\,(j=0,1,\cdots,n)$ 表达 $S(x)$，由于 $S(x)$ 在区间 $[x_j,x_{j+1}]$ 上是三次多项式，故 $S''(x)$ 在 $[x_j,x_{j+1}]$ 上是线性函数，可表示为

$$S''(x)=M_j\frac{x_{j+1}-x}{h_j}+M_{j+1}\frac{x-x_j}{h_j}. \tag{5.5}$$

对 $S''(x)$ 积分两次并利用 $S(x_j)=y_j$ 及 $S(x_{j+1})=y_{j+1}$，可定出积分常数，于是得三次样条表达式

$$S(x)=M_j\frac{(x_{j+1}-x)^3}{6h_j}+M_{j+1}\frac{(x-x_j)^3}{6h_j}+\left(y_j-\frac{M_jh_j^2}{6}\right)\frac{x_{j+1}-x}{h_j}+\left(y_{j+1}-\frac{M_{j+1}h_j^2}{6}\right)\frac{x-x_j}{h_j}$$

$$(j=0,1,\cdots,n-1) \tag{5.6}$$

这里 $M_j,\ j=0,1,\cdots,n$，是未知的．为了确定 $M_j,\ j=0,1,\cdots,n$，对 $S(x)$ 求导得

$$S'(x)=-M_j\frac{(x_{j+1}-x)^2}{2h_j}+M_{j+1}\frac{(x-x_j)^2}{2h_j}+\frac{y_{j+1}-y_j}{h_j}-\frac{M_{j+1}-M_j}{6}h_j \tag{5.7}$$

由此可求得

$$S'(x_j+0)=-\frac{h_j}{3}M_j-\frac{h_j}{6}M_{j+1}+\frac{y_{j+1}-y_j}{h_j}.$$

类似地可求出 $S(x)$ 在区间 $[x_{j-1},x_j]$ 上的表达式，从而得

$$S'(x_j-0)=\frac{h_{j-1}}{6}M_{j-1}+\frac{h_{j-1}}{3}M_j+\frac{y_j-y_{j-1}}{h_{j-1}},$$

利用 $S'(x_j-0)=S'(x_j+0)$ 可得

$$\mu_jM_{j-1}+2M_j+\lambda_jM_{j+1}=d_j \qquad (j=1,2,\cdots n-1),$$

其中

$$\mu_j=\frac{h_{j-1}}{h_{j-1}+h_j},\lambda_j=\frac{h_j}{h_{j-1}+h_j},\ j=0,1,\cdots,n$$

$$d_j=6\frac{f[x_j,x_{j+1}]-f[x_{j-1},x_j]}{h_{j-1}+h_j}=6f[x_{j-1},x_j,x_{j+1}].$$

对第一种边界条件，可导出两个方程

$$
\begin{cases}
2M_0 + M_1 = \dfrac{6}{h_0}(f[x_0,x_1] - f'_0), \\
M_{n-1} + 2M_n = \dfrac{6}{h_{n-1}}(f'_n - f[x_{n-1},x_n]).
\end{cases}
$$

如果令

$$
\lambda_0 = 1,\ d_0 = \frac{6}{h_0}(f[x_0,x_1] - f'_0),\ \mu_n = 1,\ d_n = \frac{6}{h_{n-1}}(f'_n - f[x_{n-1},x_n]),
$$

那么有

$$
\begin{pmatrix}
2 & \lambda_0 & & & \\
\mu_1 & 2 & \lambda_1 & & \\
 & \ddots & \ddots & \ddots & \\
 & & \mu_{n-1} & 2 & \lambda_{n-1} \\
 & & & \mu_n & 2
\end{pmatrix}
\begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ M_{n-1} \\ M_n \end{pmatrix}
=
\begin{pmatrix} d_0 \\ d_1 \\ \vdots \\ d_{n-1} \\ d_n \end{pmatrix}. \qquad (5.8)
$$

对第二种边界条件，直接得端点方程

$$
M_0 = f_0'',\ M_n = f_n''.
$$

如果令$\lambda_0 = \mu_n = 0$，$d_0 = 2f_0''$，$d_n = 2f_n''$，则也有对应的矩阵形式.

对于第三种边界条件，可得

$$
M_0 = M_n, \qquad \lambda_n M_1 + \mu_n M_{n-1} + 2M_n = d_n,
$$

可以写成矩阵的形式

$$
\begin{pmatrix}
2 & \lambda_1 & & & \mu_1 \\
\mu_2 & 2 & \lambda_2 & & \\
 & \ddots & \ddots & \ddots & \\
 & & \mu_{n-1} & 2 & \lambda_{n-1} \\
\lambda_n & & & \mu_n & 2
\end{pmatrix}
\begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_{n-1} \\ M_n \end{pmatrix}
=
\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n-1} \\ d_n \end{pmatrix}. \qquad (5.9)
$$

（5.8）和（5.9）是关于$M_j\ (j=0,1,\cdots,n)$的三对角方程组，M_j在力学上解释为细梁在x_j截面处的弯矩，称为$S(x)$的距，方程组（5.8）和（5.9）称为三弯矩方程.（5.8）和（5.9）的系数矩阵中元素λ_j, μ_j已完全确定.并且满足$\lambda_j \ge 0, \mu_j \ge 0, \lambda_j + \mu_j = 1$. 因此系数矩阵为严格对角占优阵，从而（5.8）和（5.9）有唯一解.

例 4　已知未知函数$y = f(x)$的一组数据

x_j	0.2	0.4	0.6	0.8	1.0
y_j	0.9798	0.8033	0.4177	0.6386	1.1843

求满足上述插值条件及第一类边界条件

$$S'(0.2)=0.2027, \quad S'(1.0)=1.5574$$

的三次插值样条函数$S(x)$.

解 $h_j=x_{j+1}-x_j=0.2, \quad j=0,1,2,3$，

$$x_0=0.2, \quad x_1=0.4, \quad x_2=0.6, \quad x_3=0.8, \quad x_4=1.0;$$

$$y_0=0.9798, \quad y_1=0.8033, \quad y_2=0.4177, \quad y_3=0.6386, \quad y_4=1.1843;$$

$$y_0'=0.2027, \quad y_4'=1.5574$$

故

$$\lambda_0=1, \quad \lambda_1=\lambda_2=\lambda_3=0.5, \quad \mu_1=\mu_2=\mu_3=0.5, \quad \mu_4=1,$$

$$d_0=-32.556, \quad d_1=-15.6825, \quad d_2=45.4875, \quad d_3=24.36, \quad d_4=-35.133,$$

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0.5 & 2 & 0.5 & 0 & 0 \\ 0 & 0.5 & 2 & 0.5 & 0 \\ 0 & 0 & 0.5 & 2 & 0.5 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} M_0 \\ M_1 \\ M_2 \\ M_3 \\ M_4 \end{pmatrix} = \begin{pmatrix} -32.556 \\ -15.6825 \\ 45.4875 \\ 24.36 \\ -35.133 \end{pmatrix},$$

解得

$$M_0=-10.9416, \quad M_1=-10.6728, \quad M_2=22.2677, \quad M_3=12.5768, \quad M_4=-23.8549,$$

从而可得

$$S(x)=0.7186+2.4179x-5.6052x^2+0.2240x^3, \quad x\in[0.2,\ 0.4);$$
$$S(x)=-1.0239+15.4866x-38.2769x^2+27.4504x^3, \quad x\in[0.4,\ 0.6);$$
$$S(x)=6.6498-22.8818x+25.6703x^2-8.0758x^3, \quad x\in[0.6,\ 0.8);$$
$$S(x)=18.0592-65.6669x+79.1518x^2-30.3597x^3, \quad x\in[0.8,\ 1.0].$$

图形见(图 2—3).

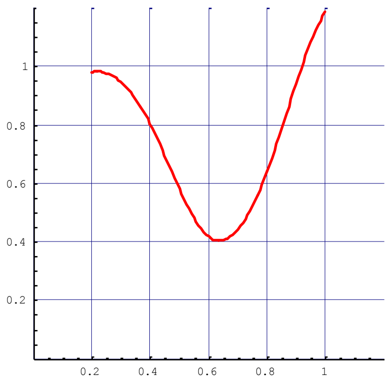


图 2—3

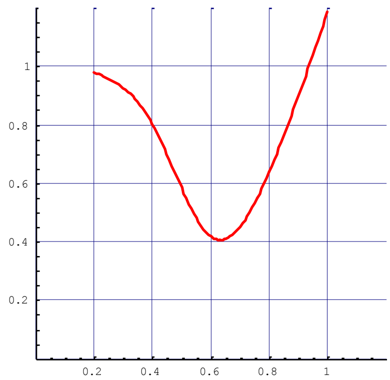


图 2—4

例 5 已知未知函数$y=f(x)$的一组数据

x_j	0.2	0.4	0.6	0.8	1.0
y_j	0.9798	0.8033	0.4177	0.6386	1.1843

求满足上述插值条件及第二类边界条件

$$S''(0.2)=S''(1.0)=0$$

的三次插值样条函数$S(x)$.

解　$h_j=x_{j+1}-x_j=0.2,\quad j=0,1,,$

$$x_0=0.2,\quad x_1=0.4,\quad x_2=0.6,\quad x_3=0.8,\quad x_4=1.0;$$

$$y_0=0.9798,\quad y_1=0.8033,\quad y_2=0.4177,\quad y_3=0.6386,\quad y_4=1.1843;$$

$$y_0''=y_4''=0$$

$$\begin{pmatrix}2&\lambda_1&0\\ \mu_2&2&\lambda_2\\ 0&\mu_3&2\end{pmatrix}\begin{pmatrix}M_1\\ M_2\\ M_3\end{pmatrix}=\begin{pmatrix}d_1\\ d_2\\ d_3\end{pmatrix}$$

故

$$\mu_2=\mu_3=0.5,\quad \lambda_1=\lambda_2=0.5,\quad d_1=-15.6825,\quad d_2=45.4875,\quad d_3=24.36$$

从而得

$$\begin{pmatrix}2&0.5&0\\ 0.5&2&0.5\\ 0&0.5&2\end{pmatrix}\begin{pmatrix}M_1\\ M_2\\ M_3\end{pmatrix}=\begin{pmatrix}-15.6825\\ 45.4875\\ 24.36\end{pmatrix}$$

有

$$M_0=0,\quad M_1=-14.0296,\quad M_2=24.7532,\quad M_3=5.9917,\quad M_4=0$$

解得

$$S(x)=1.1563-1.1818x+7.0148x^2-11.6913x^3,\ x\in[0.2,\ 0.4);$$

$$S(x)=-1.6604+19.3071x-45.7975x^2+32.3190x^3,\ x\in[0.4,\ 0.6);$$

$$S(x)=8.6976-32.4827x+40.5189x^2-15.6346x^3,\ x\in[0.6,\ 0.8);$$

$$S(x)=3.2492-12.0510x+14.9792x^2-4.9931x^3,\ x\in[0.8,\ 1.0].$$

图形见(图 2-4).

2.5.3 误差界与收敛性

三次样条函数的收敛性与误差估计比较复杂，这里不加证明地给出一个主要结果.

定理 4　设 $f(x)\in C^4[a,b], S(x)$ 为满足第一种或第二种边界条件的三次样条函数，

令 $h=\max\limits_{0\le i\le n-1} h_i, h_i=x_{i+1}-x_i(i=0,1,\cdots,n-1)$, 则有估计式

$$\max_{a\le x\le b}\left|f^{(k)}(x)-S^{(k)}(x)\right|\le C_k\max_{a\le x\le b}\left|f^{(4)}(x)\right|h^{4-k},\quad k=0,1,2,$$

其中 $C_0=\dfrac{5}{384}, C_1=\dfrac{1}{24}, C_2=\dfrac{3}{8}$.

这个定理不但给出了三次样条插值函数 $S(x)$ 的误差估计，并且 $h\to 0$ 时，$S(x)$ 及其一阶导数 $S'(x)$和二阶导数$S''(x)$ 均分别一致收敛于 $f(x), f'(x)$及$f''(x)$.

2.6 曲线拟合的最小二乘法

在实际问题中，常要从一组观测数据 $\{(x_i,y_i),i=0,1,\cdots,m\}$ 去预测函数 $f(x)$ 的近似表达式，且从给定的一组实验数据出发，寻求 $f(x)$ 的一个逼近函数 $y=S(x)$，使得逼近函数从总体上来说与已知函数的偏差按某种度量能达到最小而又不一定过全部的点 $\{(x_i,y_i),i=0,1,\cdots,m\}$，即是最小二乘曲线拟合。

用最小二乘法求拟合曲线时，首先要确定 $S(x)$ 的形式. 这不单纯是数学问题，还与所研究问题的运动规律及所得观测数据 $(x_{i,},y_i)$ 有关，通常要从问题的运动规律及给定数据描图，确定 $S(x)$ 的形式，并通过实际计算选出较好的结果.

1. 直线拟合

假设所给数据{ $(x_i,y_i),i=0,1,\cdots,m$ }的分布大致成一条直线。虽然并不要求所作的拟合直线 $S_1(x)=a+bx$ 严格地通过所有的数据点{ $(x_i,y_i),i=0,1,\cdots,m$ }. 但希望它能尽可能地从所给数据点附近通过。记误差

$$\delta_i=S_1(x_i)-y_i=a-bx_i-y_i, i=0,1,\cdots,m.$$

在原始数据给定以后，误差仅依赖于 a,b 的选取。因此，把误差的大小作为衡量 a,b
好坏的主要标志。通常要求使误差的平方和达到最小，即要求

$$Q=\sum_{i=0}^{m}\delta_i^2=\sum_{i=0}^{m}[a+bx_i-y_i]^2 \tag{6.1}$$

为最小。

由微积分中求极值的方法知，使 Q 达到极值的参数 a,b 应满足

$$\frac{\partial Q}{\partial a}=0,\quad \frac{\partial Q}{\partial b}=0$$

即有

$$\begin{cases}\dfrac{\partial Q}{\partial a}=2\sum_{i=0}^{m}(a+bx_i-y_i)=0\\ \dfrac{\partial Q}{\partial b}=2\sum_{i=0}^{m}(a+bx_i-y_i)x_i=0\end{cases}\tag{6.2}$$

（6.2）称为正规方程组。解方程组（6.2）可以得出 a,b

直线方程 $S_1(x)=a+bx$ 便可以确定。

2. 多项式拟合

有时所给数据点用直线拟合不合适，这时可考虑用多项式拟合。对于给定的数据 $\{(x_i,y_i),i=0,1,\cdots,m\}$，寻找 n 次多项式 $S_n(x)=\sum_{j=0}^{n}a_jx^j$． 使误差的平方和

$$Q=\sum_{i=0}^{m}[S_n(x_i)-y_i]^2.\tag{6.3}$$

这转化为求多元函数

$$Q(a_0,a_1,\cdots,a_n)=\sum_{i=0}^{m}[\sum_{j=0}^{n}a_jx_i^j-y_i]^2\tag{6.4}$$

的极小点 $(a_0^*,a_1^*,\cdots,a_n^*)$ 问题．由多元函数极值的必要条件，有

$$\frac{\partial Q}{\partial a_k}=2\sum_{i=0}^{m}[\sum_{j=0}^{n}a_jx_i^j-y_i]x_i^k=0,\qquad (k=0,1,\cdots,n).$$

化简得 $a_0,a_1,\cdots,a_n$ 满足的正规方程组

$$\begin{cases}a_0(m+1)+a_1\sum_{i=0}^{m}x_i+a_2\sum_{i=0}^{m}x_i^2+\cdots+a_n\sum_{i=0}^{m}x_i^n=\sum_{i=0}^{m}y_i\\ a_0\sum_{i=0}^{m}x_i+a_1\sum_{i=0}^{m}x_i^2+a_2\sum_{i=0}^{m}x_i^3+\cdots+a_m\sum_{i=0}^{m}x_i^{n+1}=\sum_{i=0}^{m}x_iy_i\\ \cdots\cdots\\ a_0\sum_{i=0}^{m}x_i^n+a_1\sum_{i=0}^{m}x_i^{n+1}+a_2\sum_{i=0}^{m}x_i^{n+2}+\cdots+a_m\sum_{i=0}^{m}x_i^{2n}=\sum_{i=0}^{m}x_i^ny_i\end{cases}.\tag{6.5}$$

定理 5 正规方程组（6.5）有唯一解。

证明： 反证法。如不然，则对应的齐次方程

$$\sum_{j=0}^{n} a_j \sum_{i=0}^{m} x_i^{j+k} = 0 \quad (k = 0,1,\cdots,n).$$

有非零解，从而有

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

因此有 $\sum_{j=0}^{n} a_j x_i^j = 0, i = 1,2,\cdots m$，即拟合多项式有 m 个零点 $x_i, i = 1,2,\cdots m$ 。当 $m > n$ 时，

由代数基本定理必有 $\sum_{j=0}^{n} a_j x_i^j \equiv 0,$ 从而 $a_j, j = 0,1,\cdots,n$ ，这就导出矛盾。

定理 6 设 $a_0, a_1, \cdots, a_n$ 为正规方程组（6.5）的解，则 n 次多项式 $S_n(x) = \sum_{j=0}^{n} a_j x^j$ 必是所求多项式。

有时根据给定数据图形，其拟合函数 $y = S(x)$ 表面上不是线性形式，但通过变换仍可化为线性模型．例如，$S(x) = ae^{bx}$，若两边取对数得

$$\ln S(x) = \ln a + bx,$$

它就是形如（6.2）的线性模型，具体做法见例 8.

例 6 已知一组实验数据如下，求它的拟合曲线．

x_i	1	2	3	4	5
f_i	4	4.5	6	8	8.5

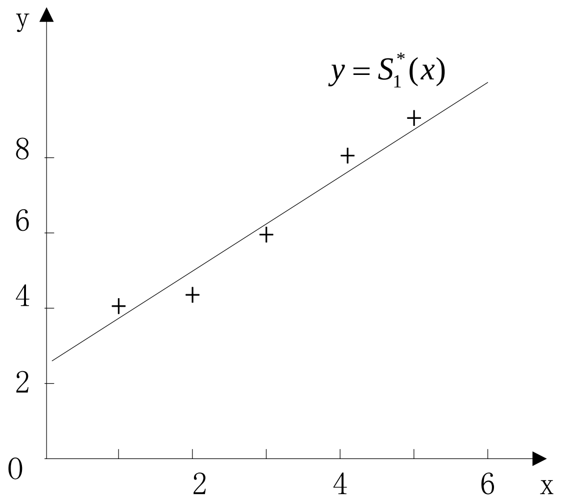


图 2—5

解 根据所给数据，在坐标纸上标出，见图 2-5.从图中看到各点在一条直线附近，故可选择线性函数作拟合曲线，即令 $S_1(x)=a_0+a_1x,$

由（6.5）得的方程组

$$\begin{cases}8a_0+22a_1=47,\\22a_0+74a_1=145.5.\end{cases}$$

解得 $a_0=2.77, a_1=1.13.$ 于是所求拟合曲线为

$$S_1^*(x)=2.77+1.13x\,.$$

例 7 设数据 $(x_i,y_i)(i=0,1,2,3,4)$ 由表 2-3 给出，表中第 4 行为 $\ln y_i=\overline{y_i}$，可以看出数学模型为 $y=ae^{bx}$，用最小二乘法确定 a 及 b .

解 根据给定数据 $(x_i,y_i)(i=0,1,2,3,4)$ 描图可确定拟合曲线方程为 $y=ae^{bx}$，它不是线性形式．两边取对数得 $\ln y=\ln a+bx$，若令 $\overline{y}=\ln y, A=\ln a,$ 则得 $\overline{y}=A+bx, \varphi=\{1,x\}$．为确定 A,b，先将 (x_i,y_i) 转化为 $(x_i,\overline{y_i})$，数据表见表 2-3.

表 2-3 计算结果

i	0	1	2	3	4
x_i	1.00	1.25	1.50	1.75	2.00
y_i	5.10	5.79	6.53	7.45	8.46
$\overline{y_i}$	1.629	1.756	1.876	2.008	2.135

由（6.5）得的方程组

$$\begin{cases} 5A + 7.50b = 9.404 \\ 7.50A + 11.875b = 14.422. \end{cases}$$

解得 $A = 1.122, b = 0.505, a = e^{A} = 3.071.$ 于是得最小二乘拟合曲线为

$$y = 3.071e^{0.505x}.$$

2.7 数值实验及程序

2.7.1 Lagrange 插值数值实验及程序

给定 $n+1$ 个插值节点下标如下表：

x	x_0	x_1	……	x_n
y	y_0	y_1	……	y_n

则可构造 n 次插值多项式 $L_n(x) = \sum_{k=0}^{n} y_k l_k(x)$，

其中插值基函数

$$l_k(x) = \frac{(x - x_0)\cdots(x - x_{k-1})(x - x_{k+1})\cdots(x - x_n)}{(x_k - x_0)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)} \qquad (k = 0,1,\cdots,n)$$

```
matlab 程序如下：（Lagran_.m）
% Lagrange 插值
% x 为横坐标向量，y 为纵坐标向量，x 与 y 必须同维；
% 计算 Lagrange 插值多项式在 xi（可为矩阵）处的值.
function fi=Lagran_(x,y,xi)
fi=zeros(size(xi)); % fi 为 Lagrange 插值多项式在 xi 处的值；初值为全零阵.
n=length(x);                % n 为插值节点的个数.
for i=1:n
    z=ones(size(xi));   %全 1 阵
    for j=1:n
        if i~=j
            z=z.*(xi-x(j))/(x(i)-x(j));   %插值基函数在 xi 处的值
        end
    end
    fi=fi+z*y(i);
end
return
```

计算实例：

例 8 已给 sin0.32=0.314567,sin0.34=0.333487,sin0.36=0.352274, 利用 Lagrange 插值计算 sin0.3367 的值

输入：

```
x=[0.32,0.34,0.36];
y=[0.314567,0.333487,0.352274];
fi=Lagran_(x,y,0.3367)
```

输出：

```
fi =
0.33037436203750
```

这个程序对于给定的插值节点，就能计算出对应的 Lagrange 插值多项式在相应点处的值. 计算速度较快，适用于插值节点较多的情形. 也可用 matlab 中的函数 polyfit(x,y,xi)来代替.

上面的程序只能计算出 Lagrange 插值多项式在相应节点处的值，并没有给出 Lagrange 插值多项式的具体形式，因此无法进行误差估计.

2.7.2 Runge 现象数值实验

Runge（龙格）现象是指当插值节点较多时，有可能出现构造的高次插值多项式 $L_n(x)$ 近似效 $f(x)$ 果并不好，甚至相差很大的情形．这是因为对任意的插值节点，当 $n\to\infty$ 时，$L_n(x)$ 不一定收敛到 $f(x)$ 的原因.

下面通过一个具体的例子，利用插值余项来说明 Runge 现象的产生.

例 9：$f(x)=\dfrac{1}{(1+x^2)}$，它在 $[-5,5]$ 上各阶导数均存在，在 $[-5,5]$ 上取 $n+1$ 个等距节点所构造的 Lagrange 插值多项式．分别对 $n=3,9,16,21$ 的情形进行讨论，画出相应图形，计算出误差估计值 $\left|R_n(x)\right|$.

matlab 程序如下：（文件 L_R.m）

```
% 画图
x1 = linspace(-5,5,3);
y1 = 1./(1+x1.^2);
c1 = polyfit(x1,y1,2);
t1 = linspace(x1(1),x1(end),100);
p1 = polyval(c1,t1);
subplot(2,2,1);
plot(t1,p1,x1,y1,'o')
title('3 个插值节点')
x2 = linspace(-5,5,9);
y2 = 1./(1+x2.^2);
```

```
c2 = polyfit(x2,y2,8);
t2 = linspace(x2(1),x2(end),100);
p2 = polyval(c2,t2);
subplot(2,2,2);
plot(t2,p2,x2,y2,'o')
title('9 个插值节点')
x3 = linspace(-5,5,15);
y3 = 1./(1+x3.^2);
c3 = polyfit(x3,y3,14);
t3 = linspace(x3(1),x3(end),100);
p3 = polyval(c3,t3);
subplot(2,2,3);
plot(t3,p3,x3,y3,'o')
title('15 个插值节点')
x4 = linspace(-5,5,21);
y4 = 1./(1+x4.^2);
c4 = polyfit(x4,y4,20);
t4 = linspace(x4(1),x4(end),100);
p4 = polyval(c4,t4);
subplot(2,2,4);
plot(t4,p4,x4,y4,'o')
title('21 个插值节点')
%  计算误差
syms x;
n=20;
f='1/(1+x^2)';
df=diff(f,1);
cdf=char(df); % converts the symbolic expression into a string

%  计算 f(x)的 n+1 阶导数的绝对值的最大值
for i=1:n+1
        df=diff(df,1);
        cdfn=char(df);
        x=fzero(cdfn,0);
        M(i)=abs(eval(cdf));
        cdf=cdfn;
end

%  计算(x-x0)(x-x1)……(x-xn)/(n+1)!的最大值
t=linspace(-5,5,10000);
for n=0:20
        h=10/(n+1);
        x=[-5:h:5];
        c=poly(x);
```

```
    r(n+1)=max(polyval(c,t));
    r(n+1)=r(n+1)/prod([1:n+1]);
end

% 打印相应的误差估计
w=M([3,9,15,21]).*r([3,9,15,21]);
fprintf('\n3 个节点时的误差估计为：  %f\n',w(1))
fprintf('\n9 个节点时的误差估计为：  %f\n',w(2))
fprintf('\n16 个节点时的误差估计为：%f\n',w(3))
fprintf('\n21 个节点时的误差估计为：%f\n',w(4))
```

计算结果：

3 个节点时的误差估计为： 54.034246

9 个节点时的误差估计为： 16800.898699

16 个节点时的误差估计为：16706459.342416

21 个节点时的误差估计为：22877132737.290627

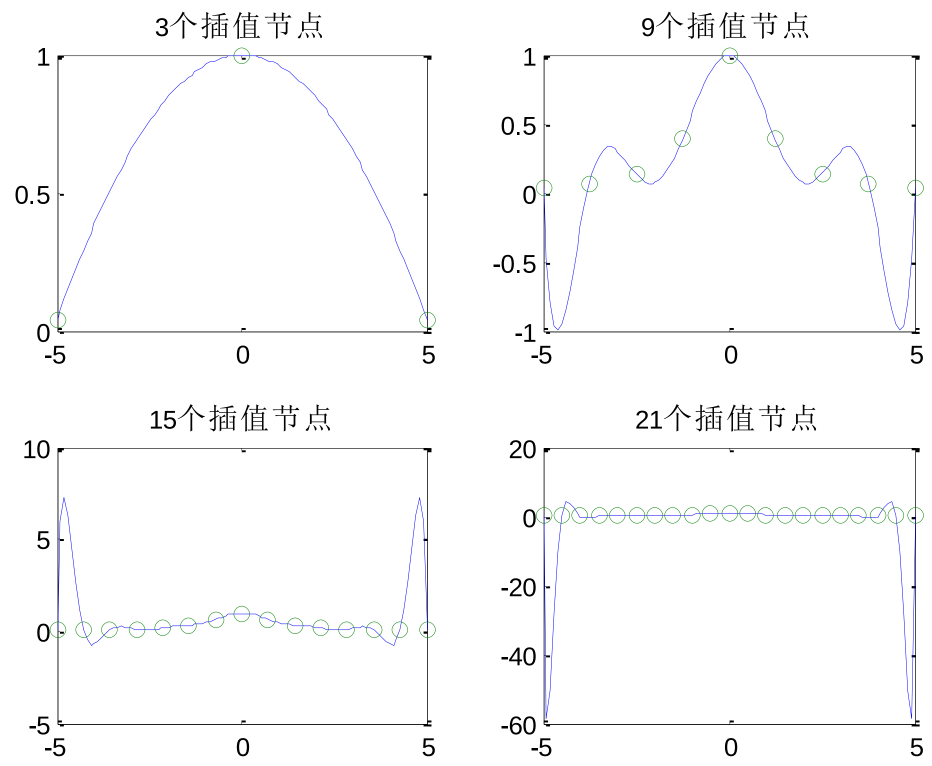


图 2-1 Runge 现象

由图形可见，当 x 越接近两个端点 ± 5 时，误差就越大. 另外，当 n 越大时，截断误差限 $R_n(x)$ 增长得非常快，这说明有的情形下，高次插值多项式 $L_n(x)$ 近似 $f(x)$ 效果并不好，甚至可以说非常差. 为了解决这个问题，可以考虑分段低次插值多项式.

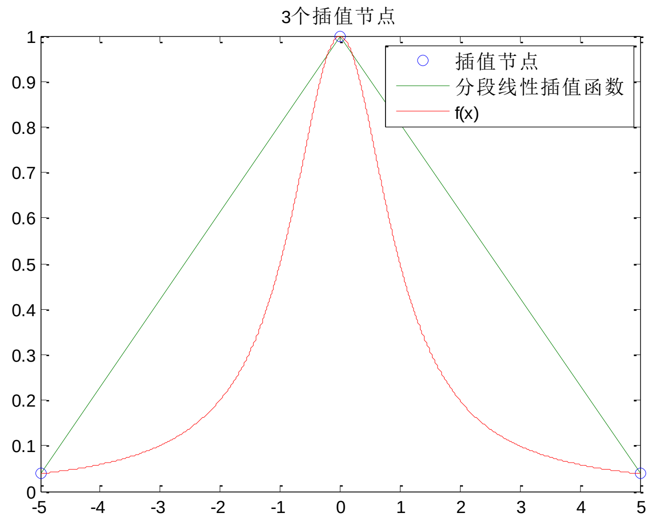
2.7.3 分段线性插值数值实验及程序

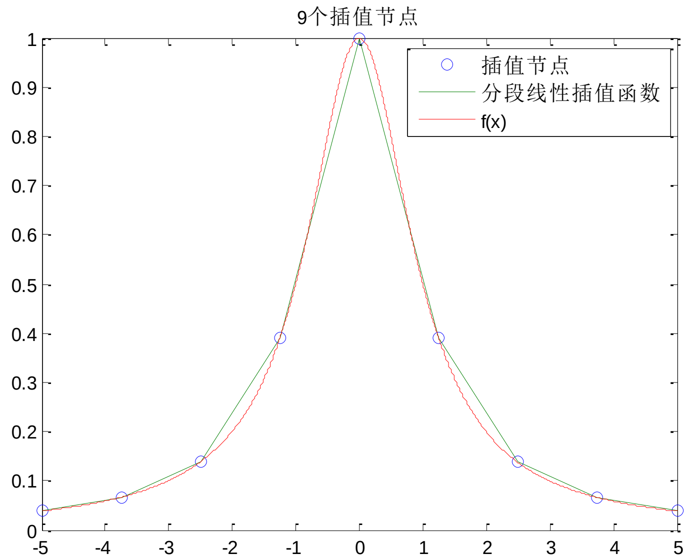
分段线性插值就是通过插值点用折线段连接起来逼近 $f(x)$，其截断误差限 $R(x) \le \frac{M_2}{8}h^2$，其中 $M_2 = \max_{a \le x \le b}|f''(x)|$，$h$ 为步长．利用分段线性插值来考虑 2.5 节实验中的例题．

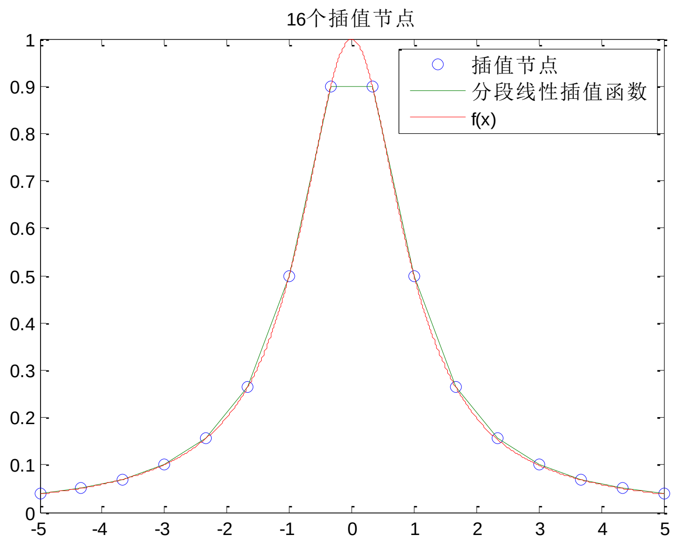
matlab 程序如下：（fd.m）

```
% 分段线性插值(等步长)
x=-5:0.00001:5; y =   1./(1+s.^2); % 画图
x1 = linspace(-5,5,3);y1 =   1./(1+x1.^2);
subplot(2,2,1); plot(x1,y1,'o',x1,y1,'-',x,y)
legend('插值节点','分段线性插值函数','f(x)')
title('3 个插值节点')
x2 = linspace(-5,5,9);y2 =   1./(1+x2.^2); subplot(2,2,2);plot(x2,y2,'o',x2,y2,'-',x,y)
legend('插值节点','分段线性插值函数','f(x)');title('9 个插值节点')
x3 = linspace(-5,5,16);y3 =   1./(1+x3.^2);subplot(2,2,3);plot(x3,y3,'o',x3,y3,'-',x,y)
legend('插值节点','分段线性插值函数','f(x)');title('16 个插值节点')
x4 = linspace(-5,5,21);y4 =   1./(1+x4.^2);subplot(2,2,4);plot(x4,y4,'o',x4,y4,'-',x,y)
legend('插值节点','分段线性插值函数','f(x)')
title('21 个插值节点')
% 计算插值多项式在 xi=4.8 的值，节点数 n=16.
xi=4.8; n=length(x4); h=(x4(n)-x4(1))/(n-1);
for i=0:n-1
    if xi<=x4(1)+i*h
        b=x4(1)+i*h;t=i;          break;
    end
end
if i>n-1
    fprintf('ERROR');
else
   fi=(xi-b)/(-h)*y4(t)+(xi-b+h)/h*y4(t+1)
end
```

计算结果如图 2—2：


3个插值节点
插值节点
分段线性插值函数
f(x)


9个插值节点
插值节点
分段线性插值函数
f(x)


16个插值节点
插值节点
分段线性插值函数
f(x)

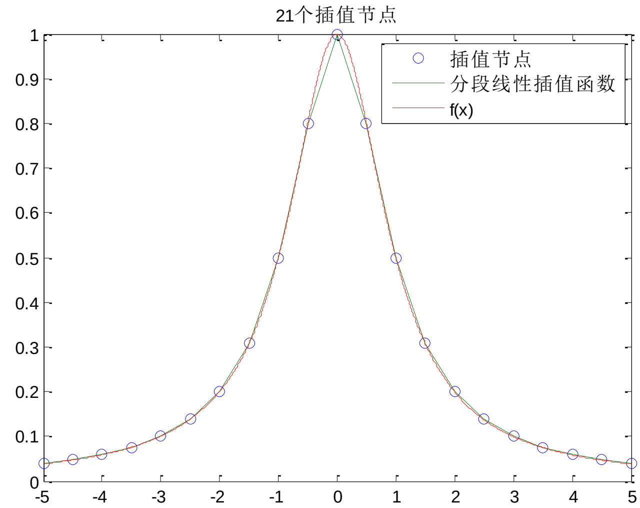


图 2—2 分段线性插值

考察在端点附近的逼近效果，接近端点 5 附近 $x_i = 4.8$ 处，真实值 $f(4.8) = 1/(1+4.8^2) = 0.04159733777038$.

当 $n = 21$ 时，利用分段线性插值多项式计算的值为 fi=0.04190045248869，逼近效果较好.

习　题　二

1．当 $x = 1,-1,2$ 时，$f(x) = 0,-3,4$，求 $f(x)$ 的二次插值多项式

2．给出 $f(x) = \ln(x)$ 的数值表

x	0.4	0.5	0.6	0.7	0.8
$\ln x$	-0.916291	-0.693147	-0.510826	-0.356675	-0.223144

用线性插值及二次插值计算 $\ln 0.54$ 的近似值.

3．给出 $\cos x$, $0^\circ \le x \le 90^\circ$ 的函数表，步长 $h = 1' = \left(1/60\right)^\circ$，若函数表具有 5 位有效数字，研究用线性插值求 $\cos x$ 的近似值时的总误差界.

4．设 x_j 为互异节点 $(j = 0,1,\cdots,n)$，求证：

i）$\sum_{j=0}^{n} x_j^k l_j(x) \equiv x^k \quad (k = 0,1,\cdots,n)$；

ii） $\sum_{j=0}^{n}(x_j-x)^k l_j(x)\equiv 0 \quad (k=0,1,\cdots,n)$.

5. 给定数据表： $i=1,2,3,4,5$ ，

x_i	1	2	4	6	7
$f(x_i)$	4	1	0	1	1

求 4 次牛顿插值多项式，并写出插值余项。

6．若 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}+a_nx^n$ 有 n 个不同实数根 $x_1,x_2,\cdots,x_n$ ，证明：

$$\sum_{j=1}^{n}\frac{x_j^k}{f'(x_j)}=\begin{cases}0,0\le k\le n-2;\\ a_n^{-1},k=n-1.\end{cases}$$

7．证明 n 阶差商有下列性质：

i）若 $F(x)=cf(x),则F[x_0,x_1,\cdots x_n]=cf[x_0,x_1,\cdots x_n]$；

ii）若 $F(x)=f(x)+g(x)$ ，则 $F[x_0,x_1,\cdots x_n]=f[x_0,x_1,\cdots x_n]+g[x_0,x_1,\cdots x_n]$.

8． $f(x)=x^7+x^4+3x+1$ ，求 $f[2^0,2^1,\cdots,2^7]及f[2^0,2^1,\cdots,2^8]$.

9．证明两点三次 Hermite 插值余项是

$R_3(x)=f^{(4)}(\xi)(x-x_k)^2(x-x_{k+1})^2/4!$ ， $\xi\in(x_k,x_{k+1})$, 并由此求出分段三次埃而米插值的误差限.

10．求一个次数不高于 4 次的多项式 $P(x)$ 使它满足 $P(0)=P'(0)=0$ ， $P(1)=P'(1)=1$，

$P(2)=1$.

11．设 $f(x)=1/(1+x^2), -5\le x\le 5$ ，求在整数节点处的 $I_h(x)$ 值.

12．求 $f(x)=x^4$ 在 $[a,b]$ 上的 $I_h(x)$ ，并估计误差.

13．给定数据表如下

x_j	0.25	0.30	0.39	0.45	0.53
y_j	0.5000	0.5477	0.6245	0.6708	0.7280

试求三次样条插值 $S(x)$ ，并满足条件：

i） $S'(0.25)=1.0000, S'(0.53)=0.6868$；　ii） $S''(0.25)=S''(0.53)=0$.

14．已知实验数据如下：

x_i	19	25	31	38	44
y_i	19.0	32.3	49.0	73.3	97.8

用最小二乘法求形如 $y=a+bx^2$ 的经验公式，并计算均方误差.

15. 有函数如下表，要求用公式 $y=a+bx^3$ 拟合所给数据，试确定拟合公式中的 a 和 b。

x_i	−3	−2	−1	0	1	2	3
y_i	−1.76	0.42	1.20	1.34	1.43	2.25	4.38