现代数值计算方法

第四章 插值法与最小二乘拟合

Back
Close

第四章 插值法与最小二乘拟合

已知函数 $y = f(x)$ 的一批数据 $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$, 而函数的表达式未知, 要从某类函数 (如多项式函数、样条函数等) 中求得一个函数 $\varphi(x)$ 作为 $f(x)$ 的近似, 这类数值计算问题称为数据建模. 有时尽管 $y = f(x)$ 有表达式, 但比较复杂, 我们也利用该方法建立一个近似模型.

数据建模有两大类方法: 一类是插值方法, 要求所求函数 $\varphi(x)$ 严格遵从数据 $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$; 另一类是拟合方法, 允许函数 $\varphi(x)$ 在数据点上有误差, 但要求达到某种误差指标最小化. 其中, 以误差向量的 2-范数为误差指标的数据拟合称为最小二乘拟合. 一般而言, 插值方法比较适合数据准确或数据量较小的情形, 而拟合方法

Back
Close

则比较适合数据有误差或数据量较大的情形.

§4.1 多项式插值

§4.1.1 插值多项式的概念

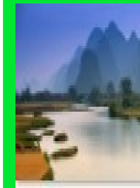
在众多的函数中, 多项式最简单、最容易计算. 因此, 已知函数 $y=f(x)$ 在 $n+1$ 个互不相同的点处的函数值 $y_i=f(x_i),\ i=0,1,\cdots,n$, 为求 $y=f(x)$ 的近似式, 首先考虑的自然应当是选取 n 次多项式

$$P_n(x)=a_0+a_1x+\cdots+a_nx^n, \tag{4.1}$$

使 $P_n(x)$ 满足条件

$$P_n(x_i)=y_i,\quad i=0,1,\cdots,n. \tag{4.2}$$

函数 $f(x)$ 称为被插函数, $P_n(x)$ 称为插值多项式, 条件 (4.2) 称为插值



条件, $x_0, x_1, \cdots, x_n$ 称为插值节点. 这种求函数近似式的方法称为插值法.

满足插值条件 (4.2) 的插值多项式 $P_n(x)$ 是唯一存在的. 事实上, 插值条件 (4.2) 可看成未知数是 $a_0, a_1, \cdots, a_n$ 的线性方程组

$$\begin{cases} a_0 + x_0 a_1 + \cdots + x_0^n a_n = y_0, \\ a_0 + x_1 a_1 + \cdots + x_1^n a_n = y_1, \\ \cdots\cdots\cdots\cdots\cdots \\ a_0 + x_n a_1 + \cdots + x_n^n a_n = y_n. \end{cases} \tag{4.3}$$

因系数行列式为范德蒙行列式

$$D = \begin{vmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{vmatrix} = \prod_{0 \le i < j \le n} (x_j - x_i) \neq 0,$$

Back
Close

知 (4.3) 必有唯一解.

§4.1.2 插值多项式的截断误差

可以证明, 如果被插函数 $y = f(x)$ 在包含插值节点 $x_0, x_1, \cdots, x_n$ 的区间 $[a, b]$ 上存在 $n+1$ 阶导数, 则在区间 $[a, b]$ 任意点 x 处, 被插函数 $f(x)$ 与插值多项式 $P_n(x)$ 的截断误差为 (利用泰勒公式):

$$R_n(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega(x), \tag{4.4}$$

其中

$$\omega(x) = (x - x_0)(x - x_1)\cdots(x - x_n) = \prod_{j=0}^{n}(x - x_j), \tag{4.5}$$

ξ 介于 x 与节点 $x_0, x_1, \cdots, x_n$ 之间.

事实上, 当 x 为节点时, (4.4) 两边皆为零, 等式显然成立. 下面假

定 x 不是节点. 作辅助函数

$$\varphi(t) = R_n(t) - \frac{R_n(x)}{\omega(x)}\omega(t).$$

不难发现

$$\varphi(x) = R_n(x) - \frac{R_n(x)}{\omega(x)}\omega(x) = 0,$$

$$\varphi(x_i) = f(x_i) - P_n(x_i) - \frac{R_n(x)}{\omega(x)}\omega(x_i) = 0,$$

$$(i = 0, 1, \cdots, n),$$

即 $\varphi(t)$ 存在 $n+2$ 个零点 $x, x_0, x_1, \cdots, x_n$. 由微分学的罗尔中值定理知 $\varphi'(t)$ 存在 $n+1$ 个零点. 同样, 对 $\varphi'(t)$ 使用罗尔定理, 知 $\varphi''(t)$ 存在 n 个零点. 依此递推最后得 $\varphi^{(n+1)}(t)$ 存在 1 个零点, 记为 ξ (介于 x

Back
Close

与 $x_0, x_1, \cdots, x_n$ 之间). 直接计算得

$$\begin{aligned}\varphi^{(n+1)}(t) &= R_n^{(n+1)}(t) - \frac{R_n(x)}{\omega(x)}\omega^{(n+1)}(t) \\ &= f^{(n+1)}(t) - \frac{R_n(x)}{\omega(x)}(n+1)!,\end{aligned}$$

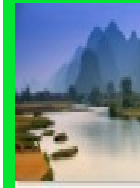
从而由 $\varphi^{(n+1)}(\xi) = 0$ 立刻得到 (4.4) 式.

由误差公式 (4.4) 可知, 如果 $f^{(n+1)}(x)$ 在 $[a, b]$ 上有界, 即存在常数 M, 使得 $|f^{(n+1)}(x)| \le M$, 则必有

$$|R_n(x)| \le \frac{M}{(n+1)!}|x - x_0| \cdot |x - x_1| \cdots |x - x_n|. \qquad (4.6)$$

例 4.1 已知 $\omega(x) = \prod_{i=0}^{n}(x - x_i)$, 求证

$$\omega'(x_k) = \prod_{i=0, i\ne k}^{n}(x_k - x_i), \quad k = 1, 2, \cdots, n.$$



Back
Close

证 因为

$$\omega(x)=\prod_{i=0}^{n}(x-x_i)=(x-x_k)\prod_{i=0,i\neq k}^{n}(x-x_i),\quad k=1,2,\cdots,n.$$

求导数得

$$\omega'(x)=\prod_{i=0,i\neq k}^{n}(x-x_i)+(x-x_k)\frac{\mathrm{d}}{\mathrm{d}x}\Big[\prod_{i=0,i\neq k}^{n}(x-x_i)\Big].$$

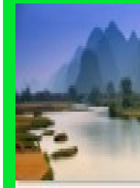
由此即得

$$\omega'(x_k)=\prod_{i=0,i\neq k}^{n}(x_k-x_i),\quad k=1,2,\cdots,n.$$

证毕. $\square$

§4.1.3 拉格朗日插值及其通用程序

前面已经讨论过, 满足插值条件 (4.2) 的插值多项式 (4.1) 是唯一



Back
Close

存在的, 它的系数可以通过求解线性方程组 (4.3) 得到. 但由于求解线性方程组的计算量较大, 且当 n 较大时, 方程组 (4.3) 是一个病态方程组, 求解不可靠. 我们可以通过 "基函数法" 得到拉格朗日插值多项式, 从而不必解线性方程组, 避免了范德蒙矩阵的病态现象.

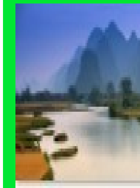
1. 线性插值

设已知 x_0, x_1 及 $y_0 = f(x_0)$, $y_1 = f(x_1)$, $L_1(x)$ 为不超过 1 次的多项式且满足 $L_1(x_0) = y_0$, $L_1(x_1) = y_1$. 几何上, $L_1(x)$ 为经过 (x_0, y_0), (x_1, y_1) 两点的直线, 从而得到

$$L_1(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0). \tag{4.7}$$

为了推广到高阶插值问题, 我们将 (4.7) 式变形为对称形式

$$L_1(x) = l_0(x)y_0 + l_1(x)y_1, \tag{4.8}$$



其中

$$l_0(x) = \frac{x - x_1}{x_0 - x_1}, \quad l_1(x) = \frac{x - x_0}{x_1 - x_0}$$

均为 1 次多项式, 且满足

$$l_0(x_0) = 1, \quad l_0(x_1) = 0,$$

$$l_1(x_0) = 0, \quad l_1(x_1) = 1.$$

上面的关系式可以统一写成

$$l_i(x_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \qquad i,\ j = 0, 1. \tag{4.9}$$

2. 抛物插值

设已知 $x_0,\ x_1,\ x_2$ 及 $y_i = f(x_i),\ i = 0, 1, 2,\ L_2(x)$ 为不超过 2 次的多项式, 且满足插值条件 $L_2(x_i) = y_i\ (i = 0, 1, 2)$. 由 $L_1(x)$ 的表达式

Back
Close

(4.8) 猜测到, $L_2(x)$ 应有下述表达式

$$L_2(x) = l_0(x)y_0 + l_1(x)y_1 + l_2(x)y_2, \tag{4.10}$$

其中, $l_0(x)$, $l_1(x)$, $l_2(x)$ 均为 2 次多项式且满足 (4.9) $(i, j = 0, 1, 2)$. 可以验证, 由 (4.10) 给出的 $L_2(x)$ 满足插值条件. 事实上,

$$L_2(x_0) = l_0(x_0)y_0 + l_1(x_0)y_1 + l_2(x_0)y_2 = 1 \times y_0 + 0 \times y_1 + 0 \times y_2 = y_0.$$

同理, $L_2(x_1) = y_1$, $L_2(x_2) = y_2$. 因此, 猜测式(4.10) 是正确的, 于是问题转化为求 $l_i(x)$ $(i = 0, 1, 2)$.

因 $l_0(x)$ 是 2 次多项式且 $l_0(x_1) = l_0(x_2) = 0$, 故有

$$l_0(x) = c(x - x_1)(x - x_2).$$

再由 $l_0(x_0) = c(x_0 - x_1)(x_0 - x_2) = 1$, 得

$$c = \frac{1}{(x_0 - x_1)(x_0 - x_2)},$$

Back
Close

从而

$$l_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}. \tag{4.11}$$

同理,

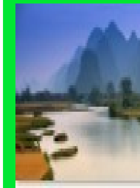
$$l_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}, \quad l_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}. \tag{4.12}$$

3. n 阶拉格朗日插值

设已知 $x_0, x_1, \cdots, x_n$ 及 $y_i = f(x_i)\,(i = 0, 1, \cdots, n)$, $L_n(x)$ 为不超过 n 次的多项式, 且满足插值条件 $L_n(x_i) = y_i\,(i = 0, 1, \cdots, n)$. 由对 $L_2(x)$ 的构造经验, 可设

$$L_n(x) = \sum_{i=0}^{n} l_i(x) y_i = l_0(x) y_0 + l_1(x) y_1 + \cdots + l_n(x) y_n,$$

其中, $l_i(x)$ $(i = 0,\ 1,\ \cdots,\ n)$ 均为 n 次多项式且满足 (4.9) $(i,\ j = 0,\ 1,\ \cdots,\ n)$. 不难验证, 这样构造出的 $L_n(x)$ 满足插值条件. 因此问题归结



为求 $l_i(x)\,(i=0,1,\cdots,n)$ 的表达式. 因 $x_j(j\neq i)$ 是 n 次多项式 $l_i(x)$ 的 n 个根, 故可设

$$l_i(x)=c(x-x_0)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)=c\prod_{j=0,j\neq i}^{n}(x-x_j).$$

再由

$$l_i(x_i)=c\prod_{j=0,j\neq i}^{n}(x_i-x_j)=1,\quad i=0,1,\cdots,n,$$

得

$$c=\frac{1}{\prod\limits_{j=0,j\neq i}^{n}(x_i-x_j)}.$$

故有

$$L_n(x)=\sum_{i=0}^{n}l_i(x)y_i,\quad l_i(x)=\prod_{j=0,j\neq i}^{n}\frac{x-x_j}{x_i-x_j}.\tag{4.13}$$

公式 (4.13) 称为 n 阶拉格朗日插值公式, 其中 $l_i(x)\,(i=0,1,\cdots,n)$ 称为 n 阶拉格朗日插值的基函数.

例 4.2 已知函数表 $\sin\frac{\pi}{6}=0.5000$, $\sin\frac{\pi}{4}=0.7071$, $\sin\frac{\pi}{3}=0.8660$, 分别由线性插值和抛物插值求 $\sin\frac{2\pi}{9}$ 的近似值, 并估计其精度.

解 (1) 线性插值只需要两个节点, 根据余项公式选取前两个节点.

$$\begin{aligned}\sin\frac{2\pi}{9} &\approx L_1\Big(\frac{2\pi}{9}\Big)=\frac{\frac{2\pi}{9}-\frac{\pi}{4}}{\frac{\pi}{6}-\frac{\pi}{4}}\times 0.5000+\frac{\frac{2\pi}{9}-\frac{\pi}{6}}{\frac{\pi}{4}-\frac{\pi}{6}}\times 0.7071\\ &=\frac{1}{3}\times 0.5000+\frac{2}{3}\times 0.7071=0.6381.\end{aligned}$$

截断误差为

$$\Big|R_1\Big(\frac{2\pi}{9}\Big)\Big|=\Big|\frac{(\sin x)''}{2!}\Big(\frac{2\pi}{9}-\frac{\pi}{6}\Big)\Big(\frac{2\pi}{9}-\frac{\pi}{4}\Big)\Big|\leq\frac{1}{2}\times\frac{\pi}{18}\times\frac{\pi}{36}=7.615\times 10^{-3},$$

Back
Close

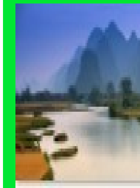
得 $\varepsilon = 7.615 \times 10^{-3} < 0.5 \times 10^{-1}$, 因此计算结果至少有 1 位有效数字.

(2) 由 (4.10), (4.11) 和 (4.12) 得

$$
\begin{aligned}
\sin\frac{2\pi}{9} \approx L_2\left(\frac{2\pi}{9}\right) &= \frac{(\frac{2\pi}{9}-\frac{\pi}{4})(\frac{2\pi}{9}-\frac{\pi}{3})}{(\frac{\pi}{6}-\frac{\pi}{4})(\frac{\pi}{6}-\frac{\pi}{3})} \times 0.5000 \\
&\quad + \frac{(\frac{2\pi}{9}-\frac{\pi}{6})(\frac{2\pi}{9}-\frac{\pi}{3})}{(\frac{\pi}{4}-\frac{\pi}{6})(\frac{\pi}{4}-\frac{\pi}{3})} \times 0.7071 \\
&\quad + \frac{(\frac{2\pi}{9}-\frac{\pi}{6})(\frac{2\pi}{9}-\frac{\pi}{4})}{(\frac{\pi}{3}-\frac{\pi}{6})(\frac{\pi}{3}-\frac{\pi}{4})} \times 0.8660 \\
&= \frac{2}{9} \times 0.5 + \frac{8}{9} \times 0.7071 - \frac{1}{9} \times 0.866 = 0.6434.
\end{aligned}
$$

截断误差为

$$
\begin{aligned}
\left|R_2\left(\frac{2\pi}{9}\right)\right| &= \left|\frac{(\sin x)'''}{3!}\left(\frac{2\pi}{9}-\frac{\pi}{6}\right)\left(\frac{2\pi}{9}-\frac{\pi}{4}\right)\left(\frac{2\pi}{9}-\frac{\pi}{3}\right)\right| \\
&\le \frac{1}{6} \times \frac{\pi}{18} \times \frac{\pi}{36} \times \frac{\pi}{36} = 8.861 \times 10^{-4},
\end{aligned}
$$



Back
Close

得 $\varepsilon = 8.861 \times 10^{-4} < 0.5 \times 10^{-2}$, 因此计算结果至少有 2 位有效数字.

4. 拉格朗日插值公式的通用程序

根据拉格朗日插值公式 (4.13), 编写下列程序, 可对于给定的数据求得插值点的插值结果. 注意, 该程序并不能输出插值多项式的表达式.

- 拉格朗日插值 MATLAB 程序

```
%malagr.m
function yy=malagr(x,y,xx)
%用途: 拉格朗日插值法求解
%格式: yy=malagr(x,y,xx)  x是节点向量,y是节点对应的函
%        数值向量,xx是插值点(可以是多个),yy返回插值结果
m=length(x); n=length(y);
```

```
if m~=n, error('向量x与y的长度必须一致'); end
s=0;
for i=1:n
    t=ones(1,length(xx));
    for j=1:n
        if j~=i
            t=t.*(xx-x(j))/(x(i)-x(j));
        end
    end
    s=s+t*y(i);
end
yy=s;
```

例 4.3 用上面的程序 malagr.m 求解例 4.2.

解 在 MATLAB 命令窗口执行下列命令：

```
>> x=pi*[1/6 1/4];
>> y=[0.5 0.7071]; xx=2*pi/9;
>> yy1=malagr(x,y,xx)
yy1 =
     0.6381
>> x=pi*[1/6 1/4 1/3];
>> y=[0.5 0.7071 0.8660]; xx=2*pi/9;
>> yy2=malagr(x,y,xx)
yy2 =
     0.6434
```

作业：P90: 4.1; 4.7.