插值与拟合

温丹苹
邮箱：dpwen@nju.edu.cn
办公室：工管院协鑫楼306

为什么插值

* 事物的运动变化规律常常用函数$y = f(x)$来描述。

* 但在实际问题中，大多数情形下$y = f(x)$的**准确表达式是未知的**，只能通过试验观测的方法得到函数的一些值。然后用一个**较简单**的便于计算的函数来近似表示未知函数。

* 另一种情形，有时$\boldsymbol{y = f(x)}$**的准确表达式是已知的**，但较复杂，需用一个简单函数来近似地代替$f(x)$。

▸ 用简单函数近似表示复杂函数或未知函数一组值的问题就是**插值与拟合问题**。

为什么插值

数据较少
数据不全

例：已测得在某处海洋不同深度处的水温如下：

深度（M）	466	741	950	1422	1634
水温（^{o}C）	7.04	4.28	3.40	2.54	2.13

根据这些数据，希望合理地估计出其它深度（如 500、600、800米…）处的水温。

数学工具：插值

什么是插值

已知函数 $y = f(x)$ 在 $[a, b]$ 上有定义，且已经测得在点

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

处的函数值为 $y_0 = f(x_0)$，…， $y_n = f(x_n)$

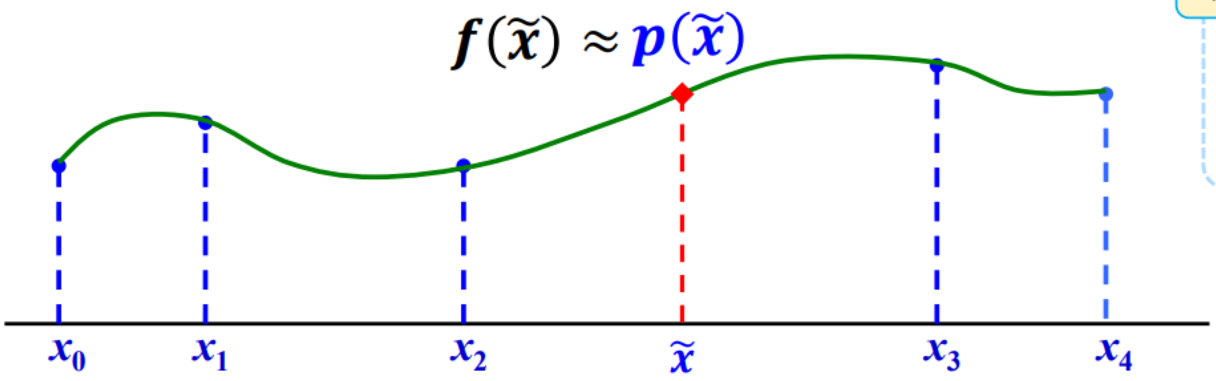
如果存在一个 简单易算 的函数 $p(x)$，使得

$$p(x_i) = f(x_i),\quad i = 0, 1, \ldots, n$$

则称 $p(x)$ 为 $f(x)$ 的 插值函数。 求插值函数 $p(x)$ 的方法就称为 插值法。

- [a, b] 为插值区间，x_i为插值节点，$p(x_i) = f(x_i)$为插值条件.
- 插值节点无需递增排列，但必须确保互不相同！

插值图示



记号说明

x_i — 插值节点

$\tilde{x}$ — 插值点

* 求出**插值函数$p(x)$**后，对于任意给定的一点$\tilde{x}$ ，我们就可以用$p(\tilde{x})$来近似$f(\tilde{x})$的值，这就是插值的目的。
* 插值是一种近似方法。
* 在实际应用中，我们感兴趣的往往是**某些点的值**，而不一定是插值函数

常用的插值方法

* **多项式插值:** $p(x)$ 为多项式，多项式最常用的插值函数。
* **分段多项式插值:** $p(x)$为分段多项式
* **三角插值:** $p(x)$为三角函数
* **有理插值:** $p(x)$为有理函数
* ……

- 不同插值方法的区别在于插值函数 p(x) 的选取。
- 我们主要介绍前两种插值方法，其他插值方法的原理是类似的。

目录

7.1.1 多项式插值概念

已知函数 $y = f(x)$ 在 $[a, b]$ 上 $n + 1$ 个点

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

处的函数值为 $y_0 = f(x_0)$，…，$y_n = f(x_n)$

求次数 不超过 n 的多项式

$$p(x) = c_0 + c_1 x + \cdots + c_n x^n$$

n次插值多项式

使得

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

- **注意:** $p(x)$的次数有可能小于n.
- n次插值多项式**几何意义：** 过平面上已知的n+1个互异点$(x_i, y_i)(i = 0,1,2,\ldots,n)$, 作一条次数不超过$n$次的多项式曲线。

7.1.2 多项式插值存在唯一性

- 对于n次插值多项式的问题，必须解决如下问题：

  (1)满足插值条件$p_n(x_i) = y_i$的多项式 是否存在？如果存在，$p_n(x)$是否唯一？

  (2) $p_n(x)$近似表示$f(x)$,误差如何？

  (3)怎样求出插值多项式$p_n(x)$？

7.1.2 多项式插值存在唯一性

定理： 多项式 $p_n(x)$存在且唯一。

证明： 设给定了 $y=f(x)$在$[a,b]$上的 $n+1$ 个互异点的值 $y_i=f(x_i)$ $(i=0,1,\cdots n)$，
$P_n(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$。将插值条件 $P_n(x_i)=y_i$ 代入多项式$P_n(x)$，得到

$$\begin{cases} a_0+a_1x_0+a_2x_0^2+\cdots+a_nx_0^n=y_0, \\ a_0+a_1x_1+a_2x_1^2+\cdots+a_nx_1^n=y_1, \\ \qquad\qquad \cdots \\ a_0+a_1x_n+a_2x_n^2+\cdots+a_nx_n^n=y_n. \end{cases}$$

7.1.2 多项式插值存在唯一性

定理： 多项式 $p_n(x)$存在且唯一。

证明（续）： 此方程组含 $n+1$个方程，$n+1$个未知数 $a_0, a_1, \cdots, a_n$。其系数行列式是范德蒙行列式

$$V = \begin{vmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix} = \prod_{0 \le j < i \le n} \left(x_i - x_j\right)$$

由于 $x_0, x_1, \cdots, x_n$ 互异，所以 $V \neq 0$，根据线性方程组的克莱姆法则，方程组存在唯一的解 $a_0, a_1, \cdots, a_n$。因此插值多项式 $P_n(x)$ 存在且唯一。

7.1.2 多项式插值存在唯一性

定理： 多项式 $p_n(x)$存在且唯一。

* 该定理的证明过程也给出了一种求$p_n(x)$的方法，**但需要求解一个线性方程组**，后面将给出更简单的计算方法。

7.1.3 线性插值（一次多项式）

例（线性插值）：求一次多项式 $p(x)$，满足

$$p(x_0)=y_0\,,\quad p(x_1)=y_1$$

$p(x_0)=y_0$
$p(x_1)=y_1$

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

点斜式

重新整理

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

我们注意到，$p(x)$ 是两个一次多项式的线性组合

令 $l_0(x)=\dfrac{x-x_1}{x_0-x_1}$，$l_1(x)=\dfrac{x-x_0}{x_1-x_0}$

则 $p(x)=y_0l_0(x)+y_1l_1(x)$

能观察到什么？

$l_0(x), l_1(x)$ 满足： $l_0(x_0)=1,\ l_0(x_1)=0,\ l_1(x_0)=0,\ l_1(x_1)=1$

7.1.3 抛物线插值（二次多项式）

例（抛物线插值）：求二次多项式 $p(x)$，满足

$$p(x_0)=y_0\ ,\quad p(x_1)=y_1\ ,\quad p(x_2)=y_2$$

思路：构造三个二次多项式 $l_0(x), l_1(x), l_2(x)$，满足

$$l_0(x_0)=1,\ \ l_0(x_1)=0,\ \ l_0(x_2)=0$$
$$l_1(x_0)=0,\ \ l_1(x_1)=1,\ \ l_1(x_2)=0$$
$$l_2(x_0)=0,\ \ l_2(x_1)=0,\ \ l_2(x_2)=1$$

$$\Rightarrow\ \boxed{p(x)=y_0 l_0(x)+y_1 l_1(x)+y_2 l_2(x)}$$

- 将问题转化为：如何构造 $l_0(x)$，$l_1(x)$，$l_2(x)$？

7.2.1 基函数插值法

记

$\mathrm{H}_n(x)$ = {次数不超过 n 的多项式的全体} → $n+1$ 维线性空间

设 $z_0(x), z_1(x), \dots, z_n(x)$ 是 $\mathrm{H}_n(x)$ 的一组基，则插值多项式可表示为

$$p(x) = a_0 z_0(x) + a_1 z_1(x) + \cdots + a_n z_n(x)$$

需要解决的两个问题

① 寻找合适的基函数

② 确定插值多项式在这组基下的线性表出系数

- 这种通过基函数来构造插值函数的方法就是基函数插值法。
- 存在唯一性定理中以$1, x, x^2, \dots, x^n$为插值基函数。

7.2.2 Lagrange插值

* Lagrange 插值（法）

用**Lagrange插值基函数**来计算插值多项式的方法。

- 已知$y=f(x)$的如下函数值表：

x_k	x_0	x_1	…	x_{i-1}	x_i	x_{i+1}	…	x_n
y_k	0	0	…	0	1	0	…	0

求满足插值条件$l_k(x_i)=y_i(i=0,1,2,\dots,n)$的n次插值多项式$l_k(x)$.

- 根据函数值表知，所要求的$l_k(x)$需满足以下两点：

（1）$l_k(x)$是n次多项式；

（2）$l_k(x_i)=\begin{cases}1, i=k\\0, i\neq k\end{cases}, \ (i=0,1,2,\dots,n)$

7.2.2 Lagrange插值

∗ Lagrange 插值（法）

用Lagrange插值基函数来计算插值多项式的方法。

定义：设 $l_k(x)$ 是 n 次多项式，且在节点 $x_0, x_1, \ldots, x_n$ 上满足

$$l_k(x_i) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \qquad i = 0, 1, 2, \ldots, n$$

则称 $l_k(x)$ 为节点 $x_0, x_1, \ldots, x_n$ 上的 n 次 Lagrange 插值基函数

（k=0, 1, …, n）

∗ $l_0(x),\ l_1(x),\ \cdots, l_n(x)$构成$H_n(x)$的一组基;

∗ $l_0(x),\ l_1(x),\ \cdots, l_n(x)$由插值节点唯一确定。

7.2.2 Lagrange插值

根据定义，可设

$$l_k(x) = \alpha_k (x - x_0)\cdots(x - x_{k-1})(x - x_{k+1})\cdots(x - x_n)$$

其中 α_k 是待定系数。将 $l_k(x_k) = 1$ 代入可求得

$$\alpha_k = \frac{1}{(x_k - x_0)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)}$$

$$l_k(x) = \frac{(x - x_0)\cdots(x - x_{k-1})(x - x_{k+1})\cdots(x - x_n)}{(x_k - x_0)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)} = \prod_{i=0, i\neq k}^{n} \frac{x - x_i}{x_k - x_i}$$

（k=0, 1, …, n）

7.2.2 Lagrange插值

由于 $l_0(x), l_1(x), \ldots, l_n(x)$ 构成 $H_n(x)$ 的一组基，因此可设插值多项式：

$$p(x) = a_0 l_0(x) + a_1 l_1(x) + \cdots + a_n l_n(x)$$

将插值条件 $p(x_i) = y_i$， $i = 0, 1, \ldots, n$ 代入，可得 $a_i = y_i$， $i = 0, 1, \ldots, n$

$$\boxed{p(x) = y_0 l_0(x) + y_1 l_1(x) + \cdots + y_n l_n(x)} \triangleq L_n(x)$$

$L_n(x)$ 就称为 $f(x)$ 的 Lagrange 插值多项式。

$$L_n(x) = \sum_{k=0}^{n} y_k l_k(x) = \sum_{k=0}^{n} y_k \prod_{i=0, i\neq k}^{n} \frac{x - x_i}{x_k - x_i}$$

* **Lagrange 插值的优点**：简单，可以直接把插值多项式写出来，易于计算机实现。

7.2 Lagrange插值

＊ 线性插值（一次多项式）

线性插值多项式（一次插值多项式）： $n=1$

$$L_1(x)=y_0l_0(x)+y_1l_1(x)=y_0\frac{x-x_1}{x_0-x_1}+y_1\frac{x-x_0}{x_1-x_0}$$

过二点的直线

＊ 抛物线插值（二次多项式）

抛物线插值多项式（二次插值多项式）： $n=2$

$$L_2(x)=y_0l_0(x)+y_1l_1(x)+y_2l_2(x)$$
$$=y_0\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}+y_1\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}+y_2\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

过三点的抛物线

- n 次插值多项式的次数有时会低于 n，比如二次插值中，如果三点共线， 则$L_2(x)$为直线.

7.2.3 Lagrange插值举例

例：已知函数 $y = \ln(x)$ 的函数值如下

x	0.4	0.5	0.6	0.7	0.8
$\ln(x)$	−0.9163	−0.6931	−0.5108	−0.3567	−0.2231

试分别用 线性插值 和 抛物线插值 计算 $\ln(0.54)$ 的近似值。

插值节点的选取：为提高计算精度，通常选取所需插值的点 x 邻近的节点

解：(1) 线性插值，取 $x_0=0.5$, $x_1=0.6$ 得

$$L_1(x) = y_0 \frac{x-x_1}{x_0-x_1} + y_1 \frac{x-x_0}{x_1-x_0} = 0.1823x - 1.6046$$

将 $x=0.54$ 代入可得：

$$\ln(0.54) \approx L_1(0.54) = -0.6202$$

(2) 抛物线插值，取 $x_0=0.4$, $x_1=0.5$, $x_2=0.6$, 可得

$$\ln(0.54) \approx L_2(0.54) = -0.6153$$

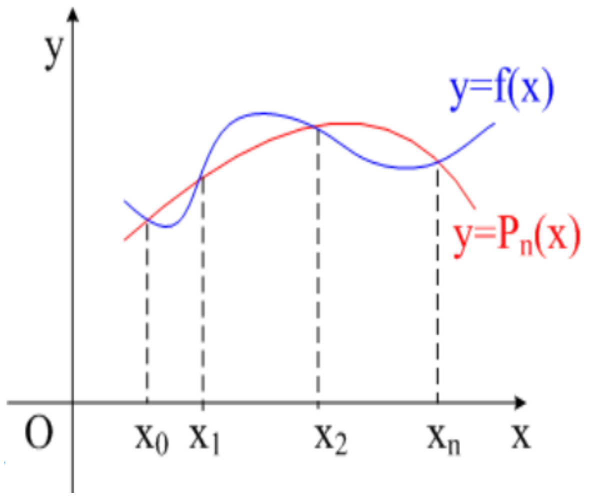
- $\ln(0.54)$的精确值为：−0.616186…
- 抛物线插值的误差比线性插值要小一些。

Demo_7_1_ Interp_Lagrange.m

思考：

是不是插值多项式次数越高，误差越小？

7.2.4 误差估计



- 在[a,b]上用$P_n(x)$近似表示$f(x)$，在插值节x_i点处是没有误差的，即$f(x_i) = P_n(x_i)$，$i = 0,1,2,\dots,n$
- 在其他x处，一般$f(x)$与$P_n(x)$不相等。
- 定义$\boldsymbol{R_n(x)}$为**插值多项式的余项或截断误差**。

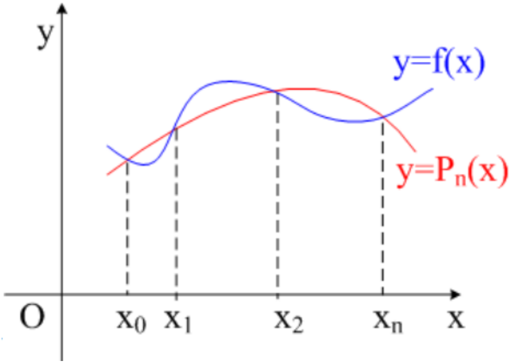
* **插值余项：**

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

7.2.4 误差估计

＊ 插值余项：

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



定理： 设 $f(x) \in C^n[a, b]$（n 阶连续可微），且 $f^{(n+1)}(x)$ 在 (a, b) 内存在，则对 $\forall x \in [a, b]$，存在 $\xi_x \in (a, b)$ 使得

$$R_n(x) = f(x) - L_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}\omega_{n+1}(x)$$

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

注：余项中的ξ_x是与x有关的。
当$x = x_i$时，结论显然成立，因此只要证明$x \neq x_i$的情形。

7.2.4 误差估计

＊ 简要证明过程：

由插值条件可知： $R_n(x_i)=0,\ i=0, 1, \ldots, n$

因此，$R_n(x)$ 可写成 $\boxed{R_n(x) = K(x)\omega_{n+1}(x)}$

对任意给定的 $x\in[a, b]\ (x \neq x_i , i=0, 1, \ldots, n)$，构造 辅助函数

$$\boxed{\varphi(t) \triangleq R_n(t) - K(x)\omega_{n+1}(t) = f(t) - L_n(t) - K(x)\omega_{n+1}(t)}$$

则 $\varphi(t)$ 在 $[a, b]$ 中有 $n+2$ 个互不相同的零点：$x, x_0, \ldots, x_n$

条件：$f(x) \in C^n[a, b]$，且 $f^{(n+1)}(x)$ 在 (a, b) 内存在 ⟹ $\varphi(t)$ 也具有此性质

由罗尔定理可知 $\varphi'(t)$ 在 (a, b) 内至少有 $n+1$ 个不同的零点。

罗尔定理 设 $f(x) \in C[a,b]$，且在 (a,b) 内可微；若 $f(a) = f(b) = 0$，则必存在一点 $\xi \in (a,b)$，使得 $f'(\xi) = 0$。

再用一次罗尔定理，可知 $\varphi''(t)$ 在 (a, b) 内至少有 n 个不同的零点。

以此类推，可知 $\varphi^{(n+1)}(t)$ 在 (a, b) 内至少有一个零点，设为 $\xi_x \in (a, b)$，有

$$\boxed{\varphi^{(n+1)}(\xi_x) = 0}$$

又

$$\begin{aligned}\varphi^{(n+1)}(t) &= R_n^{(n+1)}(t) - K(x)\omega_{n+1}^{(n+1)}(t) \\ &= f^{(n+1)}(t) - L_n^{(n+1)}(t) - K(x)(n+1)! \\ &= f^{(n+1)}(t) - K(x)(n+1)!\end{aligned}$$

$\Rightarrow$ $K(x) = \dfrac{f^{(n+1)}(\xi_x)}{(n+1)!}$ $\Rightarrow$ $\boxed{R_n(x) = \dfrac{f^{(n+1)}(\xi_x)}{(n+1)!}\omega_{n+1}(x)}$

7.2.4 误差估计

＊ 线性插值和抛物线插值的余项：

线性插值的余项（两点插值， $n=1$）

$$R_1(x)=\frac{1}{2}f''(\xi_x)(x-x_0)(x-x_1)$$

抛物线插值的余项（三点插值， $n=2$）

$$R_2(x)=\frac{1}{6}f'''(\xi_x)(x-x_0)(x-x_1)(x-x_2)$$

7.2.4 误差估计

- 余项公式只有当 $f(x)$ 的高阶导数存在时才能使用
- ξ_x 与 x 有关，通常无法确定, 实际使用中通常是估计其上界

若 $\max\limits_{a\le x\le b}\left|f^{(n+1)}(x)\right| = M_{n+1}$，则 $\left|R_n(x)\right| \le \dfrac{M_{n+1}}{(n+1)!}\prod\limits_{i=0}^{n}\left|x-x_i\right|$

- 计算点 x 上的近似值时，应尽量选取与 x 靠近插值节点

例：已知函数 $y=\ln(x)$ 的函数值如下

x	0.4	0.5	0.6	0.7	0.8
$\ln(x)$	−0.9163	−0.6931	−0.5108	−0.3567	−0.2231

试估计用 线性插值 和 抛物线插值 计算 $\ln(0.54)$ 时的误差。

解：(1) 线性插值余项

$$R_1(x)=\frac{f^{(2)}(\xi_x)}{2}(x-x_0)(x-x_1),\qquad x_0=0.5,\ x_1=0.6,\ \xi_x\in(0.5,0.6)$$

$$\left|f^{(2)}(\xi_x)\right|=\left|-\xi_x^{-2}\right|\le 4$$

$$\Rightarrow\quad \left|R_1(0.54)\right|\le\left|2(0.54-0.5)(0.54-0.6)\right|=0.048$$

7.2.4 误差估计

(2) 抛物线插值余项

$$R_2(x)=\frac{f^{(3)}(\xi_x)}{3!}(x-x_0)(x-x_1)(x-x_2)$$

$x_0=0.4,\ x_1=0.5,\ x_2=0.6,\ \xi_x\in(0.4,\ 0.6)$ $\Rightarrow$ $\left|f^{(3)}(\xi_x)\right|\le\left|2\xi_x^{-3}\right|=31.25$

$\Rightarrow$ $\left|R_2(0.54)\right|\le\dfrac{31.25}{3!}\left|(0.54-0.4)(0.54-0.5)(0.54-0.6)\right|=0.00175$

＊ 对比： $\left|R_1(0.54)\right|\le\left|2(0.54-0.5)(0.54-0.6)\right|=0.048$

＊ 抛物线插值通常优于线性插值，但绝不是次数越高就越好！

7.2.4 误差估计

例：函数 $f(x)=\dfrac{1}{1+x^2}$ ，插值区间 [-5, 5]，取等距插值节点，试画出插值多项式 $L_n(x)$ 的图像。

Demo_7_2_ Interp_Lagrange.m

7.2.4 误差估计

例：设 $f(x)\in C^2[a,b]$ （二阶连续可导），证明:

$$\max_{a\le x\le b}\left|f(x)-\left[f(a)+\frac{f(b)-f(a)}{b-a}(x-a)\right]\right|\le\frac{1}{8}M_2(b-a)^2$$

其中

$$M_2=\max_{a\le x\le b}|f''(x)|$$

证明： 易知 $L_1(x)\triangleq f(a)+\dfrac{f(b)-f(a)}{b-a}(x-a)$

是 $f(x)$ 关于点 $x_0=a, x_1=b$ 的线性插值多项式，由插值余项公式可知

$$|f(x)-L_1(x)|=\left|\frac{f^{(2)}(\xi_x)}{2}(x-a)(x-b)\right|\le\frac{1}{2}M_2|(x-a)(b-x)|\le\frac{1}{8}M_2(b-a)^2$$

$x\in[a,b]$

7.2.5 Lagrange插值基函数的重要性质

性质一： 若 $f(x)$ 是一个次数 $\le n$ 的多项式，则有 $f^{(n+1)}(x) \equiv 0$，故

$$R_n(x) = f(x) - L_n(x) \equiv 0$$

即 n 次插值多项式对于次数 $\le n$ 的多项式是精确成立的。

性质二： 设 $f(x) = x^k$，$k \le n$，则有

$$R_n(x) = x^k - \sum_{j=0}^{n} x_j^k l_j(x) = 0 \iff \boxed{x^k = \sum_{j=0}^{n} x_j^k l_j(x)}$$

特别地，当 $k = 0$ 时有 $\boxed{\sum_{j=0}^{n} l_j(x) = 1}$

7.2.6 Lagrange插值优缺点

* **优点：** 拉格朗日插值法利用插值基函数直接表示出了插值多项式，格式整齐规范，结构紧凑，便于理解记忆和理论分析。
* **缺点：** 当节点增加时，希望构造更高次的插值函数时，所有的基函数都要重新计算，不太方便。

- Q & A
- 谢谢