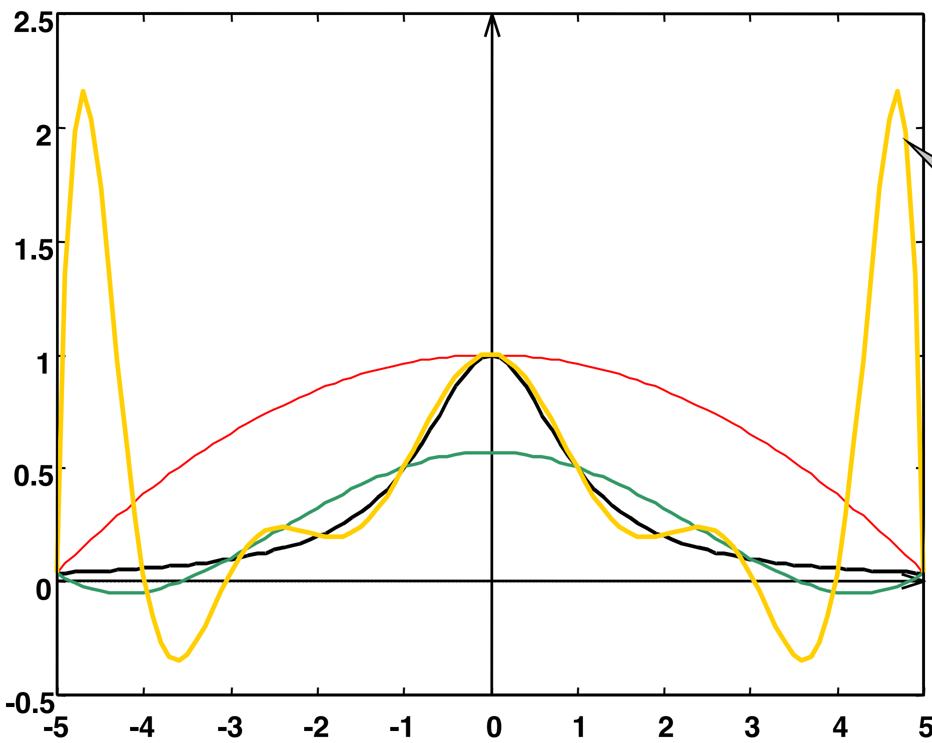
第一章　插值

1.7分段插值

1.8样条函数

分段低次插值

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



$P_n(x) \not\rightarrow f(x)$

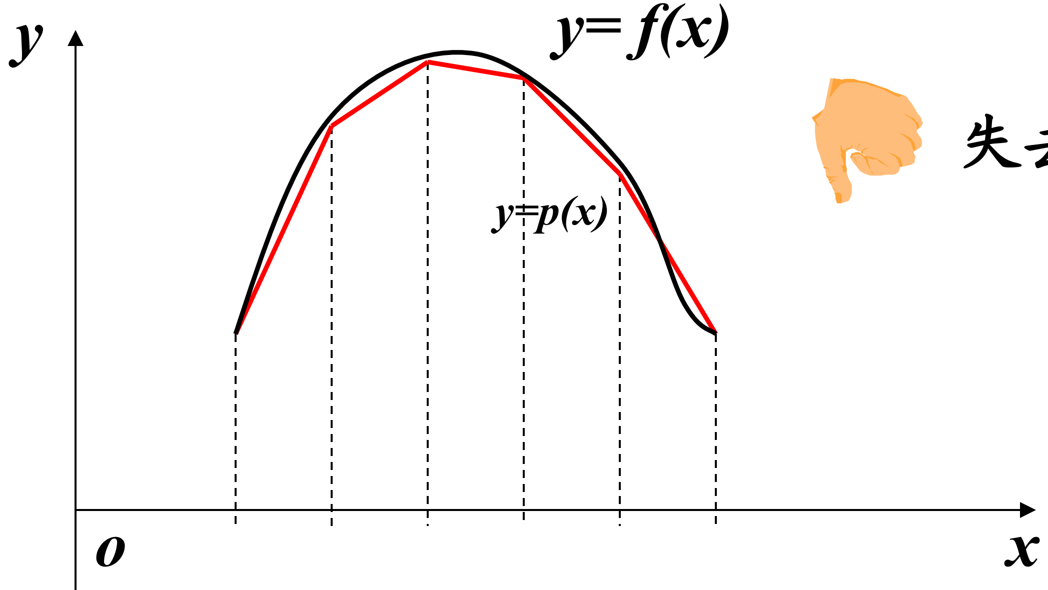
n 越大，
端点附近抖动
越大，称为
Runge 现象

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$f(x)$:

$$f(x) \approx P_1(x) = \frac{x - x_{i+1}}{x_i - x_{i+1}} y_i + \frac{x - x_i}{x_{i+1} - x_i} y_{i+1} \qquad x \in [x_i, x_{i+1}]$$

记 $h = \max |x_{i+1} - x_i|$，易证：当 $h \to 0$ 时，$P_1^h(x) \xrightarrow{一致} f(x)$



失去了原函数的光滑性。

分段线性插值的余项

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

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

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

分段Hermite插值

分段线性插值函数 $P_1^h(x)$ 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 y_k 外还给出导数值 $y'_k=m_k(k=0,1,\cdots,n)$，这样就可构造一个导数连续的分段插值函数 $P_3^h(x)$，它满足条件:

1. $P_3^h(x)\in C^1[a,b]$（$C^1[a,b]$代表区间[a, b]上一阶导数连续的函数集合）。
2. $P_3^h(x)=y_k, P_3'^h(x)=y'_k, (k=0,1,\cdots,n)$。
3. $P_3^h(x)$在每个小区间$[x_k,x_{k+1}]$上是三次多项式。

分段Hermite插值

由两点三次hermite插值多项式。可知，$P_3^h(x)$在区间$[x_k, x_{k+1}]$上表达式为

$$
\begin{aligned}
P_3^h(x) = & \left(\frac{x - x_{k+1}}{x_k - x_{k+1}}\right)^2 \left(1 + 2\frac{x - x_k}{x_{k+1} - x_k}\right) y_k \\
& + \left(\frac{x - x_k}{x_{k+1} - x_k}\right)^2 \left(1 + 2\frac{x - x_{k+1}}{x_k - x_{k+1}}\right) y_{k+1} \\
& + \left(\frac{x - x_{k+1}}{x_k - x_{k+1}}\right)^2 (x - x_k) y'_k \\
& + \left(\frac{x - x_k}{x_{k+1} - x_k}\right)^2 (x - x_{k+1}) y'_{k+1}
\end{aligned}
$$

分段Hermite插值的余项

$$\left|f(x)-s_3(x)\right| \le \max_{x_i \le x \le x_{i+1}} \left|f(x)-s_3(x)\right|$$

$$\le \left| \max_{x_i \le x \le x_{i+1}} \frac{f^4\left(\xi_i\right)}{4\,!}(x-x_i)^2(x-x_{i+1})^2 \right|$$

$$\le \frac{\max\limits_{x_i \le x \le x_{i+1}} \left|f^4\left(x\right)\right|}{4\,!} \left(\frac{x_j - x_i}{2}\right)^4 \le \frac{h_i^4}{384} \max_{x_i \le x \le x_{i+1}} \left|f^4\left(x\right)\right|$$

分段Hermite插值

给定 $x_0, \dots, x_n$； $y_0, \dots, y_n$； $y_0', \dots, y_n'$

导数一般不易得到。

余项

$$\left|f(x)-s_3(x)\right| \le \frac{\max\limits_{x_i \le x \le x_{i+1}} \left|f^4(x)\right|}{4\,!}\left(\frac{x_j - x_i}{2}\right)^4 \le \frac{h_i^4}{384}\max_{x_i \le x \le x_{i+1}}\left|f^{(4)}(x)\right|$$

分段低次插值

基本思想：用分段低次多项式来代替单个多项式。

具体作法：(1) 把整个插值区间分割成多个小区间;

(2) 在每个小区间上作低次插值多项式;

(3) 将所有插值多项式拼接整一个多项式。

优点：公式简单、运算量小、稳定性好、收敛性 ...

缺点：节点处的导数不连续，失去原函数的光滑性。

样条函数插值问题的提出

如果采用分段多项式插值,则由于插值基函数只是局部活跃(它们的支集是局部紧致的),结点上的误差可以被控制在小的范围内,因而也带来了内在的高度稳定性.这是分段插值的一大优势!

许多实际问题希望插值函数具有较高阶的整体光滑性.此时,高次**Hermite**插值或分段高次**Hermite**插值可以利用(注意:分段高次**Lagrange**插值和**Newton**插值等是做不到的,在插值结点上它们只能保证插值函数连续).

样条函数插值问题的提出

但高次Hermite插值在许多场合中看不中用！

- 提高Hermite插值多项式的次数就要增加约束条件——给出插值结点处被插函数及其直到足够高阶导数之值.
- 作为约束条件的所有数据都是通过观测得到的,而观测总难免有误差.

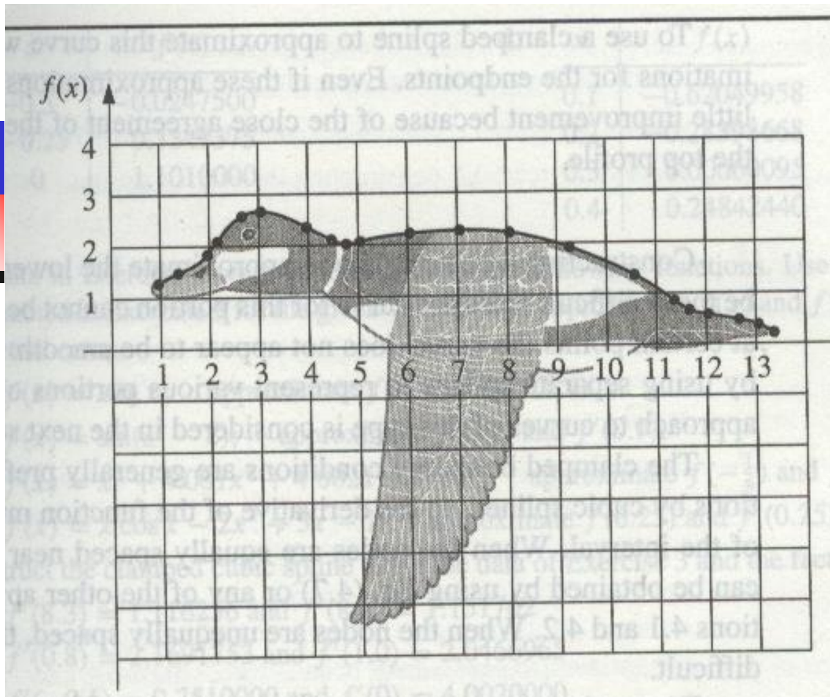
于是　高次插值不仅增添了数据准备和计算的困难,也将导致更大的误差.

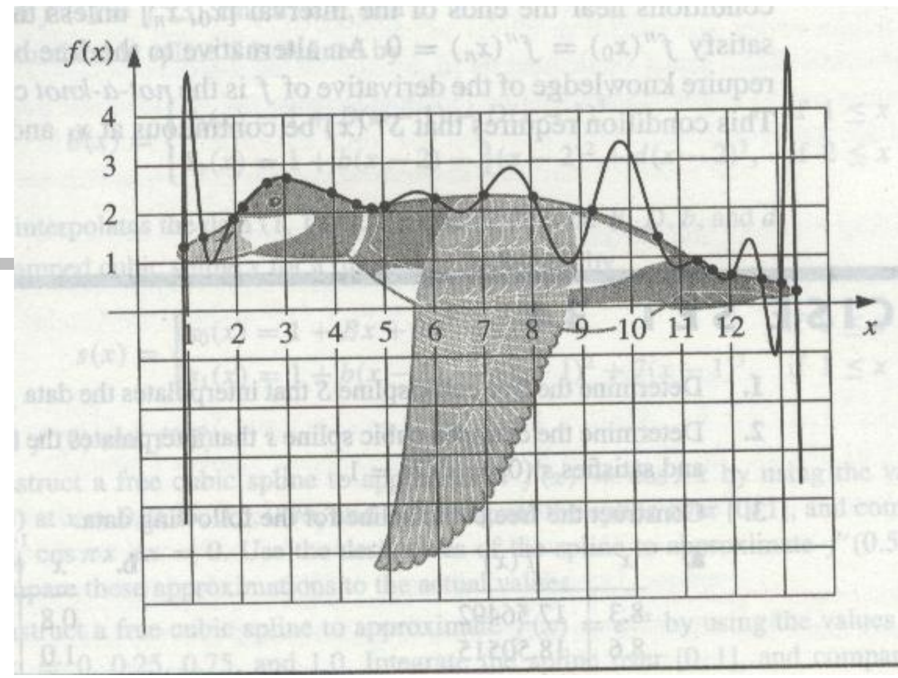
样条函数插值问题的提出

还有许多应用不仅要求插值函数具有足够高阶的整体光滑性，还要求在某些结点处转折灵活. 例如若干点处加载集中力的杆、梁或板弯曲. 这就导致本节要讨论的样条函数(Spline)插值.

数学里的样条(Spline)一词来源于它的直观几何背景：绘图员或板金工人常用弹性木条或金属条加压铁(构成样条!)来绘制或者放样成光顺曲线或者曲面. 但它之所以成为数值分析的标志性成果之一并且在数学物理的广泛领域获得非常成功的应用。

- 相同数据3次样条插值与**Lagrangr**插值效果比较





- **Cubic Spline Interpolation Lagrangr Interpolation**

样条函数插值

样条函数

由一些按照某种光滑条件分段拼接起来的多项式组成的函数。

最常用的样条函数为三次样条函数，即由三次多项式组成，满足处处有二阶连续导数。

定义 设节点 $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$，若函数 $s(x) \in C^2[a,b]$ 在每个小区间 $[x_i, x_{i+1}]$ 上是三次多项式，则称其为三次样条函数。如果同时满足 $s(x_i) = f(x_i)$ $(i = 0, 1, 2, \ldots, n)$，则称 $s(x)$ 为 $f(x)$ 在 $[a, b]$ 上的三次样条函数。

三次样条函数的确定

节点： $x_0 < x_1 < \cdots < x_{n-1} < x_n$

函数值： $y_i = f(x_i)$ $(i = 0, 1, 2, \ldots, n)$

$s(x)$ 满足：$s(x_i) = y_i$ $(i = 0, 1, 2, \ldots, n)$

由定义可设：
$$s(x) = \begin{cases} s_1(x), & x \in [x_0, x_1] \\ s_2(x), & x \in [x_1, x_2] \\ \quad \vdots & \\ s_n(x), & x \in [x_{n-1}, x_n] \end{cases}$$

其中 $s_k(x)$ 为$[x_{k-1}, x_k]$上的三次多项式，且满足

$$s_k(x_{k-1}) = y_{k-1}, \; s_k(x_k) = y_k \quad (k = 1, 2, \ldots, n)$$

边界条件

$$s(x) \in C^2[a,b] \Rightarrow s'(x_k^-) = s'(x_k^+),\ s''(x_k^-) = s''(x_k^+)$$

$$\boxed{s_k'(x_k^-) = s_{k+1}'(x_k^+),\ s_k''(x_k^-) = s_{k+1}''(x_k^+)}$$

$(k = 1, 2, \ldots, n\text{-}1)$

每个 $s_k(x)$ 均为三次多项式，有4个待定系数，所以共有 $4n$ 个待定系数，需 $4n$ 个方程才能确定。前面已经得到 $2n+2(n-1) = 4n-2$ 个方程，还缺 2 个方程！

- 实际问题通常对样条函数在端点处的状态有要求，即所谓的边界条件。

边界条件

- 第一类边界条件：给定函数在端点处的一阶导数，即

$$\boxed{s'(x_0)=f_0',\ s'(x_n)=f_n'}$$

- 第二类边界条件：给定函数在端点处的二阶导数，即

$$\boxed{s''(x_0)=f_0'',\ s''(x_n)=f_n''}$$

当 $s''(x_0)=s''(x_n)=0$ 时，称为自然边界条件，此时的样条函数称为自然样条函数。

- 第三类边界条件：设 $f(x)$ 是周期函数，并设 x_n-x_0 是一个周期，于是 $s(x)$ 满足

$$\boxed{s'(x_0)=s'(x_n),\ s''(x_0)=s''(x_n)}$$

三次样条函数的计算

- 设出 $s(x)$ 在各个节点处的二阶导数值，即

$$\boxed{s''(x_j) = M_j} \quad (j = 0, 1, 2, \ldots, n)$$

考虑区间 $[x_{j-1}, x_j]$，在此区间上， $s(x) = s_j(x)$ 是三次多项式，故 $s_j''(x)$ 为线性函数，且

$$s_j''(x_{j-1}) = s''(x_{j-1}) = M_{j-1},\ s_j''(x_j) = s''(x_j) = M_j$$

利用线性插值公式，即可得 $s_j''(x)$ 的表达式：

$$s_j''(x) = \frac{x_j - x}{h_j} M_{j-1} + \frac{x - x_{j-1}}{h_j} M_j, \quad h_j = x_j - x_{j-1}$$

积分两次后即可得 $s_j(x)$ 的表达式：

三次样条函数的计算

$$s_j(x)=\frac{(x_j-x)^3}{6h_j}M_{j-1}+\frac{(x-x_{j-1})^3}{6h_j}M_j+c_1x+c_2$$

将插值条件 $s_j(x_{j-1})=y_{j-1}$, $s_j(x_j)=y_j$ 代入可确定积分常数 c_1 和 c_2，整理上式得:

$$\boxed{\begin{aligned}s_j(x)=&\frac{(x_j-x)^3}{6h_j}M_{j-1}+\frac{(x-x_{j-1})^3}{6h_j}M_j\\&+\left(y_{j-1}-\frac{M_{j-1}h_j^2}{6}\right)\frac{x_j-x}{h_j}+\left(y_j-\frac{M_jh_j^2}{6}\right)\frac{x-x_{j-1}}{h_j}\end{aligned}}$$

只需确定 $M_0, M_1, \cdots, M_n$ 即可给出 $s(x)$ 的表达式。

三次样条函数的计算

$s(x)$ 在各个节点处的一阶导数存在 $s'(x_j^-) = s'(x_j^+)$

即有 $\boxed{s'_j(x_j^-) = s'_{j+1}(x_j^+)}$

对 $s_j(x)$ 求导得：

$$\boxed{\begin{aligned} s'_j(x) = & -\frac{(x_j - x)^2}{2h_j}M_{j-1} + \frac{(x - x_{j-1})^2}{2h_j}M_j \\ & + \frac{y_j - y_{j-1}}{h_j} - \frac{M_j - M_{j-1}}{6}h_j \end{aligned}}$$

$$\boxed{\begin{aligned} & \frac{h_j}{2}M_j - \frac{h_j}{6}(M_j - M_{j-1}) + \frac{y_j - y_{j-1}}{h_j} \\ & \qquad = -\frac{h_{j+1}}{2}M_j - \frac{h_{j+1}}{6}(M_{j+1} - M_j) + \frac{y_{j+1} - y_j}{h_{j+1}} \end{aligned}}$$

三次样条函数的计算

整理后得关于 M_{j-1}, M_j 和 M_{j+1} 的方程：

$$\mu_j M_{j-1} + 2M_j + \lambda_j M_{j+1} = d_j$$ $\boxed{\text{三弯矩方程}}$

其中

$$\begin{cases} \mu_j = \dfrac{h_j}{h_j + h_{j+1}}, \quad \lambda_j = \dfrac{h_{j+1}}{h_j + h_{j+1}}, \quad \boxed{\mu_j + \lambda_j = 1} \\ d_j = \dfrac{6}{h_j + h_{j+1}} \left(\dfrac{y_{j+1} - y_j}{h_{j+1}} - \dfrac{y_j - y_{j-1}}{h_j} \right) = 6f[x_{j-1}, x_j, x_{j+1}] \end{cases}$$

$j = 1, 2, \ldots, n\text{-}1$

共 n-1 个方程，附加边界条件，补充两个方程后，即可确定 n+1 个未知量 $M_0, M_1, \cdots, M_n$。

第一类边界条件

- 第一类边界条件： $s'(x_0)=y_0'$， $s'(x_n)=y_n'$

  直接代入 $s_j(x)$ 的一阶导数表达式即得

$$2M_0+M_1=6\big((y_1-y_0)/h_1-y_0'\big)/h_1\equiv d_0$$

$$M_{n-1}+2M_n=6\big(y_n'-(y_n-y_{n-1})/h_n\big)/h_n\equiv d_n$$

与前面的 n-1 个方程联立得 n+1 阶线性方程组:

系数矩阵严格对角占优，故矩阵可逆，方程组存在唯一解。

$$\begin{bmatrix} 2 & 1 & & & & \\ \mu_1 & 2 & \lambda_1 & & & \\ & \mu_2 & 2 & \lambda_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \mu_{n-1} & 2 & \lambda_{n-1} \\ & & & & 1 & 2 \end{bmatrix}\begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ \vdots \\ M_{n-1} \\ M_n \end{bmatrix}=\begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}$$

第二类边界条件

- 第二类边界条件：$s''(x_0)=y_0''$，$s''(x_n)=y_n''$

  直接可得 $\boxed{M_0=y_0'',\ M_n=y_n''}$

  前面方程中只含 n-1 个未知量，即可得 n-1 阶线性方程组:

$$\begin{bmatrix} 2 & \lambda_1 & & & \\ \mu_2 & 2 & \lambda_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \mu_{n-2} & 2 & \lambda_{n-2} \\ & & & \mu_{n-1} & 2 \end{bmatrix}\begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_{n-2} \\ M_{n-1} \end{bmatrix}=\begin{bmatrix} d_1-\mu_1 y_0'' \\ d_2 \\ \\ d_{n-2} \\ d_{n-1}-\lambda_{n-1}y_n'' \end{bmatrix}$$

系数矩阵严格对角占优，方程组存在唯一解。

第三类边界条件

- 第三类边界条件： $s'(x_0)=s'(x_n),\ s''(x_0)=s''(x_n)$

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

其中 $\lambda_n=h_1/(h_1+h_n),\mu_n=h_n/(h_1+h_n),$

$$d_n=6\big((y_1-y_0)/h_1-(y_n-y_{n-1})/h_n\big)/(h_1+h_n)$$

与前面的 n-1 个方程联立得 n 阶线性方程组：

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

系数矩阵严格对角占优，方程组存在唯一解。

具体计算过程

- 综上所述，满足插值条件 $s(x_j)=y_j$ 和某一类边界条件的三次样条函数存在且唯一！

- 具体计算过程
  - ✓ 根据插值条件 $s(x_j)=y_j$ 和给定的边界条件列出相应得方程组；
  - ✓ 解出该线性方程组的解 $M_0, M_1, \cdots, M_n$；

    具体求解方法参见第五章和第六章

  - ✓ 将 $M_0, M_1, \cdots, M_n$ 代入 $s_j(x)$ 的表达式，写出三次样条函数 $s(x)$ 在整个插值区间上的分段表达式。

曲线拟合的最小二乘法

高云

曲线拟合的最小二乘法

已知一个函数的数值表

x	x_1	x_2	…	x_m
y	y_1	y_2	…	y_m

求一个简单易算的近似函数 $p(x) \approx f(x)$。

但是 (1) m 通常很大;

(2) y_i 本身是测量值，不准确，即 $y_i \neq f(x_i)$。

这时没必要使 $p(x_i) = y_i$，而只要 $p(x_i) - y_i$ 总体上尽可能小。

常见做法:

- 使 $\max\limits_{1\le i\le m} |p(x_i) - y_i|$ 最小
- 使 $\sum\limits_{i=1}^{m} |p(x_i) - y_i|$ 最小
- 使 $\sum\limits_{i=1}^{m} |p(x_i) - y_i|^2$ 最小 → **最小二乘法**

太复杂☹

不可导，求解困难

曲线拟合的最小二乘问题

- 曲线拟合的最小二乘问题

  已知函数值表 $(x_i, f(x_i))$，在函数空间 Φ 中求 $s^*(x)$，使得

$$\sum_{i=0}^{m} \omega_i [s^*(x_i) - f(x_i)]^2 = \min_{s(x)\in\Phi} \sum_{i=0}^{m} \omega_i [s(x_i) - f(x_i)]^2$$

  其中 ω_i 是点 x_i 处的权。

- 这个问题实质上是最佳平方逼近问题的离散形式。可以将求连续函数的最佳平方逼近函数的方法直接用于求解该问题。

代数多项式拟合

$\Phi = P_n = \mathrm{span}\{1, x, ..., x^n\}$，则相应的法方程为

$$\begin{bmatrix} \sum_{i=0}^{m}\omega_i & \sum_{i=0}^{m}\omega_i x_i & \cdots & \sum_{i=0}^{m}\omega_i x_i^n \\ \sum_{i=0}^{m}\omega_i x_i & \sum_{i=0}^{m}\omega_i x_i^2 & \cdots & \sum_{i=0}^{m}\omega_i x_i^{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=0}^{m}\omega_i x_i^n & \sum_{i=0}^{m}\omega_i x_i^{n+1} & \cdots & \sum_{i=0}^{m}\omega_i x_i^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{m}\omega_i f_i \\ \sum_{i=0}^{m}\omega_i x_i f_i \\ \vdots \\ \sum_{i=0}^{m}\omega_i x_i^n f_i \end{bmatrix}$$

对称矩阵

此时 $s^*(x) = \sum_{k=0}^{n} a_k^* x^k$ 称为数据拟合多项式，上述拟合称为多项式拟合。

举例（二）

- 例：求下面数据表的二次最小二乘拟合多项式。

x_i	0	0.25	0.50	0.75	1.00
$f(x_i)$	1.0000	1.2840	1.6487	2.1170	2.7183

解：设二次拟合多项式为 $p_2(x)=a_0+a_1x+a_2x^2$

可得法方程
$$\begin{bmatrix} 5 & 2.5 & 1.875 \\ 2.5 & 1.875 & 1.5625 \\ 1.875 & 1.5625 & 1.3828 \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}=\begin{bmatrix} 8.7680 \\ 5.4514 \\ 4.4015 \end{bmatrix}$$

解得 $a_0=1.0052, a_1=0.8641, a_2=0.8437$

所以此数据组的二次最小二乘拟合多项式为

$$p_2(x)=1.0052+0.8641x+0.8437x^2$$

注：(1) 若题目中没有给出各点的权值 ω_i，默认为 $\omega_i=1$。
(2) 该方法不适合 n 较大时的情形。（病态问题）

矛盾方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

无解，则称方程组是矛盾的或不相容的

残差平方和

$$I(x_1, x_2, \ldots, x_n) = \sum_{k=1}^{m} [b_k - (a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n)]^2$$

求 $(x_1, x_2, \ldots x_n)$ 使得残差平方和极小。

方法：对 x_i 分别求偏导数，令其等于 0，得法方程。

举例（四）

- 例：确定下面的矛盾方程的最小二乘解。

$$\begin{cases} 2x+3y=1 \\ x-4y=-9 \\ 2x-y=-1 \end{cases}$$

解：令

$$I(x,y)=(2x+3y-1)^2+(x-4y+9)^2+(2x-y+1)^2$$

求偏导数得
$$\begin{cases} \dfrac{\partial I}{\partial x}=18x+18=0 \\ \dfrac{\partial I}{\partial y}=52y-80=0 \end{cases} \Rightarrow \begin{cases} x=-1 \\ y=20/13 \end{cases}$$

```
A = [2,3; 1,-4; 2,-1]; b = [1;-9;-1];
x = A\b; % 若是矛盾方程，则给出最小二乘解。
```