插值与拟合

Optimization and operational model

成都信息工程大学学生数学建模协会

数值计算方法_华北理工大学_中国大学MOOC(慕课) (icourse163.org)

Hermite插值方法思路 - 知乎 (zhihu.com)

2021数值分析与科学计算实践内容（Python）随课程进度更新..._哔哩哔哩_bilibili

马昌凤 林伟川《现代数值计算方法》

清风数学建模课程 B站“数学建模学习交流”

【Python】利用Python拟合函数曲线 - CloudSir - 博客园 (cnblogs.com)

主讲人：熊诗豪

目录

CONTENTS

01/ 插值与拟合介绍

Mathematical modeling and its application

什么是插值与拟合

在数据分析中，原始数据往往产生有缺失值，需要根据通过已知数据进行数据处理才能有效分析，而有时候现有的数据是缺失严重的，不足以支撑分析的进行，这时就需要利用一些数学方法，拟构一些新的但比较合理的数据来满足分析要求，因此产生插值算法和拟合算法

一般而言，插值算法比较适合数据量较小或数据较准确的情况，而拟合算法比较适合数据量较大或数据有较大偏差的情况

02 / 算法原理

Mathematical model

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

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

故有

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

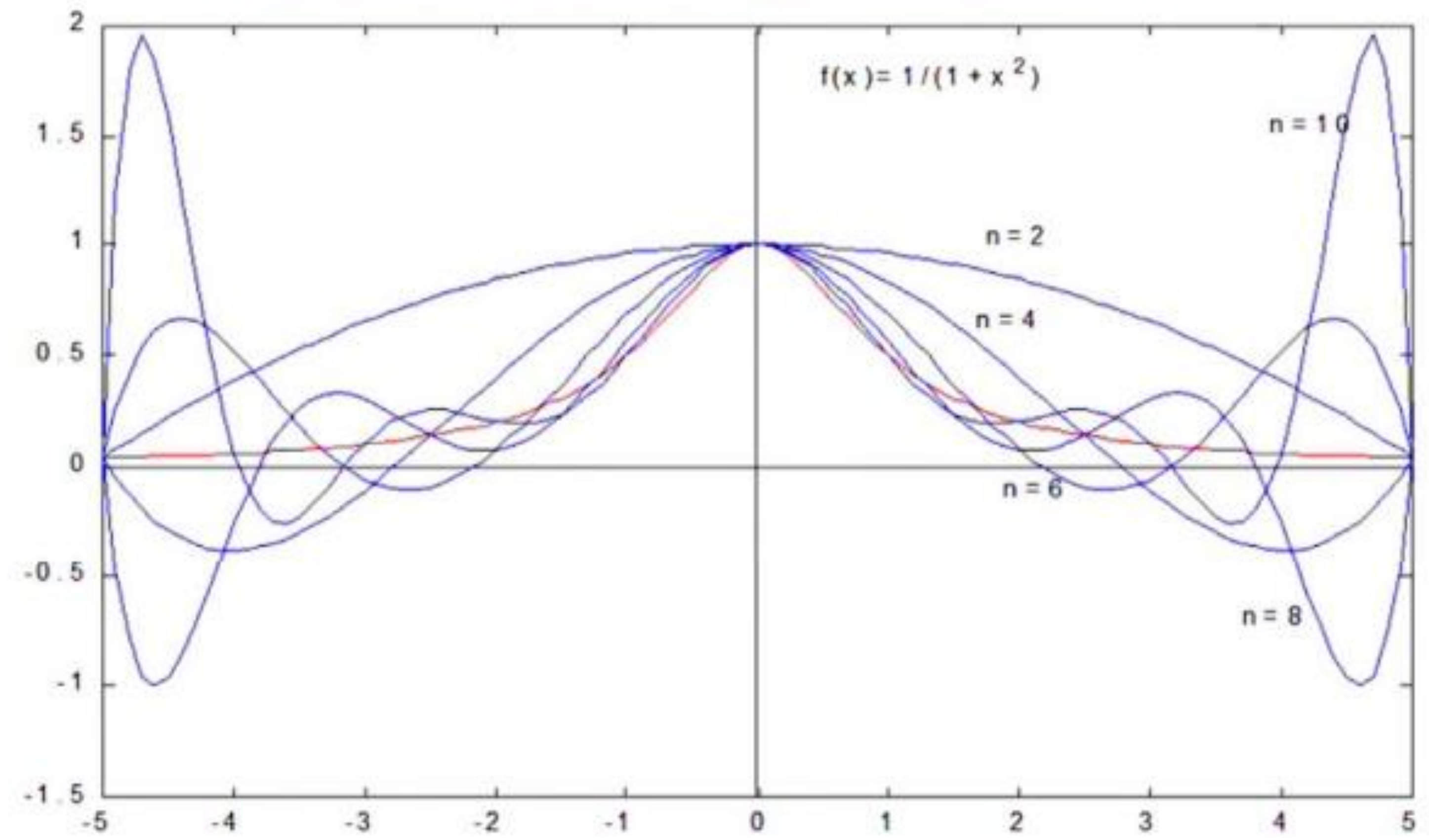
公式 (4.13) 称为 n 阶拉格朗日插值公式, 其中 $l_i(x)\,(i = 0, 1, \cdots, n)$ 称为 n 阶拉格朗日插值的基函数.

摘自《现代数值计算方法》

通过公式推导，可以看出拉格朗日是n次多项式插值，而在分析过程应注意龙格现象

龙格现象通常出现于高次多项式插值，即在两端波动极大，产生明显的震荡。在不熟悉曲线运动趋势的前提下，不要轻易使用高次插值。

不同次数的拉格朗日插值多项式的比较图



牛顿插值

差商定义：

$$f[k]=f(x_k)$$

$$f[x_{k-1},x_k]=\frac{f[x_k]-f[x_{k-1}]}{x_k-x_{k-1}}$$

得到高次差商递推公式

$$f[x_{n-j},x_{n-j+1},\cdots,x_n]=\frac{f[x_{n-j+1},\cdots,x_n]-f[x_{n-j},\cdots,x_{n-1}]}{x_n-x_{n-j}}$$

牛顿插值公式：

$$\begin{aligned}f(x)=&f(x_0)+f[x_0,x_1](x-x_0)\\&+f[x_0,x_1,x_2](x-x_0)(x-x_1)+\cdots\\&+f[x_0,x_1,\cdots,x_{n-2},x_{n-1}](x-x_0)(x-x_1)\cdots(x-x_{n-3})(x-x_{n-2})\\&+f[x_0,x_1,\cdots,x_{n-1},x_n](x-x_0)(x-x_1)\cdots(x-x_{n-2})(x-x_{n-1})\end{aligned}$$

而拉格朗日插值计算没有递推关系，因此在增加差商关系，对拉格朗日插值进行递推计算，这种插值方法称为牛顿插值，同时在牛顿插值中也存在龙格现象。与拉格朗日插值相比，牛顿插值具有继承性和易变动性。

但是从本质上和拉格朗日插值方法相同，都不光滑。

Hermite插值

Hermite 插值原理（简明） 具有节点的导数值约束的插值

Hermite 插值法

设函数 $f(x)$ 在区间 $[a, b]$ 上有 n+1个互异节点 $a=x_0<x_1<x_2<\ldots<x_n=b$，定义在$[a,b]$上函数 $f(x)$ 在节点上满足:

$$f(x_i)=y_i,\ f'(x_i)=y_i' \quad (i=0,1,2,\ldots n) \qquad (\ 2n+2 \text{ 个条件})$$

可唯一确定一个次数不超过 $2n+1$ 的多项式 $H_{2n+1}(x)=H(x)$ 满足：

$$H(x_j)=y_j, \qquad H'(x_j)=m_j \quad (j=0,1,\cdots n).$$

其余项为：

$$R(x)=f(x)-H(x)=\frac{f^{(2n+2)}(\xi)}{(2n+2)!}\omega_{2n+2}(x)$$

Hermite插值产生原因：

为了解决两点之间用直线连接进行插值导致的曲线不光滑的现象

摘自 数值计算方法_华北理工大学_中国大学MOOC(慕课)

样条插值

介绍：

一种以可变样条来作出一条经过一系列点的光滑曲线的数学方法。插值样条是由一些多项式组成的，每一个多项式都是由相邻的两个数据点决定的，这样，任意的两个相邻的多项式以及它们的导数在连接点处都是连续的。

简单理解，就是每两个点之间确定一个函数，这个函数就是一个样条，函数不同，样条就不同，所以定义中说 可变样条，然后把所有样条分段结合成一个函数，就是最终的插值函数。

原理：

对于待插值函数$f(x)$，已知节点$x_0,x_1,\dots,x_n$ 处的函数值，将相邻两节点进行分段，获得n个插值小区间，

在每个区间内使用k次多项式$S_i(x)$插值，使其满足插值条件与$k-1$阶平滑性：

$$S_i(x_i)=f(x_i),S_i(x_{i+1})=f(x_{i+1}),0\leqslant i\leqslant n-1$$

$$S_i^{(j)}(x_{i+1})=S_i^{(j)}(x_{i+1}),0\leqslant i\leqslant n-2,1\leqslant j\leqslant k-1$$

拟合原理

数据拟合又称曲线拟合，俗称**拉曲线**，是一种把现有数据透过数学方法来代入一条数式的表示方式。科学和工程问题可以通过诸如采样、实验等方法获得若干离散的数据，根据这些数据，我们往往希望得到一个连续的函数（也就是曲线）或者更加密集的离散方程与已知数据相吻合，这过程就叫做**拟合(fitting)**。

简单理解就是一个完整函数使其尽量贯穿所有数据点

感谢聆听

现实世界的奥秘等你探索和发现，

体验数学魅力， 让你收益终身！