

实验数据的插值

袁略真

3130103964

生物信息学

浙江大学

2016 年 3 月 13 日

1 n次插值

1.1 插值的含义

当两物理量 x, y 之间存在函数关系 $y=f(x)$,而其具体的函数关系式并不知道,为寻求其函数关系,常通过实验测得一组实验数据 x_0, \dots, x_n 及其对应的函数值 y_0, \dots, y_n .插值根据这些对应关系,寻求函数 $f(x)$ 的近似表达式,它常是 x 的多项式.插值满足的条件是:

1. 多项式的待定系数不超过数据个数
2. 多项式经过全部所给数据点

1.2 插值的应用场合

1. 在已知数据具有高精度的情况下,甚至数据是通过推导绝对正确的情况下,获取给定数据点之外的函数取值.当已知值存在显著误差,则应考虑拟合.
2. 插值,如本堂课讲的用多项式进行插值,在计算上比较容易,计算结果又与目标函数值十分接近.
3. 有时,目标函数是已知的,如 $\sin(x)$.计算机软件中它的值是通过多项式逼近来计算的¹.

1.3 使用n次拉格朗日多项式插值

n 次插值多项式需要 $n+1$ 个插值点.插值多项式和插值基函数由一下两式表示.

$$y(x) = \sum_{j=0}^n A_j(x) y_j$$

$$A_j(x) = \prod_{i=0, i \neq j}^n \frac{x - x_i}{x_j - x_i}$$

¹书<数值方法(MATLAB版)(第四版)>,(John H.Mathews, Kurtis D.Fink著)第4章.

2 程序流程

Input: $x_0, x_n, n, file$

Output: a matrix of x , predicted value y , the real value y_{real}

for $i \leftarrow 0$ **to** N **do**

 | calculate x_i, y_i , and store in the vector x , and y ;

end

for $x_t \leftarrow x_0$ **to** x_n , *sample 100 points* **do**

$result \leftarrow 0$;

for $j \leftarrow 0$ **to** n **do**

$A_j \leftarrow 1$;

for $i \leftarrow 0$ **to** n **do**

if $i \neq j$ **then**

 | $A_j^* = (x_t - x[i]) / (x[j] - x[i])$;

end

end

$result += A_j^* y[j]$;

end

 print $x_t, result, y(x_t)$;

end

Algorithm 1: 拉格朗日插值法

3 模拟及结果

3.1 模拟的参数设置

固定 $x_0=5, x_n=95$, file name = "result". 令 $n=2, 4, 6$ 分别运行3轮.

3.2 模拟的图形结果

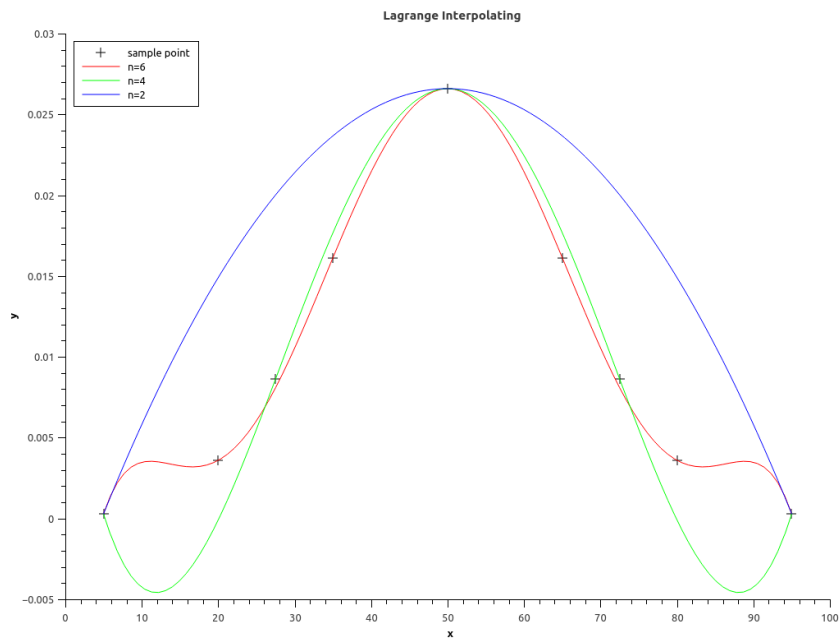


图 1: 拉格朗日插值法分别在 $n=2, 4, 6$ 下的插值曲线. 其中目标函数是正态分布的概率密度函数, 均值 50.0, 方差 15.0. 如图例所示, 黑十字表示在 3 轮插值中先后出现的样本点, 蓝色曲线表示仅有 3 样本点时抛物线型曲线, 绿色、红色曲线则分别对应 4、6 个样本点时曲线.