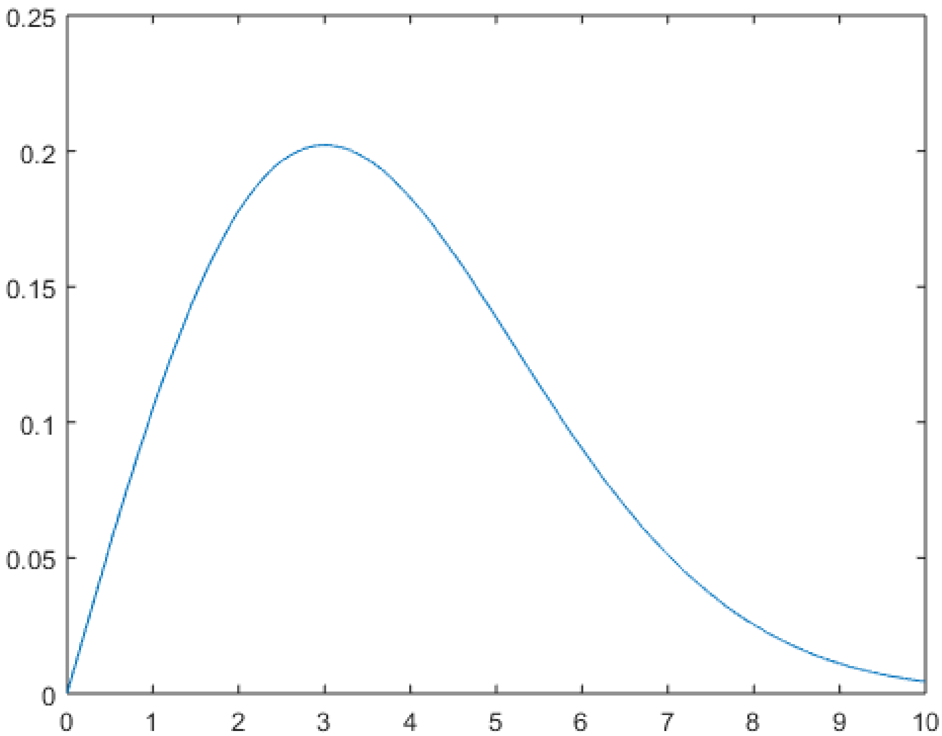
Project

Numerical Analysis 2019

问题：

如图是服从某分布的概率密度曲线，我们将重构下面的曲线。



1. 给定采样点（1, 0.1051）,(4, 0.1827) , (7, 0.0511) , (10, 0.0043)，至少采用两种方案对该曲线进行重构。报告中应给出重构的关键步骤以及最终结果。

   插值法：(递归的)拉格朗日插值多项式、牛顿分插多项式、三次分段多项式近似，埃尔米特插值

   其实，两者都是通过给定 n+1 个互异的插值节点，让你求一条 n 次代数曲线近似地表示待插值的函数曲线．这就叫做代数插值啦．Lagrange 插值代数和 Newton 法插值都属于代数插值的范畴．Lagrange 插值和 Newton 法插值的结果和余项都是一致的，因为都是利用 n 次多项式插值嘛，当然一样啦．区别：Lagrange 插值法是通过构造 n+1 个 n 次基本多项式，然后线性组合（结果当然也是 n 次的多项式啦）而得到的．而 Newton 法插值是通过求各阶差商，递推得到的一个 f(x)=f(x0)+(x-x0)f[x0,x1]+(x-x0)(x-x1)f[x0,x1,x2]....+(x-x0)...(x-x(n-1))f[x0,x1...xn]这样的公式，代进去就可以得到啦（其实一楼概括的很深入吧，抱歉我还没达到那种境界，呵呵）. 还有，Lagrange 插值法在求每个基本多项式的时候要用到所有那些结点，因此如果需要再多加进去一个结点的话，需要重新求出基本多项式才可，而这需要大量的工程，于是数学家们就发明了

Newton 法啦，你看上面的那个式子，如果再加进去一个结点是不是只要在它后面再加上一个（x-x0)(x-x1)...(x-x(n-1))(x-xn)f[x0,x1...xn,x(n+1)]

就行了呢？

2. 实际上，上述概率密度曲线对应于参数为 3 的瑞利分布，即 $f(x)=\frac{x}{\sigma^2}\exp(-\frac{x^2}{2\sigma^2})$ ，其中 $\sigma=3$（这里我们限制 $x\in[0,10]$）。计算并比较（1）中结果在 L2 范数(求积分)下的误差。如果采用 L_∞ 范数（差的绝对值最大）计算误差，结果又会怎样？

Lagrange L_infinity:0.019209604129851177
Netwon L_infinity:0.019209604129851163
Cubic L_infinity:4.533178363547163e-06
Hermite L_infinity:0.019209604129851177
Lagrange L_2:(0.000878331778560006, 6.806553079181522e-16)
Netwon L_2:(0.0008783317785600052, 6.806536204697026e-16)
Cubic L_2:(0.0008214155424958413, 1.3610321962839917e-09)
Hermite L_2:(0.000878331778560006, 6.806553079181522e-16)

3. （1）中给出的数据点是最优的吗？如果不是，应该如何选取用于重构曲线的数据点？请给出你的设计思路与解决方案。

	X 整数部分 3 位	Y 小数部分前两位 5 位
点 1	1	
点 2	3	
点 3	4	
点 4	6	
点 5	7	
	00 4	5
点 6	8	
点 7	10	

设计实验比较：数据点少但是精度高 VS 数据点多精度低

4. 考虑如下场景：甲在 A 地生成了如上图所示的曲线，想要把它发送给在 B 地的乙。我们

假设 AB 两地之间的通信容量是 60 bits。即甲只能用 60 位二进制数据描述发送给乙的数据。

我们先来考虑两种极端的情况。第一种是甲将所有的数据（x，y）全部发送给乙，而这样做产生的数据量将远远超过 60 bits，因此该方案不可行。另一种极端的情况是甲和乙都知道传输的数据是一个瑞利分布的概率密度曲线，因此甲只需要发送数值 3（σ 的取值），乙就可以根据 $f(x)=\frac{x}{\sigma^2}\exp(-\frac{x^2}{2\sigma^2})$ 来重构这条曲线。3 对应于二进制 11，因此甲只需要用两个 bit 来描述这条曲线。

在实际情况下，甲和乙都不知道函数 $f(x)$ 的解析式，因此甲发送若干个数据点（以二进制的形式），乙则根据接收到的数据点对该曲线进行重构。在 60 bits 的限制条件下，甲又该如何利用这 60 bits？请给出甲的解决方案和乙重构的结果。报告中应包含但不限于甲选择该方案的理由以及对乙重构结果的分析。

附注：

1、第 1、2 小题侧重你对课堂知识的利用（可能需要你编程来解决），第 3、4 小题则侧重你的设计思路。

2、L2 范数下的误差可以表达为 $E=\int(f(x)-\hat{f}(x))^2\mathrm{d}x$， L_∞ 范数下的误差可以表达为 $E=\sup\{\left|f(x)-\hat{f}(x)\right|\}$

3、这里给出（4）中的一种方案，帮助大家理解。假设甲、乙约定采用 $a\times10^b$ 的形式来描述任意一个数，其中 a、b 分别采用 5 个 bit 描述，而且描述 b 的 5 个 bit 中第一个表示符号：0 表示正号，1 表示负号。另外，甲、乙约定用 20 个 bit 来描述一个数据点：前 10 个 bit 描述 x，后 10 个 bit 描述 y。上述“约定”即甲、乙之间传输数据的一份协议。它唯一确定了一组数据的编码与解码方式，在传输数据之前，甲、乙双方都知道这份协议的内容。

现在甲选取（1, 0.11）,(4, 0.18) , (7, 0.05)这三个数据点，并将其编码：

00001,00000, 01011,10010，00100,00000,10010,10010，00111,00000,00101,10010（其中加了逗号是为了方便阅读）

乙收到上面这串序列后，就可以根据协议无差错地恢复（1, 0.11）,(4, 0.18) , (7, 0.05)这三个数据点。

提交内容（主要是报告）：

1. 报告。要求：包括但不限于设计思路，运行结果，和相应的分析（结论）。

   结果可重现。

   语言简练。

2. 代码。要求：上传的代码可以运行。

   有详细的注释。

上传：

将内容打包，上传至学在浙大课程网站。

命名方式："学号_姓名_Project"

截止时间：2020 年 1 月 16 日 23:59。

```
ax2.plot(x, H(x)-f(x),'--',label='Hermite')
ax2.plot(x, N(x)-f(x),'g',label='Newton')
ax2.plot(x, L(x)-f(x),'r--',label='Lagrange')
ax2.plot(x_cubic, l(x_cubic)-f(x_cubic),'b--',label='cubic',lw=0.5)
plt.legend(loc=1)
ax2.set_ylabel(r'$absolute error$', fontsize=18)
```