数值分析上机报告3

段浩东　**1500017705**

上机习题3.1

题目

设$f(x) = lnx$，分别取$h = 1/10^k, (k = 1, 2, \dots, 10)$，用下列三个公式计算$f'(0.7)$

$$f'(x) = \frac{f(x+h)-f(x)}{h}$$

$$f'(x) = \frac{f(x+h)-f(x-h)}{2h}$$

$$f'(x) = \frac{f(x-2h)-8f(x-h)+8f(x+h)-f(x+2h)}{12h}$$

1. 用双精度计算，列表比较三个公式的计算误差
2. 请推导出第三个公式及其截断误差的阶
3. 从这里我们可以得出什么结论

解答

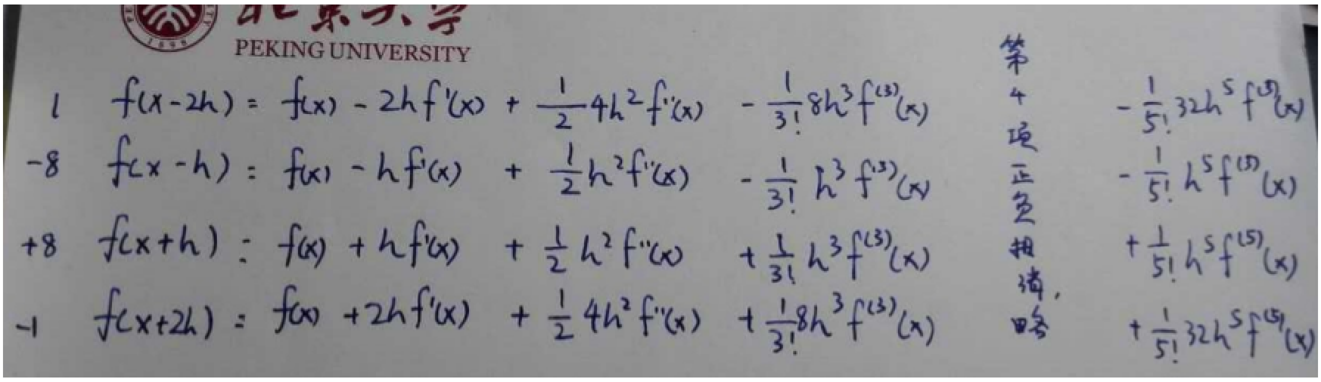
1

equation\k	1.00000e-001	1.00000e-002	1.00000e-003	1.00000e-004	1.00000e-005	1.00000e-006	1.00000e-007	1.00000e-008	1.00000e-009	1.00000e-010
eq1	-9.32575e-002	-1.01079e-002	-1.01944e-003	-1.02031e-004	-1.02040e-005	-1.02038e-006	-1.02892e-007	-7.51937e-010	-5.62631e-008	2.76804e-007
eq2	9.83893e-003	9.71936e-005	9.71819e-007	9.71802e-009	9.18323e-011	2.52192e-011	-7.51937e-010	7.57474e-009	-5.62631e-008	2.76804e-007
eq3	-5.13166e-004	-4.76339e-008	-4.71978e-012	-3.15969e-013	-6.23723e-012	5.29747e-011	-9.36974e-010	1.17381e-008	-8.86446e-008	4.61841e-007

三个公式的截断误差如图所示

我们将会发现，三个公式的效果各有不同，而且误差随着k的增大都是先降后升，降的部分是由近似公式的误差阶所主导，而升的部分是因为这些格式都存在舍入误差ϵ/h，h越小，这一项在误差之中的贡献越大。三个格式达到最小误差时的k也有所不同，这取决于其截断误差阶的区别。

2



PEKING UNIVERSITY

$$
\begin{array}{lll}
1 & f(x-2h) = f(x) - 2hf'(x) + \frac{1}{2}4h^2f''(x) - \frac{1}{3!}8h^3f^{(3)}(x) & -\frac{1}{5!}32h^5f^{(5)}(x) \\
-8 & f(x-h) = f(x) - hf'(x) + \frac{1}{2}h^2f''(x) - \frac{1}{3!}h^3f^{(3)}(x) & -\frac{1}{5!}h^5f^{(5)}(x) \\
+8 & f(x+h) = f(x) + hf'(x) + \frac{1}{2}h^2f''(x) + \frac{1}{3!}h^3f^{(3)}(x) & +\frac{1}{5!}h^5f^{(5)}(x) \\
-1 & f(x+2h) = f(x) + 2hf'(x) + \frac{1}{2}4h^2f''(x) + \frac{1}{3!}8h^3f^{(3)}(x) & +\frac{1}{5!}32h^5f^{(5)}(x)
\end{array}
$$

第4项正负相消，略

而格式三最终的结果为

$$f'(x) - \frac{1}{30}h^4f^{(5)}(x)$$

因此格式三的误差阶为$O(h^4)$，这一点在此前求出的数值误差结果中也可以大致加以验证。

3

根据课堂所学内容，若一个数值微分格式的误差阶为$O(h^n)$，则其加上舍入误差之后的误差大致为$Mh^n + \epsilon/h$，其中epsilon为机器精度。因此，我们为了得到最小的误差，应该取h使得上式最小，实验的结果也验证了这一点。

上机习题3.2

题目

设$f(x) = e^{-x/4},\ 0 \le x \le 1$。取$h = 1/10$，$x_i = ih\ (i = 0, 1, \cdots, 10)$，分别用显式中心格式和隐式格式计算$f(x)$在节点$x_1, x_2, \cdots, x_9$处的一阶导数值并与精确值作比较。

解答

显式中心格式的计算公式为

$f'(x) = \frac{f(x+h)-f(x-h)}{2h} + O(h^2)$ (3,1)

而对于隐式格式：

先将$f(x-h), f(x+h)$分别展开到6阶和7阶，得：

$f'(x) = \frac{f(x+h)-f(x-h)}{2h} - \frac{h^2}{6}f'''(x) + O(h^4)$ (1)
$f''(x) = \frac{f(x+h)-2f(x)+f(x-h)}{h^2} - \frac{h^2}{12}f^{(4)}(x) + O(h^4)$ (2)

再对2式两边求一次导，代入1式，可得方程组：

$$m_i = \frac{f_{i+1} - f_{i-1}}{2h} - \frac{1}{6}(m_{i+1} - 2m_i + m_{i-1}),\ 1 \le i \le n-1.$$

补充边界条件：

$m_0 = f'(x0), m_n = f'(x_n)$ 即可进行求解，结果如下：

x	显示结果	隐式结果	真值	显式误差	隐式误差
0.1	-0.24385288	-0.24382050	-0.24382748	2.540e-05	6.979e-06
0.2	-0.23783213	-0.23780923	-0.23780736	2.477e-05	1.870e-06
0.3	-0.23196003	-0.23193537	-0.23193587	2.416e-05	5.035e-07
0.4	-0.22623292	-0.22620950	-0.22620935	2.356e-05	1.413e-07
0.5	-0.22064721	-0.22062416	-0.22062423	2.298e-05	6.446e-08
0.6	-0.21519941	-0.21517711	-0.21517699	2.241e-05	1.137e-07
0.7	-0.20988612	-0.20986386	-0.20986426	2.186e-05	3.932e-07
0.8	-0.20470401	-0.20468414	-0.20468269	2.132e-05	1.456e-06
0.9	-0.19964985	-0.19962362	-0.19962905	2.080e-05	5.435e-06

其精确值由$f'(x)=-f(x)/4$计算而出。我们可以看出，利用隐式格式计算得到的微分近似值，其精确度比显式格式要高两阶左右。这是由于，显式格式的截断误差为$O(h^2)$，而隐式格式的截断误差为$O(h^4)$，所以隐式格式的精确度相对而言比显式格式要好，且隐式格式具有更好的数值稳定性。但隐式格式的求解一般需要明确知道两端的导数值，因而较显式格式使用了更多的信息。

上机习题**3.4**

题目

已知积分

$$\int_0^1 \frac{4}{1+x^2} = \pi$$

成立，我们可以通过对上面给定被积函数求数值积分来计算π的近似值。

（1）分别使用复合中点公式、复合梯形公式和复合Simpson公式计算π的近似值。选择不同的h，对每种求积公式，试将误差刻画为h的函数，并比较各方法的精度。是否存在某个h值，当小于这个值之后再继续减小h的值，计算不再有所改进？为什么？

（2）实现Romberg求积方法，并重复上面的计算。

（3）使用自适应求积方法重复上面的计算。

解答

复合中点公式、复合梯形公式和复合Simpson公式即是在将积分区间进行分段，并应用中点公式、梯形公式和Simpson公式。若要近似求函数$f(x)$在区间$[a,b]$上的积分，记步长$h=(b-a)/n$，子区间端点$x_i=a+ih\ (0\le i\le n)$，则有积分近似公式

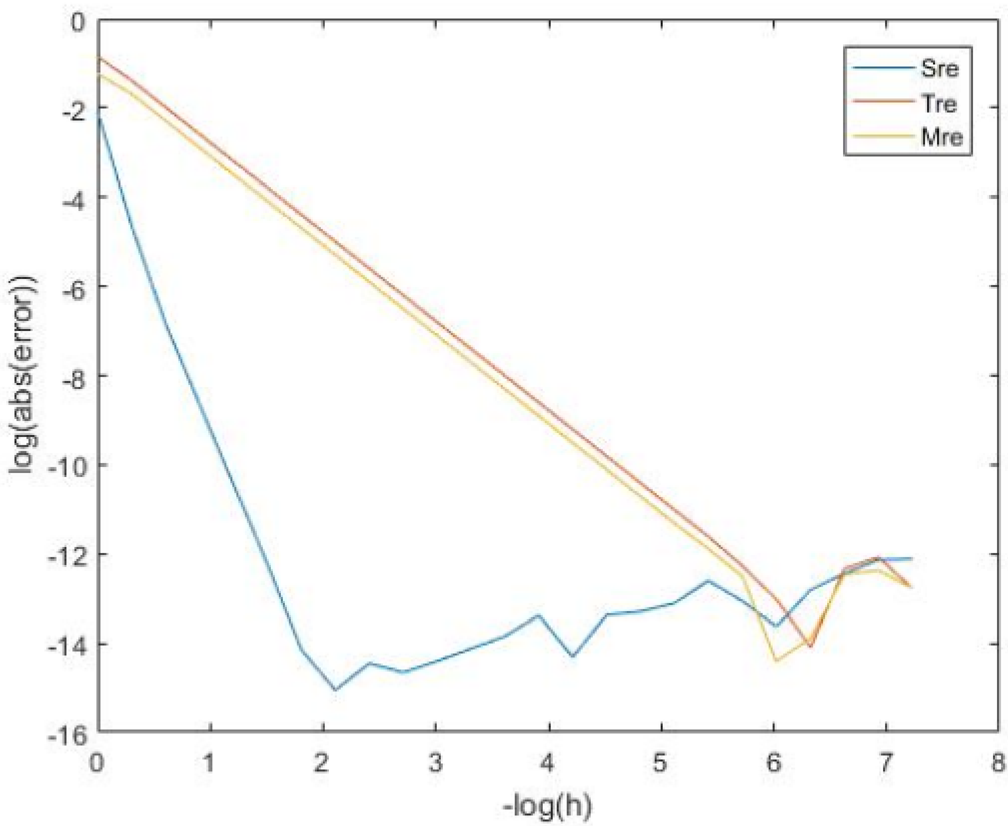
$$\int_a^b f(x)dx \approx h\sum_{i=0}^{n} f(x_{i+\frac{1}{2}}) \triangleq M(h),$$

$$\int_a^b f(x)dx \approx \frac{h}{2}\sum_{i=0}^{n-1}[f(x_i)+f(x_{i+1})] \triangleq T(h)$$

$$\int_a^b f(x)dx \approx \frac{h}{6}\sum_{i=0}^{n-1}[f(x_i)+4f(x_{i+\frac{1}{2}})+f(x_{i+1})] \triangleq S(h)$$

这三者的误差分别为$O(h^2),O(h^2),O(h^4)$

利用三个公式，针对不同的步长h，计算积分$\int_0^1 \frac{4}{1+x^2}dx$来作为$\pi$的近似值，结果如图：

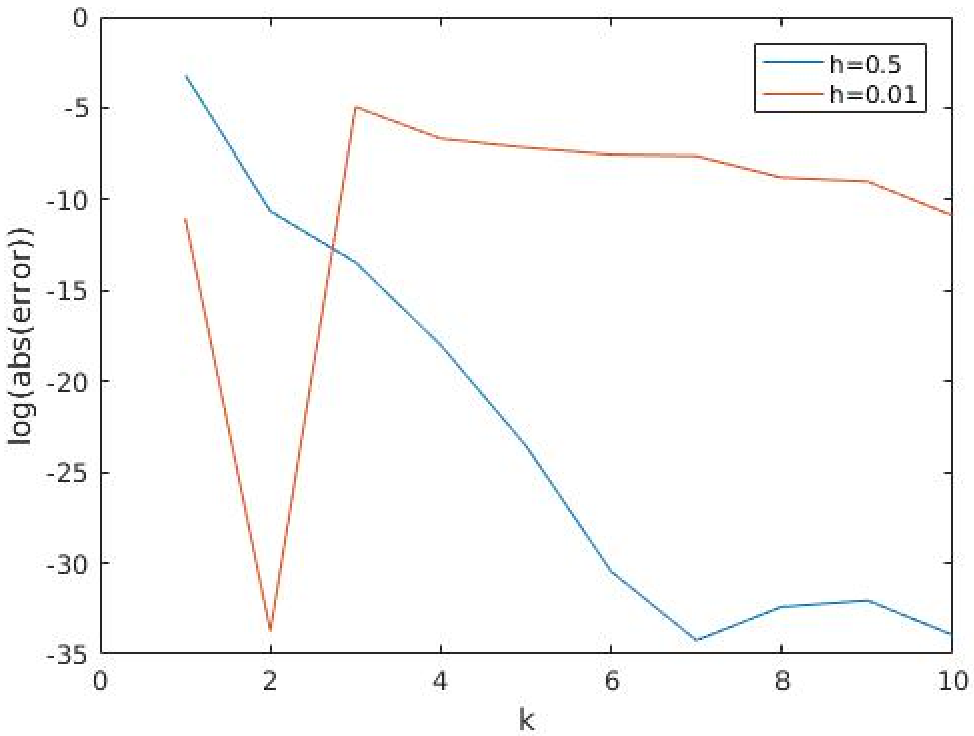


三条曲线分别为利用三种方法得到的误差对数与h对数的曲线，可以看出，随着步长的减小，中点公式和梯形公式的逼近误差都在不断减小。Simpson公式的收敛速度很快，在$h=10^{-2}$时就达到了14位有效数字的良好精度，但当h进一步减小时，计算结果并没有得到改进。而且观察误差曲线的斜率，我们发现一个现象：误差曲线的斜率似乎与方法的误差阶不尽相同，这一现象的原因仍有待探索。

在三种方法的h小到小一定程度的时候，误差均不再有改进，这样的原因是因为这时误差的主要贡献者变为了机器精度带来的舍入误差ϵ/h 。

根据Euler-Maclaurin公式，复合梯形公式的截断误差可表示为
$\int_a^b f(x)dx - T_1(h) = c_2h^2 + c_4h^4 + \cdots$
利用Richardson外推技术，将步长减半后有
$\int_a^b f(x)dx - T_1\left(\frac{h}{2}\right) = c_2\frac{h^2}{4} + c_4\frac{h^4}{16} + \cdots$
消去h^2，有
$\int_a^b f(x)dx - \frac{4T_1(h/2)-T_1(h)}{3} = O(h^4)$
另外，一般地，我们有：

$T_{k+1}(h) = \frac{4^kT_k(h/2)-T_k(h)}{4^k-1}, \quad k=1,2,\cdots$
利用这一递推，我们即可进行romberg求积，选取h=0.5及0.01，绘出了误差曲线与k的关系：



我们可以发现，在h相对取值较小的情况下，仅进行了两次外推就取得了非常高的精度，大大提升了计算精度，而这一曲线之后的上升可理解为由设入误差的放大所致。

自适应求积方法的出发点在于，在函数的导数变化幅度较大的子区间上，应考虑使用更小的步长作估计，从而才能达到更加精确的结果。而导数变化不大的区间可以采用更大的步长来估计，从而减少后者区间上不必要的计算。换言之，就是根据区间的特点而选择不同的步长。

具体计算中我们将这样实现自适应积分：对于整个积分区间上的积分，我们先设定一个容差ϵ，表示我们期望的计算结果与真值间所能容许的误差。另外，我们设定初始步长h，对于每一个在区间[a,b]上的积分，步长为h，容差为ϵ的积分，我们先以h为步长用某种方法在[a,b]上积分，再以h/2为步长分别在区间[a,(a+b)/2],[(a+b)/2,b]上进行积分并求和。若前者与后者的差值大于ϵ，则分别以h/2为步长，$\epsilon/2$为容差对$[a,(a+b)/2]$, $[(a+b)/2,b]$进行积分并求和，否则则取得到的值为[a,b]区间的积分值。

利用这一方法，为了避免出现无穷递归，我们还需要设定一个递归下界，如步长小于某一值时不再继续进行划分。

对一些数值使用此方法进行了实验：

ϵ	**value**
10^{-1}	3.14156862745098042211
10^{-4}	3.14159250245870680374
10^{-7}	3.14159265660246589391
10^{-10}	3.14159265357082162495

可以看出，在一定的精度范围内，这一方式都有着与设定相符的效果。

上机习题3.8.3

题目

已知积分

$$\int_{-\infty}^{+\infty} \exp(-x^2)\cos x dx = \sqrt{\pi}\exp\left(\frac{-1}{4}\right).$$

利用Gauss-Hermite求积公式在不同阶对应的积分节点和权数计算该积分。

解答

带权数值积分公式$\int_a^b \rho(x)f(x)dx \approx \sum_{k=1}^n A_k f(x_k)$的代数精度不超过$2n-1$阶，而Gauss求积公式通过适当选取$A_k$和$x_k$，使得代数精度恰为$2n-1$阶。具体选取方法为，积分节点$x_k$ $(1 \le k \le n)$为n阶正交多项式的零点，求积系数$A_k = \int_a^b \rho(x)l_k(x)dx$
其中$l_k(x)$为关于x_k的Lagrange插值基函数。

- 节点与权数数值来源：http://www.pitt.edu/~dejong/dbook/code/gauss_hermite.txt
- 计算结果

```
N=3      Value=1.38203307138804665222     Error=0.00164462434490375564
N=4      Value=1.38032975716125538845     Error=-0.00005868988188750812
N=5      Value=1.38039007593565621335     Error=0.00000162889251331677
N=6      Value=1.38038841005073376067     Error=-0.00000003699240913591
N=7      Value=1.38038844775407842924     Error=0.00000000071093553267
N=8      Value=1.38038844703130036962     Error=-0.00000000001184252696
N=9      Value=1.38038844704331653546     Error=0.00000000000017363888
N=10     Value=1.38038844704314134226     Error=-0.00000000000000155431
N=11     Value=1.38038844704314422884     Error=0.00000000000000133227
N=12     Value=1.38038844704314400680     Error=0.00000000000000111022
N=13     Value=1.38038844704314311862     Error=0.00000000000000022204
N=14     Value=1.38038844704314334066     Error=0.00000000000000044409
N=15     Value=1.38038844704314267453     Error=-0.00000000000000022204
N=16     Value=1.38038844704314445089     Error=0.00000000000000155431
N=17     Value=1.38038844704314311862     Error=0.00000000000000022204
N=18     Value=1.38038844704314223044     Error=-0.00000000000000066613
N=19     Value=1.38038844704314356271     Error=0.00000000000000066613
N=20     Value=1.38038844704314178635     Error=-0.00000000000000111022
```

上机习题**3.11**

题目

求解积分方程$\int_0^1 (s^2+t^2)^{1/2}u(t)\mathrm{d}t = \frac{(s^2+1)^{3/2}-s^3}{3}$

（1）分别用$n=3,\cdots,15$个等距积分节点t_j，使用复合Simpson公式离散积分，s_i也取同样的点，使用列主元Gauss消去法求解离散所得的线性代数方程组$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}$。与已知的唯一解析解$u(t)=t$相比较，$n$为多大时求得的结果最好？解释原因。

（2）对上面的每个n，计算线性代数方程组的系数矩阵$\boldsymbol{A}$的条件数，考察条件数与n的关系。

解答

取等距积分节点$t_j = j/n\ (0 \le j \le n)$，步长$h = 1/n$，则利用Simpson公式可得方程左端近似的离散积分为

$$
\begin{aligned}
S(h) &= \frac{h}{6}\sum_{j=0}^{n-1}\left[(s^2+t_j^2)^{\frac{1}{2}}u(t_j) + 4(s^2+t_{j+\frac{1}{2}}^2)^{\frac{1}{2}}u(t_{j+\frac{1}{2}}) + (s^2+t_{j+1}^2)^{\frac{1}{2}}u(t_{j+1})\right] \\
&= \frac{h}{6}\left[(s^2+t_0^2)^{\frac{1}{2}}u(t_0) + 2\sum_{j=0}^{n-1}(s^2+t_j^2)^{\frac{1}{2}}u(t_j) + (s^2+t_n^2)^{\frac{1}{2}}u(t_n) + 4\sum_{j=0}^{n-1}(s^2+t_{j+\frac{1}{2}}^2)^{\frac{1}{2}}u(t_{j+\frac{1}{2}})\right].
\end{aligned}
$$

如果对s也取同样的节点$s_i = i/n\ (0 \le i \le n)$，则可将原方程近似为一个含$2n+1$个变元的线性方程组

$$
\frac{h}{6}\left[r_{i,0}u(t_0) + 2\sum_{j=0}^{n-1}r_{i,j}u(t_j) + r_{i,n}u(t_n) + 4\sum_{j=0}^{n-1}r_{i,j+\frac{1}{2}}u(t_{j+\frac{1}{2}})\right] = \frac{(s_i^2+1)^{3/2}-s_i^3}{3}, \quad (i = 0, \frac{1}{2}, \cdots, n).
$$

其中$r_{i,j} = (s_i^2+t_j^2)^{\frac{1}{2}} = \sqrt{i^2+j^2}/n$。下举例列出n=3时的这一方程组：

列出 n=3 时的技手：

令 $k(s,t) = (s^2+t^2)^{1/2}$，$f(s) = \dfrac{(s^2+1)^{3/2}-s^3}{3}$

$s=0$ 时：$\frac{1}{2}\left[\frac{1}{6}k(0,0)u(0) + \frac{2}{3}k(0,\frac{1}{4})u(\frac{1}{4}) + \frac{1}{3}k(0,\frac{1}{2})u(\frac{1}{2}) + \frac{2}{3}k(0,\frac{3}{4})u(\frac{3}{4}) + \frac{1}{6}k(0,1)u(1)\right] = f(0)$

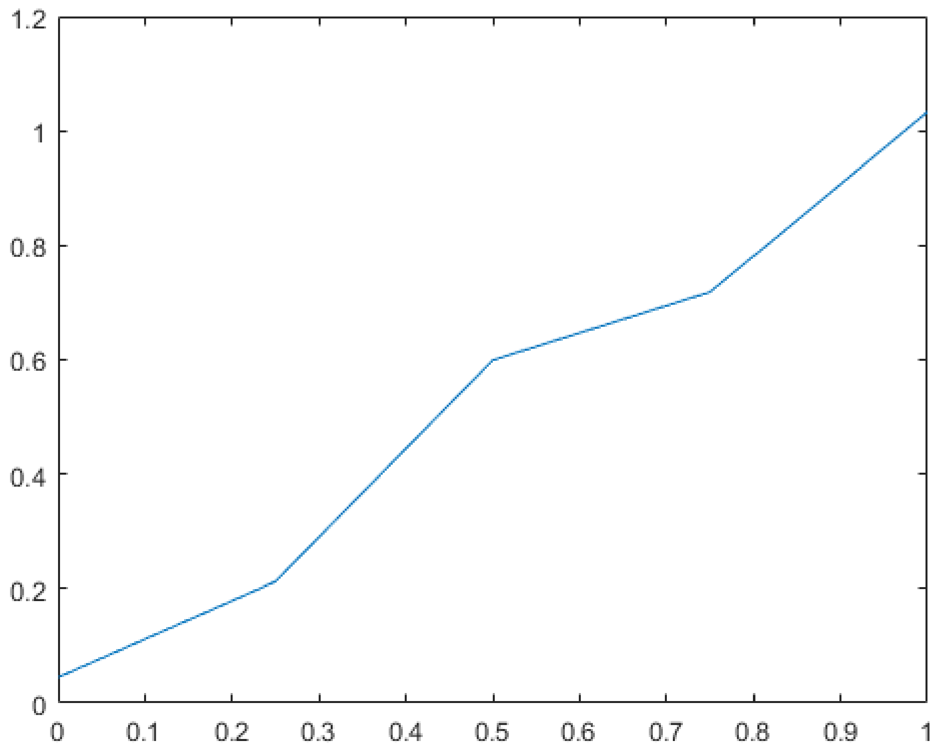
∵选用Simpson公式，∴实际变为了5个节点。

方程组形如：

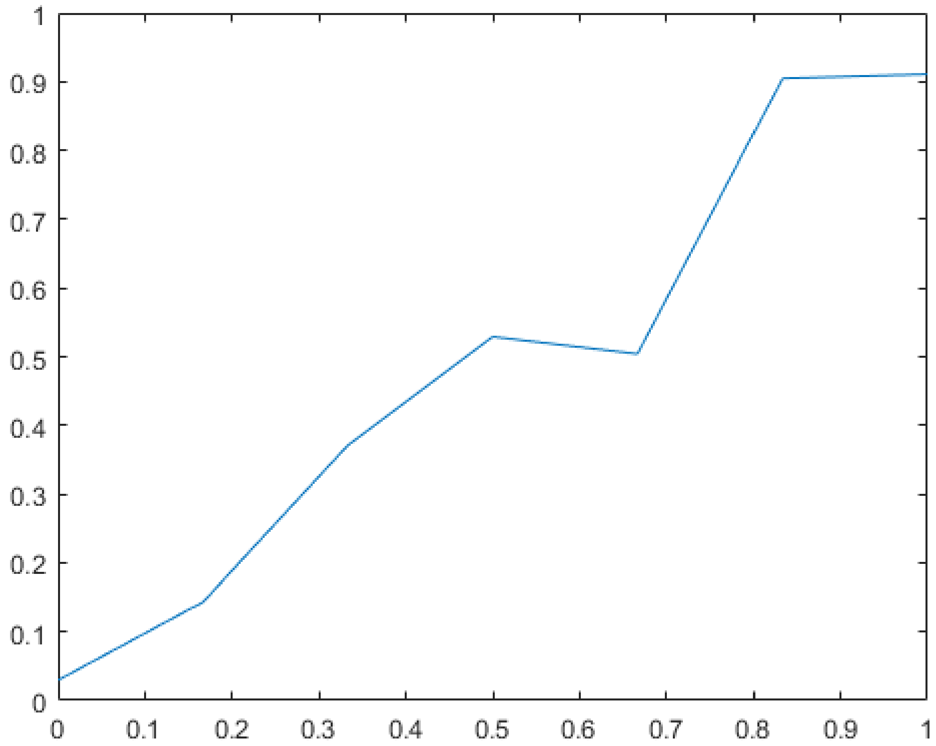
$$
\begin{bmatrix}
\frac{1}{12}k(0,0) & \frac{1}{3}k(0,\frac{1}{4}) & \frac{1}{6}k(0,\frac{1}{2}) & \frac{1}{3}k(0,\frac{3}{4}) & \frac{1}{12}k(0,1) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\frac{1}{12}k(1,0) & \frac{1}{3}k(1,\frac{1}{4}) & \frac{1}{6}k(1,\frac{1}{2}) & \frac{1}{3}k(1,\frac{3}{4}) & \frac{1}{12}k(1,1)
\end{bmatrix}
\begin{bmatrix} u(0) \\ u(\frac{1}{4}) \\ u(\frac{1}{2}) \\ u(\frac{3}{4}) \\ u(1) \end{bmatrix}
=
\begin{bmatrix} f(0) \\ f(\frac{1}{4}) \\ f(\frac{1}{2}) \\ f(\frac{3}{4}) \\ f(1) \end{bmatrix}
$$

解这些不同n时的方程组，得到结果如下：

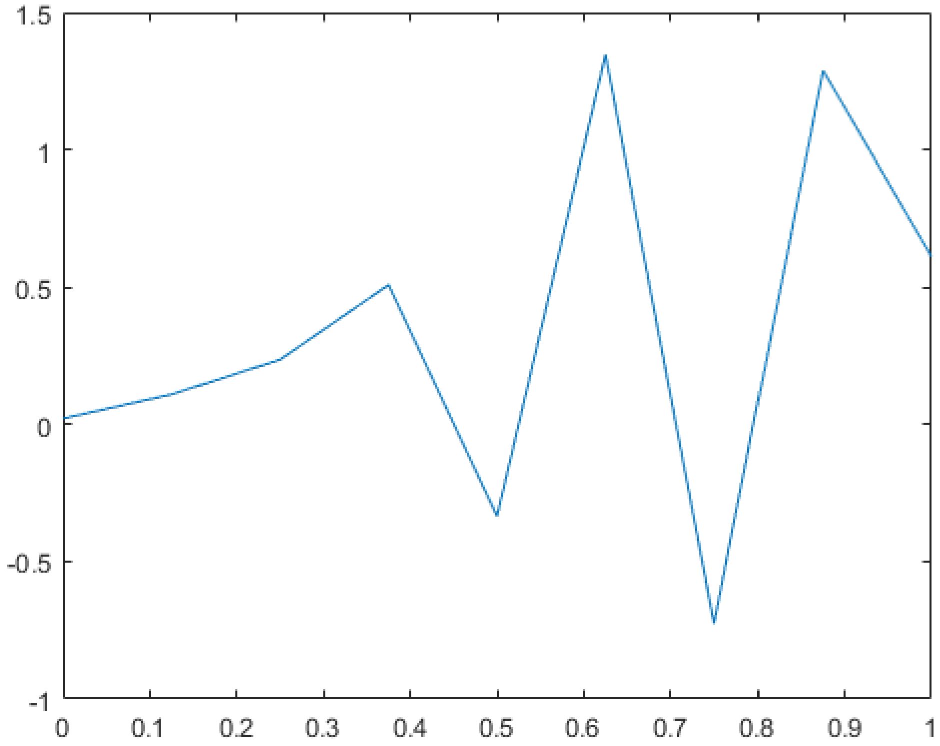
n=3：


1.2
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

n=4:


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

n=5:



- 经计算发现，n=3时求得的效果最好，猜测可能是因为n较大时矩阵的条件数过大，在n较大的情况下不具数值稳定性，影响了求解的准确度。计算矩阵的条件数如下：

  N=3 cond=1.8334e+04
  N=4 cond=3.3103e+06
  N=5 cond=5.3249e+08
  N=6 cond=8.1512e+10
  N=7 cond=1.2158e+13

- 可以发现，随着N的增长，条件数成指数增长，验证了我们之前的推测。