《数值分析》课程简介

徐岩

中国科学技术大学数学科学学院

yxu@ustc.edu.cn

http://staff.ustc.edu.cn/~yxu/

计算数学

- 现代的科学技术发展十分迅速，他们有一个共同的特点，就是都有大量的数据问题，例如求解大型线性方程组（1000000个未知量）。
- 计算问题可以是现代社会各个领域普遍存在的共同问题，研究计算问题的解决方法和有关数学理论问题的一门学科就叫做计算数学。
- 计算数学属于应用数学的范畴，它主要研究有关的数学和逻辑问题怎样由计算机加以有效解决。

计算数学(续)

- 科学计算的兴起是20世纪后半叶最重要的科技进步之一。为把信息和数据变成知识，从而探索科学未知，促进技术创新，加强国防建设，保障国家安全，计算将起不可替代的重要作用。
- 大规模计算提出的世界性难题已形成科学计算的学科前沿。求解由实际问题得到的复杂的偏微分方程不仅计算规模大，更由于非线性、多尺度、长时间、不适定、多区域、高病态等特点使计算格外困难。现有的算法远不能满足需求，这正是目前大规模科学计算必须解决的关键科学问题。

科大计算数学现状

- 教师：
  - 陈发来、邓建松、刘利刚、李新、童伟华、张举勇、陈仁杰：计算机辅助几何设计与计算机图形学
  - 张梦萍、徐岩、夏银华、段雅丽、张瑞、徐宽、蒋琰：大规模科学计算
  - 杨周旺：优化理论、计算机图形学
  - 陈先进：非线性偏微分方程不稳定多解的分析与计算
- 全校相关院系专业
  - 五系、七系：计算流体
  - 九系：计算机辅助设计
  - 十一系：高性能并行计算

目前的课程体系

- 本科阶段
  - 数值分析、数值代数、偏微分方程数值解(有限差分方法)、有限元方法、计算机图形学、小波分析
- 研究生阶段
  - 时间依赖问题的差分方法、高级有限元、非线性方程数值方法、计算流体力学、计算机辅助几何设计、样条函数与逼近论、多变量函数逼近论、计算代数几何、高级几何建模与图形学

研究内容

实际问题 $\Longrightarrow$ 数学模型 $\Longrightarrow$ 计算方法 $\Longrightarrow$ 计算机求解

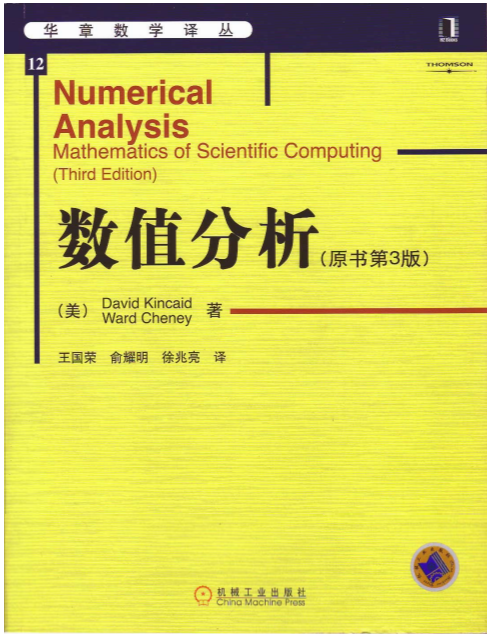
$\Downarrow$

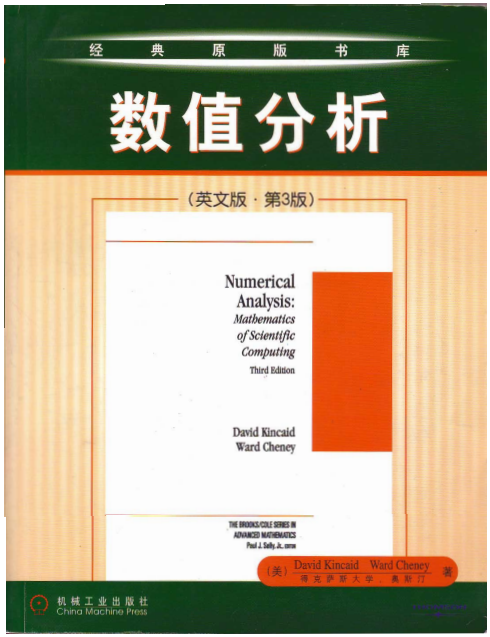
是数学的一个分支，它提出、发展、分析并应用科学计算中的方法于若干领域，如分析学、线性代数、几何学、逼近论、函数方程、优化问题和微分方程等。可以简单地认为计算方法是讨论如何求解微积分和线性代数中的计算问题。

- 其它领域，如物理学、自然和生物科学、工程、经济、金融科学也经常提出问题，而问题的解决同样需要科学计算
- 也称为科学计算，数值数学(Numerical Mathematics), 计算方法
- 主要研究对所设计的数值方法进行算法稳定性、精度和计算复杂性的分析

选用教材

David Kincaid, Ward Cheney, Numerical Analysis: Mathematics of Scientific Computing, Third Edition, Brooks/Cole, 2002. 机械工业出版社影印和翻译，价格均为￥75.00





参考读物

- S.D. Conte, C. De Boor, Elementary Numerical Analysis: An Algorithmic Approach, Mcgraw-Hill College, 1980.
- A. Quarteroni, etal., Numerical Mathematics, Springer, 2000。科学出版社影印
- S. T. Karris, Numerical Analysis : Using MATLAB and Spreadsheets (Second Edition), Orchard Publications, 2003。高等教育出版社影印
- E. Süli, etal., An Introduction to Numerical Analysis, Cambridge, 2003
- K. E. Atkinson, etal., Theoretical Numerical Analysis: A Functional Analysis Framework, Springer, 2001

讲义下载：`http://staff.ustc.edu.cn/~yxu/nmbook.zip`

讲授内容

最常用的数学模型的最基本的数值分析方法

- 函数插值与逼近 $\Longrightarrow$ 函数逼近论、样条函数
- 数值微分与积分
- 常微分方程数值解
- 非线性方程求解
- 数值代数 $\Longrightarrow$ 线性方程组解法（直接法、迭代法），矩阵的特征值和特征向量

考核方式

- 平时理论作业(15%)
  每周交一次作业
- 平时编程作业(15%)
  按指定时间通过Blackboard系统提交作业
- 课堂测验和期终理论考试(70%)

作业要求

- 请于每次编程作业布置后的第一个星期一晚上(23:00前)上传至Blackboard系统
- 书面作业每周一上课时交（无需上传Blackboard系统）
- 编程作业格式要求为：
  - 附件：只允许有一个附件，请把多个文件(源文件，头文件，说明文件)用winzip或winrar压缩到同一个文件中。请不要发送.exe文件（会被拒收）。
- 对于每个程序，请给出使用说明文件。说明文件类似于物理实验的实验报告，用word或者latex编写，内容包括：程序思路说明，编译命令或环境，使用说明，实验结果以及其它需要说明以便帮助助教判定的内容。
- 编程可以用任何语言：（C, C++, Matlab, Mathematica, Delphi, Fortran, Python等）不允许使用内置函数完成主要功能
- 结果输出要求小数点后至少12位。

《数值分析》误差简介

徐岩

中国科学技术大学数学科学学院

yxu@ustc.edu.cn

http://staff.ustc.edu.cn/~yxu/

内容

- 误差和有效数字
- 约束误差
- 一些例子

- x^*：精确值
- x：近似值

绝对误差

- 绝对误差 = 精确值−近似值 = $x^* - x$
- 绝对误差可正可负

相对误差

- 相对误差 = $\dfrac{\text{绝对误差}}{\text{精确值}} = \dfrac{x^*-x}{x^*}$
- 相对误差 = $\dfrac{\text{绝对误差}}{\text{近似值}} = \dfrac{x^*-x}{x}$

误差的来源

- 原始误差：模型误差（忽略次要因素，如空气阻力）和原始数据误差
  - 数学模型的误差
  - 物理模型的误差
- 舍入误差：计算误差，计算机仅能表示表示有限位数据引起

误差的来源

- 截断误差：方法误差，算法本身引起
  - $\ln(1+x)=\sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k}$，在实际计算中，常常使用有限项近似无穷项

$$\ln(1+x)=\sum_{k=1}^{N}(-1)^{k+1}\frac{x^k}{k}$$

被舍弃的余项

$$(-1)^k\frac{x^{k+1}}{(k+1)(1+\theta x)^{k+1}},0<\theta<1$$

即为截断误差。

误差的运算

- 加减

$$(x^* \pm y^*) - (x \pm y) = e_x \pm e_y$$

$$\frac{e_x \pm e_y}{x^* \pm y^*}$$ 两相近数相减，相对误差增大

- 相乘

$$\begin{aligned}(x^* \cdot y^*) - (x \cdot y) &= x^*(y * -y) + y(x^* - x) = ye_x + x^* e_y \\ &= \max\{|x^*|, |y|\}(|e_x| + |e_y|)\end{aligned}$$

- 相除

$$\left|\frac{x^*}{y^*} - \frac{x}{y}\right| = \left|\frac{x^*y - y^*x}{yy^*}\right| = \left|\frac{-x^*(y^* - y) + y(x^* - x)}{yy^*}\right|$$

$$= \left|\frac{-x^*e_y + ye_x}{yy^*}\right|$$ 小数作除数，绝对误差增大

有效位数

定义： 当x的误差为某一位的半个单位，则这一位到第一个非零的位数称为x的有效位数。

- 有效位的多少直接影响到近似值的绝对误差和相对误差
- 3.28和0.00587均有3位有效数字
- 4.0和4.000分别有2位和4位有效数字

误差控制

- 选择收敛稳定的数值方法
- 提高数值计算的精度：单精度、双精度
- 避免2个非常接近的数相减
- 多个数求和时，从小数加到大数

例

在计算机中求函数$f(x) = 1 - \cos(x)$在x接近0点的值。
$\cos(x) \simeq 1$当x接近0点时，此时如果用此公式直接计算，很容易在0点附近失去精度。如果采用如下公式，

$$1 - \cos(x) = \frac{(1 - \cos(x))(1 + \cos(x))}{1 + \cos(x)} = \frac{\sin^2(x)}{1 + \cos(x)}$$

则可以避免这样的问题

例

在计算机中求函数$f(x)=\sqrt{x+1}-\sqrt{x}$在x取比较大的值.

$$\begin{aligned} f(12345) &= \sqrt{12346}-\sqrt{12345} \\ &= 111.113-111.108 \\ &= 0.005 \end{aligned}$$

但实际上，$f(12345)=0.00450003262627751$.

例

$$s_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$$

应该用下面的方式来计算

$$s_n = \frac{1}{n} + \cdots + \frac{1}{2} + 1$$

例

$$A_n = \int_0^1 \frac{x^n}{x+5} dx$$

我们有

$$A_n + 5A_{n-1} = \int_0^1 \frac{x^n + 5x^{n-1}}{x+5} dx = \int_0^1 x^{n-1} dx = \frac{1}{n}$$

构造方法如下

1. $A_n = \frac{1}{n} - 5A_{n-1}$, $A_0 = ln\frac{6}{5}$,记作$\hat{A}_n$
2. $A_{n-1} = \frac{1}{5}(\frac{1}{n} - A_n)$, $A_8 = 0.019$,记作$\tilde{A}_n$

一些例子

n	A_n	$\hat{A}_n$	$\tilde{A}_n$
0	0.182	0.182	0.182
1	0.088	0.090	0.088
2	0.058	0.050	0.058
3	0.0431	0.083	0.0431
4	0.0343	-0.165	0.0343
5	0.0284	1.025	0.0284
6	0.024	-4.958	0.024
7	0.021	24.933	0.021

- 对格式1，如果前一步有误差，则被放大5倍加到后一步
- 这样的格式称为不稳定格式。

上机作业

级数计算[Hamming (1962)]

$$\varphi(x) = \sum_{k=1}^{\infty} \frac{1}{k(k+x)}$$

x取值$x = 0.0, 0.1, 0.2, \cdots, 1.0, 10.0, 20.0, \cdots, 300.0$共41个值，要求误差小于$10^{-6}$，并给出相应的$k$的取值上界。

输出格式：三列：

x, $\varphi(x)$, k

《数值分析》之

非线性方程求根

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

非线性方程求根

非线性科学是当今科学发展的一个重要研究方向，在许多应用问题中能发现非线性方程的例子。

- 在光的衍射理论中，我们需要方程

$$x - \tan x = 0$$

  的根。

- 在行星轨道的计算中，我们需要开普勒方程

$$x - a \sin x = b$$

  的根，其中a和b取任意值。

非线性方程的求根也成了一个不可缺少的内容。

通常非线性方程的根的情况非常复杂：

- 解不唯一

$$\begin{cases} \sin(\frac{\pi}{2}x) = y \\ y = \frac{1}{2} \end{cases}$$

  有无穷组解。

- 只在某个区域内可能解存在唯一，而且经常很简单的形式得不到精确解。例如：

$$e^x - \cos(\pi x) = 0.$$

- 当用计算机求函数的近似零点时，即使精确解是唯一的，还是会出现许多近似解。

根据不同的需要，非线性方程求根问题可以提出以下三类不同的要求：

- 已知某根近似，要求把根精确化。
- 确定全部的根，或者给定区域上的全部根。
- 判定给定区域上方程根的个数。

主要方法

- 对分法
- 不动点方法
- 牛顿法
- 割线法

对分法

定理

若f是区间$[a,b]$上的连续函数，且$f(a)f(b)<0$，则f在(a,b)内必有一个零点，即$\exists x\in(a,b),\ f(x)=0$.

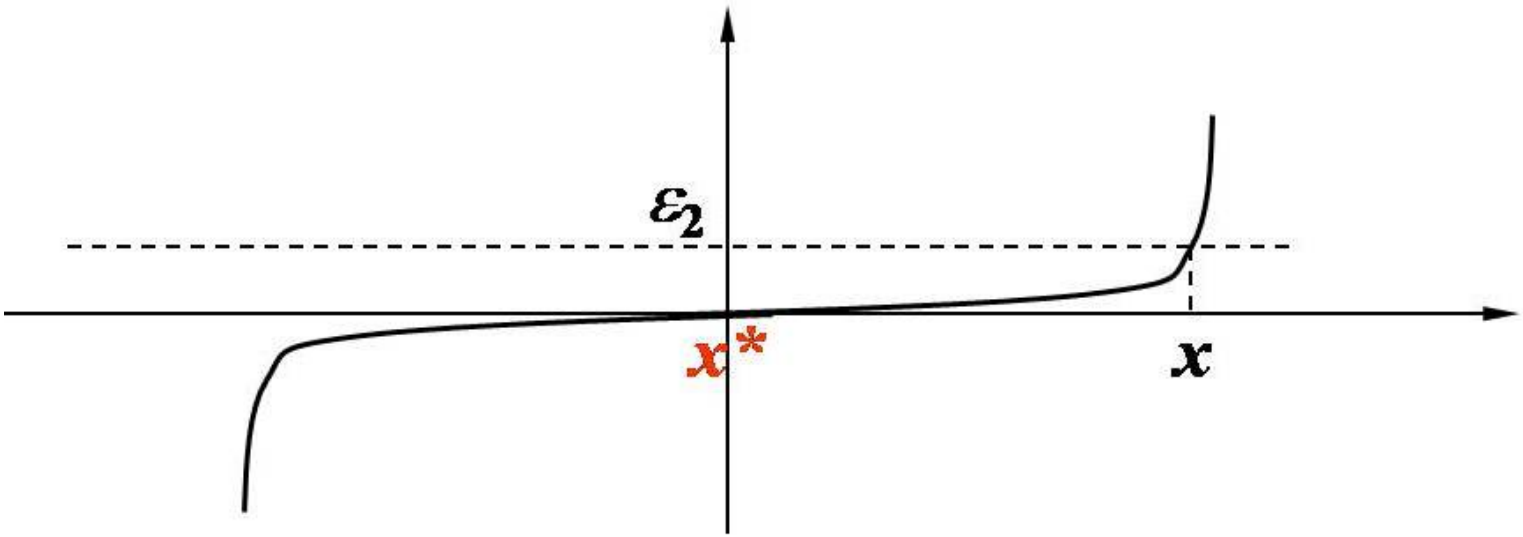
算法

1. $[a_0,b_0]=[a,b]$, $x=\frac{a_0+b_0}{2}$
2. if $f(x)\geq 0$, then $a_{n+1}=a_n$, $b_{n+1}=x$
3. else $a_{n+1}=x$, $b_{n+1}=b_n$

```
While(|a-b|>eps)
        x=(a+b)/2
       f(x)
      若(|f(x)|<eps) x为解
      若f(x)*f(b)<0 修正区间为[x,b]
      若f(a)*f(x)<0 修正区间为[a,x]
End while
```

对分法

- 每次缩小一倍的区间，收敛速度为1/2，较慢
- 只能求一个根，使用条件限制较大
- 不能保证x的精度


ε_2
x^*
x

$$f(x)=0 \Longleftrightarrow x=\varphi(x)$$
$$f(x)\text{的根} \Longleftrightarrow \varphi(x)\text{的不动点}$$

一个数学问题常常能归结成为一个求函数不动点的问题。

思路

从一个初值x^0出发，计算

$$x^1=\varphi(x^0), x^2=\varphi(x^1), \cdots, x^{k+1}=\varphi(x^k), \cdots,$$

若$\{x^k\}_0^\infty$收敛，即存在x^*使得$\lim\limits_{k\to\infty} x^k=x^*$，且$\varphi$连续，则由

$$\lim_{k\to\infty} x^{k+1}=\lim_{k\to\infty}\varphi(x^*)$$

可知$x^*=\varphi(x^*)$，即x^*是φ的不动点，也就是f的根。

不动点方法基本步骤

- 给出方程的局部等价形式, $f(x) = 0 \Longleftrightarrow x = \varphi(x)$.
- 取合适的初值x^0，产生迭代序列$x^{k+1} = \varphi(x^k)$.
- 求极限 $\lim\limits_{k\to\infty} x^k = x^*$, 该值为方程的根.

问题

迭代序列 $\lim\limits_{k\to\infty} x^k$是否收敛?

不动点定理

$\varphi(x)$是定义在$[a,b]$上的函数，若$\varphi(x)$满足

- $\forall x \in [a,b],\ \varphi(x) \in [a,b]$.
- $\varphi(x)$在$[a,b]$上可导，且存在正数$L<1$, $\forall x \in [a,b]$, 有$|\varphi'(x)| \leq L$.

则有，

1. $\exists 1\ x^*$, s.t. $x^* = \varphi(x^*)$, 称x^*为$\varphi(x)$的不动点。
2. 迭代格式$x^{k+1} = \varphi(x^k)$对任意的初值$x^0 \in [a,b]$均收敛于$\varphi(x)$的不动点x^*。
3. 误差估计式为

$$|x^* - x^k| \leq \frac{L^k}{1-L}|x^1 - x^0|$$

证明

- 不动点的存在唯一性

做辅助函数$\psi(x)=x-\varphi(x)$,则有$\psi(a)\leq 0,\psi(b)\geq 0$，则

$$\exists x^*,\ s.t.\ \psi(x^*)=0,\ i.e. x^*=\varphi(x^*).$$

若$x^{**}=\varphi(x^{**})$，则有

$$|x^*-x^{**}|=|\varphi(x^*)-\varphi(x^{**})|=|\varphi'(\xi)||x^*-x^{**}|\leq L|x^*-x^{**}|,\xi\in[a,b]$$

由$L<1$可知$x^*=x^{**}$.

- 迭代格式收敛

当$x_0 \in [a,b]$时，可用数学归纳法证明，迭代序列$\{x_k\} \subseteq [a,b]$，于是由微分中值定理

$$\begin{aligned} x^{k+1} - x^* &= \varphi(x^k) - \varphi(x^*) = \varphi'(\xi)(x^k - x^*), \xi \in [a,b] \\ |x^{k+1} - x^*| &\le L|x^k - x^*| = L|\varphi(x^{k-1}) - x^*| \\ &\le L^2|x^{k-1} - x^*| \le \cdots \le L^{k+1}|x^0 - x^*| \end{aligned}$$

因为$L < 1$,因此有当$k \to \infty$时，$L^{k+1} \to 0$,则

$$x_{k+1} \to x^*.$$

即迭代格式$x^{k+1} = \varphi(x^k)$收敛。

- 误差估计

$$|x^{k+1}-x^k| = |\varphi(x^k)-\varphi(x^{k-1})| \le L|x^k-x^{k-1}| \le \cdots L^k|x^1-x^0|$$

设k固定，对任意正整数p，有

$$\begin{aligned}|x^{k+p}-x^k| &\le |x^{k+p}-x^{k+p-1}|+\cdots+|x^{k+1}-x^k| \\ &\le (L^{k+p-1}+L^{k+p-2}+\cdots+L^k)|x^1-x^0| \\ &= \frac{L^k}{1-L}|x^1-x^0|\end{aligned}$$

由p的任意性，$\lim\limits_{p\to+\infty} x^{k+p}=x^*$，故有

$$|x^*-x^k| \le \frac{L^k}{1-L}|x^1-x^0|$$

注

构造满足定理条件的等价形式一般难于做到。要构造收敛迭代格式有两个要素：

- 等价形式
- 初值选取

例

代数方程$x^3 - 2x - 5 = 0$。

1. $x = \sqrt[3]{2x+5}$, $\varphi'(x) = \frac{1}{3}\frac{1}{(2x+5)^{\frac{2}{3}}}$, 迭代格式$x^{k+1} = \sqrt[3]{2x^k + 5}$
2. $x = \frac{x^3-5}{2}$, $\varphi'(x) = \frac{3x^2}{2}$, 迭代格式$x^{k+1} = \frac{(x^k)^3-5}{2}$
3. $x = \frac{2x+5}{x^2}$, $\varphi'(x) = -\frac{2(5+x)}{x^3}$, 迭代格式$x^{k+1} = \frac{2x^k+5}{x_k^2}$

计算下列函数的不动点

$$f(x) = 4 + \frac{1}{3}\sin(2x)$$

由中值定理，我们有

$$|f(x)-f(y)| = \frac{1}{3}|\sin(2x)-\sin(2y)| = \frac{2}{3}|\cos(2\xi)||x-y|, \quad \xi \in (a,b)$$

取初值$x = 4$,执行20次迭代.

x=4, M=20

for k=1,M

$x = 4 + \frac{1}{3}\sin(2x)$

1	4.3297861
2	4.2308951
3	4.2736338
⋮	⋮
14	4.2614830
15	4.2614840
16	4.2614836
⋮	⋮
20	4.2614837

Newton迭代法

将$f(x)$在初值处作Taylor展开

$$f(x) = f(x^0) + f'(x^0)(x - x^0) + \frac{f''(x^0)}{2}(x - x^0)^2 + \cdots$$

当x与x^0很接近，取线性部分作为$f(x)$的近似，有

$$f(x^0) + f'(x^0)(x - x^0) \approx 0$$

若$f'(x^0) \neq 0$,则有

$$x = x^0 - \frac{f(x^0)}{f'(x^0)}$$

可以归纳定义迭代格式为

$$x^{n+1} = x^n - \frac{f(x^n)}{f'(x^n)}$$

Newton迭代法

Newton迭代的等价方程为：

$$f(x)=0 \Longleftrightarrow x=\varphi(x)=x-\frac{f(x)}{f'(x)}$$

因此有

$$\varphi'(x)=\left(x-\frac{f(x)}{f'(x)}\right)'=\frac{f(x)f''(x)}{(f'(x))^2}$$

- 若$f(x)$在$x=a$处为单根，则

$$f(a)=0,\ f'(a)\neq 0$$

故有$\varphi'(a)=0$. 所以，迭代格式收敛.

收敛速度

$$
\begin{aligned}
x^{n+1} - a &= \varphi(x^n) - \varphi(a) \\
&= (x^n - a)\varphi'(a) + \frac{(x^n - a)^2}{2}\varphi''(\xi_n) \\
&= \frac{(x^n - a)^2}{2}\varphi''(\xi_n) \\
&\approx \frac{(x^n - a)^2}{2}\varphi''(a)
\end{aligned}
$$

Newton迭代是二阶迭代方法。

例

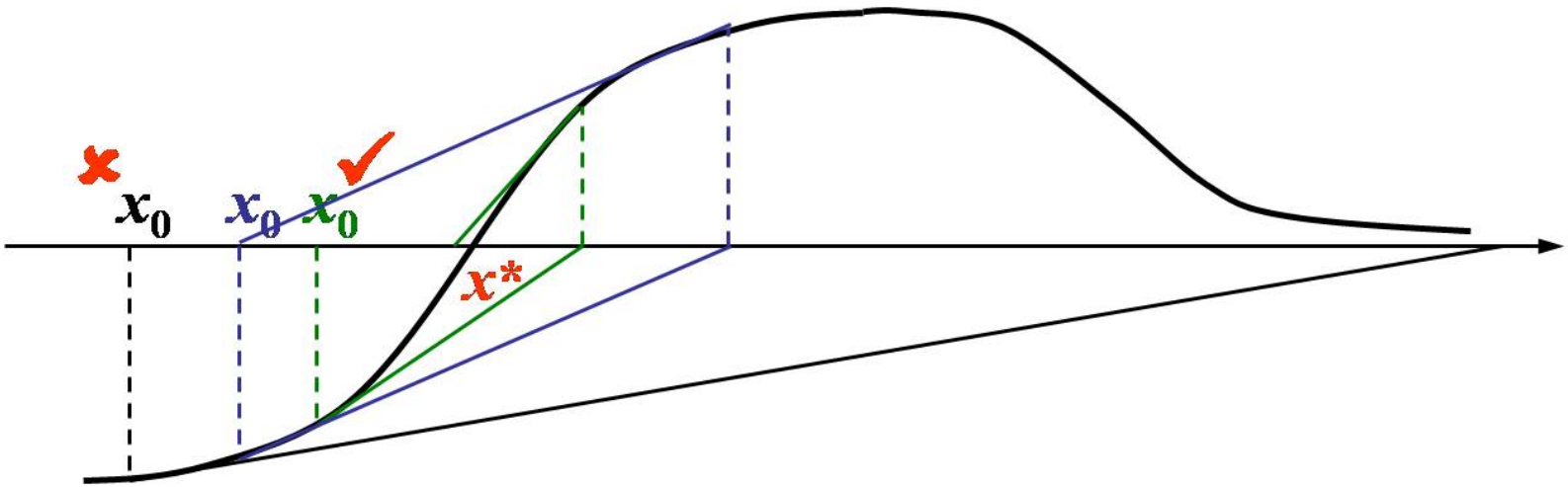
用Newton迭代法求方程$e^x - 1.5 - \arctan x = 0$的负零点。取初值$x^0 = -7$.

解：$f'(x) = e^x - \frac{1}{x^2+1}$, Newton迭代格式为

$$x^{n+1} = x^n - \frac{e^{x^n} - 1.5 - \arctan(x^n)}{e^{x^n} - \frac{1}{(x^n)^2+1}}$$

n	x	$f(x)$
0	-7.0	-0.702×10^{-1}
1	-10.677096176640013992969843 86	-0.226×10^{-1}
2	-13.279167375632712908597863 19	-0.437×10^{-2}
3	-14.053655854269238734748317 53	-0.239×10^{-3}
4	-14.101109956866413476163127 06	-0.800×10^{-6}
5	-14.101269770939415946215795 06	-0.901×10^{-11}
6	-14.101269772739968425083003 14	-0.114×10^{-20}
7	-14.101269772739968425311551 22	0.000

- Newton迭代格式的一般仅应用于求解方程的实系数方程的实根
- Newton迭代格式的收敛速度快，格式简单，应用广泛
- Newton迭代格式的收敛性依赖于初值x^0的选取。



弦截法

- Newton法的一个缺点是它需要求零点的函数导数。
- 将Newton迭代中的导数，用差商代替

$$f'(x^n) = \frac{f(x^n) - f(x^{n-1})}{x^n - x^{n-1}}$$

- 由此得到弦截法

$$x^{n+1} = x^n - f(x^n)\frac{x^n - x^{n-1}}{f(x^n) - f(x^{n-1})}$$

- 弦截法是2步格式。收敛速度比Newton迭代慢，收敛阶为$1 + \frac{\sqrt{5}}{2} \approx 1.618$。

非线性方程组

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0 \\ f_2(x_1, x_2, \cdots, x_n) = 0 \\ \cdots\cdots \\ f_n(x_1, x_2, \cdots, x_n) = 0 \end{cases}$$

写成向量形式

$$F(\boldsymbol{x}) = 0$$

这里，$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$, $F(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_n(\boldsymbol{x}))^T$.

非线性方程组的Newton方法

直接推广Newton迭代为：

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k - (J(\boldsymbol{x}^k))^{-1} F(\boldsymbol{x}^k)$$

这里

$$J(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

实际中，用解方程组的形式

$$J(\boldsymbol{x}^k)(\boldsymbol{x}^{k+1} - \boldsymbol{x}^k) = F(\boldsymbol{x}_k).$$

在$\boldsymbol{x}$的邻域中，若$\|J(\boldsymbol{x})\|_\infty \leq L < 1$，而初始值充分接近于解，则迭代收敛。

上机作业

- 分别编写用Newton迭代和弦截法求根的通用程序
- 用如上程序计算下述函数的根
  1. $f(x) = 2x^4 + 24x^3 + 61x^2 - 16x + 1$
  2. $f(x) = \tan x - x$
  3. $f(x) = 1 + x^2 + e^x \cos x$

  分别取初值x_0为0.1,0.2,0.9,9.0.
- 输出形式如下：

x_0	迭代次数	数值结果	数值误差

- 简单分析你得到的数据

《数值分析》之

函数逼近

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

概念

- 实际中，函数$f(x)$多样，复杂，通常只能观测到一些离散数据；或者函数$f(x)$过于复杂而难以运算。这时我们要用近似函数$\varphi(x)$来逼近$f(x)$。
- 函数逼近：在科学计算中用到大量复杂函数，对它们直接进行微分、积分等计算不是很容易，而插值是一种最简单的逼近方法。
- 在计算中函数通常是用离散点表示的，特别是函数来自于微分方程的数值解。插值是复原函数的直接方法。
- 在计算机图形学中需要处理由大量离散点表示的曲线和曲面，这称为曲线和曲面拟合(fitting)，其核心步骤就是函数插值。

发展概况

- 单变量函数插值：从十八世纪开始，单变量多项式插值就得到了系统发展，Lagrange, Hermite, Newton等人都有重要贡献。至今已相当成熟，出现了多项式插值、三角函数插值、样条插值，并且催生了如样条函数理论等在现代大规模科学计算和计算机辅助设计中扮演主要角色的理论体系
- 多变量函数插值：至今还处于发展阶段。比较成熟的是张量积形式的插值理论，而其它形式的插值以及多变量样条理论还不是很完善

定义

$f(x)$ 为定义在区间$[a,b]$上的函数，$x_0, x_1, \cdots, x_n$为区间上n+1个互不相同的点，Φ为给定的某一函数类。求Φ上的函数$\varphi(x)$,满足

$$\varphi(x_i) = f(x_i), i = 0, 1, \cdots, n$$

则称$\varphi(x)$为$f(x)$关于节点$x_0, x_1, \cdots, x_n$在Φ上的插值函数。
称$x_0, x_1, \cdots, x_n$为插值节点，称$(x_i, f(x_i))$为插值点。

- 插值函数是否存在唯一?
- 如何构造插值函数?
- 插值函数的误差如何估计?

代数多项式

- 易于运算
- 无穷光滑
- 易于计算积分和导数
- 通常选择代数多项式作为插值基函数

Theorem

若x_i两两不同，则对任意给定的y_i，存在唯一的次数至多是n次的多项式p_n使得$p_n(x_i) = y_i,\ i = 0, \ldots, n.$

证明：在幂基$\{1, x, \ldots, x^n\}$下待定多项式p的形式为

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

由插值条件$p(x_i) = y_i,\ i = 0, \ldots, n,$ 得到如下方程组

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

系数矩阵为Vandermonde矩阵，其行列式非零，因此方程组有唯一解。

多项式插值定理（续）

- 插值矩阵方法(待定系数法)：不具有实用价值，但具有很高的理论价值
- 注意Vandermonde矩阵是病态的，因此不推荐应用该方法计算插值多项式的形式。
- 对于给定的问题，插值多项式存在唯一。但是可以用不同的方法给出插值多项式的不同表示形式。

不同形式的插值多项式

- Lagrange插值
- Newton插值
- Hermite插值

Lagrange插值

- Lagrange基函数: 由多项式插值定理存在函数$\ell_i(x)$满足$\ell_i(x_j) = \delta_{ij}$. 实际上,

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

- Lagrange插值多项式:

$$L_n(x) = \sum_{k=0}^{n} y_k \ell_k(x)$$

- 如果插值节点相同，给定的数据点$\{y_i\}$有多组，那么采用Lagrange插值是相当有效的。但求值算法不是很直接。
- Lagrange插值的缺点: 无承袭性。增加一个节点，所有的基函数都要重新计算

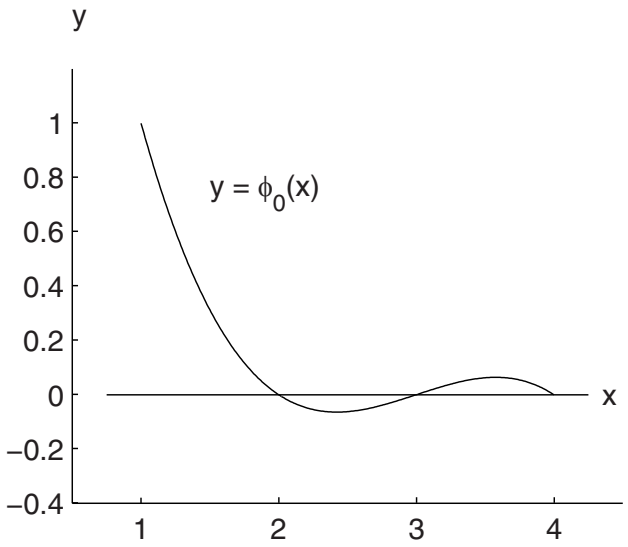
线性插值

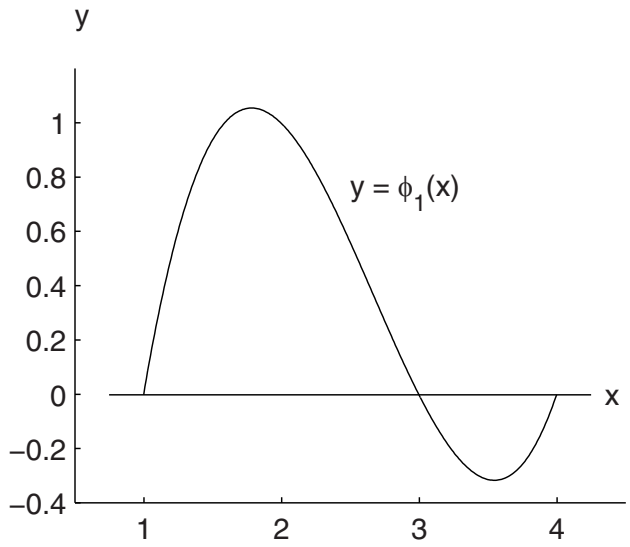
$$\ell_0(x) = \frac{x - x_1}{x_0 - x_1}, \quad \ell_1(x) = \frac{x - x_0}{x_1 - x_0},$$
$$L_1(x) = f(x_0)\ell_0(x) + f(x_1)\ell_1(x)$$

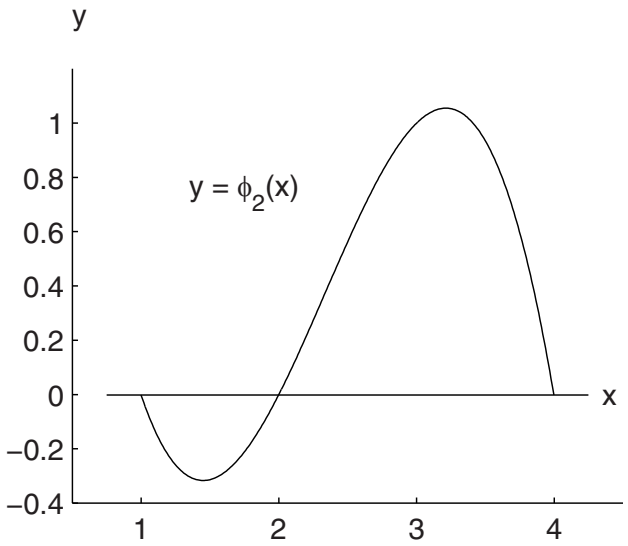
二次插值

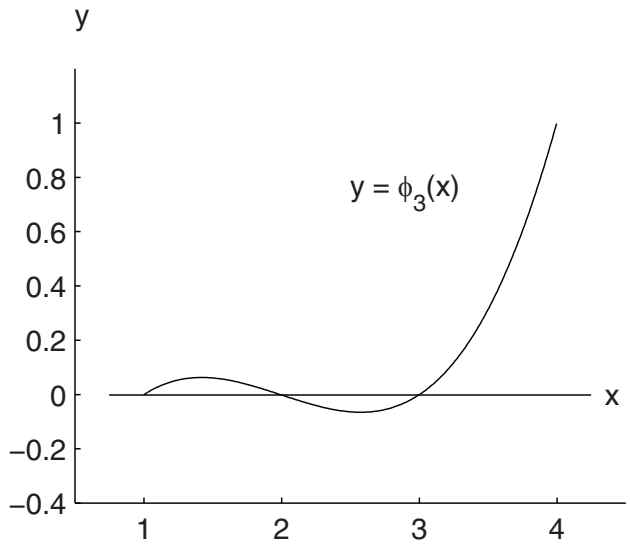
$$\ell_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)},$$
$$\ell_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)},$$
$$\ell_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)},$$
$$L_2(x) = f(x_0)\ell_0(x) + f(x_1)\ell_1(x) + f(x_2)\ell_2(x)$$

三次插值基函数


y
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
y = ϕ0(x)
x
1
2
3
4


y
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
y = ϕ1(x)
x
1
2
3
4


y
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
y = ϕ2(x)
x
1
2
3
4


y
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
y = ϕ3(x)
x
1
2
3
4

x	5	-7	-6	0
y	1	-23	-54	-954

结点为$5,-7,-6,0$，所以基函数是

$$\ell_0(x)=\frac{(x+7)(x+6)x}{(5+7)(5+6)5}=\frac{1}{660}x(x+6)(x+7)$$

$$\ell_1(x)=\frac{(x-5)(x+6)x}{(-7-5)(-7+6)(-7)}=-\frac{1}{84}x(x-5)(x+6)$$

$$\ell_2(x)=\frac{(x-5)(x+7)x}{(-6-5)(-6+7)(-6)}=-\frac{1}{66}x(x-5)(x+7)$$

$$\ell_3(x)=\frac{(x-5)(x+7)(x+6)}{(0-5)(0+7)(0+6)}=-\frac{1}{210}(x-5)(x+6)(x+7)$$

所以Lagrange型插值多项式
为$L_3(x)=\ell_0(x)-23\ell_1(x)-54\ell_2(x)-954\ell_3(x)$

算法

- 计算Lagrange基函数

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

```
fx=0.0
for(i=0;i<=n;i++) {
    tmp=1.0;
    for(j=0;j<i;j++)
        tmp=tmp*(x-x[j])/(x[i]-x[j]);
    for(j=i+1;j<=n;j++)
        tmp=tmp*(x-x[j])/(x[i]-x[j]);
    fx=fx+tmp*y[i];
}
return fx;
```

多项式插值误差定理

Theorem

设$f \in C^{n+1}[a,b]$，多项式p是f在不同结点$x_0, x_1, \ldots, x_n$上的插值多项式，$\deg p \leqslant n$。则对$[a,b]$中每个x，都有$\xi_x \in (a,b)$使得

$$f(x) - p(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i)$$

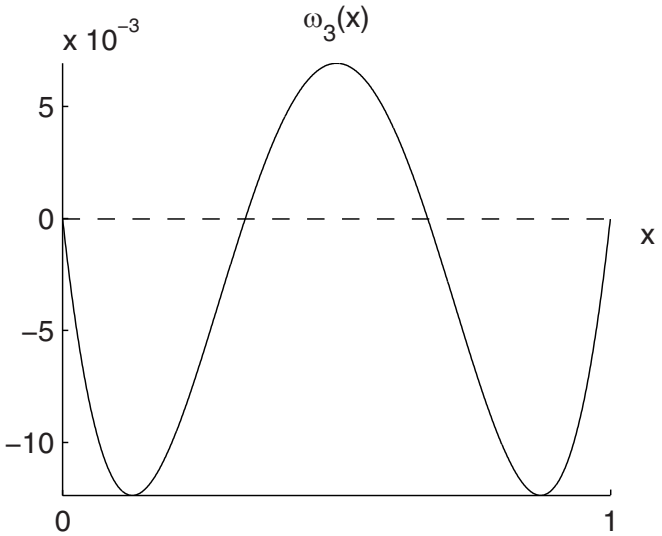
证明：当x与某个结点重合时结论成立。对其它情形，固定x, 令

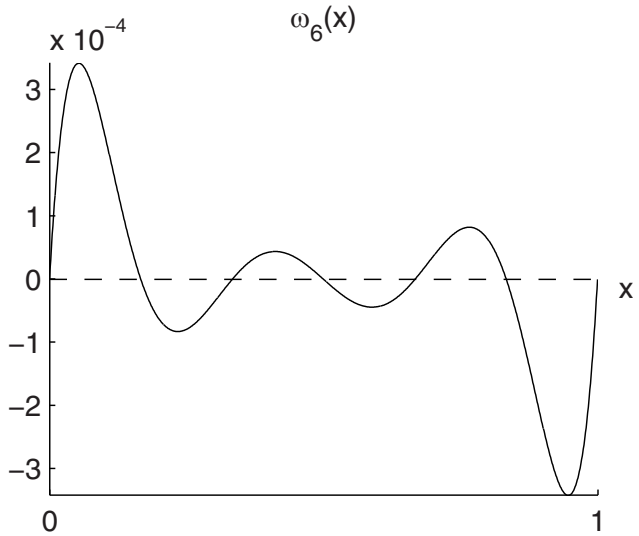
$$\phi(t) = f(t) - p(t) - \frac{f(x) - p(x)}{w(x)} w(t), w(t) = \prod_{i=0}^{n} (t - x_i)$$

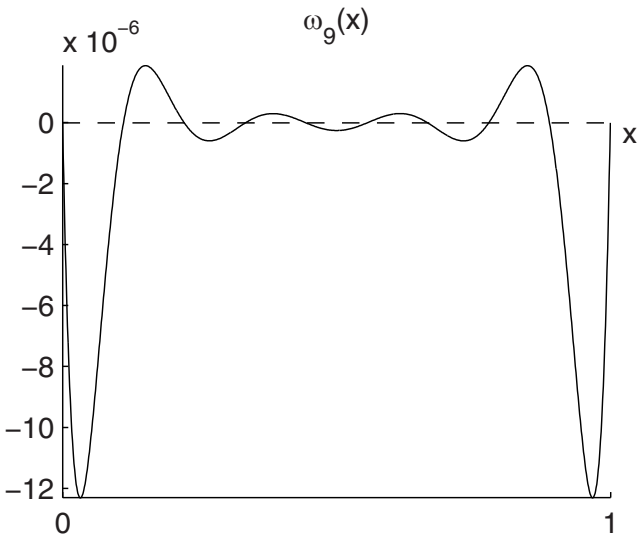
则$\phi(t)$在$[a,b]$内有$n+2$个零点，从而存在$\xi_x \in (a,b)$使得$\phi^{(n+1)}(\xi_x) = 0$，即

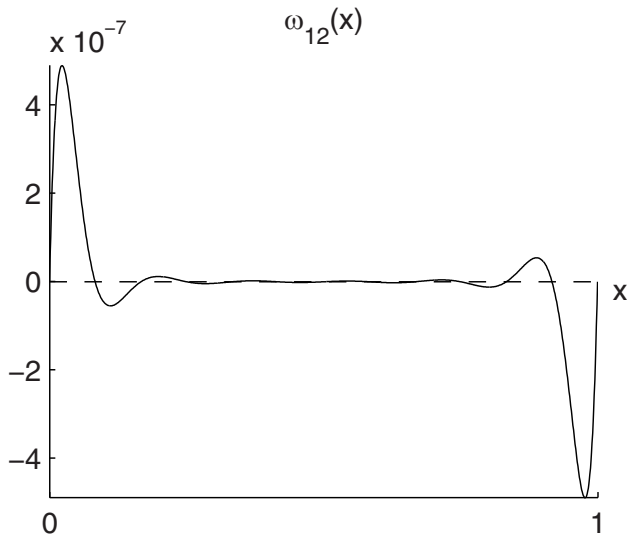
$$0 = f^{(n+1)}(\xi_x) - (n+1)! \frac{f(x) - p(x)}{w(x)}.$$

多项式插值误差定理$w(x) = \prod_{i=0}^{n}(x - x_i)$









误差估计的应用

- 可以应用于基于多项式构造的各种近似算法(如数值积分)的误差分析
- 建立插商与导数之间的关系：在结点确定的区间内存在一点ξ使得

$$f[x_0, \ldots, x_n] = \frac{f^{(n)}(\xi)}{n!}$$

Lagrange插值多项式性质

当$f(x)=x^k, k=0,1,\cdots,n$关于节点$x_0,x_1,\ldots,x_n$上的Lagrange插值多项式就是其本身，因此Lagrange基函数$\ell_i(x)$满足

$$L_n(x)=\sum_{i=0}^{n}\ell_i(x)x_i^k=x^k, k=0,1,\cdots,n$$

令$k=0$,得到

$$\sum_{i=0}^{n}\ell_i(x)=1$$

多项式插值事后误差估计

- 给定$x_0, x_1, \cdots, x_{n+1}$
- 取$x_0, x_1, \cdots, x_n$，构造$L_n(x)$
- 取$x_1, x_2, \cdots, x_{n+1}$，构造$\widetilde{L}_n(x)$

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi_1)}{(n+1)!}(x - x_0)(x - x_1)\cdots(x - x_n)$$

$$f(x) - \widetilde{L}_n(x) = \frac{f^{(n+1)}(\xi_2)}{(n+1)!}(x - x_1)(x - x_2)\cdots(x - x_{n+1})$$

- 近似$f^{(n+1)}(\xi_1) \approx f^{(n+1)}(\xi_2)$,有

$$\frac{f(x) - L_n(x)}{f(x) - \widetilde{L}_n(x)} \approx \frac{x - x_0}{x - x_{n+1}}$$

$$\Longrightarrow f(x) - L_n(x) \approx \frac{x - x_0}{x_0 - x_{n+1}}(L_n(x) - \widetilde{L}_n(x))$$

- 插值误差可以用2组插值函数的差来估计

上机作业1

对函数

$$f(x) = \frac{1}{1+x^2}, x \in [-5, 5]$$

构造Lagrange插值多项式$p_L(x)$，插值节点取为:

1. $x_i = 5 - \frac{10}{N} i,\ i = 0, 1, \cdots, N$
2. $x_i = -5\cos(\frac{2i+1}{2N+2}\pi),\ i = 0, 1, \cdots, N$ (Chebyshev point)

并计算如下误差

$$\max_i \{|f(y_i) - p(y_i)|,\ y_i = \frac{i}{10} - 5,\ i = 0, 1, \cdots, 100\}$$

对$N = 5, 10, 20, 40$比较以上两组节点的结果，并在一张图中画出$N = 10$时$f(x)$数值计算结果。

输出形式如下：

N=5

Max Error of grid (1) : XXXXXXXXXXXXXXXXXX

Max Error of grid (2) : XXXXXXXXXXXXXXXXXX

N=10

Max Error of grid (1) : XXXXXXXXXXXXXXXXXX

Max Error of grid (2) : XXXXXXXXXXXXXXXXXX

N=20

Max Error of grid (1) : XXXXXXXXXXXXXXXXXX

Max Error of grid (2) : XXXXXXXXXXXXXXXXXX

N=40

Max Error of grid (1) : XXXXXXXXXXXXXXXXXX

Max Error of grid (2) : XXXXXXXXXXXXXXXXXX

算法

- 计算Lagrange基函数

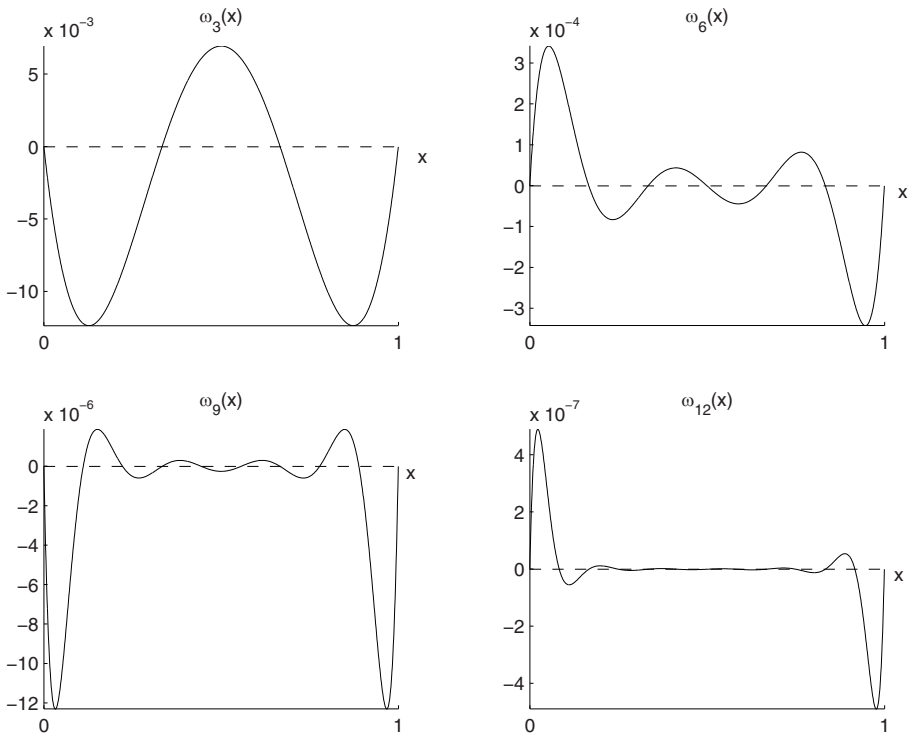
$$\ell_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j}$$

```
fx=0.0
for(i=0;i<=n;i++) {
    tmp=1.0;
    for(j=0;j<i;j++)
        tmp=tmp*(x-x[j])/(x[i]-x[j]);
    for(j=i+1;j<=n;j++)
        tmp=tmp*(x-x[j])/(x[i]-x[j]);
    fx=fx+tmp*y[i];
}
return fx;
```

多项式插值误差定理

设$f \in C^{n+1}[a,b]$，多项式p是f在不同结点$x_0, x_1, \ldots, x_n$上的插值多项式，$\deg p \leqslant n$。则对$[a,b]$中每个x，都有$\xi_x \in (a,b)$使得

$$f(x) - p(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i)$$



最佳结点的选取

- 即如何选取结点x_i, 使得$w(x) = (x - x_0)\dots(x - x_n)$ 在$[a, b]$上的绝对值最大值最小?
- 为了简单起见，不妨令$[a, b] = [-1, 1]$. 转而考虑一般的首一n次多项式$p(x)$使得它在$[-1, 1]$上的绝对值最大值最小。
- 需要用到(第一类)Tchebyshev多项式*。

*Tchebyshev (1821.5.16–1894.12.8), 俄罗斯数学家。1850年证明了Bertrand猜测，即n与$2n$之间必有至少一个素数，也接近证明了素数定理。同时他在概率论、正交函数和积分理论方面有重要贡献。其英文名有时也写作Chebyshev

(第一类)Tchebyshev多项式

有两种等价的定义方式

- 递归定义：

$$T_0(x) = 1, \quad T_1(x) = x,$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geqslant 1.$$

- 解析形式定义：

$$T_n(x) = \cos(n \arccos x)$$

Tchebyshev多项式性质

- $|T_n(x)| \leqslant 1,\ -1 \leqslant x \leqslant 1$
- $T_n\left(\cos\frac{j\pi}{n}\right) = (-1)^j,\ j = 0, \ldots, n$
- $T_n\left(\cos\frac{(2j-1)\pi}{2n}\right) = 0,\ j = 1, \ldots, n$
- $2^{1-n}T_n$是一个首一多项式

首一多项式定理

Theorem

设$p(x)$为一个n次首一多项式，则

$$\max_{-1\leqslant x\leqslant 1} |p(x)| \geqslant 2^{1-n}$$

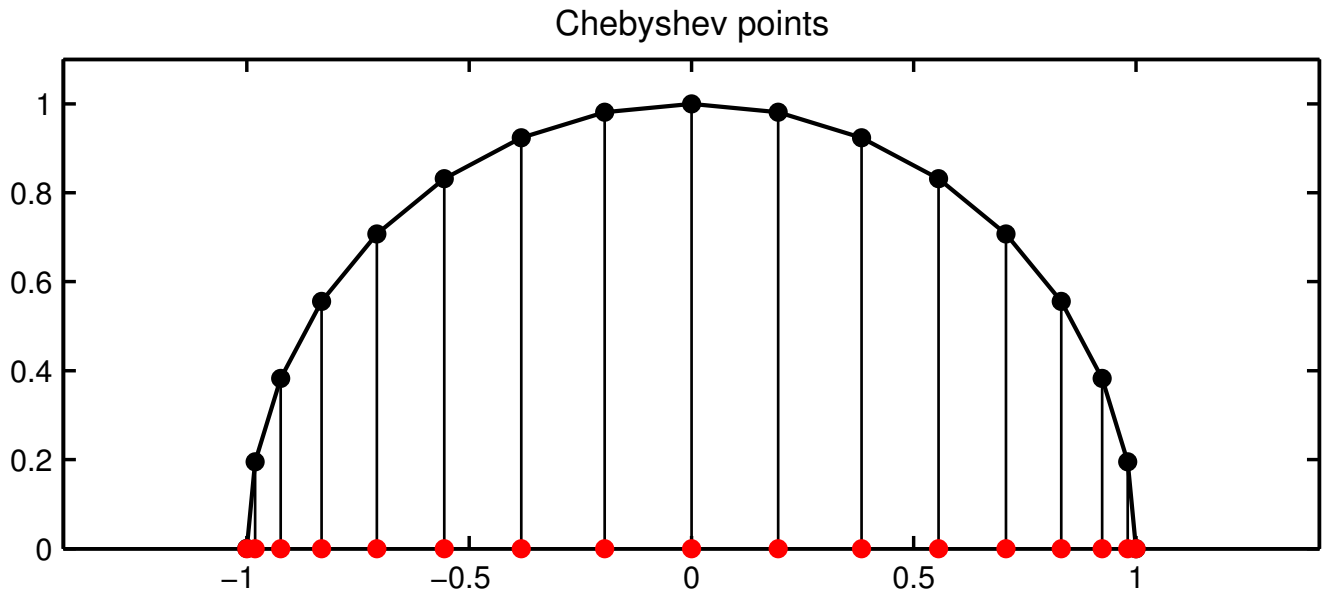
证明：反证法。设对任意$x \in [-1,1]$, $|p(x)| < 2^{1-n}$.
令$q(x) = 2^{1-n}T_n$, $x_i = \cos(i\pi/n)$,

$$(-1)^i p(x_i) \leqslant |p(x_i)| < (-1)^i q(x_i)$$

即$(-1)^i\big(q(x_i) - p(x_i)\big) > 0$, $i = 0,\ldots,n$. 这说明在区间$[-1,1]$上，多项式$q-p$的符号在正负之间变动了$n+1$次，即它在$(-1,1)$之间至少有n个根，而这是不可能的，因为$q-p$的次数至多是$n-1$. □

最佳结点选取

- Tchebyshev结点：结点x_i是Tchebyshev多项式$T_{n+1}(x)$的根



- 插值误差

$$|f(x) - p(x)| \leqslant \frac{1}{2^n(n+1)!} \max_{|t|\leqslant 1} |f^{(n+1)}(t)|$$

关于收敛性的定理

Theorem (Faber's定理)

对任意给定的结点组

$$a \leqslant x_0^{(n)} < x_1^{(n)} < \cdots < x_n^{(n)} \leqslant b, \quad n \geqslant 0 \tag{1}$$

在区间$[a, b]$上存在一个连续函数f使得f在这组结点上的插值多项式不能一致收敛于f.

Theorem

若$f \in C[a, b]$，则存在(1)式中那样的一组结点，使得f在这组结点上的插值多项式p_n满足

$$\lim_{n\to\infty} \|f - p_n\|_\infty = 0$$

正线性算子

- $C[a,b]$上的正线性算子L是指它满足
  1. 线性性：

$$L(af+bg)=aLf+bLg,\quad a,b\in\mathbb{R},\quad f,g\in C[a,b]$$

  2. 正性：若$f\geqslant 0$, 则$Lf\geqslant 0$
- 正线性算子的著名例子来自于Serge Bernstein在1912年定义的如下算子: 在$C[0,1]$中，

$$(B_nf)(x)=\sum_{i=0}^{n}f\left(\frac{k}{n}\right)B_k^n(x),\qquad B_k^n(x)=\binom{n}{k}x^k(1-x)^{n-k}$$

这里的$\{B_k^n(x)\}$称为Bernstein基函数。

Bohman-Korovkin定理

Theorem

设$L_n(n \geqslant 1)$是定义在$C[a,b]$上的一个正线性算子序列，其中每个算子在相同的空间中取值。若对于函数$f(x)=1,x,x^2$，$\|L_nf-f\|_\infty \to 0$成立，则对所有的$f \in C[a,b]$此结论也成立。

证明：若L为正线性算子，则由$f \geqslant g$可知$Lf \geqslant Lg$, 进一步有$L(|f|) \geqslant |Lf|$. 记$h_k(x)=x^k$, $k=0,1,2$. 再定义α_n, β_n, γ_n如下：

$$\alpha_n = L_nh_0 - h_0, \quad \beta_n = L_nh_1 - h_1, \quad \gamma_n = L_nh_2 - h_2$$

由定理的假设可知

$$\|\alpha_n\|_\infty \to 0, \quad \|\beta_n\|_\infty \to 0, \quad \|\gamma_n\|_\infty \to 0$$

下面证明对于任意$f \in C[a,b]$以及任意的$\varepsilon > 0$, 存在m使得当$n > m$时$\|L_nf-f\|_\infty < 3\varepsilon$.

由于f在紧区间上连续，从而一致连续，所以存在$\delta > 0$, 使得对于区间$[a, b]$中所有的x和y, 当$|x - y| < \delta$时，$|f(x) - f(y)| < \varepsilon$. 令$c = 2\|f\|_\infty/\delta^2$, 则有当$|x - y| \geqslant \delta$时，

$$|f(x) - f(y)| \leqslant 2\|f\|_\infty \leqslant 2\|f\|_\infty \frac{(x-y)^2}{\delta^2} = c(x-y)^2$$

从而对于$[a, b]$内的任意x, y有

$$|f(x) - f(y)| \leqslant \varepsilon + c(x-y)^2$$

上述不等式重写为：

$$|f - f(y)h_0| \leqslant \varepsilon h_0 + c[h_2 - 2yh_1 + y^2 h_0]$$

从而根据正线性算子的定义有：

$$|L_n f - f(y)L_n h_0| \leqslant \varepsilon L_n h_0 + c[L_n h_2 - 2yL_n h_1 + y^2 L_n h_0]$$

进一步用y代替x，

$$
\begin{aligned}
&|(L_n f)(y) - f(y)(L_n h_0)(y)| \\
&\leqslant \varepsilon(L_n h_0)(y) + c[(L_n h_2)(y) - 2y(L_n h_1)(y) + y^2(L_n h_0)(y)] \\
&= \varepsilon[1 + \alpha_n(y)] + c[y^2 + \gamma_n(y) - 2y(y + \beta_n(y)) + y^2(1 + \alpha_n(y))] \\
&= \varepsilon + \varepsilon\alpha_n(y) + c\gamma_n(y) - 2cy\beta_n(y) + cy^2\alpha_n(y) \\
&\leqslant \varepsilon + \varepsilon\|\alpha_n\|_\infty + c\|\gamma\|_\infty + 2c\|h_1\|_\infty\|\beta_n\|_\infty + c\|h_2\|_\infty\|\alpha_n\|_\infty
\end{aligned}
$$

因此存在m, 当$n \geqslant m$时，$\|L_n f - f \cdot L_n h_0\|_\infty \leqslant 2\varepsilon$. 因此必要时再增大$m$有

$$
\begin{aligned}
\|L_n f - f\|_\infty \leqslant& \|L_n f - f \cdot L_n h_0\|_\infty + \|f \cdot L_n h_0 - f \cdot h_0\|_\infty \\
\leqslant& 2\varepsilon + \|f\|_\infty\|\alpha_n\|_\infty \leqslant 3\varepsilon
\end{aligned}
$$

□

Bernstein算子的情形

- h_0: $(B_n h_0)(x) = \sum_{k=0}^{n} B_k^n(x) = 1$
- h_1:

$$(B_n h_1)(x) = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k} = x$$

- h_2:

$$(B_n h_2)(x) = \frac{n-1}{n} x^2 + \frac{x}{n} \to x^2$$

从而Bohman-Korovkin定理此时给出了Weierstrass定理：即有界闭区间上的连续函数可以被多项式一致逼近。

《数值分析》之

函数逼近

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

Newton插值

- Lagrange插值的缺点: 无承袭性。增加一个节点，所有的基函数都要重新计算
- 承袭性: $N_{n+1}(x) = N_n(x) + q_{n+1}(x)$
  - $N_n(x)$是利用$\{x_0, x_1, \cdots, x_n\}$插值得到的$n$阶多项式
  - $N_{n+1}(x)$是利用$\{x_0, x_1, \cdots, x_{n+1}\}$插值得到的$n+1$阶多项式
  - 增加一个节点，仅需在原有n个节点的多项式基础上添加多项式$q_{n+1}(x)$

如何构造

- 由$N_{n+1}(x_i) = N_n(x_i) = f(x_i), i = 0, \cdots, n$ 可知，$q_{n+1}(x)$有$\{x_0, x_1, \cdots, x_n\}$这$n+1$个零点
  则有$q_{n+1}(x) = a_{n+1}(x - x_0)(x - x_1)\cdots(x - x_n)$,其中$a_{n+1}$为实数
- $N_n(x) = N_{n-1}(x) + q_n(x)$
  则有$q_n(x) = a_n(x - x_0)(x - x_1)\cdots(x - x_{n-1})$,其中$a_n$为实数
- $N_1(x) = N_0(x) + q_1(x)$
  则有$q_1(x) = a_1(x - x_0)$,其中a_1为实数

Newton插值多项式

$$N_n(x) = a_0 + a_1(x - x_0) + \cdots + a_n(x - x_0)\cdots(x - x_{n-1})$$

确定系数a_n

$$
\begin{aligned}
&N_n(x_0) = a_0 = f(x_0),\\
&N_n(x_1) = a_0 + a_1(x_1 - x_0) = f(x_1),\\
&N_n(x_2) = a_0 + a_1(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1) = f(x_2),\\
&\qquad \cdots\\
&N_n(x_n) = a_0 + a_1(x_n - x_0) + \cdots + a_n(x_n - x_0)\cdots(x_n - x_{n-1}) = f(x_n)
\end{aligned}
$$

由此可得

$$
\begin{aligned}
a_0 &= f(x_0),\\
a_1 &= \frac{f(x_1) - f(x_0)}{x_1 - x_0}\\
a_2 &= \frac{1}{x_2 - x_1}\left(\frac{f(x_2) - f(x_0)}{x_2 - x_0} - a_1\right)\\
a_3 &= \frac{1}{x_3 - x_2}\left(\frac{f(x_3) - f(x_0)}{x_3 - x_0}\frac{1}{x_3 - x_1} - a_2\right)
\end{aligned}
$$

差商定义

定义

- 一阶差商

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

- k阶差商

  设$\{x_0, x_1, \cdots, x_k\}$互不相同，$f(x)$关于$\{x_0, x_1, \cdots, x_k\}$的$k$阶插商为

$$f[x_0, x_1 \cdots, x_k] = \frac{f[x_1, \cdots, x_k] - f[x_0, \cdots, x_{k-1}]}{x_k - x_0}$$

差商算法

$$
\begin{array}{llllll}
x_0 & f(x_0) & & & & \\
x_1 & f(x_1) & f[x_0, x_1] & & & \\
x_2 & f(x_2) & f[x_1, x_2] & f[x_0, x_1, x_2] & & \\
\cdots & & & & & \\
x_n & f(x_n) & f[x_{n-1}, x_n] & f[x_{n-2}, x_{n-1}, x_n] & f[x_0, x_1, \cdots, x_n]
\end{array}
$$

x	5	-7	-6	0
$f(x)$	1	-23	-54	-954

结点为$5, -7, -6, 0$，

$$
\begin{array}{rrrrr}
5 & 1 & & & \\
-7 & -23 & 2 & & \\
-6 & -54 & -31 & 3 & \\
0 & -954 & -150 & -17 & 4
\end{array}
$$

$c_0 = 1,\ c_1 = 2,\ c_2 = 3,\ c_3 = 4$, 所以插值多项式为

$$p_3(x) = 1 + 2(x-5) + 3(x-5)(x+7) + 4(x-5)(x+7)(x+6)$$

Newton插值多项式的表示

Newton插值多项式表示为

$$N_n(x) = f(x_0)+f[x_0,x_1](x-x_0)+\cdots+f[x_0,x_1,\cdots,x_n](x-x_0)\cdots(x-x_{n-1})$$

$$\begin{aligned}
a_0 &= f(x_0),\\
a_1 &= \frac{f(x_1)-f(x_0)}{x_1-x_0} = f[x_0,x_1]\\
a_2 &= \frac{1}{x_2-x_1}\left(\frac{f(x_2)-f(x_0)}{x_2-x_0}-a_1\right)\\
&= \frac{1}{x_2-x_1}(f[x_2,x_0]-f[x_1,x_0]) = f[x_0,x_1,x_2]\\
&\cdots\\
a_n &= f[x_0,x_1,\cdots,x_n]
\end{aligned}$$

算法

- 计算Newton多项式的值

```
for(i=1;i<=n;i++) !计算差商表
{
    for(j=n;j>=i;j--)
        y[j]=(y[j]-y[j-1])/(x[j]-x[j-i]);
}

fx=y[n]; !求Newton多项式的值
for(i=n;i>=1;i--)
{
    fx=y[i-1]+(x-x[i-1])fx;
}
```

差商性质

- k阶差商$f[x_0, x_1, \cdots, x_k]$可由$f(x_0), f(x_1), \cdots, f(x_k)$的线性表示
  - 由多项式插值的唯一性，知$N_k(x) = L_k(x)$.
  - x^k的系数相同
  - $f[x_0, x_1, \cdots, x_k] = \sum_{i=0}^{n} \frac{f(x_i)}{(x_i - x_0)\cdots(x_i - x_{i-1})(x_i - x_{i+1})\cdots(x_i - x_n)}$
- 对称性：若$i_0, i_1, \cdots, i_k$为$0, 1, \cdots, k$的任意排列，则有

$$f[x_0, x_1, \cdots, x_k] = f[x_{i_0}, x_{i_1}, \cdots, x_{i_k}]$$

- 若$f(x)$为m次多项式，则$f[x_0, x_1, \cdots, x_{k-1}, x]$为$m - k$次多项式。
- 函数差商与函数导数的关系

$$f[x, x_0, x_1, \cdots, x_n] = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x)$$

Newton插值多项式的误差

多项式插值误差定理对于Newton插值多项式同样成立，故有对$[a,b]$中每个x，都有$\xi_x \in (a,b)$使得

$$f(x) - N_n(x) = R_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i)$$

而

$$R_n(x) = f[x, x_0, x_1 \cdots, x_n] \prod_{i=0}^{n} (x - x_i)$$

故有

$$f[x, x_0, x_1 \cdots, x_n] = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x)$$

上机作业2

对函数

$$f(x)=\frac{1}{1+25x^2}, x\in[-1,1]$$

构造牛顿插值多项式$p_N(x)$，插值节点取为:

1. $x_i=1-\frac{2}{N}i,\ i=0,1,\cdots,N$
2. $x_i=-\cos(\frac{2i+1}{2N+2}\pi),\ i=0,1,\cdots,N$ (Chebyshev point)

并计算如下误差

$$\max_i\{|f(y_i)-p(y_i)|,\ y_i=\frac{i}{50}-1,\ i=0,1,\cdots,100\}$$

对$N=5,10,20,40$比较以上两组节点的结果，并在一张图中画出$N=20$时$f(x)$数值计算结果。

输出形式如下：

N=5

Max Error of grid (1) : XXXXXXXXXXXXXXXXXX

Max Error of grid (2) : XXXXXXXXXXXXXXXXXX

N=10

Max Error of grid (1) : XXXXXXXXXXXXXXXXXX

Max Error of grid (2) : XXXXXXXXXXXXXXXXXX

N=20

Max Error of grid (1) : XXXXXXXXXXXXXXXXXX

Max Error of grid (2) : XXXXXXXXXXXXXXXXXX

N=40

Max Error of grid (1) : XXXXXXXXXXXXXXXXXX

Max Error of grid (2) : XXXXXXXXXXXXXXXXXX

例1 给定f在x_0和x_1上的函数值和一阶导数值(共四个条件)，要求一个三次多项式$p(x)$, 在给定结点上与给定信息吻合。

待定

$$p(x) = a + b(x - x_0) + c(x - x_0)^2 + d(x - x_0)^2(x - x_1)$$

则可知

$$\begin{aligned} a &= f(x_0) \\ b &= f'(x_0) \\ a + b(x_1 - x_0) + c(x_1 - x_0)^2 &= f(x_1) \\ b + 2c(x_1 - x_0) + d(x_1 - x_0)^2 &= f'(x_1) \end{aligned}$$

因此插值多项式存在唯一。

例2 求一多项式p，使得$p(0)=0,\ p(1)=1,\ p'(1/2)=2$.
由于给定了三个条件，因此试用二次多项式：$p(x)=a+bx+cx^2$. 由$p(0)=0\Longrightarrow a=0$. 而另外两个条件有

$$1=p(1)=b+c$$
$$2=p'(1/2)=b+c$$

因此不存在二次多项式满足插值条件。因此考虑三次多项式，$p(x)=a+bx+cx^2+dx^3$. 此时解不唯一：$d=-4$, $b+c=5$, $a=0$.

Hermite 插值问题

Hermite插值指的是对一个函数在一组结点上的函数值和导数值进行插值。
给定函数f以及结点$x_0, \ldots, x_n$，求多项式p:

$$p^{(j)}(x_i) = f^{(j)}(x_i), \quad j = 0, 1, i = 0, \ldots, n$$

- 多项式插值空间的维数，
- 共有2(n+1)个条件
- 多项式最高次数为2n+1

Hermite 插值问题（续）

$$H(x)=\sum_{i=0}^{n} h_i(x)f(x_i)+\sum_{i=0}^{n} g_i(x)f'(x_i)$$

问题变为求解插值基函数$\{h_i(x)\}_i^n$, $\{g_i(x)\}_i^n \in P^{2n+1}(x)$,满足

$$\begin{cases} h_i(x_j)=\delta_{ij} \\ h_i'(x_j)=0 \end{cases}, \quad \begin{cases} g_i(x_j)=0 \\ g_i'(x_j)=\delta_{ij} \end{cases},$$

	h_0	$\cdots$	h_n	g_0	$\cdots$	g_n
x_0	1	$\cdots$	0	0	$\cdots$	0
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	0	$\cdots$	1	0	$\cdots$	0
x_0'	0	$\cdots$	0	1	$\cdots$	0
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n'	0	$\cdots$	0	0	$\cdots$	1

Hermite 插值基函数

$$h_i(x) = \left(1 - 2(x - x_i)\sum_{i \neq j} \frac{1}{x_i - x_j}\right) \ell_i^2(x)$$

$$g_i(x) = (x - x_i)\ell_i^2(x)$$

当取2个节点时的Hermite插值多项式基函数为

$$h_0(x) = \left(1 - 2\frac{x - x_0}{x_0 - x_1}\right)\left(\frac{x - x_1}{x_0 - x_1}\right)^2$$

$$h_1(x) = \left(1 - 2\frac{x - x_1}{x_1 - x_0}\right)\left(\frac{x - x_0}{x_1 - x_0}\right)^2$$

$$g_0(x) = (x - x_0)\left(\frac{x - x_1}{x_0 - x_1}\right)^2$$

$$g_1(x) = (x - x_1)\left(\frac{x - x_0}{x_1 - x_0}\right)^2$$

例 给定$f(-1)=0,\ f(1)=4,\ f'(-1)=2,\ f'(1)=0,$
求Hermite插值多项式，并计算$f(0.5)$
解：

$$H_3(x)=h_0(x)\cdot 0+h_1(x)\cdot 4+g_0(x)\cdot 2+g_1(x)\cdot 0$$

只需计算$h_1(x)$和$g_0(x)$

$$h_1(x)=\left(1-2\frac{x-1}{1+1}\right)\left(\frac{x+1}{1+1}\right)^2=\frac{1}{4}(2-x)(x+1)^2$$

$$g_0(x)=(x+1)\left(\frac{x-1}{-1-1}\right)^2=\frac{1}{4}(x+1)(x-1)^2$$

$$H_3(x)=(2-x)(x+1)^2+\frac{1}{2}(x+1)(x-1)^2$$

$$H_3(0.5)=3.5625$$

Theorem (Hermite插值误差估计定理)

若$f \in C^{2n+2}[a,b]$，$[a,b]$内的插值结点为$x_0,\ldots,x_n$，$p(x)$为相应的*Hermite*插值多项式，$\deg p \leqslant 2n+1$，则对于任意$x \in [a,b]$，存在$\xi_x \in (a,b)$使得

$$f(x)-p(x)=\frac{f^{(2n+2)}(\xi_x)}{(2n+2)!}\prod_{i=0}^{n}(x-x_i)^2$$

证明方法与无重结点的多项式插值误差估计定理完全类似。

Hermite 插值问题推广

给定函数f以及结点$x_0, \ldots, x_n$，求多项式p:

$$p^{(j)}(x_i) = f^{(j)}(x_i), \quad j = 0, \ldots, k_i - 1, i = 0, \ldots, n$$

Theorem (Hermite插值定理)

存在唯一的次数不超过$m = k_0 + \cdots + k_n - 1$的多项式满足上述插值条件。

证明：通过在幂基$\{1, x, \ldots, x^m\}$下待定多项式的系数，得到一个线性方程组$Au = b$, 其中A为$(m+1) \times (m+1)$阶矩阵(称为广义Vandermonde矩阵). 为证其有唯一解，只要证$Au = 0$仅有零解，即满足$p^{(j)}(x_i) = 0$的次数不超过m的多项式只能是零多项式。这可以通过统计p的零点数得证。 □

- 为了简化记号，把插值结点重记为$t_0, \ldots, t_m$，其中$t_0 = t_1 = \cdots = t_{k_1-1} = x_0$, ...
- 记f在结点$t_0, t_1, \ldots, t_m$上次数不超过m的插值多项式的x^m项系数为$f[t_0, \ldots, t_m]$.

Theorem (Newton插值多项式定理)

满足插值条件的多项式可以写为

$$p(x) = \sum_{j=0}^{n} f[x_0, \ldots, x_j] \prod_{i=0}^{j-1} (x - x_i)$$

证明：归纳法。 □

高阶差商的性质

- 差商是结点的对称函数
- $f[x_0,\ldots,x_0]=f^{(n)}(x_0)/n!$
- 设$x_0\leqslant x_1\leqslant\cdots\leqslant x_n$,

$$f[x_0,x_1,\ldots,x_n]=\begin{cases}\frac{f[x_1,x_2,\ldots,x_n]-f[x_0,x_1,\ldots,x_{n-1}]}{x_n-x_0} & x_n\neq x_0\\ \frac{f^{(n)}(x_0)}{n!} & x_n=x_0\end{cases}$$

- 消去性质：$f[x_0,\ldots,x_n]=\{(x-x_{n+1})f(x)\}[x_0,\ldots,x_n,x_{n+1}]$
- Leibnitz法则:

$$(fg)[x_0,\ldots,x_n]=\sum_{k=0}^{n}f[x_0,\ldots,x_k]g[x_k,\ldots,x_n]$$

用Newton方法确定一个多项式，满足

$$p(1) = 2, p'(1) = 3, p(2) = 6, p'(2) = 7, p''(2) = 8$$

解：

已知信息

1	2	3	?	?	?
1	2	?	?	?	
2	6	7	4		
2	6				
2	6				

用Newton方法确定一个多项式，满足

$$p(1)=2, p'(1)=3, p(2)=6, p'(2)=7, p''(2)=8$$

解：

新信息的计算

1	2				
1	2	3			
2	6	4	1		
2	6	7	3	2	
2	6	7	4	1	-1

从而所求多项式为

$$p(x)=2+3(x-1)+(x-1)^2+2(x-1)^2(x-2)-(x-1)^2(x-2)^2$$

《数值分析》之

函数逼近

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

Runge现象

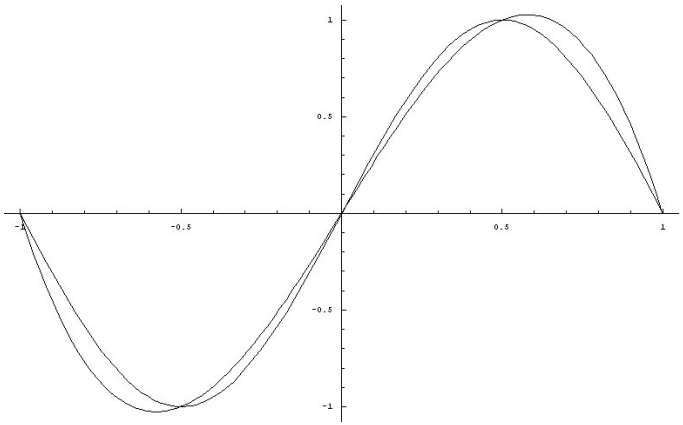
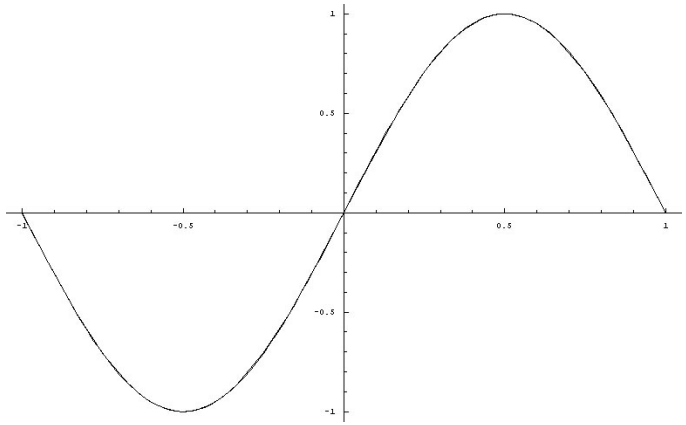
$$f(x) = \frac{1}{1+x^2}, x \in [-5, 5]$$

等距插值函数$p_n(x)$.

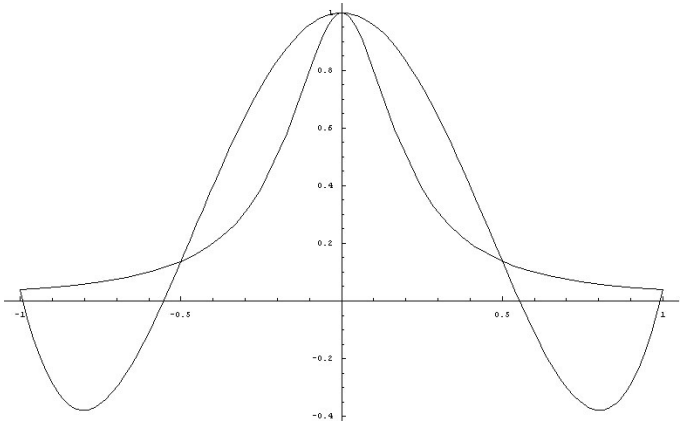
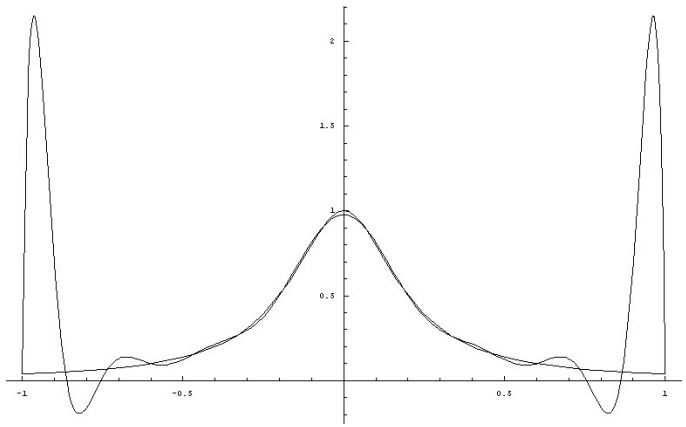
Degree n	Max error
2	0.65
4	0.44
6	0.61
8	1.04
10	1.92
12	3.66
14	7.15
16	14.25
18	28.74
20	58.59
22	121.02
24	252.78

Runge现象

- $f(x) = \sin \pi x$, $[-1, 1]$上的5个和16个等距结点

- $f(x) = \frac{1}{(25x^2+1)}$, $[-1, 1]$上的5个和16个等距结点

此例由Runge在1901年给出

Runge现象（续）

- n 越大，端点附近抖动越大
- 等距高次插值，数值稳定性差，本身是病态的。
- 考虑使用分段插值

分段线性插值

对给定区间$[a,b]$作分割

$$a = x_0 < x_1 < \cdots < x_n = b$$

在每个小区间$[x_i, x_{i+1}]$上，作线性插值函数$p(x) = p_i(x)$

$$p_i(x) = \frac{x - x_{i+1}}{x_i - x_{i+1}} f(x_i) + \frac{x - x_i}{x_{i+1} - x_i} f(x_{i+1}),\ x \in [x_i, x_{i+1}]$$

满足

- $p(x)$连续，即保证函数在$x = x_0, x_1, \cdots, x_n$点连续
- $p(x)$在每个小区间上为一个不高于1次的多项式

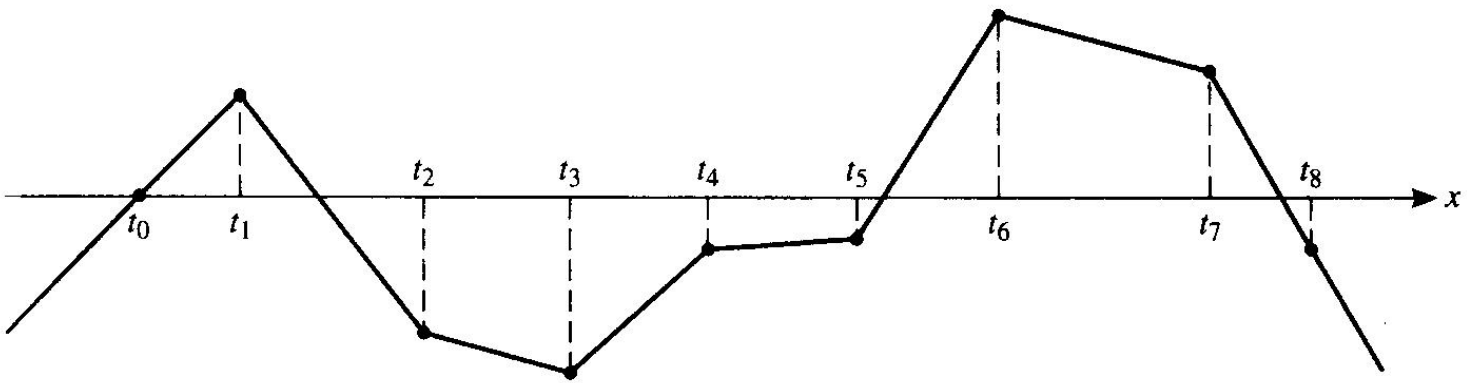
当$x \in [x_i, x_{i+1}]$时，

$$f(x) - p(x) = f(x) - p_i(x) = \frac{f^{(2)}}{2}(x - x_i)(x - x_{i+1})$$
$$\leq \frac{M_2}{8}(x_{i+1} - x_i)^2, \quad M_2 = \max_{a \leq x \leq b} |f''(x)|$$

由此得

$$|f(x) - p(x)| = \max_n \{\frac{M_2}{8}(x_{i+1} - x_i)^2\} = \frac{M_2}{8}h^2$$

当$h \to 0$时，$p(x)$收敛于$f(x)$.



- 局部性质，如果修改了某节点$(x_i, f(x_i))$的值，仅在相邻的两个区间$[x_{i-1}, x_i]$, $[x_i, x_{i+1}]$需要改动
- 插值节点处仅连续，不光滑
- 类似，可以作二重Hermite插值

样条函数

- 样条函数是一类分段（片）光滑、并且在各段交接处也有一定光滑性的函数。简称样条(spline)。
- 样条一词来源于工程绘图人员为了将一些指定点连接成一条光顺曲线所使用的工具，即富有弹性的细木条或薄钢条。由这样的样条形成的曲线在连接点处具有连续的坡度与曲率。分段低次多项式、在分段处具有一定光滑性的函数插值就是模拟以上原理发展起来的，它克服了高次多项式插值可能出现的振荡现象，具有较好的数值稳定性和收敛性，由这种插值过程产生的函数就是多项式样条函数。

历史与应用概述

- 样条函数的研究始于20世纪中叶的概率论和数理统计中的数据拟合，到了60年代它与计算机辅助设计相结合，在外形设计方面得到成功的应用。
- 样条理论已成为函数逼近的有力工具。它的应用范围也在不断扩大，不仅在数据处理、数值微分、数值积分、微分方程和积分方程数值解等数学领域有广泛的应用，而且与最优控制、变分问题、统计学、计算几何与泛函分析等学科均有密切的联系

定义

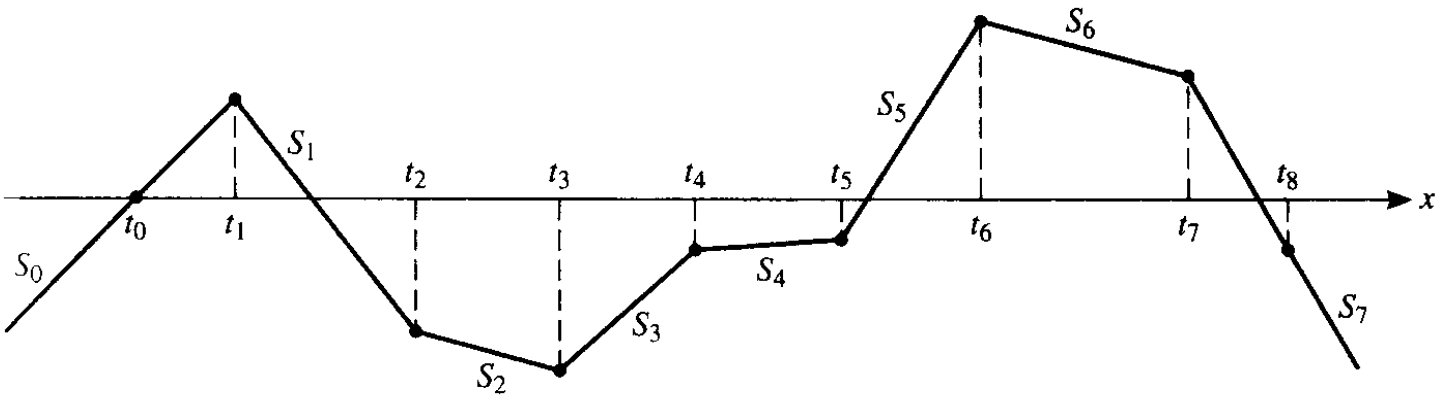
- 给定$n+1$个点$t_0 < t_1 < \cdots < t_n$ (称为结点, knots)，并指定一个非负整数$k \geqslant 0$. 在这些结点上定义的一个k次样条函数(spline function)是指满足下列条件的函数S:
  1. 在每个区间$[t_{i-1}, t_i)$上，S是一个次数不超过k的多项式。
  2. 在$[t_0, t_n]$上S有$(k-1)$阶连续导数。
- 当固定结点以及次数k，那么所有的S在通常函数运算的意义下构成一个线性空间。自然的问题是这个样条空间的维数是多少？基函数是什么？这就是样条理论的核心。其中对基函数的一般要求是：
  1. 非负性
  2. 单位剖分
  3. 局部支集

低次样条

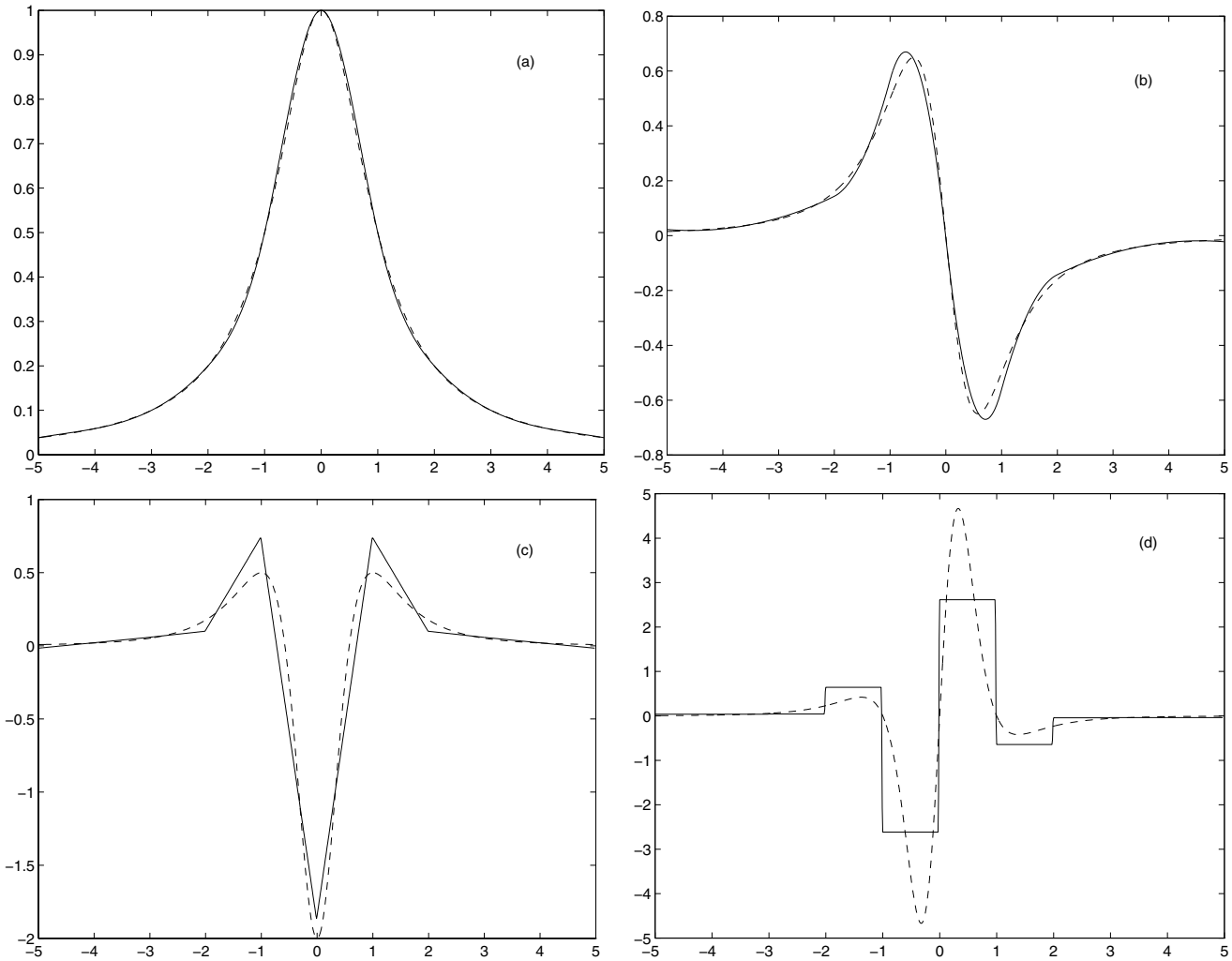
- $k = 0$: 分段常数

$$S(x) = \begin{cases} S_0(x) = c_0 & x \in [t_0, t_1) \\ S_1(x) = c_1 & x \in [t_1, t_2) \\ \vdots & \vdots \\ S_{n-1}(x) = c_{n-1} & x \in [t_{n-1}, t_n] \end{cases}$$

- $k = 1$: 分段折线函数，零阶连续



$\frac{1}{1+x^2}$：三次样条插值



函数值，一阶导数，二阶导数，三阶导数。

三次样条插值

- 给定数据点$(x_0, y_0), \ldots, (x_n, y_n)$, 其中$x_0 < \cdots < x_n$. 计算一个以$x_0, \ldots, x_n$为结点的三次样条函数$S(x)$使得,
  - $S(x_i) = y_i, \quad i = 0, 1, \cdots, n$
  - 在每个小区间$[x_i, x_{i+1}]$上至多为三次多项式
  - $S(x)$在$[a, b]$上有连续的二阶导数
- 条件分析:
  1. 待定每个区间上的三次多项式形式, 共有$4n$个自由系数
  2. 每个子区间上的三次多项式在两个端点有两个函数值约束, 即$S(x_i) = y_i$, $S(x_{i+1}) = y_{i+1}$, 因此共有$2n$个约束
  3. 在每一个内结点上, S的一阶和二阶导数连续, 分别给出$n-1$个约束

$2n + n - 1 + n - 1 = 4n - 2$

因此还有两个(??, 有待确认)自由度, 通常在边界处给出

三次样条插值的计算

- 记待求样条在结点的二阶导数值为$M_i = S''(x_i)$
- $S_i''(x)$是区间$[x_i, x_{i+1}]$上的线性函数，因此

$$S_i''(x) = \frac{M_i}{h_i}(x_{i+1} - x) + \frac{M_{i+1}}{h_i}(x - x_i), \quad h_i = x_{i+1} - x_i$$

- 积分两次，再结合$S_i(x_i) = y_i$和$S_i(x_{i+1}) = y_{i+1}$:

$$\begin{aligned} S_i(x) =& \frac{M_i}{6h_i}(x_{i+1} - x)^3 + \frac{M_{i+1}}{6h_i}(x - x_i)^3 \\ &+ \left(\frac{y_{i+1}}{h_i} - \frac{M_{i+1}h_i}{6}\right)(x - x_i) + \left(\frac{y_i}{h_i} - \frac{M_i h_i}{6}\right)(x_{i+1} - x) \end{aligned}$$

三次样条插值的计算(续)

- 可以用S'的连续性确定M_i, $i = 1, \ldots, n-1$. 由前页结果，

$$S_i'(x_i) = -\frac{h_i}{3}M_i - \frac{h_i}{6}M_{i+1} - \frac{y_i}{h_i} + \frac{y_{i+1}}{h_i}$$

$$S_{i-1}'(x_i) = \frac{h_{i-1}}{6}M_{i-1} + \frac{h_{i-1}}{3}M_i - \frac{y_{i-1}}{h_{i-1}} + \frac{y_i}{h_{i-1}}$$

根据$S_{i-1}'(x_i) = S_i'(x_i)$可得

$$h_{i-1}M_{i-1} + 2(h_i + h_{i-1})M_i + h_iM_{i+1} = \frac{6}{h_i}(y_{i+1} - y_i) - \frac{6}{h_{i-1}}(y_i - y_{i-1})$$

对$i = 1, \ldots, n-1$成立

- 从而对$M_0, \ldots, M_n$给出了$n-1$个线性条件。可以任意指定M_0和M_n以求得$M_1, \ldots, M_{n-1}$, 从而确定相应的样条

三次自然样条

- 当取$M_0 = M_n = 0$时，得到的样条称为三次自然样条(natural cubic spline)
- 此时求解$M_1, \ldots, M_{n-1}$对应的线性方程组为

$$\begin{pmatrix} u_1 & h_1 & & & & \\ h_1 & u_2 & h_2 & & & \\ & h_2 & u_3 & h_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & h_{n-3} & u_{n-2} & h_{n-2} \\ & & & & h_{n-2} & u_{n-1} \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ M_3 \\ \vdots \\ M_{n-2} \\ M_{n-1} \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_{n-2} \\ v_{n-1} \end{pmatrix}$$

其中$h_i = t_{i+1} - t_i$, $u_i = 2(h_i + h_{i+1})$, $b_i = 6(y_{i+1} - y_i)/h_i$, $v_i = b_i - b_{i-1}$. 系数阵是对称的、三对角的、对角占优的

其他边界条件

- 固定边界: 当取$S'(x_0) = m_0$, $S'(x_n) = m_n$时，此时边界处满足

$$2M_0 + M_1 = \frac{6}{h_0}\left(\frac{y_1 - y_0}{h_0} - m_0\right) = v_0$$

$$M_{n-1} + 2M_n = \frac{6}{h_{n-1}}\left(m_n - \frac{y_n - y_{n-1}}{h_{n-1}}\right) = v_n$$

  此时求解$M_0, \ldots, M_n$对应的$n+1$阶线性方程组为

- 周期边界： $m_0 = m_n$. $M_0 = M_n$.

样条函数求值

当确定系数$M_0, \ldots, M_n$后，对应的三次样条函数在任意点x处的值可如下计算：

1. 确定x是在下述哪个区间中(最有效方法是什么？)：

$$(-\infty, t_1), [t_1, t_2), \ldots, [t_{n-2}, t_{n-1}), [t_{n-1}, +\infty)$$

假设区间指标为i

2. 重写$S_i(x)$的表达式为

$$S_i(x) = y_i + (x - t_i)[C_i + (x - t_i)[B_i + (x - t_i)A_i]]$$

其中

$$A_i = \frac{1}{6h_i}(M_{i+1} - M_i)$$

$$B_i = \frac{M_i}{2}$$

$$C_i = -\frac{h_i}{6}M_{i+1} - \frac{h_i}{3}M_i + \frac{1}{h_i}(y_{i+1} - y_i)$$

Theorem (三次自然样条最优性定理)

设$f \in C^2[a,b]$, $a = t_0 < t_1 < \cdots < t_n = b$。若$S$是$f$在结点$t_0, \ldots, t_n$上的三次自然插值样条，则

$$\int_a^b \left(S''(x)\right)^2 dx \leqslant \int_a^b \left(f''(x)\right)^2 dx$$

证明：令$g = f - S$，则$g(t_i) = 0$, $i = 0, \ldots, n$, 并且

$$\int_a^b (f'')^2 dx = \int_a^b (S'')^2 dx + \int_a^b (g'')^2 dx + 2\int_a^b S''g''dx$$

根据分部积分：

$$\begin{aligned}\int_a^b S''g''dx &= \sum_{i=1}^n \int_{t_{i-1}}^{t_i} S''g''dx \\ &= \sum_{i=1}^n \left\{(S''g')(t_i) - (S''g')(t_{i-1}) - \int_{t_{i-1}}^{t_i} S'''g'dx\right\}\end{aligned}$$

$$
\begin{aligned}
&= -\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} S''' g' dx \\
&= -\sum_{i=1}^{n} c_i \int_{t_{i-1}}^{t_i} g' dx \\
&= -\sum_{i=1}^{n} c_i [g(t_i) - g(t_{i-1})] \\
&= 0
\end{aligned}
$$

□

自然样条最优性分析

- 由于$y = f(x)$定义曲线的曲率为$|f''(x)|\left[1 + \left(f'(x)\right)^2\right]^{-3/2}$, 因此可以认为三次自然样条是一条具有最小近似曲率的曲线
- 定理证明中应用到了下述和式：

$$\sum_{i=1}^{n} \left\{ (S''g')(t_i) - (S''g')(t_{i-1}) \right\} = (S''g')(b) - (S''g')(a)$$

证明中这个和式为零。实际上只要它非负，定理仍然成立。令$S'(a) = f'(a)$, $S'(b) = f'(b)$也是一种选择的方法。

高次自然样条

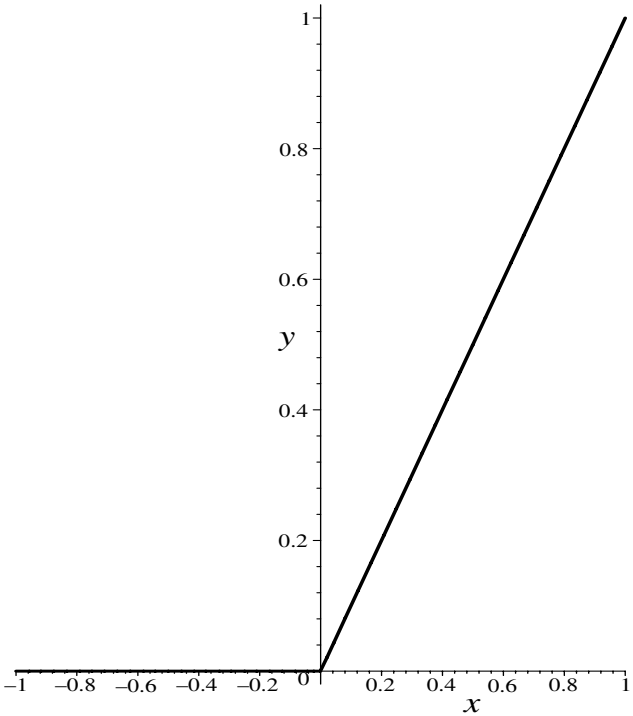
- 三次自然样条是完全来自于工程实际，我们可以对此进行推广，定义奇数次自然样条
- 给定结点集$t_0 < t_1 < \cdots < t_n$, 一个$2m+1$次自然样条是一个函数$S \in C^{2m}(\mathbb{R})$, 在每一个区间$[t_i, t_{i+1}]$内它是一个次数不超过$2m+1$的多项式，而在区间$(-\infty, t_0)$和$(t_n, +\infty)$内为一个次数至多为m的多项式。在给定结点上的全体$2m+1$次自然样条构成的线性空间记为$\mathcal{N}^{2m+1}(t_0, t_1, \ldots, t_n)$或$\mathcal{N}_n^{2m+1}$
- 三次自然样条符合上述定义，此时在区间$(-\infty, t_0)$和$(t_n, +\infty)$内延拓定义为线性多项式

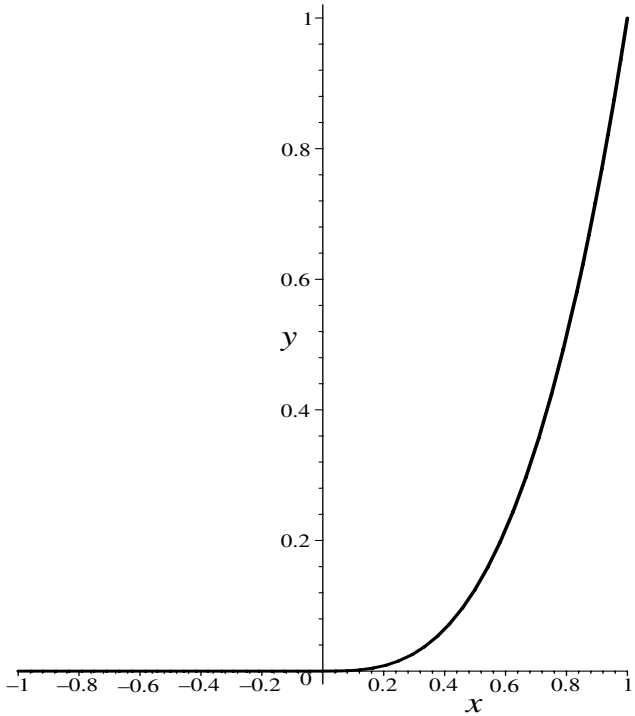
自然样条的表示

- 截断幂函数：

$$x_+^n = \begin{cases} x^n & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

它是属于C^{n-1}函数类的





Theorem

$\mathcal{N}_n^{2m+1}$中的每个元素S可以表示为

$$S(x) = \sum_{i=0}^{m} a_i x^i + \sum_{i=0}^{n} b_i (x - t_i)_+^{2m+1}$$

其中对于$0 \leqslant j \leqslant m$, $\sum_{i=0}^{n} b_i t_i^j = 0$.

证明

- 在$(-\infty, t_0)$内，$a_0, \ldots, a_m$被唯一确定。
- 在$[t_i, t_{i+1})$内S为一个$2m+1$次多项式，其中t_i点的直至$2m$阶导数已确定，因此存在b_i使得

$$S(x) = \sum_{i=0}^{m} a_i x^i + \sum_{i=0}^{n} b_i (x - t_i)_+^{2m+1}$$

- 在区间$(t_n, +\infty)$上，S为次数$\leqslant m$的多项式，因此

$$0 = S^{(m+1)}(x) = \sum_{i=0}^{n} b_i (2m+1)(2m)\cdots(m+1)(x-t_i)^m, \quad x > t_n$$

$$\begin{aligned} 0 &= \sum_{i=0}^{n} b_i (x - t_i)^m = \sum_{i=0}^{n} b_i \sum_{j=0}^{m} \binom{m}{j} x^{m-j} (-t_i)^j \\ &= \sum_{j=0}^{m} \left(\sum_{i=0}^{n} b_i t_i^j \right) (-1)^j \binom{m}{j} x^{m-j} \end{aligned}$$

自然样条唯一性定理

Theorem

给定结点$t_0 < t_1 < \cdots < t_n$, $0 \leqslant m \leqslant n$, 则存在唯一的$2m+1$次自然样条在这些结点上取给定值。

证明：根据自然样条表示的定理得到如下方程组

$$\begin{cases} S(t_i) = \sum_{j=0}^{m} a_j t_i^j + \sum_{j=0}^{n} b_j (t_i - t_j)_+^{2m+1} = \lambda_i, \quad 0 \leqslant i \leqslant n \\ \sum_{j=0}^{n} b_j t_j^i = 0, \quad 0 \leqslant i \leqslant m \end{cases}$$

有$m+n+2$个方程，$m+n+2$个未知数。为证方程组有唯一解，只需要证对应的齐次方程仅有零解。

假设对$i = 0, 1, \ldots, n$, $S(t_i) = 0$，下证

$$I = \int_{t_0}^{t_n} \left(S^{(m+1)}(x)\right)^2 dx = 0 \tag{1}$$

实际上，

$$I = S^{(m+1)}(x)S^{(m)}(x)\Big|_{t_0}^{t_n} - \int_{t_0}^{t_n} S^{(m)}(x)S^{(m+2)}(x)dx$$

$$= -\int_{t_0}^{t_n} s^{(m)}(x)S^{(m+2)}(x)dx = \cdots = (-1)^m \int_{t_0}^{t_n} S'(x)S^{(2m+1)}(x)dx$$

$$= (-1)^m \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} c_i S'(x)dx = (-1)^m \sum_{i=1}^{n} c_i(S(t_i) - S(t_{i-1})) = 0$$

因(1)式可知$S^{m+1}(x) \equiv 0$, 即S为一个次数至多m的多项式，其有零点$t_0, \ldots, t_n$，$n + 1 > m$，因此$S(x)$为零函数 □

自然样条最优性定理

Theorem

设$m \leqslant n$, $f \in C^{m+1}[a,b]$, S为在结点$a = t_0 < t_1 < \cdots < t_n = b$上插值$f$的$2m+1$次自然样条，则

$$\int_a^b \left(S^{(m+1)}(x)\right)^2 dx \leqslant \int_a^b \left(f^{(m+1)}(x)\right)^2 dx$$

证明：令$g = f - S$，则可以证明

$$\int_a^b g^{(m+1)}(x) S^{(m+1)}(x) dx = 0$$

根据这一正交性可以证明所需结论。 □

B样条

- B样条(B-splines)是给定样条空间的一组特殊基，从而其它样条函数都可以写成它的线性组合
- B样条具有许多很好的性质，如：
  - 局部支集
  - 非负性
  - 单位剖分，即所有的基函数的和恒等于1
  - 定义和计算简单
- 在B样条理论中，有许多方法可以作为定义B样条，如借助于截断幂函数的均差方法；借助于多重线性组合的递归方法和Blossoming方法[1]，等等

[1]在R. Goldman所著《金字塔算法》一书的前言中提到，“太早介绍blossoming，会出现一些问题。一是blossoming太有用了。晚接触它，可以多学一些其它的方法；二是只有亲眼看到blossoming可以取代那些毫无关联的方法技巧后，才会真正体会到blossoming的用处。”

递归定义

- 0次B样条：

$$B_i^0(x) = \begin{cases} 1 & t_i \leqslant x < t_{i+1} \\ 0 & \text{其它} \end{cases}$$

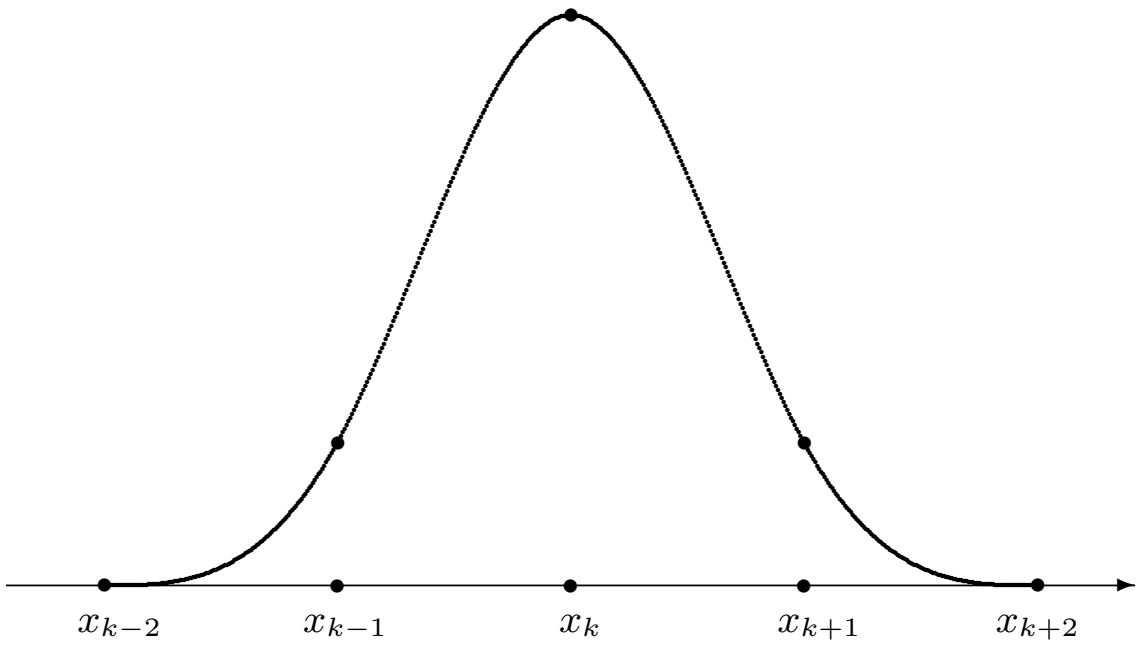
- $k \geqslant 1$:

$$B_i^k(x) = \frac{x - t_i}{t_{i+k} - t_i} B_i^{k-1}(x) + \frac{t_{i+k+1} - x}{t_{i+k+1} - t_{i+1}} B_{i+1}^{k-1}(x)$$

例：一次B样条

$$B_i^1(x) = \begin{cases} 0 & x < t_i \text{或} x \geqslant t_{i+2} \\ \frac{x-t_i}{t_{i+1}-t_i} & t_i \leqslant x < t_{i+1} \\ \frac{t_{i+2}-x}{t_{i+2}-t_{i+1}} & t_{i+1} \leqslant x < t_{i+2} \end{cases}$$

例：三次B样条



B样条性质

- $B_i^k(x)$的支集为$[t_i, t_{i+k+1}]$
- $B_i^k(x) \geqslant 0.$
- $\sum\limits_{i=-\infty}^{\infty} B_i^k(x) = 1$
- $$\frac{d}{dx} B_i^k(x) = \frac{k}{t_{i+k} - t_i} B_i^{k-1}(x) - \frac{k}{t_{i+k+1} - t_{i+1}} B_{i+1}^{k-1}(x)$$
- $$\int_{-\infty}^{x} B_i^k(s)ds = \frac{t_{i+k+1} - t_i}{k+1} \sum_{j=i}^{\infty} B_j^{k+1}(x)$$
- $\{B_j^k, \ldots, B_{j+k}^k\}$在$(t_{k+j}, t_{k+j+1})$上线性无关，
  $\{B_{-k}^k, \ldots, B_{n-1}^k\}$在$(t_0, t_n)$上线性无关

样条函数求值

- B样条函数构成相应样条空间的一组基，若

$$S(x) = \sum_i c_i^k B_i^k(x)$$

则称c_i^k为相应于$B_i^k(x)$的控制系数

- 为了计算函数$S(x)$在一点的值，先计算出每个基函数在点x的值，再进行线性组合。这是一种效率很低的算法。实际上B样条函数标准的求值算法，是de Boor算法

de Boor算法

确定指标m使得$t_m \leqslant x < t_{m+1}$

$$\begin{array}{llll} c_m^k & c_m^{k-1} & \cdots & c_m^1 \quad c_m^0 = S(x) \\ c_{m-1}^k & c_{m-1}^{k-1} & \cdots & c_{m-1}^1 \\ \vdots & \vdots & \ddots & \\ c_{m-k+1}^k & c_{m-k+1}^{k-1} & & \\ c_{m-k}^k & & & \end{array}$$

其中

$$c_i^{j-1} = \frac{1}{t_{i+j} - t_i}\left((x - t_i)c_i^j + (t_{i+j} - x)c_{i-1}^j\right)$$

上机作业

对函数

$$f(x) = e^x, x \in [0, 1]$$

构造等距节点的样条插值函数，对以下两种类型的样条函数

1. 一次分片线性样条
2. 满足$S'(0) = 1,\ S'(1) = e$的三次样条

并计算如下误差

$$\max_i \{|f(x_{i-\frac{1}{2}}) - S(x_{i-\frac{1}{2}})|,\ \ i = 1, \cdots, N\}$$

这里$x_{i-\frac{1}{2}}$为每个小区间的中点。对$N = 5, 10, 20, 40$比较以上两组节点的结果。讨论你的结果。利用公式计算算法的收敛阶。

$$Ord = \frac{\ln(Error_{old}/Error_{now})}{\ln(N_{now}/N_{old})}$$

输出形式如下：

n	Method (1) error	order	Method (2) error	order
5		–		–
10				
20				
40				

《数值分析》之

函数逼近

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

向量范数

向量范数定义

映射: $\|\cdot\|: R^n \to [0,+\infty)$满足:

1. 非负性: $\forall x \in R^n,\ \|x\| \geq 0,\ x = 0 \Leftrightarrow \|x\| = 0.$
2. 齐次性: $\forall x \in R^n,\ a \in R,\ \|ax\| = |a|\|x\|$
3. 三角不等式: $\forall x, \mathbf{y} \in R^n,\ \|x + \mathbf{y}\| \leq \|x\| + \|\mathbf{y}\|$

称该映射为向量的一种范数.

常见向量范数

1. 1范数: $\|x\|_1 = \sum_{i=1}^{n} |x_i|$
2. 2范数: $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$
3. ∞范数: $\|x\|_\infty = \max_{1 \leq i \leq n} \{|x_i|\}$

范数的等价性

设$R_1(x)$和$R_2(x)$为任意两种范数，则存在与x无关的正常数C_1和C_2，使得

$$C_1 R_2(x) \leq R_1(x) \leq C_2 R_2(x), \quad \forall x \in R^n$$

常用范数的等价关系

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$$
$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$
$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$$

离散内积

定义：函数f，g的关于离散点列$\{x_i\}_{i=0}^n$的离散内积为：

$$(f,g)_h = \sum_{i=0}^{n} f(x_i)g(x_i)$$

离散范数

定义：函数f 的离散范数为

$$\|f\|_h = \sqrt{(f,f)_h}$$

该种内积，范数的定义与向量的2范数一致。

还可以定义函数的离散范数为：

$$\|f\|_h = \max_{1\le i\le n}\{f(x_i)\}, \quad \|f\|_h = \sum_{i=1}^{n}|f(x_i)|$$

曲线拟合

- 给出一组离散点，确定一个函数逼近原函数
- 离散数据通常是由观察或者测试得到的，不可避免会有误差
- 需要一种新的逼近原函数的手段：
  1. 不要求过所有的点（可以消除误差影响）
  2. 尽可能表现数据的趋势，靠近这些点
- 需要在给定函数空间Φ上找到函数ϕ，使得ϕ到f的距离最小。函数$\phi(x)$称为$f(x)$在空间Φ上的拟合曲线。
- 曲线拟合在实际中有广泛的应用，特别是在实验、统计等方面。
  1. 根据实验或观察得到数据，将数据在平面上标出，然后确定拟合曲线的类型
  2. 拟合曲线的类型已知，需要确定曲线的具体参数

曲线拟合的最小二乘问题

定义

$f(x)$为定义在区间$[a,b]$上的函数，$\{x_i\}_{i=0}^m$为区间上$m+1$个互不相同的点，Φ为给定的某一函数类。求Φ上的函数$\phi(x)$满足$f(x)$和$\phi(x)$在给定的$m+1$点上的距离最小，如果这种距离取为$2-$范数的话，称为最小二乘问题。即：求$\phi(x)\in\Phi$，使得

$$R_2=\sqrt{\sum_{i=0}^{m}(\phi(x_i)-f(x_i))^2}$$

最小。

最小二乘问题的求解

设$\Phi = \mathrm{span}\{\varphi_0, \varphi_1, \cdots \varphi_n\}$,

$$\phi(x) = a_0\varphi_0(x) + a_1\varphi_1(x) \cdots + a_n\varphi_n(x)$$

则最小二乘问题为

$$\|f(x) - (a_0\varphi_0(x) + a_1\varphi_1(x) \cdots + a_n\varphi_n(x))\|_h$$

关于系数$\{a_0, a_1, \cdots, a_n\}$最小。

$$
\begin{aligned}
&\|f(x)-(a_0\varphi_0(x)+a_1\varphi_1(x)\cdots+a_n\varphi_n(x))\|_h^2\\
=&\|f\|_h^2-2(f,a_0\varphi_0(x)+a_1\varphi_1(x)\cdots+a_n\varphi_n(x))_h\\
&+\|a_0\varphi_0(x)+a_1\varphi_1(x)\cdots+a_n\varphi_n(x)\|_h^2\\
=&\|f\|_h^2-2\sum_{k=0}^{n}a_k(f,\varphi_k)_h+\sum_{i,k=0}^{n}a_ia_k(\varphi_i,\varphi_k)_h\\
=&Q(a_0,a_1,\cdots,a_n)
\end{aligned}
$$

由于它关于系数$\{a_0,a_1,\cdots,a_n\}$最小,因此有

$$
\frac{\partial Q}{\partial a_i}=0,\quad i=0,1,\cdots,n
$$

$$
\text{i.e.}\quad \sum_{k=0}^{n}a_k(\varphi_i,\varphi_k)_h=(f,\varphi_i)_h,\quad i=0,1,\cdots,n
$$

写成矩阵形式有：

$$\begin{pmatrix} (\varphi_0,\varphi_0)_h & \cdots & (\varphi_0,\varphi_n)_h \\ \vdots & \ddots & \vdots \\ (\varphi_n,\varphi_0)_h & \cdots & (\varphi_n,\varphi_n)_h \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} (f,\varphi_0)_h \\ \vdots \\ (f,\varphi_n)_h \end{pmatrix}$$

称为拟合曲线的法方程。由$\{\varphi_0,\varphi_1,\cdots,\varphi_n\}$的线性无关性，知道该方程存在唯一解。

注

法方程的系数矩阵是病态的，即在实际求解中，舍入误差会引起解的较大误差，因此在计算机上可用双精度计算。

线性拟合

取Φ为线性多项式空间，函数空间的基为$\{1, x\}$，拟合曲线为$y = a + bx$, 则法方程为

$$\begin{pmatrix} (1,1)_h & (1,x)_h \\ (x,1)_h & (x,x)_h \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} (f,1)_h \\ (f,x)_h \end{pmatrix}$$

二次拟合

取Φ为二次多项式空间，函数空间的基为$\{1, x, x^2\}$，拟合曲线为$y = a_0 + a_1 x + a_2 x^2$, 则法方程为

$$\begin{pmatrix} (1,1)_h & (1,x)_h & (1,x^2)_h \\ (x,1)_h & (x,x)_h & (x,x^2)_h \\ (x^2,1)_h & (x^2,x)_h & (x^2,x^2)_h \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} (f,1)_h \\ (f,x)_h \\ (f,x^2)_h \end{pmatrix}$$

形如ae^{bx}拟合

取函数空间$\Phi = \{ae^{bx}, a, b \in R\}$，该函数空间并不构成线性空间，不易得到平方误差极小意义下的拟合曲线$y = ae^{bx}$。但由

$$\ln y = \ln a + bx$$

可以先做$y^* = a^* + bx$,由此得到

$$y = e^{y^*}$$

则法方程为

$$\begin{pmatrix} (1,1)_h & (1,x)_h \\ (x,1)_h & (x,x)_h \end{pmatrix} \begin{pmatrix} a^* \\ b \end{pmatrix} = \begin{pmatrix} (f,1)_h \\ (f,x)_h \end{pmatrix}$$

求如下数据的最小二乘拟合曲线

x_i	1	2	3	4	5	6	7	8	9	10
y_i	2	3	5	7	11	13	17	19	23	29

- 线性拟合

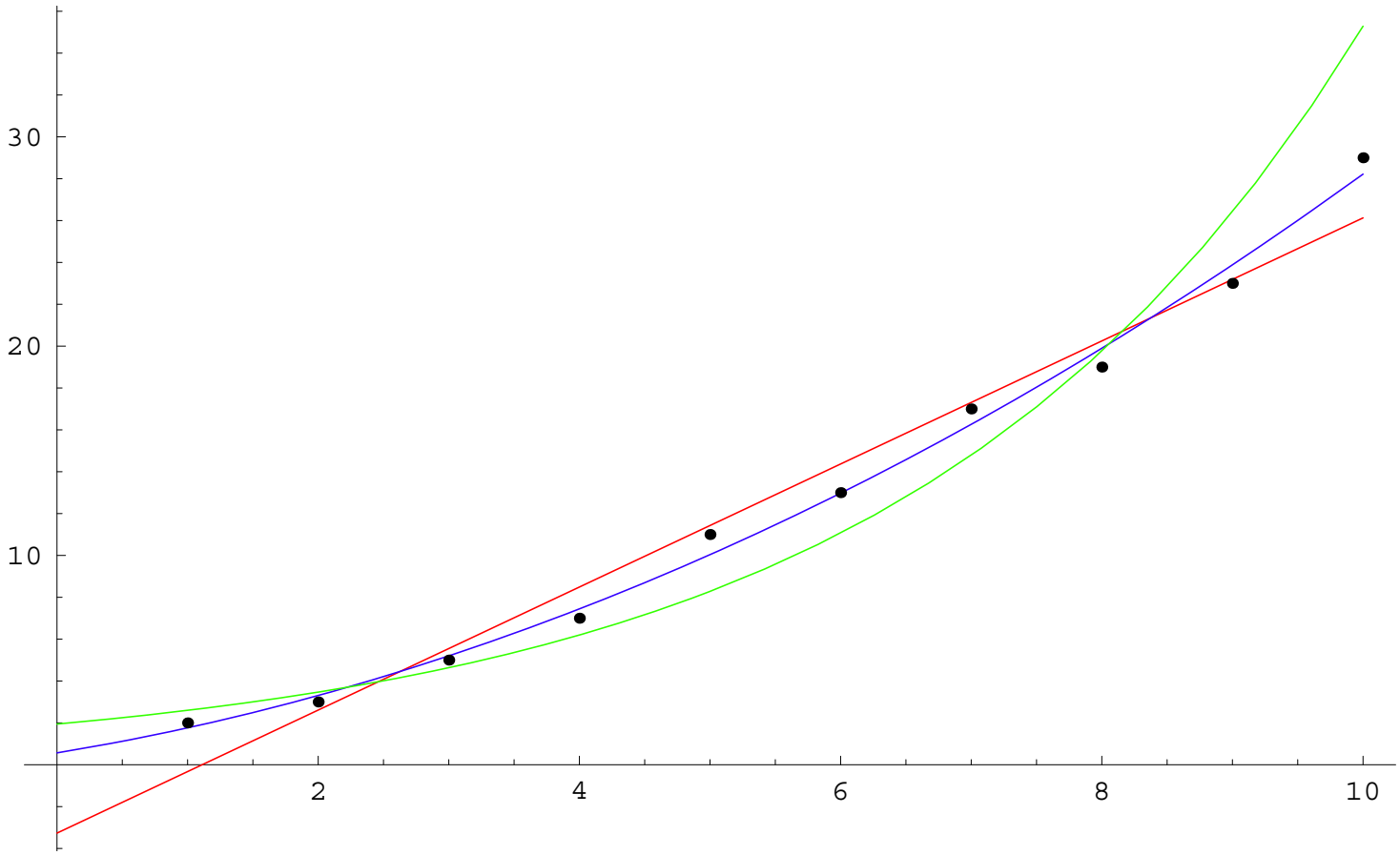
$$y = -3.26667 + 2.93939x$$

- 二次拟合

$$y = 0.566667 + 1.02273x + 0.174242x^2$$

- ae^{bx}拟合

$$y^* = 0.664723 + 0.289876x$$


30
20
10
2
4
6
8
10

矛盾方程组

给定数据序列(x_i, y_i), $i = 1, 2, \cdots, m$，作拟合直线$p(x) = a + bx$。如果要求直线$p(x)$通过这些点，则有

$$p(x_i) = a + bx_i = y_i, \quad i = 1, 2, \cdots, m$$

写成矩阵形式有

$$\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

若秩$(A, b) \neq$秩A，则方程组$Ax = b$无解，该方程组被称为矛盾方程组。

矛盾方程组的求解

求解一个矛盾方程组

$$\underset{m\times n}{A}\quad\underset{n\times 1}{x}\quad=\quad\underset{m\times 1}{b}$$

$m > n$, 计算的是在均方误差$\|Ax-b\|_h$极小意义下的解，也就是最小二乘问题。

定理

1. A为$m\times n$矩阵，b为$m\times 1$列向量，$A^TAx = A^Tb$成为方程$Ax = b$的法方程，法方程恒有解。
2. $A^TAx = A^Tb \Longleftrightarrow \|Ax-b\|_h = \min\limits_{\mathbf{y}\in R^n}\|A\mathbf{y}-b\|_h$

曲线拟合与矛盾方程组的求解

对离散数据$(x_i, y_i), i = 1, 2, \cdots, m$作$n$次多项式曲线拟合，即求解

$$Q(a_0, a_1, \cdots, a_n) = \sum_{i=1}^{m} (a_0 + a_1 x_i + \cdots + a_n x_i^n - y_i)^2$$

的极小问题与求解矛盾方程组$A\alpha = \mathbf{y}$是等价的，其中

$$A = \begin{pmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & \cdots & x_m^n \end{pmatrix}, \quad \alpha = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{pmatrix}$$

函数逼近

- 函数逼近是与函数插值理论同时发展起来的。其经典问题为：已知区间$[a,b]$上的连续函数f，对某一固定整数n，找一个次数至多是n次的多项式p，使其与f的偏差最小。这里的偏差可以有不同的定义，如最大值或平方积分。
- 一般的函数逼近问题是：给定一个赋范函数空间E以及它的一个子空间G。若$f \in E$，计算$p \in G$使得$\|f-p\|$最小。同样这里的$\|\cdot\|$定义也可以有多种选择。所得到的p称为f的在$\|\cdot\|$ 意义下的最佳逼近
- 函数逼近中比较成熟的理论是单变量函数的最小二乘理论和Tchebyshev理论

最佳逼近的存在性和唯一性

Theorem

若G是E的一个有限维子空间，则E的每一个元素在G中至少有一个最佳逼近。

证明：给定$f \in E$，则f在G中最佳逼近的候选者g必定在下述集合中：

$$K = \{g \in G : \|g - f\| \leqslant \|f\|\}$$

K为有界闭集，而G是有限维的，因此K是紧集。而泛函$g \mapsto \|f - g\|$是连续的，因此根据紧集上的连续实值函数能达到下确界得证定理。 □

Theorem

设G是内积空间E的子空间。对$f \in E, g \in G$，下列性质等价：

1. g是G中f的一个最佳逼近
2. $f - g \perp G$

证明：若$f - g \perp G$，对任一$h \in G$，

$$\|f-h\|^2 = \|(f-g)+(g-h)\|^2 = \|f-g\|^2 + \|g-h\|^2 \geqslant \|f-g\|^2$$

反之，设g是f的一个最佳逼近。再设$h \in G, \lambda > 0$,

$$\begin{aligned} 0 &\leqslant \|f-g+\lambda h\|^2 - \|f-g\|^2 \\ &= \lambda\{2\langle f-g, h\rangle + \lambda\|h\|^2\} \end{aligned}$$

令$\lambda \to 0+$, 得到$\langle f-g, h\rangle \geqslant 0$. 类似地，$\langle f-g, -h\rangle \geqslant 0$。所以$\langle f-g, h\rangle = 0$，即$f - g \perp G$. □

最佳逼近元是唯一的

Theorem

设G是内积空间E的子空间，则$f \in E$在G中的最佳逼近元是唯一的。

证明：若g_1和g_2同时是f在G中的最佳逼近元，而且$g_1 \neq g_2$，则$\|g_1 - g_2\| > 0$以及$f - g_1 \perp g_2$.

$$\|f-g_2\|^2 = \|(f-g_1)+(g_1-g_2)\|^2 = \|f-g_1\|^2+\|g_1-g_2\|^2 > \|f-g_1\|^2$$

这与g_2也为最佳逼近矛盾。 □

- 设$\{u_1, \ldots, u_n\}$是子空间G的一组基，为了计算f在G中的最佳逼近u，待定$u = \sum_{j=1}^{n} c_j u_j$. $u - f \perp G$等价于$\langle u - f, u_i \rangle = 0,\ i = 1, \ldots, n.$ 由此得到方程组

$$\sum_{j=1}^{n} c_j \langle u_j, u_i \rangle = \langle f, u_i \rangle$$

这是一个含有n个未知数，n个线性方程的方程组。其系数矩阵$G = (\langle u_i, u_j \rangle)$称为Gram矩阵。

Theorem

*Gram*矩阵为对称正定阵。

计算函数$f(x)=\sin x$在空间$\mathrm{span}\{x,x^3,x^5\}$中的最佳逼近。所用范数为

$$\|f\|=\left(\int_{-1}^{1}f^2(x)dx\right)^{1/2}$$

解：令$g_1(x)=x$, $g_2(x)=x^3$, $g_3(x)=x^5$. 待定最佳逼近元为$g(x)=c_1x+c_2x^3+c_3x^5$, 则由$\langle g-f,g_i\rangle=0$得到如下方程组：

$$c_1\langle g_1,g_i\rangle+c_2\langle g_2,g_i\rangle+c_3\langle g_3,g_i\rangle=\langle f,g_i\rangle, i=1,2,3$$

即

$$\begin{pmatrix}1/3 & 1/5 & 1/7\\ 1/5 & 1/7 & 1/9\\ 1/7 & 1/9 & 1/11\end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ c_3\end{pmatrix}=\begin{pmatrix}\sin 1-\cos 1\\ -3\sin 1+5\cos 1\\ 65\sin 1-101\cos 1\end{pmatrix}$$

系数矩阵为Hilbert矩阵，一个著名的病态矩阵。

标准正交基

- 根据幂基计算最佳逼近元，计算过程的稳定性不好。而下面的定理说明标准正交基的优势

Theorem

设G的标准正交基为$\{g_1, g_2, \ldots, g_n\}$, $f \in E$.
则$g = \sum_{i=1}^{n} c_i g_i$为$f$在$E$中最佳逼近当且仅当$c_i = \langle f, g_i \rangle$.

证明：$g = \sum_{i=1}^{n} c_i g_i$为$f$在$E$中最佳逼近$\Longleftrightarrow f - g \perp G \Longleftrightarrow$
$f - g \perp g_i$, $i = 1, 2, \ldots, n$.

$$\begin{aligned}\left\langle f - \sum_{i=1}^{n} c_i g_i, g_j \right\rangle &= \langle f, g_j \rangle - \sum_{i=1}^{n} c_i \langle g_i, g_j \rangle \\ &= \langle f, g_j \rangle - c_j = 0\end{aligned}$$

标准正交基(续)

- 可以应用Gram-Schmidt过程把一般的基转化为标准正交基
- 前例中$\{x, x^3, x^5\}$的转化结果为$\{x/\sqrt{2/3}, (5x^3 - 3x)/(2\sqrt{2/7}), (63x^5 - 70x^3 + 15x)/(8\sqrt{2/11})\}$
- 如果内积定义满足$\langle fg, h\rangle = \langle f, gh\rangle$，从单项式函数$1, x, \ldots$出发，应用Gram-Schimidt过程的结果称为正交多项式
- 常用的内积

$$\langle f, g\rangle = \int_a^b f(x)g(x)w(x)dx$$

满足上述要求

Theorem

如下定义的多项式序列是正交的：

$$p_n(x) = (x - a_n)p_{n-1}(x) - b_n p_{n-2}(x), \quad n \geqslant 2$$

其中 $p_0(x) = 1$, $p_1(x) = x - a_1$,

$$a_n = \frac{\langle xp_{n-1}(x), p_{n-1}(x)\rangle}{\langle p_{n-1}(x), p_{n-1}(x)\rangle}$$

$$b_n = \frac{\langle xp_{n-1}(x), p_{n-2}(x)\rangle}{\langle p_{n-2}(x), p_{n-2}(x)\rangle}$$

所用的内积满足 $\langle fg, h\rangle = \langle f, gh\rangle$

证明：由归纳定义可知每个p_n都是首一n次多项式，因此a_n和b_n的定义中分母不为零。

下面对n归纳证明：$\langle p_n, p_i\rangle = 0,\ i = 0, 1, \ldots, n-1.$

$n = 0$没有需要证明的。$n = 1$时由a_1的定义可以验证成立.

设$n-1$时成立，$n \geqslant 2$. 那么可以直接验证

$$\langle p_n, p_{n-1}\rangle = \langle p_n, p_{n-2}\rangle = 0$$

对$i = 0, 1, \ldots, n-3,$

$$\begin{aligned}
\langle p_n, p_i\rangle &= \langle xp_{n-1}, p_i\rangle - a_n\langle p_{n-1}, p_i\rangle - b_n\langle p_{n-2}, p_i\rangle \\
&= \langle p_{n-1}, xp_i\rangle \\
&= \begin{cases} \langle p_{n-1}, p_{i+1} + a_{i+1}p_i + b_{i+1}p_{i-1}\rangle = 0 & i \geqslant 1 \\ \langle p_{n-1}, p_1 + a_1 p_0\rangle = 0 & i = 0 \end{cases}
\end{aligned}$$

□

Legendre多项式

- 当内积定义为

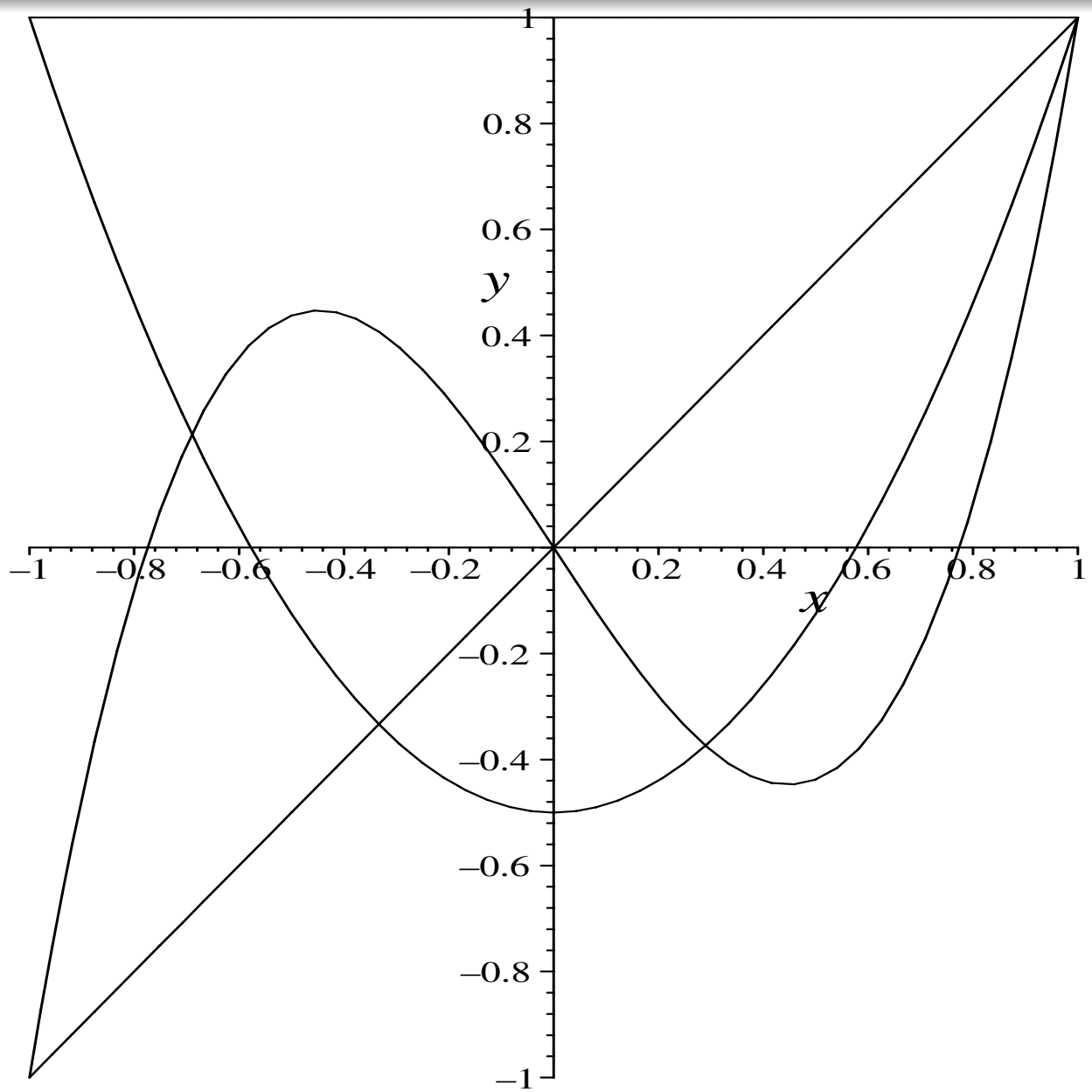
$$\int_{-1}^{1} f(x)g(x)dx$$

时生成的正交多项式称为Legendre多项式

- 前几个多项式为

$$p_0(x) = 1$$
$$p_1(x) = x$$
$$p_2(x) = x^2 - \frac{1}{3}$$
$$p_3(x) = x^3 - \frac{3}{5}x$$
$$p_4(x) = x^4 - \frac{6}{7}x^2 + \frac{3}{35}$$
$$p_5(x) = x^5 - \frac{10}{9}x^3 + \frac{5}{21}x$$

Legendre多项式



Tchebyshev多项式与Jacobian多项式

- 应用内积

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\frac{dx}{\sqrt{1-x^2}}$$

时对应的正交多项式为Tchebyshev多项式

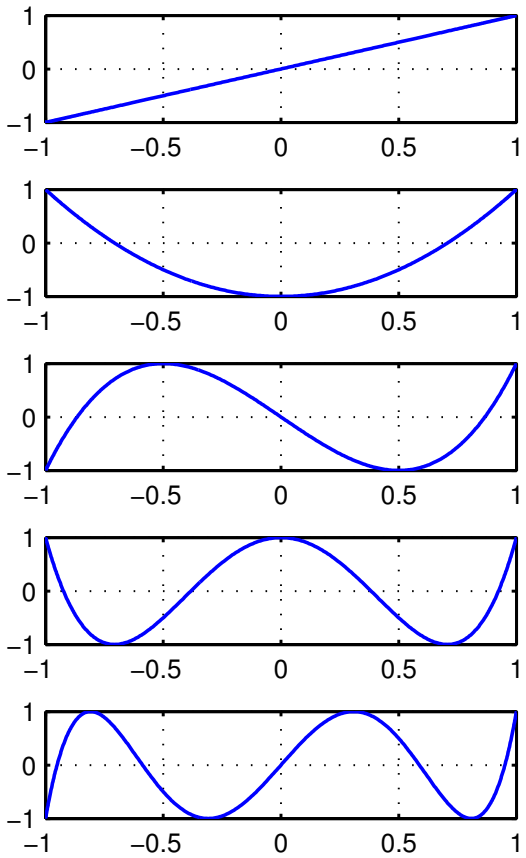
- 应用内积

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)(1-x)^{\alpha}(1+x)^{\beta}dx$$

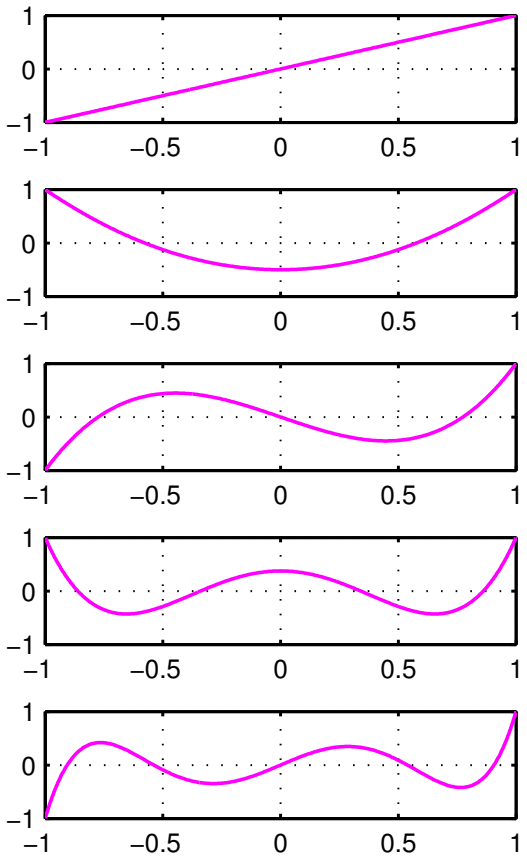
时对应的正交多项式为Jacobian多项式

Tchebyshev多项式与Legendre多项式

Chebyshev



Legendre



计算方法

- 给定$u = \sum_{k=0}^{n} c_k p_k$，其中$p_i$为某一正交多项式，那么可以应用下述算法计算$u(x)$的值：

$$
\begin{aligned}
&d_{n+2} \leftarrow 0;\ d_{n+1} \leftarrow 0 \\
&\texttt{for}\ \ k = n\ \texttt{to}\ \ 0\ \texttt{step}\ \ -1\ \texttt{do} \\
&\quad d_k \leftarrow c_k + (x - a_{k+1})d_{k+1} - b_{k+2}d_{k+2} \\
&\texttt{end do}
\end{aligned}
$$

- 有效性验证：

$$
\begin{aligned}
u(x) &= \sum_{k=0}^{n} [d_k - (x - a_{k+1})d_{k+1} + b_{k+2}d_{k+2}]p_k(x) \\
&= d_0 p_0(x) + d_1[p_1(x) - (x - a_1)p_0(x)] \\
&+ \sum_{k=2}^{n} d_k[p_k(x) - (x - a_k)p_{k-1}(x) + b_k p_{k-2}(x)] \\
&= d_0
\end{aligned}
$$

Theorem

前面定义的正交多项式p_n是所有的首一n次多项式中范数最小的。

证明：任意首一n次多项式可以写作$q = p_n - \sum_{i=0}^{n-1} c_i p_i$. $\|q\|$具有最小范数相当于在Π_{n-1}空间中寻找p_n的最佳逼近。因此应当有$q \perp \Pi_{n-1}$, 从而需选取$c_i = 0$. □

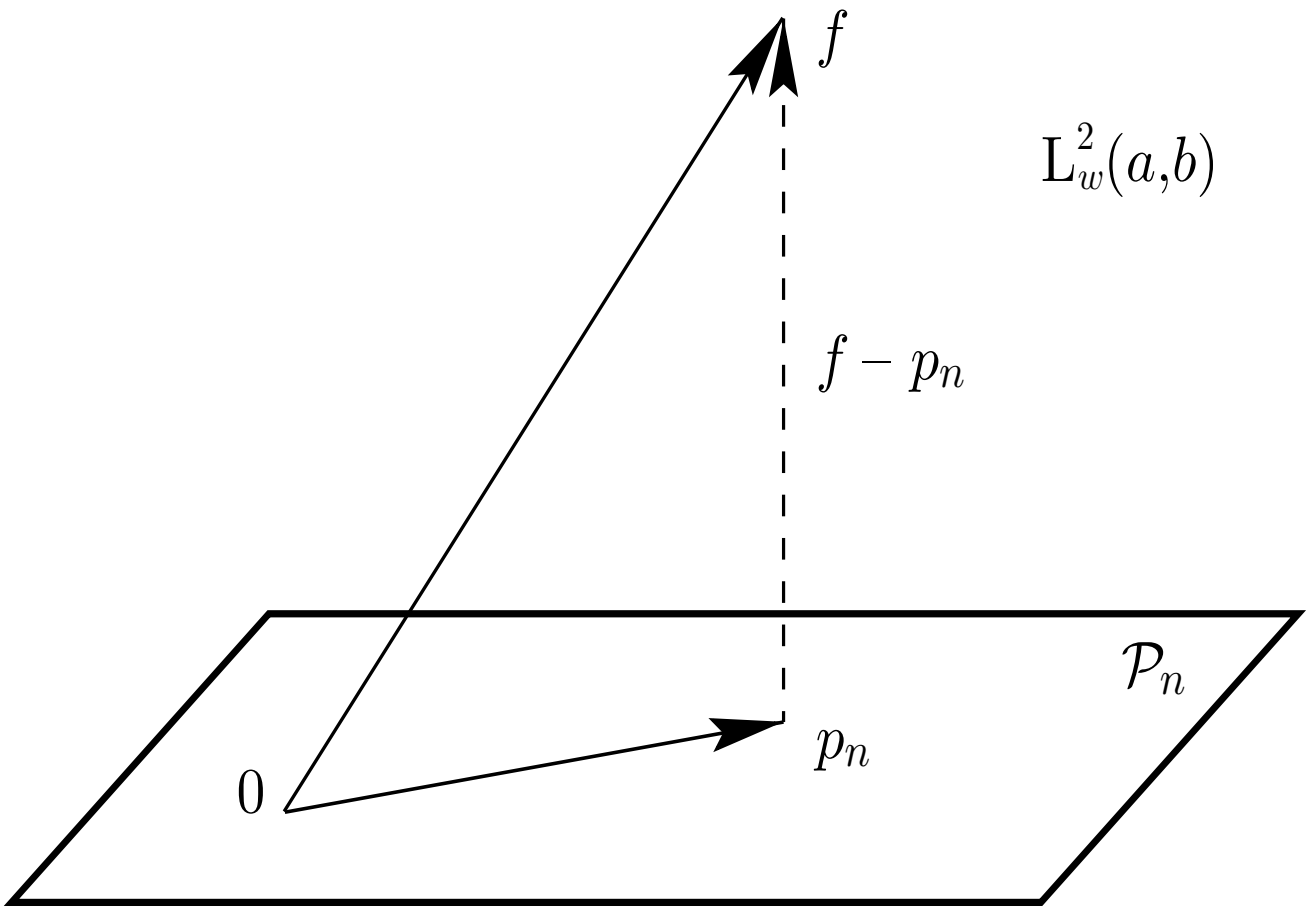
正交投影

- 给定空间G的一组标准正交基$[u_1, \ldots, u_n]$, 定义正交投影算子：

$$P_n f = \sum_{i=1}^{n} \langle f, u_i \rangle u_i$$

- 算子P_n具有如下性质：
  1. P_n为E到G的线性映射
  2. $P_n^2 = P_n$, 因此称为投影算子
  3. $f - P_n f \perp G$
  4. $P_n f$是f在G中的最佳逼近
  5. 每个P_n都是自伴的，即$\langle P_n f, g \rangle = \langle f, P_n g \rangle$

正交投影



最佳逼近

- 考虑定义在给定拓扑空间X上的全体实值连续函数形成的空间$C(X)$，这里X为紧的Hausdorff空间
- 定义范数为

$$\|f\| = \max_{x \in X} |f(x)|$$

则$C(X)$成为一个赋范空间(从而是Banach空间)

- $C(X)$中的最佳逼近问题为：给定$f \in C(X)$以及$C(X)$的一个有限维子空间G, 计算$g \in G$使得

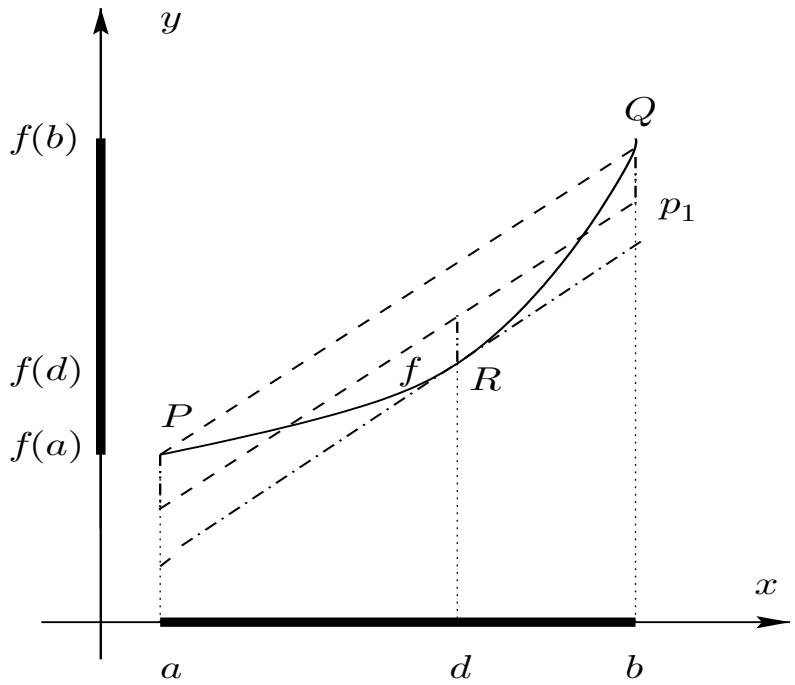
$$\|f - g\| = \operatorname{dist}(f, G) := \inf_{\bar{g} \in G} \|f - \bar{g}\|$$

- 由上节“最佳逼近存在性定理”可知g是存在的

例:用Π_1中元素在区间$[a,b]$上最佳逼近$f(x)\in C^2[a,b]$

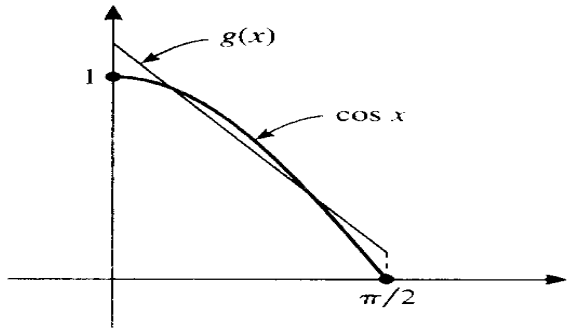
- 线性函数为了成为最佳逼近存在3个极大偏差的点，它们是a, b以及其间的一点d, 记极大偏差为δ, 最佳逼近为$p(x)$, 则应有

$$
\begin{aligned}
p(a)-f(a)&=\delta\\
p(d)-f(d)&=-\delta\\
p(b)-f(b)&=\delta\\
p'(d)-f'(d)&=0
\end{aligned}
$$



- 用Π_1中元素在区间$[0, \pi/2]$上最佳逼近$f(x) = \cos x$

如左所示线性函数为了成为最佳逼近，还需要向下移动一些，从而减小极大偏差。此外，其斜率还可以调整。为了成为最佳逼近，存在3个极大偏差的点，它们是$0, \pi/2$以及其间的一点ξ, 记极大偏差为δ, 最佳逼近为$g(x)$, 则应有



$$
\begin{aligned}
g(0) - f(0) &= \delta \\
g(\xi) - f(\xi) &= -\delta \\
g(\pi/2) - f(\pi/2) &= \delta \\
g'(\xi) - f'(\xi) &= 0
\end{aligned}
$$

《数值分析》之

函数逼近

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

- Lambert于1770年给出：

$$\arctan x = \cfrac{x}{1+\cfrac{x^2}{3+\cfrac{4x^2}{5+\cfrac{9x^2}{7+\cfrac{16x^2}{9+\cdots}}}}}, \quad |x| < 1$$

也写作

$$\arctan x = \frac{x}{1+}\frac{x^2}{3+}\frac{4x^2}{5+}\frac{9x^2}{7+}\frac{16x^2}{9+}\cdots, \quad |x| < 1$$

渐近分式

- 在连分式中第n项后终止的表达式$f_n(x)$称为原连分式的n次渐近分式, 如对前例,

$$f_n(x) = \frac{x}{1+}\frac{x^2}{3+}\frac{4x^2}{5+}\cdots\frac{(n-1)^2x^2}{2n-1}$$

- 渐近效果示例：$x = 1/\sqrt{3}$, $\arctan x = \pi/6 \approx 0.5235987756$, $f_2(x) = 0.519615$, $f_3(x) = 0.523892$, $f_4(x) = 0.523577$, $f_5(x) = 0.523600$, $f_6(x) = 0.523599$, $f_7(x) = 0.523599$

连分式的计算

- 连分式的计算不像无穷级数的计算那样简单，后者只需要用部分和代替即可，而部分和的计算很容易形成递归形式
- 连分式：$C = \dfrac{a_1}{b_1+}\dfrac{a_2}{b_2+}\dfrac{a_3}{b_3+}\cdots$ 是由序列$\{a_n\}_{n=1}^{\infty}$ 和$\{b_n\}_{n=1}^{\infty}$ 确定的
- 令

$$C_n = \frac{a_1}{b_1+}\frac{a_2}{b_2+}\frac{a_3}{b_3+}\cdots\frac{a_{n-1}}{b_{n-1}+}\frac{a_n}{b_n}$$

则C_n为给定连分式的一个近似。我们的目标是找到一个计算C_n的渐近公式

递归公式

- 定义

$$\begin{cases} A_0 = 0, \quad A_1 = a_1 \\ A_n = b_n A_{n-1} + a_n A_{n-2} \quad n \geqslant 2 \end{cases}$$

$$\begin{cases} B_0 = 1, \quad B_1 = b_1 \\ B_n = b_n B_{n-1} + a_n B_{n-2} \quad n \geqslant 2 \end{cases}$$

则

$$C_n = \frac{A_n}{B_n}$$

级数到连分式的转换

- 数学中许多重要的特殊函数都有连分式展开。

Theorem

$$\sum_{k=1}^{\infty}\frac{1}{x_k}=\frac{1}{x_1-}\frac{x_1^2}{x_1+x_2-}\frac{x_2^2}{x_2+x_3-}\cdots\frac{x_{n-1}^2}{x_{n-1}+x_n-}\cdots$$

证明：归纳法。 □

- 连分式的表示是不唯一的

《数值分析》之

函数逼近

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

周期函数的插值

- 多项式函数不适合插值周期函数
- 如果函数的周期是2π, 那么$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots$的线性组合是比较适当的插值函数
- Fourier分析：若f是周期为2π的函数，具有连续的一阶导数，那么

$$\frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx)$$

一致收敛于f, 其中

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos ktdt$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin ktdt$$

周期函数的插值: $f(x) = e^{\sin x} \sin x$

n	$\|f - p_n\|_\infty$	n	$\|f - p_n\|_\infty$
1	$1.16E+00$	8	$2.01E-07$
2	$2.99E-01$	9	$1.10E-08$
3	$4.62E-02$	10	$5.53E-10$
4	$5.67E-03$	11	$2.50E-11$
5	$5.57E-04$	12	$1.04E-12$
6	$4.57E-05$	13	$4.01E-14$
7	$3.24E-06$	14	$2.22E-15$

内积与伪内积

- 复Hilbert空间$L_2[-\pi,\pi]$中的内积定义为

$$\langle f,g\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\overline{g(x)}dx$$

$\{E_k(x)=e^{ikx}\}$构成它的一组标准正交基

- 定义

$$\langle f,g\rangle_N=\frac{1}{N}\sum_{j=0}^{N-1}f(2\pi j/N)\overline{g(2\pi j/N)}$$

此时由$\langle f,f\rangle_N=0$无法得出$f=0$, 但它满足内积定义的其它性质, 如：非负性，共扼对称性，线性性，因此称为伪内积

- 伪范数

$$\|f\|_N=\sqrt{\langle f,f\rangle_N}$$

- $\|f\|_N=0$ 当且仅当$f(2\pi j/N)=0,\ j=0,\ldots,N-1$

Theorem

对任意$N \geqslant 1$,

$$\langle E_k, E_m \rangle_N = \begin{cases} 1 & N | k-m \\ 0 & \text{其它} \end{cases}$$

证明：

$$\langle E_k, E_m \rangle_N = \frac{1}{N} \sum_{j=0}^{N-1} E_k\left(\frac{2\pi j}{N}\right) \overline{E_m\left(\frac{2\pi j}{N}\right)} = \frac{1}{N} \sum_{j=0}^{N-1} \left(e^{2\pi i(k-m)/N}\right)^j$$

如果$N|k-m$, 则$e^{2\pi i(k-m)/N} = 1$, 因此得证。若$N \nmid k-m$, 则可以应用几何数列求和公式：

$$\langle E_k, E_m \rangle_N = \frac{e^{2\pi i(k-m)} - 1}{e^{2\pi i(k-m)/N} - 1} = 0$$

指数多项式

- 一个次数至多是n次的指数多项式指的是下列形式的函数

$$P(x) = \sum_{k=0}^{n} c_k e^{ikx} = \sum_{k=0}^{n} c_k (e^{ix})^k$$

Theorem

基函数$\{E_0, E_1, \ldots, E_{N-1}\}$关于伪内积是标准正交的

Theorem

在等距结点$x_j = 2\pi j/N$上插值给定函数f的次数不超过$N-1$的指数多项式由下式唯一确定：

$$P = \sum_{k=0}^{N-1} c_k E_k, \quad c_k = \langle f, E_k \rangle_N$$

证明：存在性验证

$$\begin{aligned}
\sum_{k=0}^{N-1} c_k E_k(x_v) &= \sum_{k=0}^{N-1} \langle f, E_k \rangle_N E_k(x_v) = \sum_{k=0}^{N-1} \frac{1}{N} \sum_{j=0}^{N-1} f(x_j) \overline{E_k(x_j)} E_k(x_v) \\
&= \sum_{j=0}^{N-1} f(x_j) \frac{1}{N} \sum_{k=0}^{N-1} \overline{E_j(x_k)} E_v(x_k) = \sum_{j=0}^{N-1} f(x_j) \langle E_v, E_j \rangle_N \\
&= f(x_v) \qquad (E_k(x_v) = E_v(x_k))
\end{aligned}$$

唯一性证明

设$\sum_{k=0}^{N-1} a_k E_k$是在$x_j = 2\pi j/N,\ j = 0, 1, \ldots, N-1$上插值$f$的指数多项式，则

$$\sum_{k=0}^{N-1} a_k E_k(x_j) = f(x_j), \qquad j = 0, \ldots, N-1$$

两边同乘以$\overline{E_n(x_j)}$，再对j求和，则有

$$\sum_{k=0}^{N-1} a_k \sum_{j=0}^{N-1} E_k(x_j)\overline{E_n(x_j)} = \sum_{j=0}^{N-1} f(x_j)\overline{E_n(x_j)}$$

此即

$$\sum_{k=0}^{N-1} a_k \langle E_k, E_n\rangle_N = \langle f, E_n\rangle_N$$

从而有$a_n = \langle f, E_n\rangle_N = c_n$ □

指数多项式的计算

- 直接根据$c_k = \langle f, E_k \rangle_N$计算所有的$c_k$，需要$\mathcal{O}(N^2)$次乘法和加法
- 快速Fourier变换(FFT)把这个计算成本降到$\mathcal{O}(N \log N)$.

N	N^2	$N \log_2 N$
1 024	1 048 576	10 240
4 096	16 777 216	49 152
16 384	268 435 456	229 375

Theorem

设p和q是次数不超过$n-1$的指数多项式，使得对点$x_j = \pi j/n$有

$$p(x_{2j}) = f(x_{2j}), \quad q(x_{2j}) = f(x_{2j+1}), \quad j = 0, 1, \ldots, n-1$$

则在点$x_0, x_1, \ldots, x_{2n-1}$上插值$f$的次数不超过$2n-1$的指数多项式由下式给出：

$$P(x) = \frac{1}{2}(1 + e^{inx})p(x) + \frac{1}{2}(1 - e^{inx})q(x - \pi/n)$$

证明：由于e^{inx}是n次的，所以$P(x)$的次数不超过$2n-1$. 插值性可以直接验证。 □

指数多项式的系数定理

Theorem

对于指数多项式定理中给出的多项式系数设为

$$p=\sum_{j=0}^{n-1}\alpha_j E_j \quad q=\sum_{j=0}^{n-1}\beta_j E_j \quad P=\sum_{j=0}^{2n-1}\gamma_j E_j$$

则对于$0\leqslant j\leqslant n-1$, 有

$$\gamma_j=\frac{1}{2}\alpha_j+\frac{1}{2}e^{-ij\pi/n}\beta_j$$
$$\gamma_{j+n}=\frac{1}{2}\alpha_j-\frac{1}{2}e^{-ij\pi/n}\beta_j$$

运算次数

- 设$R(n)$表示计算点集$\{2\pi j/n : 0 \leqslant j \leqslant n-1\}$上插值多项式的系数所需的最小乘法运算次数
- $R(2n) \leqslant 2R(n) + 2n$
  - 分别用n次运算计算出$\frac{1}{2}\alpha_j$和$\frac{1}{2}e^{-ij\pi/n}\beta_j$
- $R(2^m) \leqslant m2^m$
  - 归纳法：$R(2^{m+1}) = R(2 \cdot 2^m) \leqslant 2R(2^m) + 2 \cdot 2^m$

指数多项式求值

- 给定指数多项式

$$p(x) = \sum_{j=0}^{n-1} a_j E_j(x)$$

计算它在$t - 2k\pi/n,\ k = 0, 1, \ldots, n-1$上的值

- 令$x_k = 2k\pi/n$,

$$\begin{aligned} p(t - x_k) &= \sum_{j=0}^{n-1} a_j E_j(t - x_k) = \sum_{j=0}^{n-1} a_j e^{ij(t-x_k)} \\ &= \sum_{j=0}^{n-1} a_j E^{ijt} \overline{E_k(x_j)} = n\langle g, E_k\rangle_n \end{aligned}$$

其中g是一个满足$g(x_j) = a_j e^{ijt},\ j = 0, 1, \ldots, n-1$的函数。这样对$g$进行FFT，得到系数值$\langle g, E_k\rangle_n$，乘以$n$以后就得到$p(t - x_k)$

《数值分析》之

函数逼近

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

多变量插值问题

- 多变量函数的光滑插值问题是相当困难的，此时会出现一些在单变量插值理论中所没有的异常特征。两变量插值与多于两变量的插值情形类似。
- 问题：给定xy平面内的插值点集合，记为

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

假设这n个点不相同。每个点(x_i, y_i)指定一个实数c_i. 我们的目的是寻找一个光滑并且容易计算的函数F满足

$$F(x_i, y_i) = c_i$$

例：二元二次多项式插值

- 考虑所有总次数不超过二的二元多项式全体，它构成的线性空间维数为6，因此在用这个空间构造插值多项式时，自然是给定六个插值结点(x_i, y_i)和函数值c_i. 在幂基$1, x, y, x^2, xy, y^2$下待定系数时对应的系数矩阵为

$$\begin{pmatrix} 1 & x_1 & y_1 & x_1^2 & x_1y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2y_2 & y_2^2 \\ 1 & x_3 & y_3 & x_3^2 & x_3y_3 & y_3^2 \\ 1 & x_4 & y_4 & x_4^2 & x_4y_4 & y_4^2 \\ 1 & x_5 & y_5 & x_5^2 & x_5y_5 & y_5^2 \\ 1 & x_6 & y_6 & x_6^2 & x_6y_6 & y_6^2 \end{pmatrix}$$

- 这个系数矩阵的行列式为零当且仅当给定的六个插值节点在一条二次曲线上。因此这时的插值问题的存在性和唯一性不是很显然
- 特别地，如果六个结点共线，那么只有当多项式的次数达到五次时才会有解

研究问题与应用

- 讨论对给定的插值基函数，如何选择插值结点，使插值多项式存在
- 在数据处理问题中，要求对给定点的值，构造恰当的插值基函数以及插值多项式
- 在代数学中把插值问题解释为代数形式的中国剩余定理，实际解决这一问题涉及到构造性代数几何。已有人应用其中的方法解决了对给定插值结点组构造插值基的问题
- 多元插值比一元插值有更广泛的应用前景。如：二元函数插值解决的是空间中曲面构造的问题，这是CAD 中形体设计所需要的。

张量积形式

- 考察单变量多项式插值的Lagrange形式：给定结点x_1, x_2, ..., x_p以及函数f, 那么插值相当于定义了一个线性算子P:

$$(Pf)(x) = \sum_{i=1}^{p} f(x_i)u_i(x), \quad u_i(x) = \prod_{\substack{j=1 \\ j\neq i}}^{p} \frac{x - x_j}{x_i - x_j}$$

- 算子P也可以推广作用在多变量函数上。如果f是x, y的函数，那么

$$(\overline{P}f)(x, y) = \sum_{i=1}^{p} f(x_i, y)u_i(x)$$

相当于在垂直线L_i上插值f的一个二元函数，
即$(\overline{P}f)(x_i, y) = f(x_i, y)$, 其中

$$L_i := \{(x_i, y) : -\infty < y < +\infty\}$$

- 假设在y方向有另外的插值结点$y_1, y_2, \ldots, y_q$, 那么可以类似定义线性算子

$$(Qf)(y) = \sum_{i=1}^{q} f(y_i)v_i(y),$$

$$(\overline{Q}f)(x, y) = \sum_{i=1}^{q} f(x, y_i)v_i(y)$$

其中$v_i(y)$为相应的Lagrange插值基函数

- 定义新的算子如下：

$$(\overline{PQ}f)(x, y) = \overline{P}(\overline{Q}f)(x, y) = \sum_{i=1}^{p}(\overline{Q}f)(x_i, y)u_i(x)$$

$$= \sum_{i=1}^{p}\sum_{j=1}^{q} f(x_i, y_j)v_j(y)u_i(x)$$

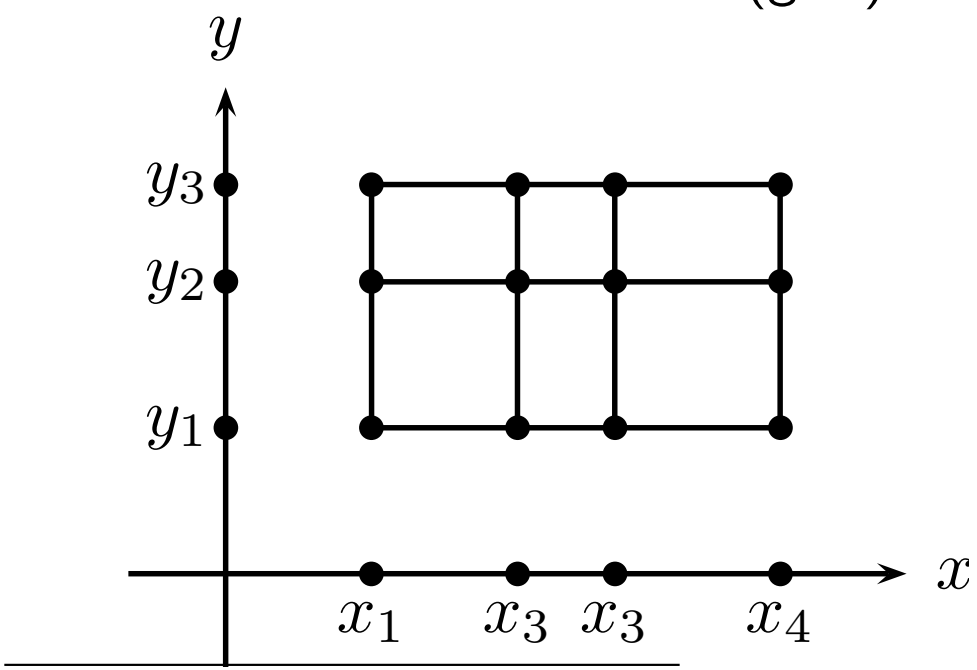
那么$(\overline{PQ}f)(x, y)$在结点(x_i, y_j)上插值于函数$f(x, y)$

- 实际上，在前面出现的基函数$u_i(x)$和$v_j(y)$可以不必是相应的Lagrange基函数，只要满足

$$u_i(x_j) = \delta_{ij}, \qquad v_i(y_j) = \delta_{ij}$$

即可，不过此时定义的结果$(\overline{PQ}f)(x, y)$并不一定是多项式

- 类似于(x_i, y_j), $i = 1, \ldots, p$, $j = 1, \ldots, q$, 形式的结点形成的阵列称为Cartesian[1]网格(grid)



[1]R. Descartes (1596–1650)的拉丁文写法为Cartesius

张量积插值

- 算子$(\overline{PQ}f)(x,y)$给出了在特殊结点上构造插值函数的一种方法，这个算子称为P和Q的张量积(tensor-product): $P \otimes Q$
- 此时，如果基本算子中采用Lagrange基函数，则相当于在Cartesian网格插值结点时应用二元多项式空间

$$\mathrm{span}\{x^i y^j : 0 \leqslant i \leqslant p-1, 0 \leqslant j \leqslant q-1\}$$

进行插值。这个空间的维数为pq，空间中多项式的次数可以记为(k,l), 其中x和y出现的最高次数分别为k和l

  - 如$p=2$, $q=2$, 则相当于给定了矩形区域四个顶点处的高度，唯一确定一个次数(1,1)的函数（称为双线性(bilinear)函数）。它是一张二次曲面（哪一种呢？）

Boolean和

- 应用算子$P \otimes Q$可以定义一个新的算子如下：

$$\begin{aligned}[(P \oplus Q)f](x,y) &= (\overline{P}f)(x,y) + (\overline{Q}f)(x,y) - (\overline{PQ}f)(x,y) \\ &= \sum_{i=1}^{p} f(x_i,y)u_i(x) + \sum_{j=1}^{q} f(x,y_j)v_j(y) \\ &\quad - \sum_{i=1}^{p}\sum_{j=1}^{q} f(x_i,y_j)u_i(x)v_j(y)\end{aligned}$$

- 这个算子称为P和Q的Boolean和，满足

$$[(P \oplus Q)f](x_i,y) = f(x_i,y), [(P \oplus Q)f](x,y_j) = f(x,y_j)$$

因此$[(P \oplus Q)f](x,y)$在所有的水平线和竖直线上插值f.

- 基于算子$P \oplus Q$可以构造在CAGD中具有重要历史地位的Coons曲面片。1964年MIT的教授Steven A. Coons提出了被后人称为超限插值(可以应用算子$P \oplus Q$进行解释)的新思想，通过插值四条任意的边界曲线来构造曲面。
- Coons方法和Bézier方法是CAGD最早的开创性工作。计算机图形学的最高奖是以Coons的名字命名的，而获得第一届（1983）和第二届(1985)Steven A. Coons 奖的，恰好是Ivan E. Sutherland和Pierre Bézier

- 由所有

$$\sum_{i=1}^{m} a_i(x)b_i(y), \qquad a_i(x) \in \Pi_k, b_i(y) \in \Pi_l$$

构成的二元多项式空间记为$\Pi_k \otimes \Pi_l$, 称为两个一元多项式空间的张量积。在这个空间中会出现总次数为$k+l$的$x^k y^l$项，但其它$k+l$次项不会出现，因此这种多项式空间并没有充分利用以总次数为限的多项式表示

- 总次数不超过k的二元多项式表示为

$$\sum_{0 \leqslant i+j \leqslant k} c_{ij} x^i y^j = \sum_{i=0}^{k} \sum_{j=0}^{k-i} c_{ij} x^i y^j$$

其全体构成的空间记为$\Pi_k(\mathbb{R}^2)$

- 空间的一组基为$x^i y^j,\ 0 \leqslant i+j \leqslant k$
  1. 它们显然生成$\Pi_k(\mathbb{R}^2)$
  2. 为证线性无关，假设

$$\sum_{i=0}^{k}\sum_{j=0}^{k-i} c_{ij}x^i y^j = \sum_{i=0}^{k}\left(\sum_{j=0}^{k-i} c_{ij}y^j\right)x^i = 0$$

这里$\sum_{j=0}^{k-i} c_{ij}y^j$可以看作是$x^i$的系数，因此由$\{1, x, \ldots, x^k\}$线性无关可得

$$\sum_{j=0}^{k-i} c_{ij}y^j = 0, \qquad i = 0, \ldots, k$$

从而得出所有系数$c_{ij} = 0$

- 空间的维数为$\binom{k+2}{2} = \frac{(k+1)(k+2)}{2}$

$\Pi_k(\mathbb{R}^2)$插值问题

- 本节开始的例子说明了应用$\Pi_k(\mathbb{R}^2)$中元素进行插值，插值结点的选取需要非常仔细
- 实际上，假设给定了n个函数$u_1, u_2, \ldots, u_n$，并且给定$\mathbb{R}^2$中n个结点$p_i = (x_i, y_i)$. 那么为了构造插值函数，需要求解线性方程组，对应的系数矩阵为$(u_j(p_i))_{n\times n}$. 若对给定的结点集，该矩阵非奇异，那么让前两个结点在$\mathbb{R}^2$平面中作连续移动，但从不重合，也不与其它结点重合，那么存在一种移动方式，两个结点相当于交换了位置，从而系数矩阵的行列式反号。根据上述移动过程与矩阵的行列式之间的连续性，可知存在一组结点集，对应的行列式为零
- 从而可知在$C(\mathbb{R}^2)$中根本没有n维子空间适合在任意n个结点的集合上进行插值。1918年Haar观察到了这一事实

插值任意数据的可能性

Theorem

空间$\Pi_k(\mathbb{R}^2)$是可以对$\mathbb{R}^2$中任意$k+1$个不同结点集上的任意数据插值的。

证明：假设被插值函数为f, 插值结点是(x_i, y_i), $i = 0, 1, \ldots, k$，则存在线性函数$\ell(x, y) = ax + by + c$使得$k+1$个数$t_i = \ell(x_i, y_i)$两两不同。根据单变量多项式插值理论，存在$p \in \Pi_k(\mathbb{R})$, 使得$p(t_i) = f(x_i, y_i)$. 显然$p \circ \ell \in \Pi_k(\mathbb{R}^2)$而且满足插值条件：

$$(p \circ \ell)(x_i, y_i) = p(\ell(x_i, y_i)) = p(t_i) = f(x_i, y_i)$$

□

形如$g \circ \ell$ $(\ell \in \Pi_1(\mathbb{R}^2))$的函数称为岭(ridge)函数，因为$g \circ \ell$在每条直线$\ell(x, y) = \lambda$上是常数，从而其图形是一张直纹面。

Newton格式

- 单变量多项式插值中的Newton格式是首先在$x_1, \ldots, x_n$上构造插值f的多项式p, 然后通过给p添加一项，使之在$x_1, \ldots, x_n, x_{n+1}$上插值$f$
- 上述过程可以抽象为：设X是一个集合，f为定义在X上的实值函数。设$\mathcal{N}$为结点集。如果p是$\mathcal{N}$上任一插值f的函数，而且q是任一在$\mathcal{N}$上取值为零的函数。如果$q(\xi) \neq 0$, 则$p^* = p + cq$给出了在$\mathcal{N} \cup \{\xi\}$上插值f的函数
- 进一步地一般化：设q是X到$\mathbb{R}$的函数，Z是它的零点集。若在$\mathcal{N} \cap Z$上p插值f, 并且在$\mathcal{N} \backslash Z$上r插值$(f - p)/q$, 则在N上$p + qr$插值f

Shepard插值

- 1968年由D. Shepard给出
- 设给定的插值结点为$p_i \in \mathbb{R}^2$, $i = 1, 2, \ldots, n$. 选取$\mathbb{R}^2 \times \mathbb{R}^2$上的一个实值函数$\phi$满足唯一性条件：

$$\phi(p, q) = 0 \text{ 当且仅当} p = q$$

  - 如$\phi(p, q) = \|p - q\|^\mu$, $\mu > 0$
- 类似于单变量Lagrange插值基函数的构造方式，定义

$$u_i(p) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{\phi(p, p_j)}{\phi(p_i, p_j)}, \qquad j = 1, 2, \ldots, n$$

  这些函数具有“基性质”：$u_i(p_j) = \delta_{ij}$
- 在结点集上插值f的函数为

$$F = \sum_{i=1}^{n} f(p_i) u_i$$

Shepard插值的变体

- 此时要求ϕ是一个非负函数，并设

$$v_i(p) = \prod_{\substack{j=1 \\ j\neq i}}^{n} \phi(p, p_j), \quad v(p) = \sum_{i=1}^{n} v_i(p), \quad w_i(p) = \frac{v_i(p)}{v(p)}$$

- 当$i \neq j$时$v_i(p_j) = 0$, 在其它除$p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_n$之外的所有点$p$上$v_i(p) > 0$。从而$v(p) > 0$，这样$w_i$有定义。根据函数的定义方式，我们有$w_i(p_j) = \delta_{ij}$, $0 \leqslant w_i(p) \leqslant 1$, $\sum_{i=1}^{n} w_i(p) = 1$. 所以下述方程定义了插值函数：

$$F = \sum_{i=1}^{n} f(p_i) w_i = \sum_{i=1}^{n} f(p_i) \frac{v_i}{v}$$

- $n = 1$的情形如何?

Shepard变体的特点

- 以下两条性质表明插值函数F继承了被插值函数的某些特征：
  - 如果数据是非负的，那么插值函数F也是非负的。
  - 如果f是常值函数，那么$F \equiv f$
- 如果ϕ可微，那么F在每个结点上都呈现出一个平坦点
  - 每个结点都是w_i的极值点，从而偏导数等于零。这样F在结点处的偏导数也等于零

令

$$\phi(x,y)=\|x-y\|^{\mu},\qquad \mu>0$$

- 当$\mu>1$时该函数可微，当$0<\mu\leqslant 1$时不可微

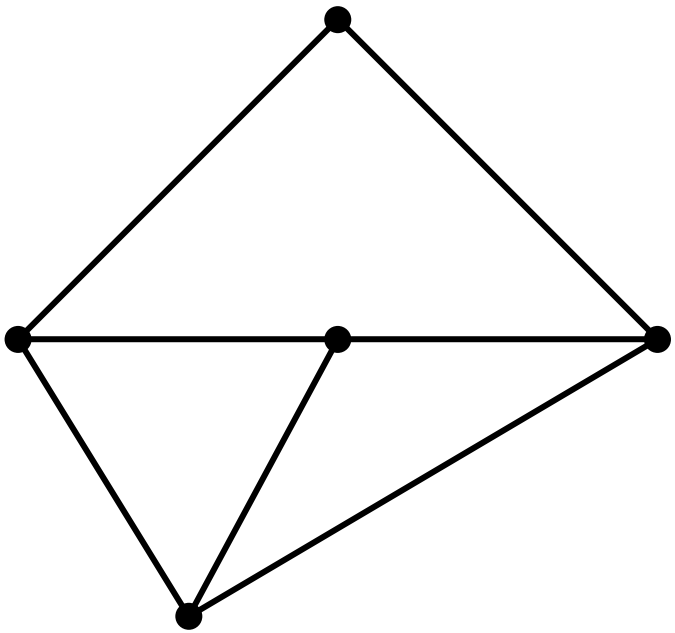
此时w_i的定义有下列等价形式：

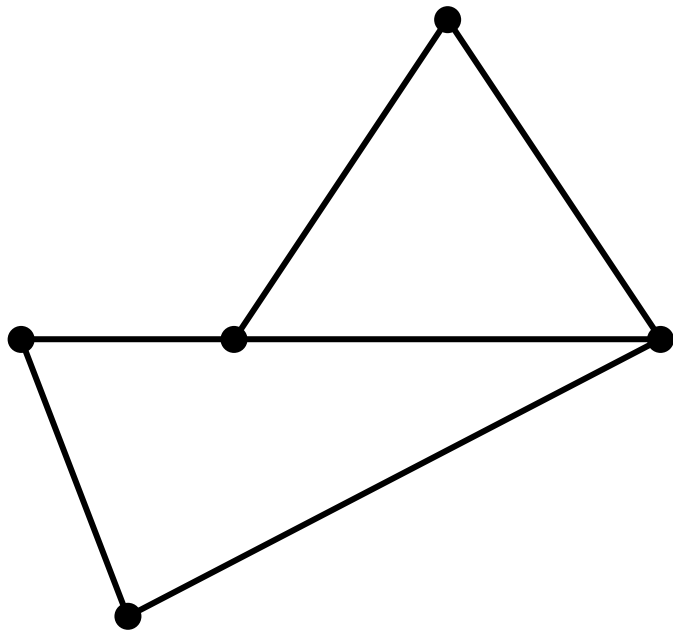
$$w_i(x)=\frac{\prod\limits_{\substack{j=1\\ j\neq i}}^{n}\|x-x_j\|^{\mu}}{\sum\limits_{k=1}^{n}\prod\limits_{\substack{j=1\\ j\neq k}}^{n}\|x-x_j\|^{\mu}}=\frac{\|x-x_i\|^{-\mu}}{\sum\limits_{j=1}^{n}\|x-x_j\|^{-\mu}}$$

后一等式中会出现∞/∞情形，需要小心应用

三角剖分

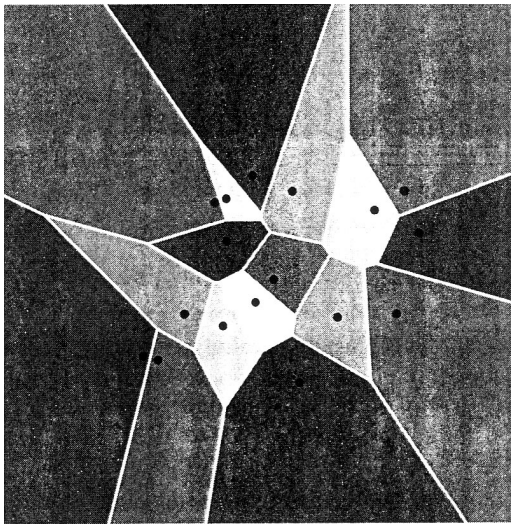
- 基于给定结点，另外一种构造插值函数的方法就是三角剖分(triangulation)，即连接结点，形成一族三角形 T_1, T_2, ..., T_m. 这些三角形满足下述规则：
  1. 每个插值结点必须是某个三角形 T_i 的顶点
  2. 每个三角形的顶点必须是结点
  3. 如果某个结点在某个三角形内，那么它一定是这个三角形的顶点
- 两个不符合规则的三角剖分

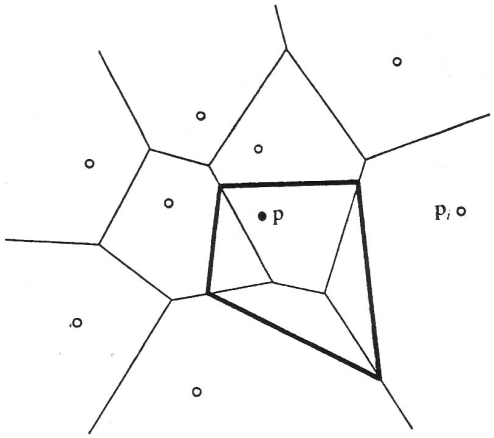


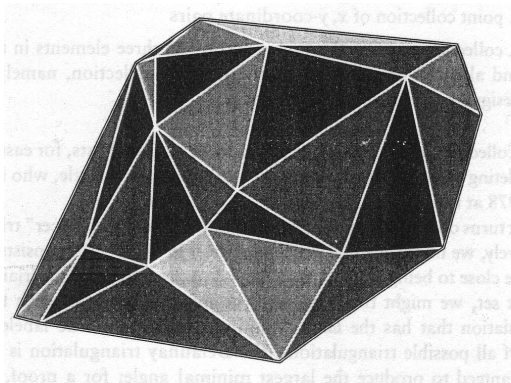


Dirichlet镶嵌与Delaunay三角剖分

为了获得给定点集的三角剖分，可以采用下述方法：

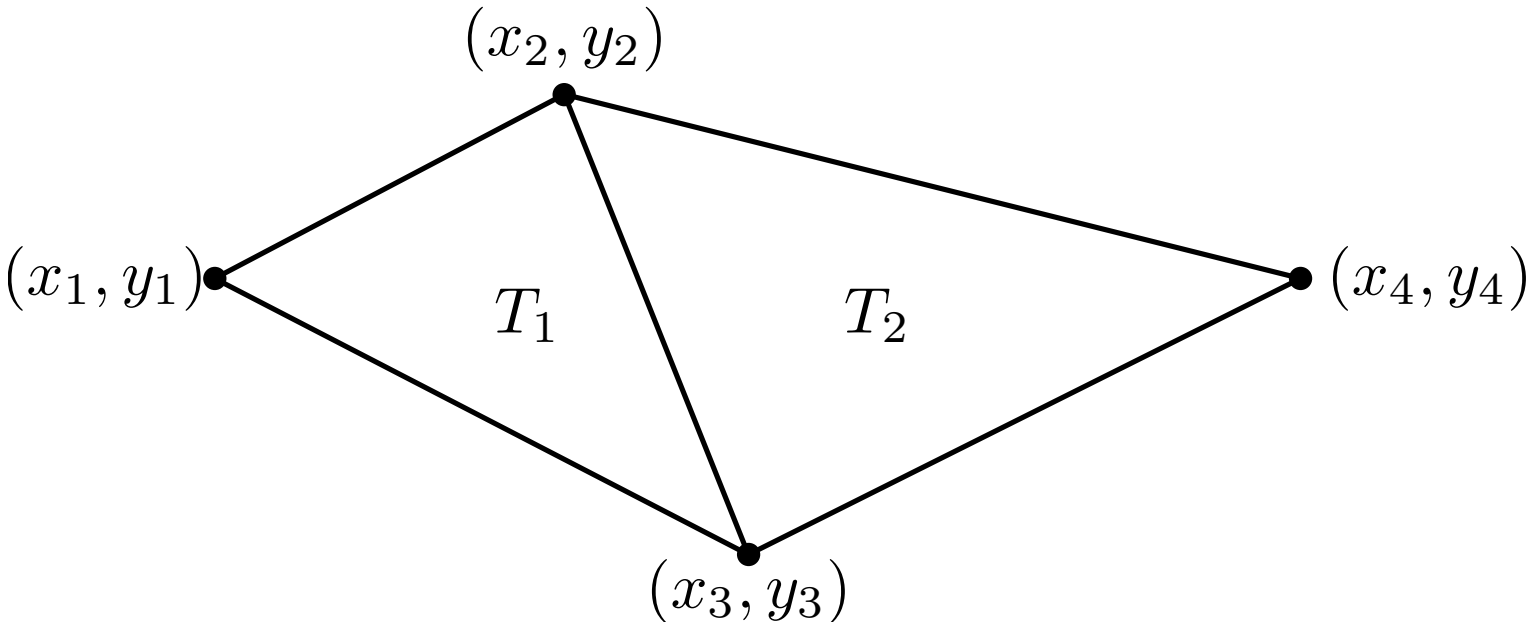






分片线性插值函数

- 对于三角剖分上最简单的插值函数就是分片线性函数，它在所有三角形的所有顶点上插值函数f，在任意三角形T_i上为一个线性函数$\ell_i(x,y) = a_i x + b_i y + c_i$，这个函数由它在三角形三个顶点的函数值唯一确定(为什么？).



- 在两个三角形的公共边界上，由于此时每个三角形中的线性函数限制在公共边上为一个单变量线性函数，由其在两个顶点的值唯一确定，因此两个线性函数沿公共边界连续拼接。从而所有三角形上的线性函数组合在一起是连续的。

三角形网格

- 三角网格在许多计算机图形应用领域中是最通用的曲面表示方式。由于它的简单和灵活，在一些注重处理性能的领域，三角网格甚至取代了传统的CAD曲面表示，如NURBS曲面
- CPU和图形硬件性能的稳定增长，廉价的内存，以及三维扫描仪的广泛使用，导致了产生大量的高精度的几何数据，其中包含几百万三角片的模型在当今已经很多

经典的最小二乘方法

- 问题：给定一个集合X作为插值函数和被插值函数f的定义域，以及其中的一组结点$x_1, \ldots, x_n$. 插值函数所在空间由$u_1, \ldots, u_m$生成，其中m相对于n很小，因此可能无法构造出插值函数。取而代之，我们希望找到系数$c_1, \ldots, c_m$, 极小化下述表达式：

$$\sum_{i=1}^{n} \left(f(x_i) - \sum_{j=1}^{m} c_j u_j(x_i) \right)^2 w_i$$

其中$w_i \geqslant 0$为权因子

- 如果记$\langle f, g \rangle = \sum_{i=1}^{n} f(x_i) g(x_i) w_i$, 那么根据内积空间中的逼近理论，$f - \sum_{i=1}^{m} c_j u_j \perp u_i$, $i = 1, 2, \ldots, m$刻划了极小化问题解的特征，从而导出方程：

$$\sum_{j=1}^{m} c_j \langle u_j, u_i \rangle = \langle f, u_i \rangle, \qquad i = 1, 2, \ldots, m$$

移动最小二乘的定义

- 移动最小二乘与经典最小二乘的区别在于允许权因子w_i是x的函数。记

$$\langle f, g\rangle_x = \sum_{i=1}^{n} f(x_i)g(x_i)w_i(x)$$

则相应的方程为

$$\sum_{j=1}^{m} c_j\langle u_j, u_i\rangle_x = \langle f, u_i\rangle_x, \qquad i = 1, 2, \ldots, m$$

最终的逼近函数为

$$g(x) = \sum_{j=1}^{m} c_j(x)u_j(x)$$

- 因为方程随x一起变化，因此当m较大时，方程难以求解。通常$m \leqslant 10$

权函数的选择

- 如果$w_i(x)$在x_i处相当很大，那么g在x_i处几乎插值。如果当x离x_i很远时$w_i(x)$很快减少为零，那么远离x_i的结点对$g(x_i)$几乎没有影响
  - $w_i(x) = \|x - x_i\|^{-2}$, 其中$\|\cdot\|$为任何范数，但欧氏范数是常用的
- 若$m = 1$, 而且$u_1(x) \equiv 1$, 那么导致Shepard 方法。此时，记$c_1(x) = c(x)$, $u(x) = u_1(x)$, 那么由$c(x)\langle u, u\rangle_x = \langle f, u\rangle_x$可解出$c$, 从而逼近函数为

$$g(x) = c(x)u(x) = c(x) = \frac{\langle f, u\rangle_x}{\langle u, u\rangle_x} = \frac{\sum_{i=1}^n f(x_i)w_i(x)}{\sum_{j=1}^n w_j(x)}$$

《数值分析》之

数值微分

徐岩

中国科学技术大学数学系

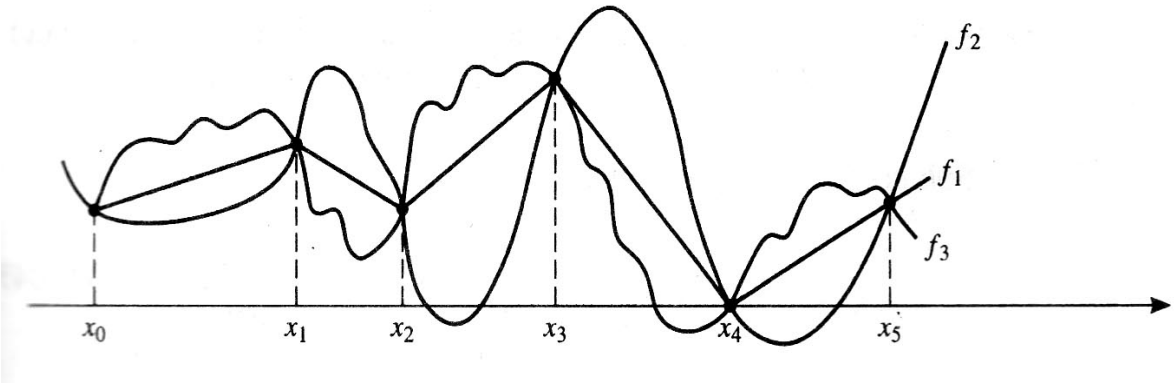
yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

内容

1. 数值微分：基于Taylor展开、基于多项式插值、Richardson外推
2. 数值积分：通过多项式插值、待定系数法、复化数值积分、Romberg积分、Guass积分

- 只给定函数f在$n+1$个点$x_0,\ldots,x_n$上的值，如何利用这些信息计算导数$f'(c)$和积分$\int_a^b f(x)dx$的值？注意过这些点的函数可以有很多，如下图所示：



基于Taylor展开计算数值微分

- Taylor展开：

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(\xi)$$

其中$\xi \in (x, x+h)$. 为了使等式成立，需要f与f'在闭区间$[x, x+h]$上连续，而且对应的开区间上存在f''

- 重新整理：

$$f'(x) = \frac{1}{h}[f(x+h) - f(x)] - \frac{h}{2}f''(\xi)$$

因此可利用的数值公式与误差项都出现在上述公式中。误差项由两部分组成：h的幂次以及f的高阶导数。因此当$h \to 0$时，误差中的h项使得整体表达式收敛于零

- $-(h/2)f''(\xi)$称为截断误差。在数值计算中，截断误差与舍入误差起着同样重要的作用

例

应用上述公式计算$f(x)=\cos x$在$x=\pi/4$点的导数，这里取$h=0.01$。它的精确度是多少？

- 数值导数值

$$\begin{aligned} f'(x) &\approx \frac{1}{h}[f(x+h)-f(x)] \\ &= \frac{1}{0.01}[0.700000476-0.707106781] \\ &= -0.71063051 \end{aligned}$$

- 精度：

$$\left|\frac{h}{2}f''(\xi)\right| = 0.005|\cos\xi| \leqslant 0.005$$

- 实际上，由于$\xi\in(\pi/4,\pi/4+h)$，所以$|\cos\xi|<0.707107$，从而给出误差界为0.0035355.
- 真正的误差为

$$-\sin\frac{\pi}{4}+0.71063051=0.003523729$$

精度分析

- 在前面的数值导数计算公式中，从截断误差$-(h/2)f''(\xi)$的表达式可见：为了精确计算$f'(x)$, 步长h必须很小
- 下面进行一个实验，其中令h通过给定的一个序列收敛到零，分别计算出相应的$f'(x)$的近似值。这里$f(x) = \arctan x$, $x = \sqrt{2}$. 精确结果应当是$f'(x) = 1/(1 + x^2)$在$x = \sqrt{2}$点的值$1/3$
  - 运行Mathematica程序“数值微分_Taylor展开.nb”，通过改变其中的m以得到不同的效果

In[1]:=
```
f[x_] := ArcTan[x]; m = 8; s = N[Sqrt[2], m]; h = 1; M = 26; F1 = N[f[s], m];
For[k = 0, k <= M, k++, F2 = N[f[s + h], m]; d = N[F2 - F1, m]; r = N[d / h, m];
 Print[k, "\t", h, "\t", F2, "\t", F1, "\t", d, "\t", r]; h = N[h / 2, m]]
```

```
0    1    1.1780972    0.95531662    0.2227806    0.2227806
1    0.50000000    1.0893836    0.95531662    0.1340670    0.2681340
2    0.25000000    1.0297268    0.95531662    0.0744102    0.2976406
3    0.12500000    0.99464439    0.95531662    0.0393278    0.314622
4    0.062500000    0.97555095    0.95531662    0.0202343    0.323749
5    0.031250000    0.96558170    0.95531662    0.0102651    0.328483
6    0.015625000    0.96048682    0.95531662    0.0051702    0.33089
7    0.0078125000    0.95791122    0.95531662    0.0025946    0.33211
8    0.0039062500    0.95661631    0.95531662    0.0012997    0.33272
9    0.0019531250    0.95596706    0.95531662    0.0006504    0.33303
10    0.00097656250    0.95564199    0.95531662    0.0003254    0.3332
11    0.00048828125    0.95547934    0.95531662    0.0001627    0.3333
12    0.00024414063    0.95539799    0.95531662    0.0000814    0.3333
13    0.00012207031    0.95535731    0.95531662    0.0000407    0.333
14    0.000061035156    0.95533696    0.95531662    0.0000203    0.333
15    0.000030517578    0.95532679    0.95531662    0.0000102    0.333
16    0.000015258789    0.95532170    0.95531662    5.1×10^-6    0.33
17    7.6293945×10^-6    0.95531916    0.95531662    2.5×10^-6    0.33
18    3.8146973×10^-6    0.95531789    0.95531662    1.3×10^-6    0.33
19    1.9073486×10^-6    0.95531725    0.95531662    6.×10^-7    0.3
20    9.5367432×10^-7    0.95531694    0.95531662    3.×10^-7    0.3
21    4.7683716×10^-7    0.95531678    0.95531662    2.×10^-7    0.3
22    2.3841858×10^-7    0.95531670    0.95531662    0.×10^-8    0.3
23    1.1920929×10^-7    0.95531666    0.95531662    0.×10^-8    0.×10^-1
24    5.9604645×10^-8    0.95531664    0.95531662    0.×10^-8    0.×10^-1
25    2.9802322×10^-8    0.95531663    0.95531662    0.×10^-8    0.×10^-1
26    1.4901161×10^-8    0.95531662    0.95531662    0.×10^-8    0.
```

精度的减法相消

- 从实验中可以看到，当h趋向于零时，d的有效数字逐渐减少，直到最后$d = 0, r = 0,$ 因此并没有使得数值导数的精度越来越高
- 当运算中字长为8位，那么在$k = 11, 12$时得到最佳的结果0.33330000. 此时$d = f(x + h) - f(x)$有四位有效数字，随着k的增加，d中有效数字的个数在减少，而$r = d/h$的有效数字个数不会比d的更多。因此当h很小时，舍入误差使得当h趋向于零时不能得到高的精度
- 当然，如果计算过程中字长变大，那么得到的精度就会更高

高阶公式

- 无论如何，前面给出公式的误差估计只是h的一次方，我们可以通过对Taylor展开进行简单的处理，得到更高阶的数值导数计算公式
- 实际上，由于

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{3!}f'''(\xi_1)$$
$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{3!}f'''(\xi_2)$$

所以两式相减得到

$$f'(x) = \frac{1}{2h}\left[f(x+h) - f(x-h)\right] - \frac{h^2}{12}\left[f'''(\xi_1) + f'''(\xi_2)\right]$$

高阶公式的误差估计

- 只要f'''存在，前面公式的误差项是可用的
- 进一步，根据导数的介值定理[1]，存在一点$\xi \in (x-h, x+h)$，$f'''(\xi) = (f'''(\xi_1) + f'''(\xi_2))/2$。因此前页的数值微分公式可以重写为

$$f'(x) = \frac{1}{2h}\big[f(x+h) - f(x-h)\big] - \frac{h^2}{6}f'''(\xi)$$

[1]若f是$[a,b]$上的实值可微函数，再设$f'(a) < \lambda < f'(b)$，那么存在一点$x \in (a,b)$，$f'(x) = \lambda$.

例

设$f(x) = \arctan x$, $x = \sqrt{2}$, 用高阶公式重新计算$f'(x)$的值。正确值为$1/3$

- 通过数值实验，此时在$k = 8, 9, 10$时得到最高的五位精度
- 减法相消现象仍然存在，即当$k > 10$后，舍入误差使得精确位数在减少

In[2]:=
```
f[x_] := ArcTan[x]; m = 8; s = N[Sqrt[2], m]; h = 1; M = 26;
For[k = 0, k <= M, k++, F2 = N[f[s + h], m]; F1 = N[f[s - h], m]; d = N[F2 - F1, m];
 r = N[d / 2 / h, m]; Print[k, "\t", h, "\t", F2, "\t", F1, "\t", d, "\t", r]; h = N[h / 2, m]]
```

0	1	1.1780972	0.3926991	0.7853982	0.3926991
1	0.50000000	1.0893836	0.7406126	0.3487710	0.3487710
2	0.25000000	1.0297268	0.86112983	0.1685969	0.3371939
3	0.12500000	0.99464439	0.91106988	0.0835745	0.3342980
4	0.062500000	0.97555095	0.93385414	0.0416968	0.333574
5	0.031250000	0.96558170	0.94474460	0.0208371	0.333394
6	0.015625000	0.96048682	0.95006969	0.0104171	0.333348
7	0.0078125000	0.95791122	0.95270283	0.0052084	0.33334
8	0.0039062500	0.95661631	0.95401213	0.0026042	0.33333
9	0.0019531250	0.95596706	0.95466498	0.0013021	0.33333
10	0.00097656250	0.95564199	0.95499095	0.0006510	0.33333
11	0.00048828125	0.95547934	0.95515382	0.0003255	0.3333
12	0.00024414063	0.95539799	0.95523523	0.0001628	0.3333
13	0.00012207031	0.95535731	0.95527593	0.0000814	0.3333
14	0.000061035156	0.95533696	0.95529627	0.0000407	0.333
15	0.000030517578	0.95532679	0.95530645	0.0000203	0.333
16	0.000015258789	0.95532170	0.95531153	0.0000102	0.333
17	7.6293945×10^{-6}	0.95531916	0.95531407	5.1×10^{-6}	0.33
18	3.8146973×10^{-6}	0.95531789	0.95531535	2.5×10^{-6}	0.33
19	1.9073486×10^{-6}	0.95531725	0.95531598	1.3×10^{-6}	0.33
20	9.5367432×10^{-7}	0.95531694	0.95531630	$6.\times10^{-7}$	0.3
21	4.7683716×10^{-7}	0.95531678	0.95531646	$3.\times10^{-7}$	0.3
22	2.3841858×10^{-7}	0.95531670	0.95531654	$2.\times10^{-7}$	0.3
23	1.1920929×10^{-7}	0.95531666	0.95531658	$0.\times10^{-8}$	0.3
24	5.9604645×10^{-8}	0.95531664	0.95531660	$0.\times10^{-8}$	$0.\times10^{-1}$
25	2.9802322×10^{-8}	0.95531663	0.95531661	$0.\times10^{-8}$	$0.\times10^{-1}$
26	1.4901161×10^{-8}	0.95531662	0.95531661	$0.\times10^{-8}$	$0.\times10^{-1}$

差商

- 向前差商

$$f'(x_0) \approx \frac{f(x_0+h)-f(x_0)}{h}, \quad R(x) = -\frac{h}{2}f''(\xi) = O(h)$$

- 向后差商

$$f'(x_0) \approx \frac{f(x_0)-f(x_0-h)}{h}, \quad R(x) = \frac{h}{2}f''(\xi) = O(h)$$

- 中心差商

$$f'(x_0) \approx \frac{f(x_0+h)-f(x_0-h)}{2h}, \quad R(x) = -\frac{h^2}{6}f'''(\xi) = O(h^2)$$

数据误差

- 数值微分对于函数值的计算误差或测量误差非常敏感，因为此时函数值的误差被乘以$1/(2h)$
- 因此当计算由有误差数据确定的导数时，需要进行数据光滑化处理（去噪）
  - 即尽可能多得用当前点周围点信息把数据噪音去掉，因此在实际应用中，为了计算数据导数，通常是根据当前点以及周围点进行数据拟合，然后根据拟合出来的表达式计算当前点的导数
- 数值积分公式对数据误差不是很敏感

设定最佳步长（事后估计法）

- 设$D(h), D(h/2)$分别为步长为$h, h/2$的差商公式
- $$f'(x) - D(h) = O(h), \quad f'(x) - D(h/2) = O(h/2)$$
$$\Rightarrow \frac{f'(x) - D(h)}{f'(x) - D(h/2)} = \frac{O(h)}{O(h/2)} \approx 2$$
$$\Rightarrow f'(x) - D(h) = 2f'(x) - 2D(h/2)$$
$$\Rightarrow f'(x) - D(h/2) = D(h/2) - D(h)$$
- 当
$$|D(h) - D(h/2)| < \varepsilon$$
时的步长$h/2$就是合适的步长

高阶导数公式

- 根据高阶的Taylor展开，我们可以得到高阶导数的计算公式。
- 如

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f^{(4)}(\xi_1)$$

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f^{(4)}(\xi_2)$$

两式相加，得到

$$f''(x) = \frac{1}{h^2}\left[f(x+h) - 2f(x) + f(x-h)\right] - \frac{h^2}{12}f^{(4)}(\xi)$$

其中$\xi \in (x-h, x+h)$

- 这个公式常用于二阶微分方程的数值求解中

通过多项式插值计算数值导数

- 假设函数f在$x_0, x_1, \ldots, x_n$上的值已知，那么f在这些结点上存在唯一的插值多项式，从而有

$$f(x) = \sum_{k=0}^{n} f(x_i)\ell_i(x) + \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) w(x)$$

其中$\ell_i(x)$为Lagrange插值基函数，$w(x) = \prod_{i=0}^{n}(x - x_i)$

- 对其求导：

$$f'(x) = \sum_{i=0}^{n} f(x_i)\ell_i'(x) + \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) w'(x)$$
$$+ \frac{1}{(n+1)!} w(x) \frac{d}{dx} f^{(n+1)}(\xi_x)$$

- 如果我们是在结点处计算数值导数，即不妨设$x = x_\alpha$，由于$w(x_\alpha) = 0$，则结果得到简化：

$$f'(x_\alpha) = \sum_{i=0}^{n} f(x_i)\ell_i'(x_\alpha) + \frac{1}{(n+1)!} f^{(n+1)}(\xi_{x_\alpha}) w'(x_\alpha)$$

- 而

$$w'(x) = \sum_{i=0}^{n} \prod_{\substack{j=0 \\ j\neq i}}^{n} (x - x_j) \Longrightarrow w'(x_\alpha) = \prod_{\substack{j=0 \\ j\neq \alpha}}^{n} (x_\alpha - x_j)$$

所以带有误差项的数值微分公式为

$$f'(x_\alpha) = \sum_{i=0}^{n} f(x_i)\ell_i'(x_\alpha) + \frac{1}{(n+1)!} f^{(n+1)}(\xi_{x_\alpha}) \prod_{\substack{j=0 \\ j\neq \alpha}}^{n} (x_\alpha - x_j)$$

- 此公式特别适用于非等距结点

给出当$n=2, \alpha=1$时上述公式的显式表达

- 此时，三个Lagrange插值基函数为

$$\ell_0(x)=\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}$$
$$\ell_1(x)=\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}$$
$$\ell_2(x)=\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

- 它们的导数分别是

$$\ell_0'(x)=\frac{2x-x_1-x_2}{(x_0-x_1)(x_0-x_2)}$$
$$\ell_1'(x)=\frac{2x-x_0-x_2}{(x_1-x_0)(x_1-x_2)}$$
$$\ell_2'(x)=\frac{2x-x_0-x_1}{(x_2-x_0)(x_2-x_1)}$$

- 计算在$x = x_1$点的值，我们有

$$\ell_0'(x_1) = \frac{x_1 - x_2}{(x_0 - x_1)(x_0 - x_2)}$$
$$\ell_1'(x_1) = \frac{2x_1 - x_0 - x_2}{(x_1 - x_0)(x_1 - x_2)}$$
$$\ell_2'(x_1) = \frac{x_1 - x_0}{(x_2 - x_0)(x_2 - x_1)}$$

- 因而带有误差项的数值微分公式是

$$\begin{aligned} f'(x_1) =& f(x_0)\frac{x_1 - x_2}{(x_0 - x_1)(x_0 - x_2)} \\ &+ f(x_1)\frac{2x_1 - x_0 - x_2}{(x_1 - x_0)(x_1 - x_2)} \\ &+ f(x_2)\frac{x_1 - x_0}{(x_2 - x_0)(x_2 - x_1)} \\ &+ \frac{1}{6}f'''(\xi_{x_1})(x_1 - x_0)(x_1 - x_2) \end{aligned}$$

例：等距情形

- 在$n=2, \alpha=1$时，前面的数值微分公式简化为

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

这就是前面给出的高阶公式

Richardson外推

- Richardson外推(extrapolation)技术是通过巧妙应用Taylor级数改进数值微分公式的精度
- 函数$f(x)$的Taylor级数为

$$f(x+h)=\sum_{k=0}^{\infty}\frac{1}{k!}h^k f^{(k)}(x)$$

$$f(x-h)=\sum_{k=0}^{\infty}\frac{1}{k!}(-1)^k h^k f^{(k)}(x)$$

两式相减，消去了所有的偶次项：

$$f(x+h)-f(x-h)=2hf'(x)+\frac{2}{3!}h^3 f'''(x)+\frac{2}{5!}h^5 f^{(5)}(x)+\cdots$$

- 重新整理得：

$$L = \varphi(h) + a_2 h^2 + a_4 h^4 + a_6 h^6 + \cdots$$

其中$L = f'(x)$, $\varphi(h) = \frac{1}{2h}[f(x+h) - f(x-h)]$,
$a_k = -\frac{2}{(k-1)!} f^{(k-1)}(x)$

- 在上述公式中，我们需要$h > 0$. 在$h = 0$时我们得不到任何信息。对每个$h > 0$, $a_2 h^2 + a_4 h^4 + \cdots$给出了误差
- 实际上，通过取不同的h，我们可以进一步消去误差项中的低次项。如根据在h和$h/2$的表达式，

$$\begin{aligned} L &= \varphi(h) + a_2 h^2 + a_4 h^4 + a_6 h^6 + \cdots \\ 4L &= 4\varphi(h/2) + a_2 h^2 + a_4 h^4/4 + a_6 h^6/16 + \cdots \\ \hline 3L &= 4\varphi(h/2) - \varphi(h) - 3a_4 h^4/4 - 15a_6 h^6/16 - \cdots \end{aligned}$$

得到

$$L = \frac{4}{3}\varphi(h/2) - \frac{1}{3}\varphi(h) - a_4 h^4/4 - 5a_6 h^6/16 - \cdots$$

- 采用公式

$$
\begin{aligned}
L &= \frac{4}{3}\varphi(h/2) - \frac{1}{3}\varphi(h) \\
&= \frac{2}{3h}[f(x+h/2) - f(x-h/2)] - \frac{1}{6h}[f(x+h) - f(x-h)]
\end{aligned}
$$

计算前面例子中的导数

- 当$k = 4, 5, 6, 7$时得到六位数的精度

```
In[1]:= f[x_] := ArcTan[x]; m = 8; s = N[Sqrt[2], m]; h = 1; M = 30; d = Table[0, {M}]; F1 = N[f[s], m];
        For[k = 0, k ≤ M, k++, d[[k]] = N[(f[s + h] - f[s - h]) / 2 / h, m]; h = N[h / 2, m]]; For[k = 1,
         k <= M, k++, r = N[d[[k]] + (d[[k]] - d[[k - 1]]) / 3, m]; Print[k, "\t", d[[k]], "\t", r];]
```

1	0.3487710	0.3341283
2	0.3371939	0.3333348
3	0.3342980	0.3333327
4	0.3335745	0.333333
5	0.333394	0.333333
6	0.333348	0.333333
7	0.333337	0.333333
8	0.33333	0.33333
9	0.33333	0.33333
10	0.33333	0.33333
11	0.33333	0.3333
12	0.3333	0.3333
13	0.3333	0.3333
14	0.3333	0.333
15	0.333	0.333
16	0.333	0.333
17	0.333	0.333
18	0.33	0.33
19	0.33	0.33
20	0.33	0.33
21	0.33	0.3
22	0.3	0.3
23	0.3	0.3
24	0.3	$0.\times 10^{-1}$
25	$0.\times 10^{-1}$	$0.\times 10^{-1}$
26	$0.\times 10^{-1}$	$0.\times 10^{-1}$

进一步外推

- 令

$$\psi(h) = \frac{4}{3}\varphi(h/2) - \frac{1}{3}\varphi(h)$$

那么可以应用$\psi(h)$在h和$h/2$的取值进一步消去低阶项

- 实际上，

$$L = \psi(h) + b_4 h^4 + b_6 h^6 + \cdots$$

$$\frac{16L = 16\psi(h/2) + b_4 h^4 + b_6 h^6/4 + \cdots}{15L = 16\psi(h/2) - \psi(h) - 3b_6 h^6/4 - \cdots}$$

- 从而令

$$\theta(h) = \frac{16}{15}\psi(h/2) - \frac{1}{15}\psi(h)$$

得到

$$L = \theta(h) + c_6 h^6 + c_8 h^8 + \cdots$$

- 类似地，

$$L = \frac{64}{63}\theta(h/2) - \frac{1}{63}\theta(h) - \frac{3}{252}c_8 h^8 - \cdots$$

Richardson外推算法

- 上述过程可以执行任意多步，得到不断增加精度的公式
- 执行M步的Richardson外推算法为
  1. 选取h的一个初值，如$h=1$，并且计算$M+1$个数

$$D(n,0)=\varphi(h/2^n),\quad n=0,1,\ldots,M$$

  2. 执行下列公式的计算

$$D(n,k)=\frac{4^k}{4^k-1}D(n,k-1)-\frac{1}{4^k-1}D(n-1,k-1)$$

  这里$k=1,2,\ldots,M$, $n=k,k+1,\ldots,M$
  3. $D(M,M)$就是所需要的结果
- 在上述算法中，$D(0,0)=\varphi(h)$, $D(1,0)=\varphi(h/2)$, $D(1,1)=\psi(h)$

结果的精度

- 根据Richardson外推算法的计算过程,

$$
\begin{aligned}
D(n,0) &= L + \mathcal{O}(h^2) \\
D(n,1) &= L + \mathcal{O}(h^4) \\
D(n,2) &= L + \mathcal{O}(h^6) \\
D(n,3) &= L + \mathcal{O}(h^8)
\end{aligned}
$$

- 我们将证明

$$
D(n,k-1) = L + \mathcal{O}(h^{2k}), \qquad \text{当 } h \to 0
$$

Richardson外推定理

Theorem

在算法中定义的$D(n,k)$满足下列形式的等式

$$D(n,k-1)=L+\sum_{j=k}^{\infty}A_{j,k}(h/2^n)^{2j}$$

证明：当$k=1$时，由$D(n,0)$的定义以及

$$L=\varphi(h)+a_2h^2+a_4h^4+a_6h^6+\cdots$$

可知定理成立：

$$D(n,0)=\varphi(h/2^n)=L-\sum_{j=1}^{\infty}a_{2j}(h/2^n)^{2j}$$

因此可设$A_{j,1}=-a_{2j}$

现在对k进行归纳证明。假设$k-1$时定理成立，那么根据算法中$D(n,k)$的定义以及归纳假设，

$$\begin{aligned} D(n,k) &= \frac{4^k}{4^k-1}\left[L+\sum_{j=k}^{\infty}A_{j,k}\left(\frac{h}{2^n}\right)^{2j}\right] \\ &\quad -\frac{1}{4^k-1}\left[L+\sum_{j=k}^{\infty}A_{j,k}\left(\frac{h}{2^{n-1}}\right)^{2j}\right] \\ &= L+\sum_{j=k}^{\infty}A_{j,k}\frac{4^k-4^j}{4^k-1}\left(\frac{h}{2^n}\right)^{2j} \end{aligned}$$

从而我们可以定义

$$A_{j,k+1}=A_{j,k}\frac{4^k-4^j}{4^k-1}$$

显然$A_{k,k+1}=0$, 定理所需要的形式成立。 □

三角阵列

- 在算法中的$D(n,k)$形成如下的三角阵列

$$
\begin{array}{ccccccc}
D(0,0) & & & & & & \\
D(1,0) & D(1,1) & & & & & \\
D(2,0) & D(2,1) & D(2,2) & & & & \\
\vdots & \vdots & \vdots & \ddots & & & \\
D(M,0) & D(M,1) & D(M,2) & \cdots & D(M,M) &
\end{array}
$$

H.W.

编程实现用Richardson外推计算$f'(x)$的值, $h = 1$。函数$f(x)$分别取

- $\ln x$, $x = 3$, $M = 3$
- $\tan x$, $x = \sin^{-1}(0.8)$, $M = 4$
- $\sin(x^2 + \frac{1}{3}x)$, $x = 0$, $M = 5$.

输出相应的三角阵列

$$
\begin{array}{llll}
D(0,0) & & & \\
D(1,0) & D(1,1) & & \\
D(2,0) & D(2,1) & D(2,2) & \\
\vdots & \vdots & \vdots & \ddots \\
D(M,0) & D(M,1) & D(M,2) & \cdots \quad D(M,M)
\end{array}
$$

《数值分析》之

数值积分

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

数值积分问题

- 数值积分是应用函数在给定区间上的定义，计算出在该区间上的积分（近似）值
- 为什么需要数值积分?
  - 从数学上说，有些初等函数的原函数不是初等函数
  - 有些时候我们只知道函数在给定区间上有限个点处的函数值
  - 虽然有些函数的原函数可以得到，但是求值过于复杂
- 数值积分的策略是用另一个函数替代原来的被积函数，而前者的积分是很容易计算的
  - 多项式（来自于多项式插值或逼近）以及样条函数是常用的选择

通过多项式插值计算数值积分

- 目标是计算积分

$$\int_a^b f(x)dx$$

- 选取$[a,b]$中的结点$x_0, x_1, \ldots, x_n$, 应用Lagrange插值过程：
  1. Lagrange插值基函数为

$$\ell_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j}, \qquad i = 0, 1, \ldots, n$$

  2. 在结点上插值f的次数最多是n的多项式为

$$p(x) = \sum_{i=0}^{n} f(x_i)\ell_i(x)$$

积分公式

- 近似积分：

$$\int_a^b f(x)dx \approx \int_a^b p(x)dx = \sum_{i=0}^{n} f(x_i) \int_a^b \ell_i(x)dx$$

- 从而得到一个适用于所有f的公式：

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i)$$

其中

$$A_i = \int_a^b \ell_i(x)dx$$

该公式对所有的n阶多项式是精确成立的。

- 如果结点是等距的，那么上述公式称为Newton-Cotes公式

梯形法则

- 当$n=1$, $x_0=a$, $x_1=b$时，对应的公式称为梯形法则：
  - 此时
$$\ell_0(x)=\frac{b-x}{b-a}, \qquad \ell_1(x)=\frac{x-a}{b-a}$$
  - 从而
$$A_0=A_1=\frac{b-a}{2}$$
  - 相应的求积分式为
$$\int_a^b f(x)dx \approx \frac{b-a}{2}\left[f(a)+f(b)\right]$$
- 该公式对所有的线性多项式函数精确成立

梯形法则的误差

- 对一般函数，梯形法则的误差为$-\frac{1}{12}(b-a)^3 f''(\xi), \xi \in (a,b)$

1. 多项式插值中的误差

$$E(x) = f(x) - p(x) = f''(\xi_x)(x-a)(x-b)/2$$

2. 对其进行积分，并应用积分中值定理[1]

$$\begin{aligned}\int_a^b E(x)dx &= -\frac{1}{2}\int_a^b f''(\xi_x)(x-a)(b-x)dx \\ &= -\frac{1}{2}f''(\xi)\int_a^b (x-a)(b-x)dx = -\frac{1}{12}(b-a)^3 f''(\xi)\end{aligned}$$

[1]令u, v为区间$[a,b]$上的连续函数，$v \geqslant 0$. 那么存在$\xi \in (a,b)$使得

$$\int_a^b u(x)v(x)dx = u(\xi)\int_a^b v(x)dx$$

Runge现象

$$\int_{-5}^{5} \frac{1}{1+x^2}\,dx$$

n	I_n
1	0.38462
2	6.79487
3	2.08145
4	2.37401
5	2.30769
6	3.87045
7	2.89899
8	1.50049
9	2.39862
10	4.67330
11	3.24477
12	−0.31294
13	1.91980
14	7.89954
15	4.15556

复化梯形法则

- 把单个区间上的积分公式应用在区间被划分以后的每个子区间上，便得到复化(composite, 复合）法则或公式
- 对于梯形法则，把区间$[a, b]$划分为

$$a = x_0 < x_1 < \cdots < x_n = b$$

在每个子区间上应用梯形法则，得到复化梯形法则：

$$\begin{aligned}\int_a^b f(x)dx &= \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)dx \\ &\approx \frac{1}{2}\sum_{i=1}^{n}(x_i - x_{i-1})\left[f(x_{i-1}) + f(x_i)\right]\end{aligned}$$

- 复化梯形法则相当于用分片线性多项式(即连接结点处函数值的折线)替换被积函数f得到的数值积分公式
- 对于等距节点$x_i = a + ih$, $h = (b-a)/n$, 复化梯形法则具有形式

$$\int_a^b f(x)dx \approx \frac{h}{2}\left[f(a) + 2\sum_{i=1}^{n-1} f(a+ih) + f(b)\right]$$

复化梯形法则的误差

- 复化梯形法则的误差为

$$-\frac{1}{12}(b-a)h^2 f''(\xi)$$

其中$\xi \in (a,b)$。证明中用到了下述事实：

1. 在(a,b)中存在一点ξ使得

$$f''(\xi) = \frac{1}{n}\sum_{i=1}^{n} f''(\xi_i)$$

这里$\xi_i \in (x_{i-1}, x_i)$

2. $n = (b-a)/h$

待定系数法

- 从上节可知，公式

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i)$$

对于所有次数不超过n的多项式精确成立，其中A_i是第i个Lagrange插值基函数在区间$[a,b]$上的积分

- 反过来，如果我们知道上述公式对所有次数不超过n的多项式精确成立，那么是否必定有下式成立呢?

$$A_i = \int_a^b \ell_i(x)dx$$

- 答案是肯定的，因为上述公式对于任何ℓ_i精确成立，从而

$$\int_a^b \ell_j(x)dx = \sum_{i=0}^{n} A_i \ell_j(x_i) = A_j$$

新的求解方法

- 因此当加上条件"对所有次数不超过n的多项式精确成立"后，上述公式是唯一确定的，从而我们不必要通过计算Lagrange插值基函数的积分来确定系数，而是采用更有效直接的待定系数法
- 例如，重新推导前面的例题，即确定下式中的A_0, A_1和A_2:

$$\int_0^1 f(x)dx \approx A_0 f(0) + A_1 f(1/2) + A_2 f(1)$$

  - 由于公式对所有次数不超过2的多项式精确成立，把$f(x) = 1, x, x^2$作为试用函数，得到

$$1 = \int_0^1 dx = A_0 + A_1 + A_2$$

$$\frac{1}{2} = \int_0^1 xdx = \frac{1}{2}A_1 + A_2$$

$$\frac{1}{3} = \int_0^1 x^2 dx = \frac{1}{4}A_1 + A_2$$

- 从而得到联立方程组的解为

$$A_0 = \frac{1}{6}, \quad A_1 = \frac{2}{3}, \quad A_2 = \frac{1}{6}$$

- 由于公式是线性的，因此对所有次数不超过2的多项式都精确成立
- 值得指出的是，由于积分节点是对称分布的，而且由于所有系数的和为1，因此这里只要算出x在$[0,1]$上的积分值，就可以算出所有的系数

Simpson法则

- 如果把例题中的$[0,1]$区间换为一般的$[a,b]$, 得到著名的Simpson法则

$$\int_a^b f(x)dx \approx \frac{b-a}{6}\left[f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right]$$

- 从公式的推导过程可知，它对所有次数不超过2的多项式精确成立。而出人意料的是它对所有次数不超过3的多项式也精确成立，原因在于：

$$\frac{1}{4}(b^4-a^4)=\int_a^b x^3dx=\frac{b-a}{6}\left[a^3+4\left(\frac{a+b}{2}\right)^3+b^3\right]$$

Simpson法则的误差

- 误差项的准确表示为

$$-\frac{1}{90}\left(\frac{b-a}{2}\right)^5 f^{(4)}(\xi)$$

其中$\xi \in (a,b)$.

- 下面应用Taylor展开，可以证明误差是$\mathcal{O}(h^5)$, $h=(b-a)/2$
  1. 重写积分公式如下：

$$\int_a^{a+2h} f(x)dx \approx \frac{h}{3}[f(a)+4f(a+h)+f(a+2h)]$$

  2. 右端项的Taylor展开结果为

$$2hf(a)+2h^2f'(a)+\frac{4}{3}h^3f''(a)+\frac{2}{3}h^4f'''(a)+\frac{100}{3\cdot 5!}h^5f^{(4)}+\cdots$$

3. 设

$$F(x) = \int_a^x f(t)dt$$

根据微积分基本定理，$F' = f$. 再根据Taylor展开，积分公式的左端为

$$\begin{aligned} F(a+2h) =& 2hf(a) + 2h^2 f'(a) + \frac{4}{3}h^3 f''(a) + \frac{2}{3}h^4 f'''(a) \\ &+ \frac{32}{5!}h^5 f^{(4)}(a) + \cdots \end{aligned}$$

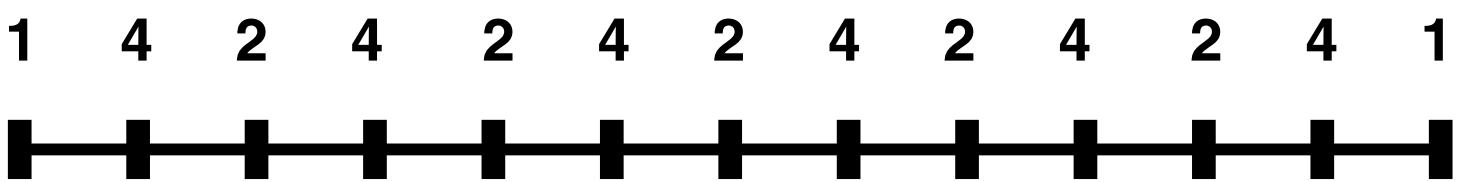
4. 把上述两个Taylor展开结合在一起，有

$$\int_a^{a+2h} f(x)dx = \frac{h}{3}[f(a)+4f(a+h)+f(a+2h)] - \frac{1}{90}h^5 f^{(4)}(a) - \cdots$$

复化Simpson法则

- 前述的Simpson法则可以应用到有偶数个子区间的复化情形
- 设n为偶数，$x_i = a + ih,\ i = 0, 1, \ldots, n,\ h = \frac{b-a}{n}$，则复化Simpson法则为

$$\begin{aligned}\int_a^b f(x)dx &= \sum_{i=1}^{n/2} \int_{x_{2i-2}}^{x_{2i}} f(x)dx \\ &= \frac{h}{3} \sum_{i=1}^{n/2} [f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})] \\ &= \frac{h}{3} \left[f(x_0) + 2 \sum_{i=2}^{n/2} f(x_{2i-2}) + 4 \sum_{i=1}^{n/2} f(x_{2i-1}) + f(x_n) \right]\end{aligned}$$



- 误差为$-\frac{1}{180}(b-a)h^4 f^{(4)}(\xi),\ \xi \in (a, b)$

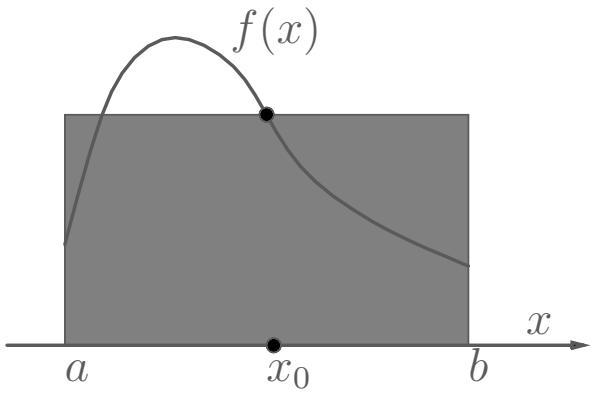
计算$\pi = \int_0^1 \frac{1}{1+x^2} dx$.

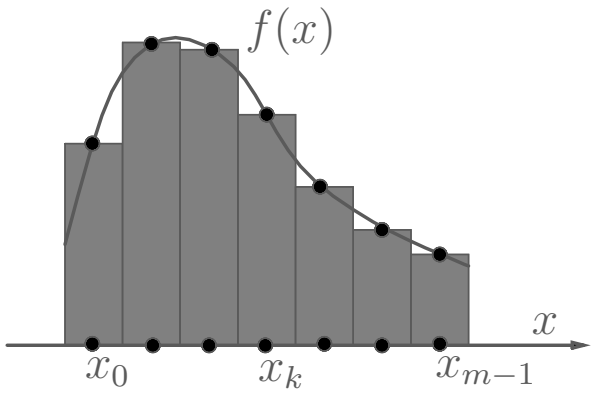
解:

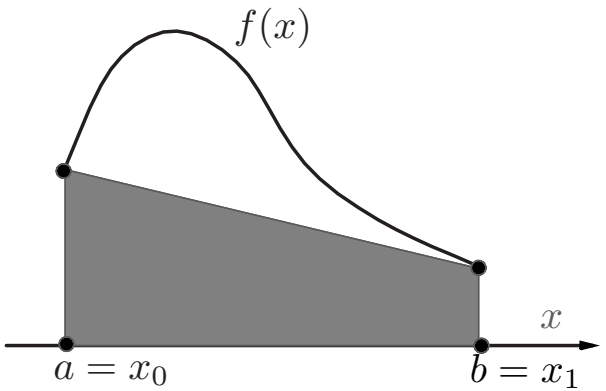
$$T_8 = \frac{1}{16}\left[f(0) + 2\sum_{k=1}^{7} f(x_k) + f(1)\right]$$
$$= 3.138988494$$
$$S_4 = \frac{1}{24}\left[f(0) + 4\sum_{odd} f(x_k) + 2\sum_{even} f(x_k) + f(1)\right]$$
$$= 3.141592502$$

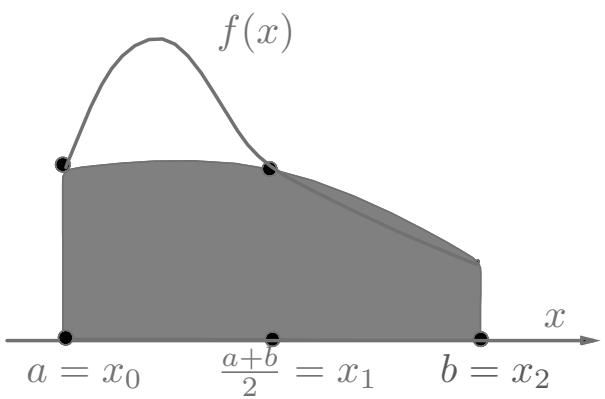
其中$x_k = k/8$.

几个常用积分公式









一般积分公式

- 可以把前面的数值积分公式一般化为离散点上的函数值的线性组合

$$I(f) \equiv \int_a^b f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i) \equiv I_n(f)$$

要求公式对于所有次数不超过n的多项式精确成立。

- 数值积分有k阶代数精度是指：

$$I_n(x^i) = I(x^i), \quad i = 0, \cdots, k, \quad I_n(x^{k+1}) \neq I(x^{k+1}).$$

对任意次数不高于k次的多项式f(x)，数值积分没有误差

确定下面公式中的系数，使得当f为次数不超过3的多项式时精确成立：

$$\int_{-\pi}^{\pi} f(x)\cos x\,dx \approx A_0 f\left(-\frac{3}{4}\pi\right)+A_1 f\left(-\frac{1}{4}\pi\right)+A_2 f\left(\frac{1}{4}\pi\right)+A_3 f\left(\frac{3}{4}\pi\right)$$

- 由于公式要对所有次数不超过3的多项式精确成立，因此把$f(x)=x^i,\ i=0,1,2,3$代入得到四个线性方程。而根据对称性，应有$A_0=A_3,\ A_1=A_2$，所以只需要应用两个条件如下：

$$0=\int_{-\pi}^{\pi}\cos x\,dx = 2A_0+2A_1$$

$$-4\pi=\int_{-\pi}^{\pi}x^2\cos x\,dx = 2A_0(3\pi/4)^2+2A_1(\pi/4)^2$$

得到解为$A_1=A_2=-A_0=-A_3=4/\pi$

区间变换

- 经过变量的线性变换，我们可以从某一个区间上的数值积分公式导出其它区间上的数值积分公式，而且两个公式同时对次数不超过同样m的多项式精确成立
- 假设有一个数值积分公式

$$\int_c^d f(t)dt \approx \sum_{i=0}^{n} A_i f(t_i)$$

已知，它对所有次数不超过m的多项式精确成立

- 为了导出在区间$[a, b]$上的公式，定义t 的线性函数

$$\lambda(t) = \frac{b-a}{d-c}t + \frac{ad-bc}{b-c}$$

满足$\lambda(c) = a$, $\lambda(d) = b$, 而且其间线性变化

- 对积分

$$\int_a^b f(x)dx$$

作变量代换$x = \lambda(t)$, $dx = \frac{b-a}{d-c}dt$,

$$\int_a^b f(x)dx = \frac{b-a}{d-c}\int_c^d f(\lambda(t))dt \approx \frac{b-a}{d-c}\sum_{i=0}^{n} A_i f(\lambda(t_i))$$

- 从而得到$[a, b]$区间上的数值积分公式

$$\int_a^b f(x)dx \approx \frac{b-a}{d-c}\sum_{i=0}^{n} A_i f\left(\frac{b-a}{d-c}t_i + \frac{ad-bc}{d-c}\right)$$

由于当f为多项式时，$f(t)$与$f(\lambda(t))$具有相同的次数，因此对于次数不超过m的多项式，新公式也精确成立

由于前面的数值积分公式本质上是来自于多项式插值，因此基于多项式插值的误差公式我们也可以给出数值积分的误差估计

- 如果p是在点$x_0, x_1, \ldots, x_n$上插值$f \in C^{(n+1)}[a, b]$的次数不超过n的多项式，那么

$$f(x) - p(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i)$$

- 从而我们有

$$\int_a^b f(x) - \sum_{i=0}^{n} A_i f(x_i) = \frac{1}{(n+1)!} \int_a^b f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i) dx$$

- 如果在$[a,b]$上$|f^{(n+1)}(x)| \leqslant M$, 则有

$$\left| \int_a^b f(x) - \sum_{i=0}^{n} A_i f(x_i) \right| \leqslant \frac{M}{(n+1)!} \int_a^b \prod_{i=0}^{n} |x - x_i| dx$$

- 类似于多项式插值理论中（第一类）Tchebyshev多项式的零点使得

$$\max_{x \in [a,b]} \prod_{i=0}^{n} |x - x_i|$$

达到最小，哪么如何选取x_i使得

$$\int_a^b \prod_{i=0}^{n} |x - x_i| dx$$

最小?

第二类Tchebyshev多项式

- 函数

$$U_{n+1}(x) = \frac{\sin((n+2)\theta)}{\sin\theta}, \quad x = \cos\theta$$

  称为第二类Tchebyshev多项式

- 递推表示为

$$U_0(x) = 1, \quad U_1(x) = 2x$$
$$U_{n+1}(x) = 2xU_n - U_{n-1}, \quad n \geqslant 1$$

- 它满足如下正交关系：

$$\int_{-1}^{1} U_n(x)U_m(x)\sqrt{1-x^2}dx = \delta_{mn}\frac{\pi}{2}$$

- 与第一类Tchebyshev多项式的关系为 $T_n'(x) = nU_{n-1}(x)$

- $U_{n+1}(x)$是次数为$n+1$的多项式，首项系数为2^{n+1}
- 它的零点都在$[-1,1]$内，分别是

$$x_i = \cos\frac{(i+1)\pi}{n+2}, i = 0, 1, \ldots, n$$

- 从而有

$$\frac{U_{n+1}(x)}{2^{n+1}} = (x - x_0)(x - x_1)\cdots(x - x_n)$$

$$\int_{-1}^{1} \left|(x-x_0)(x-x_1)\cdots(x-x_n)\right| dx = \frac{1}{2^n}$$

理由在于：

$$\begin{aligned}
\int_{-1}^{1} |U_{n+1}(x)| dx &= \int_0^{\pi} |\sin(n+2)\theta| d\theta \\
&= \sum_{i=0}^{n+1} \int_{\pi i/(n+2)}^{\pi(i+1)/(n+2)} (-1)^i \sin(n+2)\theta d\theta \\
&= \sum_{i=0}^{n+1} (-1)^{i+1} \left(\frac{\cos(n+2)\theta}{n+2} \Big|_{\pi i/(n+2)}^{\pi(i+1)/(n+2)} \right) \\
&= 2
\end{aligned}$$

极值性质定理

Theorem

在所有的n次首一多项式p中，使得

$$\int_{-1}^{1} |p(x)|dx$$

最小的多项式是$2^{-n}U_n(x)$

证明：首先证明下述正交关系：

$$I = \int_{-1}^{1} U_m(x)\,\mathrm{sign}[U_n(x)]dx = 0, \qquad 0 \leqslant m < n$$

实际上，在积分I中进行变量代换$\cos\theta = x$得到

$$
\begin{aligned}
I &= \int_0^{\pi} \sin(m+1)\theta \,\mathrm{sign}\left[\frac{\sin(n+1)\theta}{\sin\theta}\right] d\theta \\
&= \sum_{k=0}^{n}(-1)^k \int_{k\varphi}^{(k+1)\varphi} \sin(m+1)\theta d\theta \qquad \varphi = \frac{\pi}{n+1} \\
&= \frac{1}{m+1}\sum_{k=0}^{n}(-1)^{k+1}[\cos(m+1)(k+1)\varphi - \cos(m+1)k\varphi]
\end{aligned}
$$

为了证明最后的求和项等于零，需要借助于复数的Euler公式，转化为幂级数求和。[2]

[2]实际上，在一些数学手册中可以很容易找到上述求和表达式，从而得证所需要的结论。如在“Table of Integrals, Series, and Products, 6th Ed.”中的1.343给出了如下公式

$$
\sum_{k=1}^{n}(-1)^k \cos kx = -\frac{1}{2} + \frac{(-1)^n \cos\left(\frac{2n+1}{2}x\right)}{2\cos x/2}
$$

令$\alpha = (m+1)\varphi + \pi$, 则有

$$
\begin{aligned}
(m+1)I &= \sum_{k=0}^{n}[\cos(k+1)\alpha + \cos k\alpha] \\
&= \mathrm{Re}\left\{\sum_{k=0}^{n}[e^{i\alpha(k+1)} + e^{i\alpha k}]\right\} \\
&= \mathrm{Re}\frac{e^{i\alpha(n+2)} - e^{i\alpha} + e^{i\alpha(n+1)} - 1}{e^{i\alpha} - 1} \\
&= \frac{\mathrm{Re}[(e^{-i\alpha} - 1)(e^{i\alpha(n+2)} - e^{i\alpha} + e^{i\alpha(n+1)} - 1)]}{|e^{i\alpha} - 1|^2}
\end{aligned}
$$

最后一项的分母为实数，分子为(经过简单的计算后)

$$
\mathrm{Re}(e^{i\alpha n} - e^{i\alpha(n+2)} + e^{i\alpha} - e^{-i\alpha}) = \cos n\alpha - \cos(n+2)\alpha = 0
$$

下面完成定理的证明。令p为任意的首一n次多项式，则p有表示

$$p = 2^{-n}U_n + a_{n-1}U_{n-1} + \cdots + a_0 U_0$$

从而根据前面的正交关系，

$$\begin{aligned}\int_{-1}^{1} |p|dx &\geqslant \int_{-1}^{1} p \operatorname{sign} U_n dx = \frac{1}{2^n}\int_{-1}^{1} U_n \operatorname{sign} U_n dx \\ &= 2^{-n}\int_{-1}^{1} |U_n|dx\end{aligned}$$

□

更紧致的误差估计

- 对于$[-1,1]$区间上的数值积分，如果数值积分的$n+1$个结点来自于U_{n+1}的零点，那么

$$\left| \int_{-1}^{1} f(x)dx - \sum_{i=0}^{n} A_i f(x_i) \right| \leqslant \frac{M}{(n+1)!2^n}$$

- 如果积分区间为$[a,b]$，那么相应的积分结点可以是$[-1,1]$区间中的仿射变换，即为

$$x_i = \frac{a+b}{2} + \frac{b-a}{2} \cos \frac{(i+1)\pi}{n+2}, \quad i = 0,1,\ldots,n$$

上机作业

对函数

$$f(x) = \frac{1}{1+25x^2}, x \in [-1,1]$$

构造Lagrange插值多项式$p_L(x)$，插值节点取为:

$$1.\ x_i = 1 - \frac{2}{N}i,\ i = 0,1,\cdots,N$$

$$2.\ x_i = -\cos(\frac{i+1}{N+2}\pi),\ i = 0,1,\cdots,N$$

利用$\int_{-1}^{1} p_L(x)dx$计算积分$\int_{-1}^{1} f(x)dx$的近似值，并计算如下误差

$$|\int_{-1}^{1} p_L(x)dx - \int_{-1}^{1} f(x)dx|,$$

对$N = 5,10,15,20,25,30,35,40$比较以上两组节点的结果。

输出形式如下：

N	$\int_{-1}^{1} p_L(x)dx$	$\int_{-1}^{1} f(x)dx$	$\lvert \int_{-1}^{1} p_L(x)dx - \int_{-1}^{1} f(x)dx \rvert$
5			
10			
15			
20			
25			
30			
35			
40			

奇异积分的计算

- 主要考虑如下几类奇异积分
  1. 被积函数$f(x)$在某点具有有限跳跃，即存在$c \in [a, b]$，$f(c+) - f(c-)$为非零有限值
  2. 无界函数的积分
  3. 积分区间无界
- 上述函数的积分不能通过直接采用多项式插值逼近原函数，然后用多项式的积分代替被积函数的积分，因为插值误差的界是无限的

间断函数的积分

- 令c为间断点，那么

$$I(f) = \int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

  此时可以在$[a, c-]$和$[c+, b]$上应用前面的数值积分公式，以得到$I(f)$的近似值
- 当在积分区间中具有有限个此类间断点时，可以类似处理
- 当间断点不预先知道时，需要对函数的图形进行分析，或者采用自适应积分的方法以确定间断点的存在

无界函数的积分

- 假设 $\lim\limits_{x\to a+} f(x)=\infty$ (当f在$x\to b-$时为无穷可类似处理; 如果是在区间内点为无界，那么可以从该内点把区间分开，再分开处理)
- 假设

$$f(x)=\frac{\phi(x)}{(x-a)^{\mu}},\quad 0\leqslant\mu<1$$

  这里$\phi(x)$以M为界
- 那么

$$|I(f)|\leqslant M\lim_{t\to a+}\int_t^b\frac{1}{(x-a)^{\mu}}dx=M\frac{(b-a)^{1-\mu}}{1-\mu}$$

- 对于任意ε: $0 < \varepsilon < b - a$, 积分可写为$I(f) = I_1 + I_2$,

$$I_1 = \int_a^{a+\varepsilon} \frac{\phi(x)}{(x-a)^\mu} dx, \qquad I_2 = \int_{a+\varepsilon}^b \frac{\phi(x)}{(x-a)^\mu} dx$$

- I_2为正常的积分，按通常的数值积分公式计算
- 为了计算I_1, 把$\phi(x)$在$x = a$点进行Taylor展开：

$$\phi(x) = \Phi_p(x) + \frac{(x-a)^{p+1}}{(p+1)!}\phi^{(p+1)}(\xi_x), \quad p \geqslant 0$$

其中

$$\Phi_p(x) = \sum_{k=0}^{p} \frac{(x-a)^k}{k!}\phi^{(k)}(a)$$

- 从而有

$$I_1 = \varepsilon^{1-\mu} \sum_{k=0}^{p} \frac{\varepsilon^k \phi^{(k)}(a)}{k!(k+1-\mu)} + \frac{1}{(p+1)!} \int_a^{a+\varepsilon} (x-a)^{p+1-\mu} \phi^{(p+1)}(\xi_x) dx$$

- 因此当用第一项(有限和)代替I_1时，对应的误差E_1有如下估计

$$|E_1| \leqslant \frac{\varepsilon^{p+2-\mu}}{(p+1)!(p+2-\mu)} \max_{a \leqslant x \leqslant a+\varepsilon} |\phi^{(p+1)}(x)|$$

- 对于固定的p, 右端项为ε的增函数。而当$\varepsilon < 1$，并且$\phi^{(p+1)}(x)$ 随着p的增加，变化不是很快时，那么右端项随着p增加而减小
- 实际应用时，在确定计算I_1和I_2的数值公式参数时，需要保证两者的误差几乎相当

无界区间

- 考虑积分

$$I(f) = \int_a^{\infty} f(x)dx$$

- $I(f)$存在的一个充分条件是

$$\exists \rho > 0, \text{s.t.} \quad \lim_{x \to +\infty} x^{1+\rho} f(x) = 0$$

第一种方法

- 为了计算$I(f)$, 把它分为两部分：$I(f) = I_1 + I_2$, 其中

$$I_1 = \int_a^c f(x)dx, \qquad I_2 = \int_c^\infty f(x)dx$$

这里要选择恰当的c，使得I_2对整个积分的贡献可以被忽略。即如果希望数值计算出来的积分值与$I(f)$的误差不超过δ, 那么要选择c使得$|I_2| \leqslant \delta/2$

- 作业：在计算积分

$$\int_0^\infty \cos^2(x)e^{-x}dx$$

时为了使得误差不超过$\delta = 10^{-3}$, 那么c应取什么值？

第二种方法

- 把积分在$c > 0$点分开，此时不需要保证第二个积分足够小
- 为了计算I_2, 引入变量代换$x = 1/t$:

$$I_2 = \int_c^{\infty} f(x)dx = \int_0^{1/c} \frac{f(1/t)}{t^2} dt := \int_0^{1/c} g(t)dt$$

- 如果$g(t)$在区间$[0, 1/c]$内连续，那么可以采用通常的数值积分方法；否则可以采用无界函数的积分方法

- 还有一种方法，需要用到正交多项式，放在Gauss积分一节中讲述

上机作业

- 分别编写用复化Simpson积分公式和复化梯形积分公式计算积分的通用程序
- 用如上程序计算积分

$$I(f) = \int_0^4 sin(x)dx$$

取节点$x_i, i = 0, \cdots, N$, N为$2^k, k = 1, \cdots, 12$,并分析误差

- 利用公式计算算法的收敛阶。

$$Ord = \frac{\ln(Error_{old}/Error_{now})}{\ln(N_{now}/N_{old})}$$

输出形式如下：

N	复化Simpson error	order	复化梯形error	order
2		–		–
4				
8				
16				
32				
64				
128				
256				
512				
1024				
2048				
4096				

《数值分析》之

数值积分

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

重积分的计算

- 在微积分中，二重积分的计算是用化为累次积分的方法进行的。
- 计算二重数值积分也同样采用累次积分的计算过程。简化起见，我们仅讨论矩形区域上的二重积分。
- 对非矩形区域的积分，大多可以变化为矩形区域上的累次积分。

$$\int_a^b \int_c^d f(x,y)dydx$$

a,b,c,d 为常数，f 在D 上连续。将它变为化累次积分

$$\begin{aligned}
&\int_a^b \int_c^d f(x,y)dydx \\
=&\int_a^b \left(\int_c^d f(x,y)dy \right) dx \\
=&\int_c^d \left(\int_a^b f(x,y)dx \right) dy
\end{aligned}$$

二重积分的复化梯形公式

- 做等距节点，x轴，y轴分别有$h = \frac{b-a}{m}$, $k = \frac{d-c}{n}$
- 将x作为常数,先计算$\int_c^d f(x,y)dy$,有

$$\int_c^d f(x,y)dy \approx \frac{k}{2}\left(f(x,y_0) + 2\sum_{i=1}^{n-1} f(x,y_i) + f(x,y_n)\right)$$

二重积分的复化梯形公式

- 再将y作为常数，在x方向，计算上式的每一项的积分

$$\int_a^b f(x,y_0)dx \approx \frac{h}{2}\left(f(x_0,y_0)+2\sum_{j=1}^{m-1}f(x_j,y_0)+f(x_m,y_0)\right)$$

$$\int_a^b f(x,y_n)dx \approx \frac{h}{2}\left(f(x_0,y_n)+2\sum_{j=1}^{m-1}f(x_j,y_n)+f(x_m,y_n)\right)$$

$$\int_a^b \sum_{i=1}^{n-1} f(x,y_i) = \sum_{i=1}^{n-1}\int_a^b f(x,y_i)$$

$$\approx \sum_{i=1}^{n-1}\frac{h}{2}\left(f(x_0,y_i)+2\sum_{j=1}^{m-1}f(x_j,y_i)+f(x_m,y_i)\right)$$

二重积分的复化梯形公式

$$\int_a^b \int_c^d f(x,y)dydx \approx \frac{hk}{4}\left(f(x_0,y_0)+f(x_0,y_n)+f(x_m,y_0)+f(x_m,y_n) \right.$$
$$+2\sum_{j=1}^{m-1} f(x_j,y_0)+2\sum_{j=1}^{m-1} f(x_j,y_n)+2\sum_{i=1}^{n-1} f(x_0,y_i)+2\sum_{i=1}^{n-1} f(x_n,y_i)$$
$$\left. +4\sum_{i=1}^{n-1}\sum_{j=1}^{m-1} f(x_j,y_i) \right)$$

- 系数，在积分区域的四个角点为1/4，4个边界为1/2，内部节点为1
- 误差

$$-\frac{(b-a)(d-c)}{12}\left(h^2\frac{\partial^2}{\partial x^2}f(\xi,\eta)+k^2\frac{\partial^2}{\partial y^2}f(\bar{\xi},\bar{\eta}) \right)$$

二重积分的复化Simpson公式

$$\int_a^b \int_c^d f(x,y)dydx \approx hk \sum_{i=0}^{n} \sum_{j=0}^{m} \omega_{i,j} f(x_j, y_i)$$

- 系数

$$U = \{u_0, u_1, \cdots, u_m\} = \{\frac{1}{3}, \frac{4}{3}, \frac{2}{3}, \frac{4}{3}, \cdots, \frac{2}{3}, \frac{4}{3}, \frac{1}{3}\}$$

$$V = \{v_0, v_1, \cdots, v_n\} = \{\frac{1}{3}, \frac{4}{3}, \frac{2}{3}, \frac{4}{3}, \cdots, \frac{2}{3}, \frac{4}{3}, \frac{1}{3}\}$$

$$\omega_{i,j} = u_i v_j$$

- 误差

$$-\frac{(b-a)(d-c)}{180}\left(h^4 \frac{\partial^4}{\partial x^4} f(\xi,\eta) + k^4 \frac{\partial^4}{\partial y^4} f(\bar{\xi},\bar{\eta})\right)$$

《数值分析》之

数值积分

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

数值积分结点的选择

- 数值积分公式

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i)$$

在一旦结点取定后，只要加上条件“对次数不超过n的多项式精确成立"，那么组合系数就唯一确定

- 在Simpson法则中，虽然积分公式是利用“次数不超过2的多项式精确成立"确定的，但最后却对三次多项式也确精成立。因此自然的问题是：能否通过选择结点，使得积分公式更好?

数值积分结点的选择

1. 是否存在所有组合系数都相等的积分公式?

$$\int_a^b f(x)dx \approx c\sum_{i=0}^{n} f(x_i)$$

2. 通过对$n+1$个结点的选取，使得积分公式对尽可能高的多项式精确成立? 目标是对“次数不超过2n+1的多项式精确成立”

Tchebyshev积分公式

- 在数值积分公式中，如果所有的组合系数是相同的，那么可以大大减少计算的乘法次数。这称为Tchebyshev积分公式(Tchebyshev's quadrature formulas)
- 为了简单起见，通常的积分区间取为$[-1,1]$

$$\int_{-1}^{1} f(x)dx \approx \frac{2}{n+1}\sum_{i=0}^{n} f(x_i)$$

- 只有当$n = 0,1,2,3,4,5,6,8$时才存在这样的积分公式
- 当$n = 0,1,2,3,4$时，结点可以解析写出，其它情形只有数值解

解析结点

- $n = 0$, $x_0 = 0$
- $n = 1$, $x_0 = -\frac{\sqrt{3}}{3} \approx -0.57735$, $x_1 = \frac{\sqrt{3}}{3}$
- $n = 2$, $x_0 = -\frac{\sqrt{2}}{2} \approx -0.707107$, $x_1 = 0$, $x_2 = \frac{\sqrt{2}}{2}$
- $n = 3$, $x_0 = -\sqrt{\frac{\sqrt{5}+2}{3\sqrt{5}}} \approx -0.794654$,
  $x_1 = -\sqrt{\frac{\sqrt{5}-2}{3\sqrt{5}}} \approx -0.187592$, $x_2 = \sqrt{\frac{\sqrt{5}-2}{3\sqrt{5}}}$, $x_3 = \sqrt{\frac{\sqrt{5}+2}{3\sqrt{5}}}$
- $n = 4$, $x_0 = -\sqrt{\frac{5+\sqrt{11}}{12}} \approx -0.832497$,
  $x_1 = -\sqrt{\frac{5-\sqrt{11}}{12}} \approx -0.374541$, $x_2 = 0$, $x_3 = \sqrt{\frac{5-\sqrt{11}}{12}}$,
  $x_4 = \sqrt{\frac{5+\sqrt{11}}{12}}$

数值结点

- $n = 5,\ x_i = \pm 0.2666354015, \pm 0.4225186537, \pm 0.8662468181$
- $n = 6,$
  $x_i = 0, \pm 0, 3239118105, \pm 0.5296567752, \pm 0.8838617007$
- $n = 8,\ x_i = 0, \pm 0.167906, \pm 0.528762, \pm 0.601019, \pm 0.911589$

在两点数值积分公式中，如果积分点也作为未知量，则有4个未知量，可以列出4个方程：（以$f(x)$在$[-1,1]$为例）

$$2=\int_{-1}^{1}dx=A_0+A_1, \qquad 0=\int_{-1}^{1}xdx=A_0x_0+A_1x_1$$

$$\frac{2}{3}=\int_{-1}^{1}x^2dx=A_0x_0^2+A_1x_1^2, \qquad 0=\int_{-1}^{1}x^3dx=A_0x_0^3+A_1x_1^3$$

可解出

$$A_0=1,\ A_1=1,\ x_0=-\frac{\sqrt{3}}{3},\ x_1=\frac{\sqrt{3}}{3}$$

数值积分公式

$$\int_{-1}^{1}f(x)dx=f(-\frac{\sqrt{3}}{3})+f(\frac{\sqrt{3}}{3})$$

Theorem (Gauss积分定理)

设w是正的权函数，q是一个$n+1$次非零多项式，并且与n次非零多项式是关于w正交的，即对任意n次非零多项式p都有

$$\int_a^b q(x)p(x)w(x)dx = 0$$

若$x_0, x_1, \ldots, x_n$是q的零点，则下述积分公式对所有$2n+1$次非零多项式f精确成立：

$$\int_a^b f(x)w(x)dx \approx \sum_{i=0}^{n} A_i f(x_i), \quad A_i = \int_a^b w(x) \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{x - x_j}{x_i - x_j} dx$$

$2n+1$次非零多项式f, 用q去除f, 得到商p和余式r, p, r 为n次非零多项式，则有

$$f = qp + r,$$

因此$f(x_i) = r(x_i)$. 根据正交定义，以及积分公式对所有n次非零多项式精确成立，可得

$$\int_a^b f(x)w(x)dx = \int_a^b rwdx = \sum_{i=0}^{n} A_i r(x_i) = \sum_{i=0}^{n} A_i f(x_i)$$

Gauss积分公式计算

对于给定的n，积分区间$[a, b]$和权函数w，可以如下确定Gauss积分公式:

- 采用正交多项式递推公式，计算出$n+1$次关于w正交的多项式，并且计算出它的零点
  - 对于较大的n，需要用数值方法求出所有的根
- 系数A_i可以采用待定系数法确定
- 在许多数学手册中给出了各种类型的积分公式结点以及系数的表格，使用时可以直接查手册

n	x_k	A_k
1	0	2
2	±0.5773502692	1
3	±0.7745966692 0	0.5555555556 0.8888888889
4	±0.8611363116 ±0.3399810436	0.3478548451 0.6521451549
5	±0.9061798459 ±0.5384693101 0	0.2369268851 0.4786286705 0.5688888889

n	x_k	A_k
6	±0.9324695142 ±0.6612093865 ±0.2386191861	0.1713244924 0.3607615730 0.4679139346
7	±0.9491079123 ±0.7415311856 ±0.4058451514 0	0.1294849662 0.2797053915 0.3818300505 0.4179591837
8	±0.9602898565 ±0.7966664774 ±0.5255324099 ±0.1834346425	0.1012285363 0.2223810345 0.3137066459 0.3626837834

正交多项式的零点

- 在Gauss积分定理中需要$n+1$次多项式q的根都落在区间$[a,b]$内，并且都是单根。

Theorem (符号变化次数定理)

设w是$C[a,b]$中正的权函数，并且f是$C[a,b]$中与Π_n关于w正交的非零元，那么f在$[a,b]$上至少变号$n+1$次

证明：由于$1 \in \Pi_n$, 所以$\int_a^b f(x)w(x)dx = 0$, 从而f至少变号一次。假设f在$[a,b]$上变号r次，$r \leqslant n$, 并且变号点t_i满足

$$a = t_0 < t_1 < t_2 < \cdots < t_r < t_{r+1} = b$$

则多项式$p(x) = (x-t_1)\cdots(x-t_r) \in \Pi_n$与$f$在每个区间$[t_i, t_{i+1}]$上恒同号或恒反号，从而$\int_a^b f(x)p(x)w(x)dx \neq 0$, 与正交性矛盾。 □

Legendre多项式

n次多项式

$$L_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n$$

称为Legendre多项式, $\{L_n(x)\}$为$[-1,1]$上的正交多项式系，即

$$(L_n(x), L_m(x)) = \int_{-1}^{1} L_n(x)L_m(x)dx = 0, \quad m \neq n$$

性质

- $L_n(x)$在$(-1,1)$上有n个相异的实根
- $L_n(x)$在$[-1,1]$上正交于任何一个不高于$n-1$次的多项式，即若$P(x)$为一个不高于$n-1$次的多项式，则

$$(L_n(x), P(x)) = \int_{-1}^{1} L_n(x)P(x)dx = 0$$

Legendre多项式

$$
\begin{aligned}
p_0(x) &= 1 \\
p_1(x) &= x \\
p_2(x) &= \frac{3}{2}x^2 - \frac{1}{2} \\
p_3(x) &= \frac{5}{2}x^3 - \frac{3}{2}x \\
p_4(x) &= \frac{35}{8}x^4 - \frac{15}{4}x^2 + \frac{3}{8} \\
p_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x)
\end{aligned}
$$

- Gauss原创性工作的情形：$w(x) = 1$, $[a, b] = [-1, 1]$
- $n = 1$:

$$\int_{-1}^{1} f(x)dx \approx f(-1/\sqrt{3}) + f(1/\sqrt{3})$$

$x_i = \pm 1/\sqrt{3}$为二次Legendre多项式$p_2(x) = \frac{1}{2}(3x^2 - 1)$的零点

- $n = 4$:

$$\int_{-1}^{1} f(x)dx \approx A_0 f(x_0) + A_1 f(x_1) + A_2 f(x_2) + A_3 f(x_3) + A_4 f(x_4)$$

其中x_i为五次Legendre多项式

$$p_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$

的零点

Lemma

在Gauss积分公式中，组合系数为正数，而且它们的和是$\int_a^b w(x)dx$

证明：对于给定的n, Gauss积分公式中的结点是$n+1$次多项式q的零点，其中q关于w与n次非零多项式正交。对于固定的j, 令$p = q/(x - x_j)$, 则$\deg p^2 \leqslant 2n$, 所以积分公式对它精确成立，即

$$0 < \int_a^b p^2(x)w(x)dx = \sum_{i=0}^{n} A_i p^2(x_i) = A_j p^2(x_j)$$

因此可得$A_j > 0$. 由于积分公式对于$f(x) \equiv 1$精确成立，所以

$$\int_a^b w(x)dx = \sum_{i=0}^{n} A_i$$

Theorem

若f在$[a,b]$上连续，则当$n \to \infty$时近似积分公式

$$\int_a^b f(x)w(x)dx \approx \sum_{i=0}^{n} A_{n,i} f(x_{n,i}), \qquad n \geqslant 0$$

收敛于积分

证明：根据Weierstrass定理，对于任意$\varepsilon > 0$, 存在多项式p满足$|f(x) - p(x)| < \varepsilon$, $x \in [a,b]$. 对于任一整数n, $2n > \deg p$, 则n次Gauss积分公式对于p精确成立，从而有(接下页)

$$
\begin{aligned}
&\left| \int_a^b f(x)w(x)dx - \sum_{i=0}^{n} A_{n,i} f(x_{n,i}) \right| \\
\leqslant &\left| \int_a^b f(x)w(x)dx - \int_a^b p(x)w(x)dx \right| \\
&+ \left| \sum_{i=0}^{n} A_{n,i} p(x_{n,i}) - \sum_{i=0}^{n} A_{n,i} f(x_{n,i}) \right| \\
\leqslant &\int_a^b |f(x) - p(x)| w(x)dx + \sum_{i=0}^{n} A_{n,i} |p(x_{n,i}) - f(x_{n,i})| \\
\leqslant &\varepsilon \int_a^b w(x)dx + \varepsilon \sum_{i=0}^{n} A_{n,i} = 2\varepsilon \int_a^b w(x)dx
\end{aligned}
$$

带误差项的Gauss积分定理

Theorem

考虑带误差项的*Gauss*积分公式：

$$\int_a^b f(x)w(x)dx = \sum_{i=0}^{n-1} A_i f(x_i) + E$$

若$f \in C^{2n}[a,b]$, 则有

$$E = \frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b q^2(x)w(x)dx$$

其中$\xi \in (a,b)$, $q(x) = (x - x_0)\cdots(x - x_{n-1})$

证明：应用Hermite插值，存在次数不超过$2n-1$的多项式p, 满足

$$p(x_i) = f(x_i), p'(x_i) = f'(x_i), i = 0, 1, \ldots, n-1$$

这个插值的误差公式为

$$f(x) - p(x) = \frac{1}{(2n)!} f^{(2n)}(\zeta(x)) q^2(x)$$

因此

$$\begin{aligned}
&\int_a^b f(x)w(x)dx - \int_a^b p(x)w(x)dx \\
=&\frac{1}{(2n)!} \int_a^b f^{(2n)}(\zeta(x)) q^2(x) w(x) dx \\
=&\frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b q^2(x) w(x) dx
\end{aligned}$$

再根据p的次数不超过$2n-1$即可得所需要的有误差项的积分公式 □

Gauss积分的应用：无界积分区间情形的积分

- 假设考虑的区间为$[0,+\infty)$
- 为了计算积分

$$\int_0^{\infty} f(x)dx$$

引入权因子e^{-x}, 把积分变形为

$$\int_0^{\infty} \varphi(x)e^{-x}dx$$

- 从而我们考虑在$[0,+\infty)$上的关于权e^{-x}正交的多项式，以它们的零点应用Gauss积分公式

Laguerre多项式

- 即在$[0,+\infty)$上的关于权e^{-x}正交的多项式
- 其定义为

$$\mathcal{L}_n(x) = e^x \frac{d^n}{dx^n}(e^{-x}x^n) \qquad n \geqslant 0$$

- 满足递推关系：

$$\mathcal{L}_{n+1}(x) = (2n+1-x)\mathcal{L}_n(x) - n^2\mathcal{L}_{n-1}(x), \quad n \geqslant 0$$
$$\mathcal{L}_{-1} = 0, \quad \mathcal{L}_0 = 1$$

$(-\infty,+\infty)$上的积分与Hermite多项式

- 如果考虑$(-\infty,+\infty)$上的积分，那么引入权因子e^{-x^2}
- 在$(-\infty,+\infty)$上关于e^{-x^2}正交的多项式称为Hermite多项式
  1. 定义：
  $$\mathcal{H}_n(x)=(-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2},\qquad n\geqslant 0$$
  2. 递推关系：
  $$\mathcal{H}_{n+1}(x)=2x\mathcal{H}_n(x)-2n\mathcal{H}_{n-1}(x),\quad n\geqslant 0$$
  $$\mathcal{H}_{-1}=0,\quad \mathcal{H}_0=1$$

Gauss型求积公式

- Gauss-Legendre求积公式
  区间$[-1,1]$上权函数$w(x)=1$的Gauss型求积公式,称为Gauss-Legendre求积公式,其Gauss点为Legendre多项式的零点.
- Gauss-Laguerre求积公式
  区间$[0,+\infty)$上的权函数为$w(x)=e^{-x}$的Gauss型求积公式,称为Gauss-Laguerre求积公式, 其Gauss点为Laguerre多项式的零点.
- Gauss-Hermite求积公式
  区间$(-\infty,+\infty)$上的权函数为$w(x)=e^{-x^2}$的Gauss型求积公式,称为Gauss-Hermite求积公式, 其Gauss点为Hermite多项式的零点.

上机作业

- 利用复化梯形积分公式和复化3点Gauss积分公式计算积分的通用程序计算下列积分

$$I_1(f) = \int_0^1 e^{-x^2} dx,\ I_2(f) = \int_0^4 \frac{1}{1+x^2} dx,$$
$$I_3(f) = \int_0^{2\pi} \frac{1}{2+\cos(x)} dx$$

取节点$x_i, i = 0, \cdots, N$, N为$2^k, k = 1, \cdots, 7$,给出如下的误差表格,其中阶为$\frac{\ln(Error_{old}/Error_{now})}{\ln(N_{now}/N_{old})}$.

	$I_1(f)$		$I_2(f)$		$I_3(f)$	
N	误差	阶	误差	阶	误差	阶
2						
4						
8						
16						

- 简单分析你得到的数据

《数值分析》之

数值积分

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

递推梯形法则

- 用$T_n(f)$表示在长度为$h=(b-a)/n$的n个子区间上积分$I=\int_a^b f(x)dx$的复化梯形法则. 对于区间$[a,b]=[0,1]$,

$$
\begin{aligned}
T_1(f) =& \frac{1}{2}f(0)+\frac{1}{2}f(1)\\
T_2(f) =& \frac{1}{4}f(0)+\frac{1}{2}f\left(\frac{1}{2}\right)+\frac{1}{4}f(1)\\
T_4(f) =& \frac{1}{8}f(0)+\frac{1}{4}\left[f\left(\frac{1}{4}\right)+f\left(\frac{1}{2}\right)+f\left(\frac{3}{4}\right)\right]+\frac{1}{8}f(1)\\
T_8(f) =& \frac{1}{16}f(0)+\frac{1}{8}\left[f\left(\frac{1}{8}\right)+f\left(\frac{1}{4}\right)+f\left(\frac{3}{8}\right)+f\left(\frac{1}{2}\right)+f\left(\frac{5}{8}\right)\right.\\
&\left.+f\left(\frac{3}{4}\right)+f\left(\frac{7}{8}\right)\right]+\frac{1}{16}f(1)
\end{aligned}
$$

- 为了计算$T(2n)$，可以利用$T(n)$计算中已有的结果，从而只需要计算那些出现在$T(2n)$，没有出现在$T(n)$中的项。

$$
\begin{aligned}
T_2(f) &= \frac{1}{2}T_1(f) + \frac{1}{2}f\left(\frac{1}{2}\right) \\
T_4(f) &= \frac{1}{2}T_2(f) + \frac{1}{4}\left[f\left(\frac{1}{4}\right) + f\left(\frac{3}{4}\right)\right] \\
T_8(f) &= \frac{1}{2}T_4(f) + \frac{1}{8}\left[f\left(\frac{1}{8}\right) + f\left(\frac{3}{8}\right) + f\left(\frac{5}{8}\right) + f\left(\frac{7}{8}\right)\right]
\end{aligned}
$$

- 令$h = \frac{b-a}{2n}$，则在一般区间$[a, b]$上的一般公式为

$$
\begin{aligned}
T_{2n}(f) &= \frac{1}{2}T_n(f) + h[f(a+h) + f(a+3h) + \cdots + f(a+(2n-1)h)] \\
&= \frac{1}{2}T_n(f) + h\sum_{i=1}^{n} f(a+(2i-1)h)
\end{aligned}
$$

积分的自适应计算

- 函数变化有急有缓，为了照顾变化剧烈部分的误差，我们需要加密格点。
- 对于变化缓慢的部分，加密格点会造成计算的浪费。
- 以此我们介绍一种算法，可以自动在变化剧烈的地方加密格点计算，而变化缓慢的地方，则取稀疏的格点。

事后误差估计

- 复化梯形公式

$$I(f) - T_n(f) = -\frac{b-a}{12}h^2 f''(\xi), \quad \text{n等分区间}$$

$$I(f) - T_{2n}(f) = -\frac{b-a}{12}(\frac{h}{2})^2 f''(\eta), \quad \text{2n等分区间}$$

- 近似有：$f''(\eta) \approx f''(\xi)$
- 从而得到

$$I(f) - T_{2n}(f) \approx \frac{1}{3}(T_{2n}(f) - T_n(f))$$

- 误差可以用2组复化梯形公式的差来估计

外推

由前面的事后误差估计式

$$I(f) - T_{2n}(f) \approx \frac{1}{3}(T_{2n}(f) - T_n(f)),$$

则

$$I(f) \approx T_{2n}(f) + \frac{1}{3}(T_{2n}(f) - T_n(f)) = S_{2n}(f),$$

可以用低阶的公式组合后成为一个高阶的公式，截断误差由$O(h^2)$提高到$O(h^4)$。这种手段称为外推算法。

记$I(h)$为以步长为h的数值积分公式，有

$$I(f) - I(h) = ch^m + O(h^{m+1}),$$

$$I(f) - I(\frac{h}{2}) = c(\frac{h}{2})^m + O((\frac{h}{2})^{m+1}),$$

$$I(f) - I(\frac{h}{2}) \approx \frac{I(\frac{h}{2}) - I(h)}{2^m - 1},$$

$$I(f) \approx I(\frac{h}{2}) + \frac{I(\frac{h}{2}) - I(h)}{2^m - 1}.$$

基于Euler-Maclaurin公式，可以建立起新的积分外推公式。

- Euler-Maclaurin公式: 对于$f \in C^{2m}[0,1]$,

$$\int_0^1 f(t)dt = \frac{1}{2}[f(0)+f(1)] + \sum_{k=1}^{m-1} A_{2k}[f^{(2k-1)}(0) - f^{(2k-1)}(1)] - A_{2m}f^{(2m)}(\xi_0)$$

其中$\xi_0 \in (0,1)$, $k!A_k$称为Bernoulli常数，由下式定义:

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} A_k x^k$$

- 进行变量代换，Euler-Maclaurin公式变为

$$\int_{x_i}^{x_{i+1}} f(x)dx = \frac{h}{2}[f(x_i) + f(x_{i+1})] + \sum_{k=1}^{m-1} A_{2k} h^{2k} [f^{(2k-1)}(x_i) - f^{(2k-1)}(x_{i+1})] - A_{2m} h^{2m+1} f^{(2m)}(\xi_i)$$

- 令$x_i = a + ih,\ i = 0, 1, \ldots, 2^n,\ h = (b-a)/2^n$, 进行求和：

$$\int_a^b f(x)dx = \frac{h}{2} \sum_{i=0}^{2^n-1} [f(x_i) + f(x_{i+1})] + \sum_{k=1}^{m-1} A_{2k} h^{2k} [f^{(2k-1)}(a) - f^{(2k-1)}(b)] - A_{2m}(b-a) h^{2m} f^{(2m)}(\xi)$$

- 从而有

$$I = T(2^n) + c_2 h^2 + c_4 h^4 + \cdots + c_{2m-2} h^{2m-2} + c_{2m} h^{2m} f^{(2m)}(\xi)$$

这样我们可以应用Richardson外推技术，得到公式：

Euler-Maclaurin定理

若$I(f) = I^{(m)}(h) + O(h^{2m})$为$2m$阶公式，则

$$I^{(m+1)}(\frac{h}{2}) = I^{(m)}(\frac{h}{2}) + \frac{I^{(m)}(\frac{h}{2}) - I^{(m)}(h)}{2^{2m} - 1} + O(h^{2m+2})$$

Romberg 积分就是不断地用如上定理组合低阶公式为高阶公式，进而计算积分

$$R(n, 0) = T_{2^n}(f)$$

$$R(n, m) = R(n, m-1) + \frac{1}{4^m - 1}[R(n, m-1) - R(n-1, m-1)]$$

Romberg算法

- 对一个适当的M, 按如下公式计算$R(i,j)$, $i = 0, 1, \ldots, M$, $j = 0, 1, \ldots, M$:

$$R(0,0) = \frac{1}{2}(b-a)[f(a)+f(b)]$$

$$R(n,0) = \frac{1}{2}R(n-1,0) + h_n \sum_{i=1}^{2^{n-1}} f(a+(2i-1)h_n)$$

$$R(n,m) = R(n,m-1) + \frac{1}{4^m - 1}[R(n,m-1) - R(n-1,m-1)]$$

其中$h_0 = b-a$, $h_n = h_{n-1}/2$

- 在精致的算法中，应当选取适度的M, 并且加上一个自动终止程序，当达到指定的误差标准时停止计算

Romberg阵列

$$
\begin{array}{lllllll}
R(0,0) & & & & & & \\
R(1,0) & R(1,1) & & & & & \\
R(2,0) & R(2,1) & R(2,2) & & & & \\
R(3,0) & R(3,1) & R(3,2) & R(3,3) & & & \\
R(4,0) & R(4,1) & R(4,2) & R(4,3) & R(4,4) & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \\
R(M,0) & R(M,1) & R(M,2) & R(M,3) & R(M,4) & \cdots & R(M,M)
\end{array}
$$

收敛性定理

- 在Romberg算法中，为了应用Euler-Maclaurin公式，我们需要$f \in C^{2m}[a,b]$, 从而$R(n,m)$收敛于f的积分，并且误差为$\mathcal{O}(h^{2m})$
- 如果f只是连续，我们有下面的定理

Theorem

若$f \in C[a,b]$, 则*Romberg*阵列中每一列都收敛于f的积分，即对每个m,

$$\lim_{n\to\infty} R(n,m) = \int_a^b f(x)dx := I$$

采用归纳法。对于第一列，它是积分I的梯形估计。而具有k个子区间的梯形法则可以写成

$$\frac{1}{2}h\sum_{i=0}^{k-1}f(a+ih)+\frac{1}{2}h\sum_{i=1}^{k}f(a+ih)$$

这是两个Riemann和的平均。由于$h=(b-a)/k$, 所以当$k\to\infty$时子区间的长度趋向于零。从而根据Riemann积分理论，两个Riemann和都趋向于I, 从而它们的平均值也趋向于I. 这就证明了

$$\lim_{n\to\infty}R(n,0)=I$$

假设对于$m-1$结论成立，则由于

$$R(n,m)=R(n,m-1)+\frac{1}{4^m-1}[R(n,m-1)-R(n-1,m-1)]$$

所以

$$\lim_{n\to\infty}R(n,m)=\frac{4^m}{4^m-1}I-\frac{1}{4^m-1}I=I$$

《数值分析》之

数值积分

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

Bernoulli多项式

- Bernoulli多项式是由下列等式定义

$$\sum_{k=0}^{n} \binom{n+1}{k} B_k(t) = (n+1)t^n$$

- 最初的几个Bernoulli多项式是

$$B_0(t) = 1$$
$$B_1(t) = t - \frac{1}{2}$$
$$B_2(t) = t^2 - t + \frac{1}{6}$$
$$B_3(t) = t^3 - \frac{3}{2}t^2 + \frac{1}{2}t$$

Bernoulli多项式性质

1. $B_n' = nB_{n-1},\ (n \geqslant 1).$
2. $B_n(t+1) - B_n(t) = nt^{n-1},\ (n \geqslant 2).$
3. $B_n(t) = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} B_k(0) t^{n-k}.$
4. $B_n(1-t) = (-1)^n B_n(t).$

Theorem

函数$G(t) = B_{2n}(t) - B_{2n}(0)$在开区间$(0,1)$中没有零点。

证明：有性质2和4，令$t = 0$得到

$$B_n(0) = B_n(1) = (-1)^n B_n(0)$$

从而有$B_3(0) = B_5(0) = B_7(0) = \cdots = 0$.
反证法。假设$G(t)$在开区间$(0,1)$中有一个零点。
由$G(0) = G(1) = 0$,由Rolle中值定理知$G'(t)$在$(0,1)$中有2个零点.
$\because G'(t) = B'_{2n}(t) = 2nB_{2n-1}(t), \Longrightarrow B_{2n-1}(t)$在$(0,1)$中有2个零点.
又$\because B_{2n-1}(0) = B_{2n-1}(1) = 0$,
$\Longrightarrow B'_{2n-1}(t) = (2n-1)B_{2n-2}(t)$在$(0,1)$中有3个零点.
由此可知,对所有奇数指标$k < 2n$, B_k在$(0,1)$中至少有2个零点.
因此, B_3除0,1两个零点外,在$(0,1)$中还有2个零点. 而B_3为三次多项式, 这显然是不可能的.

Euler-Maclaurin公式

Theorem

对于$f \in C^{2m}[0,1]$,

$$\int_0^1 f(t)dt = \frac{1}{2}[f(0)+f(1)] + \sum_{k=1}^{m-1} \frac{b_{2k}}{(2k)!}[f^{(2k-1)}(0) - f^{(2k-1)}(1)] + R$$

其中

$$b_k = B_k(0)$$

$$R = -\frac{b_{2m}}{(2m)!} f^{(2m)}(\xi),\ (0 < \xi < 1).$$

《数值分析》之

数值积分

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

线性泛函

- 线性空间上的一个线性泛函是指由线性空间到数域(一般为$\mathbb{R}$)的一个线性映射
- 若线性空间为$C[a,b]$, 常用的两种线性泛函为
  1. 定积分泛函：
  $$\varphi(f) = \int_a^b f(x)dx$$
  2. 在数值计算中最基本的泛函是如下的点赋值泛函：取定$x \in [a,b]$,
  $$\hat{x}(f) = f(x)$$
  定义了线性泛函$\hat{x}$. 利用点泛函的线性组合，得到
  $$\psi = \sum_{i=0}^{n} c_i \hat{x}_i, \qquad \text{即}\psi(f) = \sum_{i=0}^{n} c_i f(x_i)$$
- 点泛函的线性组合是数值计算中可直接计算泛函中最一般的类型，其它泛函要用这样的ψ进行逼近

逼近泛函

- 1940年到1970年间，Arthur Sard发展了逼近泛函的相关理论，而且最终与自然样条有着有趣的联系
- 被逼近泛函的定义如下：

$$\varphi(f) = \sum_{i=0}^{N} \left\{ \int_a^b \alpha_i(x) f^{(i)}(x) dx + \sum_{j=1}^{n} \beta_{ij} f^{(i)}(z_{ij}) \right\}$$

其中$z_{ij} \in [a, b]$, $\alpha_i(x)$在$[a, b]$上分段连续, $f \in C^N[a, b]$

Peano核

- 上页定义的线性泛函的m阶Peano核是如下函数：

$$K_m(t) = \frac{1}{m!}\varphi_x[(x-t)_+^m]$$

其中$m \geqslant N$, φ_x表示泛函作用到关于x的函数上，x_+^m就是截断幂函数

$$x_+^m = \begin{cases} x^m & x \geqslant 0 \\ 0 & x < 0 \end{cases}$$

- 如果对任意$f \in W$, $\varphi(f) = 0$, 则称泛函φ零化空间W

考虑如下定义的泛函：

$$\varphi(f) = \int_0^{\pi} f'(x)\cos x\, dx$$

这是前面一般形式泛函在$N = 1$, $\alpha_1(x) = \cos x$, $n = 0$时的情形。确定φ的Peano核K_1

- $$\frac{d}{dx}(x-t)_+^m = m(x-t)_+^{m-1}, \qquad m \geqslant 1$$
- 对于本题，

$$\begin{aligned} K_1(t) = \varphi_x[(x-t)_+^1] &= \int_0^{\pi} (\cos x)\frac{d}{dx}(x-t)_+ dx \\ &= \int_0^{\pi} (\cos x)(x-t)_+^0\, dx \\ &= \int_t^{\pi} \cos x\ dx = \sin t \end{aligned}$$

Peano核定理

Theorem

若前面定义的一般泛函φ零化Π_m, 则对所有的$f \in C^{m+1}[a,b]$,

$$\varphi(f) = \int_a^b K_m(t) f^{(m+1)}(t) dt$$

其中$m \geqslant N$

证明：带积分余项的Taylor展开定理为

$$f(x) = \sum_{k=0}^{m} \frac{1}{k!} f^{(k)}(a)(x-a)^k + r(x),$$

$$r(x) = \frac{1}{m!} \int_a^x f^{(m+1)}(t)(x-t)^m dt$$

由于φ零化Π_m, 所以$\varphi(f) = \varphi(r)$. 而r可以重写为

$$r(x) = \frac{1}{m!}\int_a^b f^{(m+1)}(t)(x-t)_+^m dt$$

因此

$$\varphi(r) = \frac{1}{m!}\int_a^b f^{(m+1)}(t)\varphi_x[(x-t)_+^m]dt$$

□

注意把φ_x移到积分号内，需要用到积分交换顺序以及积分与求导交换顺序等微积分定理, 所定义的φ是满足这些条件的

Sard在1963年给出的例：求出如下定义泛函φ

$$\varphi(f) = \int_0^1 f(x)x^{-1/2}dx$$

的形式为$\psi(f) = c_1 f(0) + c_2 f(1)$的逼近，它对于$\Pi_1$精确成立。给出逼近误差。

- 我们要求$\varphi - \psi$零化Π_1, 因此可以用待定系数法确定系数：

$$\varphi(1) - \psi(1) = \int_0^1 x^{-1/2}dx - (c_1 + c_2) = 2 - c_1 - c_2 = 0$$

$$\varphi(x) - \psi(x) = \int_0^1 \sqrt{x}dx - c_2 = \frac{2}{3} - c_2 = 0$$

因此$c_1 = \frac{4}{3}$, $c_2 = \frac{2}{3}$

- 泛函$\varphi-\psi$的Peano核K_1为

$$
\begin{aligned}
(\varphi_x-\psi_x)(x-t)_+^1 &= \int_0^1 (x-t)_+ x^{-1/2}dx - \frac{4}{3}(0-t)_+ - \frac{2}{3}(1-t)_+ \\
&= \int_t^1 (x-t)x^{-1/2}dx - \frac{2}{3}(1-t) \\
&= \frac{4}{3}t(\sqrt{t}-1)
\end{aligned}
$$

- 因此根据Peano核定理：

$$
\int_0^1 f(x)x^{-1/2}dx - \left[\frac{4}{3}f(0)+\frac{2}{3}f(1)\right] = \int_0^1 \frac{4}{3}t(\sqrt{t}-1)f''(t)dt
$$

当$f\in C^2[0,1]$时，这里等号右边项即为误差。进一步应用积分中值定理：

$$
\int_0^1 \frac{4}{3}t(\sqrt{t}-1)f''(t)dt = f''(\xi)\int_0^1 \frac{4}{3}t(\sqrt{t}-1)dt = -\frac{2}{15}f''(\xi)
$$

Sard意义下的最佳逼近

- 如果φ和ψ是在Π_m上相同的两个泛函，那么根据Cauchy-Scharwz不等式

$$|\varphi(f)-\psi(f)| \leqslant \left\|K_m\right\|_2 \left\|f^{(m+1)}\right\|_2$$

- 在前面示例中，如果ψ的系数不能完全由$\varphi-\psi$零化Π_m得到，那么通过极小化$\left\|K_m\right\|_2^2=\int_a^b[K_m(t)]^2dt$来选取这些参数，由此得到的泛函称为Sard意义下$\varphi$的一个最佳逼近
- Schoenberg发现可以用自然样条得到这种最佳逼近

Sard逼近的Schoenberg定理

Theorem

设φ为如前定义的一般线性泛函。给定结点$a = t_0 < t_1 < \cdots < t_n = b$, $n > N$. 在所有形如$\sum_{i=0}^{n} c_i \hat{t}_i$, 并且在$\Pi_m$上与$\varphi$相同的泛函中，*Sard*意义下$\varphi$的最佳逼近是$\varphi \circ L$, 其中$L(f)$是在给定结点上插值$f$的$2m+1$次自然样条

证明：令

$$\psi = \sum_{i=0}^{n} c_i \hat{t}_i$$

并且假设对任意$p \in \Pi_m$, $\psi(p) = \varphi(p)$, 即$\varphi - \psi$零化Π_m。
记K_m为$\varphi - \psi$的Peano核。
如果f是给定结点上的$2m+1$次自然样条，那么$Lf = f$. 而次数$\leqslant m$的多项式也是自然样条，因此对$p \in \Pi_m$, $Lp = p$. 所以$\varphi - \varphi \circ L$也零化$\Pi_m$. 我们的目标就是证明$\psi = \varphi \circ L$.

设$\overline{K}_m$为$\varphi - \varphi \circ L$的Peano核，那么下面证明

$$\int_a^b \left[\overline{K}_m(t)\right]^2 dx \leqslant \int_a^b [K_m(t)]^2 dt$$

就可以完成证明。
泛函$\theta = \varphi \circ L - \psi = (\varphi - \psi) - (\varphi - \varphi \circ L)$的Peano核为$\overline{\overline{K}}_m = K_m - \overline{K}_m$, 它具有形式

$$\overline{\overline{K}}_m(t) = \frac{1}{m!}\theta_x[(x-t)_+^m]$$

如果我们能证明$\langle \overline{\overline{K}}_m, \overline{K}_m \rangle = 0$, 那么就可以得到所需要的结论。
实际上，为此首先证明满足$g^{(m+1)} = \overline{\overline{K}}_m$的函数$g$是自然样条，所以$Lg = g$, 从而

$$\int_a^b \overline{\overline{K}}_m \overline{K}_m dt = \int_a^b \overline{K}_m g^{(m+1)} dt = (\varphi - \varphi \circ L)(g) = 0$$

证明最后的关键点：g是自然样条

- 设$s_0, s_1, \ldots, s_n$是自然样条空间的一组插值基函数，即满足$s_i(t_j) = \delta_{ij}$, 那么L具有形式

$$Lf = \sum_{i=0}^{n} f(t_i)s_i$$

- 因此可得θ的形式为

$$\theta(f) = \varphi(Lf) - \psi(f) = \sum_{i=0}^{n} f(t_i)\varphi(s_i) - \sum_{i=0}^{n} c_i f(t_i) = \sum_{i=0}^{n} \gamma_i f(t_i)$$

- 所以$\overline{\overline{K}}_m$的形式为

$$\overline{\overline{K}}_m(t) = \frac{1}{m!} \sum_{i=0}^{n} \gamma_i (t_i - t)_+^m$$

- 令函数g满足$g^{(m+1)} = \overline{\overline{K}}_m$, 则可证$g$为$2m+1$次自然样条
  1. g为$2m+1$次样条：来自于$\overline{\overline{K}}_m$为$m$次样条
  2. $t \geqslant b$时$g^{(m+1)}(t) = 0$, 这是由于$\overline{\overline{K}}_m(t) = 0,\ t \geqslant b$
  3. $t \leqslant a$时$g^{(m+1)}(t) = 0$，因为此时

$$\overline{\overline{K}}_m(t) = \frac{1}{m!}\theta_x[(x-t)^m]$$

而θ零化$\prod_m$, 所以有此结论

□

$$\varphi(f) = \int_{-1}^{1} f(x)dx$$
$$\psi(f) = c_1 f(-1) + c_2 f(0) + c_3 f(1)$$

两者在Π_1上一致，ψ为φ在Sard意义下的最佳逼近，确定ψ的形式

- 此时L的形式为

$$(Lf)(x) = a_0 + a_1 x + b_0(x+1)_+^3 + b_1(x)_+^3 + b_2(x-1)_+^3$$

其中

$$a_0 = \frac{1}{4}[-f(-1) + 6f(0) - f(1)]$$
$$a_1 = \frac{1}{4}[-5f(-1) + 6f(0) - f(1)]$$
$$b_0 = b_2 = -b_1/2 = \frac{1}{4}[f(-1) - 2f(0) + f(1)]$$

- 因此最佳公式为

$$\begin{aligned}\psi(f) = \varphi(Lf) &= \int_{-1}^{1} (Lf)(x)dx \\ &= 2a_0 + 4b_0 + \frac{1}{4}b_1 \\ &= \frac{3}{8}f(-1) + \frac{5}{4}f(0) + \frac{3}{8}f(1)\end{aligned}$$

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

微分方程

- 微分方程是研究自然科学和社会科学中的事物、物体和现象运动、演化和变化规律的最为基本的数学理论和方法。
- 物理、化学、生物、工程、航空航天、医学、经济和金融领域中的许多原理和规律都可以描述成适当的常微分方程，如牛顿的运动定律、万有引力定律、机械能守恒定律，能量守恒定律、人口发展规律、生态种群竞争、疾病传染、遗传基因变异、股票的涨伏趋势、利率的浮动、市场均衡价格的变化等，对这些规律的描述、认识和分析就归结为对相应的常微分方程描述的数学模型的研究。
- 因此，微分方程的理论和方法不仅广泛应用于自然科学，而且越来越多的应用于社会科学的各个领域。
- 一般情况下，这些微分方程都需要用数值方法去求解。

初值问题

- 如下方程称为初值问题：

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

其中y是x的未知函数

- 例：

$$\begin{cases} y' = y\tan(x+3) \\ y(-3) = 1 \end{cases}$$

要在一个包含初始点$x_0 = -3$的区间上确定$y(x)$. 这个方程的解析解是$y(x) = \sec(x+3)$. 由于$\sec x$在$x = \pm\pi/2$时变成无穷，因此解仅对$-\pi/2 < x+3 < \pi/2$成立

- 例子说明：对于初值问题，如果$f(x, y)$只在(x_0, y_0)的一个邻域内连续，那么解的存在区域肯定局限在这个邻域内。但是，即使$f(x, y)$是处处连续的，那么解的存在区域也可能是有限的.

- 例：

$$\begin{cases} y' = 1 + y^2 \\ y(0) = 0 \end{cases}$$

1. 直接性态分析: 解曲线从$x = 0$出发，斜率为1, 因此$y(x)$在$x = 0$时递增，所以$1 + y^2$也递增，所以我们可以期望$y(x)$有一个垂直渐近线
2. 实际上，方程的解析解为$y(x) = \tan x$, 所以在$x = \pi/2$处出现渐近线

Theorem (初值问题的第一存在性定理)

若f在中心为(x_0, y_0)的矩形

$$R = \{(x, y) : |x - x_0| \leqslant \alpha, |y - y_0| \leqslant \beta\}$$

内连续，则在$|x - x_0| \leqslant \min(\alpha, \beta/M)$内初值问题有解$y(x)$, 其中$M$是$|f(x, y)|$在矩形$R$内的最大值

讨论下述初值问题

$$\begin{cases} y' = (x + \sin y)^2 \\ y(0) = 3 \end{cases}$$

解的存在区域

- $f(x, y) = (x + \sin y)^2$, $(x_0, y_0) = (0, 3)$
- 在矩形

$$\{(x, y) : |x - x_0| \leqslant \alpha, |y - y_0| \leqslant \beta\}$$

内f满足

$$|f(x, y)| \leqslant (\alpha + 1)^2 := M$$

- 所以该问题的解在整个实轴上存在，因为可以通过选择足够大的β, 使得$\min(\alpha, \beta/M)$为任意正数

唯一性

- 在初值问题中，即使f为连续函数，那么也有可能出现多解的情形
- 例：

$$\begin{cases} y' = y^{2/3} \\ y(0) = 0 \end{cases}$$

显然$y(x) \equiv 0$是这个问题的解，而$y(x) = x^3/27$也是一个解

- 因此我们需要对f再多做一些假设

Theorem (初值问题解的唯一性定理)

若$f(x, y)$和$\partial f(x, y)/\partial y$在矩形

$$R = \{(x, y) : |x - x_0| \leqslant \alpha, |y - y_0| \leqslant \beta\}$$

内连续，则初值问题在区间$|x - x_0| < \min(\alpha, \beta/M)$内有唯一解

Theorem (初值问题解的第二存在性定理)

若f在$a \leqslant x \leqslant b, -\infty < y < +\infty$内连续并且满足不等式

$$|f(x, y_1) - f(x, y_2)| \leqslant L|y_1 - y_2| \tag{1}$$

则初值问题在区间$[a, b]$上有唯一解

- 不等式(1)称为关于第二个变量的Lipschitz条件。对于单变量函数，它简化为

$$|g(x_1) - g(x_2)| \leqslant L|x_1 - x_2|$$

这个条件比连续性更强，但它比可导弱些。因为可导肯定满足Lipschitz条件

证明函数

$$g(x) = \sum_{i=1}^{n} a_i |x - w_i|$$

满足Lipschitz条件

- 容易验证有下述不等式：

$$\begin{aligned}
|g(x_1) - g(x_2)| &= \left| \sum_{i=1}^{n} a_i |x_1 - w_i| - \sum_{i=1}^{n} a_i |x_2 - w_i| \right| \\
&= \left| \sum_{i=1}^{n} a_i \left[|x_1 - w_i| - |x_2 - w_i| \right] \right| \\
&\leqslant \sum_{i=1}^{n} |a_i| \Big| |x_1 - w_i| - |x_2 - w_i| \Big| \\
&\leqslant \sum_{i=1}^{n} |a_i| \cdot |x_1 - x_2|
\end{aligned}$$

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

求解目标

- 对于微分方程，很少能直接得到显式解，通常要采用数值方法
- 常微分方程的解是一个函数，但是，计算机没有办法对函数进行运算。
- 常微分方程的数值解并不是求函数的近似，而是求解函数在某些节点的近似值。
- 通常要求构造下列形式的函数值表格：

x_0	x_1	x_2	x_3	$\cdots$	x_m
y_0	y_1	y_2	y_3	$\cdots$	y_m

其中y_i是在x_i的精确解$y(x_i)$的近似计算值

- 因此常微分方程数值解的目标就是产生上面那样的表格

考察初值问题

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

我们对区间做等距分割：$x_i = hi,\ h = (b-a)/m$. 设解函数在节点的近似为$\{y_i\}$，则

$$y'|_{x=x_i} = f(x, y)|_{x=x_i}$$

由数值微分公式，我们有

$$\frac{y_{i+1} - y_i}{h} \approx f(x_i, y_i)$$

$$y_{i+1} = y_i + hf(x_i, y_i)$$

可以看到，给出初值，就可以用上式求出所有的$\{y_i\}$.

基本步骤如下：

- 对区间作分割：$a = x_0 < x_1 < \cdots < x_n = b$ 求$y(x)$在x_i的近似值y_i, 称为分割上的格点函数
- 由微分方程出发，建立求格点函数的差分方程。这个方程满足：
  - 解存在唯一
  - 稳定，收敛
  - 相容
- 解差分方程，求出格点函数

数值离散方法

为了考察数值方法提供的数值解，是否有实用价值，需要知道如下几个结论：

- 收敛性问题
  步长充分小时，所得到的数值解能否逼近问题的真解
- 误差估计
- 稳定性问题
  舍入误差，在以后各步的计算中，是否会无限制扩大

Euler方法（向前差商公式）

做等距分割，利用数值微分代替导数项，建立差分方程。

- 形式为：

$$y_{n+1} = y_n + hf(x_n, y_n)$$

- 显示格式：由y_n直接算出y_{n+1}
- 优点：不需要对f求导数
- 缺陷：为了得到满意的精度，需要较小的h
- 由于该方法只需要在存在性定理成立的基础上就可以采用，因此具有理论上的重要意义

误差

在微分方程数值解中会出现若干种类型的误差。一种分类方法如下：

- 局部截断误差
- 局部舍入误差
- 整体截断误差
- 整体舍入误差
- 总误差

Euler方法的收敛性

局部截断误差

在假设$y_i = y(x_i)$，即第i步计算是精确的前提下，考虑的截断误差$R_i = y(x_{i+1}) - y_{i+1}$ 称为局部截断误差

- 局部截断误差:

$$y(x_{n+1}) = y(x_n) + hf(x_n, y(x_n)) + \frac{h^2}{2}y''(\xi_n)$$

$$hT_{n+1} = \frac{h^2}{2}y''(\xi_n) = \mathcal{O}(h^2)$$

- 这个误差在逐步计算过程中会传播，积累。
- 这类误差出现在数值解的每一步
- 局部截断误差是所选用方法固有的，与舍入误差完全无关

局部舍入误差

- 局部舍入误差是在每一步计算过程中由于计算机的有限精度而引起的误差，它的值与计算机的字长有关(即与浮点机器数尾数中的位数有关)
- 在向机器中输入数据时会发生舍入误差，算术运算后也会发生舍入误差
- 通常的舍入模式是舍入到最接近数：选择实数左右两边较近的那个机器数。在距离相同时，采用舍入到偶数
- 也可以采用其它的舍入模式：向零舍入(也称截断)，向$+\infty$舍入，向$-\infty$舍入

整体截断误差

- 许多局部截断误差的全体累积起构成整体截断误差
- 整体舍入误差是前面步骤中局部舍入误差的累积
- 总误差是整体截断误差和整体舍入误差的和
  - 即使所有的计算都是精确的值，这个误差还是会出现
  - 它与方法有关，而与执行计算的计算机无关
  - 若局部截断误差是$\mathcal{O}(h^{p+1})$, 则整体截断误差必定是$\mathcal{O}(h^p)$

精度

若某算法的局部截断误差为$\mathcal{O}(h^{p+1})$，则称该算法有p阶精度。

整体截断误差(续)

$$
\begin{aligned}
|e_{n+1}| &= |y(x_{n+1}) - y_{n+1}| \\
&\le |y(x_n) - y_n| + h|f(x_n, y(x_n)) - f(x_n, y_n)| + h|T_{n+1}| \\
&\le |e_n| + hL|y(x_n) - y_n| + h|T_{n+1}| \\
&\le (1 + hL)|e_n| + hT, \qquad T = \max_j |T_j| \\
&\le (1 + hL)((1 + hL)|e_{n-1}| + hT) + hT \\
&\le \cdots \\
&\le (1 + hL)^{n+1}|e_0| + ((1 + hL)^n + \cdots + (1 + hL) + 1)hT \\
&= (1 + hL)^{n+1}|e_0| + \frac{1 - (1 + hL)^{n+1}}{1 - (1 + hL)}hT \\
&\le (1 + hL)^{n+1}|e_0| + \frac{(1 + hL)^{n+1}}{hL}hT \\
&\le (1 + hL)^{n+1}(|e_0| + \frac{T}{L})
\end{aligned}
$$

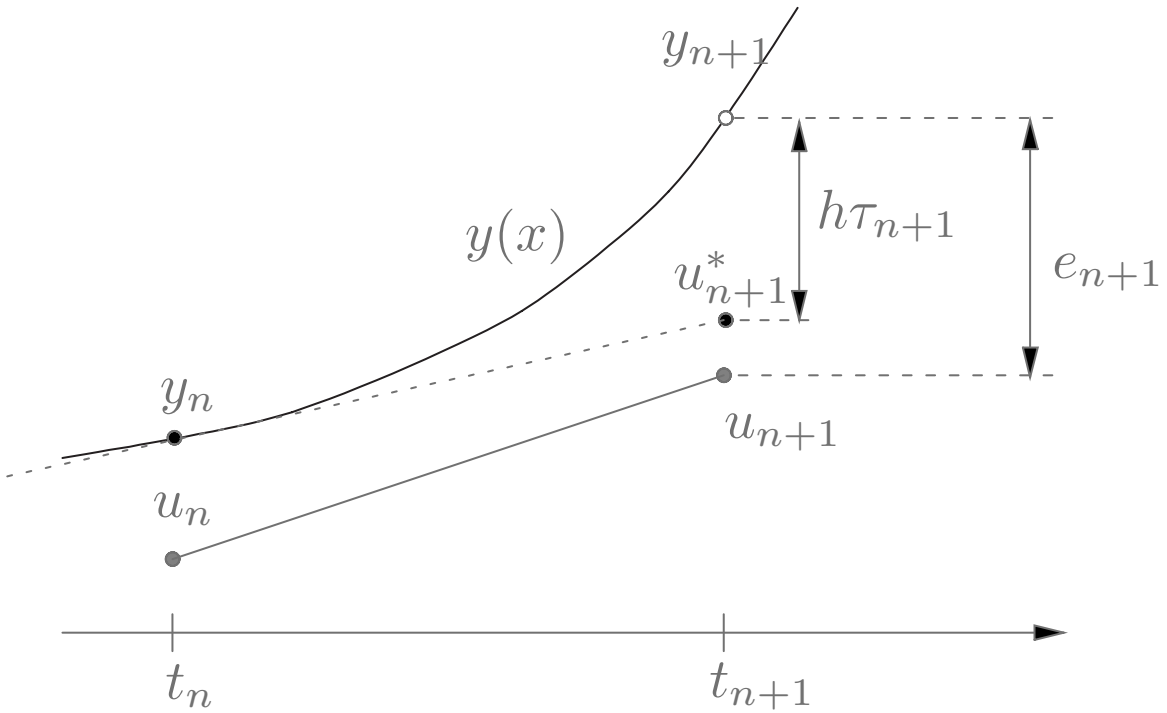
整体截断误差(续)

由$(1+x)^n < e^{nx}$,

$$|e_{n+1}| \leq e^{(n+1)hL}\frac{T}{L}$$

且$T = \mathcal{O}(h)$,故有

$$e_{n+1} = \mathcal{O}(h)$$



稳定性

- 误差在以后各步的计算中不会无限制扩大。
- 考虑简单情况：仅初值有误差，而其他计算步骤无误差。
- 设$\{z_i\}$是初值有误差后的计算值，则

$$y_{n+1} = y_n + hf(x_n, y_n)$$
$$z_{n+1} = z_n + hf(x_n, z_n)$$

所以，我们有

$$\begin{aligned}
|e_{n+1}| &= |y_{n+1} - z_{n+1}| \\
&\leq |e_n| + h|f(x_n, y_n) - f(x_n, z_n)| \\
&\leq |e_n| + hL|y_n - z_n| \\
&= |e_n|(1 + hL) \\
&\leq \cdots \leq |e_0|(1 + hL)^{n+1} \\
&\leq |e_0| \exp((n+1)hL)
\end{aligned}$$

- 向前差商公式关于初值是稳定的。当初始误差充分小，以后各步的误差也充分小

Euler方法（向后差商公式）

- 形式为：

$$y_{n+1} = y_n + hf(x_{n+1}, y_{n+1})$$

- 隐式格式：通常$f(x,y)$是关于y_{n+1}的非线性方程，需要通过迭代法求得y_{n+1}

$$y_{n+1}^{(0)} = y_n + hf(x_n, y_n)$$
$$y_{n+1}^{(k+1)} = y_n + hf(x_{n+1}, y_{n+1}^{(k)}),\ k = 0, 1, 2\cdots$$

直到

$$|y_{n+1}^{(k+1)} - y_{n+1}^{(k)}| < \varepsilon$$

Euler方法（中心差商公式）

- 形式为：

$$y_{n+1} = y_{n-1} + 2hf(x_n, y_n)$$

- 是多步，2阶格式，该格式不稳定

基于数值积分的公式

对微分方程

$$\frac{dy}{dx} = f(x, y)$$

做积分，则：

$$y(x_{n+1}) - y(x_n) = \int_{x_n}^{x_{n+1}} f(x, y)dx$$

$$y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} f(x, y)dx$$

$$= y(x_n) + \int_{x_n}^{x_{n+1}} y'(x)dx$$

矩形公式

用矩形积分公式近似计算$\int_{x_n}^{x_{n+1}} y'(x)dx$.

- 取$y'(x) \approx y'(x_n) = f(x_n, y(x_n))$

$$\int_{x_n}^{x_{n+1}} y'(x)dx \approx (x_{n+1} - x_n)y'(x_n) = hf(x_n, y(x_n))$$

即为向前Euler公式。

- 取$y'(x) \approx y'(x_{n+1}) = f(x_{n+1}, y(x_{n+1}))$

$$\int_{x_n}^{x_{n+1}} y'(x)dx \approx (x_{n+1} - x_n)y'(x_{n+1}) = hf(x_{n+1}, y(x_{n+1}))$$

即为向后Euler公式。

梯形公式

用梯形积分公式近似计算$\int_{x_n}^{x_{n+1}} y'(x)dx$.

$$\begin{aligned}\int_{x_n}^{x_{n+1}} y'(x)dx &\approx \frac{1}{2}(x_{n+1}-x_n)(y'(x_{n+1})+y'(x_n)) \\ &= \frac{h}{2}(f(x_n,y(x_n))+f(x_{n+1},y(x_{n+1})))\end{aligned}$$

得到梯形公式

$$y(x_{n+1}) = y(x_n) + \frac{h}{2}(f(x_n,y(x_n))+f(x_{n+1},y(x_{n+1})))$$

- 局部截断误差: $-\frac{h^3}{2}f''(\xi)$
- 误差估计: $e_{n+1} = \mathcal{O}(h^2)$
- 隐式方法，要用迭代法求解

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

高精度格式-Taylor级数方法

方法的要点是$y(x)$的Taylor级数展开：

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2!}y''(x) + \frac{h^3}{3!}y'''(x) + \frac{h^4}{4!}y^{(4)}(x) + \cdots$$

因此对于固定的x和h, 为了计算出$y(x+h)$的值，我们只需要知道在x点$y(x)$的各阶导数值

$$\begin{aligned} y'(x) &= f(x,y) \\ y''(x) &= f_x(x,y) + f_y(x,y)y'(x) \\ y'''(x) &= \cdots \end{aligned}$$

所以，可以构造格式

$$\begin{aligned} y(x_{n+1}) = y(x_n) &+ hf(x_n, y(x_n)) \\ &+ \frac{h^2}{2}\left(f_x(x_n, y(x_n)) + f_y(x_n, y(x_n))f(x_n, y(x_n))\right) \end{aligned}$$

为了应用Taylor级数方法，我们需要假定f的各阶偏导数存在，例如

$$\begin{cases} y' = \cos x - \sin y + x^2 \\ y(-1) = 3 \end{cases}$$

- 这些导数值可以从给定的微分方程和初值条件中得到：

$$\begin{aligned} y' &= \cos x - \sin y + x^2 \qquad (\text{已知条件}) \\ y'' &= -\sin x - y' \cos y + 2x \\ y''' &= -\cos x - y'' \cos y + (y')^2 \sin y + 2 \\ y^{(4)} &= \sin x - y''' \cos y + 3y'y'' \sin y + (y')^3 \cos y \end{aligned}$$

  我们当然还可以继续下去。如果我们决定仅应用Taylor展开中到h^4之前的项，那么其它项共同构成方法的截断误差，所对应的方法称为四阶方法
- 注意求导中要应用$d \sin y(x)/dx$的链式法则
- 当然可以执行各种代换，使得右边不出现y的导数y', y'', y''', ...。但如果是按上面给出的次序应用这些公式的话，就不必进行这种代换
- 运行Mathematica程序ode_taylor.nb

局部截断误差的累加

- 在上面的算法的每一步中，因为不包含Taylor级数中涉及$h^5, h^6, \ldots$的项，所以局部截断误差是$\mathcal{O}(h^5)$
- 因此当$h \to 0$时，局部截断误差类似于Ch^5。但我们并不知道C是多大
- 不过此例中$h = 0.01$, 因此$h^5 = 10^{-10}$, 每一步中的误差粗略地具有10^{-10}的量级，因此几百步后这此小的误差累加起来，可能不太会损坏精度
- 另外，在每一步中，$y(x_k)$的估计值y_k中已包含误差，进一步地计算继续增加这些误差，因此在得到的数值解中，不要盲目地采用所有的数字

解的验证

- 因此我们需要给出一种方法，来确定最终解的有效数字到底是多少?
- 在此例中我们有$y_{200} = 6.42194$. 以这个值作为同样方程的初值，并且取$h = -0.01$, 重复前面的求解过程，得到$x = -1.0$时解为3.00000, 它与原来的初值几乎相同，因此我们可以认为原来的解具有六位精度

误差的数值估计

- 在n阶方法中，Taylor级数展开到h^n项，那么有如下的误差估计

$$E_n = \frac{1}{(n+1)!} h^{n+1} y^{(n+1)}(x+\theta h), \qquad 0 < \theta < 1$$

- 因此可以用简单的有限差分逼近估计这个误差。例如，对上前例，$n = 4$, $h = 0.01$, 那么

$$E_4 \approx \frac{1}{5!} h^5 \frac{y^{(4)}(x+h) - y^{(4)}(x)}{h} = \frac{h^4}{120}[y^{(4)}(x+h) - y^{(4)}(x)]$$

Taylor级数方法

优点

- 方法概念简单，并且具有高精度的潜力。如果能很容易地得到$y(x)$的20阶导数，则没有什么能阻止我们使用20阶的方法。应用这样高的阶，同样的精度情形下可以采用较大的步长，如$h = 0.2$. 穿过给定区间需要的步数变少，从而有可能减小计算量
- 可以应用符号计算系统执行非数值类型的计算，从而把相当复杂的表达式的微分和积分转换到这些系统中进行。这些系统还可以把计算表达式转化为所需要的代码

Taylor级数方法

缺点

- 依赖于给定的微分方程的反复求导，因此在解曲线经过的$x-y$平面的区域内函数$f(x,y)$必须具有所需要的偏导数。而这样的条件对于解的存在性是不必要的
- 需要对问题进行初步的分析工作。从而在这个步骤中造成的误差可能被忽略而且始终不被发现
- 对于各阶求导必须单独编程，增加了编程的复杂性以及编程错误出现的可能性，代码的可读性下降

延迟型微分方程

- 在一些实际问题中有一类特殊类型的微分方程，称为延迟型微分方程(delay differential equation)或具有延迟变量的微分方程(differential equation with retarded argument)
- 人口模型以及混合问题通常具有这种特征，即$y'(x)$的值与y在x的前面值上的函数值有关
- 例如：

$$y'(x) = f(y(x-1))$$

若知道y在$x-1$上值，微分方程就能够计算$y'(x)$的值。为了从$x=0$开始积分微分方程，我们需要在$x=-1$开始的$y(x)$的变化情况。因此必须提供$y(x)$在区间$[-1,0]$上的值作为初值：

$$\begin{cases} y'(x) = y(x-1) & x \geqslant 0 \\ y(x) = x^2 & -1 \leqslant x \leqslant 0 \end{cases}$$

- 上例中第二个等式给出所需要的$y(x)$的值。若x限定在区间$[0,1]$中，则$x-1$在$[-1,0]$中，因此

$$\begin{cases} y'(x) = y(x-1) = (x-1)^2 & 0 \leqslant x \leqslant 1 \\ y(0) = 0 \end{cases}$$

这是一个通常的ODE，通过积分可以得到解为

$$y(x) = \frac{1}{3}(x-1)^3 + \frac{1}{3}, \qquad 0 \leqslant x \leqslant 1$$

- 如果解被延拓到下一个区间$[1,2]$上，则可以类似处理。此时，对于$x \in [1,2]$, 我们有

$$\begin{cases} y'(x) = y(x-1) = \frac{1}{3}(x-2)^3 + \frac{1}{3} & 1 \leqslant x \leqslant 2 \\ y(1) = \frac{1}{3} \end{cases}$$

也可以得到显式解。类似计算可以一直持续下去

复杂情形

- 对于复杂的方程，如

$$y'(x) = \sin[y(x-1)^3] + \log[y(x) + x^5]$$

我们需要借助于数值方法在每一个区间上求解：Taylor级数方法

- 例如，考虑

$$\begin{cases} y'(x) = 2y(x-1) + y(x) & x > 0 \\ y(x) = x^3 & -1 \leqslant x \leqslant 0 \end{cases}$$

- 为了在区间$[0,1]$中求解，采用如下截断的Taylor展开：

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2}y''(x) + \frac{h^3}{6}y'''(x)$$

以步长h向前进行求解

- 我们需要提供下述导数表达式：

$$\begin{aligned}
y'(x) &= 2y(x-1) + y(x) = 2(x-1)^2 + y(x) \\
y''(x) &= 2y'(x-1) + y'(x) = 4(x-2)^2 + 2(x-1)^2 + y'(x) \\
y'''(x) &= 2y''(x-1) + y''(x) = 8(x-3)^2 + 8(x-2)^2 \\
&\quad + 2(x-1)^2 + y''(x)
\end{aligned}$$

- 基于上述信息，可以得到$[0,1]$中的离散点上$y(x)$的值。同时为了在下一区间内使用，需要存放在这些离散点上的$y'(x)$, $y''(x)$和$y'''(x)$的值。若不改变h的值，那么可以在每个区间上应用适当的存储值类似处理

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

Taylor级数方法的回顾

- 对于ODE数值解的Taylor级数方法，为了进行程序设计，需要事先进行一些分析，确定求导的最高阶数
- 例如，若想对一般的ODE初值问题

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

应用四阶Taylor级数方法，那么就需要进行逐次微分，得到y'', y'''和$y^{(4)}$

- Runge-Kutta方法也是采用了Taylor级数展开，但是它通过对$f(x, y)$的恰当组合，避免了上述困难

二阶Runge-Kutta方法

- $y(x+h)$的Taylor级数展开：

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2!}y''(x) + \frac{h^3}{3!}y'''(x) + \cdots$$

- 根据微分方程，我们得到

$$\begin{aligned} y'(x) &= f \\ y''(x) &= f_x + f_y y' = f_x + f_y f \\ y'''(x) &= f_{xx} + f_{xy} f + (f_x + f_y f) f_y + f(f_{yx} + f_{yy} f) \\ &\vdots \end{aligned}$$

这里下标表示偏导数

- $y(x+h)$的Taylor展开到前三项，可以写为

$$\begin{aligned} y(x+h) &= y + hf + \frac{1}{2}h^2(f_x + ff_y) + \mathcal{O}(h^3) \\ &= y + \frac{1}{2}hf + \frac{1}{2}h[f + hf_x + hff_y] + \mathcal{O}(h^3) \end{aligned}$$

- 应用两变量函数的Taylor展开，可以消去前几项中的偏导数。事实上，

$$f(x+h, y+hf) = f + hf_x + hff_y + \mathcal{O}(h^2)$$

从而有

$$y(x+h) = y + \frac{1}{2}hf + \frac{1}{2}hf(x+h, y+hf) + \mathcal{O}(h^3)$$

Heun公式

- 从而步进求解的公式可以写为

$$y(x+h) = y(x) + \frac{h}{2}f(x,y) + \frac{h}{2}f(x+h, y+hf(x,y))$$

或等价地写为

$$y(x+h) = y(x) + \frac{1}{2}(F_1 + F_2)$$

其中

$$F_1 = hf(x,y), \qquad F_2 = hf(x+h, y+F_1)$$

- 上述公式称为二阶Runge-Kutta方法,也称Heun公式

二阶Runge-Kutta方法的一般化

- 一般来说，二阶Runge-Kutta公式具有形式

$$y(x+h) = y + w_1 hf + w_2 hf(x+\alpha h, y+\beta hf) + \mathcal{O}(h^3)$$

其中w_1, w_2, α和β是需要配置的参数

- 借助于两变量函数的Taylor展开，上式可改写为

$$y(x+h) = y + w_1 hf + w_2 h[f + \alpha h f_x + \beta h f f_y] + \mathcal{O}(h^3)$$

因此结合$y(x+h)$真正的Taylor展开式

$$y(x+h) = y + hf + \frac{1}{2}h^2(f_x + f f_y) + \mathcal{O}(h^3)$$

可得对参数强加的条件：

$$w_1 + w_2 = 1, w_2\alpha = w_2\beta = \frac{1}{2} \Longrightarrow \alpha = \beta = \frac{1}{2w_2} = \frac{1}{2(1-w_1)}$$

各种版本的二阶Runge-Kutta方法

- $w_1 = w_2 = 1/2$, $\alpha = \beta = 1$对应的就是Heun方法
- $w_1 = 0$, $w_2 = 1$, $\alpha = \beta = 1/2$对应的就是修正的Euler方法：

$$y(x+h) = y(x) + F_2$$

其中

$$F_1 = hf(x, y)$$
$$F_2 = hf\left(x + h/2, y + F_1/2\right)$$

- 二阶以及其它阶的Runge-Kutta方法都具有一个显著特点，那就是只需要计算$f(x, y)$在一些点的值，而不需要计算f_x, f_y等导数的表达式以及取值

四阶Runge-Kutta方法

- 推导高阶Runge-Kutta方法的过程是相当令人乏味的，但是推导出来的结果一般是相当简明的
- 经典的四阶Runge-Kutta方法是

$$y(x+h) = y(x) + \frac{1}{6}(F_1 + 2F_2 + 2F_3 + F_4)$$

其中

$$\begin{aligned} F_1 &= hf(x, y) \\ F_2 &= hf\left(x + h/2, y + F_1/2\right) \\ F_3 &= hf\left(x + h/2, y + F_2/2\right) \\ F_4 &= hf\left(x + h, y + F_3\right) \end{aligned}$$

低存储量格式：计算完Fi，直接带上去计算，并进一步计算Fi+1（意思是只用一个存储量就能实现，不用全都存下来

- 这个方法被称为四阶方法，是因为它应用了Taylor展开中直到四阶的所有项。误差为$\mathcal{O}(h^5)$

下面给出一个例子，用以验证方法的有效性

- 在区间$[1,3]$上，采用步长$h=1/128$的四阶Runge-Kutta方法求解下列初值问题：

$$\begin{cases} y' = \frac{xy-y^2}{x^2} \\ y(1) = 2 \end{cases}$$

- 其精确解为

$$y(x) = \frac{x}{\ln x + 1/2}$$

- 结果“ode_runge_kutta_4.pdf”

其它的低阶Runge-Kutta公式

- Runge-Kutta方法的构造设想比较清晰，但是其中系数所满足的非线性方程相当复杂，而求解这个非线性方程组更困难。通常这组方程的解并不唯一，因此可以有几种不同的同阶Runge-Kutta方法
- 高阶Runge-Kutta方法通常用于要求高精度的常微分方程数值求解
- 常用的三阶Runge-Kutta方法有三个：
  1. 第一个：

$$y(x+h) = y(x) + \frac{1}{6}(F_1 + 4F_2 + F_3)$$

其中

$$\begin{aligned} F_1 &= hf(x, y) \\ F_2 &= hf(x + h/2, y + F_1/2) \\ F_3 &= hf(x + h, y - F_1 + 2F_2) \end{aligned}$$

- (续)
  2. 第二个

$$y(x+h) = y(x) + \frac{1}{4}(F_1 + 3F_3)$$

其中

$$\begin{aligned} F_1 &= hf(x, y) \\ F_2 &= hf(x + h/3, y + F_1/3) \\ F_3 &= hf(x + 2h/3, y + 2F_2/3) \end{aligned}$$

  3. 第三个

$$y(x+h) = y(x) + \frac{1}{9}(2F_1 + 3F_2 + 4F_3)$$

其中

$$\begin{aligned} F_1 &= hf(x, y) \\ F_2 &= hf(x + h/2, y + F_1/2) \\ F_3 &= hf(x + 3h/4, y + 3F_2/4) \end{aligned}$$

其它的四阶Runge-Kutta公式

1. $$y(x+h) = y(x) + \frac{1}{8}(F_1 + 3F_2 + 3F_3 + F_4)$$

其中

$$\begin{aligned} F_1 &= hf(x, y) \\ F_2 &= hf(x + h/3, y + F_1/3) \\ F_3 &= hf(x + 2h/3, y + F_1/3 + F_2) \\ F_4 &= hf(x + h, y + F_1 - F_2 + F_3) \end{aligned}$$

2. 第二个: Runge-Kutta-Gill方法

$$y(x+h) = y(x) + \frac{1}{6}\Big(F_1 + (2-\sqrt{2})F_2 + (2+\sqrt{2})F_3 + F_4\Big)$$

其中

$$\begin{aligned}
F_1 &= hf(x,y)\\
F_2 &= hf(x+h/2, y+F_1/2)\\
F_3 &= hf\Big(x+h/2, y+\frac{\sqrt{2}-1}{2}F_1 + (1-\sqrt{2}/2)F_2\Big)\\
F_4 &= hf\Big(x+h, y-\frac{\sqrt{2}}{2}F_2 - (1+\sqrt{2}/2)F_3\Big)
\end{aligned}$$

误差

- 在Runge-Kutta方法的第一步，$y(x_0+h)$的值是由算法计算出来的。另一方面，存在一个我们不知道的正确解$y^*(x_0+h)$. 那么在该步中的局部截断误差是

$$y^*(x_0+h)-y(x_0+h)$$

- Runge-Kutta方法的理论指出这个截断误差对小的h值具有类似于Ch^{n+1}的性态，其中n是方法的阶，C是一个与h无关但与x_0和函数y^*有关的数

局部截断误差的估计

- 以四阶Runge-Kutta方法为例。为估计Ch^5, 我们假定当x从x_0变到x_0+h时C不变
- 设v是在x_0+h上的近似解的值，它是从x_0出发取长度为h的步长一步得到的；设u 是在x_0+h上的另一种数值近似解，它是从x_0出发取长度为$h/2$步长进行两步计算得到的。u, v都是可计算的(computable).
- 从而有

$$y^*(x_0+h) = v + Ch^5$$
$$y^*(x_0+h) = u + 2C(h/2)^5$$

两式相减，得到

$$\text{局部截断误差} = Ch^5 = \frac{u-v}{1-2^{-4}} \approx u - v$$

自适应策略

- 在Runge-Kutta方法的计算机实现中，为了保证近似的截断误差处于特定的容限之下，我们可以通过计算$|u-v|$的值来帮助判断
- 若$|u-v|$超出特定的容限，可以减小步长(通常减半)来改善局部截断误差
- 另一方面，若局部截断误差远远小于特定的容限，那么可以使步长加倍

高阶Runge-Kutta方法的不足

- 同阶的Runge-Kutta方法会有不同的版本，因此Runge-Kutta方法是一族方法
- 在各阶Runge-Kutta方法中，需要的函数求值数目相对于阶数的增加要变得更迅速，如下表所示

函数求值数	1	2	3	4	5	6	7	8
Runge-Kutta方法的阶	1	2	3	4	4	5	6	6

- 因此高阶Runge-Kutta方法并不比经典的四阶Runge-Kutta方法具有更大的吸引力

Runge-Kutta-Fehlberg方法

- 为了尝试在Runge-Kutta方法中设计一个自动调整步长的方法，Fehlberg仔细研究了一个具有五次函数求值的四阶方法和一个具有六次函数求值的五阶方法，通过巧妙选取方法的参数，得到了一个只需要六次函数求值的自适应Runge-Kutta方法
- 方法的基本思路就是让四阶方法的五次函数求值包含在五阶方法的六次函数求值中，从而把本来的11次函数求值减少为只有六次。同时应用两种方法得到结果的差做为对局部截断误差的估计

RKF方法中的六次函数求值

- 六次函数求值：

$$
\begin{aligned}
F_1 &= hf(x, y) \\
F_2 &= hf(x + h/4, y + F_1/4) \\
F_3 &= hf\left(x + \frac{3}{8}h, y + \frac{3}{32}F_1 + \frac{9}{32}F_2\right) \\
F_4 &= hf\left(x + \frac{12}{13}h, y + \frac{1932}{2197}F_1 - \frac{7200}{2197}F_2 + \frac{7296}{2197}F_3\right) \\
F_5 &= hf\left(x + h, y + \frac{439}{216}F_1 - 8F_2 + \frac{3680}{513}F_3 - \frac{845}{4104}F_4\right) \\
F_6 &= hf\left(x + \frac{1}{2}h, y - \frac{8}{27}F_1 + 2F_2 - \frac{3544}{2565}F_3 \right. \\
&\qquad \left. + \frac{1859}{4104}F_4 - \frac{11}{40}F_5\right)
\end{aligned}
$$

RKF方法中的四阶和五阶方法

- 五阶方法

$$y(x+h) = y(x) + \sum_{i=1}^{6} a_i F_i$$
$$= y(x) + \frac{16}{135}F_1 + \frac{6656}{12825}F_3 + \frac{28561}{56430}F_4 - \frac{9}{50}F_5 + \frac{2}{55}F_6$$

注意其中F_2项的系数为零

- 四阶方法

$$\bar{y}(x+h) = y(x) + \sum_{i=1}^{6} b_i F_i$$
$$= y(x) + \frac{25}{216}F_1 + \frac{1408}{2565}F_3 + \frac{2197}{4104}F_4 - \frac{1}{5}F_5$$

注意其中F_2和F_6项的系数为零

RKF45 vs RKF54

- 如果以四阶方法计算的结果作为最终结果，那么称为RKF45方法
- 反之，如果以五阶方法计算的结果作为最终结果，那么称为RKF54方法
- 一般认为RKF54方法较好，但也存在精度被减少的情形，因此两种方法孰优孰劣，没有定论

自适应步长选择的策略

假设要构造一个自适应算法，试图使局部截断误差e的值限定在预先给定的容限δ之内

- 第一种策略：假设$y(x) - \bar{y}(x)$估计的局部截断误差为Ch^5。当加倍步长导致$C(2h)^5 < \delta/4$时，对步长进行加倍是合理的。因此当局部截断误差小于$\delta/128$时，对步长进行加倍。另一方面，按自然的方法对步长进行减半，即一旦误差大于δ，就对步长进行减半
- 第二种策略：在每一步计算后，采用如下通用公式确定下一步的步长：

$$h \leftarrow 0.9h\left(\frac{\delta}{|e|}\right)^{\frac{1}{1+p}}$$

其中p是一对Runge-Kutta方法中第一个公式的阶。这个公式既可能增加步长，也可能减小步长

RKF方法的推广

- 由一对阶分别为p和q的Runge-Kutta方法构成，其中两者共享高阶方法的函数求值，这种方法称为嵌入式Runge-Kutta方法(embedded Runge-Kutta procedures). 一般采用$q = p + 1$
- 通常是用来求解非刚性初值问题的有效方法
- 现在已推导出了大量带有复杂系数的高阶Runge-Kutta方法
- 除非采用的是自适应格式，经典的四阶Runge-Kutta方法仍是最常用的

程序作业

- 应用RKF45或RKF54方法，设计实现自适应方法，求解如下常微分方程初值问题：

$$\begin{cases} y' = e^{yx} + \cos(y - x), \\ y(1) = 3 \end{cases}$$

- 初值步长取为$h = 0.01$. 在自适应方法中步长的选取采用第二种策略。
- 在解溢出前终止。
- 程序的输出：
  1. 解的范围: $[1, ?]$
  2. 提示输入一个介于上述范围的值，应用简单的两点线性插值计算出对应的函数值。

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

单步法与多步法

- 前面几节中给出的求解初值问题的Taylor级数方法和Runge-Kutta方法具有一个共同特点，那就是解从x前进到$x+h$时，新计算出来的函数值只与x时刻的函数值有关，因此称为**单步法**。
  - 即若$x_0, x_1, \ldots, x_i$是沿x轴的点，那么y_{i+1}(即$y(x_{i+1})$的近似值)仅与y_i有关，而与$y_{i-1}, y_{i-2}, \ldots, y_0$的信息无关
- 如果在每一步利用解的某些先前的值，则可以设计出更有效的方法，这些方法称为**多步法**。

多步法的基本原理

- 我们的目标是数值求解下列初值问题：

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

- 若真解为$y(x)$, 那么对上述方程的积分可以得到

$$\int_{x_n}^{x_{n+1}} y'(x)dx = y(x_{n+1}) - y(x_n)$$

- 前提条件：在某时刻，已经得到了x轴上点$x_0, x_1, \ldots, x_n$ (可以是不等距的)处的函数值。

基于数值积分的构造法

- 在x_{n+1}处函数值的公式为

$$y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} f(x, y(x))dx$$

右边的积分可以用数值积分格式逼近，结果是一个逐步生成近似解的公式

- 注意：这里的数值积分格式与前一章中的有些不同
  1. 前一章的格式：为了计算$[a, b]$上的积分，积分节点一般取在$[a, b]$内
  2. 这里为了计算$[x_n, x_{n+1}]$上面的积分，积分节点是前面的点$x_n, x_{n-1}, \cdots, x_{n-q}$，在$[x_n, x_{n+1}]$外

基于数值积分的构造法

- 若积分$\int_{x_{n-p}}^{x_{n+1}} y'(x)dx$用节点$x_n, x_{n-1}, \cdots, x_{n-q}$作为积分点，则得到显格式，$q+1$阶$r+1$步格式。$r=\max(p,q)$

$$\int_{x_n}^{x_{n+1}} y'(x)dx \approx h(a_0 y'(x_n) + a_1 y'(x_{n-1}) + \cdots + a_q y'(x_{n-q}))$$

利用$y'(x_n) = f(x_n, y(x_n))$，有

$$y(x_{n+1}) = y(x_{n-p}) + h\sum_{j=0}^{q} a_j f(x_{n-j}, y(x_{n-j}))$$

- 积分系数

$$ha_j = \int_{x_{n-p}}^{x_{n+1}} l_j(x)dx$$

- 局部截断误差

$$\int_{x_{n-p}}^{x_{n+1}} \frac{y^{(q+2)}(\xi)}{(q+1)!}\omega_q(x)dx$$

基于数值积分的构造法

- 若积分$\int_{n-p}^{x_{n+1}} y'(x)dx$用节点$x_{n+1}, x_n, \cdots, x_{n+1-q}$作为积分点，则得到隐格式，$q+1$阶$r+1$步格式。$r=\max(p,q)$

$$\int_{x_n}^{x_{n+1}} y'(x)dx \approx h(a_0 y'(x_{n+1}) + a_1 y'(x_n) + \cdots + a_q y'(x_{n-q+1}))$$

利用$y'(x_n) = f(x_n, y(x_n))$，有

$$y(x_{n+1}) = y(x_{n-p}) + h\sum_{j=0}^{q} a_j f(x_{n+1-j}, y(x_{n+1-j}))$$

- 积分系数

$$ha_j = \int_{x_{n-p}}^{x_{n+1}} l_j(x)dx$$

- 局部截断误差

$$\int_{x_{n-p}}^{x_{n+1}} \frac{y^{(q+2)}(\xi)}{(q+1)!}\omega_q(x)dx$$

线性多步法

线性多步法的一般形式为

$$a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = h(b_k f_n + b_{k-1} f_{n-1} + \cdots + b_0 f_{n-k})$$

- 这称为k步方法
- 系数a_i, b_i已知
- y_i表示解在x_i上的近似，$x_i = x_0 + ih$, f_i表示$f(x_i, y_i)$
- 假定y_0, y_1, ..., y_{n-1}的值已知，采用上式计算y_n. 因此可以假定$a_k \neq 0$
- 若$b_k = 0$, 方法称为**显式**的(explicit)，此时y_n可直接由上式简单计算出来
- 若$b_k \neq 0$, 则右端项f_n中包含未知数y_n, 因此称为**隐式**(implicit) 方法

建立$p=1, q=2$的显格式

- $p=1$,积分区间为$\int_{x_{n-1}}^{x_{n+1}} y'(x)dx$
- q=2，显格式，积分节点为x_n, x_{n-1}, x_{n-2}

$$ha_0 = \int_{x_{n-1}}^{x_{n+1}} \frac{(x-x_{n-1})(x-x_{n-2})}{(x_n-x_{n-1})(x_n-x_{n-2})}dx = \frac{7}{3}h$$

$$ha_1 = \int_{x_{n-1}}^{x_{n+1}} \frac{(x-x_n)(x-x_{n-2})}{(x_{n-1}-x_n)(x_{n-1}-x_{n-2})}dx = -\frac{2}{3}h$$

$$ha_2 = \int_{x_{n-1}}^{x_{n+1}} \frac{(x-x_n)(x-x_{n-1})}{(x_{n-2}-x_{n-1})(x_{n-2}-x_{n-1})}dx = \frac{1}{3}h$$

- 局部截断误差为

$$T_{n+1} = \int_{x_{n-1}}^{x_{n+1}} \frac{y^{(4)}(\xi)}{3!}(x-x_n)(x-x_{n-1})(x-x_{n-2})dx = \frac{1}{3}h^4 y^{(4)}(\eta)$$

建立$p=2, q=2$的隐格式

- $p=2$,积分区间为$\int_{x_{n-2}}^{x_{n+1}} y'(x)dx$
- q=2，隐格式，积分节点为x_{n+1}, x_n, x_{n-1}

$$ha_0 = \int_{x_{n-2}}^{x_{n+1}} \frac{(x-x_n)(x-x_{n-1})}{(x_{n+1}-x_n)(x_{n+1}-x_{n-1})}dx = \frac{3}{4}h$$

$$ha_1 = \int_{x_{n-2}}^{x_{n+1}} \frac{(x-x_{n+1})(x-x_{n-1})}{(x_n-x_{n+1})(x_n-x_{n-1})}dx = 0$$

$$ha_2 = \int_{x_{n-2}}^{x_{n+1}} \frac{(x-x_{n+1})(x-x_n)}{(x_{n-1}-x_{n+1})(x_{n-1}-x_n)}dx = \frac{9}{4}h$$

- 局部截断误差为

$$T_{n+1} = \int_{x_{n-1}}^{x_{n+1}} \frac{y^{(4)}(\xi)}{3!}(x-x_n)(x-x_{n+1})(x-x_{n-1})dx = -\frac{3}{8}h^4 y^{(4)}(\eta)$$

Adams-Bashforth公式

- 公式类型为

$$y_{n+1} = y_n + af_n + bf_{n-1} + cf_{n-2} + \cdots$$

其中$f_i = f(x_i, y_i)$. 这类公式称为Adams-Bashforth公式

- 基于等距节点$x_i = y_0 + ih,\ i = 1, 2, \ldots, n$的五阶Adams-Bashforth公式为

$$y_{n+1} = y_n + \frac{h}{720}\big(1901f_n - 2774f_{n-1} + 2616f_{n-2} - 1274f_{n-3} + 251f_{n-4}\big)$$

五阶Adams-Bashforth公式的推导

- 系数来自于下述数值积分

$$\int_{x_n}^{x_{n+1}} f(x, y(x))dx \approx h(Af_n + Bf_{n-1} + Cf_{n-2} + Df_{n-3} + Ef_{n-4})$$

为了确定系数A, B, C, D, E，要求当被积函数在Π_4中公式精确成立

- 不失一般性，假设$x_n = 0,\ h = 1$
- 通过取Π_4中的五个Newton基作为测试函数，可以求解出所需要的系数。注意这并不是Newton-Cotes公式，因为积分节点在积分区间外

- 五个Newton基函数为

$$
\begin{aligned}
p_0(x) &= 1 \\
p_1(x) &= x \\
p_2(x) &= x(x+1) \\
p_3(x) &= x(x+1)(x+2) \\
p_4(x) &= x(x+1)(x+2)(x+3)
\end{aligned}
$$

代入下式

$$\int_0^1 p_n(x)dx = Ap_n(0)+Bp_n(-1)+Cp_n(-2)+Dp_n(-3)+Ep_n(-4)$$

得到确定系数A, B, C, D, E的五个方程。

- 这个方程组为

$$\begin{cases} A+B+C+D+E=1 \\ -B-2C-3D-4E=1/2 \\ 2C+6D+12E=5/6 \\ -6D-24E=9/4 \\ 24E=251/30 \end{cases}$$

从而得到前面列出的解。

- 上述过程称为**待定系数法**。原则上，它可以用来确定高阶和各种各样其他情况的类似公式。
- 另外，由于$y'=f$, 因此令$y=p_n$并且应用$f=p_n'$可以得到类似的方程组以确定公式中的系数。
- 在数值实践中，很少直接使用Adams-Bashforth公式，而是把它与其它公式联合起来，因为其中的积分是采用外推方法计算的。

Adams-Moulton公式

- 应用待定系数法可以构造如下所示的公式，其中包含f_{n+1}项：

$$y_{n+1} = y_n + af_{n+1} + bf_n + cf_{n-1} + \cdots$$

- 例如，五阶的Adams-Moulton公式如下：

$$y_{n+1} = y_n + \frac{h}{720}(251f_{n+1} + 646f_n - 264f_{n-1} + 106f_{n-2} - 19f_{n-3})$$

- 这类公式不能直接用于步进求解，因为y_{n+1}同时出现在公式的两边。
  - f_i表示$f(x_i, y_i)$，因此包含f_{n+1}的项仅在y_{n+1}知道后才能计算。
  - 有两种方法解决这个问题：预估–校正、迭代泛函

预估–校正方法

- 我们可以用Adams-Bashforth公式预估y_{n+1}的**试验**值y^*_{n+1}，然后使用Adams-Moulton公式计算y_{n+1}的**校正**值，即组合计算公式为

$$y^*_{n+1} = y_n + \frac{h}{720}(1901f_n - 2774f_{n-1} + 2616f_{n-2} - 1274f_{n-3} + 251f_{n-4})$$

$$y_{n+1} = y_n + \frac{h}{720}\left(251f(x_{n+1}, y^*_{n+1}) + 646f_n - 264f_{n-1} + 106f_{n-2} - 19f_{n-3}\right)$$

预估–校正过程的启动

- 在应用上述的预估–校正过程中，第一步只是知道y_0，还没有足够的信息进行上述预估–校正过程，所以需要采用特殊的过程进行方法启动。
  - 在知道y_1，y_2，y_3，y_4的值后才能启动公式。
- 采用Runge-Kutta方法得到上述值是一种不错的选择。
  - 通常同阶的公式一起使用，即这里采用五阶的Runge-Kutta方法与Adams-Bashforth, Adams-Moulton公式一起使用。

不动点理论与迭代函数

- Adams-Moulton公式说明y_{n+1}是某个映射的不动点，这里的映射定义为

$$\phi(z) = \frac{251}{720} hf(x_{n+1}, z) + C$$

- 所以泛函迭代法可以作为计算y_{n+1}的一种方法。
  - 根据泛函迭代理论，在适当的假设下，由$z_{k+1} = \phi(z_k)$确定的序列收敛于ϕ的不动点
  - 若ξ是ϕ的不动点，即所求的y_{n+1}值，那么迭代应当从中心在ξ的区间中一点z_0出发，在这个区间中每点满足$|\phi'(z)| < 1$. 这里需要假定ϕ'是连续的。此时

$$\phi'(z) = \frac{251}{720} h \frac{\partial f(x_{n+1}, z)}{\partial z}$$

    通过变小步长h, 可以使得上式变得足够小。
  - 实际中，通常只需要一到两步迭代就可以找到y_{n+1}的足够精度的近似值。

程序作业

初值问题

$$x' = \frac{t - e^{-t}}{x + e^{x}}$$
$$x(0) = 0$$

该方程的真解由等式

$$x^2 - t^2 + 2e^{x} - 2e^{-t} = 0$$

隐式给出。

- 当$t = 1$时，数值求解等式$x^2 - t^2 + 2e^{x} - 2e^{-t} = 0$，将这一数值解作为参考的准确解。

- 利用Adams-Bashforth公式计算方程在$t = 1$的数值解，利用Runge − Kunta格式得到初值，取节点$x_i, i = 0, \cdots, N$, N为$2^k, k = 3, \cdots, 8$,给出如下的误差表格,其中阶为

$$\frac{\ln(Error_{old}/Error_{now})}{\ln(N_{now}/N_{old})}$$

N	误差	阶
8		
16		
32		
64		

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

线性多步法的分析

- 本节余下部分讨论一般的线性多步法理论。
- 线性多步法的一般形式为

$$a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = h(b_k f_n + b_{k-1} f_{n-1} + \cdots + b_0 f_{n-k})$$

  - 这称为k步方法
  - 系数a_i, b_i已知
  - y_i表示解在x_i上的近似，$x_i = x_0 + ih$, f_i表示$f(x_i, y_i)$
  - 假定$y_0, y_1, \ldots, y_{n-1}$的值已知，采用上式计算$y_n$. 因此可以假定$a_k \neq 0$
  - 若$b_k = 0$, 方法称为**显式**的(explicit)，此时y_n可直接由上式简单计算出来
  - 若$b_k \neq 0$, 则右端项f_n中包含未知数y_n, 因此称为**隐式**(implicit) 方法

阶

- 微分方程数值解的精度在很大程度上是由使用的算法的**阶**确定的
- 阶表明方法所模拟的Taylor级数解中有多少项被考虑。
  - 例如，在Adams-Bashforth公式中，它之所以被称为五阶的，是因为它近似地产生与带有$h, \ldots, h^5$的Taylor级数方法相同的精度
  - 从而在利用Adams-Bashforth公式求解的每一步中误差都可以期望是$\mathcal{O}(h^6)$
- 下面的论述将把阶的概念变得更精确

线性泛函

- 定义线性泛函如下：

$$Ly = \sum_{i=0}^{k} \Big(a_i y(ih) - h b_i y'(ih) \Big)$$

这里为了简化记号，设$k = n$, 并且第一个值是在$x = 0$，而不是$x = n - k$

- 上述泛函可作用到任何可微的函数y上。在下面的分析中，假定y是由$x = 0$的Taylor级数表示。利用y的Taylor级数，可以把L表示为下列形式：

$$Ly = d_0 y(0) + d_1 h y'(0) + d_2 h^2 y''(0) + \cdots$$

- 为了计算系数d_i, 写出y和y'的Taylor级数：

$$y(ih) = \sum_{j=0}^{\infty} \frac{(ih)^j}{j!} y^{(j)}(0) \qquad y'(ih) = \sum_{j=0}^{\infty} \frac{(ih)^j}{j!} y^{(j+1)}(0)$$

- 然后代入到泛函L的定义式中，按h的幂次重新整理，得到d_i的值如下：

$$
\begin{aligned}
d_0 &= \sum_{i=0}^{k} a_i \\
d_1 &= \sum_{i=0}^{k} (i a_i - b_i) \\
d_2 &= \sum_{i=0}^{k} \left(\frac{1}{2} i^2 a_i - i b_i \right) \\
&\vdots \\
d_j &= \sum_{i=0}^{k} \left(\frac{i^j}{j!} a_i - \frac{i^{j-1}}{(j-1)!} b_i \right), \quad j \geqslant 1
\end{aligned}
$$

Theorem

线性多步法的下列三个性质等价：

1. $d_0 = d_1 = \cdots = d_m = 0$
2. 对每个次数不超过m的多项式p有$Lp = 0$
3. 对一切$y \in C^{m+1}$, Ly是$\mathcal{O}(h^{m+1})$

证明：若性质1成立，则泛函为

$$Ly = d_{m+1} h^{m+1} y^{(m+1)}(0) + \cdots$$

而对于次数不超过m的多项式p, $p^{(j)}(0) = 0$, $j > m$。所以$Lp = 0$, 即性质2成立。

假设性质2成立。若$y \in C^{m+1}$, 则由Taylor定理，记$y = p + r$, 其中p是一个次数不超过m的多项式，函数r在0点的前m阶导数为零，从而

$$Ly = Lr = d_{m+1} h^{m+1} r^{(m+1)}(\xi) = \mathcal{O}(h^{m+1})$$

所以性质3成立。
最后，若性质3成立，则必定有$d_0 = d_1 = \cdots = d_m = 0$，即性质1成立。 □

- 因此阶的严格定义是唯一使得

$$d_0 = d_1 = \cdots = d_m = 0 \neq d_{m+1}$$

成立的自然数m.

分析由下式确定的Milne方法的阶是多少？

$$y_n - y_{n-2} = \frac{1}{3}h(f_n + 4f_{n-1} + f_{n-2})$$

- $a_0 = -1,\ a_1 = 0,\ a_2 = 1,\ b_0 = 1/3,\ b_1 = 4/3,\ b_2 = 1/3$
- 因此有

$$\begin{aligned} d_0 &= a_0 + a_1 + a_2 = 0 \\ d_1 &= -b_0 + (a_1 - b_1) + (2a_2 - b_2) = 0 \\ d_2 &= (a_1/2 - b_1) + (2a_2 - 2b_2) = 0 \end{aligned}$$

- $$d_3 = \left(\frac{1}{6}a_1 - \frac{1}{2}b_1\right) + \left(\frac{4}{3}a_2 - 2b_2\right) = 0$$
$$d_4 = \left(\frac{1}{24}a_1 - \frac{1}{6}b_1\right) + \left(\frac{2}{3}a_2 - \frac{4}{3}b_2\right) = 0$$
$$d_5 = \left(\frac{1}{120}a_1 - \frac{1}{24}b_1\right) + \left(\frac{4}{15}a_2 - \frac{2}{3}b_2\right) = -\frac{1}{90}$$

所以方法是四阶的。

说明

- 如果其它特性不相上下的话，我们可能更喜欢高阶方法。
- 为了产生一个$2k$阶的k步方法，那么考虑下述$2k+1$个方程：

$$d_0 = d_1 = \cdots = d_{2k} = 0$$

这是一个有$2k+2$个未知数的$2k+1$个齐次线性方程构造的方程组，因此必定有非平凡解

- 1956年Dahlquist证明了对于$a_k \neq 0$存在一个解
- 但是多步法的首要特征是稳定性。Dahlquist证明了一个稳定k步法不能有大于$k+2$的阶

- 定义V表示所有无穷复数序列组成的集合，即V中元素具有形式

$$y = (y_1, y_2, y_3, \ldots), \quad y_i \in \mathbb{C}$$

实际上，y可以看作是正整数$\mathbb{N} = \{1, 2, 3, \ldots\}$上的复值函数。用$y_n$代替函数$y$在自变量$n$处的值$y(n)$只是为了方便。

- 在V中定义通常的加法和数乘，V就成为线性空间，这个空间是无限维的。
- 考虑线性算子$L : V \to V$, 其中最重要的一类就是移位算子，记为E, 定义为

$$Ey = (y_2, y_3, y_4, \ldots)$$

即$(Ey)_n = y_{n+1}$

线性差分算子

- 移位算子可以连续复合在一起，例如$(EEy)_n = y_{n+2}$，$(E^k y)_n = y_{n+k}$.
- 由E的幂次有限线性组合表示的线性算子，称为线性差分算子，它的形式为

$$L = \sum_{i=0}^{m} c_i E^i$$

  其中E^0为恒等算子。
- V的所有线性差分算子构成从V到V的所有线性算子形成的空间的子空间，E的幂次就是这个空间的一组基。
- 此处我们主要研究线性差分方程$Ly = 0$的所有解。显然集合$\{y : Ly = 0\}$是V的一个线性子空间，称为L的零空间。当找到这个子空间的一组基后，我们就认为求解出了方程$Ly = 0$.

特征多项式

- L是E的一个多项式，记$L = p(E)$, 其中p是一个多项式，称为L的特征多项式，定义为

$$p(\lambda) = \sum_{i=0}^{m} c_i \lambda^i$$

- 例：当$c_0 = 2$, $c_1 = -3$, $c_2 = 1$, 其它$c_i = 0$, 对应的线性差分方程为

$$(E^2 - 3E^1 + 2E^0)y = 0 \quad \text{或者}$$
$$y_{n+2} - 3y_{n+1} + 2y_n = 0, \quad n \geqslant 1, \quad \text{或者}$$
$$p(E)y = 0, \quad p(\lambda) = \lambda^2 - 3\lambda + 2$$

- 很容易得到上述方程的解。实际上，可以任意选择y_1, y_2，那用应用$y_{n+2} - 3y_{n+1} + 2y_n = 0$就可以迭代确定后面的分量。例如

$$(1, 0, -2, -6, -14, -30, \ldots)$$
$$(1, 1, 1, 1, \ldots)$$
$$(2, 4, 8, 16, \ldots)$$

其中第一个很难看出通项，而后两个解的形式为$y_n = \lambda^n$，$\lambda = 1, 2$. 而这两个数就是特征多项式的根

- 是否存在其它形式为λ^n的解呢？把$y_n = \lambda^n$代入$y_{n+2} - 3y_{n+1} + 2y_n = 0$可得

$$\lambda^n(\lambda - 1)(\lambda - 2) = 0$$

因此此类形式的其它解只可能是$(0, 0, 0, \ldots)$

- 实际上，由$u_n = 1$和$v_n = 2^n$定义的解形成了零空间的一组基。实际上，设y是任意解，下面求解常数α, β使得$y = \alpha u + \beta v$。这个等式即

$$y_n = \alpha u_n + \beta v_n$$

特别地，对于$n = 1, 2$, 有

$$y_1 = \alpha + 2\beta, \quad y_2 = \alpha + 4\beta$$

方程组有唯一解α, β. 对于其它的n:

$$\begin{aligned} y_n &= 3y_{n-1} - 2y_{n-2} \\ &= 3(\alpha u_{n-1} + \beta v_{n-1}) - 2(\alpha u_{n-2} + \beta v_{n-2}) \\ &= \alpha(3u_{n-1} - 2u_{n-2}) + \beta(3v_{n-1} - 2v_{n-2}) \\ &= \alpha u_n + \beta v_n \end{aligned}$$

Theorem (零空间定理)

若λ是多项式p的一个根，则$(\lambda, \lambda^2, \lambda^3, \ldots)$是差分方程$p(E)y = 0$的一个解。若$p$的所有根为单根，则差分方程的每个解是这些特解的一个线性组合。

证明：若λ为任意复数，$u = (\lambda, \lambda^2, \lambda^3, \ldots)$，则有$(Eu)_n = \lambda u_n$，即$Eu = \lambda u$. 从而有$E^i u = \lambda^i u$. 由此可得

$$p(E)u = \left(\sum_{i=0}^{m} c_i E^i\right) u = \sum_{i=0}^{m} c_i (E^i u) = \sum_{i=0}^{m} c_i \lambda^i u = p(\lambda)u$$

所以若$p(\lambda) = 0$, 则有$p(E)u = 0$.

设多项式p的所有根λ_k都是单根，则对每个根λ_k, 差分方程$p(E)y = 0$有一个解$u^{(k)} = (\lambda_k, \lambda_k^2, \lambda_k^3, \ldots)$. 设$y$是差分方程的任意解，下面把它表示成$y = \sum_{k=1}^{m} a_k u^{(k)}$. 实际上，取这级数的前$m$项，得到

$$y_i = \sum_{k=1}^{m} a_k \lambda_k^i, \qquad i = 1, 2, \ldots, m$$

方程组的系数阵为Vandermonde矩阵，因此可得到唯一的$a_1, \ldots, a_m$使得上式成立。令

$$z = y - \sum_{k=1}^{m} a_k u^{(k)}$$

那么有$p(E)z = 0$，由此可得$z = 0$. □

- 若λ是p的一个k重根，那么下述序列是差分方程$p(E)y = 0$的解：

$$
\begin{aligned}
y(\lambda) &= (\lambda, \lambda^2, \lambda^3, \ldots) \\
y'(\lambda) &= (1, 2\lambda, 3\lambda^2, \ldots) \\
y''(\lambda) &= (0, 2, 6\lambda, \ldots) \\
&\vdots \\
y^{(k-1)}(\lambda) &= \frac{d^{k-1}}{d\lambda^{k-1}}(\lambda, \lambda^2, \lambda^3, \ldots)
\end{aligned}
$$

Theorem (零空间的基定理)

设p是一个多项式，并且$p(0) \neq 0$, 则可以得到$p(E)$的零空间一组基为：对于p的每个k重根λ, 有相应的k个解$y(\lambda)$, $y'(\lambda)$, ..., $y^{(k-1)}(\lambda)$，其中$y(\lambda) = (\lambda, \lambda^2, \lambda^3, \ldots)$

求差分方程的通解

$$4y_n + 7y_{n-1} + 2y_{n-2} - y_{n-3} = 0$$

- $p(\lambda) = 4\lambda^3 + 7\lambda^2 + 2\lambda - 1 = (\lambda+1)^2(4\lambda-1)$. p有一个二重根-1和单根$1/4$. 所以基解为

$$\begin{aligned} y(-1) &= (-1, 1, -1, 1, \ldots) \\ y'(-1) &= (1, -2, 3, -4, \ldots) \\ y(1/4) &= \left(\frac{1}{4}, \frac{1}{16}, \frac{1}{64}, \cdots\right) \end{aligned}$$

从而通解为

$$\begin{aligned} y &= \alpha y(-1) + \beta y'(-1) + \gamma y(1/4) \\ &= \alpha(-1)^n + \beta n(-1)^{n-1} + \gamma(1/4)^n \end{aligned}$$

稳定的差分方程

- 如果对于V中元素$y = (y_1, y_2, \ldots)$存在常数c使得对所有的n, 有$|y_n| \leqslant c$, 即

$$\sup_n |y_n| < \infty$$

  则称y有界。
- 若形如$p(E)y = 0$的差分方程的解有界，则称此差分方程是稳定的。

Theorem

对于一个满足$p(0) \neq 0$的多项式p, 下述条件是等价的：

1. 差分方程$p(E)y = 0$是稳定的
2. p的所有根满足$|z| \leqslant 1$, 并且所有重根满足$|z| < 1$

线性多步法的理论分析

- 线性多步法的一般形式为

$$a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = h(b_k f_n + b_{k-1} f_{n-1} + \cdots + b_0 f_{n-k})$$

- 设想初值问题的数值解是由不同步长计算得到的，
用$y(h,x)$表示在步长h时得到的数值解。精确解记为$y(x)$。
线性多步法称为收敛的，是指对于区间$[x_0, x_m]$中的任意x,

$$\lim_{h \to 0} y(h,x) = y(x)$$

这里的前提条件就是初始值满足同样的定义，即

$$\lim_{h \to 0} y(h, x_0 + nh) = y_0, \quad 0 \leqslant n < k$$

以及函数f满足基本的存在性定理的假设。

收敛性和相容性

- 线性多步法相应的两个多项式是

$$p(z) = a_k z^k + a_{k-1} z^{k-1} + \cdots + a_0$$
$$q(z) = b_k z^k + b_{k-1} z^{k-1} + \cdots + b_0$$

- 若p的所有根位于圆盘$|z| \leqslant 1$中，而且模为1的根是单根，则方法是稳定的
- 若$p(1) = 0,\ p'(1) = q(1)$, 则方法是相容的

Theorem (线性多步法的稳定性和相容性定理)

线性多步法收敛的充要条件就是这个多步法是稳定的和相容的。

这个定理的必要性证明比较简单。充分性相当复杂，不给出。

收敛$\Longrightarrow$稳定

若方法不稳定，则或者p有一个根λ满足$|\lambda| > 1$或者p有一个根λ满足$|\lambda| = 1$, 并且$p'(\lambda) = 0$.
考虑初值问题(其解为$y(x) \equiv 0$):

$$\begin{cases} y' = 0, \\ y(0) = 0 \end{cases}$$

那么线性多步法是由等式

$$a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = 0$$

确定。这是一个线性差分方程，它的一个解是$y_n = h\lambda^n$, 其中λ就是p的一个根。若$|\lambda| > 1$, 则对所有$0 \leqslant n < k$, 我们有

$$|y(h, nh)| = h|\lambda^n| < h|\lambda|^k \to 0, \quad h \to 0$$

满足了前面的限制条件$\lim\limits_{h\to 0} y(h, x_0 + nh) = y_0$。

但是它违背收敛条件$\lim_{h\to 0} y(h,x) = y(x)$，因为若$x = nh$, 则$h = x/n$, 并且

$$|y(h,x)| = |y(h,nh)| = x|\lambda|^n/n \to \infty$$

另外一方面，若$|\lambda| = 1$且$p'(\lambda) = 0$, 则上述差分方程的一个解是$y_n = hn\lambda^n$. 这时同样满足限制条件：

$$|y(h,nh)| = hn|\lambda|^n = hn < hk \to 0, \quad h \to 0, 0 \leqslant n < k$$

但违背收敛条件，因为

$$|y(h,x)| = (x/n)n|\lambda|^n = x \neq 0$$

- 考虑问题

$$\begin{cases} y' = 0, \\ y(0) = 1 \end{cases}$$

其精确解为$y \equiv 1$. 线性多步法仍然是

$$a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = 0$$

取$y_0 = y_1 = \cdots = y_{k-1} = 1$得到一个解，然后利用线性多步法得到后面的$y_k$值。因为方法收敛，所以$\lim\limits_{n\to\infty} y_n = 1$. 把它代入到多步法的定义中，得到$a_k + a_{k-1} + \cdots + a_n = 0$,
即$p(1) = 0$.

- 再考虑初值问题

$$\begin{cases} y' = 1, \\ y(0) = 0 \end{cases}$$

其精确解为$y = x$. 多步法表示为

$$a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = h[b_k + b_{k-1} + \cdots + b_0]$$

由于方法收敛，从而是稳定的，因而$p(1) = 0$, $p'(1) \neq 0$. 下面验证$y_n = (n+k)h\gamma$, $\gamma = q(1)/p'(1)$给出上面非齐次差分方程的解，实际上，

$$\begin{aligned} & h\gamma[a_k(n+k) + a_{k-1}(n+k-1) + \cdots + a_0 n] \\ = & nh\gamma(a_k + a_{k-1} + \cdots + a_0) + h\gamma[ka_k + (k-1)a_{k-1} + \cdots + a_1] \\ = & nh\gamma p(1) + h\gamma p'(1) \\ = & h\gamma p'(1) = hq(1) = h(b_k + b_{k-1} + \cdots + b_0) \end{aligned}$$

- 因为

$$\lim_{h\to 0}(n+k)h\gamma = 0, \quad n = 0, 1, \ldots, k-1$$

所以这个数值解中的开始值与初值$y(0) = 0$相容。此时的收敛条件要求当$nh = x$时

$$\lim_{n\to\infty} y_n = x$$

因此我们有

$$\lim_{n\to\infty}(n+k)h\gamma = x$$

而$\lim_{n\to\infty} kh = 0$, 所以得到$\gamma = 1$即$p'(1) = q(1)$

5.2.1 例：Milne方法

$$y_n - y_{n-2} = \frac{h}{3}(f_n + 4f_{n-1} + f_{n-2})$$

- 这是一个四阶的隐式方法，它由下述两个多项式来描述：

$$p(z) = z^2 - 1$$
$$q(z) = \frac{1}{3}z^2 + \frac{4}{3}z + \frac{1}{3}$$

- p的根为$+1$和-1, 都是单根，而且$p'(z) = 2z$, 所以$p'(1) = 2 = q(1)$，从而相容性和稳定性条件满足，即Milne方法是收敛的。

局部截断误差

- 目标是分析多步法中产生的局部截断误差，即假定在所有前面的值$y_{n-1}, y_{n-2}, \ldots$是准确的假设下，利用给定的多步法得到y_n所产生的误差: $y(x_n) - y_n$. 这个误差是由差分方程近似微分方程而导致的。其中不包含舍入误差

Theorem (线性多步法的局部截断误差定理)

若多步法是m阶的，$y \in C^{m+2}$, 而且$\partial f/\partial y$连续，则

$$y(x_n) - y_n = \frac{d_{m+1}}{a_k} h^{m+1} y^{(m+1)}(x_{n-k}) + \mathcal{O}(h^{m+2})$$

其中系数d_k定义见第4节。

证明：只要证明$n = k$时的等式就可以了。利用第4节中定义的线性泛函L, 我们有

$$Ly = \sum_{i=0}^{k} \big(a_i y(x_i) - h b_i y'(x_i)\big) = \sum_{i=0}^{k} \big(a_i y(x_i) - h b_i f(x_i, y(x_i))\big)$$

另一方面，数值解满足等式：

$$0 = \sum_{i=0}^{k} \big(a_i y_i - h b_i f(x_i, y_i)\big)$$

因为我们已假定$y_i = y(x_i)$, $i < k$, 所以从上面两式相减得到

$$Ly = a_k(y(x_k) - y_k) - h b_k \big(f(x_k, y(x_k)) - f(x_k, y_k)\big)$$

对前一结果的最后一项应用中值定理，得到

$$\begin{aligned} Ly &= a_k(y(x_k) - y_k) - hb_k \frac{\partial f}{\partial y}(x_k, \xi)(y(x_k) - y_k) \\ &= (a_k - hb_k F)(y(x_k) - y_k) \end{aligned}$$

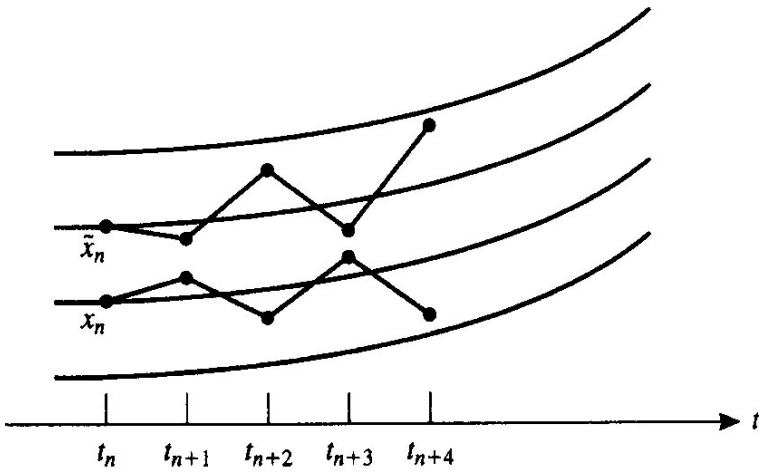
其中ξ位于$y(x_k)$和y_k之间，$F = \partial f(x_k, \xi)/\partial y$. 若使用的方法是$m$阶的，则

$$Ly = d_{m+1} h^{m+1} y^{(m+1)}(x_0) + \mathcal{O}(h^{m+2})$$

把上述两式组合起来，即证明了定理。这里略去了分母中的$hb_k F$.

整体截断误差

- 目标是建立起微分方程数值求解中的整体截断误差界。
- 在求解过程中任何给定的x_n步上，设计算出来的解为y_n，它不同于真解$y(x_n)$，差$y(x_n) - y_n$就是整体截断误差。
- 整体截断误差不只是前面点上出现的所有局部截断误差之和。因为在求解过程中每一步必须使用前面步上计算的近似纵坐标作为初值，而纵坐标是有误差的，所以数值过程实际上试图跟踪的是错误的解曲线
- 因此我们需要了解改变初值对解曲线上后面纵坐标的影响。



变分方程

- 考虑初值问题：

$$\begin{cases} y' = f(x, y), \\ y(0) = s \end{cases}$$

这里$f_y = \partial f / \partial y$连续，并且在$0 \leqslant x \leqslant T$, $y \in \mathbb{R}$定义的区域内满足$f_y(x, y) \leqslant \lambda$.

- 解是x的函数，但也与初值s有关，所以记为$y(x; s)$. 定义$u(x) = \partial y(x; s)/\partial s$.
- 对初值问题中u关于s的微分可得到一个微分方程(称为变分方程)为

$$\begin{cases} u'(x) = f_y(x, y)u, \\ u(0) = 1 \end{cases}$$

在初值问题

$$\begin{cases} y' = y^2, \\ y(0) = s \end{cases}$$

中显式地求出u.

- 这里$f(x,y) = y^2$, $f_y = 2y$, 因此变分方程为

$$\begin{cases} u' = 2yu, \\ u(0) = 1 \end{cases}$$

- 初值问题的解是$y(x) = \dfrac{s}{1-sx}$, 因此变分方程变为

$$\begin{cases} u'(x) = 2s(1-sx)^{-1}u(x), \\ u(0) = 1 \end{cases}$$

其解为

$$u(x) = \frac{1}{(1-sx)^2}$$

Theorem

若$f_y \leqslant \lambda$, 则变分方程的解满足不等式

$$|u(x)| \leqslant e^{\lambda x}, \quad x \geqslant 0$$

证明：从变分方程得到

$$u'/u = f_y = \lambda - \alpha(x)$$

其中$\alpha(x) \geqslant 0$. 对上式进行积分，得到

$$\log|u| = \lambda x - \int_0^x \alpha(\tau)d\tau = \lambda x - A(x)$$

由于$A(x) \geqslant 0$, 所以$\log|u| \leqslant \lambda x$, 即$|u| \leqslant e^{\lambda x}$ □

Theorem

若初值问题用初值s和$s+\delta$求解，则解曲线在x上差别至多为$|\delta|e^{\lambda x}$

证明：根据u的定义，对变分方程采用中值定理，再根据变分方程定理，得到

$$\begin{aligned} & |y(x;s)-y(x;s+\delta)| \\ =& \left|\frac{\partial}{\partial s}y(x,s+\theta\delta)\right||\delta| \\ =& |u(x)|\cdot|\delta| \leqslant |\delta|e^{\lambda x} \end{aligned}$$

□

整体截断误差界定理

Theorem

若在$x_1, x_2, \ldots, x_n$上的局部截断误差在数量上不超过δ, 则在x_n上的整体截断误差不超过

$$\delta \frac{e^{n\lambda h} - 1}{e^{\lambda h} - 1}$$

证明：设在$x_1, x_2, \ldots$上数值解的局部截断误差为$\delta_1, \delta_2, \ldots$. 在计算$y_2$时初始条件有一个$\delta_1$的误差，由初值问题解曲线定理，在解曲线上这个误差在x_2的影响至多是$|\delta_1|e^{\lambda h}$. 把这个值加到x_2的截断误差上，因此x_2的整体截断误差至多为$|\delta_1|e^{\lambda h} + |\delta_2|$. 这个误差在$x_3$上的影响不大于$(|\delta_1|e^{\lambda h} + |\delta_2|)e^{\lambda h}$, 把这个值加到$x_3$的截断误差上。以这个方式继续下去，得到在$x_n$上的整体截断误差不大于

$$\sum_{k=1}^{n} |\delta_k| e^{(n-k)\lambda h} \leqslant \delta \sum_{k=0}^{n-1} e^{k\lambda h} = \delta \frac{e^{n\lambda h} - 1}{e^{\lambda h} - 1}$$

整体截断误差逼近定理

Theorem

若数值解中局部截断误差是$\mathcal{O}(h^{m+1})$, 则整体截断误差是$\mathcal{O}(h^m)$.

证明：在整体截断误差界定理中，设δ是$\mathcal{O}(h^{m+1})$。因为e^z-1是$\mathcal{O}(z)$, $nh=x$, 所以整体截断误差的阶减少一。 □

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

刚性问题

- 在用微分方程描述的一个变化过程中，若往往又包含着多个相互作用但变化速度相差十分悬殊的子过程，这样一类过程就认为具有“刚性”。描述这类过程的微分方程初值问题称为“刚性问题”。
- 例如，宇航飞行器自动控制系统一般包含两个相互作用但效应速度相差十分悬殊的子系统，一个是控制飞行器质心运动的系统，当飞行器速度较大时，质心运动惯性较大，因而相对来说变化缓慢；另一个是控制飞行器运动姿态的系统，由于惯性小，相对来说变化很快，因而整个系统就是一个刚性系统。
- 用来描述这些过程的微分方程初值问题都是刚性问题。刚性问题解答的难度就在于其快变子系统的干扰，当我们试图在慢变区间上求解刚性问题时，尽管快变分量的值已衰减到微不足道，但这种快速变化的干扰仍严重影响数值解的稳定性和精度，给整个计算带来很大的实质性的困难。

$$y'(x) = \begin{pmatrix} 0 & 10 \\ -100 & -1001 \end{pmatrix} y(x)$$

$$\Rightarrow \quad Z'(x) = \begin{pmatrix} -1 & 0 \\ 0 & -1000 \end{pmatrix} Z(x)$$

$$Z_1(x) = Z_1(0)e^{-x} \quad Z_2(x) = Z_2(0)e^{-1000x} \simeq 0$$

精度由$Z_1(x)$决定, 稳定性由$Z_2(x)$决定.

考察初值问题

$$y'(x) = -30y(x), \quad y(0) = 1$$

在区间$[0, 0.5]$上的解。分别用欧拉显、隐式格式和改进的欧拉格式计算数值解。

节点x_i	欧拉显式	欧拉隐式	改进欧拉法	精确解$y = e^{-30x}$
0.0	1.0000	1.0000	1.0000	1.0000
0.1	-2.0000	2.5×10^{-1}	2.5000	4.9787×10^{-2}
0.2	4.0000	6.25×10^{-2}	6.2500	2.4788×10^{-3}
0.3	-8.0000	1.5625×10^{-2}	1.5625×10^{1}	1.2341×10^{-4}
0.4	1.6×10^{1}	3.9063×10^{-3}	3.9063×10^{1}	6.1442×10^{-6}
0.5	-3.2×10^{1}	9.7656×10^{-4}	9.7656×10^{1}	3.0590×10^{-7}

差分方程的绝对稳定性

考虑最简单的模型：只有初值产生误差，看看这个误差的传播。对于一般的差分方程

$$a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = h(b_k f_n + b_{k-1} f_{n-1} + \cdots + b_0 f_{n-k})$$

- 由初始误差产生了差分解的误差，实际上是同一差分方程，取不同初值所得到的2组差分解之间的差。
- 这个差不仅与差分方程本身有关，而且与微分方程本身有关。
- 如果微分方程本身是不稳定，那就没理由要求这2组解充分接近。
- 差分方程的稳定性概念是建立在微分方程稳定的基础上的。

把这个典型微分方程规定为：

$$\frac{dy}{dx} = \lambda y, \quad (Re\lambda < 0)$$

差分方程运用到如上的微分方程后，可以得到

$$a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = \lambda h(b_k y_n + b_{k-1} y_{n-1} + \cdots + b_0 y_{n-k})$$

对于给定的初始误差$e_0, e_1, \cdots, e_{k-1}$，误差方程具有一样的形式

$$a_k e_n + a_{k-1} e_{n-1} + \cdots + a_0 e_{n-k} = \lambda h(b_k e_n + b_{k-1} e_{n-1} + \cdots + b_0 e_{n-k})$$

绝对稳定性

差分方程称为绝对稳定的，若差分方程作用到微分方程

$$\frac{dy}{dx} = \lambda y, \quad (Re\lambda < 0)$$

时，对任意的初值，总存在左半复平面上的一个区域，当λh在这个区域时，差分方程的解趋于0。这个区域称为稳定区域

Euler显式公式的稳定性

$$y_{n+1} = y_n + \lambda h y_n$$

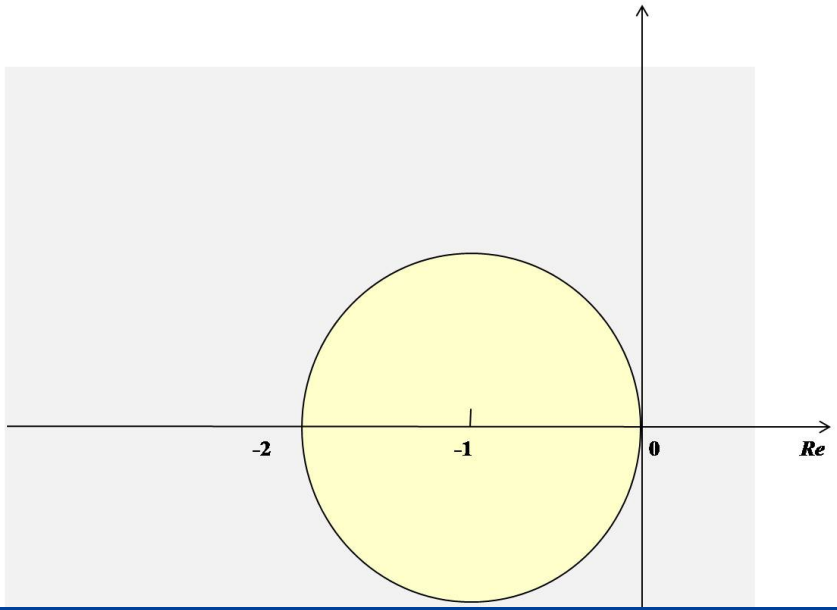
误差方程：

$$e_{n+1} = e_n + \lambda h e_n$$

令$\mu = \lambda h$, 绝对稳定区域为

$$\left| \frac{e_{n+1}}{e_n} \right| = |1 + \mu| < 1$$

绝对稳定区域为以-1为中心的圆盘



显式Runge-Kutta方法的稳定性

二阶Runge-Kutta方法

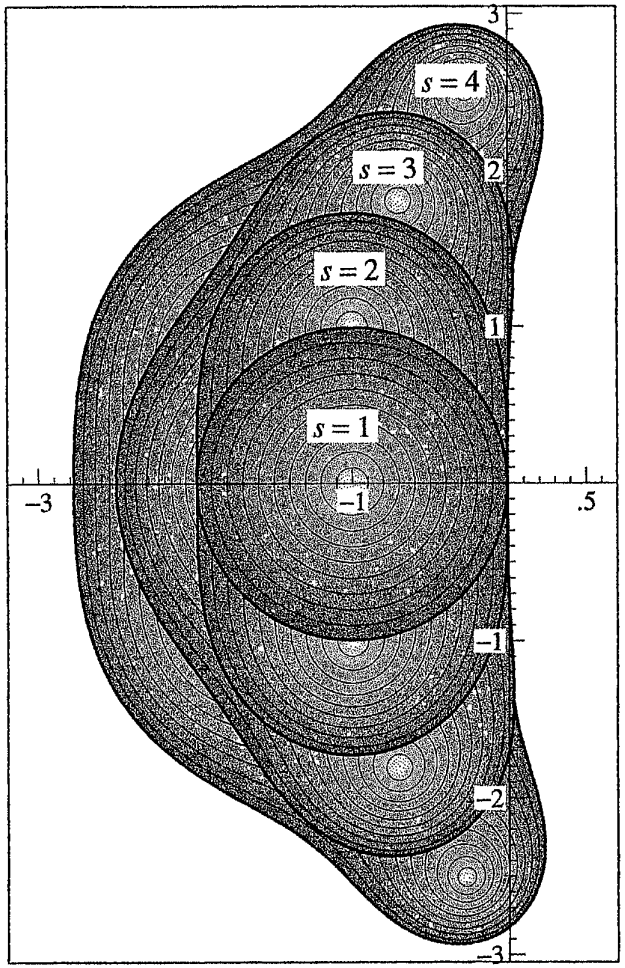
$$y_{n+1} = y_n + h\lambda y_n + \frac{h^2}{2}\lambda^2 y_n = (1 + \mu + \frac{\mu^2}{2})y_n$$

$$y_n = (1 + \mu + \frac{\mu^2}{2})^n y_0$$

绝对稳定区域为 $\Leftrightarrow$ $|1 + \mu + \frac{\mu^2}{2}| < 1$

显式Runge-Kutta方法的稳定性

1–4 阶显式Runge-Kutta方法的绝对稳定区域为



Euler隐式公式的稳定性

$$y_{n+1} = y_n + \lambda h y_{n+1}$$

$$y_{n+1} = \frac{1}{1-\lambda h} y_n$$

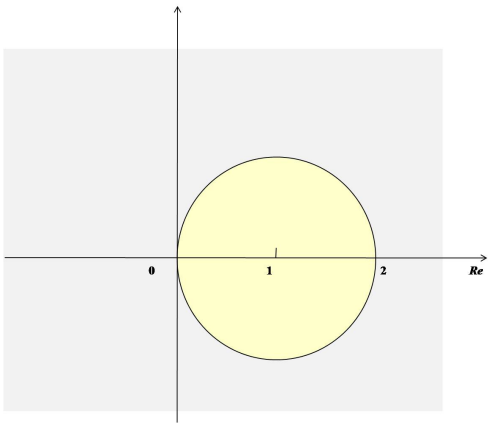
误差方程：

$$e_{n+1} = \frac{1}{1-\lambda h} e_n$$

绝对稳定区域为

$$|1-\mu| > 1$$

因此，绝对稳定区域包含$Re(\mu) < 0$的左半平面,此时方法称为A-stable.



隐式梯形公式的稳定性

$$y_{n+1} = y_n + \frac{h}{2}[f(x_n, y_n) + f(x_{n+1}, y_{n+1})]$$

$$y_{n+1} = y_n + \frac{h}{2}[\lambda y_n + \lambda y_{n+1}]$$

$$y_{n+1} = \frac{1 + \frac{\mu}{2}}{1 - \frac{\mu}{2}} y_n$$

绝对稳定区域为

$$\left|\frac{1 + \frac{\mu}{2}}{1 - \frac{\mu}{2}}\right| < 1$$

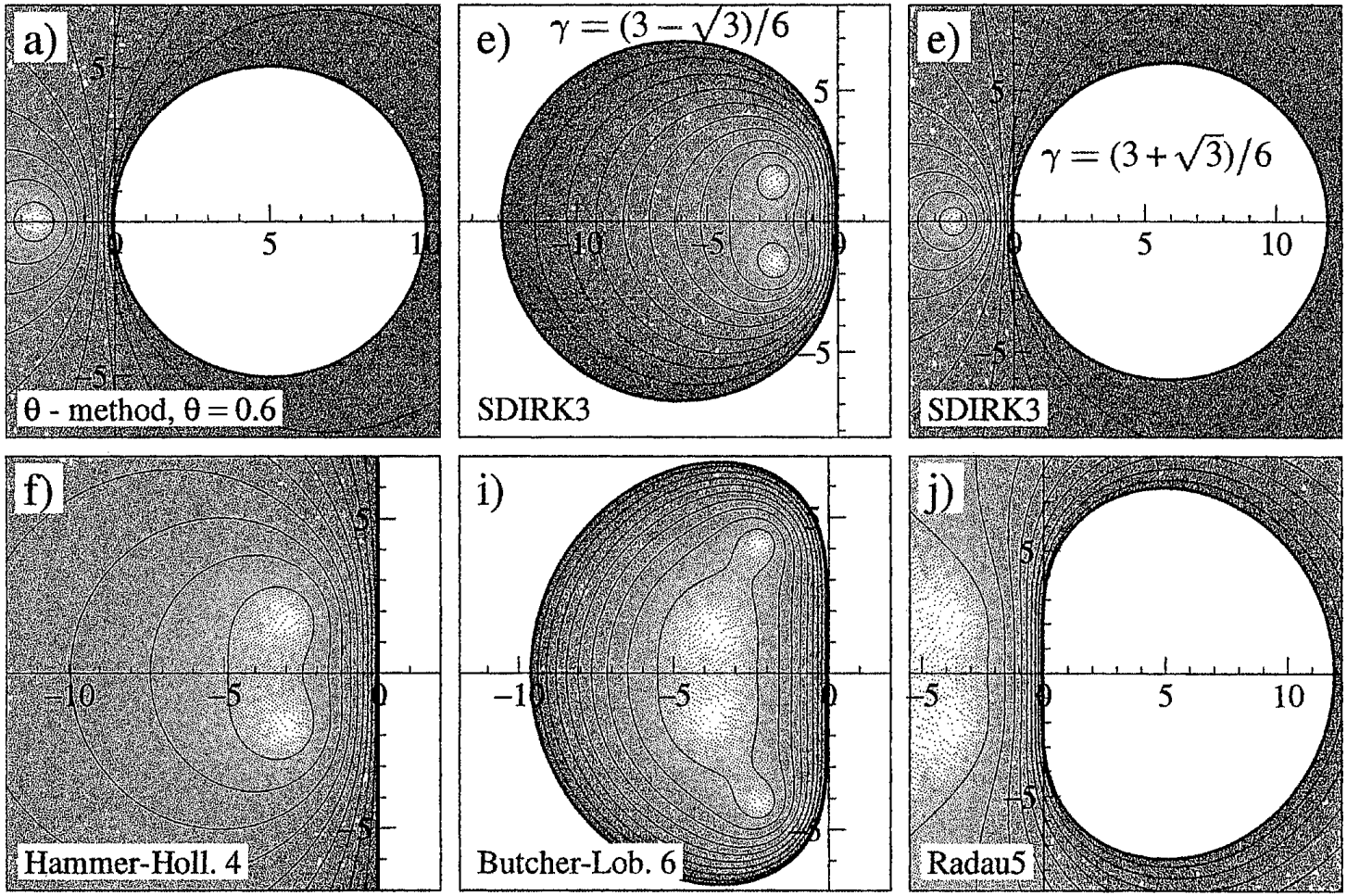
$$\mu = 2(a + bi) \quad |1 + a + bi|^2 \leq |1 - a - bi|^2$$

$$(1 + a)^2 + b^2 \leq (1 - a)^2 + b^2$$

$$a < 0 \quad Re(\mu) < 0$$

方法为A-stable。

隐式Runge-Kutta方法的稳定性



二阶Adam-Bashforth方法

$$y_{n+1} = y_n + \frac{h}{2}\left(3f(y_n) - f(y_{n-1})\right)$$

$$y_{n+1} = y_n + \frac{h\lambda}{2}\left(3y_n - y_{n-1}\right) = (1 + \frac{3}{2}\mu)y_n - \frac{\mu}{2}y_{n-1}$$

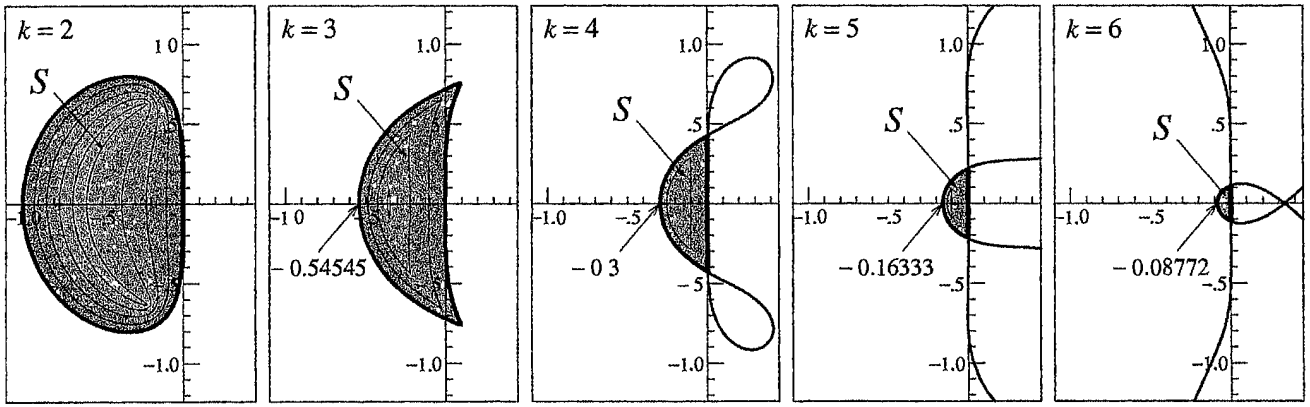
特征方程

$$z^2 - (1 + \frac{3}{2}\mu)z + \frac{\mu}{2} = 0$$

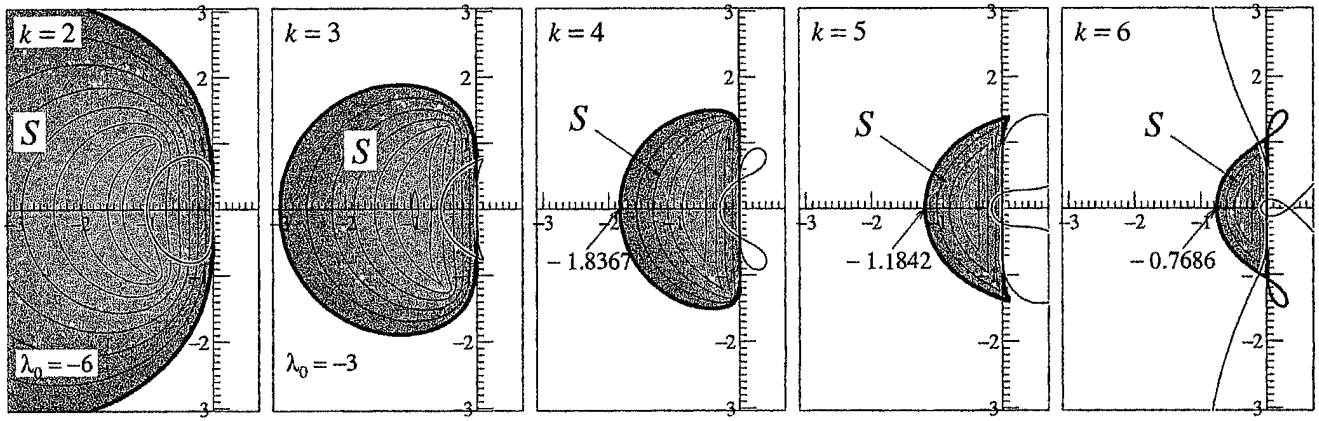
绝对稳定性要求特征方程的2个根在$|z| < 1$中.

多步方法的稳定性

显式多步方法



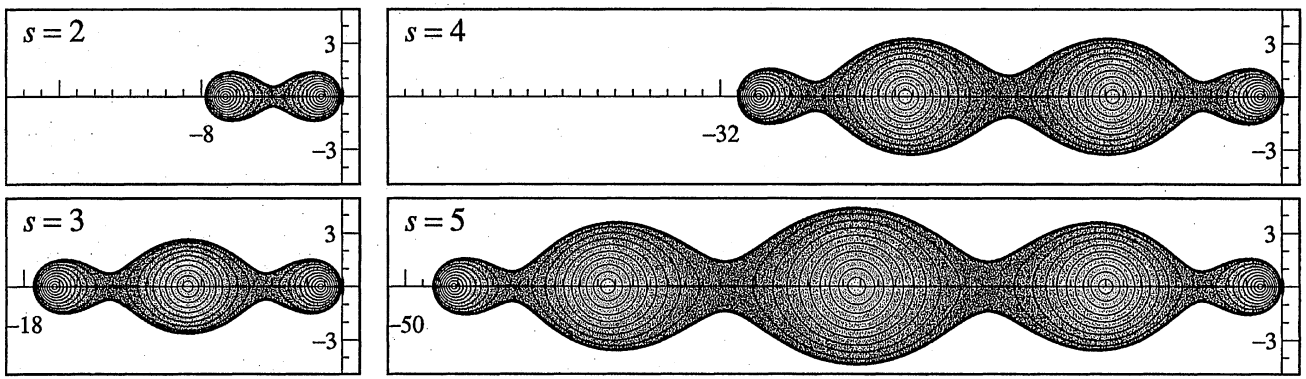
隐式多步方法



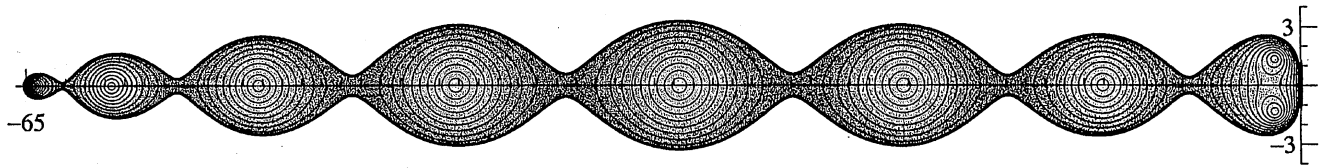
Theorem (Dahlquist)

- A-stable的线性多步方法必是隐式方法
- A-stable的线性多步方法至多为二阶精度.

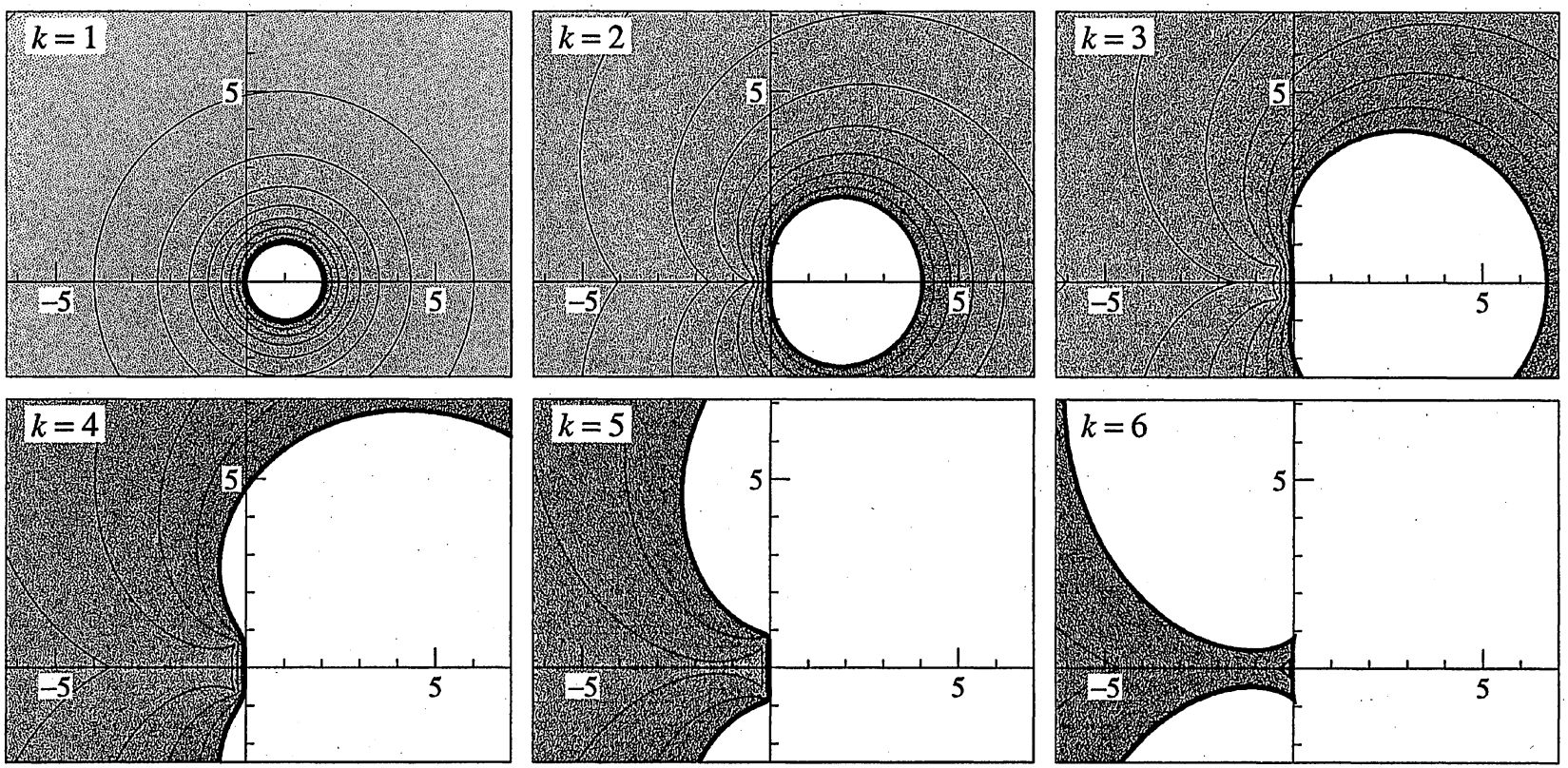
Chebyshev方法



Zolotarev方法



向后差商方法



程序作业

- 画出五阶Adams-Bashforth公式和五阶Adams-Moulton公式的绝对稳定性区域
- 可以使用Mathematica绘制,使用ImplicitPlot函数

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

初值问题与边值问题

- 我们现在可以求解如下方程：

$$\begin{cases} y'' = f(x, y, y') \\ y(a) = \alpha, \quad y'(a) = \beta \end{cases}$$

- 因为取$y_1 = y$, $y_2 = y'$可以把它转换成一阶方程组的形式

$$\begin{cases} y_1' = y_2, & y_1(a) = \alpha \\ y_2' = f(x, y_1, y_2), & y_2(a) = \beta \end{cases}$$

从而可以应用前面的步进方法进行求解。

- 然而如果问题改为

$$\begin{cases} y'' = f(x, y, y') \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

则前面方法失效。

边值问题的求解困难

- 步进方法不适合用于求解边值问题，因为没有完整的初值，数值求解无法开始。
- 前面是一个典型的两点边值问题。此类问题的求解难度要比初值问题大很多。
- 只有极个别的两点边值问题不需要用数值方法求解。如

$$\begin{cases} y'' = -y \\ y(0) = 3, \quad y(\pi/2) = 7 \end{cases}$$

方程的通解为$y(x) = A\sin x + B\cos x$，从而可以应用两点边值确定A, B，以得到方程的解为$y(x) = 7\sin x + 3\cos x$.

- 如果其中的微分方程通解不知道的话，刚才的方法无效。我们的目标是给出可处理任何两点边值问题的数值方法。

存在性和唯一性

- 一般来说，只假设f是一个“好”的函数并不能保证解的存在性。如

$$\begin{cases} y'' = -y \\ y(0) = 3, \quad y(\pi) = 7 \end{cases}$$

其中同前得到通解后，应用边值条件确定组合系数时，得到矛盾的方程组$3 = B$和$7 = -B$，因此问题无解。

- 关于两点边值问题解的存在性定理是相当复杂的。下面是Keller给出的一个结果。

Theorem (边值问题解的存在性定理)

当$\partial f/\partial y$连续、非负且在不等式$0 \leqslant x \leqslant 1, -\infty < y < +\infty$定义的无限带内有界时，边值问题

$$\begin{cases} y'' = f(x, y) \\ y(0) = 0, \quad y(1) = 0 \end{cases}$$

有唯一解

证明下述边值问题有唯一解：

$$\begin{cases} y'' = (5y + \sin 3y)e^x \\ y(0) = y(1) = 0 \end{cases}$$

- 这里

$$\frac{\partial f}{\partial y} = (5 + 3\cos 3y)e^x$$

它在无限带$0 \leqslant x \leqslant 1, -\infty < y < +\infty$内是连续的，而且它以$8e$为上界。另外，由于$3\cos 3y \geqslant -3$, 所以它是非负的。因此上述定理所需要的条件满足。

变量代换

- 前节定理讨论的是一种特殊情形。但是通过简单的变量代换，就可以把更一般的问题化为这里的特殊情形。
- 假设原问题为

$$\begin{cases} y'' = f(x, y) \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

令$x = a + (b - a)s$, $z(s) = y(a + \lambda s)$, $\lambda = b - a$. 则有$z'(s) = \lambda y'(a + \lambda s)$, $z''(s) = \lambda^2 y''(a + \lambda s)$. 同样地，$z(0) = y(a) = \alpha$, $z(1) = y(b) = \beta$, 于是若y是上述边值问题的解，则z是下述边值问题的解：

$$\begin{cases} z'' = \lambda^2 f(a + \lambda s, z(s)) \\ z(0) = \alpha, \quad z(1) = \beta \end{cases}$$

反之亦然，即若y是后者的解，
则$y(x) = z((x - a)/(b - a))$是前者的解。

两点边值问题第一定理

Theorem

考查下列两点边值问题：

$$1.\begin{cases} y'' = f(x,y) \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

$$2.\begin{cases} z'' = g(x,z) \\ z(0) = \alpha, \quad z(1) = \beta \end{cases}$$

其中$g(p,q) = (b-a)^2 f(a+(b-a)p, q)$. 若$z$是问题2的解，则函数$y(x) = z((x-a)/(b-a))$是问题1的解；反之，若$y$是问题1的解，则$z(x) = y(a+(b-a)x)$是问题2的解。

齐次化

- 为了简化两点边值问题

$$\begin{cases} y'' = g(x, y) \\ y(0) = \alpha, \quad y(1) = \beta \end{cases}$$

为一个具有齐次边值的问题，从y中减去一个在0和1取值为α和β的线性函数。

Theorem (两点边值问题的第二定理)

考查下列两点边值问题：

$$1. \begin{cases} y'' = g(x, y) \\ y(0) = \alpha, \quad y(1) = \beta \end{cases}$$

$$2. \begin{cases} z'' = h(x, z) \\ z(0) = 0, \quad z(1) = 0 \end{cases}$$

其中$h(p, q) = g(p, q + \alpha + (\beta - \alpha)p)$. 若$z$是问题2的解，则函数$y(x) = z(x) + \alpha + (\beta - \alpha)x$是问题1的解；反之，若$x$是问题1的解，则$z(x) = y(x) - (\alpha + (\beta - \alpha)x)$是问题2的解。

例

说明下列问题有唯一解:

$$\begin{cases} y'' = [5y - 10x + 35 + \sin(3y - 6x + 21)]e^x \\ y(0) = -7, \quad y(1) = -5 \end{cases}$$

- 边界值非齐次，不能直接应用Keller定理。首先齐次化，设

$$z(x) = y(x) - \ell(x), \quad \ell(x) = -7 + 2x$$

则

$$\begin{aligned} z'' = y'' &= [5y - 10x + 35 + \sin(3y - 6x + 21)]e^x \\ &= \{5(z + \ell) - 10x + 35 + \sin[3(z + \ell) - 6x + 21]\}e^x \\ &= \{5z + \sin 3z\}e^x \end{aligned}$$

新变量z的边界值为齐次的，根据前面的例题，此问题解存在唯一。

把下列问题转化为$[0,1]$区间上的齐次边界值问题：

$$\begin{cases} u'' = u^2 + 3 - x^2 + ux \\ u(3) = 7, \quad u(5) = 9 \end{cases}$$

- 由第一定理，此问题的等价问题为

$$\begin{cases} y'' = g(x,y) \\ y(0) = 7, \quad y(1) = 9 \end{cases}$$

其中$g(x,y) = 4f(3+2x,y)$
$= 4[y^2 + 3 - (3+2x)^2 + (3+2x)y]$。再由第二定理，另一个等价问题为

$$\begin{cases} z'' = h(x,z) \\ z(0) = 0, \quad z(1) = 0 \end{cases}$$

其中

$$h(x,z) = g(x, z+7+2x)$$
$$= 4[(z+7+2x)^2 + 3 - (3+2x)^2 + (z+7+2x)(3+2x)]$$

唯一解定理

Theorem (边值问题唯一解定理)

设f为(x,s)的连续函数，其中$0 \leqslant x \leqslant 1,\ -\infty < s < +\infty$. 假如在这个区域上

$$|f(x,s_1) - f(x,s_2)| \leqslant k|s_1 - s_2|, \quad k < 8$$

则两点边值问题

$$\begin{cases} y'' = f(x,y) \\ y(0) = y(1) = 0 \end{cases}$$

在$C[0,1]$中有唯一解。

- 证明：采用Green公式和Banach压缩映射定理。

例

证明下列问题有唯一解：

$$\begin{cases} y'' = 2e^{x\cos y} \\ y(0) = y(1) = 0 \end{cases}$$

- 这里$f(x,s) = 2e^{x\cos s}$，由中值定理，

$$|f(x,s_1) - f(x,s_2)| = \left|\frac{\partial f}{\partial s}(x,s_3)\right||s_1 - s_2|$$

其中所需要的导数满足

$$\left|\frac{\partial f}{\partial s}\right| = |2e^{x\cos s}(-x\sin s)| \leqslant 2e < 8$$

从而由定理，问题的解唯一。

有限差分法: 基本想法

- 同样考虑的是如下两点边值问题：

$$\begin{cases} y'' = f(x, y, y') \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

- 把区间$[a, b]$离散化为$a = x_0 < x_1 < \cdots < x_{n+1} = b$. 虽然不需要是均匀分布，但为了简化形式，后面假设

$$x_i = a + ih, \quad h = \frac{b-a}{n+1}, \quad i = 0, 1, \ldots, n+1$$

- 导数的近似计算公式为

$$y'(x) = \frac{1}{2h}\big(y(x+h) - y(x-h)\big) - \frac{1}{6}h^2 y'''(\xi)$$
$$y''(x) = \frac{1}{h^2}\big(y(x+h) - 2y(x) + y(x-h)\big) - \frac{1}{12}h^2 y^{(4)}(\tau)$$

离散形式

- 用y_i表示$y(x_i)$的近似值。把边值问题中的导数用前面的数值公式代替，则有如下离散形式

$$\begin{cases} y_0 = \alpha \\ \dfrac{1}{h^2}(y_{i-1} - 2y_i + y_{i+1}) = f\left(x, y_i, (y_{i+1} - y_{i-1})/(2h)\right), \\ \qquad\qquad i = 1, 2, \ldots, n \\ y_{n+1} = \beta \end{cases}$$

- 未知数为$y_1, \ldots, y_n$，方程个数为n. 若f为y_i的非线性形式，那么这些方程是非线性的，求解将变得非常困难。

- 现在假定f关于y和y'是线性的，即

$$f(x,y,y') = u(x) + v(x)y + w(x)y'$$

则上述方程组成为线性方程组，形式为

$$\begin{cases} y_0 = \alpha \\ \left(-1 - \frac{1}{2}hw_i\right)y_{i-1} + (2 + h^2 v_i)y_i \\ \qquad + \left(-1 + \frac{1}{2}hw_i\right)y_{i+1} = -h^2 u_i, \qquad i = 1, 2, \ldots, n \\ y_{n+1} = \beta \end{cases}$$

其中$u_i = u(x_i)$, $v_i = v(x_i)$, $w_i = w(x_i)$

- 引进缩写

$$a_i = -1 - \frac{1}{2}hw_{i+1}, d_i = 2 + h^2 v_i, c_i = -1 + \frac{1}{2}hw_i, b_i = -h^2 u_i$$

则方程组的矩阵形式为

$$\begin{pmatrix} d_1 & c_1 & & & & \\ a_1 & d_2 & c_2 & & & \\ & a_2 & d_3 & c_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-2} & d_{n-1} & c_{n-1} \\ & & & & a_{n-1} & d_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} b_1 - a_0\alpha \\ b_2 \\ b_3 \\ \vdots \\ b_{n-1} \\ b_n - c_n\beta \end{pmatrix}$$

- 系数矩阵为三对角的，所以可用特殊的Gauss消去法求解。
- 特别地，当h足够小，而且$v_i > 0$时, 矩阵是对角占优的，因为

$$|2 + h^2 v_i| > \left|1 + \frac{1}{2} h w_i\right| + \left|1 - \frac{1}{2} h w_i\right| = 2$$

- 在后面的收敛性分析中需要下面这个等式：

$$\begin{aligned} |d_i| - |c_i| - |a_{i-1}| &= 2 + h^2 v_i - \left(1 - \frac{1}{2} h w_i\right) - \left(1 + \frac{1}{2} h w_i\right) \\ &= h^2 v_i \end{aligned}$$

收敛性分析

- 下面证明当$h \to 0$时，离散解收敛于边值问题的解。为了知道边值问题

$$\begin{cases} y'' = f(x, y, y') \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

是否有唯一解，引用下面Keller给出的一个定理。

Theorem

边值问题

$$\begin{cases} y'' = f(x, y, y') \\ c_{11}y(a) + c_{12}y'(a) = c_{13} \\ c_{21}y(b) + c_{22}y'(b) = c_{23} \end{cases}$$

若满足下列条件，则在$[a, b]$上有唯一解

1. f及其一阶偏导数f_x, f_y, $f_{y'}$在域$D = [a, b] \times \mathbb{R} \times \mathbb{R}$上连续；
2. 在D上$f_y > 0$, $|f_y| \leqslant M$, $|f_{y'}| \leqslant M$；
3. $|c_{11}| + |c_{12}| > 0$, $|c_{21}| + |c_{22}| > 0$, $|c_{11}| + |c_{21}| > 0$, $c_{11}c_{12} \leqslant 0 \leqslant c_{21}c_{22}$

- 因此在线性问题中我们假设$u, v, w \in C[a, b]$, $v > 0$. 这样我们所考虑的两点边值问题有唯一解。
- 用$y(x)$表示问题的真解，y_i表示离散问题的解。这里y_i与h有关。我们将估计$|y(x_i) - y_i|$，并指出当$h \to 0$时它也趋向于零。
- $y(x)$满足下列方程组

$$\begin{aligned}&\frac{1}{h^2}(y(x_{i-1}) - 2y(x_i) + y(x_{i+1})) - \frac{1}{12}h^2 y^{(4)}(\tau_i)\\ &= u_i + v_i y(x_i) + w_i\left[\frac{1}{2h}(y(x_{i+1}) - y(x_{i-1}) - \frac{1}{6}h^2 y'''(\xi_i)\right]\end{aligned}$$

- 另外一方面，离散解满足下述方程

$$\frac{1}{h^2}(y_{i-1} - 2y_i + y_{i+1}) = u_i + v_i y_i + \frac{1}{2h}w_i(y_{i+1} - y_{i-1}))$$

- 两式相减，并记$e_i = y(x_i) - y_i$，则有

$$\frac{1}{h^2}(e_{i-1} - 2e_i + e_{i+1}) = v_i e_i + \frac{1}{2h} w_i (e_{i+1} - e_{i-1}) + h^2 g_i$$

其中

$$g_i = \frac{1}{12} y^{(4)}(\tau_i) - \frac{1}{6} y'''(\xi_i)$$

- 合并同类项，并且两边同乘以$-h^2$后，得到

$$\left(-1 - \frac{1}{2}hw_i\right)e_{i-1} + (2 + h^2 v_i)e_i + \left(-1 + \frac{1}{2}hw_i\right)e_{i+1} = -h^4 g_i$$

即

$$a_{i-1}e_{i-1} + d_i e_i + c_i e_{i+1} = -h^4 g_i$$

- 设$\lambda = \|e\|_\infty$，并且指标i满足

$$|e_i| = \|e\|_\infty = \lambda$$

这里e为向量$e = (e_1, e_2, \ldots, e_n)$. 这样从上式我们有

$$|d_i||e_i| \leqslant h^4|g_i| + |c_i||e_{i+1}| + |a_{i-1}||e_{i-1}|$$

- 因此有

$$\begin{aligned} |d_i|\lambda &\leqslant h^4\|g\|_\infty + |c_i|\lambda + |a_{i-1}|\lambda \\ \lambda(|d_i| - |c_i| - |a_{i-1}|) &\leqslant h^4\|g\|_\infty \\ h^2 v_i \lambda &\leqslant h^4\|g\|_\infty \\ \|e\|_\infty &\leqslant h^2 \frac{\|g\|_\infty}{\inf v(x)} \end{aligned}$$

其中$\|g\|_\infty \leqslant \frac{1}{12}\|y^{(4)}\|_\infty + \frac{1}{6}\|y'''\|_\infty$。而$\|g\|_\infty / \inf v(x)$是一个与$h$无关的项，因此当$h \to 0$时$\|e\|_\infty$是$\mathcal{O}(h^2)$。

配置法：基本想法

- 配置法所提供的思路可以用来解决应用数学中的许多问题。
- 假设给定一个线性算子L (例如，积分算子或者微分算子)，并且希望求解方程

$$Lu = w$$

其中w已知，u未知。

- 配置法求解此类问题的思路为：
  1. 选取某个基向量组$\{v_1, v_2, \ldots, v_n\}$，然后待定向量

$$u = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

  2. 为了尝试求解$Lu = w$, 把u的待定形式代入，得到

$$Lu = \sum_{j=1}^{n} c_j L v_j$$

从而得到

$$\sum_{j=1}^{n} c_j L v_j = w$$

- 一般来说，无法从

$$\sum_{j=1}^{n} c_j L v_j = w$$

中解出系数$c_1, c_2, \ldots, c_n$，但我们可以使之几乎成立。

- 在配置法中，向量u, w, v_j定义在相同的区域上。我们可以要求函数w与$\sum_{j=1}^{n} c_j L v_j$在n个给定点上的值相同，即

$$\sum_{j=1}^{n} c_j (L v_j)(x_i) = w(x_i), \quad i = 1, 2, \ldots, n$$

- 这是一个由n个方程，n个未知数构成的线性方程组。因此可以计算出所需要的系数。当然我们应该选择函数v_j和点x_i使得上述线性方程组对应的矩阵非奇异。

例：Sturm-Liouville边值问题

- 问题的描述为

$$\begin{cases} u'' + pu' + qu = w \\ u(0) = 0, \quad u(1) = 0 \end{cases}$$

其中p, q, w已知，并且在$[0,1]$上连续。未知函数u也定义在区间$[0,1]$上，但期望它是二阶连续的。

- 定义

$$Lu \equiv u'' + pu' + qu$$

- 定义向量空间：

$$V = \{u \in C^2[0,1] : u(0) = u(1) = 0\}$$

我们的目标是在V中寻找$Lu = w$的一个解。

- 如果从V中取一组基函数$\{v_1, v_2, \ldots, v_n\}$, 则齐次边界条件自然满足。一种选择是

$$v_{jk}(x) = x^j(1-x)^k, \quad j, k \geqslant 1$$

容易验证这组基函数满足

$$\begin{aligned} v'_{jk} &= jv_{j-1,k} - kv_{j,k-1}, \\ v''_{jk} &= j(j-1)v_{j-2,k} - 2jkv_{j-1,k-1} + k(k-1)v_{j,k-2} \end{aligned}$$

- 因此很容易写出Lv_{jk}的表达式。从而可以采用配置法求解前面的问题。

三次B样条

- 下面考虑稍微更一般的问题：

$$\begin{cases} u'' + pu' + qu = w \\ u(a) = \alpha, \quad u(b) = \beta \end{cases}$$

这时可能更好的基函数选择是B样条。

- 为了使基函数具有二阶连续导数，考虑三次B样条，并且为了简化记号，取$x_{i+1} - x_i = h$. 并且用样条结点作为配置点。
- 设n是采用的基函数个数。为了确定n个系数，需要n个条件。其中包括两个端点条件：

$$\sum_{j=1}^{n} c_j v_j(a) = \alpha, \quad \sum_{j=1}^{n} c_j v_j(b) = \beta$$

- 而由于维数为n的三次样条空间，有$n-2$个结点，因此恰好取这些内结点作为配置结点，从而得到另外的$n-2$个条件:

$$\sum_{j=1}^{n} c_j (Lv_j)(x_i) = w(x_i), \quad i = 1, 2, \ldots, n-2$$

其中

$$h = \frac{b-a}{n-3}, \quad x_i = a + (i-1)h$$

这样我们有$a = x_1 < x_2 < \cdots < x_{n-2} = b$. 另外，为了定义完全的B样条基函数，需要对这些结点进行扩充。

- 此时对应的系数矩阵是带状的，可以考虑如何充分利用这一稀疏性质以提高效率。

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

打靶法:基本想法

- 考查初值问题

$$\begin{cases} y'' = f(x, y, y') \\ y(a) = \alpha, \quad y'(a) = z \end{cases}$$

对任意的z，我们都可以应用前面的数值方法进行求解，记解为y_z。
因此对给定的β，当我们能选择z使得$y_z(b) = \beta$时，那么得到的y_z就是下列两点边值问题的解：

$$\begin{cases} y'' = f(x, y, y') \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

- 即首先猜测$y'(a)$的一个值，得到近似解，测试是否有$y(b) = \beta$. 若$y(b) \neq \beta$, 则修改猜测值，继续进行求解和测试。这个过程称为打靶(shooting).

求解非线性方程组

- 因此打靶法可以认为是求解下列的非线性方程：

$$\phi(z) \equiv y_z(b) - \beta$$

这里的$y_z(x)$函数没有显式定义，只是满足对任意给定z, 可以计算出对应的函数值，即新的边值与期望边值的差。

- 因此我们可以采用“数值代数”中求解非线性方程的方法进行求解。
  1. 二分法
  2. 割线法
  3. Newton法
  4. ……

割线法

- 割线法复习：对于方程$\phi(z)=0$以及给定的两个初值$\phi(z_1)$和$\phi(z_2)$，那么根是下述迭代的极限：

$$z_n = z_{n-1} - \frac{z_{n-1}-z_{n-2}}{\phi(z_{n-1})-\phi(z_{n-2})}\phi(z_{n-1})$$

- 当已经得到的z值使得$\phi(z)$几乎为零时，则停止这个迭代过程，并利用插值多项式去估计较好的零点值。
  - 假设$\phi(z_1),\ldots,\phi(z_n)$很小，这里我们的目标是构造一个多项式$p(x)$满足$p(\phi(z_i))=z_i,\ i=1,2,\ldots,n.$ 则下一个估计值是由$p(0)=z_{n+1}$确定。
  - 这相当于用多项式逼近ϕ的反函数。方法成功的前提是ϕ的根在一个邻域内有一个可微的反函数。

说明

- 打靶法是非常耗时的方法，因此下面考查如何可以更有效地应用打靶法得出所需要的数值解。
  - 显然应该充分发掘$y'(a)$的任何信息。因为高精度在打靶法的第一步基本上是被浪费的，因此可以考虑用大步长求解初值问题。只有当$\phi(z)$值几乎是零时才使用较小的步长。
- 有一类问题，对其应用割线法可以一步得到精确解。实际上，当ϕ为线性函数时就会发生这种情况。同样当微分方程是线性的时候也会出现这种情况。

- 在线性情况下，两点边值问题具有形式

$$\begin{cases} y'' = u(x) + v(x)y + w(x)y' \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

其中$u(x)$, $v(x)$, $w(x)$在区间$[a, b]$上连续。

- 假设已经用两个不同的初始条件两次求解上式相应的初值问题，得到解y_1和y_2，即

$$\begin{cases} y_1(a) = \alpha & y_1'(a) = z_1 \\ y_2(a) = \alpha & y_2'(a) = z_2 \end{cases}$$

- 考虑y_1和y_2的一个线性组合：

$$y(x) = \lambda y_1(x) + (1-\lambda)y_2(x)$$

其中λ为一个参数。容易验证$y(x)$满足微分方程以及第一个初值条件，即$y(a) = \alpha$. 可以选择参数λ使得$y(b) = \beta$, 即为了满足

$$\beta = y(b) = \lambda y_1(b) + (1-\lambda)y_2(b)$$

可得

$$\lambda = \frac{\beta - y_2(b)}{y_1(b) - y_2(b)}$$

从而对应的$y(x)$即为给定两点边值问题的解。

实现技巧

- 在计算机中实际上述想法时，我们可以通过下述方法同时得到y_1和y_2
  1. 所考虑的初值问题分别为

$$\begin{cases} y'' = f(x, y, y') \\ y(a) = \alpha, \quad y'(a) = 0 \end{cases} \qquad \begin{cases} y'' = f(x, y, y') \\ y(a) = \alpha, \quad y'(a) = 1 \end{cases}$$

  其中$f(x, y, y') = u(x) + v(x)y + w(x)y'$. 第一个的解为$y_1$，第二个的解为$y_2$
  2. 为产生x不显式出现的一阶方程组，令$y_0 = x$, $y_3 = y_1'$, $y_4 = y_2'$，因而带有初值的微分方程组为

$$\begin{cases} y_0' = 1 & y_0(a) = a \\ y_1' = y_3 & y_1(a) = \alpha \\ y_2' = y_4 & y_2(a) = \alpha \\ y_3' = f(y_0, y_1, y_3) & y_3(a) = 0 \\ y_4' = f(y_0, y_1, y_4) & y_4(a) = 1 \end{cases}$$

  3. 对$a = x_0 \leqslant x_i \leqslant x_m = b$，离散函数近似值$y_1(x_i)$和$y_2(x_i)$应当存放在内存中。$\lambda$的值应用前面的公式计算，然后再分别计算出$x_i$上相应的$y$值。

Theorem

若u, v, w是闭区间$[a, b]$上的连续函数，则对任何实数对α和α'，初值问题

$$\begin{cases} y'' = u(x) + v(x)y + w(x)y' \\ y(a) = \alpha, \quad y'(a) = \alpha' \end{cases}$$

在$[a, b]$上有唯一解。

Theorem

非齐次方程

$$y'' - vy - wy' = u$$

的每个解可以表示成$y_0 + c_1 y_1 + c_2 y_2$的形式，其中y_0为上述方程的特解，而y_1和y_2构成齐次方程

$$y'' - vy - wy' = 0$$

的线性无关的解集。

Theorem

若线性两点边值问题

$$\begin{cases} y'' = u(x) + v(x)y + w(x)y' \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

有解，而且若y_1不是解，那么则有$y_1(b) - y_2(b) \neq 0$, 而且y是所需要的解。*(*这里的y_1, y_2定义就是相应的初值问题的解，在a点导数值分别为*0*和*1)*

证明：设y_0, y_1, y_2分别是下列初值问题的解：

$$\begin{array}{lll} y_0'' = u + vy_0 + wy_0' & y_0(a) = \alpha & y_0'(a) = 0 \\ y_1'' = vy_1 + wy_1' & y_1(a) = 1 & y_1'(a) = 0 \\ y_2'' = vy_2 + wy_2' & y_2(a) = 0 & y_2'(a) = 1 \end{array}$$

由二阶线性微分方程理论，定理中给定的微分方程通解为

$$y_0 + c_1y_1 + c_2y_2$$

其中c_1, c_2为任意常数。

前面讨论的y_1和y_2是通解的特殊情况。它们由下式给出：

$$y_1 = y_0 + z_1 y_2, \quad y_2 = y_0 + z_2 y_2$$

我们已经假定定理中的两点边值问题有解，那么存在c_1, c_2使得

$$\alpha = y_0(a) + c_1 y_1(a) + c_2 y_2(a)$$
$$\beta = y_0(b) + c_1 y_1(b) + c_2 y_2(b)$$

其中第一个式即为$c_1 = 0$. 于是c_2应当满足的式子为

$$\beta = y_0(b) + c_2 y_2(b)$$

若$y_1(b) - y_2(b) \neq 0$, 则前面定义的y就是所需要的解。
若$y_1(b) - y_2(b) = 0$, 此即$y_2(b) = 0$, 从而$y_0(b) = \beta$, 所以y_1是所需的解。 □

Newton方法

- 现在讨论如何应用Newton方法求解两点边值问题。
- 设y_z为下列问题的解

$$\begin{cases} y_z'' = f(x, y_z, y_z') \\ y_z(a) = \alpha, \quad y_z'(a) = z \end{cases}$$

我们要选择z使得$\phi(z) \equiv y_z(b) - \beta = 0$.

- 关于ϕ的Newton公式是

$$z_{n+1} = z_n - \frac{\phi(z_n)}{\phi'(z_n)}$$

导数计算

- 为了确定ϕ'，对两点边值问题关于z求偏导，得到

$$\begin{cases} \dfrac{\partial y_z''}{\partial z} = \dfrac{\partial f}{\partial y_z}\dfrac{\partial y_z}{\partial z} + \dfrac{\partial f}{\partial y_z'}\dfrac{\partial y_z'}{\partial z} \\ \dfrac{\partial}{\partial z}y_z(a) = 0, \quad \dfrac{\partial}{\partial z}y_z'(a) = 1 \end{cases}$$

- 令$v = \partial y_z/\partial z$, 上式简化为

$$\begin{cases} v'' = f_{y_z}(x, y_z, y_z')v + f_{y_z'}(x, y_z, y_z')v' \\ v(a) = 0, \quad v'(a) = 1 \end{cases}$$

这是一个初值问题，称为第一变分方程。它可以与关于y_z的初值问题一起求解。然后利用$v(b)$得到$\phi'(z)$:

$$v(b) = \frac{\partial y_z(b)}{\partial z} = \phi'(z)$$

从而可以应用Newton方法求解问题。

多重打靶法

- 多重打靶法(multiple shooting)是打靶法的一个重要发展。其基本策略是把给定的区间$[a,b]$分成子区间，并试图在每个小段上求解整体问题。
- 下面以区间$[a,b]$被分成$[a,c]$和$[c,b]$的情况说明多重打靶法。此时考虑的问题仍然为

$$\begin{cases} y'' = f(x,y,y') \\ y(a) = \alpha, \quad y(b) = \beta \end{cases}$$

- 在每个子区间上，求解下列两个初值问题，得到解为y_1和y_2.

$$\begin{cases} y_1'' = f(x,y_1,y_1') \quad y_1(a) = \alpha \quad y_1'(a) = z_1, \quad a \leqslant x \leqslant c \\ y_2'' = f(x,y_2,y_2') \quad y_2(b) = \beta \quad y_2'(b) = z_2, \quad c \leqslant x \leqslant b \end{cases}$$

这里z_1和z_2是所配置的参数。y_2的数值解按x递减方向进行。

- 下面的目标是调整参数z_1和z_2直到分段函数

$$y(x) = \begin{cases} y_1(x) & a \leqslant x \leqslant c \\ y_2(x) & c \leqslant x \leqslant b \end{cases}$$

变成问题的解。因此需要y和y'在c点上满足：

$$y_1(c) - y_2(c) = 0, \quad y_1'(c) - y_2'(c) = 0$$

- 通常选择z_1和z_2可以实现这一目标。可以采用二维Newton方法处理这一问题。
- 对k个子区间的多重打靶法将涉及到k个子函数。每个子函数通过求解一个初值问题得到。这k个子函数的初值构成一个有$2k$个参数的集合。在区间的$k-1$个内分点上的连续性得到$2k-2$个条件，再加上端点条件，正好参数个数与条件个数匹配。同样采用非线性方程组迭代求解。

《数值分析》之

常微分方程数值方法

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu

常微分方程组的数值解法

- 一阶微分方程组的标准形式为

$$\begin{cases} y_1' = f_1(x, y_1, y_2, \ldots, y_n) \\ y_2' = f_2(x, y_1, y_2, \ldots, y_n) \\ \quad\vdots \\ y_n' = f_n(x, y_1, y_2, \ldots, y_n) \end{cases}$$

- 在这个方程组中，要确定n个未知函数$y_1, y_2, \ldots, y_n$. 它们是单变量x的函数，记号y_i'表示dy_i/dx

- $$\begin{cases} y_1' = y_1 + 4y_2 - e^x \\ y_2' = y_1 + y_2 + 2e^x \end{cases}$$
- 通解为

$$y_1(t) = 2ae^{3x} - 2be^{-x} - 2e^x$$
$$y_2(t) = ae^{3x} + be^{-x} + \frac{1}{4}e^x$$

其中a, b为任意常数。

初值条件

- 在估计可能具有唯一解的明确定义的物理问题中，微分方程组会伴有确定通解中任意常数的辅助条件。
- 前例中的初始条件可以是

$$y_1(0) = 4, \qquad y_2(0) = \frac{5}{4}$$

则解为

$$y_1(t) = 4e^{3x} + 2e^{-x} - 2e^{x}$$
$$y_2(t) = 2e^{3x} - e^{-x} + \frac{1}{4}e^{x}$$

- 一般的方程组的初值问题是由n个微分方程连同给定的在$x = x_0$的初值所组成。

向量记号

- 可以采用向量记号来改写方程组。设Y表示分量为$y_1, y_2, \ldots, y_n$的列向量，因此Y是$\mathbb{R}$或其中一个区间到$\mathbb{R}^n$的一个映射。
- 类似地，设F表示具有分量$f_1, f_2, \ldots, f_n$的列向量，每个分量是$\mathbb{R}^{n+1}$或它的一个子集上的一个函数
- 方程组可以写为

$$Y' = F(x, Y)$$

方程组的初值问题还包括向量$Y(X_0)$的数值。

高阶微分方程

- 高阶微分方程可以转换为一阶微分方程组。假设给定下列形式的单个微分方程

$$y^{(n)} = f(x, y, y', y'', \ldots, y^{(n-1)})$$

它可以转化为

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \quad \vdots \\ y_n' = f(x, y_1, y_2, \ldots, y_n) \end{cases}$$

- 这种转换后就可以在某些软件上采用求解常微分方程组的过程求解高阶微分方程
- 高阶方程组也可以类似转换为一阶微分方程组

Taylor级数方法

- 完全类似于标量方程讨论的Taylor级数方法。
- 对每个函数写出如下的截断Taylor级数

$$y_i(x+h) \approx y_i(x) + hy_i'(x) + \frac{h^2}{2!}y_i''(x) + \cdots + \frac{h^n}{n!}y_i^{(n)}(x)$$

向量记号为

$$Y(x+h) \approx Y(x) + hY'(x) + \frac{h^2}{2!}Y''(x) + \cdots + \frac{h^n}{n!}Y^{(n)}(x)$$

其中出现的导数可以从微分方程组中得到。

- 通常当这些导数用于计算机编程时，需要采用特定的次序进行计算。

对下列初值问题采用三阶Taylor级数方法，$h = -0.1$, 在区间$[-2, 1]$上计算

$$\begin{cases} y_1' = y_1 + y_2^2 - x^3 & y_1(1) = 3 \\ y_2' = y_2 + y_1^3 + \cos x & y_2(1) = 1 \end{cases}$$

- 需要的高阶导数是

$$\begin{aligned} y_1'' &= y_1' + 2y_2y_2' - 3x^2 \\ y_2'' &= y_2' + 3y_1^2y_1' - \sin x \\ y_1''' &= y_1'' + 2y_2y_2'' + 2(y_2')^2 - 6x \\ y_2''' &= y_2'' + 6y_1(y_1')^2 + 3y_1^2y_1'' - \cos x \end{aligned}$$

- 执行Mathematica程序“odes_taylor_3.nb”查看结果

自控方程组

- 可以通过引入新变量，使得方程组中不出现变量x。具体做法是令$(y_1)_0 = x$, 然后再加入一个新方程$(y_1)_0' = 1$就可以实现这种转化。从而方程组可以写为

$$Y' = F(Y)$$

这种方程组称为自控的(autonomous)

Runge-Kutta方法

- 当方程组具有形式$Y' = F(x, Y)$时，可以采用Runge-Kutta方法求解。
- 经典的向量形式的四阶Runge-Kutta公式是

$$Y(x+h) = Y(x) + \frac{1}{6}(F_1 + 2F_2 + 2F_3 + F_4)$$

其中

$$\begin{aligned} F_1 &= hF(x, Y) \\ F_2 &= hF(x + h/2, Y + F_1/2) \\ F_3 &= hF(x + h/2, Y + F_2/2) \\ F_4 &= hF(x + h, Y + F_3) \end{aligned}$$

- 类似地，也可以给出Runge-Kutta-Fehlberg公式。

- 多步法也可以推广到方程组
- 例如，Adams-Bashforth-Moulton预估–校正方法的向量形式为

$$
\begin{aligned}
Y^*(x+h) = Y(x) + \frac{h}{720}\Big[&1901F(Y(x)) - 2774F(Y(x-h)) \\
&+ 2616F(Y(x-2h)) - 1274F(Y(x-3h)) \\
&+ 251F(Y(x-4h))\Big]
\end{aligned}
$$

$$
\begin{aligned}
Y(x+h) = Y(x) + \frac{h}{720}\Big[&251F(Y^*(x+h)) + 646F(Y(x)) \\
&- 264F(Y(x-h)) + 106F(Y(x-2h)) \\
&- 19F(Y(x-3h))\Big]
\end{aligned}
$$

这时需要提供用一个单步过程(如五阶Runge-Kutta方法)提代起始值：$Y(x_0+h)$, $Y(x_0+2h)$, $Y(x_0+3h)$, $Y(x_0+4h)$

《数值分析》之

考试说明

徐岩

中国科学技术大学数学系

yxu@ustc.edu.cn

https://faculty.ustc.edu.cn/yxu/

定理证明

1. §6.1.1 多项式插值定理
2. §6.1.3 多项式插值误差定理
3. §6.3.3 Hermite插值误差估计定理
4. §6.8.6 正交多项式定理
5. §7.2.7 第二类Tchebyshev多项式的极值性质定理
6. §7.3.2 带误差项的Gauss积分公式定理
7. §8.5.2 收敛的多步法必定是稳定的和相容的
8. §8.9.3 有限差分法的收敛阶为$\mathcal{O}(h^2)$.

多项式插值与逼近

- 插值多项式的分类(Lagrange, Newton, Hermite)、计算、误差
- 均差(差商)的定义与应用: 有无重节点
- 两类Tchebyshev多项式的性质、递推公式，以及一般正交多项式的性质、计算公式和应用
- 最小二乘逼近：基于L_2逼近与线性方程组理论的最小二乘方法有什么关系？最佳逼近元的计算
- 最佳逼近的Tchebyshev理论：对于定理证明不做要求，但是会应用特征定理及等价条件。

数值微分与积分

- 数值微分的计算方法
- 数值积分的代数精度，格式构造，复合公式构造
- Gauss积分公式：格式确定、性质、误差
- Romberg积分与Richardson外推技术

常微分方程数值解

- 初值问题解的存在性与唯一性定理
- 初值问题求解方法(单步法，多步法)，误差分类
- 二阶Runge-Kutta方法的建立，一般形式
- 绝对稳定区域的判定
- 线性多步法的建立，理论分析：精度，稳定性和相容性(收敛性)
- 两点边值问题的求解思路

说明

- 考试教室: 5503
- 请携带一卡通或身份证
- 各种类型的手机、电脑、平板电脑、具有网络传输功能的电子手表、计算器均不允许携带
- 考试时参考书、电子产品请关机后放到书包里，书包需放到教室前部。
- 考试时间：2024年5月15日14:00-16:00
- 习题课：2023年5月11日和5月13日正常上课时间。