西安交通大學

专业学位研究生教育系列教材

计算方法

基本内容与解题方法

凌永祥

西安交通大学出版社
XI'AN JIAOTONG UNIVERSITY PRESS

西安交通大學

专业学位研究生教育系列教材

计算方法
基本内容与解题方法

凌永祥

西安交通大学出版社
XI'AN JIAOTONG UNIVERSITY PRESS

内容提要

本书是与《计算方法》(专业学位研究生教育系列教材)配套的辅导书。内容包括了数值分析的基础、线性方程组解法、非线性方程求解、多项式插值和数值微积分等内容。基本上与所配套的教材相一致,但也略有推广。

本书可供学习《计算方法》课程的研究生和本科生参考使用。也可供从事数值计算的科技工作人员和工程技术人员参考。

图书在版编目(CIP)数据

计算方法基本内容与解题方法/凌永祥编著. —西安:西安交通大学出版社,2009.3

ISBN 978-7-5605-3023-9

Ⅰ.计… Ⅱ.凌… Ⅲ.计算方法-研究生-教材 Ⅳ.0241

中国版本图书馆 CIP 数据核字(2009)第 003423 号

书　　名 计算方法基本内容与解题方法
编　　著 凌永祥
责任编辑 叶　涛

出版发行 西安交通大学出版社
(西安市兴庆南路 10 号　邮政编码 710049)
网　　址 http://www.xjtupress.com
电　　话 (029)82668357　82667874(发行中心)
(029)82668315　82669096(总编办)
传　　真 (029)82668280
印　　刷 西安市新城区兴庆印刷厂

开　　本 727 mm×960 mm　1/16　**印张** 8.375　**字数** 153 千字
版次印次 2009 年 3 月第 1 版　2009 年 3 月第 1 次印刷
书　　号 ISBN 978-7-5605-3023-9/O·290
定　　价 15.00 元

读者购书、书店添货、如发现印装质量问题,请与本社发行中心联系、调换。
订购热线:(029)82665248　(029)82665249
投稿热线:(029)82664954
读者信箱:jdlgy@yahoo.cn

版权所有　侵权必究

前言

本书是为学习《计算方法》(专业学位系列教材)编写的习题辅导书。

初学《计算方法》的人,大多会认为这是一门比较难学的课程。确实,它的内容涉及到基本的微积分(例如数值积分、数值微分)、线性代数、微分方程等高等数学的许多分支,但是,其思想方法与过去学习过的数学课程大不相同,因此在演算《计算方法》的习题时往往会感到束手无策。本书的目的就是努力通过对《计算方法》课程内容的总结和例题的分析,帮助《计算方法》课程的学习者通过解习题,进一步理清思路,了解算法的实质。

大学里开设《高等数学》课程有两个重要目的:一是学习、研究和掌握现代科学技术必须具有的一个基本工具,二是培养训练分析问题的能力,形成严密的逻辑思维方法。随着现代计算技术的发展,《高等数学》的工具功能正通过《计算方法》得到不断的延伸;而严密的逻辑思维、分析问题的能力在学习《计算方法》的过程中也得到了应用。许多课程在刚开始学习的时候,常常感到似乎是一团乱麻,但当课程结束后,回过头来仔细清理一下,就会发现在一门课程(尤其是数学课)中实际上贯穿着一条清晰的线索。例如《计算方法》中就有多项式插值贯穿了数值微积分与微分方程数值解,而解方程(包括线性和非线性方程)的迭代法实际上也是统一的。

本书努力通过对《计算方法》课程内容的总结和一些例题的分析、比较,使知识条理化,系统化,使《计算方法》介绍的算法,不再是抽象的理论推导,而是实际可用的算法。知道这些算法可以解决什么问题,从而在遇到数学软件中的各种方法时,知道它们有什么用;反过来,遇到实际问题时也可以联想到可以找哪类方法去解决。

本书包括了数值分析的基础、线性方程组解法、非线性方程求解、多项式插值和数值微积分等内容。基本上与所配套的教材相一致,但也略有推广;由于原教材没有介绍微分方程的内容,所以本书也没有将这方面的习题内容列入。如果需要,将来再版时再考虑加入相应内容。

本书的内容,作者已酝酿了很久,但由于种种原因,一直没有动手编写。这次由于西安交通大学研究生院和西安交通大学出版社的支持,特别是本书编辑叶涛先生的鼓励,才将本书编写完成,作者在此对他们表示由衷的感谢。由于第一次编写习题辅导,必定存在不少缺点和错误,衷心希望同行专家和广大读者提出批评和建议,以便将来再版时改正与提高。

凌永祥

2008 年 12 月于西安交通大学

目　录

此为试读,需要完整PDF请访问: www.ertongbook.com

第 1 章　数值方法基础，近似数与误差

1.1　基本内容

1.1.1　近似数与误差

数值方法一般指在电子计算机上进行的数值计算过程。由于计算机中数的存储单元和运算器中的字长都是有限的，因此，无论是数的输出、输入，或数的运算，总会产生误差.

(1)误差

若准确数记为 x，近似数记为 $\tilde{x}$，则记

绝对误差

$$\delta x = x - \tilde{x} \tag{1.1}$$

相对误差

$$\Delta x = \frac{|\delta x|}{|x|} \approx \frac{|\delta x|}{|\tilde{x}|} \tag{1.2}$$

而这些误差的最大范围称为误差界：

绝对误差界

$$\delta = \max|\delta x| \tag{1.3}$$

相对误差界

$$\Delta = \max\{\Delta x\} \tag{1.4}$$

(2)浮点数系

浮点数系 $F(\beta, t, L, U)$ 是指一个 β 进制，尾数字长为 t 位，首数（或指数）的下、上限分别为 L、U 的如下形式的实数和零的集合

$$\pm\left(\frac{d_1}{\beta}+\frac{d_2}{\beta^2}+\cdots+\frac{d_t}{\beta^t}\right)\times\beta^l \tag{1.5}$$

其中 $1 \leqslant d_1 < \beta$，$0 \leqslant d_j < \beta$，$j = 2, 3, \cdots, t$，$L \leqslant l \leqslant U$.

浮点数系 $F(\beta, t, L, U)$ 中数字的总数有限，共有

$$2(\beta-1)\beta^t(U-L+1)+1$$

个数,若此浮点数系中最小和最大的正数记为 m、M,则任何一个绝对值在$[m, M]$中的实数 x,按“四舍五入”的方式,在浮点数系中都有一个对应的浮点数 $fl(x)$,使两者的绝对误差最小

$$|x-fl(x)|\leqslant\frac{1}{2}\times\beta^{l-t} \tag{1.6}$$

由于 $m=\frac{1}{2}\times\beta^{L-1}$, $M<\beta^{U}$,因此当$|x|<\frac{1}{2}\times\beta^{L-1}$时 $fl(x)=0$,此时 x 就在计算机中显示为“0”,在计算机运算中称为“机器零”(或“下溢”),而当$|x|\geqslant\beta^{U}$时,$fl(x)=\infty$,在计算机运算中被称为“机器无穷大”(或“上溢”).

(3) 有效数字

“有效数字”是指一个近似数的“有意义”的数字的数位,通常都在十进制数系中讨论,设 $x=\pm(0.\alpha_1\alpha_2\cdots\alpha_t\cdots)\times10^{l}\in\mathbf{R}$,其中 $\alpha_i\in\{0, 1, \cdots, 9\}$, $\alpha_1\neq0$, 近似数 $\tilde{x}=\pm0.\alpha_1\alpha_2\cdots\tilde{\alpha}_t\times10^{l}$,若$|x-\tilde{x}|\leqslant\frac{1}{2}\times10^{l-t}$,则称 $\tilde{x}$ 是 x 的具有 t 位有效数字的近似数,或简称 $\tilde{x}$ 具有 t 位有效数字. 事实上,比较式(1.6)可知,t 位有效数字就是一般的实数在十进制浮点数系 $F(10, t, L, U)$中相应数字.

事实上,$\tilde{\alpha}_t$ 有两种可能:若 $\alpha_{t+1}\leqslant4$, 则 $\tilde{\alpha}_t=\alpha_t$; 若 $\alpha_{t+1}\geqslant5$, 则 $\tilde{\alpha}_t=\alpha_t+1$,对于具有 t 位有效数字的近似数 $\tilde{x}$,由于$|x|\geqslant0.\alpha_1\times10^{l}$,因此相对误差界 Δ 满足:

$$\Delta x=\frac{|x-\tilde{x}|}{|x|}\leqslant\frac{1}{2\alpha_1}\times10^{1-t}=\Delta \tag{1.7}$$

(4) 数值计算应注意的事项

(i) 避免相近数相减,以免有效数字损失

因为在浮点数系中运算时,参加运算的数字的数位有限,两个相近数若其前 s 位都相同,那么相减后就会只剩下 $t-s$ 位有效数,如果 $t-s$ 与 t 相比很小时,这样的运算结果显然损失了大部分有效数字,它必然会影响以后它参与运算的结果的数值的有效数位.

(ii) 避免以绝对值很小的数作除数

若以小数 ε 为除数,由于 $\delta x=x-\tilde{x}$,则若不计除法自身产生的误差,理论结果 x/ε 与实际计算结果 $\tilde{x}/\varepsilon$ 之误差:$\delta(x/\varepsilon)=\delta x/\varepsilon$,说明原始误差 δx 经过除法运算被放大了 $1/\varepsilon$ 倍,当 ε 很小时,$1/\varepsilon$ 将是个很大的数.

(iii) 避免大小数相加减

浮点数系中的四则运算,最重要的数值运算是对尾数进行运算. 相加减的一对数字,若两者的数量差异很大,它们进入运算器中进行运算时,数量级低(即指数 l 小)的一方,将变动其指数 l,变成与数量级高的数字有相同的指数. 这样数量级低的一方,真尾数的前几位在必须增加若干个“0”的同时,必须放弃其后部同等个

数的尾数，导致若干有效数字的损失，严重时可能出现"大数吃小数"的现象. 因此，如果有一系列数字相加减，应尽可能按数字的绝对值从小到大的顺序进行运算.

（ⅳ）简化运算步骤，减小运算次数

由于每次浮点运算，通常会产生舍入误差，因此在运算前应尽可能作适当变换，简化运算步骤，减小浮点运算量，减少舍入误差发生的机会.

1.1.2　问题的性态

若记问题 f 的原始数据为 x，结果为 $f(x)$. 当原始数据发生扰动 δx，变为 $x+\delta x$ 后，问题的结果就随之变为 $f(x+\delta x)$. 称两者相对误差之比的上界为条件数：

$$\mathrm{Coud}\,(f)=\mathrm{Sup}\left\{\frac{\dfrac{|f(x+\delta x)-f(x)|}{|f(x)|}}{\dfrac{|\delta x|}{|x|}}\right\} \tag{1.8}$$

若条件数 Coud(f)大，则称问题 f 是病态的，若条件数 Coud(f)小，则称问题 f 是良态的. 事实上，条件数反映了问题结果关于原始数据发生扰动的敏感性.

1.1.3　方法的稳定性

一个求解问题的方法是经过多步计算才能取得问题的解，当初始值有误差，通过逐步计算后，若其误差的发展总能被有效地控制，则称此方法是稳定的，否则就是不稳定的.

1.1.4　正交多项式

(1)函数的正交

给定定义在(a,b)上的函数 $\rho(x)$，$\rho(x)>0\quad\forall x\in(a,\ b)$. 对于定义在$(a,b)$上的连续函数 $f(x)$，$g(x)$，称

$$(f,\ g)=\int_a^b\rho(x)f(x)g(x)\mathrm{d}x \tag{1.9}$$

为函数 f 与 g 在(a,b)上(以 $\rho(x)$ 为权函数)的内积. 显然，函数的内积具有以下性质

（ⅰ）$(f,g)=(g,f)$

（ⅱ）$(\alpha f,\ g)=\alpha(f,\ g)\quad\forall\alpha\in\mathbf{R}$

（ⅲ）$(f_1+f_2,\ g)=(f_1,\ g)+(f_2,\ g)$

若$(f,\ g)=0$，称函数 $f(x)$、$g(x)$在(a,b)上(按权函数 $\rho(x)$)正交.

(2) 正交多项式

记 $\varphi_n(x)$为最高项系数非零的 n 次多项式，如果多项式函数系$\{\varphi_k\}$满足

$$(\varphi_i, \varphi_j) = \int_a^b \rho(x)\varphi_i(x)\varphi_j(x)\mathrm{d}x = \begin{cases} 0, & i \neq j \\ A_i, & i = j \end{cases} \tag{1.10}$$

则称$\{\varphi_k\}$为(a,b)上以$\rho(x)$为权函数的正交多项式系.

(3)正交多项式系的递推

令$\varphi_{-1}(x)=0$, $\varphi_0(x)=1$,则

$$\varphi_{k+1}(x) = \left(x - \frac{(x\varphi_k, \varphi_k)}{(\varphi_k, \varphi_k)}\right)\varphi_k(x) - \frac{(\varphi_k, \varphi_k)}{(\varphi_{k-1}, \varphi_{k-1})}\varphi_{k-1}(x), \tag{1.11}$$

组成最高项系数为1的正交多项式系$\{\varphi_k(x)\}$.

正交多项式系有以下主要性质:

(ⅰ) $\varphi_k(x)$在(a,b)上有k个互异的实根;

(ⅱ) 任何k次多项式总可表示为$\{\varphi_0, \varphi_1, \cdots, \varphi_k\}$的线性组合.

1.2 例题与分析

例1-1 试给出数系$F(10, 4, -5, 5)$中的最大数、最小数和最小正数.

解 最大数:0.9999×10^5;最小数-0.9999×10^5;最小正数:0.1000×10^{-5}.

例1-2 已知$e=2.71828182845904523536028747\cdots$.求它在$F(10, 5, -5, 5)$和$F(10, 8, -5, 5)$中的浮点化数.

解 在$F(10, 5, -5, 5)$中的$fl(e)=0.2718310^1$;

在$F(10, 8, -5, 5)$中的$fl(e)=0.27182818\times10^1$.

例1-3 已知数e的以下几个近似数,它们分别有几位有效数字?相对误差界是多少?

$x_0=2.7182$, $x_1=2.7183$, $x_2=2.7182818$

解 由于$|e-x_0|=0.0000818>\frac{1}{2}\times10^{-4}$,又$|e-x_0|\leqslant\frac{1}{2}\times10^{-3}=\frac{1}{2}\times10^{1-4}$因此,$x_0$有4位有效数字.又$\alpha_1=2$,故由式(1.7)可知相对误差界

$$\Delta x_0 \leqslant \frac{1}{4}\times10^{-3}$$

由$|e-x_1|\approx0.00002\leqslant\frac{1}{2}\times10^{-4}=\frac{1}{2}\times10^{1-5}$,因此,$x_1$有5位有效数字,又$\alpha_1=2$,故由式(1.7)可知相对误差界

$$\Delta x_1 \leqslant \frac{1}{4}\times10^{-4}$$

由$|e-x_2|\approx0.00000003\leqslant\frac{1}{2}\times10^{-7}=\frac{1}{2}\times10^{1-8}$,因此,$x_2$有8位有效数字;

又 $\alpha_1=2$，故由式(1.7)可知相对误差界

$$\Delta x_2 \leqslant \frac{1}{4}\times 10^{-7}$$

讨论：比较例 1－2 与例 1－3 可知，实数 x 进入浮点数系得到 $fl(x)$．而由浮点数系 $F(\beta, t, L, U)$ 可知，$fl(x)$ 有 t 位有效数字．因此可将两者对比，便可知近似数有几位有效数字．

例 1－4　数 $\frac{(3-\sqrt{8})^3}{(3+\sqrt{8})^3}$ 与下述各式在实数意义下是相等的，

(1) $(17-6\sqrt{8})^3$　　(2) $[(17+6\sqrt{8})^3]^{-1}$

(3) $(3-\sqrt{8})^6$　　(4) $[(3+\sqrt{8})^6]^{-1}$

(5) $19601-6920\sqrt{8}$　　(6) $(19601+6920\sqrt{8})^{-1}$

试说明在浮点数系 $F(10, 4, -8, 8)$ 中，用哪个公式计算获得的结果误差最小．

解　根据本例的数字，$\sqrt{8}$ 与 3 在此浮点数系中可认为是相近数，应避免相近数相减．因此，尽可能避免用(1)(3)(5)计算．而从减少运算次数来看，(1)(2)(3)(4)均不可取．因此，应使用(6)进行计算，可使获得的结果误差最小．

例 1－5　$f(x)=[x\mathrm{e}^{\frac{x}{2}}+\ln(1-x)]/x^3$，当 $|x|\ll 1$ 时如何计算才能获得准确的结果．

解　当 $|x|$ 很小时，$f(x)$ 中的分子是两个相近的小数相减，而分母也是一个小数字，因此应尽可能避免简单地按原顺序直接计算，而应作适当的预处理后再计算．

由 Taylor 展开式：

$$x\mathrm{e}^{\frac{x}{2}}=x+x\left(\frac{x}{2}\right)+x\frac{(x/2)^2}{2!}+x\frac{(x/2)^3}{3!}+x\frac{(x/2)^4}{4!}+\cdots$$

$$\ln(1-x)=-x-\frac{x^2}{2}-\frac{x^3}{3}-\cdots-\frac{x^n}{n}-\cdots$$

因此

$$f(x)=\left[\left(\frac{1}{8}-\frac{1}{3}\right)x^3+\left(\frac{1}{48}-\frac{1}{4}\right)x^4+\left(\frac{1}{16\times 24}-\frac{1}{5}\right)x^5+\cdots\right]\Big/x^3$$

$$\approx -\frac{5}{24}-\frac{11}{48}x-\frac{379}{1920}x^2$$

此处最后略去部分的第一项是

$$\left(\frac{1}{120\times 32}-\frac{1}{6}\right)x^3=-\frac{639}{3840}x^3$$

当 $|x|\ll 1$ 时，这一部分将是相当小的数值，可考虑略去．若要求精度很高，也可以

此为试读，需要完整PDF请访问：www.ertongbook.com

考虑将这一部分也加入计算.

例 1-6 计算 $\cos(1.473)$,若只有以 0.01 为间距的余弦的数表,试估计此余弦的准确值与数表中的最近值的误差.

解 方法一 根据公式

$$\cos(x+\delta)=\cos x\cos\delta-\sin x\sin\delta$$

由于当 $|\delta|\ll 1$ 时,$\cos\delta\approx 1$, $\sin\delta\approx\delta$, 故对本例中取

$$x=1.47,\quad \delta=0.003,$$

有 $\cos(0.003)\approx 1$, $\sin(0.003)\approx 0.003$. 又 $\sin x=\sqrt{1-\cos^2 x}$,故由

$$\cos(x+\delta)-\cos x\approx -\sin x\cdot\delta\approx -\delta\sqrt{1-\cos^2 x}$$

有

$$\begin{aligned}\cos(1.473)-\cos(1.47)&\approx -0.003\times\sqrt{1-\cos^2(1.47)}\\&\approx -0.003\times\sqrt{1-0.1006)^2}\\&\approx -0.003\times\sqrt{1-0.01012}\\&\approx -0.00298\end{aligned}$$

即,误差约为 -0.00298.

方法二 由 $f(x+\delta)=f(x)+\delta f'(\xi)$,或

$$f(x+\delta)-f(x)=\delta f'(\xi),\ \xi \text{ 在 } x \text{ 与 } x+\delta \text{ 之间}.$$

当 $|\delta|\ll 1$ 时,有

$$f(x+\delta)-f(x)\approx\delta f'(x)$$

因此

$$\begin{aligned}\cos(1.473)-\cos(1.47)&\approx -0.003\times\sin(1.47)\\&\approx -0.003\times\sqrt{1-\cos^2(1.47)}\\&\approx -0.00298\end{aligned}$$

注 此处可用的仅是"cos"函数,所以用上述方式进行计算. 其次,由比较准确的计算有 $\cos 1.473\approx 0.0976405$, $\cos 1.47\approx 0.100635$. 与本例获得的结果是相符的.

例 1-7 记计算 x 的绝对误差界为 δ,试估计计算 x^k 的绝对误差界.

解 设 x 的近似为 $\tilde{x}$,由条件得

$$\tilde{x}=x+\delta x,\quad |\delta x|\leqslant\delta$$

则

$$(\tilde{x})^k=(x+\delta x)^k=x^k+kx^{k-1}\cdot\delta x+\cdots$$

由于 δ 一般较小,故可略去 $(\delta x)^2$ 及其以后诸项,因此

$$(\tilde{x})^k-x^k\approx kx^{k-1}\cdot\delta x\approx k(\tilde{x})^{k-1}\delta x$$

因此，计算 x^k 的绝对误差界是

$$|k(\tilde{x})^{k-1}\delta|$$

例 1-8 设 $x>0$，x 的近似值的相对误差界为 Δ，试估计计算 $\ln x$ 的绝对误差界.

解 设 x 的近似为 $\tilde{x}$，由条件可知

$$\left|\frac{x-\tilde{x}}{x}\right|=\left|\frac{\delta}{x}\right|\leqslant\Delta$$

又由中值定理，有

$$\ln x-\ln\tilde{x}=\frac{1}{\xi}(x-\tilde{x}),\quad \xi\text{在}x\text{与}\tilde{x}\text{之间}$$

由 $\xi\geqslant x-|\delta x|$，因此

$$|\ln x-\ln\tilde{x}|\leqslant\frac{|x-\tilde{x}|}{x-|\delta x|}=\frac{|x-\tilde{x}|}{|x|}\cdot\frac{1}{1-|\delta x/x|}\leqslant\frac{\Delta}{1-\Delta}$$

例 1-9 已知 $\sqrt{3477}\approx58.966092$，方程 $x^2-59x+1=0$ 的一个根 $x_1\approx58.983046$，若求另一根 $x_2=(59-\sqrt{3477})/2$ 应如何计算？在浮点数系 $F(10,6,-5,5)$ 计算之.

解 由于 $59-\sqrt{3477}$ 是相近数相减，故有

$$x_2=\frac{59^2-3477}{2(59+\sqrt{3477})}=\frac{2}{59+\sqrt{3477}}$$

在 $F(10,6,-5,5)$ 中计算：

$$\begin{aligned}x_2&=0.200000\times10^1/(0.590000\times10^2+0.589661\times10^2)\\&=0.200000\times10^1/(0.117966\times10^3)\\&=0.196540\times10^{-1}\end{aligned}$$

例 1-10 已知数列 $\{x_n\}$ 收敛于 $f(x)=0$ 的根 α（$x_n>\alpha$ 或 $x_n<\alpha$，$n=1,2,\cdots$），问：计算 y_n 和 z_n 宜采用哪个算式：

(1) A：$y_n=x_n-\dfrac{(x_{n+1}-x_n)^2}{x_{n+1}-2x_n+x_{n-1}}$

B：$y_n=\dfrac{x_{n+1}x_{n-1}-x_n^2}{x_{n+1}-2x_n+x_{n-1}}$

(2) A：$z_n=x_n-\dfrac{(x_n-x_{n-1})f(x_n)}{f(x_n)-f(x_{n-1})}$

B：$z_n=x_n-\dfrac{x_n-x_{n-1}}{f(x_n)/f(x_{n-1})-1}\cdot\dfrac{f(x_n)}{f(x_{n-1})}$

解 为研究本题，先考虑数列 $\{x_n\}$ 收敛的状态，设当 n 充分大时，有

$$\frac{x_n-\alpha}{x_{n-1}-\alpha}\approx q,\quad 0<|q|<1$$

$\therefore \quad x_{n-1}-x_n\approx(1-q)(x_{n-1}-\alpha)$

若记

$$\delta_n=x_{n-1}-x_n\approx(1-q)(x_{n-1}-\alpha)$$

则

$$\begin{aligned}\delta_{n+1}&=x_n-x_{n+1}\approx(1-q)(x_n-\alpha)\\&=(1-q)[(x_{n-1}-\alpha)-(x_{n-1}-x_n)]\\&\approx\delta_n-(1-q)\delta_n\end{aligned}$$

或

$$\delta_{n+1}-\delta_n\approx-(1-q)\delta_n,\quad \delta_{n+1}\approx q\delta_n$$

这说明 $\delta_{n+1}-\delta_n$ 与 δ_n，δ_{n+1}一般属同一数量级．

(1) 由前推导，求 y_n 之(A)中的除法

$$\frac{(x_{n+1}-x_n)^2}{x_{n+1}-2x_n+x_{n-1}}=\frac{\delta_{n+1}^2}{\delta_n-\delta_{n+1}}$$

可见分母与分子相比并不是一个很小的量，而是一个较大的量；而对应的(B)的除法

$$\begin{aligned}\frac{x_{n+1}x_{n-1}-x_n^2}{x_{n+1}-2x_n+x_{n-1}}&=\frac{(x_n-\delta_{n+1})(x_n+\delta_n)-x_n^2}{\delta_n-\delta_{n+1}}\\&=\frac{(\delta_n-\delta_{n+1})x_n-\delta_n\delta_{n+1}}{\delta_n-\delta_{n+1}}\end{aligned}$$

当$\{x_n\}$的极限 $\alpha\neq0$ 时，分母与分子均为小量．与(A)相比，为避免小量作分母，宜用算法(A)．

另外，(A)与(B)中的分子，(A)是$(x_{n+1}-x_n)^2=\delta^2$，是一个小量自身相乘，而(B)中 $x_{n+1}x_{n-1}-x_n^2$ 是两个相近数：$x_{n+1}x_{n-1}$与 x_n^2 相减，从此角度看也应避免．综上所述，应选用(A)计算．

(2) 由 $x_n=x_{n-1}-\delta$，及 Taylor 展开式

$$f(x_n)=f(x_{n-1})-\delta_nf'(x_{n-1})+O(\delta_n^2)$$

便有

$$\frac{f(x_n)}{f(x_{n-1})}\approx1-\delta_n\frac{f'(x_{n-1})}{f(x_{n-1})}$$

可见，在(A)中求 z_n 的分式中的分母

$$f(x_n)-f(x_{n-1})\approx-\delta_nf'(x_{n-1})$$

而在(B)中分式的分母

$$\frac{f(x_n)}{f(x_{n-1})}-1\approx-\delta\frac{f'(x_{n-1})}{f(x_{n-1})}$$

由于$\{x_n\}$收敛于 $f(x)=0$ 的解，因此当 n 充分大时，$f(x_{n-1})$是一小量，从而$-\delta f'(x_{n-1})/f(x_{n-1})$与 $f(x_n)-f(x_{n-1})\approx-\delta f'(x_{n-1})$相比是一较大的量．因此，计

算 z_n 宜用方法(B).

例 1-11 定义内积，$(f,g)=\int_0^1 \frac{1}{\sqrt{x}}f(x)g(x)\mathrm{d}x$. 试确定此内积意义下的正交多项式 $\varphi_1(x)$，$\varphi_2(x)$.

解 按 $\varphi_0(x)=1$，由递推式(1.11)：

$$\varphi_1(x)=x-\frac{(x\varphi_0,\varphi_0)}{(\varphi_0,\varphi_0)}$$

由于

$$(\varphi_0,\varphi_0)=\int_0^1\frac{1}{\sqrt{x}}\mathrm{d}x=2\sqrt{x}\Big|_0^1=2$$

$$(x\varphi_0,\varphi_0)=\int_0^1\sqrt{x}\mathrm{d}x=\frac{2}{3}x^{\frac{3}{2}}\Big|_0^1=\frac{2}{3}$$

故

$$\varphi_1(x)=x-\frac{1}{3}$$

又

$$\varphi_2(x)=(x-\frac{(x\varphi_1,\varphi_1)}{(\varphi_1,\varphi_1)})\varphi_1(x)-\frac{(\varphi_1,\varphi_1)}{(\varphi_0,\varphi_0)}\varphi_0(x)$$

由于

$$\begin{aligned}(\varphi_1,\varphi_1)&=\int_0^1\frac{1}{\sqrt{x}}(x-\frac{1}{3})^2\mathrm{d}x\\&=\int_0^1(x^{\frac{3}{2}}-\frac{2}{3}x^{\frac{1}{2}}+\frac{1}{9}x^{-\frac{1}{2}})\mathrm{d}x\\&=\left[\frac{2}{5}x^{\frac{5}{2}}-\frac{4}{9}x^{\frac{3}{2}}+\frac{2}{9}x^{\frac{1}{2}}\right]\Big|_0^1=\frac{8}{45}\end{aligned}$$

$$\begin{aligned}(x\varphi_1,\varphi_1)&=\int_0^1\sqrt{x}(x-\frac{1}{3})^2\mathrm{d}x\\&=\int_0^1(x^{\frac{5}{2}}-\frac{2}{3}x^{\frac{3}{2}}+\frac{1}{9}x^{\frac{1}{2}})\mathrm{d}x\\&=(\frac{2}{7}x^{\frac{7}{2}}-\frac{4}{15}x^{\frac{5}{2}}+\frac{2}{27}x^{\frac{3}{2}})\Big|_0^1=\frac{88}{5\times7\times27}\end{aligned}$$

因此

$$\begin{aligned}\varphi_2(x)&=(x-\frac{11}{21})(x-\frac{1}{3})-\frac{4}{45}\\&=x^2-\frac{6}{7}x+\frac{3}{35}\end{aligned}$$

1.3 练习题

1-1 试证明浮点数系 $F(\beta, t, L, U)$中的数是有限的.

1-2 在浮点数系 $F(2, 8, -7, 8)$中,

(1) 共有多少个数;

(2) 如何表示十进制数 3.125 和 59.6;

(3) 求(2)中两数的和.

1-3 已知 $\sqrt{783}\approx 27.982$,求 $x^2-56x+1=0$ 的两个根,使之具有至少 4 位有效数字.

1-4 下列公式如何计算才比较准确:

(1) $y=\dfrac{1}{2}(e^{2x}-1)$, $|x|\ll 1$

(2) $y=\dfrac{1}{1+2x}-\dfrac{1-x}{1+x}$, $|x|\ll 1$

(3) $y=\sqrt{x+\dfrac{1}{x}}-\sqrt{x-\dfrac{1}{x}}$, $x\gg 1$

(4) $y=1-\cos 2x$, $|x|\ll 1$

1-5 计算以下函数值,只有自变量以间距为 0.01 给出的数表,试估计函数的准确值与数表中与自变量最近值的表值的误差:

(1) $\tan(2.621)$

(2) $\arctan(2.621)$

(3) $\ln(1.471)$

(4) $e^{2.653}$

第 2 章　解线性方程组的直接方法、范数

2.1　基本内容

2.1.1　解线性方程组的基本问题

解线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$，即

$$\begin{cases} \alpha_{11}x_1+\alpha_{12}x_2+\cdots+\alpha_{1n}x_n=\beta_1 \\ \alpha_{21}x_1+\alpha_{22}x_2+\cdots+\alpha_{2n}x_n=\beta_2 \\ \qquad\vdots \\ \alpha_{n1}x_1+\alpha_{n2}x_2+\cdots+\alpha_{nn}x_n=\beta_n \end{cases} \tag{2.1}$$

的直接方法是指可以通过有限步得到方程组解的方法，而且在实数意义下，得到的是原方程组的准确解．常用的直接方法是 Gauss 消去法．

2.1.2　Gauss 消去法

对方程(2.1)，不失一般性，可以设 $\alpha_{11}\neq 0$，Gauss 消去法是：对式(2.1)的第 $2\sim n$ 个方程分别减去 $l_{i1}=\alpha_{i1}/\alpha_{11}$ 倍的第 1 个方程，从而形成新方程组

$$\begin{cases} \alpha_{11}x_1+\alpha_{12}x_2+\cdots+\alpha_{1n}x_n=\beta_1 \\ \begin{cases} \alpha_{22}^{(1)}x_2+\cdots+\alpha_{2n}^{(1)}x_n=\beta_2^{(1)} \\ \qquad\vdots \\ \alpha_{n2}^{(1)}x_2+\cdots+\alpha_{nn}^{(1)}x_n=\beta_n^{(1)} \end{cases} \end{cases} \tag{2.2}$$

其中 $\tilde{\alpha}_{ij}=\alpha_{ij}-l_{i1}\alpha_{1j}$，$\tilde{\beta}_i=\beta_i-l_{i1}\beta_1$，$i,\ j=2,\ 3,\ \cdots,\ n$. 显然，现在第 $2\sim n$ 个方程形成了新的 $n-1$ 元方程组，而对此 $n-1$ 元方程组，可以重复前一步的工作，直到经过 $n-1$ 步以后，形成如下形式(上三角形)的方程组

$$\begin{cases} \alpha_{11}x_1+\alpha_{12}x_2+\cdots+\alpha_{1\,n-1}x_{n-1}+\alpha_{1n}x_n=\beta_1 \\ \qquad\alpha_{22}^{(1)}x_2+\cdots+\alpha_{2\,n-1}^{(1)}x_{n-1}+\alpha_{2n}^{(1)}x_n=\beta_2^{(1)} \\ \qquad\qquad\vdots \\ \qquad\qquad\alpha_{n-1\,n-1}^{(n-2)}x_{n-1}+\alpha_{n-1\,n}^{(n-2)}x_n=\beta_{n-1}^{(n-2)} \\ \qquad\qquad\qquad\alpha_{nn}^{(n-1)}x_n=\beta_n^{(n-1)} \end{cases} \tag{2.3}$$

此为试读，需要完整PDF请访问：www.ertongbook.com