TSINGHUA UNIVERSITY

高等数值分析

蔡大用 白峰杉

清华大学出版社

高等数值分析

蔡大用　白峰杉

清华大学出版社

(京)新登字158号

内 容 简 介

本书是为“数值分析”后续课编写的教材，要求的预备知识包括微积分、线性代数和大学本科数值分析，内容上与本科阶段的数值分析前后衔接并尽可能减少重复.本书以培养学生的科学计算能力为目标，体系内容新颖，收入了计算数学传统领域中许多新算法，也涉及到计算数学的很多新兴领域，可以作为理工科研究生学位课程“高等数值分析”的教材，也可以供有同样基础的科技人员参考.

图书在版编目(CIP)数据

高等数值分析/蔡大用，白峰杉编著.—北京：清华大学出版社，1996

ISBN 7-302-02395-6

Ⅰ.高… Ⅱ.①蔡… ②白… Ⅲ.计算方法-高等学校-教材 Ⅳ.O241

中国版本图书馆CIP数据核字(96)第24490号

出版者：清华大学出版社(北京清华大学校内，邮编100084)
印刷者：北京清华园胶印厂
发行者：新华书店总店北京科技发行所
开 本：850×1168 1/32 印张：9.625 字数：233千字
版 次：1997年4月 第1版 1997年9月 第2次印刷
书 号：ISBN 7-302-02395-6/O・176
印 数：3001～5000
定 价：12.00元

光华基金会为支持学术专著和研究生教材的出版，给予我社资助，本书即为由光华基金会资助出版的专著之一.

前　　言

计算机是对本世纪科学、工程技术和人类社会生活影响最深刻的高新技术之一.它对科学技术最深刻的影响,莫过于使科学计算平行于理论分析和实验研究,成为人类探索未知科学领域和进行大型工程设计的第三种方法和手段.在独创性工作的先行性研究中,科学计算更具有突出的作用.科学计算能力是跨世纪人才不可或缺的.然而它并不是计算机本身的自然产物,而是数学与计算机有机结合的结果,它的核心内容是以现代化的计算机及数学软件为工具,以数学模型为基础进行模拟研究.近年来,它同时也成为数学科学本身发展的源泉和途径之一.高等教育中如何培养学生科学计算的能力正日益受到关注,已成为当前教育改革的核心和焦点之一.

数值分析及有关的数学基础教学环节,在培养学生科学计算能力上具有不可替代的作用.目前国内许多高等学校已将数值分析列入自然科学、工程技术乃至社会科学本科的教学计划之中.但仅靠本科的数值分析课尚不足以培养高层次人才的科学计算能力.除数学基础教育的体系内容亟待改革外,建设好研究生学位课档次的提高课程也是重要的一环.

基于上述认识,在笔者为清华大学研究生讲授"高等数值分析"课程的基础上完成此书.课程的对象是非计算数学专业的博士生和部分硕士生.假设他们学习过微积分、线性代数和大学阶段的数值分析,并具有一定的计算机的编程能力和经验.本书也可供有

同样基础的科技人员参考.

全书共 9 章,分成四个部分. 另有一个附录.

第 1、4 两章分别介绍矩阵论和近代分析的基本理论,它们不仅提供其它各章所必需的预备知识,很多结果也具有其独立意义.它可以作为大学数学基础知识的补充、延伸和深化,也是学习其它各章的基础.

第 2、3 两章是数值线性代数部分. 在 Ritz 原理和 Galerkin 原理的理论格局下介绍了近代流行的实用算法. 其中有些算法在理论上还不成熟.

第 5、6、7 章集中讨论非线性问题的理论和数值算法. 非线性问题是一个十分广泛的领域,我们在这部分集中介绍了非线性方程组的 Newton 法及其变形、非线性特征值问题、分岔问题以及离散动力系统的有关课题.

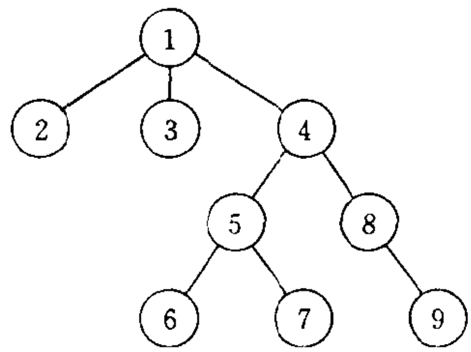
第 8、9 两章介绍有关常微分方程的初值及边值问题的算法.第 9 章里的古典变分原理和有限单元方法也为进一步学习偏微分方程数值解法做了准备.

我们历来主张学习数值分析时,主要精力应放在算法分析和算法设计上. 在应用时要注意针对具体问题筛选和比较算法,尽可能避免低水平上重复已有算法的计算机编码,建议大家尽量使用数学软件工具.

本书内容各部分有相对的独立性,各章间的关系可以用如下一棵树来表示. 教师可以据此灵活安排教学,读者可据此自学部分感兴趣的内容.

每章后面都附有一定数量的习题,目的是帮助学生巩固课堂内容. 其中有些是我们自己编撰的,肯定不尽完美.

在每章的最后均写有一段评注. 我们并没有企图在评注一栏



里完整地回顾某一专题的发展史，也没有试图完整地探讨学科的未来发展趋势. 评注中的内容均是结合笔者自己的经验发表的一管之见，难免有失偏颇之处，其目的是开阔学生的视野，也为有兴趣在某些方面进行进一步研究的学生提供一些线索. 教学中我们鼓励有余力的学生开展课题研究，收到了很好的效果.

读者会发现本书收入了计算数学传统领域中许多新算法，也涉及到计算数学的很多新兴领域. 然而标新并不是本书的出发点，内容的选择是为学生能力的培养服务的. 本书在选材上曾几经反复，在原有试用讲义的基础上，本次成书做了较大修改，力图更好体现这样的意图.

在完成本书过程中，得到我校研究生院和应用数学系很多同事的鼓励、帮助与支持. 在两年教学过程中，我们的学生对本书的试用稿也提出很多意见和建议. 在此一并向他们致谢. 最后作者对本书编辑的出色工作表示诚挚的谢意。

作　者

1996 年于清华园

目　录

矩阵论中的若干问题

矩阵是科学计算的基本工具之一.本章介绍矩阵论中若干重要结果.它们不但在后续章节中要多次用到,而且很多结果都有其独立的意义.现已有不少专著介绍这方面的内容.读者可以在文献9,11,24,28,71,72 中找到有关更详细的论述.

1.1 预备知识

本节中列出若干线性代数的基本结果,而不加证明.它们的证明可以在任何一本有关线性代数的教科书中找到.

我们用 $\mathbf{R}^n(\mathbf{C}^n)$ 代表 n 维实(复)向量空间. $\mathbf{R}^{n\times n}(\mathbf{C}^{n\times n})$ 代表 n 阶实(复)矩阵的全体. $\boldsymbol{A}^{\mathrm{T}}$ 代表矩阵的转置.而 $\boldsymbol{A}^*$ 代表矩阵经转置后,取诸元素的共轭数后得到的矩阵.

定理 1.1.1 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$,而且 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$.

则 1) $\boldsymbol{A}$ 的所有特征值都是实的.

2) 存在有正交矩阵 $\boldsymbol{Q}$,满足 $\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{I}$,使得 $\boldsymbol{A}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{Q}$,其中 $\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_i)$. $\lambda_i(i=1,2,3,\cdots,n)$ 为 $\boldsymbol{A}$ 的特征值.

定理 1.1.2 设 $\boldsymbol{A}$ 是一个 $n\times n$ 的实矩阵 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,如果满足下面三个互相等价的条件之一,则称矩阵 $\boldsymbol{A}$ 是**对称正定的**.

1) 对于任一不为零的向量 $\boldsymbol{x}\in\mathbf{R}^n$ 恒有: $(\boldsymbol{x},\boldsymbol{A}\boldsymbol{x})>0$,其中 $(\cdot,\cdot)$ 代表 $\mathbf{R}^n$ 上的内积.

2）A 的所有特征值都是正的.

3）由 A 的任意 k 行（$k=1,2,\cdots,n$）和相应 k 列交点处的元素构成的 $k\times k$ 子矩阵的行列式都大于零.

从定理 1.1.2 中的条件 1）不难得出下面推论.

推论 1.1.1 若 A 为一对称正定矩阵，对任意非奇异矩阵 W，$B=W^{T}AW$ 也是对称正定矩阵.

对于 $n\times n$ 的复矩阵 A，如果满足 $A^{*}=A$，则称 A 为一**厄米尔特矩阵**. 如 $A^{*}A=I$，则称 A 为一**酉矩阵**.

对于复矩阵完全可以平行地叙述类似于定理 1.1.1—推论 1.1.1 的结果，只需要把转置矩阵和正交矩阵分别换成厄米尔特矩阵和酉矩阵即可.

内积（或**点积**）也是矩阵运算中不可缺少的.

在 $\mathbf{R}^{n}$ 空间中两个向量 x,y 的内积定义为：

$$(x,y)=\sum_{i=1}^{n}x_iy_i. \tag{1.1.1}$$

而在 $\mathbf{C}^{n}$ 中内积的定义为：

$$(x,y)=\sum_{i=1}^{n}x_i\bar{y}_i. \tag{1.1.2}$$

关于内积的一个常用的著名不等式为 **Hölder 不等式**.

定理 1.1.3 设 x,y，为 $\mathbf{C}^{n}$ 中任意两个向量. p,q 为两个正实数，满足 $1/p+1/q=1$，则有

$$|(x,y)|\leqslant(x,x)^{1/p}\cdot(y,y)^{1/q}. \tag{1.1.3}$$

当 $p=q=2$ 时，(1.1.3)称为 **Schwartz 不等式**

$$|(x,y)|\leqslant\sqrt{(x,x)}\cdot\sqrt{(y,y)}. \tag{1.1.4}$$

关于内积有一个很重要的性质.

定理 1.1.4 设 x,y 为 $\mathbf{R}^{n}(\mathbf{C}^{n})$ 中的任意两个向量. Q 为一正

交矩阵(酉矩阵),则

$$(\boldsymbol{Qx},\boldsymbol{Qy}) = (\boldsymbol{x},\boldsymbol{y}). \qquad (1.1.5)$$

最后,我们列出两个关于一般 $n\times n$ 复矩阵的定理.

定理 1.1.5 (Jordan 定理) 设 $\boldsymbol{A}$ 为一 $n\times n$ 的复矩阵. 则存在非奇异矩阵 $\boldsymbol{W}$,使得

$$\boldsymbol{A} = \boldsymbol{W}^{-1}\boldsymbol{JW}, \qquad (1.1.6)$$

其中

$$\boldsymbol{J} = \begin{bmatrix} J_1 & & & \\ & J_2 & & 0 \\ & 0 & \ddots & \\ & & & J_s \end{bmatrix}, \qquad (1.1.7)$$

这里的 $J_i(i=1,2,\cdots,s)$都有如下特殊形状

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & 0 \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}, \qquad (1.1.8)$$

J_i 称为 A 的 **Jordan 块**.

定理 1.1.6 (Gerischgorin 定理) 设 $\boldsymbol{A}$ 为一 $n\times n$ 的复矩阵,$\lambda_i(i=1,2,\cdots,n)$为其特征值,

$$\Lambda_i = \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{i,j}|, \qquad (1.1.9)$$

在复平面上定义:

$$D_i = |z - a_{ii}| \leqslant \Lambda_i \qquad i = 1,2,\cdots,n$$

为矩阵 A 的 **Gerischgorin 圆盘**. 则对任一 λ_i 必有:

$$\lambda_i \in \bigcup_{i=1}^{n} D_i. \tag{1.1.10}$$

1.2 矩阵的分解

在算术中把一个整数分解为质因数的连乘积，它在其后的诸多应用中都起很重要的作用. 类似情形也存在于矩阵论中. 把一个给定的矩阵分解成几个"较简单的"矩阵连乘积形式，无论在理论证明和科学计算中都是十分重要的. 本节中介绍几个有关矩阵分解的重要结果.

如不加特别说明，以下矩阵都是实矩阵.

定理 1.2.1 (Cholesky 分解) 设 $\boldsymbol{A}$ 为一个 $n\times n$ 的对称正定矩阵，则存在对角线为正的下三角矩阵 $\boldsymbol{L}$，使得 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}$.

证明 证明方法是对矩阵阶数使用归纳法.

当 $n=1$ 时命题显然正确.

设 $n=k$ 时命题正确.

当 $n=k+1$ 时，我们把 $\boldsymbol{A}$ 表示为：

$$\boldsymbol{A}=\begin{bmatrix} \boldsymbol{A}_k & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_{k+1,k+1} \end{bmatrix}, \tag{1.2.1}$$

其中 $\boldsymbol{A}_k$ 为一 $k\times k$ 的矩阵. 由正定矩阵的性质(定理 1.1.2)可知 $\boldsymbol{A}_k$ 也是一个正定矩阵. 由归纳法假设可有 $\boldsymbol{A}_k=\boldsymbol{L}_k\boldsymbol{L}_k^{\mathrm{T}}$，其中 $\boldsymbol{L}_k$ 为一对角元素为正的下三角矩阵.

令

$$\boldsymbol{L}_{k+1}=\begin{bmatrix} \boldsymbol{L}_k & 0 \\ \boldsymbol{l}^{\mathrm{T}} & \lambda \end{bmatrix},$$

$\boldsymbol{l}$ 和 $\boldsymbol{\lambda}$ 分别是待定的向量和数量. 并令其满足

$$
\begin{aligned}
\boldsymbol{A} &= \boldsymbol{L}_{k+1}\boldsymbol{L}_{k+1}^{\mathrm{T}} \\
&= \begin{bmatrix} \boldsymbol{L}_k & 0 \\ \boldsymbol{l}^{\mathrm{T}} & \lambda \end{bmatrix} \begin{bmatrix} \boldsymbol{L}_k^{\mathrm{T}} & \boldsymbol{l} \\ 0 & \lambda \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{L}_k\boldsymbol{L}_k^{\mathrm{T}} & \boldsymbol{L}_k\boldsymbol{l} \\ \boldsymbol{l}^{\mathrm{T}}\boldsymbol{L}_k^{\mathrm{T}} & \boldsymbol{l}^{\mathrm{T}}\boldsymbol{l} + \lambda^2 \end{bmatrix}.
\end{aligned} \tag{1.2.2}
$$

将(1.2.1)和(1.2.2)进行比较，得到

$$\boldsymbol{A}_k = \boldsymbol{L}_k\boldsymbol{L}_k^{\mathrm{T}}, \tag{1.2.3}$$

$$\boldsymbol{L}_k\boldsymbol{l} = \boldsymbol{\alpha}, \tag{1.2.4}$$

$$\boldsymbol{l}^{\mathrm{T}}\boldsymbol{l} + \lambda^2 = a_{k+1,k+1}. \tag{1.2.5}$$

由于 $\boldsymbol{L}_k$ 的取法，(1.2.3)式显然满足. 由(1.2.4)可以唯一地确定 $\boldsymbol{l}$. 由(1.2.5)得到 $\lambda^2 = a_{k+1,k+1} - \boldsymbol{l}^{\mathrm{T}}\boldsymbol{l}$. 只要能证明 $a_{k+1,k+1} - \boldsymbol{l}^{\mathrm{T}}\boldsymbol{l} > 0$，则 λ 也可以通过开平方得到.

我们取

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{I}_k & 0 \\ -\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_k^{-1} & 1 \end{bmatrix},$$

其中 $\boldsymbol{I}_k$ 为 k 阶单位矩阵，不难算出，

$$\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{A}_k & 0 \\ 0 & a_{k+1,k+1} - \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_k^{-1}\boldsymbol{\alpha} \end{bmatrix},$$

再由(1.2.3)可知

$$
\begin{aligned}
a_{k+1,k+1} - \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_k^{-1}\boldsymbol{\alpha} &= a_{k+1,k+1} - \boldsymbol{\alpha}^{\mathrm{T}}(\boldsymbol{L}_k\boldsymbol{L}_k^{\mathrm{T}})^{-1}\boldsymbol{\alpha} \\
&= a_{k+1,k+1} - \boldsymbol{l}^{\mathrm{T}}\boldsymbol{l} = \lambda^2.
\end{aligned} \tag{1.2.6}
$$

由 $\boldsymbol{A}$ 的正定性，显然得到 $\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^{\mathrm{T}}$ 也是正定阵(定理 1.1.1 的推论). 因此，$\lambda^2 = a_{k+1,k+1} - \boldsymbol{l}^{\mathrm{T}}\boldsymbol{l} > 0$. 从而定理得证. ■

从上面的证明不难看出，如果完成了一个 $n\times n$ 阶矩阵的一个 k 阶顺序主子矩阵 $\boldsymbol{A}_k$ 的 Cholesky 分解，$\boldsymbol{A}_k = \boldsymbol{L}_k\boldsymbol{L}_k^{\mathrm{T}}$. 只要求解一个

形如(1.2.4)的方程组. 这里的 α 是由矩阵 A 的第 $k+1$ 列前面 k 个元素构成的向量，这个方程组的系数矩阵是下三角形的. 因此，很容易求解. 得到解 l 之后，再算出

$$\lambda = \sqrt{a_{k+1,k+1} - l^{\mathrm{T}} l},$$

则得到 A 的 $k+1$ 阶顺序主子式的 Cholesky 分解

$$A_{k+1} = L_{k+1} L_{k+1}^{\mathrm{T}} = \begin{bmatrix} L_k & 0 \\ l^{\mathrm{T}} & \lambda \end{bmatrix} \begin{bmatrix} L_k^{\mathrm{T}} & l \\ 0 & \lambda \end{bmatrix}.$$

按照上面的步骤取 $k=1,2,\cdots,n-1$，就可以递推地完成矩阵 A 的分解.

为了介绍另一种重要的分解，给出以下定义.

定义 1.2.1 设 u、v 是 $\mathbf{C}^n$ 中的两个列向量. σ 为一复数.

$$E(u,v,\sigma) = I - \sigma u v^*$$

称为一个**初等矩阵**，其中 I 为 $n\times n$ 的单位矩阵.

初等矩阵有两个简单性质：

1)
$$\begin{aligned} & E(u,v,\sigma)\cdot E(u,v,\tau) \\ & = E(u,v,\sigma+\tau-\sigma\tau v^* u). \end{aligned} \tag{1.2.7}$$

上式表明两个相同向量形成的初等矩阵之积仍然是个初等矩阵. 如果 $\sigma+\tau-\sigma\tau v^* u=0$，即，$\sigma^{-1}+\tau^{-1}=v^* u$，则由(1.2.7)式可得到：

$$E(u,v,\sigma) = E^{-1}(u,v,\tau). \tag{1.2.8}$$

2)
$$\det(u,v,\sigma) = 1 - \sigma v^* u. \tag{1.2.9}$$

(1.2.7)和(1.2.9)式很容易直接验证，留给读者做为练习(见习题 3).

在计算数学中很多常用的矩阵都可以用初等矩阵(或它们的乘积)表示出来.