21世纪高等院校教材

信息与计算科学专业教材系列

数值代数

（第二版）

张凯院　徐　仲　编著

科学出版社
www.sciencep.com

21世纪高等院校教材

信息与计算科学专业教材系列

数值代数

（第二版）

张凯院　徐　仲　编著

科学出版社

北　京

内 容 简 介

本书分为六章，内容包括矩阵论基础、线性方程组的迭代解法、带状线性方程组的直接解法、特殊线性方程组的递推解法、矩阵特征值问题的解法及线性矩阵方程的迭代解法。各章后均配有适量的习题，书后还附有习题答案与提示。

本书内容新颖，叙述严谨，表达流畅，可作为高等院校数学专业高年级本科生教材，也可供有关专业的研究生和从事科学计算的工程技术人员参考。

图书在版编目(CIP)数据

数值代数/张凯院，徐仲编著．—2版．—北京：科学出版社，2006
(21世纪高等院校教材·信息与计算科学专业教材系列)
ISBN 7-03-017746-0

Ⅰ.数…　Ⅱ.①张…　②徐…　Ⅲ.线性代数计算法-高等学校-教材
Ⅳ.O241.6

中国版本图书馆CIP数据核字(2006)第086194号

责任编辑：姚莉丽　于宏丽 / 责任校对：张　琪
责任印制：张克忠 / 封面设计：耕者设计工作室

科学出版社 出版
北京东黄城根北街16号
邮政编码：100717
http://www.sciencep.com

双青印刷厂 印刷

科学出版社发行　各地新华书店经销

*

2000年8月西北工业大学出版社第一版
2006年8月第　二　版　开本：B5(720×1000)
2006年8月第一次印刷　印张：13 1/2
印数：1—3 500　字数：249 000

定价：18.00元

第二版前言

本书第一版自2000年出版以来，经过5年的教学使用，在吸取使用本书的同行和读者提出的宝贵意见的基础上，结合我们近期的一些研究结果，将其部分内容作了修订、充实和完善。

本次修订的主要内容是，在第3章中增加了求解系数矩阵为三对角矩阵、周期三对角矩阵、块三对角矩阵、周期块三对角矩阵的线性代数方程组的变参数追赶法、变形追赶法、线性插值法和参数法，并将第一版中的 Hankel 方程组、Toeplitz方程组、Loewner 方程组和范德蒙德方程组的递推算法单独列为第4章。

增加和修订的具体内容如下：

(1) 增加了解块三对角方程组的二次 PE 方法（2.6.2节）；

(2) 增加了解三对角方程组的追赶法的数值稳定性分析（3.1.1节）；

(3) 增加了解三对角方程组的变参数追赶法（3.1.2节）；

(4) 增加了解周期三对角方程组的追赶法和变形及变参数追赶法（3.2节）；

(5) 增加了解块三对角方程组的追赶法和线性插值法及双参数法（3.4节）；

(6) 增加了解周期块三对角方程组的追赶法和线性插值法及三参数法（3.5节）；

(7) 修订了矩阵特征值问题的幂方法（5.1节）；

(8) 增加了解矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CXD}=\boldsymbol{F}$ 的参数迭代解法（6.3.4节）；

(9) 增加了解矩阵方程 $\boldsymbol{AX}+\boldsymbol{XB}=\boldsymbol{F}$ 的分组迭代解法（6.3.5节）；

(10) 修订了个别定理的叙述和证明；

(11) 增加了习题答案与提示。

增加和修订上述内容之后，讲授全书约需60学时。

本书第二版分为六章，其中第1、2、5、6章和第3章的3.3～3.5节由张凯院编写，第3章的3.1节与3.2节和第4章由徐仲编写，由张凯院对全书统稿。

编者对关心本书和对本书第一版提出宝贵意见的同行及读者表示衷心的感谢。

编　者

2005年10月于西北工业大学

第一版前言

数值代数是研究代数问题的数值计算方法及其有关理论的一门学科，它既涉及数学理论方面的研究，也涉及工程设计方面的应用。可以说相当一部分计算数学问题最终都要归结为线性代数方程（组）的求解问题，或者是矩阵的特征值与特征向量的计算问题。随着计算机技术的快速发展，能够进行数值计算的实际问题的维数不断增高，相应的数值计算方法也在不断地改进或更新。因此，对于需要从事科学计算的大学生、研究生以及工程技术人员来说，系统地了解和掌握数值代数的基本理论和方法，特别是近代发展起来的比较成熟的新型算法是至关重要的。

本书是在同名讲义的基础上改写而成的。内容包括矩阵论基础、线性方程组的迭代解法、带状线性方程组的直接解法、特殊线性方程组的递推解法、矩阵特征值问题的解法及线性矩阵方程的迭代解法等。

在编写内容上，我们使读者在熟悉“线性代数”和“计算方法”知识的基础上，运用矩阵理论研究算法的收敛性和计算机实现问题。收编了关于线性方程组基于 Galerkin 原理的迭代解法，块三对角方程组的 PE 解法，特殊方程组的快速算法，Lyapunov 矩阵方程的迭代解法，以及一般的线性矩阵方程的迭代-校正解法等内容。这些内容能够为数学专业的高年级学生指出一些研究方向，使他们顺利地进行毕业论文工作，同时收编的诸多算法在科学计算方面都有着广泛的应用。

本书突出了矩阵分析方法，力求结构紧凑、叙述简明。在处理书中有关内容时，我们着重介绍经典算法的相互关系和深化问题；着重介绍近代算法的基本原理和计算机实现问题。讲授全书约需 50 学时。

全书分为 5 章，第 1、2、4、5 章由张凯院编写，第 3 章由徐仲编写，统稿工作由张凯院负责。本书在编写过程中得到了西北工业大学应用数学系领导及同事们的关心和帮助；西北工业大学罗学波、叶正麟教授审阅了本书初稿并给予了肯定和鼓励；西北大学盛德成教授仔细审阅了书稿并提出了许多宝贵意见；西北工业大学教务处、出版社对本书的出版给予了支持，在此一并表示衷心的感谢。

由于作者水平所限，书中疏漏和不妥之处，恳请读者指正。

编　者

2000 年 3 月于西北工业大学

符号说明

$\mathbf{R}(\mathbf{C})$	实（复）数集合
$\mathbf{R}^n(\mathbf{C}^n)$	实（复）n 维向量集合
$\mathbf{R}^{m\times n}(\mathbf{C}^{m\times n})$	实（复）$m\times n$ 矩阵集合
$\mathbf{R}_r^{m\times n}(\mathbf{C}_r^{m\times n})$	秩为 r 的实（复）$m\times n$ 矩阵集合
$W^\perp$	子空间 W 的正交补
$\dim V$	线性空间 V 的维数
$\mathrm{span}\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m\}$	由元素 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_m$ 生成的子空间
$\mathbf{0}$	零向量或线性空间的零元素
$\boldsymbol{e}_i$	第 i 个分量为 1，其余分量为 0 的向量
$\boldsymbol{O}$	零矩阵
$\boldsymbol{I}$	单位矩阵
$\boldsymbol{E}_{ij}$	第 i 行 j 列元素为 1，其余元素为 0 的矩阵
$\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$	以 λ_1，λ_2，…，λ_n 为对角元素的对角矩阵
$\det\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的行列式
$\mathrm{tr}\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的迹
$\rho(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的谱半径
$\mathrm{adj}\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的伴随矩阵
$\mathrm{rank}\boldsymbol{A}$	矩阵 $\boldsymbol{A}$ 的秩
$R(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的值域
$N(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的零空间
$\mathrm{vec}(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 按列拉直的列向量
$\overline{\mathrm{vec}}(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 按行拉直的列向量
$\boldsymbol{A}^{\mathrm{T}}$	矩阵 $\boldsymbol{A}$ 的转置
$\boldsymbol{A}^{\mathrm{H}}$	矩阵 $\boldsymbol{A}$ 的共轭转置
$\boldsymbol{A}^{+}$	矩阵 $\boldsymbol{A}$ 的 Moore-Penrose 逆
$\boldsymbol{A}^{(1)}$	矩阵 $\boldsymbol{A}$ 的 $\{1\}$-逆
$A\{1\}$	矩阵 $\boldsymbol{A}$ 的 $\{1\}$-逆的集合
$\boldsymbol{A}\otimes\boldsymbol{B}$	矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的直积
$[\boldsymbol{x},\ \boldsymbol{y}]$	元素 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积
$\boldsymbol{x}\perp\boldsymbol{y}$	元素 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交

$\|\boldsymbol{x}\|_A$	向量 $\boldsymbol{x}$ 的椭圆范数
$\|\boldsymbol{x}\|_p$	向量 $\boldsymbol{x}$ 的 p-范数
$\|\boldsymbol{A}\|_F$	矩阵 $\boldsymbol{A}$ 的 Frobenius 范数
$V_1 \cap V_2$	集合 V_1 与 V_2 的交
$V_1 \cup V_2$	集合 V_1 与 V_2 的并
V_1+V_2	子空间 V_1 与 V_2 的和
$V_1 \oplus V_2$	子空间 V_1 与 V_2 的直和
$\partial f(\lambda)$	多项式 $f(\lambda)$ 的次数
$f(\lambda)\|g(\lambda)$	多项式 $f(\lambda)$ 整除多项式 $g(\lambda)$

目　　录

第二版前言

第一版前言

符号说明

第 1 章　矩阵论基础 ······ 1

1.1　矩阵的三角相似与对角相似 ······ 1

1.2　矩阵的 QR 分解 ······ 6

1.3　矩阵的满秩分解 ······ 15

1.4　矩阵的奇异值分解 ······ 17

1.5　矩阵的广义逆及其应用 ······ 20

1.6　矩阵的特征值估计与隔离 ······ 27

习题 1 ······ 34

第 2 章　线性方程组的迭代解法 ······ 36

2.1　古典迭代方法 ······ 36

2.2　基于变分原理的迭代方法 ······ 45

2.2.1　最速下降法 ······ 47

2.2.2　共轭梯度法 ······ 48

2.3　基于 Galerkin 原理的迭代方法 ······ 51

2.3.1　Galerkin 原理 ······ 51

2.3.2　Arnoldi 算法 ······ 52

2.3.3　GMRES 算法 ······ 55

2.4　行作用方法 ······ 58

2.5　迭代-校正加速方法 ······ 63

2.5.1　整体校正方法 ······ 63

2.5.2　基于矩阵特征值的外推方法 ······ 67

2.5.3　基于 Chebyshev 多项式的最小零偏差方法 ······ 68

2.6　块三对角方程组的迭代解法 ······ 70

2.6.1　PE 方法 ······ 70

2.6.2　二次 PE 方法 ······ 75

习题 2 ······ 77

第 3 章　带状线性方程组的直接解法 ······ 79

3.1　三对角方程组 ······ 79

3.1.1 追赶法 …… 79
3.1.2 变参数追赶法 …… 84
3.1.3 线性插值法 …… 87
3.1.4 双参数法 …… 88
3.2 周期三对角方程组 …… 90
3.2.1 追赶法 …… 90
3.2.2 变形追赶法 …… 91
3.2.3 变参数追赶法 …… 94
3.3 Hessenberg 方程组 …… 97
3.3.1 线性插值法 …… 98
3.3.2 双参数法 …… 98
3.3.3 Givens 变换法 …… 99
3.4 块三对角方程组 …… 100
3.4.1 追赶法 …… 100
3.4.2 线性插值法 …… 101
3.4.3 双参数法 …… 103
3.5 周期块三对角方程组 …… 105
3.5.1 追赶法 …… 105
3.5.2 线性插值法 …… 107
3.5.3 三参数法 …… 110
习题 3 …… 112
第 4 章 特殊线性方程组的递推解法 …… 113
4.1 Hankel 方程组 …… 113
4.2 Toeplitz 方程组 …… 116
4.3 Loewner 方程组 …… 119
4.4 范德蒙德方程组 …… 124
4.4.1 线性方程组 $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{a}=\boldsymbol{f}$ 的递推算法 …… 125
4.4.2 线性方程组 $\boldsymbol{V}\boldsymbol{y}=\boldsymbol{b}$ 的递推算法 …… 126
习题 4 …… 129
第 5 章 矩阵特征值问题的解法 …… 130
5.1 幂方法 …… 130
5.1.1 乘幂法 …… 130
5.1.2 逆幂法 …… 131
5.2 Krylov 方法 …… 132
5.2.1 矩阵多项式 …… 132
5.2.2 向量相对于矩阵的零化多项式 …… 133
5.2.3 向量相对于矩阵的零化多项式计算 …… 134

5.2.4 矩阵的最小多项式计算 …… 135
5.2.5 矩阵的特征向量计算 …… 137
5.3 Lanczos 方法 …… 139
5.3.1 Lanczos 正交化过程 …… 139
5.3.2 向量相对于矩阵的零化多项式计算 …… 140
5.3.3 矩阵的最小多项式计算 …… 141
5.3.4 实对称矩阵的 L1 算法 …… 143
5.3.5 非对称矩阵的 L2 算法 …… 144
5.4 Frame 方法 …… 147
5.5 Samuelson 方法 …… 149
习题 5 …… 153
第 6 章　线性矩阵方程的迭代解法 …… 154
6.1 线性矩阵方程解的存在性 …… 154
6.1.1 矩阵的直积 …… 154
6.1.2 线性矩阵方程解的存在性 …… 157
6.2 计算逆矩阵的迭代方法 …… 160
6.2.1 古典迭代方法 …… 160
6.2.2 Newton 迭代方法 …… 162
6.3 Lyapunov 矩阵方程的迭代解法 …… 162
6.3.1 矩阵方程 $AX+XA^H=F$ 的参数迭代解法 …… 163
6.3.2 矩阵方程 $AX+XB=F$ 的参数迭代解法 …… 166
6.3.3 矩阵方程 $AXB^T+BXA^T=F$ 的参数迭代解法 …… 171
6.3.4 矩阵方程 $AXB+CXD=F$ 的参数迭代解法 …… 176
6.3.5 矩阵方程 $AX+XB=F$ 的分组迭代解法 …… 182
6.4 线性矩阵方程的迭代-校正解法 …… 189
6.4.1 整体校正过程 …… 190
6.4.2 单列校正过程 …… 192
6.4.3 迭代-校正方法 …… 194
习题 6 …… 197
参考文献 …… 198
习题答案与提示 …… 199

第1章　矩阵论基础

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论，这些结果不仅在后续各章中多次用到，而且很多结果都有其独立的意义.

本书用 $\mathbf{R}^n$（$\mathbf{C}^n$）表示 n 维实（复）向量空间，$\mathbf{R}^{m\times n}$（$\mathbf{C}^{m\times n}$）表示 $m\times n$ 实（复）矩阵空间.

1.1　矩阵的三角相似与对角相似

为了介绍矩阵论中极为重要的 Hamilton-Cayley 定理，先来建立任意 n 阶矩阵与三角矩阵相似的理论.

定理 1.1　任意 n 阶矩阵与上三角矩阵相似.

证　设 $\boldsymbol{A}=(a_{ij})_{n\times n}$，采用数学归纳法来证明.

当 $n=1$ 时，定理显然成立. 假设 $n=k-1$ 时定理成立，下面证明 $n=k$ 时定理也成立.

设 $\boldsymbol{A}=(a_{ij})_{k\times k}$ 的 k 个特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_k$，且对应于特征值 λ_1 的一个特征向量为 $\boldsymbol{x}_1\in\mathbf{C}^k$，扩充向量 $\boldsymbol{x}_1$ 为向量空间 $\mathbf{C}^k$ 的基，并记增加的向量为 $\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k$. 令 $\boldsymbol{P}_1=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k]$，则 $\boldsymbol{P}_1$ 可逆，且有

$$\boldsymbol{AP}_1=[\boldsymbol{Ax}_1\quad \boldsymbol{Ax}_2\quad\cdots\quad\boldsymbol{Ax}_k]=[\lambda_1\boldsymbol{x}_1\quad\boldsymbol{Ax}_2\quad\cdots\quad\boldsymbol{Ax}_k].$$

由于 $\boldsymbol{Ax}_j\in\mathbf{C}^k$（$j=2,\cdots,n$），所以 $\boldsymbol{Ax}_j$ 可由 $\mathbf{C}^k$ 的基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k$ 线性表示，即

$$\boldsymbol{Ax}_j=b_{1j}\boldsymbol{x}_1+b_{2j}\boldsymbol{x}_2+\cdots+b_{kj}\boldsymbol{x}_k\qquad(j=2,\cdots,k).$$

于是

$$\boldsymbol{AP}_1=[\lambda_1\boldsymbol{x}_1\ \boldsymbol{Ax}_2\ \cdots\ \boldsymbol{Ax}_k]=[\boldsymbol{x}_1\ \boldsymbol{x}_2\ \cdots\ \boldsymbol{x}_k]\begin{bmatrix}\lambda_1 & b_{12} & \cdots & b_{1k}\\ 0 & b_{22} & \cdots & b_{2k}\\ \vdots & \vdots & & \vdots\\ 0 & b_{k2} & \cdots & b_{kk}\end{bmatrix}.$$

记 $\boldsymbol{A}_1$ 为上式右端矩阵右下角的 $k-1$ 阶子矩阵，则上式可写为

$$\boldsymbol{P}_1^{-1}\boldsymbol{AP}_1=\left[\begin{array}{c:ccc}\lambda_1 & b_{12} & \cdots & b_{1k}\\ \hdashline 0 & & & \\ \vdots & & \boldsymbol{A}_1 & \\ 0 & & & \end{array}\right].$$

易见，A_1 的 $k-1$ 个特征值为 $\lambda_2,\cdots,\lambda_k$. 对于 $k-1$ 阶矩阵 A_1，由归纳法假设，有 $k-1$ 阶可逆矩阵 Q，使得

$$Q^{-1}A_1Q=\begin{bmatrix}\lambda_2 & * & \cdots & * \\ & \lambda_3 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_k\end{bmatrix}.$$

令

$$P_2=\left[\begin{array}{c:ccc}1 & 0 & \cdots & 0 \\ \hdashline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right],\qquad P=P_1P_2,$$

则有

$$\begin{aligned}P^{-1}AP&=P_2^{-1}(P_1^{-1}AP_1)P_2\\&=\left[\begin{array}{c:ccc}1 & 0 & \cdots & 0 \\ \hdashline 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{array}\right]\left[\begin{array}{c:ccc}\lambda_1 & b_{12} & \cdots & b_{1k} \\ \hdashline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right]\left[\begin{array}{c:ccc}1 & 0 & \cdots & 0 \\ \hdashline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right]\\&=\left[\begin{array}{c:ccc}\lambda_1 & * & \cdots & * \\ \hdashline 0 & & & \\ \vdots & & Q^{-1}A_1Q & \\ 0 & & & \end{array}\right]=\begin{bmatrix}\lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_k\end{bmatrix}.\end{aligned}$$

根据归纳法原理，对于任意正整数 n，定理成立.

证毕.

例 1.1　已知 $A=\begin{bmatrix}2 & 1 & 1\\-2 & 5 & 1\\-3 & 2 & 5\end{bmatrix}$，求可逆矩阵 P，使 $P^{-1}AP$ 为上三角矩阵.

解　求出 A 的三个特征值为 $\lambda_1=\lambda_2=\lambda_3=4$，对应于 λ_1 的一个特征向量为 $x_1=(1,1,1)^{\mathrm{T}}$，以 x_1 作为 P_1 的第一列，则有

$$P_1=\begin{bmatrix}1 & 0 & 0\\1 & 1 & 0\\1 & 1 & 1\end{bmatrix},\quad P_1^{-1}=\begin{bmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & -1 & 1\end{bmatrix},\quad P_1^{-1}AP_1=\begin{bmatrix}4 & 2 & 1\\0 & 4 & 0\\0 & 1 & 4\end{bmatrix}.$$

记 $A_1=\begin{bmatrix}4 & 0\\1 & 4\end{bmatrix}$，它的两个特征值为 $\lambda_2=\lambda_3=4$，求出对应于 λ_2 的一个特征向量为 $y_1=(0,1)^{\mathrm{T}}$，以 y_1 作为 Q 的第一列，则有

$$Q=\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix},\quad Q^{-1}=\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix},\quad Q^{-1}A_1Q=\begin{bmatrix}4 & 1\\0 & 4\end{bmatrix}.$$

令

$$\boldsymbol{P}_2=\begin{bmatrix}1 & \\ & \boldsymbol{Q}\end{bmatrix},\qquad \boldsymbol{P}=\boldsymbol{P}_1\boldsymbol{P}_2=\begin{bmatrix}1&0&0\\1&0&1\\1&1&1\end{bmatrix},$$

则有

$$\begin{aligned}\boldsymbol{P}^{-1}\boldsymbol{AP}&=\boldsymbol{P}_2^{-1}(\boldsymbol{P}_1^{-1}\boldsymbol{AP}_1)\boldsymbol{P}_2\\&=\begin{bmatrix}1&\\&\boldsymbol{Q}^{-1}\end{bmatrix}\begin{bmatrix}4&2&1\\0&4&0\\0&1&4\end{bmatrix}\begin{bmatrix}1&\\&\boldsymbol{Q}\end{bmatrix}=\begin{bmatrix}4&1&2\\&4&1\\&&4\end{bmatrix}.\end{aligned}$$

需要指出，定理 1.1 中的矩阵 $\boldsymbol{P}$ 不一定唯一，最后得出的上三角矩阵也不一定唯一. 在构造矩阵 $\boldsymbol{P}$ 的过程中，为使步骤减少，选取 $\boldsymbol{P}_1$ 时可取矩阵 $\boldsymbol{A}$ 的多个线性无关的特征向量作为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots$这样能使 $\boldsymbol{P}_1^{-1}\boldsymbol{AP}_1$ 的前面多个列呈上三角状. 此外，根据定理 1.1，$\boldsymbol{A}^{\mathrm{T}}$ 与上三角矩阵相似，从而可得：任意 n 阶矩阵与下三角矩阵相似.

设多项式

$$f(\lambda)=b_0\lambda^m+b_1\lambda^{m-1}+\cdots+b_{m-1}\lambda+b_m,$$

对于 n 阶矩阵 $\boldsymbol{A}$，定义矩阵多项式

$$f(\boldsymbol{A})=b_0\boldsymbol{A}^m+b_1\boldsymbol{A}^{m-1}+\cdots+b_{m-1}\boldsymbol{A}+b_m\boldsymbol{I}$$

($\boldsymbol{I}$ 表示单位矩阵)，如果 $f(\boldsymbol{A})=\boldsymbol{O}$，那么称 $\boldsymbol{A}$ 为 $f(\lambda)$ 的矩阵根.

定理 1.2 (Hamilton-Cayley) 任意 n 阶矩阵都是其特征多项式的矩阵根.

证 设 n 阶矩阵 $\boldsymbol{A}$ 的特征多项式为

$$\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n,$$

记 $\boldsymbol{A}$ 的 n 个特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则有

$$\varphi(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

由定理 1.1 可知，存在可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{bmatrix}\lambda_1&*&\cdots&*\\&\lambda_2&\ddots&\vdots\\&&\ddots&*\\&&&\lambda_n\end{bmatrix},$$

于是

$$\begin{aligned}\varphi(\boldsymbol{P}^{-1}\boldsymbol{AP})&=(\boldsymbol{P}^{-1}\boldsymbol{AP}-\lambda_1\boldsymbol{I})(\boldsymbol{P}^{-1}\boldsymbol{AP}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{P}^{-1}\boldsymbol{AP}-\lambda_n\boldsymbol{I})\\&=\begin{bmatrix}0&*&\cdots&*\\&\lambda_2-\lambda_1&\ddots&\vdots\\&&\ddots&*\\&&&\lambda_n-\lambda_1\end{bmatrix}\begin{bmatrix}\lambda_1-\lambda_2&*&\cdots&*\\&0&\ddots&*\\&&\ddots&*\\&&&\lambda_n-\lambda_2\end{bmatrix}\end{aligned}$$

$$
\times(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda_3\boldsymbol{I})\cdots(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda_n\boldsymbol{I})
$$

$$
=\begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & * \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ & \lambda_2-\lambda_3 & * & \cdots & * \\ & & 0 & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_n-\lambda_3 \end{bmatrix}
$$

$$
\times\cdots\times\begin{bmatrix} \lambda_1-\lambda_n & * & \cdots & * \\ & \ddots & \ddots & \vdots \\ & & \lambda_{n-1}-\lambda_n & * \\ & & & 0 \end{bmatrix}=\boldsymbol{O},
$$

即 $\boldsymbol{P}^{-1}\varphi(\boldsymbol{A})\boldsymbol{P}=\boldsymbol{O}$，也就是 $\varphi(\boldsymbol{A})=\boldsymbol{O}$.

证毕.

在定理 1.1 的证明过程中，若取 $\boldsymbol{x}_1$ 为单位向量，并且扩充 $\boldsymbol{x}_1$ 为 $\mathbf{C}^k$ 的标准正交基，那么 $\boldsymbol{P}_1=[\boldsymbol{x}_1\ \boldsymbol{x}_2\cdots\ \boldsymbol{x}_k]$ 为酉矩阵. 同样，也可以选取 $\boldsymbol{Q}$ 为酉矩阵，从而 $\boldsymbol{P}$ 为酉矩阵. 进一步，若 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 的 n 个特征值都是实数，可取 $\boldsymbol{x}_1$ 为实的单位向量，并扩充 $\boldsymbol{x}_1$ 为 $\mathbf{R}^k$ 的标准正交基，得到的 $\boldsymbol{P}_1$ 为正交矩阵. 同样，也可以选取 $\boldsymbol{Q}$ 为正交矩阵，从而 $\boldsymbol{P}$ 为正交矩阵. 综上所述，可得如下的 Schur 定理.

定理 1.3　n 阶矩阵 $\boldsymbol{A}$ 酉相似于上三角矩阵.

定理 1.4　设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 的特征值都是实数，则 $\boldsymbol{A}$ 正交相似于上三角矩阵.

在计算数学领域，人们颇为关心的是，通过相似变换可以把一个矩阵变换成何种最简单的形状. 利用定理 1.3 和定理 1.4，可以导出矩阵酉（正交）相似于对角矩阵的充要条件.

定理 1.5　n 阶矩阵 $\boldsymbol{A}$ 酉相似于对角矩阵的充要条件是 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$.

证　必要性. 设酉矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$，这里 $\boldsymbol{\Lambda}$ 表示对角矩阵. 则有

$$
\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{H}},\quad \boldsymbol{A}^{\mathrm{H}}=\boldsymbol{P}\bar{\boldsymbol{\Lambda}}\boldsymbol{P}^{\mathrm{H}},
$$

$$
\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{P}\bar{\boldsymbol{\Lambda}}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{H}}=\boldsymbol{P}\bar{\boldsymbol{\Lambda}}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{H}}=\boldsymbol{P}\boldsymbol{\Lambda}\bar{\boldsymbol{\Lambda}}\boldsymbol{P}^{\mathrm{H}}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{P}\bar{\boldsymbol{\Lambda}}\boldsymbol{P}^{\mathrm{H}}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}.
$$

充分性. 由定理 1.3 可知，存在酉矩阵 $\boldsymbol{P}$，使得

$$
\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{bmatrix},
$$

记上式右端的矩阵为 $\boldsymbol{B}$，利用条件 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$，可得

$$
\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}=\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}=\boldsymbol{B}\boldsymbol{B}^{\mathrm{H}}.
$$

比较上式两端矩阵主对角线上的对应元素，可得 $b_{ij}=0$ $(i<j)$，也就是 $\boldsymbol{B}=\mathrm{diag}(b_{11},b_{12},\cdots,b_{nn})$，即矩阵 $\boldsymbol{A}$ 酉相似于对角矩阵 $\boldsymbol{B}$.

证毕.

定理 1.6 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 的 n 个特征值都是实数，则 $\boldsymbol{A}$ 正交相似于对角矩阵的充要条件是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$.

证明类似于定理 1.5，这里从略.

由定理 1.5 可得，n 阶 Hermite 矩阵酉相似于对角矩阵. 由定理 1.6 可得，n 阶实对称矩阵正交相似于对角矩阵. 因此，n 阶 Hermite 矩阵存在完备的特征向量系（$\mathbf{C}^n$ 中）. n 阶实对称矩阵也存在完备的特征向量系（$\mathbf{R}^n$ 中）. 下面介绍两个 n 阶实对称矩阵“同时”相似于对角矩阵的问题.

定理 1.7 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 阶实对称矩阵，则存在正交矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 和 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}$ 都是对角矩阵的充要条件是 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{B}\boldsymbol{A}$.

证 充分性. 设 $\boldsymbol{A}$ 的全体互异特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_r$，相应的重数为 $n_1,n_2,\cdots,n_r(n_1+n_2+\cdots+n_r=n)$. 由 $\boldsymbol{A}$ 实对称知，存在正交矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\begin{bmatrix}\lambda_1\boldsymbol{I}_{n_1} & & & \\ & \lambda_2\boldsymbol{I}_{n_2} & & \\ & & \ddots & \\ & & & \lambda_r\boldsymbol{I}_{n_r}\end{bmatrix}.$$

划分矩阵 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q}$ 为分块形式

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1r}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2r}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{B}_{r1} & \boldsymbol{B}_{r2} & \cdots & \boldsymbol{B}_{rr}\end{bmatrix},$$

其中 $\boldsymbol{B}_{ij}$ 为 $n_i\times n_j$ 矩阵 $(i,j=1,2,\cdots,r)$. 由 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{B}\boldsymbol{A}$ 及 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{I}$ 可得

$$(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q})(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q})=(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q})(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}),$$

即

$$\begin{bmatrix}\lambda_1\boldsymbol{B}_{11} & \lambda_1\boldsymbol{B}_{12} & \cdots & \lambda_1\boldsymbol{B}_{1r}\\ \lambda_2\boldsymbol{B}_{21} & \lambda_2\boldsymbol{B}_{22} & \cdots & \lambda_2\boldsymbol{B}_{2r}\\ \vdots & \vdots & & \vdots\\ \lambda_r\boldsymbol{B}_{r1} & \lambda_r\boldsymbol{B}_{r2} & \cdots & \lambda_r\boldsymbol{B}_{rr}\end{bmatrix}=\begin{bmatrix}\lambda_1\boldsymbol{B}_{11} & \lambda_2\boldsymbol{B}_{12} & \cdots & \lambda_r\boldsymbol{B}_{1r}\\ \lambda_1\boldsymbol{B}_{21} & \lambda_2\boldsymbol{B}_{22} & \cdots & \lambda_r\boldsymbol{B}_{2r}\\ \vdots & \vdots & & \vdots\\ \lambda_1\boldsymbol{B}_{r1} & \lambda_2\boldsymbol{B}_{r2} & \cdots & \lambda_r\boldsymbol{B}_{rr}\end{bmatrix}.$$

也就是

$$\lambda_i\boldsymbol{B}_{ij}=\lambda_j\boldsymbol{B}_{ij}\qquad(i,j=1,2,\cdots,r).$$

当 $i\neq j$ 时，由 $\lambda_i\neq\lambda_j$ 可得 $\boldsymbol{B}_{ij}=\boldsymbol{O}$ $(i\neq j;\ i,j=1,2,\cdots,r)$. 于是

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q}=\mathrm{diag}(\boldsymbol{B}_{11},\boldsymbol{B}_{22},\cdots,\boldsymbol{B}_{rr}).$$

因为 $Q^{\mathrm{T}}BQ$ 实对称，所以 B_{ii} ($i=1,2,\cdots,r$) 实对称，那么存在正交矩阵 T_i，使得 $T_i^{\mathrm{T}}B_{ii}T_i=D_i$ ($i=1,2,\cdots,r$)，其中 D_i 为对角矩阵. 令 $T=\mathrm{diag}(T_1,T_2,\cdots,T_r)$，则 T 为正交矩阵，从而 $P=QT$ 也为正交矩阵，且有

$$P^{\mathrm{T}}BP=T^{\mathrm{T}}(Q^{\mathrm{T}}BQ)T=\mathrm{diag}(D_1,D_2,\cdots,D_r),$$
$$P^{\mathrm{T}}AP=T^{\mathrm{T}}(Q^{\mathrm{T}}AQ)T=\mathrm{diag}(\lambda_1I_{n_1},\lambda_2I_{n_2},\cdots,\lambda_rI_{n_r}).$$

必要性. 设 n 阶正交矩阵 P 使得

$$P^{\mathrm{T}}AP=\Lambda,\quad P^{\mathrm{T}}BP=D,$$

其中 Λ 和 D 都是对角矩阵，则有

$$AB=P\Lambda P^{\mathrm{T}}PDP^{\mathrm{T}}=P\Lambda DP^{\mathrm{T}}=PD\Lambda P^{\mathrm{T}}=PDP^{\mathrm{T}}P\Lambda P^{\mathrm{T}}=BA.$$

证毕.

定理 1.7 给出了两个实对称矩阵“同时”正交相似于对角矩阵的充要条件，下面再给出两个 Hermite 矩阵“同时”合同于对角矩阵的充分条件.

定理 1.8 设 A 和 B 都是 n 阶 Hermite 矩阵，且 B 为正定矩阵，则存在可逆矩阵 P，使得 $P^{\mathrm{H}}AP$ 和 $P^{\mathrm{H}}BP$ 都是对角矩阵.

证 由 B 正定知，存在酉矩阵 Q_1，使得

$$Q_1^{\mathrm{H}}BQ_1=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}=\Lambda_1\qquad(\lambda_i>0).$$

令

$$Q_2=Q_1\begin{bmatrix}1/\sqrt{\lambda_1} & & \\ & \ddots & \\ & & 1/\sqrt{\lambda_n}\end{bmatrix},$$

则 Q_2 可逆，且有 $Q_2^{\mathrm{H}}BQ_2=I$. 对于 Hermite 矩阵 $Q_2^{\mathrm{H}}AQ_2$，存在酉矩阵 Q_3，使得

$$Q_3^{\mathrm{H}}(Q_2^{\mathrm{H}}AQ_2)Q_3=\begin{bmatrix}\mu_1 & & \\ & \ddots & \\ & & \mu_n\end{bmatrix}=\Lambda_2.$$

令 $P=Q_2Q_3$，则 P 可逆，且有

$$P^{\mathrm{H}}AP=\Lambda_2,\qquad P^{\mathrm{H}}BP=I,$$

其中 Λ_2 和 I 都是对角矩阵.

证毕.

1.2 矩阵的 QR 分解

本节以 Givens 矩阵与 Householder 矩阵为基础，讨论将 n 阶实矩阵分解为正交矩阵与上三角矩阵的乘积问题，以及 n 阶实矩阵正交相似于 Hessenberg 矩

阵的问题.

定义 1.1 设实数 c 和 s 满足 $c^2+s^2=1$，称 n 阶矩阵

$$T_{ij}(c,s)=\begin{bmatrix} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{bmatrix}\begin{matrix} \\ (i) \\ \\ (j) \\ \\ \end{matrix} \qquad (i\neq j) \tag{1.1}$$

为 Givens 矩阵（初等旋转矩阵）.

性质 1 $[T_{ij}(c,s)]^{\mathrm{T}}[T_{ij}(c,s)]=I$,

$$[T_{ij}(c,s)]^{-1}=[T_{ij}(c,s)]^{\mathrm{T}}=T_{ij}(c,-s),\quad \det[T_{ij}(c,s)]=1.$$

性质 2 Givens 变换 $y=T_{ij}x$ 只改变 x 的第 i 个和第 j 个分量.

事实上，设 $x=(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$，$y=(\eta_1,\eta_2,\cdots,\eta_n)^{\mathrm{T}}$，则由 $y=T_{ij}x$ 可得

$$\eta_i=c\xi_i+s\xi_j,\quad \eta_j=-s\xi_i+c\xi_j,\quad \eta_k=\xi_k \qquad (k\neq i,j).$$

当 $\xi_i^2+\xi_j^2\neq 0$ 时，选取

$$c=\frac{\xi_i}{\sqrt{\xi_i^2+\xi_j^2}},\qquad s=\frac{\xi_j}{\sqrt{\xi_i^2+\xi_j^2}} \tag{1.2}$$

可使 $\eta_i=\sqrt{\xi_i^2+\xi_j^2}>0$，$\eta_j=0$.

性质 3 设 $x=(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}\neq 0$，则存在有限个 Givens 矩阵的乘积 T，使得 $Tx=|x|e_1$.

证 （1）设 $\xi_1\neq 0$，依次构造

$$T_{12}:\ c=\frac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}},\qquad s=\frac{\xi_2}{\sqrt{\xi_1^2+\xi_2^2}},$$

$$T_{12}x=(\sqrt{\xi_1^2+\xi_2^2},0,\xi_3,\cdots,\xi_n)^{\mathrm{T}},$$

$$T_{13}:\ c=\frac{\sqrt{\xi_1^2+\xi_2^2}}{\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}},\qquad s=\frac{\xi_3}{\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}},$$

$$T_{13}(T_{12}x)=(\sqrt{\xi_1^2+\xi_2^2+\xi_3^2},0,0,\xi_4,\cdots,\xi_n)^{\mathrm{T}},$$

……

$$T_{1n}:\ c=\frac{\sqrt{\xi_1^2+\cdots+\xi_{n-1}^2}}{\sqrt{\xi_1^2+\cdots+\xi_n^2}},\qquad s=\frac{\xi_n}{\sqrt{\xi_1^2+\cdots+\xi_n^2}},$$

$$T_{1n}(T_{1,n-1}\cdots T_{13}T_{12}x)=(\sqrt{\xi_1^2+\cdots+\xi_n^2},0,\cdots,0)^{\mathrm{T}},$$

令 $T=T_{1,n}T_{1,n-1}\cdots T_{13}T_{12}$，则有 $Tx=|x|e_1$.

（2）设 $\xi_1=\cdots=\xi_{k-1}=0$，$\xi_k\neq 0$（$1<k\leqslant n$），则由 $T_{1,k}$ 开始即可.