现代数学基础

23 矩阵计算六讲

■ 徐树方 钱江 著

高等教育出版社
HIGHER EDUCATION PRESS

■ 学科类别：数学

academic.hep.com.cn

ISBN 978-7-04-031966-8

9 787040 319668 >

定价 49.00 元

23

矩阵计算六讲

■ 徐树方　钱江　著

高等教育出版社·北京
HIGHER EDUCATION PRESS　BEIJING

内容简介

本书较系统地介绍了矩阵计算这门学科近十年来发展起来的新方法和新理论。全书共分 6 讲，内容包括：标准 Schur 分解、广义 Schur 分解和周期 Schur 分解的计算，特征值的排序问题，多项式之根的快速求法，奇异值分解的计算，求解线性方程组和特征值问题的 Krylov 子空间方法，以及求解特征值问题的共轭梯度法。

本书在选材上，在注重基础性和实用性的前提下，重点放在了反映该学科的最新进展上；在内容的处理上，在介绍方法的同时，尽可能地阐明方法的设计思想和理论依据，并对有关的结论尽可能地给出严格而又简洁的数学证明；在叙述表达上，力求清晰易读，便于教学与自学。

本书可作为综合性大学、理工科大学及高等师范院校计算数学、应用数学、工程计算等专业高年级本科生和研究生的教材或教学参考书，也可供从事科学与工程计算的科技人员参考。

图书在版编目（CIP）数据

矩阵计算六讲 / 徐树方，钱江著．—北京：高等教育出版社，2011.6

ISBN 978-7-04-031966-8

Ⅰ．①矩… Ⅱ．①徐… ②钱… Ⅲ．①矩阵－计算方法－高等学校－教材 Ⅳ．① O241.6

中国版本图书馆 CIP 数据核字（2011）第 064734 号

策划编辑 赵天夫　　责任编辑 赵天夫　　封面设计 赵 阳　　责任印制 朱学忠

出版发行 高等教育出版社
社　　址 北京市西城区德外大街 4 号
邮政编码 100120
印　　刷 涿州市星河印刷有限公司
开　　本 787×1092 1/16
印　　张 19.75
字　　数 350 000
购书热线 010-58581118

咨询电话 400-810-0598
网　　址 http://www.hep.edu.cn
http://www.hep.com.cn
网上订购 http://www.landraco.com
http://www.landraco.com.cn
版　　次 2011 年 6 月第 1 版
印　　次 2011 年 6 月第 1 次印刷
定　　价 49.00 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换

物 料 号 31966-00

前　言

矩阵计算是科学与工程计算的基石, 有关的方法和理论已经十分丰富, 许多大师级人物的专著不断问世: 早在 20 世纪 60 年代就有 Wilkinson 的 *The Algebraic Eigenvalue Problem*; 70 年代有 Stewart 的 *Introduction to Matrix Computation*; 80 年代有 Golub 和 Van Loan 的 *Matrix Computations*; 90 年代有 Demmel 的 *Applied Numerical Linear Algebra*; 进入本世纪又有 Stewart 的 *Matrix Algorithms*.

既然已经有了如此多大师级人物的专著, 为什么我们还要撰写这本教科书呢? 其原因有二: 一是这些大型的专著一般并不适合作为教科书来使用; 二是近年来又有不少新的方法和新的理论出现, 而这些方法和理论对矩阵计算的发展和应用又是十分重要的. 因此, 我们写作本书的一个主要目的就是为与计算数学有关专业的研究生和高年级本科生提供一本比较合适的教学参考书, 希望通过本书的学习, 使学生既能学到矩阵计算这门学科的典型方法和理论, 又能了解这一领域的最新进展. 此外, 我们的另一个目标是, 通过系统地阐述和展示这一领域的一些新方法的基本思路和具体实现的技巧与方法, 能够为有志从事科学与工程计算的科技工作者提供一些值得学习和借鉴的方法和理论.

全书共有六讲. 第一讲着重介绍著名的 QR 方法. QR 方法是上个世纪计算数学的重大成就之一, 是用于计算一个矩阵之 Schur 分解的. 这一讲可以看作是本书的基础, 特别是 QR 方法是本书的基石, 在后续的几讲里经常用到它. 作为 QR 方法的应用与推广, 我们还简要地介绍了计算广义 Schur 分解的 QZ 方法和计算周期 Schur 分解的周期 QZ 方法. 周期 QZ 方法是 20 世纪 90 年代才发展起来的, 是研究周期系统的重要工具. 此外, 我们也讨论了各类 Schur 标准形中对角块的重排问题.

第二讲致力于多项式之根的快速求法. 多项式的求根问题是一个古老而又在实际应用中经常遇到的数学问题, 有关的研究已经相当深入, 而且已经有不少行之有效的数值方法. 在著名的软件包 MATLAB 之中, 就有一个专门用于求多项式根的函数 **roots**. 该函数所采用的算法就是将著名的 QR 方法用到由多项式的系数所构成的友矩阵上而得到的. 由于这一方法并没有充分利用友矩阵所具有的特殊结构, 因

此用其求一个 n 次多项式的根所需的运算量为 $O(n^3)$. 直到最近这方面的研究才有了突破性的进展, 人们终于找到了一种如何充分利用友矩阵之特殊结构的方法, 使 QR 方法的代价从 $O(n^3)$ 降到了 $O(n^2)$. 本书将要介绍的快速 QR 方法是去年才发表的最新方法.

由于矩阵的奇异值分解的应用十分广泛, 因此如何更好、更快地求出给定矩阵之奇异值分解的问题就成为科学与工程计算领域中研究的一个热门课题, 吸引了众多的有关专家致力于这方面的研究, 涌现出了各种各样奇妙的方法. 早在 20 世纪 60 年代就有 Golub 和 Kahan 的工作, 巧妙地将对称 QR 方法应用到了奇异值分解的计算上, 开发出了所谓的 Golub-Kahan SVD 算法; 之后, 又有 Ming Gu 和 Eisenstat 的工作, 改进和完善了分而治之法, 使其大大地优于 Golub-Kahan SVD 算法; 最近, 又有 Drmač 和 Veselić 的工作, 将古老的 Jacobi 方法返老还童变成了一个速度快、计算精度高的实用方法, 该方法获得了“2009 SIAM Activity Group on Linear Algebra Prize”. 本书第三讲中我们将系统地介绍这些算法的设计思想和具体实现的方法与技巧.

由于目前计算机存储空间和运算速度的限制, QR 方法以及由其导出的各类方法只适用于小型矩阵特征值和特征向量的计算, 并不适用于求解大型矩阵的特征值问题. 对于大型稀疏矩阵的特征值问题, 目前较为有效和实用的一种求解方法就是 Krylov 子空间方法. 特别是近几年, 随着重新启动技术的不断完善, 使得这类方法的效率越来越高, 适用范围越来越广. 因此, 本书的第四讲我们将介绍与 Krylov 子空间方法有关的一些基本概念和重要理论, 并且详细地阐述目前这方面较为成熟的重启 Arnoldi 方法和重启 Lanczos 方法.

第五讲是第四讲的继续, 我们将着重介绍求解大型稀疏线性方程组的 Krylov 子空间方法, 内容包括: 共轭梯度法、极小剩余法、广义极小剩余法、拟极小剩余法和正交投影类方法.

众所周知, 目前共轭梯度法已经成为求解对称正定线性方程组和非线性优化问题的最常用方法之一. 然而随着 Krylov 子空间方法的兴起, 这几十年里用共轭梯度法来求解特征值问题的思想和方法已经几乎被大家遗忘了. 最近, 随着需要求解的特征值问题的规模的快速增长, 共轭梯度法因其所需存储量极少的缘故而又重新引起了大家的关注, 并且有关的方法和理论又有了新的进展. 本书第六讲我们将介绍用共轭梯度法来求解特征值问题的一些有关的基本概念和主要方法, 并且详细地阐述最近发展起来的梯度型子空间迭代法及其漂亮的收敛性理论.

本书在选材上, 在注重基础性和实用性的前提下, 重点放在了反映该学科的最新进展上; 在内容的处理上, 在介绍方法的同时, 尽可能地阐明方法的设计思想和理论依据, 并对有关的结论尽可能地给出严格而又简洁的数学证明; 在叙述表达上, 力求清晰易读, 便于教学与自学. 此外, 书中每讲都安排了几道难易程度不同的习题, 目的是为学生提供一些练习和实践的素材, 以便学生复习、巩固和拓广从课堂所学

的知识.

学习本书的先行课程是线性代数、数学分析和初等数值代数. 本书各讲相对独立, 读者可以选择其中任何一讲来阅读, 不会受其他部分内容的影响. 本书第一作者曾经用该书的原稿给北京大学计算数学专业的研究生和高年级本科生讲授过两次. 根据我们的教学实践, 如果不讲第一讲, 讲授其余五讲大约需要 60 学时.

本书第一作者有幸受美国 Texas 大学 Arlington 分校李仁仓教授的邀请于 2009 年初访问他两个月, 这段时间的学习、工作以及与李仁仓教授的讨论是促成本书的主要原因. 其实, 本书中一些新方法的有关信息都是在本书第一作者与李仁仓教授的讨论过程中获得的, 大部分相关资料都是李仁仓教授帮助收集的. 因此, 借本书出版之际, 特向李仁仓教授表示衷心的感谢!

北京大学数学科学学院科学与工程计算系的研究生梁鑫和施德才曾仔细阅读了全部书稿并指正了原稿中存在的不少错误, 梁鑫还对本书的部分算法进行了数值测试, 书中的例 2.2.1 和例 3.4.3 就是他给出的, 高等教育出版社的赵天夫编辑为本书的出版付出了辛勤的劳动, 在此一并表示诚挚的谢意.

限于水平, 书中的不当乃至错误之处难免, 恳请读者批评指正.

作者

2011 年 1 月

目 录

第一讲　Schur 分解的计算

这一讲, 我们将着重介绍计算标准 Schur 分解的 QR 方法、计算广义 Schur 分解的 QZ 方法以及上世纪 90 年代才发展起来的计算周期 Schur 分解的周期 QZ 方法, 并且讨论各类 Schur 标准形中对角块的重排问题. 这一讲可以看作是本书的基础, 特别是 QR 方法, 它是本书的基石, 在后续的几讲里经常用到.

§1.1 标准 Schur 分解的计算

首先, 我们来介绍标准 Schur 分解以及计算它的 QR 方法. 为此, 我们需要矩阵计算中的两个最常用的工具: Householder 变换和 Givens 变换.

§1.1.1　Householder 变换和 Givens 变换

Householder 变换是指如下形式的矩阵

$$H(w) = I - 2ww^*, \tag{1.1.1}$$

其中 $w \in \mathscr{C}^n$ 满足 $w^*w = 1$. 容易验证 Householder 变换具有如下基本性质:

(1) $\det H(w) = -1$;

(2) $H(w)^* = H(w) = H(w)^{-1}$;

(3) 对任意 $u \in w^{\perp}$, 有

$$H(w)(u + \beta w) = u - \beta w, \qquad \beta \in \mathscr{C};$$

(4) 设 $a, b \in \mathscr{C}^n$, $a \neq b$, 则存在 Householder 变换 $H(w)$ 使得 $H(w)a = b$ 成立的充分必要条件是

$$a^*a = b^*b \quad 和 \quad a^*b = b^*a;$$

并且在上述条件成立时, 所需的向量 w 可取作

$$w = e^{i\theta} \frac{a - b}{\sqrt{(a-b)^*(a-b)}},$$

其中 θ 为任意实数.

注 1.1.1　Householder 变换的性质 (3) 的几何解释如图 1.1.1 所示, 即 $H(w)(u+\beta w)$ 是 $u+\beta w$ 关于 w 的垂直超平面的镜像反射. 因此, 有时我们亦称 Householder 变换为**镜像变换**.

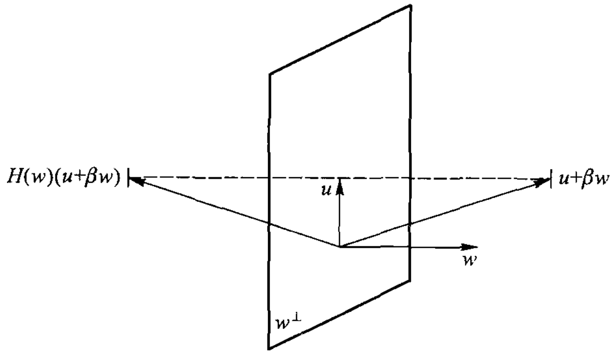

图 1.1.1

注 1.1.2　由 Householder 变换的性质 (4) 可知, 对任给的

$$a = (\alpha_1, \alpha_2, \ldots, \alpha_n)^{\top} \neq 0,$$

若令

$$\delta = \begin{cases} 1, & \text{当 } \alpha_1 = 0 \text{ 时}, \\ \bar{\alpha}_1/|\alpha_1|, & \text{当 } \alpha_1 \neq 0 \text{ 时}, \end{cases} \tag{1.1.2}$$

并取

$$w = \frac{\delta a + \|a\|_2 e_1}{\sqrt{(\delta a + \|a\|_2 e_1)^*(\delta a + \|a\|_2 e_1)}} = \frac{\delta a + \|a\|_2 e_1}{\sqrt{2\|a\|_2(\|a\|_2 + |\alpha_1|)}},$$

则有 $H(w)a = -\bar{\delta}\|a\|_2 e_1$. 而且当 a 是实向量时, 这样得到的 $H(w)$ 是实对称正交矩阵.

基于注 1.1.2, 我们有如下的算法.

算法 1.1.1 (计算 Householder 变换)　给定一个向量 $a \in \mathscr{C}^n$(这里不要求 $a \neq 0$), 下面的算法计算一个向量 w 和一个数 γ, 使得 $Ha = \gamma e_1$, 其中 $H = I - ww^*$, $\|w\|_2 = \sqrt{2}$.

function $[w, \gamma] =$ **house** (a)
　$w = a$; $\gamma = \|a\|_2$
　if $\gamma = 0$
　　$w(1) = \sqrt{2}$; 结束
　end
　if $w(1) \neq 0$

$\delta = \overline{w(1)}/|w(1)|$

else

$\delta = 1$

end

$w = (\delta/\gamma)w;\ w(1) = w(1) + 1$

$w = w/\sqrt{w(1)};\ \gamma = -\bar{\delta}\gamma$

注 1.1.3 在目前所流行的大多数教科书中, Householder 变换的计算只是针对实向量的, 而我们这里所给出的算法 1.1.1 既能用于实向量又能用于复向量.

注 1.1.4 在实际应用中, 由该算法得到的 Householder 变换 H 主要作用于某一给定的矩阵 A: 一种情况是用 H 左乘 A, 此时有

$$HA = (I - ww^*)A = A - wv^*, \quad v^* = w^*A;$$

另一种情况是用 H 右乘 A, 此时有

$$AH = A(I - ww^*) = A - vw^*, \quad v = Aw.$$

因此, 通常我们并不需要把矩阵 H 明确地计算出来, 只需保存构成它的向量 w 即可.

作为 Householder 变换的应用, 我们来证明如下的 QR 分解定理. QR 分解有许多重要的应用, 特别地, 它是这里将要介绍的 QR 方法的基础.

定理 1.1.1 (QR 分解定理) 设 $A \in \mathscr{C}^{m\times n}\,(m \geq n)$, 则存在酉矩阵 $Q \in \mathscr{C}^{m\times m}$ 和上三角矩阵 $R \in \mathscr{C}^{n\times n}$, 使得

$$A = Q\begin{bmatrix} R \\ 0 \end{bmatrix}. \tag{1.1.3}$$

当 $A \in \mathscr{C}^{n\times n}$ 非奇异且 R 的对角元均为非负数时, 分解是唯一的. 此外, 若 A 是实矩阵, 则 Q 和 R 均可取做实的. 通常, 我们称形如 (1.1.3) 的分解为矩阵 A 的 **QR 分解**.

证明 先证存在性. 对 n 用数学归纳法. 当 $n = 1$ 时, 显然成立. 现假定我们已经证明定理对所有的 $p \times (n-1)$ 矩阵成立, 这里假设 $p \geq (n-1)$. 设 $A \in \mathscr{R}^{m\times n}$ 的第一列为 a_1, 并且无妨假定 $a_1 \neq 0$. 由注 1.1.2 知, 存在一个 Householder 变换 $Q_1 \in \mathscr{C}^{m\times m}$ 使得 $Q_1^* a_1 = \delta e_1$. 于是有

$$Q_1^* A = \begin{bmatrix} \delta & * \\ 0 & A_1 \end{bmatrix}, \quad A_1 \in \mathscr{R}^{(m-1)\times(n-1)}.$$

对于 $(m-1)\times(n-1)$ 矩阵 A_1, 由归纳法假定知, 存在一个 $m-1$ 阶酉矩阵 Q_2 和一个 $n-1$ 阶上三角矩阵 R_2, 使得

$$A_1 = Q_2\begin{bmatrix} R_2 \\ 0 \end{bmatrix}.$$

这样, 令

$$Q = Q_1 \begin{bmatrix} 1 & 0 \\ 0 & Q_2 \end{bmatrix}, \qquad R = \begin{bmatrix} \delta & * \\ 0 & R_2 \end{bmatrix},$$

则 Q 和 R 满足定理的要求. 于是, 由归纳法原理知定理的存在性得证.

再证唯一性. 假定 $A = QR = \widetilde{Q}\widetilde{R}$, 其中 $Q, \widetilde{Q} \in \mathscr{C}^{n\times n}$ 是酉矩阵, $R, \widetilde{R} \in \mathscr{C}^{n\times n}$ 是具有非负对角元的上三角阵. 由于 A 是非奇异的, 因此 $R, \widetilde{R}$ 的对角元均为正数. 这样我们有

$$\widetilde{Q}^* Q = \widetilde{R}R^{-1}$$

既是酉矩阵又是对角元均为正数的上三角矩阵, 这只能是单位矩阵. 从而, 必有 $\widetilde{Q} = Q$, $\widetilde{R} = R$, 即此时 QR 分解是唯一的.

此外, 由注 1.1.2 和存在性的证明可知, 当 A 是实矩阵时, 当然 Q 和 R 均可取做实的. □

其实这一定理的证明过程也给我们提供了一种计算给定矩阵的 QR 分解的方法, 请读者作为练习给出详细的算法.

Givens 变换是指形如

$$G(i,j;c,s) = \begin{array}{c} \left[\begin{array}{ccccccccc} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& c & & \cdots & & s && \\ &&&& 1 &&&& \\ &&& \vdots && \ddots && \vdots & \\ &&&&&& 1 && \\ &&& -\bar{s} && \cdots && \bar{c} & \\ &&&&&&&& 1 \\ &&&&&&&&& \ddots \\ &&&&&&&&&& 1 \end{array}\right] \begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{array} \\ \qquad\qquad i \qquad\qquad\qquad j \end{array} \tag{1.1.4}$$

的矩阵, 其中 $c, s \in \mathscr{C}$ 满足 $|c|^2 + |s|^2 = 1$. 有时我们亦称 Givens 变换为 (i,j) 平面的**平面旋转变换**. 显然, $G(i,j;c,s)$ 是一个酉矩阵.

设 $x = (\xi_1, \xi_2, \ldots, \xi_n)^\top \in \mathscr{C}^n$, 并设 $y = G(i,j;c,s)x = (\eta_1, \eta_2, \ldots, \eta_n)^\top$. 容易验证, 当 $|\xi_i|^2 + |\xi_j|^2 \neq 0$ 时, 若取

$$c = \frac{|\xi_i|}{\sqrt{|\xi_i|^2 + |\xi_j|^2}}, \quad s = \frac{\bar{\xi}_j}{\sqrt{|\xi_i|^2 + |\xi_j|^2}} \frac{\xi_i}{|\xi_i|}, \tag{1.1.5}$$

则有

$$\eta_k = \xi_k,\ k \neq i,\ j\,, \quad \eta_i = \frac{\xi_i}{|\xi_i|}\sqrt{|\xi_i|^2 + |\xi_j|^2}, \quad \eta_j = 0, \tag{1.1.6}$$

即可以通过 Givens 变换把向量 x 的第 j 个分量化为零.

根据 (1.1.5), 我们可以设计算法如下.

算法 1.1.2 (计算 Givens 变换) 给定两个复数 α 和 β, 下面的算法计算三个数 c, s, η, 使得

$$\begin{bmatrix} c & s \\ -\bar{s} & \bar{c} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \eta \\ 0 \end{bmatrix}, \quad |c|^2 + |s|^2 = 1.$$

function $[c,\ s,\ \eta] =$ **givens** (α, β)

if $\beta = 0$

  $c = 1;\ s = 0;\ \eta = \alpha;$ 结束

end

if $\alpha = 0$

  $c = 0;\ s = 1;\ \eta = \beta;$ 结束

end

$\mu = \alpha/|\alpha|;\ \tau = |\alpha| + |\beta|$

$\delta = \tau\sqrt{|\alpha/\tau|^2 + |\beta/\tau|^2}$

$c = |\alpha|/\delta;\ s = \mu\bar{\beta}/\delta;\ \eta = \delta\mu$

注 1.1.5 注意, 该算法产生的数 c 是非负实数. 在实际应用中, 由该算法得到的 Givens 变换 $G(i,j;c,s)$ 主要作用于某一给定的矩阵 A. 用 $G(i,j;c,s)$ 左 (或右) 乘一个矩阵 A 时, 只有矩阵 A 的第 i, j 行 (或列) 的元改变, 而其他元都保持不变.

§1.1.2 Schur 分解定理

我们先介绍一些有关的基本概念.

设 $A \in \mathscr{C}^{n\times n}$. 如果存在 $x \in \mathscr{C}^n$ 和 $\lambda \in \mathscr{C}$ 满足

$$Ax = \lambda x, \quad x \neq 0, \tag{1.1.7}$$

则称 λ 是 A 的**特征值**, x 是 A 之属于特征值 λ 的**特征向量**. 有时为了叙述简洁, 常将特征值 λ 和对应的特征向量放在一起记作 (λ, x), 并且称之为 A 的一个**特征对**. A 的特征值的全体记作 $\lambda(A)$, 称之为 A 的**谱**.

根据特征值的定义, 易知 $\lambda \in \lambda(A)$ 的充要条件是

$$\det(\lambda I - A) = 0.$$

因此, $P_A(\lambda) = \det(\lambda I - A)$ 叫做 A 的**特征多项式**. 对任意给定的 n 阶方阵 A, 其特征多项式必可分解为

$$P_A(\lambda) = \prod_{i=1}^{r}(\lambda - \lambda_i)^{n_i}, \quad \lambda_i \neq \lambda_j\ (i \neq j), \tag{1.1.8}$$

$$n_1 + n_2 + \cdots + n_r = n,$$

其中的 n_i 叫做特征值 λ_i 的**代数重数**, 记作 $n(\lambda_i) = n_i$.

对任意给定的 $\lambda \in \lambda(A)$, 定义

$$\mathcal{E}_\lambda = \{\, x \in \mathscr{C}^n \;:\; Ax = \lambda x \,\}. \tag{1.1.9}$$

容易验证 $\mathcal{E}_\lambda$ 是 $\mathscr{C}^n$ 的一个子空间, 通常称之为 A 对应于特征值 λ 的**特征子空间**. $\mathcal{E}_\lambda$ 的维数称作特征值 λ 的**几何重数**, 记作 $m(\lambda) = \dim \mathcal{E}_\lambda$, 它表示 A 之属于 λ 的线性无关特征向量的个数. 显然, 有

$$1 \leqslant m(\lambda) = n - \operatorname{rank}(\lambda I - A) \leqslant n(\lambda) \leqslant n.$$

习惯上, 人们常将代数重数简称为**重数**, 例如, 在今后的叙述中, 我们说 λ 是 A 的 p 重特征值, 是指 λ 的代数重数是 p. 一般将代数重数为 1 的特征值称作**单特征值**, 而将几何重数与代数重数相等的特征值称作**半单特征值**.

这里需指出的是, 若 λ 是 A 的 r 重特征值, 则在其谱集 $\lambda(A)$ 中就含有 r 个 λ. 例如, $A = \operatorname{diag}(1,1,2)$, 则有 $\lambda(A) = \{1,1,2\}$.

设 $A, B \in \mathscr{C}^{n\times n}$, 如果存在非奇异矩阵 $X \in \mathscr{C}^{n\times n}$ 使得 $A = XBX^{-1}$, 则称 A 与 B 是**相似的**. 显然, 若 A 与 B 相似, 则 $P_A(\lambda) = P_B(\lambda)$, 而且 (λ, y) 是 B 的一个特征对当且仅当 (λ, Xy) 是 A 的一个特征对. 因此, 理论分析和实际计算的一个重要的手段, 就是设法将一个一般的方阵 A, 通过相似变换变成一个具有特殊结构的简单矩阵 B, 然后通过 B 的特征值和特征向量所具有的特性来导出 A 的特征值和特征向量所具有的特性. 下面我们就来介绍通过酉相似变换所能达到的最简形式.

设 (λ_1, x_1) 是矩阵 A 的一个特征对, 即

$$Ax_1 = \lambda_1 x_1. \tag{1.1.10}$$

由于 x_1 是一个非零向量, 由注 1.1.2 可知, 存在一个 Householder 变换 H_1 使得

$$H_1 x_1 = \delta e_1, \quad \delta \neq 0.$$

注意到方程 (1.1.10) 等价于

$$H_1 A H_1^* H_1 x_1 = \lambda_1 H_1 x_1,$$

便有

$$H_1 A H_1^* e_1 = \lambda_1 e_1.$$

这表明矩阵 $H_1 A H_1^*$ 必有如下形式:

$$H_1 A H_1^* = \begin{bmatrix} \lambda_1 & * \\ 0 & A_1 \end{bmatrix}, \quad A_1 \in \mathscr{C}^{(n-1)\times(n-1)}.$$