生数学教学系列（工科类）　5

矩阵论简明教程

徐　仲　张凯院
陆　全　冷国伟　编著

2

GM

科学出版社

研究生数学教学系列(工科类)

矩阵论简明教程

徐　仲　张凯院
陆　全　冷国伟　编著

科学出版社

2001

内 容 简 介

本书共分七章,介绍矩阵的相似变换,范数理论,矩阵分析,矩阵分解,特征值的估计与表示,广义逆矩阵以及矩阵的直积.各章末配有习题,书末有答案或提示.与传统矩阵论教材不同的是,本书不是从较抽象的线性空间与线性变换开始,而是以较具体的矩阵相似变换理论作为基础来介绍矩阵理论的主要内容,以达到由浅入深的目的,并使读者在较短时间内掌握近现代矩阵理论相当广泛而又很基本的内容.学习过工科线性代数课程的读者均可阅读本书.

本书可作为一般院校工科硕士研究生和工程硕士生的教材,以及本科高年级学生选修课教材,也可供工程技术或研究人员自学及参考使用.

图书在版编目(CIP)数据

矩阵论简明教程/徐仲等编著.-北京:科学出版社,2001
(研究生数学教学系列(工科类))
ISBN 7-03-009660-6

Ⅰ.矩… Ⅱ.徐… Ⅲ.矩阵-研究生-教材 Ⅳ.O151.21

中国版本图书馆 CIP 数据核字(2001)第 057250 号

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
新蕾印刷厂 印刷
科学出版社发行 各地新华书店经销
*
2001 年 9 月第 一 版 开本:B5(720×1000)
2001 年 9 月第一次印刷 印张:11 3/4
印数:1—3 000 字数:206 000

定价: 15.00 元

(如有印装质量问题,我社负责调换〈新欣〉)

前　言

近年来，由于计算机的发展和普及，矩阵理论的重要性愈加显著，应用日益广泛．这是因为用矩阵理论和方法来解决现代工程技术中的各种问题，不仅表述简洁，便于进行研究，而且具有适合计算机处理的特点．可以说，矩阵理论已成为从事科学研究和工程设计的科技人员必备的数学基础．

编者多年来在西北工业大学为理工科硕士研究生讲授矩阵论课程，并在大学本科高年级学生中多次开设相应的选修课，本书是在使用多遍讲义的基础上修改而成的．

本书以大学通用的工程数学《线性代数》作为预备知识，但不涉及线性空间与线性变换等较抽象的内容．这是基于以下考虑．现有的各种矩阵论教材，无一例外将抽象的线性空间与线性变换的理论放在第一章或第二章讲授，虽然这些内容对培养学生数学素养是不可缺少的，但工科研究生，特别是工程硕士生一开始就学习这些内容相对来说比较困难，加之，这部分内容（约需讲授20学时）与后续的具体矩阵理论部分联系的不是很紧密，从课程内容的安排来看，有头重脚轻之感，且使教师对具体的矩阵理论讲解的选择余地大为缩小．本书的编写目的就是想为读者架设一座通向矩阵理论的桥梁，使读者在较短的时间内尽快地得到各自需要的矩阵知识．当然，编者并不认为完全砍掉抽象的线性空间与线性变换的理论是恰当的，只要学时许可，可以将这部分内容放到最后去讲，这样可以使学生由浅入深，由具体到抽象，在已学习了许多矩阵的知识后，再将其放到线性空间的框架内重新审视，以利于提高学生的数学素养．编者在多年的矩阵论教学过程中多次尝试采用这种方法，取得了较好的教学效果．

本书共分七章：第一章、矩阵的相似变换；第二章、范数理论；第三章、矩阵分析；第四章、矩阵分解；第五章、特征值的估计与表示；第六章、广义逆矩阵；第七章、矩阵的直积．第一章起着承上启下的作用，对于线性代数进行加深并为后续章节奠定必要的基础；第二章至第七章介绍近现代的矩阵理论和方法，这也是工科研究生和科技人员在实际中直接、大量地用到的工具．除第一、二章外，其余各章是相对独立的，不同专业可根据需要灵活选用．各章均配有一定数量的习题，书末附有习题答案与提示，讲完全书约需40～50学时．

本书第一、三、四章由徐仲编写，第二、五章由张凯院编写，第六章由陆全

编写，第七章由冷国伟编写．由徐仲对全书统稿．

鉴于本书的读者是高等院校的工科研究生、工程硕士生、大学本科高年级学生及科技工作者，因此在编写时既重视基本的理论，也注重应用．对于必要的理论推导和分析，尽量使其清晰和简明．对个别理论则不苛求推导，侧重于介绍方法和应用．

在本书编写过程中，西北工业大学研究生院、教务处和应用数学系的领导及同事们给我们以很大的鼓励和支持；航空工业总公司631研究所周天孝教授详细审阅了书稿，提出了中肯的修改意见，并给予很高的评价，编者在此一并表示衷心的感谢．

由于我们水平有限，书中错误和疏漏之处难免，恳望有关专家和读者不吝赐教．

作　者

2001年2月于西北工业大学

符 号 说 明

$\overline{\boldsymbol{A}}$	矩阵 $\boldsymbol{A}$ 的共轭
$\boldsymbol{A}^{\mathrm{T}}$	矩阵 $\boldsymbol{A}$ 的转置
$\boldsymbol{A}^{\mathrm{H}}$	矩阵 $\boldsymbol{A}$ 的共轭转置(即 $\overline{\boldsymbol{A}}^{\mathrm{T}}$)
$\boldsymbol{A}^{+}$	矩阵 $\boldsymbol{A}$ 的 Moore-Penrose 逆
$\overrightarrow{\boldsymbol{A}}$	矩阵 $\boldsymbol{A}$ 的拉直
$\boldsymbol{A}^{(i,j,\cdots,l)}$	矩阵 $\boldsymbol{A}$ 的$\{i,j,\cdots,l\}$-逆
$A\{i,j,\cdots,l\}$	矩阵 $\boldsymbol{A}$ 的$\{i,j,\cdots,l\}$-逆的集合
$\boldsymbol{A}\sim\boldsymbol{B}$	方阵 $\boldsymbol{A}$ 相似于$\boldsymbol{B}$
$\boldsymbol{A}\otimes\boldsymbol{B}$	矩阵 $\boldsymbol{A}$ 与$\boldsymbol{B}$ 的直积或 Kronecker 积
$\boldsymbol{J}$	方阵的 Jordan 标准形
$\boldsymbol{J}_i$	第 i 个 Jordan 块
$\boldsymbol{I}$	单位矩阵
$\boldsymbol{O}$	零矩阵
$\boldsymbol{0}$	零向量
$\boldsymbol{e}_i$	第 i 个分量为 1,其余分量为 0 的 n 维列向量
$\mathrm{adj}\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的伴随矩阵
$\det\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的行列式
$\mathrm{cond}(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的条件数
$\mathrm{rank}\boldsymbol{A}$	矩阵 $\boldsymbol{A}$ 的秩
$\mathrm{tr}\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的迹,$\boldsymbol{A}$ 的主对角元之和
$\rho(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的谱半径
$\Vert\boldsymbol{A}\Vert$	矩阵 $\boldsymbol{A}$ 的范数
$\mathbf{R}$	实数域
$\mathbf{R}^n$	实 n 维列向量集合,n 维实向量空间
$\mathbf{R}^{m\times n}$	实 $m\times n$ 矩阵集合
$\mathbf{R}_r^{m\times n}$	秩为 r 的实 $m\times n$ 矩阵集合
$\mathbf{C}$	复数域
$\mathbf{C}^n$	复 n 维列向量集合,n 维复向量空间

$\mathbf{C}^{m\times n}$	复 $m\times n$ 矩阵集合
$\mathbf{C}_r^{m\times n}$	秩为 r 的复 $m\times n$ 矩阵集合
$[\boldsymbol{x},\boldsymbol{y}]$	向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积
$\mathrm{diag}(a_1,a_2,\cdots,a_n)$	以 $a_1,a_2,\cdots,a_n$ 为对角元素的 n 阶对角矩阵
$\mathrm{span}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s)$	由向量 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s$ 生成的子空间
$\psi(\lambda)$	方阵 $\boldsymbol{A}$ 的特征多项式
$m_{\boldsymbol{A}}(\lambda)$	方阵 $\boldsymbol{A}$ 的最小多项式
$G_k(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的第 k 个 Gerschgorin 圆(盖尔圆)
$\sigma_i=\sqrt{\lambda_i}$	矩阵 $\boldsymbol{A}$ 的第 i 个奇异值
$\mathrm{Re}(\lambda)$	复数 λ 的实部
$\mathrm{Im}(\lambda)$	复数 λ 的虚部
$f(\lambda)\mid g(\lambda)$	多项式 $f(\lambda)$ 整除 $g(\lambda)$
$X\cap Y$	集合 X 与 Y 的交集
$X\cup Y$	集合 X 与 Y 的并集

(O-1524.0101)
责任编辑：张邦固　胡华强／封面设计：黄华斌　王　浩

研究生数学教学系列

1. 现代科学计算
2. 工程数学基础
3. 矩阵论
4. 非线性系统的理论和方法
5. 矩阵论简明教程

ISBN 7-03-009660-6
9 787030 096609 >

ISBN 7-03-009660-6/O · 1524
定　价：15.00 元

目　　录

第一章　矩阵的相似变换 …… 1
　§1.1　特征值与特征向量 …… 1
　§1.2　相似对角化 …… 5
　§1.3　Jordan 标准形介绍 …… 9
　§1.4　Hamilton-Cayley 定理 …… 19
　§1.5　向量的内积 …… 23
　§1.6　酉相似下的标准形 …… 28
　习题一 …… 35
第二章　范数理论 …… 37
　§2.1　向量范数 …… 37
　§2.2　矩阵范数 …… 43
　　一、方阵的范数 …… 43
　　二、与向量范数的相容性 …… 45
　　三、从属范数 …… 46
　　四、长方阵的范数 …… 51
　§2.3　范数应用举例 …… 52
　　一、矩阵的谱半径 …… 52
　　二、矩阵的条件数 …… 53
　习题二 …… 56
第三章　矩阵分析 …… 58
　§3.1　矩阵序列 …… 58
　§3.2　矩阵级数 …… 60
　§3.3　矩阵函数 …… 66
　　一、矩阵函数的定义 …… 67
　　二、矩阵函数值的计算 …… 68
　　三、常用矩阵函数的性质 …… 74
　§3.4　矩阵的微分和积分 …… 76
　　一、函数矩阵的微分和积分 …… 76
　　二、数量函数对矩阵变量的导数 …… 78

三、矩阵值函数对矩阵变量的导数 ······ 81
§3.5　矩阵分析应用举例 ······ 82
一、求解一阶线性常系数微分方程组 ······ 82
二、求解矩阵方程 ······ 84
三、最小二乘问题 ······ 85
习题三 ······ 87
第四章　矩阵分解 ······ 90
§4.1　矩阵的三角分解 ······ 90
一、三角分解及其存在惟一性问题 ······ 90
二、三角分解的紧凑计算格式 ······ 93
§4.2　矩阵的 QR 分解 ······ 97
一、Householder 矩阵与 Givens 矩阵 ······ 98
二、矩阵的 QR 分解 ······ 104
三、矩阵酉相似于 Hessenberg 矩阵 ······ 109
§4.3　矩阵的满秩分解 ······ 112
一、Hermite 标准形 ······ 113
二、矩阵的满秩分解 ······ 116
§4.4　矩阵的奇异值分解 ······ 118
习题四 ······ 123
第五章　特征值的估计与表示 ······ 125
§5.1　特征值界的估计 ······ 125
§5.2　特征值的包含区域 ······ 128
一、Gerschgorin 定理 ······ 128
二、特征值的隔离 ······ 131
三、Ostrowski 定理 ······ 133
§5.3　Hermite 矩阵特征值的表示 ······ 135
§5.4　广义特征值问题 ······ 138
一、广义特征值问题 ······ 138
二、广义特征值的表示 ······ 140
习题五 ······ 143
第六章　广义逆矩阵 ······ 145
§6.1　广义逆矩阵的概念 ······ 145
§6.2　{1}-逆及其应用 ······ 146
一、{1}-逆的计算及有关性质 ······ 146

二、{1}-逆的应用 …… 149
三、由{1}-逆构造其他的广义逆矩阵 …… 151
§6.3 Moore-Penrose 逆 $\boldsymbol{A}^{+}$ …… 153
一、$\boldsymbol{A}^{+}$的计算及有关性质 …… 153
二、$\boldsymbol{A}^{+}$在解线性方程组中的应用 …… 155
习题六 …… 158
第七章 矩阵的直积 …… 160
§7.1 直积的定义和性质 …… 160
§7.2 直积的应用 …… 164
一、矩阵的拉直及其与直积的关系 …… 164
二、线性矩阵方程的可解性及其求解 …… 165
习题七 …… 168
习题答案与提示 …… 169
参考文献 …… 178

第一章　矩阵的相似变换

相似变换是矩阵的一种重要变换.本章研究矩阵在相似变换下的化简问题,这是矩阵理论中的基本问题之一.本章的内容是后续各章的基础,虽然部分概念在线性代数课程中已接触过,但更多的是新的内容.为避免篇幅过长,对部分结论将不加证明,读者可以在高等代数教材中找到有关的证明.

§1.1　特征值与特征向量

工程技术中的一些问题,如振动问题和稳定性问题,常可归结为求一个方阵的特征值和特征向量.

定义 1.1　设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,如果 $\lambda\in\mathbf{C}$ 和 $\mathbf{0}\neq\boldsymbol{x}\in\mathbf{C}^n$ 使关系式

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x} \tag{1.1}$$

成立,则称 λ 为 $\boldsymbol{A}$ 的**特征值**,称 $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的对应特征值 λ 的**特征向量**.

将式(1.1)改写为

$$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\mathbf{0}$$

这是 n 个未知数 n 个方程的齐次线性方程组,它有非零解的充分必要条件是系数行列式 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$,这是以 λ 为未知数的一元 n 次方程,其最高次项 λ^n 的系数为 1(称为**首一的**).

定义 1.2　设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,称 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 为 $\boldsymbol{A}$ 的**特征矩阵**,又称 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的**特征多项式**.

显然,$\boldsymbol{A}$ 的特征值就是特征方程的根.特征方程在复数范围内恒有解,其个数为方程的次数(重根按重数计算),因此 n 阶方阵 $\boldsymbol{A}$ 有 n 个特征值.计算 n 阶方阵 $\boldsymbol{A}$ 的特征值与特征向量的步骤如下:

第一步:求 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$ 的 n 个根 $\lambda_1,\lambda_2,\cdots,\lambda_n$,它们即为 $\boldsymbol{A}$ 的全部特征值;

第二步:求解齐次方程组 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\mathbf{0}$,其非零解向量即为 $\boldsymbol{A}$ 的对应特征值 λ_i 的特征向量.

例 1.1　求下列矩阵的特征值与特征向量:

$$(1)\ \boldsymbol{A}=\begin{pmatrix}1&-2&2\\-2&-2&4\\2&4&-2\end{pmatrix};\quad(2)\ \boldsymbol{A}=\begin{pmatrix}-1&0&1\\1&2&0\\-4&0&3\end{pmatrix}.$$

解　(1)$\boldsymbol{A}$ 的特征多项式为

$$\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda-1 & 2 & -2\\ 2 & \lambda+2 & -4\\ -2 & -4 & \lambda+2\end{vmatrix}=(\lambda-2)^2(\lambda+7)$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=\lambda_2=2,\lambda_3=-7$.

当 $\lambda_1=\lambda_2=2$ 时,解方程组 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$. 由

$$2\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}1 & 2 & -2\\ 2 & 4 & -4\\ -2 & -4 & 4\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 2 & -2\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

得基础解系

$$\boldsymbol{p}_1=(-2,1,0)^{\mathrm{T}},\quad \boldsymbol{p}_2=(2,0,1)^{\mathrm{T}}$$

所以对应 $\lambda_1=\lambda_2=2$ 的全部特征向量为 $k_1\boldsymbol{p}_1+k_2\boldsymbol{p}_2$($k_1,k_2$ 不同时为 0).

当 $\lambda_3=-7$ 时,解方程组 $(-7\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$. 由

$$-7\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}-8 & 2 & -2\\ 2 & -5 & -4\\ -2 & -4 & -5\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & 1/2\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}$$

得基础解系

$$\boldsymbol{p}_3=(-1,-2,2)^{\mathrm{T}}$$

故对应 $\lambda_3=-7$ 的全部特征向量为 $k_3\boldsymbol{p}_3(k_3\neq0)$.

(2)$\boldsymbol{A}$ 的特征多项式为

$$\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda+1 & 0 & -1\\ -1 & \lambda-2 & 0\\ 4 & 0 & \lambda-3\end{vmatrix}=(\lambda-1)^2(\lambda-2)$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=\lambda_2=1,\lambda_3=2$.

当 $\lambda_1=\lambda_2=1$ 时,解方程组 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$. 由

$$\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}2 & 0 & -1\\ -1 & -1 & 0\\ 4 & 0 & -2\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & -1/2\\ 0 & 1 & 1/2\\ 0 & 0 & 0\end{pmatrix}$$

得基础解系

$$\boldsymbol{p}_1=(1,-1,2)^{\mathrm{T}}$$

所以 $k_1\boldsymbol{p}_1(k_1\neq0)$ 是对应 $\lambda_1=\lambda_2=1$ 的全部特征向量.

当 $\lambda_3=2$ 时,解方程组 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$. 由

$$2\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}3 & 0 & -1\\ -1 & 0 & 0\\ 4 & 0 & -1\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}$$

得基础解系

$$p_3=(0,1,0)^{\mathrm{T}}$$

故 $k_3p_3(k_3\neq0)$是对应 $\lambda_3=2$ 的全部特征向量.

矩阵的特征值与特征向量有如下一些性质.

定理 1.1 设 λ_i 是$A\in C^{n\times n}$的 r_i 重特征值(称 r_i 为特征值λ_i 的**代数重数**),对应 λ_i 有s_i 个线性无关的特征向量(称 s_i 为特征值λ_i 的**几何重数**),则 $1\leqslant s_i\leqslant r_i$.(证明略)

对于例 1.1 中的第 1 个矩阵 A,对应 2 重特征值 2 有 2 个线性无关的特征向量,而第 2 个矩阵 A 对应 2 重特征值 1 只有 1 个线性无关的特征向量.

定义 1.3 设 $f(\lambda)$是 λ 的多项式

$$f(\lambda)=a_s\lambda^s+a_{s-1}\lambda^{s-1}+\cdots+a_1\lambda+a_0$$

对于 $A\in C^{n\times n}$,规定

$$f(A)=a_sA^s+a_{s-1}A^{s-1}+\cdots+a_1A+a_0I$$

称 $f(A)$为**矩阵 A 的多项式**.

定理 1.2 设 $A\in C^{n\times n}$,A 的n 个特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$,对应的特征向量为 $x_1,x_2,\cdots,x_n$,又设 $f(\lambda)$为一多项式,则 $f(A)$的特征值为 $f(\lambda_1)$,$f(\lambda_2),\cdots,f(\lambda_n)$,对应的特征向量仍为 $x_1,x_2,\cdots,x_n$.如果 $f(A)=O$,则 A 的任一特征值λ_i 满足$f(\lambda_i)=0$.

证 因为 $Ax_i=\lambda_ix_i(i=1,2,\cdots,n)$,于是对正整数 k,有

$$A^kx_i=A^{k-1}(Ax_i)=\lambda_iA^{k-1}x_i=\cdots=\lambda_i^kx_i$$

故

$$\begin{aligned}f(A)x_i&=(a_sA^s+a_{s-1}A^{s-1}+\cdots+a_1A+a_0I)x_i\\&=a_sA^sx_i+a_{s-1}A^{s-1}x_i+\cdots+a_1Ax_i+a_0x_i\\&=(a_s\lambda_i^s+a_{s-1}\lambda_i^{s-1}+\cdots+a_1\lambda_i+a_0)x_i=f(\lambda_i)x_i\end{aligned}$$

当 $f(A)=O$ 时

$$0=f(A)x_i=f(\lambda_i)x_i$$

由 $x_i\neq0$ 知 $f(\lambda_i)=0$. 证毕

定理 1.3 设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是方阵A 的互不相同的特征值,$x_1,x_2,\cdots,x_s$ 是分别与之对应的特征向量,则 $x_1,x_2,\cdots,x_s$ 线性无关.

证 对 s 用数学归纳法证明.当 $s=1$ 时,因为 $x_1\neq0$,所以 x_1 线性无关,即定理成立.假定对 $s-1$ 个互不相同的特征值定理成立,下证对 s 个互不相同的特征值定理也成立.为此,设有常数 $k_1,k_2,\cdots,k_s$,使

$$k_1x_1+k_2x_2+\cdots+k_sx_s=0$$

由于 $\boldsymbol{A}\boldsymbol{x}_i=\lambda_i\boldsymbol{x}_i(i=1,2,\cdots,s)$,用 $\boldsymbol{A}$ 左乘上式得

$$k_1\lambda_1\boldsymbol{x}_1+k_2\lambda_2\boldsymbol{x}_2+\cdots+k_s\lambda_s\boldsymbol{x}_s=\boldsymbol{0}$$

从上面两个等式中消去 $\boldsymbol{x}_s$,得

$$k_1(\lambda_1-\lambda_s)\boldsymbol{x}_1+\cdots+k_{s-1}(\lambda_{s-1}-\lambda_s)\boldsymbol{x}_{s-1}=\boldsymbol{0}$$

由归纳假定,$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{s-1}$线性无关,又因为 $\lambda_i-\lambda_s\neq0(i=1,2,\cdots,s-1)$,所以 $k_1=k_2=\cdots=k_{s-1}=0$,进而可得 $k_s=0$,故 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s$ 线性无关. 证毕

定理 1.3 还可以推广为如下的定理(证明与定理 1.3 相仿,这里略去).

定理 1.4　设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是方阵 $\boldsymbol{A}$ 的互不相同的特征值,$\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{ir_i}$是对应特征值 λ_i 的线性无关的特征向量$(i=1,2,\cdots,s)$,则向量组

$$\boldsymbol{x}_{11},\cdots,\boldsymbol{x}_{1r_1},\boldsymbol{x}_{21},\cdots,\boldsymbol{x}_{2r_2},\cdots,\boldsymbol{x}_{s1},\cdots,\boldsymbol{x}_{sr_s}$$

也线性无关.

定理 1.5　设 n 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$,则

(1) $a_{11}+a_{22}+\cdots+a_{nn}=\lambda_1+\lambda_2+\cdots+\lambda_n$;

(2) $\det\boldsymbol{A}=\lambda_1\lambda_2\cdots\lambda_n$;

(3) $\boldsymbol{A}^{\mathrm{T}}$ 的特征值是 $\lambda_1,\lambda_2,\cdots,\lambda_n$,而 $\boldsymbol{A}^{\mathrm{H}}=(\bar{a}_{ji})_{n\times n}$的特征值为$\bar{\lambda}_1,\bar{\lambda}_2,\cdots,\bar{\lambda}_n$.

证　由行列式的定义知,在 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})$的展开式中,有一项是主对角线上元素的乘积$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})$,而展开式中其余各项至多包含 $n-2$ 个主对角线上的元素,这是因为,如果某一项含有 $a_{ij}(i\neq j)$,则该项就不能含有 $\lambda-a_{ii}$与$\lambda-a_{jj}$,因此这些项关于 λ 的次数最多是$n-2$,于是

$$\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots$$

又因为 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})$的 n 个根,所以

$$\begin{aligned}\det(\lambda\boldsymbol{I}-\boldsymbol{A})&=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)\\&=\lambda^n-(\lambda_1+\lambda_2+\cdots+\lambda_n)\lambda^{n-1}+\cdots\end{aligned}\tag{1.2}$$

于是

$$a_{11}+a_{22}+\cdots+a_{nn}=\lambda_1+\lambda_2+\cdots+\lambda_n$$

在式(1.2)中取 $\lambda=0$ 得

$$\det(-\boldsymbol{A})=(-1)^n\lambda_1\lambda_2\cdots\lambda_n$$

从而

$$\det\boldsymbol{A}=\lambda_1\lambda_2\cdots\lambda_n$$

最后,由

$$\det(\bar{\lambda}_i\boldsymbol{I}-\boldsymbol{A}^{\mathrm{H}})=\det(\overline{\lambda_i\boldsymbol{I}-\boldsymbol{A}})^{\mathrm{T}}=\det(\overline{\lambda_i\boldsymbol{I}-\boldsymbol{A}})=\overline{\det(\lambda_i\boldsymbol{I}-\boldsymbol{A})}=0$$

知 $\bar{\lambda}_1,\bar{\lambda}_2,\cdots,\bar{\lambda}_n$ 是 $\boldsymbol{A}^{\mathrm{H}}$ 的特征值.同理可证 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}^{\mathrm{T}}$ 的特征值.

证毕

推论 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,则 0 是 $\boldsymbol{A}$ 的特征值的充分必要条件是 $\det\boldsymbol{A}=0$.

定义 1.4 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$,称

$$a_{11}+a_{22}+\cdots+a_{nn}$$

为 $\boldsymbol{A}$ 的**迹**,记为 $\mathrm{tr}\boldsymbol{A}$.

定理 1.5(1)表明,矩阵 $\boldsymbol{A}$ 的迹等于 $\boldsymbol{A}$ 的所有特征值的和.

关于矩阵的迹有以下结论.

定理 1.6 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$,则 $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}(\boldsymbol{BA})$.

证 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$, $\boldsymbol{B}=(b_{ij})_{n\times n}$,则 $\boldsymbol{AB}$ 的对角线元素为 $\sum\limits_{k=1}^{n}a_{ik}b_{ki}$ $(i=1,2,\cdots,n)$,而 $\boldsymbol{BA}$ 的对角线元素为 $\sum\limits_{i=1}^{n}b_{ki}a_{ik}(k=1,2,\cdots,n)$,于是

$$\mathrm{tr}(\boldsymbol{AB})=\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{ik}b_{ki}\right)=\sum_{k=1}^{n}\left(\sum_{i=1}^{n}b_{ki}a_{ik}\right)=\mathrm{tr}(\boldsymbol{BA})$$ 证毕

§1.2 相似对角化

定义 1.5 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$,若存在 $\boldsymbol{P}\in\mathbf{C}_n^{n\times n}$ 使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$$

则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相似**,记为 $\boldsymbol{A}\sim\boldsymbol{B}$,称 $\boldsymbol{P}$ 为把 $\boldsymbol{A}$ 变成 $\boldsymbol{B}$ 的**相似变换矩阵**.

相似是矩阵之间的一种重要的关系.相似矩阵具有以下性质.

定理 1.7 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}\in\mathbf{C}^{n\times n}$, $f(\lambda)$是一多项式.

(1) $\boldsymbol{A}\sim\boldsymbol{A}$(反身性);

(2) 若 $\boldsymbol{A}\sim\boldsymbol{B}$,则 $\boldsymbol{B}\sim\boldsymbol{A}$(对称性);

(3) 若 $\boldsymbol{A}\sim\boldsymbol{B}$, $\boldsymbol{B}\sim\boldsymbol{C}$,则 $\boldsymbol{A}\sim\boldsymbol{C}$(传递性);

(4) 若 $\boldsymbol{A}\sim\boldsymbol{B}$,则 $\det\boldsymbol{A}=\det\boldsymbol{B}$, $\mathrm{rank}\boldsymbol{A}=\mathrm{rank}\boldsymbol{B}$;

(5) 若 $\boldsymbol{A}\sim\boldsymbol{B}$,则 $f(\boldsymbol{A})\sim f(\boldsymbol{B})$;

(6) 若 $\boldsymbol{A}\sim\boldsymbol{B}$,则 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\det(\lambda\boldsymbol{I}-\boldsymbol{B})$,即 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征多项式,从而特征值相同.

证 只证(5)和(6).设

$$f(\lambda)=a_s\lambda^s+a_{s-1}\lambda^{s-1}+\cdots+a_1\lambda+a_0$$

因为 $\boldsymbol{A}\sim\boldsymbol{B}$,所以存在 $\boldsymbol{P}\in\mathbf{C}_n^{n\times n}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$,于是

$$f(\boldsymbol{B})=a_s\boldsymbol{B}^s+a_{s-1}\boldsymbol{B}^{s-1}+\cdots+a_1\boldsymbol{B}+a_0\boldsymbol{I}$$

$$= a_s(\boldsymbol{P}^{-1}\boldsymbol{AP})^s + \cdots + a_1(\boldsymbol{P}^{-1}\boldsymbol{AP}) + a_0\boldsymbol{I}$$
$$= \boldsymbol{P}^{-1}(a_s\boldsymbol{A}^s + \cdots + a_1\boldsymbol{A} + a_0\boldsymbol{I})\boldsymbol{P} = \boldsymbol{P}^{-1}f(\boldsymbol{A})\boldsymbol{P}$$

又有

$$\det(\lambda\boldsymbol{I} - \boldsymbol{B}) = \det(\lambda\boldsymbol{I} - \boldsymbol{P}^{-1}\boldsymbol{AP}) = \det(\boldsymbol{P}^{-1}(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{P}) = \det(\lambda\boldsymbol{I} - \boldsymbol{A})$$

证毕

对角矩阵是较简单的矩阵之一,无论计算它的乘积、幂、逆矩阵和特征值等都比较方便.现在的问题是,方阵 $\boldsymbol{A}$ 能否相似于一个对角矩阵?

定义 1.6 设 $\boldsymbol{A}\in \mathbf{C}^{n\times n}$,如果 $\boldsymbol{A}$ 相似于一个对角矩阵,则称 $\boldsymbol{A}$ **可对角化**.

定理 1.8 设 $\boldsymbol{A}\in \mathbf{C}^{n\times n}$,则 $\boldsymbol{A}$ 可对角化的充分必要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

证 设 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$,其中 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)$,则由 $\boldsymbol{AP}=\boldsymbol{P\Lambda}$ 得

$$\boldsymbol{A}\boldsymbol{p}_i = \lambda_i\boldsymbol{p}_i \qquad (i = 1,2,\cdots,n)$$

可见 λ_i 是 $\boldsymbol{A}$ 的特征值,$\boldsymbol{P}$ 的列向量 $\boldsymbol{p}_i$ 是对应特征值 λ_i 的特征向量,再由 $\boldsymbol{P}$ 可逆知 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 线性无关.

反之,如果 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$,即有

$$\boldsymbol{A}\boldsymbol{p}_i = \lambda_i\boldsymbol{p}_i \qquad (i = 1,2,\cdots,n)$$

记 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)$,则 $\boldsymbol{P}$ 可逆,且有

$$\boldsymbol{AP} = (\boldsymbol{A}\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2,\cdots,\boldsymbol{A}\boldsymbol{p}_n) = (\lambda_1\boldsymbol{p}_1,\lambda_2\boldsymbol{p}_2,\cdots,\lambda_n\boldsymbol{p}_n)$$
$$= (\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n) = \boldsymbol{P}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

即有

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

故 $\boldsymbol{A}$ 可对角化. 证毕

由定理的证明过程可以看到,若 n 阶方阵 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似,则 $\boldsymbol{\Lambda}$ 的主对角线元素恰为 $\boldsymbol{A}$ 的 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$,而相似变换矩阵 $\boldsymbol{P}$ 的 n 个列向量 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 分别是对应 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的特征向量.

由定理 1.3 和定理 1.4 可以得到如下两个便于使用的条件.

推论 1 如果 n 阶方阵 $\boldsymbol{A}$ 有 n 个互不相同的特征值,则 $\boldsymbol{A}$ 可对角化.

推论 2 设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 n 阶方阵 $\boldsymbol{A}$ 的所有互不相同的特征值,其重数分别为 $r_1,r_2,\cdots,r_s$.若对应 r_i 重特征值 λ_i 有 r_i 个线性无关的特征向量 $(i=1,2,\cdots,s)$,则 $\boldsymbol{A}$ 可对角化.

例 1.2 下列哪个矩阵可对角化,哪个不可对角化?对于可对角化的矩