普通高等院校“十一五”规划教材

矩阵分析与计算

JUZHEN FENXI YU JISUAN

朱元国 饶玲 严涛 张军 李宝成 编

国防工业出版社
National Defense Industry Press

普通高等院校“十一五”规划教材

矩阵分析与计算

朱元国　饶　玲　严　涛
张　军　李宝成　编

国防工业出版社
·北京·

内 容 简 介

本书内容包括矩阵的标准型，向量范数与矩阵范数，矩阵分解，特征值的估计与计算，广义逆矩阵，矩阵函数，线性方程组的直接解法，线性最小二乘问题，线性方程组的迭代解法等内容，最后一章介绍线性空间与线性变换，是线性代数相关内容的简介。本书的特点之一是在介绍矩阵论有关基础理论的同时，引入矩阵计算的相关内容，使读者能将解决问题的精确方法与近似方法进行对比，了解到精确方法在实际计算中的缺陷以及近似方法在实际应用中的有效性。

本书可作为工科高校研究生教材，也可作为理科或管理等学科的研究生、教师及有关研究者的参考书。

图书在版编目(CIP)数据

矩阵分析与计算 / 朱元国等编. —北京：国防工业出版社，2010.8

普通高等院校“十一五”规划教材

ISBN 978-7-118-06896-2

Ⅰ. ①矩... Ⅱ. ①朱... Ⅲ. ①矩阵分析-高等学校-教材②矩阵-计算方法-高等学校-教材 Ⅳ. ①O151.21②O241.6

中国版本图书馆 CIP 数据核字(2010)第 124602 号

※

国防工業出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100048)

国防工业出版社印刷厂印刷

新华书店经售

*

开本 787×1092 1/16 **印张** 15 **字数** 342 千字

2010 年 8 月第 1 版第 1 次印刷 **印数** 1—4000 册 **定价** 28.00 元

国防书店：(010)68428422 发行邮购：(010)68414474

发行传真：(010)68411535 发行业务：(010)68472764

前　　言

随着研究生教育改革的深入发展,各校对研究生课程教学也进行了有效的改革。“矩阵分析”是所有工科高校都开设的研究生公共基础课程,因专业的要求,有的学校还开设了相关课程,如“数值代数”,“数值逼近”等,本质上就是要给学生介绍数值计算的有关理论和方法。但教学改革的方向之一是使研究生尤其是硕士研究生有较为宽广的基础知识面,当前教育部提出要大力培养应用型研究生,这为我们进行研究生课程改革提出了新的要求。一方面,要开设更多的课程与研究生有限的培养时间有冲突;另一方面,学生在理论学习与实际应用上有冲突。在矩阵分析课程的教学中,我们发现已出版的多数教材注重理论分析和问题的精确计算,学生并不能很好地将所学知识应用到自己的研究领域。在工科各学科理论研究中,需要理论分析;而在应用研究中,大多数问题需要的计算是近似计算。我们在不增加其他课程的基础上,在矩阵分析课程中增加了矩阵计算的内容,使得学生能同时学习到矩阵的理论分析和矩阵计算的知识,能将解决问题的精确方法与近似方法进行对比,了解到精确方法在实际计算中的缺陷以及近似方法在实际应用中的有效性。通过几年的教学实践,我们发现这样能使学生在素质上有一定的提高。现将在该课程的教学内容汇编成本书出版。

任何涉及数学的领域(包括工程学、最优化、经济学、控制论、电子学、网络等)都需要矩阵的知识。本书介绍矩阵分析及计算的基本概念、理论和方法,使学生掌握矩阵的理论分析及应用的概念和方法,做到理论性强、知识面广,提高学生分析问题、解决问题的能力,使学生能灵活地应用所学的矩阵论知识处理各自学科研究中出现的相关问题。通过本书的学习,使矩阵论成为学生们今后进行科学研究的强有力工具。本书包括矩阵的标准型,向量范数与矩阵范数,矩阵分解,特征值的估计与计算,广义逆矩阵,矩阵函数,线性方程组的直接解法,线性最小二乘问题,线性方程组的迭代解法等内容,最后一章介绍线性空间与线性变换,是线性代数相关内容的简介,可以作为学习本书的基础。本书可作为工科高校研究生教材,也可作为理科或管理等学科的研究生、教师及有关研究者的参考书。

本书第 1 章和第 10 章由李宝成编写,第 2 章由朱元国编写,第 3 章和第 4 章由严涛编写,第 5 章和第 6 章由张军编写,第 7 章、第 8 章及第 9 章由饶玲编写。全书由朱元国统稿。由于编者水平有限,书中难免有疏漏之处,敬请广大专家及读者批评指正。

作者

2010 年 3 月于南京理工大学

常用符号

$\boldsymbol{x},\boldsymbol{y}$	向量
$\boldsymbol{A},\boldsymbol{B}$	矩阵
$\boldsymbol{C}$	复数域
$\boldsymbol{R}$	实数域
$\boldsymbol{C}^n$	n 维复向量集合
$\boldsymbol{R}^n$	n 维实向量集合
$\boldsymbol{C}^{m\times n}$	m 行 n 列复矩阵组成的集合
$\boldsymbol{C}_r^{m\times n}$	m 行 n 列秩为 r 的复矩阵组成的集合
$\boldsymbol{R}^{m\times n}$	m 行 n 列实矩阵组成的集合
$\overline{\boldsymbol{A}}$	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
$\boldsymbol{0}$	零向量,零矩阵
$\boldsymbol{I}$	单位矩阵
$\boldsymbol{J}$	方阵的 Jordan 标准型
$\mathrm{rank}\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的秩
$\det\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的行列式
$\mathrm{tr}\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的迹($\boldsymbol{A}$ 的主对角线上元素之和)
$\mathrm{cond}(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的条件数
$\rho(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的谱半径
$\\|\boldsymbol{A}\\|$	矩阵 $\boldsymbol{A}$ 的范数
$\mathrm{diag}(a_1,a_2,\cdots,a_n)$	以 $a_1,a_2,\cdots,a_n$ 为对角元素的 $n\times n$ 对角矩阵

目　录

第 1 章　矩阵的标准形

本章主要讨论矩阵在相似变换下的标准形问题。首先给出在线性代数课程中已介绍的相似对角形,然后借助 λ 矩阵导出矩阵的 Jordan 标准形,最后讨论一些特殊矩阵在酉相似下的标准形。

1.1　矩阵的相似对角形

本节对矩阵的特征值和特征向量及相似对角化等基本内容进行复习。这些内容一般要在后续各章中用到。

1.1.1　特征值与特征向量

定义 1.1　设 $\boldsymbol{A} \in \boldsymbol{C}^{n\times n}$,若 $\lambda \in \boldsymbol{C}$,$0\neq\boldsymbol{\xi} \in \boldsymbol{C}^n$,使得

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

成立,则称 λ 为 $\boldsymbol{A}$ 的特征值,$\boldsymbol{\xi}$ 为 $\boldsymbol{A}$ 的对应特征值 λ 的特征向量。

将式(1.1)改写为

$$(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{\xi} = 0 \quad (\boldsymbol{\xi} \neq 0)$$

这表明特征向量 $\boldsymbol{\xi}$ 是齐次线性方程组

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

的非零解,由线性方程组的理论知,λ 为 $\boldsymbol{A}$ 的特征值的充分必要条件是行列式

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0 \tag{1.2}$$

式(1.2)称为矩阵 $\boldsymbol{A}$ 的特征方程,特征方程的根就是 $\boldsymbol{A}$ 的特征值。多项式 $f_A(\lambda) = \det(\lambda\boldsymbol{I} - \boldsymbol{A})$ 称为 $\boldsymbol{A}$ 的特征多项式,而矩阵 $(\lambda\boldsymbol{I} - \boldsymbol{A})$ 称作 $\boldsymbol{A}$ 的特征矩阵。

由此可得特征值与特征向量的求法如下:

(1) 从式(1.2)求 A 的特征值 λ;

(2) 对每一特征值 λ,求出齐次线性方程组 $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0$ 的基础解系,即为对应于特征值 λ 的线性无关特征向量。显然对应的任一特征向量可由它们线性表示。

1.1.2　特征值与特征向量的性质

性质 1.1　设 $\boldsymbol{A} \in \boldsymbol{C}^{n\times n}$,则

$$f_A(\lambda) = \lambda^n - (\mathrm{tr}\boldsymbol{A})\lambda^{n-1} + \cdots + (-1)^n\det\boldsymbol{A} \tag{1.3}$$

其中 $\mathrm{tr}\boldsymbol{A} = \sum_{i=1}^{n} a_{ii}$ 为矩阵 $\boldsymbol{A}$ 的迹。

事实上,在

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}$$

的展开式中,只有一项是主对角线上元素的连乘积,即

$$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn}) \tag{1.4}$$

而展开式中其余各项至多包含$(n-2)$个主对角线上的元素,所以特征多项式中含 λ 的 n 次与$(n-1)$次的项只能在主对角线上元素的连乘积,即式(1.4)中出现,它们是

$$\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}$$

在特征多项式中,令 $\lambda=0$,即可得它的常数项为

$$|-\boldsymbol{A}|=(-1)^n|\boldsymbol{A}|$$

因此,若只写出特征多项式的前两项与常数项,则有式(1.3)。

因为$f_A(\lambda)$为 n 次多项式,由代数学基本定理知道,在复数域上$f_A(\lambda)$可作如下因式分解:

$$f_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_k)^{n_k}$$

式中:$\lambda_1,\cdots,\lambda_k$ 为$f_A(\lambda)$的互异零点,从而是 $\boldsymbol{A}$ 的互异特征值;n_i 为特征值 λ_i 的重数,称为 λ_i 的代数重数,也说 λ_i 为 $\boldsymbol{A}$ 的 n_i 重特征值,$i=1,2,\cdots,k$。特别地,$n_i=1$ 时,也称作单特征值。于是有:

性质 1.2 n 阶矩阵 $\boldsymbol{A}$ 有且仅有 n 个特征值,其中 m 重特征值以 m 个计。

应当指出,5 次或 5 次以上的多项式方程一般是没有公式求解的。所以对于阶数较大的矩阵,实际上求特征值是非常困难的,因而就要研究特征值的各种近似求法。本书在第 4 章中给出的幂法和 QR 算法正是求特征值近似值的基本方法。

再由式(1.3)中根与系数的关系可得

性质 1.3 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}=(a_{ij})_{n\times n}$的 n 个特征值(λ_i 未必互异),则

$$\mathrm{tr}\boldsymbol{A}=\sum_{i=1}^{n}\lambda_i,\ \det\boldsymbol{A}=\prod_{i=1}^{n}\lambda_i$$

显然 $\det\boldsymbol{A}=0$ 当且仅当 $\boldsymbol{A}$ 具有零特征值。

设 $\boldsymbol{A}=(a_{ij})\in \boldsymbol{C}^{m\times n}$,用 $\overline{\boldsymbol{A}}=(\overline{a_{ij}})$表示以 $\boldsymbol{A}$ 的元素的共轭复数为元素组成的矩阵,而 $\boldsymbol{A}^{\mathrm{H}}=(\overline{\boldsymbol{A}})^{\mathrm{T}}$ 称为 $\boldsymbol{A}$ 的共轭转置矩阵。

容易验证矩阵的共轭转置运算具有下列性质:

(1) $\boldsymbol{A}^{\mathrm{H}}=\overline{(\boldsymbol{A}^{\mathrm{T}})}$;

(2) $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}+\boldsymbol{B}^{\mathrm{H}}$;

(3) $(k\boldsymbol{A})^{\mathrm{H}}=\bar{k}\boldsymbol{A}^{\mathrm{H}}$;

(4) $(\boldsymbol{A}\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}$;

(5) $(\boldsymbol{A}^{\mathrm{H}})^{\mathrm{H}}=\boldsymbol{A}$;

(6) 如果 $\boldsymbol{A}$ 可逆,则 $(\boldsymbol{A}^{\mathrm{H}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{H}}$。

性质 1.4 设 λ 为 n 阶矩阵 $\boldsymbol{A}$ 的特征值,则

(1) λ 也是矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的特征值;

(2) $\bar{\lambda}$ 为矩阵 $\boldsymbol{A}^{\mathrm{H}}$ 的特征值;

(3) 若 $\boldsymbol{A}$ 非奇异,λ^{-1} 为矩阵 $\boldsymbol{A}^{-1}$ 的特征值。

定义 1.2 设 $g(x)$ 为 x 的多项式,即

$$g(x)=a_m x^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$$

对于 $\boldsymbol{A}\in\boldsymbol{C}^{n\times n}$,则相应的多项式

$$g(A)=a_m\boldsymbol{A}^m+a_{m-1}\boldsymbol{A}^{m-1}+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{I}$$

称为矩阵 $\boldsymbol{A}$ 的 m 次多项式。

显然,$g(\boldsymbol{A})$ 仍为一 n 阶矩阵。矩阵多项式有许多类似于代数多项式的性质:

(1) $\boldsymbol{A}^2-\boldsymbol{I}=(\boldsymbol{A}+\boldsymbol{I})(\boldsymbol{A}-\boldsymbol{I})$;

(2) $\boldsymbol{A}^3+\boldsymbol{I}=(\boldsymbol{A}+\boldsymbol{I})(\boldsymbol{A}^2-\boldsymbol{A}+\boldsymbol{I})$;

(3) $f(x)$、$g(x)$ 为任意两多项式,则 $f(\boldsymbol{A})g(\boldsymbol{A})=g(\boldsymbol{A})f(\boldsymbol{A})$,即同一矩阵 $\boldsymbol{A}$ 的二个多项式的乘法满足交换律;

(4) 设 $\boldsymbol{A}=\mathrm{diag}(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_r)$ 为块对角阵,$g(x)$ 为一多项式,则

$$g(\boldsymbol{A})=\mathrm{diag}(g(\boldsymbol{A}_1),g(\boldsymbol{A}_2),\cdots,g(\boldsymbol{A}_r))$$

(5) 若 $\boldsymbol{P}$ 为可逆矩阵,$g(x)$ 为一多项式,则

$$g(\boldsymbol{PAP}^{-1})=\boldsymbol{P}g(\boldsymbol{A})\boldsymbol{P}^{-1}$$

性质 1.5 设 $\boldsymbol{A}\in\boldsymbol{C}^{n\times n}$,$\lambda_0$ 为 $\boldsymbol{A}$ 的特征值,$\boldsymbol{x}$ 为相应的特征向量,又 m 次多项式 $g(x)=\sum_{k=0}^{m}a_kx^k$,则 $g(\lambda_0)$ 是 $g(\boldsymbol{A})$ 的特征值,$\boldsymbol{x}$ 仍为 $g(\boldsymbol{A})$ 的属于 $g(\lambda_0)$ 的特征向量。如果 $g(\boldsymbol{A})=0$,则有 $g(\lambda_0)=0$。

注 1.1 $g(\boldsymbol{A})$ 的特征向量不一定是 $\boldsymbol{A}$ 的特征向量。例如,取

$$g(x)=x^2,\boldsymbol{A}=\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$$

则

$$g(\boldsymbol{A})=\boldsymbol{A}^2=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}$$

易得矩阵 $\boldsymbol{A}$ 的 0 特征值对应的线性无关的特征向量为 $\boldsymbol{\xi}_1=(1,0,0)^{\mathrm{T}}$,而矩阵 $\boldsymbol{A}^2$ 的 0 特征值对应的线性无关的特征向量为 $\boldsymbol{\xi}_1=(1,0,0)^{\mathrm{T}}$,$\boldsymbol{\xi}_2=(0,1,0)^{\mathrm{T}}$,显然,$\boldsymbol{\xi}_2$ 不是 $\boldsymbol{A}$ 的特征向量,因 $\boldsymbol{A\xi}_2\neq 0\cdot\boldsymbol{\xi}_2$。

下面再考察矩阵的特征向量。对于不同特征值对应的特征向量之间的关系,可由下面的性质给出。

性质 1.6 设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 为矩阵 $\boldsymbol{A}$ 的互异特征值,对应的特征向量分别为 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_s$,则 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_s$ 必线性无关。

推论 1.1 设 $\lambda_1,\cdots,\lambda_s$ 为矩阵 $\boldsymbol{A}$ 的互异特征值,若 $\boldsymbol{\xi}_{i1},\cdots,\boldsymbol{\xi}_{il_i}$ 分别为属于 λ_i 的线性无关的特征向量($i=1,\cdots,s$),则 $\boldsymbol{\xi}_{11},\cdots,\boldsymbol{\xi}_{1l_1},\cdots,\boldsymbol{\xi}_{s1},\cdots,\boldsymbol{\xi}_{sl_s}$ 也线性无关。

定义 1.3 设 $\boldsymbol{A}\in\boldsymbol{C}^{n\times n}$,$\lambda_0$ 为 $\boldsymbol{A}$ 的特征值,称齐次线性方程组

$$(\lambda_0\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=0$$

的解空间为 $\boldsymbol{A}$ 的属于 λ_0 的特征子空间,记作 V_{λ_0},而特征子空间的维数 $\dim V_{\lambda_0}$ 称为特征值 λ_0 的几何重数。

显然,$\dim V_{\lambda_0}=n-\operatorname{rank}(\lambda_0\boldsymbol{I}-\boldsymbol{A})$,也等于对应特征值 λ_0 的线性无关的特征向量的个数。

可以看出,矩阵 $\boldsymbol{A}$ 的每一特征值的代数重数和几何重数是两个不同的概念。但二者之间有如下结论。

性质 1.7 设 $\boldsymbol{A}\in\boldsymbol{C}^{n\times n}$,则 $\boldsymbol{A}$ 的任一特征值的几何重数不超过其代数重数。

注 1.2 如果矩阵 $\boldsymbol{A}\in\boldsymbol{C}^{m\times n}$ 的特征值 λ_0 为单特征值,则 $\dim V_{\lambda_0}=1$。

1.1.3 矩阵的对角化

设 $\boldsymbol{A}\in\boldsymbol{C}^{n\times n}$,是否存在一个非奇异矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 成为对角矩阵,是一个在理论和应用上都很重要的问题。为此引入定义:

定义 1.4 设 $\boldsymbol{A}$、$\boldsymbol{B}\in\boldsymbol{C}^{n\times n}$,若存在 $\boldsymbol{P}\in\boldsymbol{C}_n^{n\times n}$($\boldsymbol{C}_r^{m\times n}$ 表示秩为 r 的 $m\times n$ 阶复矩阵集合),使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$,则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似,记作 $\boldsymbol{A}\sim\boldsymbol{B}$,并称 $\boldsymbol{P}$ 为把 $\boldsymbol{A}$ 变成 $\boldsymbol{B}$ 的相似变换矩阵。

不难验证相似关系满足:

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

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

(3) 传递性:若 $\boldsymbol{A}\sim\boldsymbol{B}$,$\boldsymbol{B}\sim\boldsymbol{D}$,则 $\boldsymbol{A}\sim\boldsymbol{D}$。

相似的矩阵有许多共同的性质:

性质 1.8 设 $\boldsymbol{A}$、$\boldsymbol{B}\in\boldsymbol{C}^{n\times n}$,若 $\boldsymbol{A}\sim\boldsymbol{B}$,则

(1) $\det\boldsymbol{A}=\det\boldsymbol{B}$;

(2) $\boldsymbol{A}\simeq\boldsymbol{B}$,即 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价;

(3) $\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{B}$;

(4) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 或者都可逆,或者都不可逆,且当 $\boldsymbol{A}$,$\boldsymbol{B}$ 可逆时,有 $\boldsymbol{A}^{-1}\sim\boldsymbol{B}^{-1}$;

(5) 若 $f(x)$ 为一多项式,有 $f(\boldsymbol{A})\sim f(\boldsymbol{B})$;

(6) $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\det(\lambda\boldsymbol{I}-\boldsymbol{B})$,即相似矩阵有相同的特征多项式,从而有相同的特征值(包括重数相同);

(7) $\operatorname{tr}\boldsymbol{A}=\operatorname{tr}\boldsymbol{B}$,即相似矩阵有相同的迹。

定义 1.5 设 $\boldsymbol{A}\in\boldsymbol{C}^{n\times n}$,如果 $\boldsymbol{A}$ 相似于一个对角矩阵,则称 $\boldsymbol{A}$ 可对角化。

现在来分析,是否每一个矩阵都可以对角化?即矩阵可以对角化的某些特征。

若 $\boldsymbol{A}\sim\boldsymbol{\Lambda}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$,即存在可逆矩阵 $\boldsymbol{P}$,使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$,这等价于 $\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{\Lambda}$,

将 $\boldsymbol{P}$ 按列分块为 $\boldsymbol{P}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n)$。则有

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

即

$$(\boldsymbol{A\xi}_1,\boldsymbol{A\xi}_2,\cdots,\boldsymbol{A\xi}_n)=(\lambda_1\boldsymbol{\xi}_1,\lambda_2\boldsymbol{\xi}_2,\cdots,\lambda_n\boldsymbol{\xi}_n)$$

可得 $\boldsymbol{A\xi}_i=\lambda_i\boldsymbol{\xi}_i(i=1,2,\cdots,n)$，又 $\boldsymbol{P}$ 可逆，所以 $\boldsymbol{\xi}_i\neq 0$，这表明 λ_i 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{P}$ 的列向量 $\boldsymbol{\xi}_i$ 是对应特征值 λ_i 的特征向量。再由 $\boldsymbol{P}$ 可逆，知 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n$ 线性无关。

反之，如果 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n$，即有 $\boldsymbol{A\xi}_i=\lambda_i\boldsymbol{\xi}_i(i=1,2,\cdots,n)$。令矩阵 $\boldsymbol{P}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n)$，显然 $\boldsymbol{P}$ 可逆，且有

$$\begin{aligned}\boldsymbol{AP}&=(\boldsymbol{A\xi}_1,\boldsymbol{A\xi}_2,\cdots,\boldsymbol{A\xi}_n)=(\lambda_1\boldsymbol{\xi}_1,\lambda_2\boldsymbol{\xi}_2,\cdots,\lambda_n\boldsymbol{\xi}_n)\\&=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n)\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)=\boldsymbol{P\Lambda}\end{aligned}$$

则有 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$，故 $\boldsymbol{A}$ 可对角化。

综上可得

定理 1.1　n 阶矩阵 $\boldsymbol{A}$ 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。

由上面的分析（证明）过程可以看到，若 n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似，则 $\boldsymbol{\Lambda}$ 的对角元恰为 $\boldsymbol{A}$ 的特征值，而相似变换矩阵 $\boldsymbol{P}$ 的每一列是 $\boldsymbol{A}$ 的特征向量。需要注意的是，λ_1，$\lambda_2,\cdots,\lambda_n$ 的排列顺序必须与 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n$ 的排列顺序相对应，否则 $\boldsymbol{P}$ 就不是原来的了。

用定理 1.1 判别矩阵是否可对角化需求出 $\boldsymbol{A}$ 的特征向量，并不方便，下面给出一个不需求出特征向量的较实用的判别法。

定理 1.2　设 $\boldsymbol{A}\in \boldsymbol{C}^{n\times n}$，则 $\boldsymbol{A}$ 可对角化的充分必要条件是 $\boldsymbol{A}$ 的每一个特征值的几何重数等于其代数重数。

因 $\dim V_{\lambda_i}=n-\mathrm{rank}(\lambda_i\boldsymbol{I}-\boldsymbol{A})$，所以使用定理 1.2 时不需求出 λ_i 对应的特征向量。由定理 1.2 及推论 1.1 得下面的结论。

推论 1.2　n 阶矩阵 $\boldsymbol{A}$ 如果有 n 个互异特征值，则 $\boldsymbol{A}$ 必可对角化。

1.2　λ 矩阵及标准形、不变因子和初等因子

本节引进 λ 矩阵这个工具，为后面的把矩阵化为 Jordan 标准形做准备。先介绍多项式的最大公因式及一些性质。

设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 为复数域 $\boldsymbol{C}$ 上关于未定元 x 的多项式，若 $a_n\neq 0$，则称 $f(x)$ 是一个 n 次多项式，a_nx^n 称为 $f(x)$ 的最高次项或首项，a_n 称为首项系数，a_0 称为常数项。若 $f(x)=a$，则称 $f(x)$ 为常数多项式，当 $a\neq 0$ 时，称它为零次多项式，当 $a=0$ 时，称之为零多项式。

称多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 为首 1 多项式，如果 $f(x)$ 的首项系数为 1，即有形式：

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

定义 1.6　设 $f(x)$、$g(x)$ 是复数域 $\boldsymbol{C}$ 上的多项式，如果存在复数域 $\boldsymbol{C}$ 上的多项式 $h(x)$，使

$$f(x)=g(x)h(x)$$

则称 $g(x)$ 整除 $f(x)$，记作 $g(x) \mid f(x)$。此时，称 $g(x)$ 为 $f(x)$ 的因式（或因子），$f(x)$ 称为 $g(x)$ 的倍式。

例如，$f(x)=(x+1)^2, g(x)=2(x+1)$，则 $g(x) \mid f(x)$，因 $f(x)=\frac{1}{2}(x+1)g(x)$。

由整除的定义知道，零多项式的因式可以是任一多项式，但零多项式不能是非零多项式的因式。下面不予证明的给出整除性的几个常用性质：

(1) 若 $g(x) \mid f(x)$，则 $cg(x) \mid f(x)(c\neq 0)$，由此知非零常数是任一非零多项式的因子；

(2) 若 $g(x) \mid f(x)$，$f(x) \mid g(x)$，且 $f(x)$、$g(x)$ 都是非零多项式，则 $f(x)=cg(x)$，其中 c 为非零常数；

(3) 若 $g(x) \mid f(x)$，$f(x) \mid h(x)$，则 $g(x) \mid h(x)$；

(4) 若 $g(x) \mid f(x)$，$g(x) \mid h(x)$，则对任意的多项式 $u(x)$、$v(x)$，有

$$g(x) \mid (f(x)u(x)+h(x)v(x))$$

如果多项式 $\varphi(x)$ 满足：$\varphi(x) \mid f(x)$，$\varphi(x) \mid g(x)$，则称 $\varphi(x)$ 是 $f(x)$ 与 $g(x)$ 的一个公因式（或公因子）。

定义 1.7 设 $f(x)$、$g(x)$ 是复数域 $\boldsymbol{C}$ 上多项式，如果存在复数域 $\boldsymbol{C}$ 上多项式 $d(x)$ 满足：

(1) $d(x) \mid f(x)$，$d(x) \mid g(x)$；

(2) 若 $d_1(x)$ 是 $f(x)$ 与 $g(x)$ 的公因式，则必有 $d_1(x) \mid d(x)$。

则称 $d(x)$ 是 $f(x)$、$g(x)$ 的一个最大公因式。

若 $d_1(x)$、$d_2(x)$ 是 $f(x)$ 与 $g(x)$ 的两个最大公因式，则由定义知二者必互相整除，即 $d_1(x) \mid d_2(x)$，$d_2(x) \mid d_1(x)$，这样就有 $d_1(x)=cd_2(x)(c\neq 0)$，这说明 $f(x)$ 与 $g(x)$ 的两个最大公因式至多相差一个非零常数。用 $(f(x),g(x))$ 表示首项系数为 1 的最大公因式。显然 $(f(x),g(x))$ 是唯一确定的。特别地，若 c 为非零常数，$f(x)$ 是首 1 多项式，则有：$(f(x),c)=1$，$(f(x),0)=f(x)$。

多个多项式 $f_1(x),\cdots,f_s(x)$ 的最大公因式可同样定义，仍用 $(f_1(x),\cdots,f_s(x))$ 表示首项系数为 1 的最大公因式，如对 $s=3$，有

$$(f_1(x),f_2(x),f_3(x))=((f_1(x),f_2(x)),f_3(x))$$

一般地，在复数域上求若干个多项式的最大公因式时，可先把每个多项式分解成一次因式的方幂的乘积形式，然后取含有公共一次因式的最低方幂的乘积，即得所求的最大公因式。例如：

$$f(x)=(x-1)^3(x+2)^2(x+5), g(x)=(x-1)^2(x+2)^3(x+3), h(x)=(x-1)^2(x+2)(x+3)^5$$

则有

$$(f(x),g(x),h(x))=(x-1)^2(x+2)$$

1.2.1 λ 矩阵的概念

定义 1.8 设 $a_{ij}(\lambda)(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 为复数域 $\boldsymbol{C}$ 上的多项式，则称以

$a_{ij}(\lambda)$为元素的 $m\times n$ 阶矩阵 $\boldsymbol{A}(\lambda)=(a_{ij}(\lambda))_{m\times n}$为 λ 矩阵或多项式矩阵。

为与 λ 矩阵相区别,把以复数域 $\boldsymbol{C}$ 中的数为元素的矩阵称为数字矩阵。显然,数字矩阵可看作是特殊的 λ 矩阵,因它的元素 a_{ij}为零次多项式。方阵 $\boldsymbol{A}$ 的特征矩阵$(\lambda\boldsymbol{I}-\boldsymbol{A})$也是一个特殊的 λ 矩阵。

λ 矩阵可同数字矩阵一样定义相等、加法、数乘、乘法等运算,且与数字矩阵有相同的运算规律。对于 $n\times n$ 阶的方 λ 矩阵可同样定义行列式,而且与数字矩阵的行列式有相同的性质。有了 λ 矩阵行列式的概念,也就有了 λ 矩阵的子式、余子式、伴随矩阵等概念,进而利用子式这个概念,就可给出 λ 矩阵秩的定义:

定义 1.9　如果 λ 矩阵 $\boldsymbol{A}(\lambda)$中有一个 r 阶$(r\geqslant 1)$子式为非零多项式,而所有$(r+1)$阶子式(如果有的话)全为零多项式,则称 $\boldsymbol{A}(\lambda)$的秩为 r,记作 $\mathrm{rank}\boldsymbol{A}(\lambda)=r$ 或 $r_{A(\lambda)}=r$。零矩阵的秩规定为零。

如果 n 阶 λ 矩阵 $\boldsymbol{A}(\lambda)$秩为 n,即 $\det\boldsymbol{A}(\lambda)\neq 0$,则称 $\boldsymbol{A}(\lambda)$为满秩的或非奇异的。若 n 阶 λ 矩阵 $\boldsymbol{A}(\lambda)$的秩小于 n,也即 $\det\boldsymbol{A}(\lambda)=0$,就称 $\boldsymbol{A}(\lambda)$是降秩的或奇异的。例如,数字矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$的特征矩阵 $\boldsymbol{A}(\lambda)=\lambda\boldsymbol{I}-\boldsymbol{A}$ 就是重要的满秩矩阵,因为$f_A(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})$不是零多项式(至多有 n 个根)。

定义 1.10　设 $\boldsymbol{A}(\lambda)$为一 n 阶 λ 矩阵,如果存在一个 n 阶 λ 矩阵 $\boldsymbol{B}(\lambda)$,满足

$$\boldsymbol{A}(\lambda)\boldsymbol{B}(\lambda)=\boldsymbol{B}(\lambda)\boldsymbol{A}(\lambda)=\boldsymbol{I} \tag{1.5}$$

则称 $\boldsymbol{A}(\lambda)$是可逆的,或称 $\boldsymbol{A}(\lambda)$是单模矩阵,这里 $\boldsymbol{I}$ 为 n 阶单位矩阵。$\boldsymbol{B}(\lambda)$称为 $\boldsymbol{A}(\lambda)$的逆矩阵,记为 $\boldsymbol{A}^{-1}(\lambda)=\boldsymbol{B}(\lambda)$。满足式(1.5)的 $\boldsymbol{A}(\lambda)$、$\boldsymbol{B}(\lambda)$显然互为逆矩阵。

与数字矩阵的证法一样,若 $\boldsymbol{A}(\lambda)$可逆,则一定唯一。

定理 1.3　n 阶 λ 矩阵 $\boldsymbol{A}(\lambda)$可逆的充分必要条件是 $\boldsymbol{A}(\lambda)$的行列式是一个非零常数,即

$$\det\boldsymbol{A}(\lambda)=c\neq 0$$

显然,可逆的 λ 矩阵一定是满秩的,但满秩的 λ 矩阵未必是可逆的。因满秩的 λ 矩阵的行列式不一定等于非零常数。例如,特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 满秩,但明显不可逆。又如,对 λ 矩阵

$$\boldsymbol{A}(\lambda)=\begin{pmatrix}\lambda+1 & \lambda+3\\ \lambda^2+3\lambda & \lambda^2+5\lambda+4\end{pmatrix},\boldsymbol{B}(\lambda)=\begin{pmatrix}\lambda & 2\lambda+1 & 1\\ 1 & \lambda+1 & \lambda^2+1\\ \lambda-1 & \lambda & -\lambda^2\end{pmatrix}$$

$$\boldsymbol{D}(\lambda)=\begin{pmatrix}\lambda & 0\\ 0 & \lambda+1\end{pmatrix}$$

经简单计算可得

$$\det\boldsymbol{A}(\lambda)=4,\det\boldsymbol{B}(\lambda)=0,\det\boldsymbol{D}(\lambda)=\lambda^2+\lambda$$

所以 $\boldsymbol{A}(\lambda)$是可逆的;$\boldsymbol{B}(\lambda)$不是满秩的,当然也不是可逆的;而 $\boldsymbol{D}(\lambda)$满秩,但不可逆。

对 λ 矩阵也可以有初等变换。

定义 1.11　对 λ 矩阵 $\boldsymbol{A}(\lambda)$施行的下列三种变换称为 λ 矩阵的初等行(列)变换:

(1) 变换 $\boldsymbol{A}(\lambda)$的任意两行(列)(交换 i、j 两行(列),记作 $r_i\leftrightarrow r_j(c_i\leftrightarrow c_j)$);