NUMERICAL ALGEBRA

由同顺 编著

数值代数

NUMERICAL ALGEBRA

天津大学出版社
TIANJIN UNIVERSITY PRESS

数 值 代 数

由同顺 编著

天津大学出版社
TIANJIN UNIVERSITY PRESS

内容简介

全书共分为5章,在第1章回顾和补充介绍了线性代数基本内容后,介绍了数值代数的常用计算方法.包括求解线性代数方程组的Gauss消元法、Cholesky分解法(平方根法)、最小二乘法及消元法的舍入误差分析;求解线性代数方程组的迭代方法、最速下降法和共轭梯度法;求解代数特征值问题的各种实用的计算方法:幂法、反幂法、Jacobi方法、二分法和*QR*方法.

本书可作为理工科大学的本科生教材或教学参考书,也可供从事科学与工程计算的有关人员参考.

图书在版编目(CIP)数据

数值代数/由同顺编著.—天津:天津大学出版社,2006.8
ISBN 7-5618-2318-5

Ⅰ.数...　Ⅱ.由...　Ⅲ.线性代数　Ⅳ.0151.2

中国版本图书馆CIP数据核字(2006)第082377号

出版发行　天津大学出版社
出 版 人　杨欢
地　　址　天津市卫津路92号天津大学内(邮编:300072)
电　　话　发行部:022-27403647　邮购部:022-27402742
网　　址　www.tjup.com
短信网址　发送“天大”至916088
印　　刷　昌黎太阳红彩色印刷有限责任公司
经　　销　全国各地新华书店
开　　本　170mm×240mm
印　　张　12
字　　数　263千
版　　次　2006年8月第1版
印　　次　2006年8月第1次
印　　数　1-3 000
定　　价　18.00元

前　　言

随着计算机的飞速发展,数值计算几乎遍布所有的科学技术领域.科学与工程计算已经成为科学研究中的第三种手段.许多诸如油藏模拟、气象预报、航天问题等科学与工程计算问题最终都归结为数值代数计算问题(主要包括大规模线性方程组及代数特征问题的求解问题).因此,数值代数在现代科学与工程计算中占有特殊重要的基础地位.

本书是在作者多年为南开大学数学科学学院本科生讲授"数值代数"课程的过程中经过多次修改后逐步形成的.全书共分为5章,在第1章回顾和补充介绍了线性代数基本内容后,介绍了数值代数的常用计算方法.包括求解线性代数方程组的Gauss消元法、Cholesky分解法(平方根法)、最小二乘法及消元法的舍入误差分析;求解线性代数方程组的迭代方法、最速下降法和共轭梯度法;求解代数特征值问题的各种实用的计算方法:幂法、反幂法、Jacobi方法、二分法和 ***QR*** 方法.

本书在选材上兼顾了理论基础性和实用性,同时也注意到新近的数值方法和算法.在内容编排上,由浅入深,循序渐进,便于自学与教学.本书尽可能采用构造性方法证明定理,以便阐明相关算法的设计理念和理论依据.另外在每章后面提供了较为丰富的习题,以便学生更好地掌握所学内容以及拓展所学知识.

在本书的写作过程中得到南开大学教材建设基金的资助,也得到了南开大学数学科学学院的领导及科学计算系老师们的支持和帮助,天津大学出版社的老师为本书的出版付出了辛勤的劳动,在此,作者表示衷心的感谢.由于作者水平有限,书中难免有错误及不妥之处,敬请读者不吝赐教.

编者

2006.5

目 录

第1章　线性代数基础知识

在本章中，将复习线性代数中的一些重要概念及定义，给出一些常用结果. 在这里只对个别结论给出证明，对于其他结论的证明可在任何一本线性代数教科书中找到. 另外，本章还补充介绍了一些在一般线性代数教材中所不涉及的内容. 本章的内容基本上为以后学习数值线性代数提供了所需的线性代数的内容.

1.1　线性空间和子空间

1.1.1　线性空间

定义 1.1.1　设 V 是一个非空集合，$\mathbf{P}$ 是一个数域（实数域 $\mathbf{R}$ 或复数域 $\mathbf{C}$）. 在集合 V 的元素之间定义了一种称之为加法的代数运算，即给出一个法则，对于 V 中任意两个元素 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$，在 V 中都有唯一的一个元素 $\boldsymbol{\gamma}$ 与它们对应，称 $\boldsymbol{\gamma}$ 为 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的和，记为 $\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{\beta}$. 在数域 $\mathbf{P}$ 与集合 V 的元素之间还定义了一种称之为数量乘法的运算，即对于数域 $\mathbf{P}$ 中任一元素 k 和 V 中任一元素 $\boldsymbol{\alpha}$，在 V 中都有唯一的一个元素 $\boldsymbol{\delta}$ 与它们对应，称 $\boldsymbol{\delta}$ 为 k 与 $\boldsymbol{\alpha}$ 的数量乘积，记为 $\boldsymbol{\delta}=k\boldsymbol{\alpha}$. 如果加法和数量乘法满足如下规则：

(1) $\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\beta}+\boldsymbol{\alpha},\ \forall\,\boldsymbol{\alpha},\boldsymbol{\beta}\in V$；

(2) $(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}=\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma}),\ \forall\,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in V$；

(3) 在 V 中有一个元素 $\mathbf{0}$，对于 V 中任一元素 $\boldsymbol{\alpha}$，有 $\boldsymbol{\alpha}+\mathbf{0}=\boldsymbol{\alpha}$（称元素 $\mathbf{0}$ 为 V 的零元素）；

(4) 对于 V 中任一元素 $\boldsymbol{\alpha}$，都有 V 中元素 $\boldsymbol{\beta}$，使得 $\boldsymbol{\alpha}+\boldsymbol{\beta}=\mathbf{0}$（称 $\boldsymbol{\beta}$ 为 $\boldsymbol{\alpha}$ 的负元素，记为 $-\boldsymbol{\alpha}$）；

(5) $1\boldsymbol{\alpha}=\boldsymbol{\alpha},\ \forall\,\boldsymbol{\alpha}\in V$；

(6) $k(l\boldsymbol{\alpha})=(kl)\boldsymbol{\alpha},\ \forall\,\boldsymbol{\alpha}\in V,\ \forall\,k,l\in\mathbf{P}$；

(7) $(k+l)\boldsymbol{\alpha}=k\boldsymbol{\alpha}+l\boldsymbol{\alpha},\ \forall\,\boldsymbol{\alpha}\in V,\ \forall\,k,l\in\mathbf{P}$；

(8) $k(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta},\ \forall\,\boldsymbol{\alpha},\boldsymbol{\beta}\in V,\ \forall\,k\in\mathbf{P}$.

则称 V 是数域 $\mathbf{P}$ 上的线性空间或向量空间. 线性空间 V 中的元素也称为向量.

例 1.1.1　设 $V=\{\boldsymbol{x}\mid\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}},x_i\in\mathbf{C}\}$，数域 $\mathbf{P}=\mathbf{C}$，则对任意 $\boldsymbol{x},\boldsymbol{y}\in V,k\in\mathbf{C}$，定义

$$\boldsymbol{x}+\boldsymbol{y}=(x_1+y_1,x_2+y_2,\cdots,x_n+y_n)^{\mathrm{T}}$$

$$k\boldsymbol{x}=(kx_1,kx_2,\cdots,kx_n)^{\mathrm{T}}$$

其中 $\boldsymbol{x}^{\mathrm{T}}$ 表示列向量 $\boldsymbol{x}$ 的转置，它是行向量；$\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$.

容易看出 V 对于复数域 $\mathbf{C}$ 满足上面的 8 条运算规则.这样分量属于复数域 $\mathbf{C}$ 的全体 n 元数组 $(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$(列向量)构成复数域上的一个线性空间,称为 n 维复向量空间并记为 $\mathbf{C}^n$.类似地,将例 1.1.1 中的复数域 $\mathbf{C}$ 都换为实数域 $\mathbf{R}$,则得到 n 维实向量空间,记为 $\mathbf{R}^n$.$\mathbf{C}^n$ 和 $\mathbf{R}^n$ 是最常用的两个线性空间,以后用小写字母 $\boldsymbol{x},\boldsymbol{y}$ 等表示 $\mathbf{R}^n$ $(\mathbf{C}^n)$中的列向量.

例 1.1.2 区间$[a,b]$上的全体实连续函数,按函数的加法和数与函数的乘法,构成一个实数域上的线性空间,记为 $C[a,b]$.

定义 1.1.2 设 V 是一个线性空间,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 V 中的一组向量.如果存在数域 $\mathbf{P}$ 中不全为零的数 $k_1,k_2,\cdots,k_n$ 使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0}$$

则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是线性相关的;否则,称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是线性无关的.

定义 1.1.3 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是线性空间 V 中的一线性无关的向量组,并且 V 中的任一向量 $\boldsymbol{\alpha}$ 都可表示成它们的线性组合,即存在数域中的一组数 $k_1,k_2,\cdots,k_n$,使得

$$\boldsymbol{\alpha}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n$$

则称$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$为 V 的一个基底,称 n 元数组 $(k_1,k_2,\cdots,k_n)$为 $\boldsymbol{\alpha}$ 的关于基底$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$的坐标.

关于基底,我们有如下定理.

定理 1.1.1 设 V 是一个以$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$为基底的线性空间,则 V 的任一基底必含有 n 个向量.

定理 1.1.1 说明,线性空间的不同基底中的向量个数是不变的,这个数是线性空间的一个重要特征.

定义 1.1.4 如果线性空间 V 的基底含有 n 个向量,则称 V 是 n 维线性空间,其维数记为 $\dim V=n$;如果 V 中含有无穷多个线性无关的向量,则称 V 是无穷维线性空间;如果线性空间 V 只由零向量构成,则规定其维数为零.

应当指出,本书的以后章节都只研究或涉及有限维线性空间.

例 1.1.3 用 $e_i=(0,\cdots,0,1,0,\cdots,0)^{\mathrm{T}}(i=1,2,\cdots,n)$表示 $\mathbf{R}^n$ 或 $\mathbf{C}^n$ 空间中第 i 个分量为 1,其余分量为 0 的向量,则$\{e_1,e_2,\cdots,e_n\}$为 $\mathbf{R}^n$ 或 $\mathbf{C}^n$ 的一个基底且 $\dim\mathbf{R}^n=\dim\mathbf{C}^n=n$.$\{e_1,e_2,\cdots,e_n\}$称为 $\mathbf{R}^n$ 和 $\mathbf{C}^n$ 的标准基底,e_i 称为单位向量.

1.1.2 线性子空间

定义 1.1.5 设 V 是数域 $\mathbf{P}$ 上的线性空间,W 是 V 的一个非空子集.如果 W 中的向量关于 V 中的加法运算和数量乘法运算也构成数域 $\mathbf{P}$ 上的线性空间,则称 W 是 V 的一个线性子空间或简称子空间.

关于线性空间 V 中的一个非空子集 W 能否成为子空间,可用如下定理来判别.

定理 1.1.2 设 V 是数域 $\mathbf{P}$ 上的线性空间,W 是 V 的一个非空子集,则 W 是 V

的一个子空间的充分必要条件为对任意 $\boldsymbol{\alpha},\boldsymbol{\beta}\in W$,及 $k,l\in\mathbf{P}$,都有 $k\boldsymbol{\alpha}+l\boldsymbol{\beta}\in W$(也即 W 关于 V 的两种运算是封闭的).

由定义易知,$\{0\}$及 V 都是线性空间 V 的子空间,称为 V 的平凡子空间.如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是线性空间 V 中的一组向量,则由定理 1.1.2 可知,由此向量组的所有线性组合所构成的集合

$$\mathrm{span}\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}=\{\sum_{i=1}^{m}k_i\boldsymbol{\alpha}_i\mid k_i\in\mathbf{P},i=1,2,\cdots,m\}$$

是 V 的一个线性子空间,称为由$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$张成的子空间.

另外,根据定理 1.1.2 可知,如果 $V_1,V_2,\cdots,V_m$ 是数域 $\mathbf{P}$ 上的线性空间 V 的子空间,则它们的交 $V_1\cap V_2\cap\cdots\cap V_m=\bigcap_{i=1}^{m}V_i$ 也是 V 的子空间;它们的和

$$V_1+V_2+\cdots+V_m=\{\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_m\mid\boldsymbol{\alpha}_i\in V_i,i=1,2,\cdots,m\}$$

也是 V 的子空间.和空间 $V_1+V_2+\cdots+V_m$ 可简记为 $\sum_{i=1}^{m}V_i$.

定义 1.1.6　设 $V_1,V_2,\cdots,V_m$ 是数域 $\mathbf{P}$ 上的线性空间 V 的子空间.如果和 $\sum_{i=1}^{m}V_i$ 中每一个向量 $\boldsymbol{\alpha}$ 的的分解式

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_i\in V_i(i=1,2,\cdots,m)$$

是唯一的,则称该和为直和,记为

$$V_1\oplus V_2\oplus\cdots\oplus V_m=\bigoplus_{i=1}^{m}V_i$$

关于直和,我们有如下定理.

定理 1.1.3　设 $V_1,V_2,\cdots,V_m$ 是数域 $\mathbf{P}$ 上的线性空间 V 的子空间,则下面三个条件等价.

(1) $W=\sum_{i=1}^{m}V_i$ 是直和;

(2) $V_i\cap\sum_{j\neq i}V_j=\{0\}(i=1,2,\cdots,m)$;

(3) $\dim W=\sum_{i=1}^{m}\dim V_i$.

特别,如果线性空间 $V=V_1\oplus V_2$,则称子空间 V_1 和 V_2 互补,即 V_1 是 V_2 的补空间,V_2 也是 V_1 的补空间.

定理 1.1.4　设 V_1 是数域 $\mathbf{P}$ 上的线性空间 V 的一个子空间,则必存在 V_1 的补空间 V_2,使得 $V=V_1\oplus V_2$.

定理 1.1.4 表明,线性空间的任一线性无关组必能扩充为它的整个基底.显然这样的扩充不唯一.

1.1.3　内积空间

在本段中,把空间解析几何中的向量长度和正交的概念推广到有限维线性空间.为

此,引入内积的概念.

定义 1.1.7 设 V 是实数域 $\mathbf{R}$ 上的线性空间,所谓 V 上的一个内积是指定义在 V 上的一个二元函数,记其为 $(\boldsymbol{x},\boldsymbol{y})$,它对一切 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in V$ 以及 $k\in\mathbf{R}$,满足

(1) $(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{y},\boldsymbol{x})$;

(2) $(k\boldsymbol{x},\boldsymbol{y})=k(\boldsymbol{x},\boldsymbol{y})$;

(3) $(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{z})+(\boldsymbol{y},\boldsymbol{z})$;

(4) $(\boldsymbol{x},\boldsymbol{x})\geqslant 0$, 当且仅当 $\boldsymbol{x}=\mathbf{0}$ 时,等号成立.

定义了实内积的线性空间称为实内积空间.如果把定义 1.1.7 中的实数域 $\mathbf{R}$ 换为复数域 $\mathbf{C}$ 以及把(1)换为 $(\boldsymbol{x},\boldsymbol{y})=\overline{(\boldsymbol{y},\boldsymbol{x})}$,则 V 上的一个二元复函数 $(\boldsymbol{x},\boldsymbol{y})$ 就定义了 $\mathbf{C}$ 上的一个内积.我们称定义了复内积的线性空间为复内积空间.另外,设 $\boldsymbol{x}$ 是内积空间 V 中的任一向量,我们称 $\|\boldsymbol{x}\|=(\boldsymbol{x},\boldsymbol{x})^{\frac{1}{2}}$ 为向量 $\boldsymbol{x}$ 的范数.如果 V 中两个向量 $\boldsymbol{x},\boldsymbol{y}$ 满足 $(\boldsymbol{x},\boldsymbol{y})=0$,则称 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交.根据正交概念,引入如下定义.

定义 1.1.8 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是内积空间 V 中的一组非零向量,如果其中任意两个向量均正交,即 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=0\ (i\neq j)$,则称向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$ 是正交向量组.特别地,如果

$$(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=\delta_{ij}\equiv\begin{cases}1 & 若\ i=j\\ 0 & 若\ i\neq j\end{cases}\tag{1.1.1}$$

则称向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\}$ 是标准正交向量组.

易知,$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_iy_i$ 及 $(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_i\bar{y}_i$ 分别为向量空间 $\mathbf{R}^n$ 和 $\mathbf{C}^n$ 的内积,相应的内积空间分别称为 n 维欧氏空间及酉空间.范数 $\|\boldsymbol{x}\|_2=(\sum_{i=1}^{n}|x_i|^2)^{\frac{1}{2}}$ 称为欧几里得范数.显然,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 为 $\mathbf{R}^n$ 及 $\mathbf{C}^n$ 的一组标准正交基.

最后,引入正交补空间的概念.设 S 是内积空间 V 的一个子空间,令

$$S^{\perp}=\{\boldsymbol{x}\mid(\boldsymbol{x},\boldsymbol{y})=0,\boldsymbol{x}\in V,\forall\,\boldsymbol{y}\in S\}$$

则容易知道,$S^{\perp}$ 是内积空间 V 的一个子空间,称其为 S 的正交补空间.关于正交补空间,我们有如下定理.

定理 1.1.5 设 S 是内积空间 V 的任一子空间,则一定存在 S 的正交补空间 $S^{\perp}$,使得 $V=S\oplus S^{\perp}$.

1.2 矩阵及其运算

1.2.1 矩阵的定义

称由 $m\times n$ 个实数(或复数)排成 m 行及 n 列的数表

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

为一个 $m\times n$ 矩阵，它也可简记为$(a_{ij})_{m\times n}$或(a_{ij})，a_{ij}称为矩阵A 的(i,j)元素. 当 $m=n$ 时，$n\times n$ 矩阵也称为n 阶矩阵或n 阶方阵.

显然，$\mathbf{R}^n$或$\mathbf{C}^n$中的列向量就是$n\times 1$ 阶矩阵. 今后我们用大写字母 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{X},\boldsymbol{Y}$ 等表示矩阵，用小写字母 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{x},\boldsymbol{y}$ 等表示向量.

用 $\boldsymbol{I}_n$表示n 阶单位矩阵，其(i,j)元素为 δ_{ij}. 在不引起混淆的情况下，常常把单位矩阵简记为 $\boldsymbol{I}$，用 $\boldsymbol{O}$ 表示零矩阵，其元素全为零. 这样我们用 0 表示数字零、用 $\mathbf{0}$ 表示零向量以及用 $\boldsymbol{O}$ 表示零矩阵，用 $\boldsymbol{A}^{\mathrm{T}}$和 $\boldsymbol{A}^*$ 分别表示矩阵 $\boldsymbol{A}$ 的转置矩阵和共轭转置矩阵，即若 $\boldsymbol{A}=(a_{ij})_{m\times n}$，则 $\boldsymbol{A}^{\mathrm{T}}$和 $\boldsymbol{A}^*$ 都为 $n\times m$ 阶矩阵，它们的 (i,j)元素分别为 a_{ji}和 $\bar{a}_{ji}$，其中 $\bar{a}_{ji}$为a_{ji}的共轭复数.

1.2.2　矩阵的运算

1. 矩阵的加法与数乘

用$\mathbf{C}^{m\times n}$($\mathbf{R}^{m\times n}$)表示复数域(实数域)上 $m\times n$ 矩阵的全体. 在 $\mathbf{C}^{m\times n}$ 中，可定义矩阵的加法和数乘，即对 $\mathbf{C}^{m\times n}$ 中的任意两个矩阵$\boldsymbol{A}$ 和$\boldsymbol{B}$，$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和$\boldsymbol{P}=\boldsymbol{A}+\boldsymbol{B}$ 为 $m\times n$ 矩阵，其元素为 $c_{ij}=a_{ij}+b_{ij}$；数 λ 与 $\boldsymbol{A}$ 的数乘为$\boldsymbol{P}=\lambda\boldsymbol{A}$，其元素为 $p_{ij}=\lambda a_{ij}$. 容易验证矩阵的加法和数乘满足线性空间的定义，这样 $\mathbf{C}^{m\times n}$($\mathbf{R}^{m\times n}$)是线性空间.

设 $\boldsymbol{E}_{ij}\in\mathbf{R}^{m\times n}(\mathbf{C}^{m\times n})$，它的$(i,j)$元素为 1，其余元素都为 0，则集合

$$\{\boldsymbol{E}_{ij}\mid i=1,2,\cdots,m,j=1,2,\cdots,n\}$$

为 $\mathbf{R}^{m\times n}(\mathbf{C}^{m\times n})$的一个基底. 这样 $\dim\mathbf{R}^{m\times n}=\dim\mathbf{C}^{m\times n}=mn$. 显然，如果 $\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{m\times n}(\mathbf{C}^{m\times n})$，则 $\boldsymbol{A}=\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}a_{ij}\boldsymbol{E}_{ij}$.

2. 矩阵的乘法

设 $\boldsymbol{A}=(a_{ij})_{m\times n}$，$\boldsymbol{B}=(b_{ij})_{n\times p}$，则定义积 $\boldsymbol{C}=\boldsymbol{AB}$ 为 $m\times p$ 矩阵且其元素为

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}\,(i=1,2,\cdots,m,j=1,2,\cdots,p)$$

由矩阵乘法定义，易知$(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$. 利用矩阵乘法，则在上节定义的 $\mathbf{R}^n$ 和 $\mathbf{C}^n$ 的内积可分别表示为$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}$ 及$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^*\boldsymbol{x}$.

将矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$分块写为$\boldsymbol{A}=(\boldsymbol{A}_{pq})_{s\times t}$，其中子矩阵 $\boldsymbol{A}_{pq}$的阶数为$m_p\times n_q$($p=1,2,\cdots,s$；$q=1,2,\cdots,t$)，显然有$\sum\limits_{p=1}^{s}m_p=m$ 及$\sum\limits_{q=1}^{t}n_q=n$. 对于分块矩阵的加法和乘法，只要剖分恰当，所分块的子阵可以整体地进行运算. 本节只对矩阵的分块乘法给

以说明.

设 $\boldsymbol{A}$ 为 $m\times n$ 阶矩阵,$\boldsymbol{B}$ 为 $n\times l$ 阶矩阵且 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的分块形式为 $\boldsymbol{A}=(\boldsymbol{A}_{pk})_{s\times t}$,$\boldsymbol{B}=(\boldsymbol{B}_{kq})_{t\times r}$,如果子阵 $\boldsymbol{A}_{pk}$ 和 $\boldsymbol{B}_{kq}$($1\leqslant p\leqslant s,1\leqslant k\leqslant t,1\leqslant q\leqslant r$)可乘,则 $\boldsymbol{AB}=\boldsymbol{C}=(\boldsymbol{C}_{pq})_{s\times r}$,这里 $\mathbf{C}_{pq}=\sum\limits_{k=1}^{t}\boldsymbol{A}_{pk}\boldsymbol{B}_{kq}$($p=1,2,\cdots,s;q=1,2,\cdots,r$).特别,如果矩阵 $\boldsymbol{B}$ 按列分块,即

$$\boldsymbol{B}=(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_l)\ (\boldsymbol{b}_i\text{ 为 }\boldsymbol{B}\text{ 的第 }i\text{ 列构成的列向量})$$

则
$$\boldsymbol{AB}=\boldsymbol{A}(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_l)=(\boldsymbol{Ab}_1,\boldsymbol{Ab}_2,\cdots,\boldsymbol{Ab}_l)$$

类似地,如果矩阵 $\boldsymbol{A}$ 按行分块且记 $\boldsymbol{A}$ 的第 i 行为 $\boldsymbol{a}_i^{\mathrm{T}}$($i=1,2,\cdots,m$),则

$$\boldsymbol{AB}=\begin{bmatrix}\boldsymbol{a}_1^{\mathrm{T}}\\ \boldsymbol{a}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{a}_m^{\mathrm{T}}\end{bmatrix}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{B}\\ \boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{B}\\ \vdots\\ \boldsymbol{a}_m^{\mathrm{T}}\boldsymbol{B}\end{bmatrix}$$

下面讨论矩阵的秩和逆,为此给出它们的定义.

定义 1.2.1 设 $\boldsymbol{A}$ 为 $m\times n$ 阶矩阵,称矩阵 $\boldsymbol{A}$ 的列向量组中的最大线性无关组的向量个数为矩阵 $\boldsymbol{A}$ 的秩,记为 $\mathrm{rank}(\boldsymbol{A})$.特别,当 $\mathrm{rank}(\boldsymbol{A})=n$ 时,称 $\boldsymbol{A}$ 为列满秩.

应该注意到 $\mathrm{rank}(\boldsymbol{A}^{\mathrm{T}})=\mathrm{rank}(\boldsymbol{A})$,所以也可由矩阵 $\boldsymbol{A}$ 的行向量组中的最大线性无关组的向量个数定义矩阵 $\boldsymbol{A}$ 的秩.类似地,当 $\mathrm{rank}(\boldsymbol{A})=m$ 时,称 $\boldsymbol{A}$ 为行满秩.

定义 1.2.2 设 $\boldsymbol{A}$ 为 n 阶矩阵,如果存在 n 阶矩阵 $\boldsymbol{B}$ 使得 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$,则称 $\boldsymbol{A}$ 是可逆的或非奇异的,$\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆矩阵且记其为 $\boldsymbol{B}=\boldsymbol{A}^{-1}$.

我们用 $\det(\boldsymbol{A})$ 表示矩阵 $\boldsymbol{A}$ 的行列式,关于行列式的定义及性质,在这里不作赘述,请读者参阅有关教材.关于可逆矩阵,我们有如下结论:

(1)设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶可逆矩阵,则 $(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ 及 $(\boldsymbol{A}^{\mathrm{T}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}$;

(2)n 阶矩阵 $\boldsymbol{A}$ 可逆等价于矩阵 $\boldsymbol{A}$ 的秩等于 n(此时称 $\boldsymbol{A}$ 为满秩矩阵),也等价于 $\det(\boldsymbol{A})\neq 0$.

1.2.3 几类常用矩阵

设 $\boldsymbol{A}=(a_{ij})$ 为 n 阶方阵,如果 $a_{ij}=0(i>j)$,则称 $\boldsymbol{A}$ 为上三角阵;如果 $a_{ij}=0(i<j)$,则称 $\boldsymbol{A}$ 为下三角阵;如果 $a_{ij}=0(i\neq j)$,则称 $\boldsymbol{A}$ 为对角阵,记其为 $\boldsymbol{A}=\mathrm{diag}(a_{11},a_{22},\cdots,a_{nn})$.对于上(下)三角阵 $\boldsymbol{A}$,如果主对角线元素 $a_{ii}=1(i=1,2,\cdots,n)$,则称 $\boldsymbol{A}$ 为单位上(下)三角阵.易知,上(下)三角阵的逆仍为上(下)三角阵.

最后,给出正定矩阵的定义及简单性质.

定义 1.2.3 设 $\boldsymbol{A}$ 为 n 阶实矩阵,如果 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为实对称矩阵;如果 $\boldsymbol{A}$ 为对称矩阵且对于任意非零向量 $\boldsymbol{x}\in\mathbf{R}^n$,都有 $(\boldsymbol{Ax},\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}>0(\geqslant 0)$,则称 $\boldsymbol{A}$ 为正定矩阵(半正定矩阵);如果 $-\boldsymbol{A}$ 为正定矩阵,则称 $\boldsymbol{A}$ 为负定矩阵.

关于正定矩阵,有如下常用判别定理.

定理 1.2.1　n 阶实对称矩阵 $\boldsymbol{A}$ 是正定的充分必要条件为矩阵 $\boldsymbol{A}=(a_{ij})$ 的所有顺序主子式

$$\Delta_1=a_{11},\Delta_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\cdots,\Delta_n=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

都大于零.

1.3　值域与核空间

1.3.1　值域与核空间的定义

定义 1.3.1　设 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$,称集合 $R(\boldsymbol{A})=\{\boldsymbol{y}\in\mathbf{R}^m \mid \boldsymbol{y}=\boldsymbol{A}\boldsymbol{x},\boldsymbol{x}\in\mathbf{R}^n\}$ 为矩阵 $\boldsymbol{A}$ 的值域或像空间,也记为 $\mathrm{Im}(\boldsymbol{A})$;称集合 $\ker(\boldsymbol{A})=\{\boldsymbol{x}\in\mathbf{R}^n \mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\}$ 为 $\boldsymbol{A}$ 的核空间.

容易验证,$R(\boldsymbol{A})$ 为 $\mathbf{R}^m$ 的子空间,$\ker(\boldsymbol{A})$ 为 $\mathbf{R}^n$ 的子空间,且有

$$R(\boldsymbol{A})=\mathrm{span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n\}$$

其中 $\boldsymbol{a}_i(i=1,2,\cdots,n)$ 为 $\boldsymbol{A}$ 的第 i 列. 显然,$\dim R(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A})$.

关于值域与核空间的维数关系,有如下定理.

定理 1.3.1　设 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$,则 $\dim R(\boldsymbol{A})+\dim\ker(\boldsymbol{A})=n$.

证　由于 $\ker(\boldsymbol{A})\subset\mathbf{R}^n$,所以可把 $\ker(\boldsymbol{A})$ 的一组基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s\}$ 扩充为 $\mathbf{R}^n$ 的一组基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\alpha}_{s+1},\cdots,\boldsymbol{\alpha}_n\}$. 设 $\boldsymbol{x}$ 为 $\mathbf{R}^n$ 中任一向量,则 $\boldsymbol{x}=\sum\limits_{i=1}^{n}k_i\boldsymbol{\alpha}_i,k_i\in\mathbf{R}(i=1,2,\cdots,n)$. 注意到 $\boldsymbol{A}\boldsymbol{\alpha}_i=\boldsymbol{0}(i=1,2,\cdots,s)$,则

$$\boldsymbol{A}\boldsymbol{x}=\sum_{i=1}^{n}k_i\boldsymbol{A}\boldsymbol{\alpha}_i=\sum_{i=s+1}^{n}k_i\boldsymbol{A}\boldsymbol{\alpha}_i$$

因此 $R(\boldsymbol{A})=\mathrm{span}\{\boldsymbol{A}\boldsymbol{\alpha}_{s+1},\boldsymbol{A}\boldsymbol{\alpha}_{s+2},\cdots,\boldsymbol{A}\boldsymbol{\alpha}_n\}$,故 $r=\dim R(\boldsymbol{A})\leqslant n-s$.

现证明 $\boldsymbol{A}\boldsymbol{\alpha}_{s+1},\boldsymbol{A}\boldsymbol{\alpha}_{s+2},\cdots,\boldsymbol{A}\boldsymbol{\alpha}_n$ 线性无关,令 $\sum\limits_{i=s+1}^{n}k_i\boldsymbol{A}\boldsymbol{\alpha}_i=\boldsymbol{0}$,则有 $\boldsymbol{A}\left(\sum\limits_{i=s+1}^{n}k_i\boldsymbol{\alpha}_i\right)=\boldsymbol{0}$. 所以存在 $k_i\in\mathbf{R}(i=1,2,\cdots,s)$,使得

$$\sum_{i=s+1}^{n}k_i\boldsymbol{\alpha}_i=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s$$

根据 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的线性无关性,可得 $k_i=0(i=1,2,\cdots,n)$.

故 $\boldsymbol{A}\boldsymbol{\alpha}_{s+1},\boldsymbol{A}\boldsymbol{\alpha}_{s+2},\cdots,\boldsymbol{A}\boldsymbol{\alpha}_n$ 线性无关. 因此定理得证.

由定理 1.3.1 容易得到如下推论.

推论 1.3.1 设 $A \in \mathbf{R}^{m\times n}$ 且 $m<n$，则 $\dim\ker(\boldsymbol{A})>0$，也即齐次方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 有非零解.

定理 1.3.2 设 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$，则 $R(\boldsymbol{A})^{\perp}=\ker(\boldsymbol{A}^{\mathrm{T}})$.

证 设 $\boldsymbol{z}\in R(\boldsymbol{A})^{\perp}$，则有 $(\boldsymbol{A}\boldsymbol{x},\boldsymbol{z})=0(\forall\,\boldsymbol{x}\in\mathbf{R}^n)$. 所以 $(\boldsymbol{x},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z})=0$. 由 $\boldsymbol{x}$ 的任意性可知，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}=\boldsymbol{0}$. 故 $R(\boldsymbol{A})^{\perp}\subset\ker(\boldsymbol{A}^{\mathrm{T}})$.

另一方面，设 $\boldsymbol{z}\in\ker(\boldsymbol{A}^{\mathrm{T}})$，则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}=\boldsymbol{0}$. 所以对任意 $\boldsymbol{x}\in\mathbf{R}^n$，有

$$0=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z},\boldsymbol{x})=(\boldsymbol{z},\boldsymbol{A}\boldsymbol{x})$$

从而 $\boldsymbol{z}\in R(\boldsymbol{A})^{\perp}$. 故 $\ker(\boldsymbol{A}^{\mathrm{T}})\subset R(\boldsymbol{A})^{\perp}$，因此定理得证.

1.3.2 线性方程组的解

设线性方程组为

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \tag{1.3.1}$$

其中 $\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{m\times n}$ 为系数矩阵，$\boldsymbol{b}=(b_1,b_2,\cdots,b_m)^{\mathrm{T}}$ 为右端项，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 为未知量.

下面讨论方程组(1.3.1)的解存在的条件以及解的结构.

定理 1.3.3 线性方程组(1.3.1)的解存在的充分必要条件是 $\boldsymbol{A}$ 与其增广矩阵 $(\boldsymbol{A},\boldsymbol{b})$ 的秩相等.

证 必要性：设方程组有解，则 $\boldsymbol{b}$ 可由 $\boldsymbol{A}$ 的列向量 $\boldsymbol{a}_i$ 线性表示，所以

$$\mathrm{span}\{\boldsymbol{a}_1,\boldsymbol{a}_2\cdots,\boldsymbol{a}_n\}=\mathrm{span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n,\boldsymbol{b}\}$$

故 $$\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}((\boldsymbol{A},\boldsymbol{b})).$$

充分性：$\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}((\boldsymbol{A},\boldsymbol{b}))$，则 $\boldsymbol{b}$ 可由 $\boldsymbol{A}$ 的列向量 $\boldsymbol{a}_i$ 线性表示，即 $\boldsymbol{b}=\sum_{i=1}^{n}x_i\boldsymbol{a}_i$. 所以 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 为方程组 1.3.1 的解.

定理 1.3.4 设 $\boldsymbol{x}^*$ 为方程组 1.3.1 的特解，则方程组 1.3.1 的解集合为

$$\{\boldsymbol{x}^*\}+\ker(\boldsymbol{A})\equiv\{\boldsymbol{x}^*+\boldsymbol{y}\mid\boldsymbol{y}\in\ker(\boldsymbol{A})\}$$

且方程组 1.3.1 有唯一解的充分必要条件为 $\ker(\boldsymbol{A})=\{\boldsymbol{0}\}$.

证 设 $\boldsymbol{x}$ 为方程组 1.3.1 的任一解，则 $\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x}^*)=\boldsymbol{0}$. 故 $\boldsymbol{x}\in\{\boldsymbol{x}^*\}+\ker(\boldsymbol{A})$.

另一方面，显然 $\{\boldsymbol{x}^*\}+\ker(\boldsymbol{A})$ 中的任一元素都是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解. 因此方程组 1.3.1 的解集合为 $\{\boldsymbol{x}^*\}+\ker(\boldsymbol{A})$. 由此可知，方程组 1.3.1 有唯一解的充分必要条件为 $\ker(\boldsymbol{A})=\{\boldsymbol{0}\}$.

由上面的讨论可知，对于 $m=n$ 的情况，即系数矩阵 $\boldsymbol{A}$ 为方阵，方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有唯一解的充分必要条件为 $\mathrm{rank}(\boldsymbol{A})=n$（或 $\det(\boldsymbol{A})\neq 0$）. 由于用 Cramer 法则求解方程组是熟知的结果，所以在这里不再赘述.

1.4　矩阵的特征值

1.4.1　矩阵的特征值的定义

定义 1.4.1　设 $A\in C^{n\times n}$，如果存在 $\lambda\in C$，使方程组 $Ax=\lambda x$ 有非零解，则称 λ 为 A 的特征值，相应的非零解 $x\in C^n$ 称为矩阵 A 关于 λ 的特征向量.

显然，λ 为 A 的特征值的充分必要条件是

$$f_A(\lambda)\equiv\det(\lambda I-A)=0$$

称 $f_A(\lambda)$ 为 A 的特征多项式. 因此 A 的特征值 λ 正是其特征多项式 $f_A(\lambda)$ 的根. 由于 $f_A(\lambda)$ 是 n 次多项式，根据代数学基本定理可知，它在复数域上恰有 n 个根，也即，n 阶矩阵 A 在复数域上恰有 n 个特征值(其中重特征值按重数计算). 称由 A 的 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 构成的集合为 A 的谱，记为 $\sigma(A)$. 称

$$\rho(A)=\max_{1\leqslant i\leqslant n}|\lambda_i|$$

为 A 的谱半径.

由于 $\det(\lambda I-A^{\mathrm T})=\det(\lambda I-A)$，所以 A 与 $A^{\mathrm T}$ 有相同的特征值. 称矩阵 $B=S^{-1}AS$ 与 A 相似，其中 S 为可逆矩阵. 易知，相似矩阵有相同的特征值. 下面讨论矩阵 A 的特征值与其特征多项式的系数之间的常用关系式. 由行列式的展开式可知

$$f_A(\lambda)=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+(1\leqslant\text{次数}\leqslant n-2\text{ 的项})+(-1)^n\det(A)$$

而

$$f_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

所以

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

其中 $\mathrm{tr}(A)=\sum\limits_{i=1}^{n}a_{ii}$ 称为矩阵 A 的迹.

1.4.2　特征值的重数及特征向量系

定义 1.4.2　设 $A\in C^{n\times n}$，由关于 A 的特征值 λ 的所有特征向量及零向量构成的集合，称为 A 的对应于 λ 的特征子空间，记为 V_λ.

显然，特征子空间 V_λ 就是 $A-\lambda I$ 的核空间 $\ker(A-\lambda I)$，且它是 A 的不变子空间，即对于 V_λ 中的任一向量 v，都有 $Av\in V_\lambda$. 我们称特征子空间 V_λ 的维数 $\dim V_\lambda$ 为特征值 λ 的几何重数，而称 λ 在特征多项式 $f_A(\lambda)$ 中根的重数为特征值 λ 的代数重数. 关于特征值的几何重数和代数重数，有如下定理.

定理 1.4.1　n 阶矩阵 A 的特征值的几何重数小于或等于它的代数重数.

证　设 λ 为 A 的特征值，且其几何重数为 k，则存在 k 个线性无关的特征向量 v_1，

$\boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$，使得 $\boldsymbol{A}\boldsymbol{v}_i = \lambda \boldsymbol{v}_i (i=1,2,\cdots,k)$. 现把 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$ 扩充为 $\mathbf{C}^n$ 的一组基 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \cdots, \boldsymbol{v}_n$. 记 $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n)$，则矩阵 $\boldsymbol{V}$ 非奇异. 易知

$$\boldsymbol{V}^{-1}(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k) = \begin{bmatrix} \boldsymbol{I}_k \\ \boldsymbol{0} \end{bmatrix}$$

所以

$$\begin{aligned} \boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} &= \boldsymbol{V}^{-1}(\boldsymbol{A}\boldsymbol{v}_1, \boldsymbol{A}\boldsymbol{v}_2, \cdots, \boldsymbol{A}\boldsymbol{v}_k, \boldsymbol{A}\boldsymbol{v}_{k+1}, \cdots, \boldsymbol{A}\boldsymbol{v}_n) \\ &= \boldsymbol{V}^{-1}(\lambda\boldsymbol{v}_1, \lambda\boldsymbol{v}_2, \cdots, \lambda\boldsymbol{v}_k, \boldsymbol{A}\boldsymbol{v}_{k+1}, \cdots, \boldsymbol{A}\boldsymbol{v}_n) \\ &= \boldsymbol{V}^{-1}(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k, \boldsymbol{A}\boldsymbol{v}_{k+1}, \cdots, \boldsymbol{A}\boldsymbol{v}_n) \begin{bmatrix} \lambda\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{n-k} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{V}_{12} \\ \boldsymbol{0} & \boldsymbol{V}_{22} \end{bmatrix} \begin{bmatrix} \lambda\boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{n-k} \end{bmatrix} = \begin{bmatrix} \lambda\boldsymbol{I}_k & \boldsymbol{V}_{12} \\ \boldsymbol{0} & \boldsymbol{V}_{22} \end{bmatrix} \end{aligned}$$

其中

$$\begin{bmatrix} \boldsymbol{V}_{12} \\ \boldsymbol{V}_{22} \end{bmatrix} = \boldsymbol{V}^{-1}\boldsymbol{A}(\boldsymbol{v}_{k+1}, \boldsymbol{v}_{k+2}, \cdots, \boldsymbol{v}_n)$$

由上式可知，$\boldsymbol{A}$ 的特征值 λ 的代数重数至少为 k，故定理得证.

推论 1.4.1 如果 λ 是 $\boldsymbol{A}$ 的特征多项式的单根，则 λ 的几何重数为 1.

关于互异特征值所对应的特征向量有如下关系.

定理 1.4.2 n 阶矩阵 $\boldsymbol{A}$ 的互异特征值所对应的特征向量是线性无关的.

定理 1.4.2 的证明请参阅线性代数的教材，此定理表明，如果矩阵 $\boldsymbol{A}$ 有 n 个互异特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$，则它必有 n 个线性无关的特征向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$. 记 $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n)$，$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$，则有 $\boldsymbol{A}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda}$ 或 $\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \boldsymbol{\Lambda}$. 所以 $\boldsymbol{A}$ 相似于对角矩阵.

显然，n 阶矩阵 $\boldsymbol{A}$ 未必有 n 个线性无关的特征向量，但它的线性无关的特征向量的最大个数 r 必满足 $r \leqslant n$. 称由这 r 个线性无关的特征向量组成的向量组为矩阵 $\boldsymbol{A}$ 的特征向量系. 如果 $r < n$，则称 $\boldsymbol{A}$ 是亏损矩阵；如果 $r = n$，则称 $\boldsymbol{A}$ 具有完全特征向量系.

根据上面的讨论，容易得到如下定理.

定理 1.4.3 n 阶矩阵 $\boldsymbol{A}$ 具有完全特征向量系 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 的一个充分必要条件为

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

其中 $\lambda_i (i=1,2,\cdots,n)$ 为 $\boldsymbol{A}$ 的特征值，矩阵 $\boldsymbol{V} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n)$ 的第 i 列为 $\boldsymbol{A}$ 的关于 λ_i 的特征向量.

1.5　矩阵的标准形

由上节可知,亏损矩阵不能相似于对角矩阵.例如

$$A=\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}$$

易知,它只有一个线性无关的特征向量,所以它不能相似于对角矩阵.

为了讨论一般矩阵 $\boldsymbol{A}$ 的相似矩阵的最简单形式,先讨论正交矩阵和酉矩阵.

定义 1.5.1　如果 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 且 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{I}$,则称 $\boldsymbol{A}$ 为正交矩阵;如果 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 且 $\boldsymbol{A}^*\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{I}$,则称 $\boldsymbol{A}$ 为酉矩阵.

如果记正交矩阵 $\boldsymbol{Q}$ 的各列为 $\boldsymbol{q}_i(i=1,2,\cdots,n)$,则由 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{I}$ 可知,$(\boldsymbol{q}_i,\boldsymbol{q}_j)=\delta_{ij}$,即 $\{\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_n\}$ 为 $\mathbf{R}^n$ 的一个标准正交基,类似地,酉矩阵的各列也构成 $\mathbf{C}^n$ 的一个标准正交基.显然,正交矩阵的逆等于它的转置,酉矩阵的逆等于它的共轭转置.下一定理给出了正交矩阵的其他性质.

定理 1.5.1　(1)正交矩阵的特征值的模为 1.

(2)正交矩阵的行列式为 ±1.

(3)正交矩阵的乘积仍为正交矩阵.

证　(1)设 $\lambda\in\mathbf{C}$ 为正交矩阵 $\boldsymbol{Q}$ 的任一特征值,则存在非零向量 $\boldsymbol{x}\in\mathbf{C}^n$,使得 $\boldsymbol{Qx}=\lambda\boldsymbol{x}$, 所以

$$(\boldsymbol{Qx},\boldsymbol{Qx})=(\lambda\boldsymbol{x},\lambda\boldsymbol{x})=|\lambda|^2\|\boldsymbol{x}\|_2^2$$

而
$$(\boldsymbol{Qx},\boldsymbol{Qx})=\boldsymbol{x}^*\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Qx}=\|\boldsymbol{x}\|_2^2$$

因此 $|\lambda|=1$.

(2)注意到

$$1=\det(\boldsymbol{I})=\det(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q})=\det(\boldsymbol{Q}^{\mathrm{T}})\det(\boldsymbol{Q})=(\det(\boldsymbol{Q}))^2$$

则
$$\det(\boldsymbol{Q})=\pm1$$

(3)设 $\boldsymbol{A},\boldsymbol{B}$ 为正交矩阵,则

$$(\boldsymbol{AB})^{\mathrm{T}}(\boldsymbol{AB})=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AB}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=\boldsymbol{I}$$

同样
$$(\boldsymbol{AB})(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{I}$$

故 $\boldsymbol{AB}$ 为正交矩阵.

定理 1.5.1 关于酉矩阵仍成立,只是结论 2 改为酉矩阵的行列式的模为 1.用 $\boldsymbol{I}_{ij}$ 表示置换阵,它是将单位矩阵 $\boldsymbol{I}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)$ 的第 i 列与第 j 列交换位置所得的矩阵,即

$$\boldsymbol{I}_{ij}=(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_j,\cdots,\boldsymbol{e}_i,\cdots,\boldsymbol{e}_n)$$

显然,$\boldsymbol{I}_{ij}^{\mathrm{T}}=\boldsymbol{I}_{ij}$ 且 $\boldsymbol{I}_{ij}^{\mathrm{T}}\boldsymbol{I}_{ij}=\boldsymbol{I}$,所以 $\boldsymbol{I}_{ij}$ 为正交阵.

下一定理证明了任一矩阵 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 酉相似于上三角矩阵,即存在酉矩阵 $\boldsymbol{U}$,使得 $\boldsymbol{U}^*\boldsymbol{AU}$ 为上三角阵.虽然这种形式不是唯一的,但是它是在酉相似条件下所能得到的