矩阵理论学习指导

黄廷祝 杨传胜 主编

清华大学出版社

ISBN 978-7-302-23741-9
9 787302 237419 >
定价：19.00元

矩阵理论学习指导

黄廷祝 杨传胜 主编

清华大学出版社
北京

内 容 简 介

本书对矩阵理论的基本概念、主要结论等作了简明扼要的分类总结，针对每章主要内容给出了典型例题分析，并对各章的课后习题做了详细的解答，最后提供了6套复习题及相应解答以便读者自测参考.

本书叙述简明，概括性强，可作为理工科研究生和本科高年级学生学习“矩阵理论”课程的辅导书，也可作为从事矩阵理论教学工作的教师和有关科技工作者的参考书.

图书在版编目（CIP）数据

矩阵理论学习指导/黄廷祝，杨传胜主编. --北京：清华大学出版社，2010.10
ISBN 978-7-302-23741-9

Ⅰ.①矩…　Ⅱ.①黄…②杨…　Ⅲ.①矩阵－理论－自学参考资料　Ⅳ.①O151.21

中国版本图书馆 CIP 数据核字(2010)第 167878 号

责任编辑：陈　明
责任校对：王淑云
责任印制：孟凡玉

出版发行：清华大学出版社　　**地　址**：北京清华大学学研大厦 A 座
http://www.tup.com.cn　　**邮　编**：100084
社 总 机：010-62770175　　**邮　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：北京鑫海金澳胶印有限公司
经　销：全国新华书店
开　本：185×230　**印　张**：9.75　**字　数**：208 千字
版　次：2010 年 10 月第 1 版　**印　次**：2010 年 10 月第 1 次印刷
印　数：1～3000
定　价：19.00 元

产品编号：038977-01

前言

矩阵理论是面向理工科研究生和数学专业高年级本科生开设的一门数学基础课，它在数值分析、最优化方法、微分方程、控制理论等分支及各种工程学科有极其重要的应用. 矩阵理论已成为科技领域中不可缺少的数学工具. 因此，学习和掌握矩阵的基本理论与方法，对于初学者来说是必不可少的.

本书是研究生教材《矩阵理论》(黄廷祝、钟守铭、李正良，高等教育出版社，2003)一书的学习指导书，由相应章节内容的基本概念、主要结论、典型例题、习题解答及复习题五部分组成，书中给出了教材中所有习题的详细解答，其目的在于帮助读者学习矩阵的基本理论、掌握解题方法、提高解题技巧以及了解相关学科中的应用等.

本书对于学习矩阵理论课程的研究生、本科生以及参加博士生入学矩阵理论考试的有关人员有很好的辅导作用，对于从事矩阵理论教学的教师也有一定的参考价值.

本书由黄廷祝、杨传胜主编，具体执笔者为：李厚彪(第一章)、程光辉(第二章及复习题)、高中喜(第三、四章)、龚丽莎(第五、六章)和王转德(第七章)等.

限于水平，书中难免存在不当甚至错误之处，敬请读者批评指正，以期今后改正.

编者

电子科技大学，成都

2010 年 6 月

目录

第一章 线性代数基础

线性代数有着悠久的历史，是数学学科中的一门重要基础课，在线性规划、离散数学、管理科学、计算机科学以及物理学、化学等学科中有着极为广泛的应用.

本章主要分为线性代数的基本概念、主要结论、典型例题和习题解答几个部分，所讨论的内容既是已有线性代数知识的深化，也是全面理解和掌握本书后面内容的数学基础.

一、基本概念

（一）线性空间与线性变换

作为在研究物理学、力学中满足叠加原理的系统的一种数学抽象，线性空间与线性变换是学习矩阵理论的两个基本概念. 线性空间是对集合中的元素在进行线性运算时的共性加以概括而形成的，是 n 维向量空间的抽象和推广. 线性变换则反映了线性空间中的元素之间最基本的线性关系，是研究线性空间的主要工具.

1. 线性空间

定义 1 设 V 是一个非空集合，P 是一个数域. 如果 V 满足如下两个条件：

(1) 在 V 中定义了一个封闭的加法运算，即当 $\alpha,\beta\in V$ 时，有唯一的和 $\alpha+\beta\in V$，并且加法运算满足 4 条性质：

① $\alpha+\beta=\beta+\alpha$（交换律）；

② $(\alpha+\beta)+\upsilon=\alpha+(\beta+\upsilon)$（结合律）；

③ 存在零元素 $0\in V$，使得对于 V 中任何一个元素 α 都有 $\alpha+0=\alpha$；

④ 存在负元素，即对任一元素 $\alpha\in V$，存在有一个元素 $\beta\in V$，使得 $\alpha+\beta=0$（β 称为 α 的负元素，记为 $\beta=-\alpha$）.

(2) 在 V 中定义了一个封闭的数乘运算，即对于数域 P 中任一数 k 和任意 $\alpha\in V$，有唯一的 $k\alpha\in V$，且数乘运算满足 4 条性质：

⑤ $1\alpha=\alpha$；

⑥ $k(l\alpha)=(kl)\alpha$（结合律）；

⑦ $(k+l)\alpha=k\alpha+l\alpha$（分配律）；

⑧ $k(\alpha+\beta)=k\alpha+k\beta$（数因子分配律）.

这时，我们称 V 是数域 P 上的线性空间，其中 α,β,υ 表示 V 中的任意元素；k,l 是数域 P 中的任意数；1 是数域 P 中的单位元.

通常我们把 V 中满足以上 8 条性质且封闭的加法及数乘两种运算，统称为线性运算. 线性运算是线性空间的本质，它反映了集合中元素之间的某种代数结构.

下面是几种常见的线性空间：

(1) 实行向量空间 $\mathbf{R}^n=\{\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\mid x_i\in\mathbf{R}\}$.

(2) 复矩阵空间 $\mathbf{C}^{m\times n}=\{\boldsymbol{A}=(a_{ij})_{m\times n}\mid a_{ij}\in\mathbf{C}\}$.

(3) 复多项式空间 $P_n[t]=\{f(t)=a_0+a_1t+\cdots+a_nt^n\mid a_i\in\mathbf{C}\}$.

(4) 齐次线性方程组的解空间 $\{\boldsymbol{x}\mid\boldsymbol{Ax}=\boldsymbol{0},\boldsymbol{x}\in\mathbf{R}^n\}$，其中 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$.

2. 线性空间的基与维数

定义 2 设 V 是数域 P 上的线性空间，$x_1,x_2,\cdots,x_n(n\geqslant1)$ 是 V 中的任意 n 个向量，如果满足

(1) $x_1,x_2,\cdots,x_n$ 线性无关；

(2) V 中任何向量 x 均可由 $x_1,x_2,\cdots,x_n$ 线性表示.

则称 $x_1,x_2,\cdots,x_n$ 是 V 的一组基（或基底），并称 $x_1,x_2,\cdots,x_n$ 为 V 的一组基向量.

线性空间 V 的基向量所含向量的个数 n，称为线性空间 V 的维数，记为 $\dim V=n$，并称 V 为 n 维线性空间，简记为 V^n. 最小的有限维线性空间是零空间，其维数为 0.

3. 线性子空间

定义 3 如果数域 P 上线性空间 V 的一非空子集 W 对于 V 的两种运算也构成线性空间，则称 W 为 V 的一个线性子空间（简称子空间）.

显然，对于一个线性空间 V，这里至少存在两个子空间—— 一个是其自身，另一个是仅含零向量的零子空间. 另外，子空间之间也可进行多种运算，如交与和，其相互作用可构成一些新的子空间.

4. 线性变换

定义 4 线性空间 V 的一个变换 A 称为线性变换，如果对于 V 中任意的元素 α,β 和数域 P 中任意数 k,l，都有 $A(k\alpha+l\beta)=kA(\alpha)+lA(\beta)$.

下面给出几种常见的线性子空间.

(1) 生成子空间 $L(x_1,x_2,\cdots,x_n)$ 或者 $\operatorname{span}\{x_1,x_2,\cdots,x_n\}$　设 V 是数域 P 上的线性空间，$x_i\in V(i=1,2,\cdots,n)$，则

$$L(x_1,x_2,\cdots,x_n)=\{x=k_1x_1+k_2x_2+\cdots+k_nx_n \mid k_i\in P, i=1,2,\cdots,n\}.$$

(2) 矩阵 $\boldsymbol{A}$ 的值域 $R(\boldsymbol{A})$　设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$ 的列向量组为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$,则

$$R(\boldsymbol{A})=L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=\{\boldsymbol{y}\mid \boldsymbol{y}=\boldsymbol{Ax},\boldsymbol{x}\in\mathbf{C}^n\}.$$

(3) 矩阵 $\boldsymbol{A}$ 的零空间 $N(\boldsymbol{A})$　设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$,则 $N(\boldsymbol{A})=\{\boldsymbol{x}\mid\boldsymbol{Ax}=\boldsymbol{0},\boldsymbol{x}\in\mathbf{C}^n\}$.

(4) 线性变换 A 的值域 $R(A)$　设 A 是线性空间 V 的线性变换,则

$$R(A)=\{y\mid y=A\alpha,\alpha\in V\}.$$

(5) 线性变换 A 的核 $N(A)$　设 A 是线性空间 V 的线性变换,则

$$N(A)=\{x\mid Ax=0,x\in V\}.$$

(6) 线性变换 A 的特征子空间 V_λ　设 λ 是线性空间 V 中线性变换 A 的一个特征值,则

$$V_\lambda=\{x\mid Ax=\lambda x,x\in V\}.$$

另外,对于线性变换来说,这里也存在着许多常见的特殊形式,如正交变换(初等旋转变换和镜像变换等)、对称变换、正规变换和酉变换等. 这些变换的定义在此不再详述,可详见本章第二部分中关于这些概念的一些相应的等价条件.

(二) 空间分解与维数定理

定义 5　设 V_1,V_2 是线性空间 V 的子空间,由同时属于这两个子空间中的向量构成的子集合,叫做 V_1 与 V_2 的交,记为 $V_1\cap V_2=\{\alpha\mid\alpha\in V_1,\alpha\in V_2\}$.

定义 6　设 V_1,V_2 是线性空间 V 的子空间,所谓 V_1 与 V_2 的和,记为

$$V_1+V_2=\{\alpha=\alpha_1+\alpha_2\mid\alpha_1\in V_1,\alpha_2\in V_2\}.$$

一般地,V_1+V_2 和 $V_1\cap V_2$ 均是 V 的子空间,但一般说来,$V_1\cup V_2$ 并不是 V 的一个子空间(见本章典型例题分析中的例 2). 另外,它们一起构成的分配律一般也不成立.

定义 7　设 $V_1,V_2,\cdots,V_s$ 是线性空间 V 的子空间,如果和空间 $V_1+V_2+\cdots+V_s$ 中的每个向量 α 都有分解式

$$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_s,\quad \alpha_i\in V_i(i=1,2,\cdots,s),$$

且这种分解是唯一的,则称这个和为直和,记为 $V_1\oplus V_2\oplus\cdots\oplus V_s$.

定义 8　设 T 是酉空间(复内积空间)$V^n(\mathbf{C})$上的线性变换,W 是 $V^n(\mathbf{C})$的子空间,如果对 $\forall\xi\in W$,有 $T(\xi)\in W$,即 $T(W)\subseteq W$,则称 W 是 T 的不变子空间.

定义 9　对酉空间 $V^n(\mathbf{C})$的两个子空间 V_1 和 V_2,若有 $V_1\perp V_2$,且 $V_1+V_2=V^n(\mathbf{C})$,则称 V_2 为 V_1 的正交补子空间(或简称正交补),记为 $V_2=V_1^\perp$.

显然,如果 V_2 是 V_1 的正交补,则 V_1 也是 V_2 的正交补,且 V_1+V_2 是直和,即 $V^n(\mathbf{C})=V_1\oplus V_2$.

(三) 商空间

定义 10　设 $\alpha\in V$,如果 $\alpha'\in V$ 满足 $\alpha'-\alpha\in M$,则称 α' 与 α 模 M 同余,记为 $\alpha'\equiv\alpha(\mathrm{mod}\ M)$. 其中 V 是一个 n 维线性空间,M 是它的子空间. 显然,$\forall\alpha\in V$,则 V 的子集 $\alpha+M=$

$\{\alpha+m|m\in M\}$内的任意一向量必与α模M同余；反之，与α模M同余的向量必属于$\alpha+M$. 那么，我们称$\alpha+M$为一个模M的同余类，而α称为这个同余类的代表.

定义 11 V的所有模M同余类的全体组成的集合称为V的商集，记为$\bar{V}$.

定义 12 在商集$\bar{V}$中定义如下加法和数乘两种运算：

(1) $(\alpha+M)+(\beta+M)=(\alpha+\beta)+M$;

(2) $k(\alpha+M)=k\alpha+M, \forall k\in P$.

则商集$\bar{V}$构成了数域P上的一个线性空间，称为V对于子空间M的商空间，记为V/M.

上述定义的同余关系实际上是一种等价关系，因此可以用其对n维线性空间V进行分类. 另外，与之类似的一个概念就是如下要谈到的线性流形问题.

（四）线性流形与凸包

定义 13 设V_1是V的子空间，r_0是V的固定向量，则称集合

$$P=r_0+V_1=\{r_0+\alpha \mid \alpha\in V_1\}$$

为线性空间V的一个线性流形，且V_1的维数称为线性流形P的维数.

线性流形是一个很重要的概念，许多实际问题都可转化为线性流形上的问题，如超平面问题、非奇次线性方程组解的结构问题等.

定义 14 在$\mathbf{R}^n$中以向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$的终点为端点的线段定义为

$$\{\boldsymbol{r} \mid \boldsymbol{r}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2, k_1+k_2=1, k_1,k_2\geqslant 0\}.$$

如果在$\mathbf{R}^n$的点集M中，以任意两点为端点组成的线段都含于M中，则称M为凸集. 给定集合$A\subset\mathbf{R}^n$，$\mathbf{R}^n$中所有包含A的凸集的交集称为A的凸包. 显然，$\mathbf{R}^n$中有限点集$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$的凸包为

$$M^*=\left\{\boldsymbol{r} \mid \boldsymbol{r}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s, \sum_{i=1}^{s}k_i=1, 0\leqslant k_i\leqslant 1\right\}.$$

（五）特征值与特征向量

定义 15 设$\boldsymbol{A}\in\mathbf{C}^{n\times n}$，如果存在$\lambda\in\mathbf{C}$和非零向量$\boldsymbol{x}\in\mathbf{C}^n$，使$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，则$\lambda$叫做$\boldsymbol{A}$的特征值，$\boldsymbol{x}$叫做$\boldsymbol{A}$的属于特征值$\lambda$的特征向量.

另外，就线性变换来说，也存在与之类似，实质上等价的定义.

定义 16 设T是线性空间V的一个线性变换，如果存在$\lambda\in\mathbf{C}$和非零向量$\xi\in V$，使得$T\xi=\lambda\xi$，则λ叫做T的特征值，ξ叫做T的属于特征值λ的特征向量.

一般地，矩阵$\boldsymbol{A}$的所有特征值的全体，叫做$\boldsymbol{A}$的谱，记做$\lambda(\boldsymbol{A})$. 另外，除了上述一般特征值问题外，还有如下更一般的广义特征值问题.

定义 17 设$\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$，如果对于一个复数$\lambda$，存在非零向量$\boldsymbol{x}\in\mathbf{C}^n$，使得$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{B}\boldsymbol{x}$，则称$\lambda$为矩阵$\boldsymbol{A}$相对于$\boldsymbol{B}$的特征值，或称$\lambda$为$\boldsymbol{A}$与$\boldsymbol{B}$确定的广义特征值. 非零向量$\boldsymbol{x}$称为与$\lambda$对应的广义特征向量.

定义 18 设 $\boldsymbol{A}\in \mathbf{C}^{n\times n}$，如果存在可逆矩阵 $\boldsymbol{P}\in \mathbf{C}^{n\times n}$，使得

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

则 $\boldsymbol{A}$ 叫做可对角化矩阵.

(六) 欧氏空间上的度量

定义 19 在线性空间 $V^n(P)$ 中，若映射 (x,y)：$V^n(P)\times V^n(P)\to P$ 满足

(1) $(x,x)\geqslant 0$，$(x,x)=0$ 当且仅当 $x=0$；

(2) $(x,y)=\overline{(y,x)}$，$\forall x,y\in V^n(P)$；

(3) $(\lambda x,y)=\bar{\lambda}(x,y)$，$\forall \lambda\in P$，$\forall x,y\in V^n(P)$；

(4) $(x+y,z)=(x,y)+(x,z)$，$\forall x,y,z\in V^n(P)$，

则称 (x,y) 是 $V^n(P)$ 上的内积，定义了内积的线性空间称为内积空间.

(七) 初等矩阵及酉变换

定义 20 设 $\boldsymbol{u},\boldsymbol{v}\in \mathbf{C}^n$，$\sigma\in \mathbf{C}$，形如

$$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\sigma) = \boldsymbol{E}_n - \sigma \boldsymbol{u}\boldsymbol{v}^{\mathrm{H}}$$

的矩阵叫做初等矩阵，其中 $\boldsymbol{v}$ 的 Hermite 转置 $\boldsymbol{v}^{\mathrm{H}}$ 表示对 $\boldsymbol{v}$ 取转置后再取复共轭，$\boldsymbol{E}_n$ 表示 n 阶单位矩阵.

定义 21 设 $\boldsymbol{u}\in \mathbf{C}^n$，且 $\boldsymbol{u}^{\mathrm{H}}\boldsymbol{u}=1$，则

$$\boldsymbol{H}(\boldsymbol{u}) = \boldsymbol{E}(\boldsymbol{u},\boldsymbol{u};2) = \boldsymbol{E}_n - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}},$$

称为初等酉阵，或 Householder 变换.

定义 22 若线性空间 $V^n(\mathbf{C})$ 上的变换 T 满足

$$(T(x),T(y)) = (x,y),\ \forall x,y \in V^n(\mathbf{C}),$$

则称 T 为 $V^n(\mathbf{C})$ 上的酉变换.

(八) 酉空间的分解与投影

定义 23 设 $V^n(\mathbf{C})$ 是线性空间，如果线性变换 T:$V^n(\mathbf{C})\to V^n(\mathbf{C})$ 具有 $T^2=T$ 的性质，则称 T 是 $V^n(\mathbf{C})$ 上的投影 (也称投影算子或幂等算子).

定义 24 设 T 是 $V^n(\mathbf{C})$ 上的投影，$V^n(\mathbf{C})=R(T)\oplus N(T)$. 如果 $R^{\perp}(T)=N(T)$，则称 T 是正交投影.

(九) Kronecker 乘积

定义 25 设 $\boldsymbol{A}=(a_{ij})\in P^{m\times n}$，$\boldsymbol{B}=(b_{ij})\in P^{p\times q}$，则

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

称为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 Kronecker 乘积（也称直积或张量积）.

定义 26 设 $m\times n$ 矩阵 $\boldsymbol{A}=(\boldsymbol{A}_{c1},\boldsymbol{A}_{c2},\cdots,\boldsymbol{A}_{cn})$，令

$$\mathrm{Vec}\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{c1}\\ \boldsymbol{A}_{c2}\\ \vdots\\ \boldsymbol{A}_{cn}\end{pmatrix},$$

称 Vec 为向量化算符(俗称拉直).

（十）线性代数在科学工程中的应用

数学学科的重要性除了它对人的逻辑推理、空间想象和分析问题等能力的培养具有其他学科无法代替的作用外，更重要的是其应用的广泛性. 线性代数的应用大体上可以分为两个方面，一是在其他数学分支中的应用，另一就是在数学学科之外的应用，其应用范围不仅包括物理学、化学、天文学、生物学等传统学科，还包括现代工程技术和经济理论等学科.

对线性空间和线性变换来说，在数学学科中可用于最优化理论中的最优下降法、线性规划和最小二乘问题等，现代科学计算中的 Krylov 子空间方法等，概率论中协方差的计算等，空间解析几何中二次曲面的划分问题等；在其他科学中，还可用于计算机图形学中计算机层析 X 射线问题，化学中离子(原子和分子)的振动问题，以及经济学研究中常用的随机矩阵和非负矩阵相关问题，物理学中求解相对论问题的洛伦兹变换等.

另外，在移动通信中，多址通信技术的理论基础是基于线性代数中的正交向量等理论；在模式识别中，近邻分类模式分类法也常用到线性代数中的欧氏空间上的度量等基础知识.

空间分解与投影技术在实际问题中也越来越受到了大家的欢迎，如 20 世纪 30 年代发展起来的斜投影算子现已广泛应用于图像恢复、快速系统识别、脉冲噪声对消、误码校正编码、信道与发射字符的联合估计等领域；正交投影矩阵在移动通信的盲多用户检测的 LMS 型自适应算法中也有很好的应用等.

另外，就其在数学其他分支中的应用来说，商空间、线性流形与凸包实际上是对常见的线性方程组解空间理论的更加一般性的抽象和概括. 如相容的(即有解的)非齐次线性方程组的解实际上是一个由导出组的解空间构成的线性流形.

还如，现代工程技术中一些振动问题(如振动分子的固有频率)与稳定性问题，数学中矩阵的对角化与微分方程组的求解问题，还有卫星和喷气式飞机工程设计等一些实际问题，都可以归结为求矩阵的特征值与特征向量的问题.

最后，矩阵的 Kronecker 乘积和向量化算符相结合也有着广泛的应用，例如利用这一技术可以方便地研究一般线性矩阵方程

$$\boldsymbol{A}_1\boldsymbol{X}\boldsymbol{B}_1+\boldsymbol{A}_2\boldsymbol{X}\boldsymbol{B}_2+\cdots+\boldsymbol{A}_p\boldsymbol{X}\boldsymbol{B}_p=\boldsymbol{C}$$

的相容性及其解法等问题(见本章的主要结论部分)，其中 $\boldsymbol{A}_i\in\mathbf{C}^{m\times m}$，$\boldsymbol{B}_i\in\mathbf{C}^{n\times s}$，$\boldsymbol{C}\in\mathbf{C}^{m\times s}$ 均为已知矩阵，$\boldsymbol{X}\in\mathbf{C}^{m\times n}$ 是未知矩阵.

二、主要结论

(一) 线性空间与线性变换

线性空间和线性变换是线性代数研究的主要内容之一．特别是一些重要的不变子空间和线性变换的性质，都是线性代数中的重点知识．

1．子空间的性质

(1) 设 $V_1,V_2,\cdots,V_m$ 是线性空间 V 的真子空间，则必存在 $\alpha\in V$，使得 $\alpha\notin V_i$，$1\leqslant i\leqslant m$.

(2)（扩展性定理）设 $V_1=L(u_1,u_2,\cdots,u_m)$，令 $v_1,v_2,\cdots,v_r$ 是 V_1 中的 r 个线性无关的向量，且 $r<m$，则可以从 $u_1,u_2,\cdots,u_m$ 中去掉 r 个向量，使剩下的 $m-r$ 个向量与 $v_1,v_2,\cdots,v_r$ 合在一起仍生成子空间 V_1.

2．子空间的和与交的基与维数的求法

设 V_1 和 V_2 是 V 的子空间，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$是 V_1 的一组基，$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_l$是 V_2 的一组基，则

(1) V_1+V_2 的维数等于矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_l)$的秩 r，且 $\boldsymbol{A}$ 中 r 个线性无关的列即为 V_1+V_2 的基；

(2) 令 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_k\boldsymbol{\alpha}_k=y_1\boldsymbol{\beta}_1+y_2\boldsymbol{\beta}_2+\cdots+y_l\boldsymbol{\beta}_l$，解这个方程组求得一个基础解系

$$(x_{i1},x_{i2},\cdots,x_{ik},y_{i1},y_{i2},\cdots,y_{il})^{\mathrm{T}},\quad i=1,2,\cdots,d;d=k+l-r,$$

则 $\boldsymbol{z}_i=\sum\limits_{j=1}^{l}y_{ij}\boldsymbol{\beta}_j\,(i=1,2,\cdots,d)$ 就是 $V_1\cap V_2$ 的一组基，$V_1\cap V_2$ 的维数为 $d=k+l-r$.

3．线性变换的值域与核

(1) $R(A)=L(A\alpha_1,A\alpha_2,\cdots,A\alpha_n)$，其中 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是线性空间 V 的一组基，且 $\dim R(A)=r$（线性变换 A 所对应的矩阵的秩）；

(2) $\dim R(A)+\dim N(A)=\dim V$；

(3) 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是$R(A)$的一组基且 $A\beta_i=\alpha_i$，$1\leqslant i\leqslant n$，则

$$V=N(A)\oplus L(\beta_1,\beta_2,\cdots,\beta_n).$$

因此，一般地 V 不等于$R(A)$与 $N(A)$的直和；

(4) $R(A)$和 $N(A)$都是线性变换 A 的不变子空间；

(5) A 与 B 可换，则 B 的核与值域也是 A 的不变子空间.

4. 不变子空间

(1) 线性空间 V 的子空间 W 是线性变换 A 的不变子空间当且仅当对任意的 $\alpha \in W$ 有 $A\alpha \in W$；

(2) 设 λ 是线性变换 A 的特征值，则 A 属于 λ 的特征子空间 $V_\lambda = \{x \mid Ax = \lambda x\}$ 是 A 的不变子空间；

(3) 不变子空间的和与交仍是不变子空间；

(4) 任一空间是数乘变换的不变子空间；

(5) 设 W 是线性空间 V 的子空间且 $W = L(\alpha_1, \alpha_2, \cdots, \alpha_r)$，则 W 是线性变换 A 的不变子空间当且仅当 $A\alpha_i \in W, i = 1, 2, \cdots, r$；

(6) 设 V_1 是线性变换 A 的不变子空间，则对任一多项式 f，V_1 是 $f(A)$ 的不变子空间；

(7) 设 A 和 B 是线性变换且 $AB = BA$，V_λ 是 A 的特征子空间，则 V_λ 也是 B 的不变子空间；

(8) 设 V_1 是线性变换 A 和 B 的不变子空间，则它也是 $A+B$ 及 AB 的不变子空间；

(9) 设 A 是 n 维线性空间 V 上的线性变换，若 A 有 n 个不同的特征根，则 A 有 2^n 个不变子空间.

5. 几种常见的线性变换

1) 正交变换

(1) 下列说法等价：

① 欧氏空间 V 的线性变换 A 是正交变换；

② A 保持内积不变，即对任意的 $\alpha, \beta \in V$，有 $(A\alpha, A\beta) = (\alpha, \beta)$；

③ A 保持长度不变，即对 V 的任意元 α，有 $(A\alpha, A\alpha) = (\alpha, \alpha)$；

④ A 把一组标准正交基变为另一组标准正交基；

⑤ A 在一组标准正交基下的矩阵是正交矩阵.

(2) 欧氏空间中的一个变换，若保持内积不变，则是正交变换.

(3) 正交变换的逆和积是正交变换.

(4) A 的特征值的模等于 1.

2) 对称变换

(1) 欧氏空间 V 中的线性变换 A 是对称变换，当且仅当对任意的 $\alpha, \beta \in V$，有 $(A\alpha, \beta) = (\alpha, A\beta)$；当且仅当在一组标准正交基下变换 A 所对应的矩阵为对称矩阵.

(2) 设 A 是欧氏空间 V 中的对称变换，若 V_1 是 A 的不变子空间，则 V_1 的正交补 $V_1^\perp$ 也是 A 的不变子空间.

(3) 对称变换的所有特征值全为实数，且属于不同特征值的特征向量互相正交.

3）共轭变换与正规变换

类似于欧氏空间的定义，设 V 是复数域上的线性空间，并在其上定义了内积，则称 V 为酉空间. 这里的内积是二元复函数，满足 $(\alpha,\beta)=\overline{(\beta,\alpha)}$. 对应于欧氏空间的正交变换和对称变换，在酉空间中有酉变换和对称变换. 设 A 是酉空间中的线性变换. 若对任意的 $\alpha,\beta\in V$，有 $(A\alpha,\beta)=(\alpha,B\beta)$，则称 B 是 A 的共轭变换，并记为 $B=A^*$. 有关共轭变换的几个重要结论可归纳如下：

(1) 设线性变换 A 在一组标准正交基下的矩阵为 $\boldsymbol{A}$，则 A 的共轭变换在这组基下的矩阵为 $\bar{\boldsymbol{A}}^{\mathrm{T}}$.

(2) 共轭变换满足 $(A^*)^*=A$，$(A+B)^*=A^*+B^*$，$(AB)^*=B^*A^*$，$(kA)^*=\bar{k}A^*$.

(3) 设酉空间 V 的子空间 W 是线性变换 A 的不变子空间，则 W 的正交补 $W^{\perp}$ 是 A^* 的不变子空间.

(4) 若 $AX=\lambda X$，则 $A^*X=\bar{\lambda}X$.

(5) 若线性变换 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则 A^* 的特征值为 $\bar{\lambda}_1,\bar{\lambda}_2,\cdots,\bar{\lambda}_n$.

(6) 若线性变换 A 满足 $A^*A=AA^*$，则称 A 为正规变换. 常见的正交变换，对称变换都是正规变换.

(7) 设 A 是正规变换，则属于 A 的不同特征值的特征向量正交.

(8) 若 A 是正规变换，W 是 A 的不变子空间，则 $W^{\perp}$ 也是 A 的不变子空间.

(9) 若线性变换 A 满足 $A^*=-A$，则称 A 为酉空间的反对称变换. 显然反对称变换也是正规变换，且它满足对任意的 $\alpha,\beta\in V$，有 $(A\alpha,\beta)=-(\alpha,A\beta)$.

(10) 设 A 是酉空间的线性变换，则 A 是正规变换当且仅当 A 的每个特征向量也是 A^* 的特征向量.

（二）空间分解与维数定理

1. （维数公式）设 V_1,V_2 是线性空间 V 的子空间，则

$$\dim(V_1)+\dim(V_2)=\dim(V_1+V_2)+\dim(V_1\cap V_2).$$

2. （存在性定理）设 V_1 是 n 维线性空间 V 的一个子空间，则一定存在 V 的一个子空间 V_2，使得 $V=V_1\oplus V_2$.

3. 设 $V_1,V_2,\cdots,V_s$ 是线性空间 V 的子空间，则下列几条命题相互等价：

(1) $W=\sum\limits_{i=1}^{s}V_i$ 是直和；

(2) 零向量表式法唯一；

(3) $V_i\cap\left(\sum\limits_{j\neq i}V_j\right)=\{0\}$，$1\leqslant i\leqslant s$；

(4) $\dim(W)=\sum\limits_{i=1}^{s}\dim(V_i)$.

4. 设 T 是线性空间 V^n 的线性变换，若 V^n 可分解为 T 的不变子空间 $V_i(i=1,2,\cdots,m)$

的直和,则 T 在由 $V_1,V_2,\cdots,V_m$ 的基拼接而构成 V^n 的基下的矩阵为准对角矩阵.

5. 设 T 是线性空间 V^n 的线性变换,若 T 在 V^n 的某个基下的矩阵为准对角矩阵 $\mathrm{diag}(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m)$,则 V^n 可分解为 T 的 m 个不变子空间的直和.

6. 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$,则有

(1) $[R(\boldsymbol{A})]^{\perp}=N(\boldsymbol{A}^{\mathrm{H}})$,且 $\mathbf{C}^m=R(\boldsymbol{A})\oplus N(\boldsymbol{A}^{\mathrm{H}})$;

(2) $[R(\boldsymbol{A}^{\mathrm{H}})]^{\perp}=N(\boldsymbol{A})$,且 $\mathbf{C}^n=R(\boldsymbol{A}^{\mathrm{H}})\oplus N(\boldsymbol{A})$.

(三) 商空间

1. V 的关于子集 M 的商集 $\bar{V}$ 中的任意元素 $\alpha+M$,皆是模 M 的一个同余类,且满足如下等价关系:

(1) 反身律　$\alpha\equiv\alpha(\mathrm{mod}\ M)$;

(2) 对称律　若 $\alpha'\equiv\alpha(\mathrm{mod}\ M)$,则 $\alpha\equiv\alpha'(\mathrm{mod}\ M)$;

(3) 传递律　若 $\alpha''\equiv\alpha'(\mathrm{mod}\ M)$,$\alpha'\equiv\alpha(\mathrm{mod}\ M)$,则 $\alpha''\equiv\alpha(\mathrm{mod}\ M)$.

2. (商空间维数公式) 设 M 是 V 的子空间,则

$$\dim\left(\frac{V}{M}\right)=\dim(V)-\dim(M).$$

(四) 线性流形与凸包

1. (线性流形存在性定理)对任一 s 维线性流形 P,存在 $s+1$ 个向量 $\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$,使得

$$P=\{x=k_0\boldsymbol{\alpha}_0+k_1\boldsymbol{\alpha}_1+\cdots+k_s\boldsymbol{\alpha}_s\mid k_0+k_1+\cdots+k_s=1\},$$

且向量组 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_s-\boldsymbol{\alpha}_0$ 线性无关. 反之,设 $\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 是 $\mathbf{R}^n$ 中任意 $s+1$ 个向量,且 $k_0+k_1+\cdots+k_s=1$,则形如 $x=k_0\boldsymbol{\alpha}_0+k_1\boldsymbol{\alpha}_1+\cdots+k_s\boldsymbol{\alpha}_s$ 的所有向量组成一个维数等于向量组 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_0,\cdots,\boldsymbol{\alpha}_s-\boldsymbol{\alpha}_0$ 的秩的线性流形 P.

2. 空间 $\mathbf{R}^n$ 的两个维数分别为 k 和 h 的线性流形 P 和 Q 包含在一个维数小于等于 $k+h+1$ 的线性流形中. 若线性流形 P 和 Q 有一个公共向量 $\boldsymbol{\alpha}_0$,则 $P\cap Q$ 是一个维数大于等于 $k+h-n$ 的线性流形.

(五) 特征值与特征向量

1. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 有 r 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_r$,其代数重数分别为 $n_1,n_2,\cdots,n_r$,则必存在可逆矩阵 $\boldsymbol{P}\in\mathbf{C}^{n\times n}$,使得

$$\boldsymbol{J}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\boldsymbol{J}_1(\lambda_1),\cdots,\boldsymbol{J}_r(\lambda_r)),$$

其中 $J_i(\lambda_i)(i=1,2,\cdots,r)$ 是维数为 n_i 的 Jordan 块,则矩阵 $\boldsymbol{J}$ 叫做 $\boldsymbol{A}$ 的 Jordan 标准形.

2. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,则下列命题等价:

(1) $\boldsymbol{A}$ 是可对角化矩阵;

(2) $\mathbf{C}^n$ 存在由 $\boldsymbol{A}$ 的特征向量构成的一组基底;

(3) $\boldsymbol{A}$ 的 Jordan 标准形中的 Jordan 块都是一阶的；

(4) 矩阵 $\boldsymbol{A}$ 的每个特征值的代数重数和几何重数皆相等；

(5) $\boldsymbol{A}$ 的初等因子都是一次因式；

(6) $\boldsymbol{A}$ 的最小多项式无重根；

(7) 若 n 阶方阵 $\boldsymbol{A}$ 的谱是 $\lambda_1,\lambda_2,\cdots,\lambda_s$，则 $\mathbf{C}^n=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s}$.

3. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，$\boldsymbol{A}$ 是可对角化矩阵的几个常见的充分条件为

(1) $\boldsymbol{A}$ 有 n 个不同的特征值；

(2) $\boldsymbol{A}$ 的最后一个不变因子是不同的一次因式的乘积；

(3) $\boldsymbol{B}$ 有 n 个不同的特征值，且 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可换；

(4) $\boldsymbol{A}=\boldsymbol{BC}$，且 $\boldsymbol{B}$ 正定，$\boldsymbol{C}$ 为实对称矩阵；

(5) $\boldsymbol{A}$ 是实对称矩阵.

4. 线性变换(或矩阵)T 的不同特征值所对应的特征向量是线性无关的.

5. 设 $n\times n$ 矩阵 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$，$\boldsymbol{B}=\boldsymbol{B}^{\mathrm{H}}$，且 $\boldsymbol{B}$ 是正定的，则 $\boldsymbol{B}$ 的共轭向量系 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 具有以下性质：

(1) $\boldsymbol{x}_i\neq\boldsymbol{0}(i=1,2,\cdots,n)$；

(2) $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 线性无关；

(3) λ_i 与 $\boldsymbol{x}_i$ 满足方程 $\boldsymbol{A}\boldsymbol{x}_i=\lambda_i\boldsymbol{B}\boldsymbol{x}_i(i=1,2,\cdots,n)$；

(4) 若记 $\boldsymbol{X}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)$，则 $\boldsymbol{X}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{X}=\boldsymbol{E}$，$\boldsymbol{X}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{X}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$.

6. 广义特征值和特征向量的性质

设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$ 都是 Hermite 矩阵(即 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$，$\boldsymbol{B}=\boldsymbol{B}^{\mathrm{H}}$)，且 $\boldsymbol{B}$ 为正定矩阵，则

(1) 若有分解 $\boldsymbol{B}=\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}$，则 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{B}\boldsymbol{x}$ 等价于 $\boldsymbol{S}\boldsymbol{y}=\lambda\boldsymbol{y}$，其中 $\boldsymbol{S}=\boldsymbol{G}^{-1}\boldsymbol{A}(\boldsymbol{G}^{-1})^{\mathrm{H}}$，$\boldsymbol{y}=\boldsymbol{G}^{\mathrm{H}}\boldsymbol{x}$；

(2) 若 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{B}\boldsymbol{x}$，即 λ 是 $\boldsymbol{A}$ 相对于 $\boldsymbol{B}$ 的广义特征值，则 $\lambda\in\mathbf{R}$，且 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{y}=0$，其中 $\boldsymbol{x},\boldsymbol{y}$ 是对应不同广义特征值的广义特征向量；

(3) $R_{\boldsymbol{B}}(\boldsymbol{x})=\dfrac{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{x}}(\boldsymbol{x}\neq\boldsymbol{0})$在 $\mathbf{C}^n$ 的一维子空间上的值为常数；

(4) $R_{\boldsymbol{B}}(\boldsymbol{x})$在 $\mathbf{C}^n$ 的 k 维子空间 V^k 上的最值存在，且能够在 V^k 的子集$\{\boldsymbol{x}\mid\boldsymbol{x}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{x}=1,\boldsymbol{x}\in V^k\}$上达到；

(5) $R_{\boldsymbol{B}}(\boldsymbol{x})$的驻点是 $\boldsymbol{A}$ 相对于 $\boldsymbol{B}$ 的广义特征值，反之亦然；

(6) 若 $\boldsymbol{x}_0$ 是 $R_{\boldsymbol{B}}(\boldsymbol{x})$的驻点，则 $R_{\boldsymbol{B}}(\boldsymbol{x}_0)$是 $\boldsymbol{A}$ 相对于 $\boldsymbol{B}$ 的广义特征值；

(7) $\boldsymbol{A}$ 相对于 $\boldsymbol{B}$ 的最小和最大广义特征值分别为$\min\limits_{\boldsymbol{x}\neq\boldsymbol{0}}R_{\boldsymbol{B}}(\boldsymbol{x})$和$\max\limits_{\boldsymbol{x}\neq\boldsymbol{0}}R_{\boldsymbol{B}}(\boldsymbol{x})$；

(8) $\boldsymbol{A}$ 相对于 $\boldsymbol{B}$ 的从小到大排列的第 k 个广义特征值为$\min\limits_{V^k}[\max\limits_{0\neq\boldsymbol{x}\in V^k}R_{\boldsymbol{B}}(\boldsymbol{x})]$，$\boldsymbol{A}$ 相对于 $\boldsymbol{B}$ 的从大到小排列的第 k 个广义特征值为$\max\limits_{V^k}[\min\limits_{0\neq\boldsymbol{x}\in V^k}R_{\boldsymbol{B}}(\boldsymbol{x})]$，其中 V^k 表示 $\mathbf{C}^n$ 的任意一个 k 维子空间.

（六）初等矩阵及酉变换

1. 初等矩阵的性质

（1）$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\sigma)$的特征值为$(1-\sigma\boldsymbol{v}^{\mathrm{H}}\boldsymbol{u}),1,1,\cdots,1$.

（2）$\det \boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\sigma))=1-\sigma\boldsymbol{v}^{\mathrm{H}}\boldsymbol{u}$.

（3）当且仅当$\sigma\boldsymbol{v}^{\mathrm{H}}\boldsymbol{u}\neq 1$时，$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\sigma)$可逆，且

$$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\sigma)^{-1}=\boldsymbol{E}\left(\boldsymbol{u},\boldsymbol{v};\frac{\sigma}{\sigma\boldsymbol{v}^{\mathrm{H}}\boldsymbol{u}-1}\right).$$

特别地

$$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\sigma)^{-1}=\begin{cases}\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};-\sigma), & \text{当 }\sigma\boldsymbol{v}^{\mathrm{H}}\boldsymbol{u}=0,\\ \boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\sigma), & \text{当 }\sigma\boldsymbol{v}^{\mathrm{H}}\boldsymbol{u}=2.\end{cases}$$

（4）对任意非零向量$\boldsymbol{a},\boldsymbol{b}\in\mathbf{C}^n$，存在$\boldsymbol{u},\boldsymbol{v}$与$\sigma$，使得$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\sigma)\boldsymbol{a}=\boldsymbol{b}$.

2. 酉变换或 Householder 变换（对应矩阵 $\boldsymbol{H}(\boldsymbol{u})$）的性质

（1）$\boldsymbol{H}(\boldsymbol{u})^{\mathrm{H}}=\boldsymbol{H}(\boldsymbol{u})=\boldsymbol{H}(\boldsymbol{u})^{-1}$，$\det \boldsymbol{H}(\boldsymbol{u})=-1$；

（2）$\boldsymbol{H}(\boldsymbol{u})$是镜像变换；

（3）设$\boldsymbol{a},\boldsymbol{b}\in\mathbf{C}^n$，则存在单位向量$\boldsymbol{u}$，使得$\boldsymbol{H}(\boldsymbol{u})\boldsymbol{a}=\boldsymbol{b}$的充要条件是

$$\|\boldsymbol{a}\|_2=\|\boldsymbol{b}\|_2 \quad \text{和} \quad \boldsymbol{a}^{\mathrm{H}}\boldsymbol{b}=\boldsymbol{b}^{\mathrm{H}}\boldsymbol{a},$$

且在上式的条件下，使得$\boldsymbol{H}(\boldsymbol{u})\boldsymbol{a}=\boldsymbol{b}$成立的单位向量$\boldsymbol{u}$可取为

$$\boldsymbol{u}=\frac{\mathrm{e}^{\mathrm{i}\theta}(\boldsymbol{a}-\boldsymbol{b})}{\|\boldsymbol{a}-\boldsymbol{b}\|_2},$$

其中θ为任一实数.

（4）任一旋转必可分解为两个 Householder 镜像变换的乘积. 事实上有

$$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}=\left[\boldsymbol{E}_2-2\begin{pmatrix}0\\1\end{pmatrix}(0,1)\right]\left[\boldsymbol{E}_2-2\begin{pmatrix}\sin\dfrac{\theta}{2}\\ -\cos\dfrac{\theta}{2}\end{pmatrix}\left(\sin\frac{\theta}{2},-\cos\frac{\theta}{2}\right)\right].$$

3. 设 T 是 $V^n(\mathbf{C})$ 上的线性变换，则下列命题等价：

（1）T是酉变换；

（2）$\|T(x)\|=\|x\|$，$\forall x\in V^n(\mathbf{C})$；

（3）设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是$V^n(\mathbf{C})$的标准正交基，则$T(\varepsilon_1),T(\varepsilon_2),\cdots,T(\varepsilon_n)$也是它的标准正交基；

（4）T在任一标准正交基下的矩阵$\boldsymbol{A}$是酉矩阵，即$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{E}_n$.

4. 酉矩阵的几个性质

设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 是酉矩阵，则

(1) $(\boldsymbol{Ax},\boldsymbol{Ay})=(\boldsymbol{x},\boldsymbol{y})$，$\forall\boldsymbol{x},\boldsymbol{y}\in\mathbf{C}^n$；

(2) $\|\boldsymbol{Ax}\|=\|\boldsymbol{x}\|$，$\forall\boldsymbol{x}\in\mathbf{C}^n$；

(3) $\boldsymbol{A}^{\mathrm{H}}$ 也是酉矩阵；

(4) 若 $\boldsymbol{B}\in\mathbf{C}^{n\times n}$ 也是酉矩阵，则 $\boldsymbol{AB}$，$\boldsymbol{BA}$ 也是酉矩阵；

(5) 酉矩阵的特征值的模为 1.

（七）欧氏空间上的度量

1. 内积性质

若 (x,y) 是 $V^n(P)$ 上的内积，则

(1) $\|\lambda x\|=\lambda\|x\|$；

(2) $\|x-y\|^2+\|x+y\|^2=2\|x\|^2+2\|y\|^2$；

(3) $|(x,y)|\leqslant\|x\|\ \|y\|$（Cauchy 不等式）；

(4) $\|x+y\|\leqslant\|x\|+\|y\|$（三角不等式）.

2. n 维欧氏空间 V 中的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 线性相关的充要条件是其 Gram 行列式为零，即

$$G(\alpha_1,\alpha_2,\cdots,\alpha_k)=\begin{vmatrix}(\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_k)\\(\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_k)\\\vdots & \vdots & & \vdots\\(\alpha_k,\alpha_1) & (\alpha_k,\alpha_2) & \cdots & (\alpha_k,\alpha_k)\end{vmatrix}=0.$$

3. 如果 n 维欧氏空间中的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 线性无关，则将 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 正交化后，其 Gram 行列式不变，即

$$\begin{aligned}G(\alpha_1,\alpha_2,\cdots,\alpha_k)&=G(\beta_1,\beta_2,\cdots,\beta_k)=(\beta_1,\beta_1)(\beta_2,\beta_2)\cdots(\beta_k,\beta_k)\\&=\|\beta_1\|^2\|\beta_2\|^2\cdots\|\beta_k\|^2.\end{aligned}$$

4. 若向量组 $\alpha_1,\alpha_2,\alpha_3$ 线性无关，则 $G(\alpha_1,\alpha_2)$ 表示由向量 α_1,α_2 所构成的平行四边形面积的平方；$G(\alpha_1,\alpha_2,\alpha_3)$ 表示由向量 $\alpha_1,\alpha_2,\alpha_3$ 构成的平行六面体体积的平方.

5. 关于欧氏空间中线性无关向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 所构成的 n 维平行多面体的体积 $V(\alpha_1,\alpha_2,\cdots,\alpha_n)$，我们有下面结论：

(1) $V(\alpha_1)=\|\alpha_1\|$；

(2) $V(\alpha_1,\alpha_2,\cdots,\alpha_n)=V(\alpha_1,\alpha_2,\cdots,\alpha_{n-1})h_n$，其中 h_n 是向量 α_n 在向量 $\alpha_1,\alpha_2,\cdots,\alpha_{n-1}$ 所生成的子空间上的正交投影的长度.

(3) $V(\alpha_1,\alpha_2,\cdots,\alpha_n)=\sqrt{G(\alpha_1,\alpha_2,\cdots,\alpha_n)}=|D|$，其中 D 是给定的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$

在其生成的 n 维空间 $L(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 的某组标准正交基下的坐标的行列式.

6. 对于给定的向量 α 与线性子空间 V_1 中一切向量的夹角,以它与 α 在 V_1 上的正交投影 r 之间的夹角为最小.

(八) 酉空间的分解与投影

1. 线性变换 T 的矩阵为准对角矩阵的充要条件是 $V^n(\mathbf{C})$ 可分解为 T 的不变子空间的直和.

2. 正交补子空间

(1) 设 $\boldsymbol{A}\in\mathbf{C}^{n\times m}$, $\boldsymbol{B}\in\mathbf{C}^{n\times s}$,则 $R(\boldsymbol{A})\perp R(\boldsymbol{B})$ 的充要条件是 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{B}=\mathbf{0}$.

(2) 酉空间 $V^n(\mathbf{C})$ 的每一个子空间 V_1 都有唯一的正交补.

3. 投影

(1) 设 T 是线性空间 $V^n(\mathbf{C})$ 上的投影,则 $V^n(\mathbf{C})=R(T)\oplus N(T)$,即投影的值域和核互为直和补.

(2) 设 $V^n(\mathbf{C})=V_1\oplus V_2$,则存在投影 T,使 $R(T)=V_1$, $N(T)=V_2$.

4. 若 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 是幂等矩阵,即 $\boldsymbol{A}^2=\boldsymbol{A}$,则有

(1) $\boldsymbol{A}^{\mathrm{H}}$ 与 $\boldsymbol{E}-\boldsymbol{A}$ 也是幂等矩阵;

(2) $\boldsymbol{A}$ 的特征值为 0 或 1,且可对角化;

(3) $\mathrm{rank}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{A})$;

(4) $\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})=(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{A}=\mathbf{0}$;

(5) $\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{\alpha}\Leftrightarrow\boldsymbol{\alpha}\in R(\boldsymbol{A})$;

(6) $N(\boldsymbol{A})=R(\boldsymbol{E}-\boldsymbol{A})$, $R(\boldsymbol{A})=N(\boldsymbol{E}-\boldsymbol{A})$.

5. 正交投影

设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$, $\boldsymbol{A}^2=\boldsymbol{A}$,则下列命题等价:

(1) $\boldsymbol{A}$ 是 $\mathbf{C}^n$ 上的正交投影;

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

(3) $\mathbf{C}^n=R(\boldsymbol{A})\oplus N(\boldsymbol{A})$, $R^{\perp}(\boldsymbol{A})=N(\boldsymbol{A})$.

(九) Kronecker 乘积

1. Kronecker 乘积的运算性质

设 $\boldsymbol{A}\in P^{m\times n}$, $\boldsymbol{B}\in P^{p\times q}$, $\boldsymbol{C}\in P^{r\times s}$, $\boldsymbol{D}\in P^{k\times h}$,则

(1) 单位矩阵之积　$\boldsymbol{E}_m\otimes\boldsymbol{E}_n=\boldsymbol{E}_{mn}$;

(2) 纯量积　$\lambda(\boldsymbol{A}\otimes\boldsymbol{B})=(\lambda\boldsymbol{A})\otimes\boldsymbol{B}=\boldsymbol{A}\otimes(\lambda\boldsymbol{B})$, $\forall\lambda\in P$;

(3) 分配律　$(\boldsymbol{A}+\boldsymbol{B})\otimes\boldsymbol{C}=(\boldsymbol{A}\otimes\boldsymbol{C})+(\boldsymbol{B}\otimes\boldsymbol{C})$, $\boldsymbol{C}\otimes(\boldsymbol{A}+\boldsymbol{B})=(\boldsymbol{C}\otimes\boldsymbol{A})+(\boldsymbol{C}\otimes\boldsymbol{B})$;

(4) 结合律　$\boldsymbol{A}\otimes(\boldsymbol{B}\otimes\boldsymbol{C})=(\boldsymbol{A}\otimes\boldsymbol{B})\otimes\boldsymbol{C}$;

(5) 转置及共轭 $(\boldsymbol{A}\otimes\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\otimes\boldsymbol{B}^{\mathrm{T}},\overline{(\boldsymbol{A}\otimes\boldsymbol{B})}=\overline{\boldsymbol{A}}\otimes\overline{\boldsymbol{B}},(\boldsymbol{A}\otimes\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}\otimes\boldsymbol{B}^{\mathrm{H}}$;

(6) 混合积 当 $n=r,q=k$ 时,有 $(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})=(\boldsymbol{AC})\otimes(\boldsymbol{BD})$;

(7) 逆 当 $m=n,p=q$ 时,且 $\boldsymbol{A}^{-1},\boldsymbol{B}^{-1}$ 存在,则 $(\boldsymbol{A}\otimes\boldsymbol{B})^{-1}=\boldsymbol{A}^{-1}\otimes\boldsymbol{B}^{-1}$;

(8) 迹 当 $m=n,p=q$ 时,有 $\mathrm{tr}(\boldsymbol{A}\otimes\boldsymbol{B})=\mathrm{tr}\boldsymbol{A}\cdot\mathrm{tr}\boldsymbol{B}$;

(9) 秩 $\mathrm{rank}(\boldsymbol{A}\otimes\boldsymbol{B})=\mathrm{rank}(\boldsymbol{A})\cdot\mathrm{rank}(\boldsymbol{B})$;

(10) 行列式 当 $m=n,p=q$ 时,有 $\det(\boldsymbol{A}\otimes\boldsymbol{B})=(\det\boldsymbol{A})^{p}(\det\boldsymbol{B})^{m}$.

2. Kronecker 乘积的特征值

(1) (Stephanos 定理) 设 $\boldsymbol{A}\in\mathbf{C}^{m\times m}$ 的特征值是 $\lambda_r(r=1,2,\cdots,m)$,$\boldsymbol{B}\in\mathbf{C}^{n\times n}$ 的特征值是 $\mu_k(k=1,2,\cdots,n)$,令 $P(\boldsymbol{A},\boldsymbol{B})=\sum\limits_{i,j=1}^{k}c_{ij}\boldsymbol{A}^{i}\otimes\boldsymbol{B}^{j}$,$P(x,y)=\sum\limits_{i,j=1}^{k}c_{ij}x^{i}y^{j}$,则 $P(\boldsymbol{A},\boldsymbol{B})$ 的特征值是 $P(\lambda_r,\mu_k)$,$1\leqslant r\leqslant m,1\leqslant k\leqslant n$.

(2) 设 $\boldsymbol{A}\in\mathbf{C}^{m\times m}$ 的特征值是 $\lambda_r(r=1,2,\cdots,m)$,其相应的特征向量为 $\boldsymbol{x}_r(r=1,2,\cdots,m)$;另外设 $\boldsymbol{B}\in\mathbf{C}^{n\times n}$ 的特征值是 $\mu_k(k=1,2,\cdots,n)$,$\boldsymbol{y}_k(k=1,2,\cdots,n)$ 是其相应的特征向量,则 $\boldsymbol{A}\otimes_k\boldsymbol{B}$ 的特征值是 $\lambda_r+\mu_k$,相应的特征向量为 $\boldsymbol{x}_r\otimes\boldsymbol{y}_k$,$1\leqslant r\leqslant m,1\leqslant k\leqslant n$.

3. 向量化算符 Vec 的性质及应用

(1) $\mathrm{Vec}(k_1\boldsymbol{A}_1+k_2\boldsymbol{A}_2)=k_1\mathrm{Vec}(\boldsymbol{A}_1)+k_2\mathrm{Vec}(\boldsymbol{A}_2)$;

(2) 矩阵 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_k\in\mathbf{C}^{m\times n}$ 线性无关的充要条件是 $\mathrm{Vec}(\boldsymbol{A}_1),\mathrm{Vec}(\boldsymbol{A}_2),\cdots,\mathrm{Vec}(\boldsymbol{A}_k)$ 在 $\mathbf{C}^{m\times n}$ 中线性无关.

(3) 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n},\boldsymbol{X}\in\mathbf{C}^{n\times r},\boldsymbol{B}\in\mathbf{C}^{r\times s}$,则

$$\mathrm{Vec}(\boldsymbol{AXB})=(\boldsymbol{B}^{\mathrm{T}}\otimes\boldsymbol{A})\mathrm{Vec}(\boldsymbol{X}).$$

(4) (Kronecker 积的拟交换性) 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n},\boldsymbol{B}\in\mathbf{C}^{r\times s}$,则存在置换矩阵 $\boldsymbol{U}_1\in\mathbf{C}^{mr\times mr}$ 和 $\boldsymbol{U}_2\in\mathbf{C}^{ns\times ns}$,使得 $\boldsymbol{A}\otimes\boldsymbol{B}=\boldsymbol{U}_1(\boldsymbol{B}\otimes\boldsymbol{A})\boldsymbol{U}_2$.

(5) 设 $\boldsymbol{A}_i\in\mathbf{C}^{m\times p},\boldsymbol{X}\in\mathbf{C}^{p\times q},\boldsymbol{B}_i\in\mathbf{C}^{q\times n},\boldsymbol{F}\in\mathbf{C}^{m\times n}$,则矩阵方程 $\sum\limits_{i=1}^{l}\boldsymbol{A}_i\boldsymbol{X}\boldsymbol{B}_i=\boldsymbol{F}$ 有解的充分必要条件是 $\mathrm{Vec}(F)\in R\left(\sum\limits_{i=1}^{l}\boldsymbol{A}_i\otimes\boldsymbol{B}_i^{\mathrm{T}}\right)$.

(6) 设 $\boldsymbol{A}\in\mathbf{C}^{m\times m}$ 和 $\boldsymbol{B}\in\mathbf{C}^{n\times n}$ 的特征值分别为 $\lambda_i(i=1,2,\cdots,m)$ 和 $\mu_j(j=1,2,\cdots,n)$,$\boldsymbol{F}\in\mathbf{C}^{m\times n}$,则

① 矩阵方程 $\boldsymbol{AX}+\boldsymbol{XB}=\boldsymbol{F}$ 有唯一解的充要条件是

$$\lambda_i+\mu_j\neq 0\quad(i=1,2,\cdots,m;j=1,2,\cdots,n);$$

② 矩阵方程 $\boldsymbol{X}+\boldsymbol{AXB}+\cdots+\boldsymbol{A}^{l}\boldsymbol{X}\boldsymbol{B}^{l}=\boldsymbol{F}$ 有唯一解的充要条件是

$$1+(\lambda_i\mu_j)+\cdots+(\lambda_i\mu_j)^{l}\neq 0(i=1,2,\cdots,m;j=1,2,\cdots,n).$$

三、典型例题

例 1 设非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}(\boldsymbol{b}\neq\boldsymbol{0})$有解，其通解一般地可表示为

$$\boldsymbol{\xi}+k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r},$$

其中 $k_i(i=1,2,\cdots,n-r)$为任意实数. 试证：向量组$\boldsymbol{\xi},\boldsymbol{\xi}+\boldsymbol{\alpha}_1,\boldsymbol{\xi}+\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\xi}+\boldsymbol{\alpha}_{n-r}$是方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 所有解的极大线性无关组，但是 $\boldsymbol{Ax}=\boldsymbol{b}$ 所有解的集合不构成线性空间.

证 首先证$\boldsymbol{\xi},\boldsymbol{\xi}+\boldsymbol{\alpha}_1,\boldsymbol{\xi}+\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\xi}+\boldsymbol{\alpha}_{n-r}$线性无关. 不妨设

$$l_0\boldsymbol{\xi}+l_1(\boldsymbol{\xi}+\boldsymbol{\alpha}_1)+l_2(\boldsymbol{\xi}+\boldsymbol{\alpha}_2)+\cdots+l_{n-r}(\boldsymbol{\xi}+\boldsymbol{\alpha}_{n-r})=\boldsymbol{0},$$

即

$$(l_0+l_1+l_2+\cdots+l_{n-r})\boldsymbol{\xi}+l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_{n-r}\boldsymbol{\alpha}_{n-r}=\boldsymbol{0}.$$

用矩阵 $\boldsymbol{A}$ 左乘上式两端，并由 $\boldsymbol{A\alpha}_i=\boldsymbol{0},\boldsymbol{A\xi}=\boldsymbol{b}$，可得

$$(l_0+l_1+l_2+\cdots+l_{n-r})\boldsymbol{b}=\boldsymbol{0},$$

由 $\boldsymbol{b}\neq\boldsymbol{0}$，知 $l_0+l_1+l_2+\cdots+l_{n-r}=0$，故

$$l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_{n-r}\boldsymbol{\alpha}_{n-r}=\boldsymbol{0}.$$

又因为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系，所以

$$l_0=l_1=l_2=\cdots=l_{n-r}=0.$$

即$\boldsymbol{\xi},\boldsymbol{\xi}+\boldsymbol{\alpha}_1,\boldsymbol{\xi}+\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\xi}+\boldsymbol{\alpha}_{n-r}$线性无关.

设$\boldsymbol{\beta}$为 $\boldsymbol{Ax}=\boldsymbol{b}$ 的任意解，由题意知，存在 $k_i(i=1,2,\cdots,n-r)$使得

$$\begin{aligned}\boldsymbol{\beta}&=\boldsymbol{\xi}+k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r}\\&=(1-k_1-\cdots-k_{n-r})\boldsymbol{\xi}+k_1(\boldsymbol{\alpha}_1+\boldsymbol{\xi})+\cdots+k_{n-r}(\boldsymbol{\alpha}_{n-r}+\boldsymbol{\xi}).\end{aligned}$$

这表明$\boldsymbol{\beta}$可由$\boldsymbol{\xi},\boldsymbol{\xi}+\boldsymbol{\alpha}_1,\boldsymbol{\xi}+\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\xi}+\boldsymbol{\alpha}_{n-r}$线性表示，即$\boldsymbol{\xi},\boldsymbol{\xi}+\boldsymbol{\alpha}_1,\boldsymbol{\xi}+\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\xi}+\boldsymbol{\alpha}_{n-r}$是方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 所有解的极大线性无关组.

显然，$\boldsymbol{Ax}=\boldsymbol{b}$ 的所有解之和，解的任意常数倍都不是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解，所以 $\boldsymbol{Ax}=\boldsymbol{b}$ 所有解的集合不构成线性空间.

注 此题意在表明一个无穷集合可以有极大线性无关组，但其自身并不一定是线性空间.

例 2 设 W_1,W_2 为线性空间 V 的两个子空间，令

$$W_1\cap W_2=\{X\mid X\in W_1\text{ 且 }X\in W_2\},W_1\cup W_2=\{X\mid X\in W_1\text{ 或 }X\in W_2\}.$$

问：$W_1\cap W_2,W_1\cup W_2$ 是否分别都构成子空间？如果能构成子空间，请给出证明；如果不能，请举出反例.

证 (1) $W_1\cap W_2$ 是 V 的子空间，证明如下. 因 W_1,W_2 为 V 的线性子空间，均含有零元，因而 $W_1\cap W_2$ 是非空集.

再假定 $\alpha,\beta\in W_1\cap W_2$，则 $\alpha,\beta\in W_1,\alpha,\beta\in W_2$. 因 W_1,W_2 为子空间，故

$$\alpha+\beta\in W_1,\quad\alpha+\beta\in W_2,$$

于是 $\alpha+\beta\in W_1\cap W_2$. 同样，因 $\lambda\alpha\in W_1,\lambda\alpha\in W_2$，故 $\lambda\alpha\in W_1\cap W_2$，因而 $W_1\cap W_2$ 是 V 的子空间.

(2) $W_1\cup W_2$ 不一定是 V 的子空间. 反例如下. 取 $V=\mathbf{R}^2$，令 $W_1=\{(x,0)^{\mathrm{T}}\mid x\in\mathbf{R}\}$，$W_2=\{(0,y)^{\mathrm{T}}\mid y\in\mathbf{R}\}$，则它们均为 $\mathbf{R}^2$ 的子空间，取 $\boldsymbol{\alpha}=(1,0)^{\mathrm{T}}$，$\boldsymbol{\beta}=(0,1)^{\mathrm{T}}$，显然 $\boldsymbol{\alpha},\boldsymbol{\beta}\in W_1\cup W_2$，但 $\boldsymbol{\alpha}+\boldsymbol{\beta}=(1,0)^{\mathrm{T}}+(0,1)^{\mathrm{T}}=(1,1)^{\mathrm{T}}\notin W_1\cup W_2$，因而 $W_1\cup W_2$ 不是 $\mathbf{R}^2$ 的子空间.

注 本题第二部分实际上对任意有限个子空间也是成立的，即若 $W_1,W_2,\cdots,W_s$ 为线性空间 V 的 s 个非平凡子空间，则 V 中至少有一个向量不属于 $W_1,W_2,\cdots,W_s$ 中任何一个.

例 3 设 $\boldsymbol{A}$ 为 n 维线性空间 V 中线性变换 A 关于某基的矩阵，求证 $\operatorname{rank}(\boldsymbol{A}^2)=\operatorname{rank}(\boldsymbol{A})$ 当且仅当 $V=R(A)\oplus N(A)$.

证 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 V 的一组基，A 在这组基下的矩阵为 $\boldsymbol{A}$，则线性变换 A^2 在这组基下的矩阵为 $\boldsymbol{A}^2$，且 $R(A)=L(A\alpha_1,A\alpha_2,\cdots,A\alpha_n)=L(A\alpha_{i_1},A\alpha_{i_2},\cdots,A\alpha_{i_s})$，其中 $A\alpha_{i_1},A\alpha_{i_2},\cdots,A\alpha_{i_s}$ 是 $R(A)$ 的一组基. 于是 $R(A^2)=L(A^2\alpha_{i_1},A^2\alpha_{i_2},\cdots,A^2\alpha_{i_s})$.

设 $\operatorname{rank}(\boldsymbol{A}^2)=\operatorname{rank}(\boldsymbol{A})$，则 $\dim R(A)=\dim R(A^2)$. 于是 $A^2\alpha_{i_1},A^2\alpha_{i_2},\cdots,A^2\alpha_{i_s}$ 是 $R(A^2)$ 的一组基. 任取 $\alpha\in R(A)$ 且 $\alpha\in N(A)$，则 $\alpha=a_1A\alpha_{i_1}+a_2A\alpha_{i_2}+\cdots+a_sA\alpha_{i_s}$，又 $A\alpha=0$，于是 $a_1A^2\alpha_{i_1}+a_2A^2\alpha_{i_2}+\cdots+a_sA^2\alpha_{i_s}=0$. 因为 $A^2\alpha_{i_1},A^2\alpha_{i_2},\cdots,A^2\alpha_{i_s}$ 线性无关，我们有 $a_i=0,1\leqslant i\leqslant s$. 因此 $\alpha=0$，即有 $R(A)\cap N(A)=\{0\}$，又 $\dim R(A)+\dim N(A)=\dim V$，于是必有 $V=R(A)\oplus N(A)$.

反之，设 $V=R(A)\oplus N(A)$. 任取 $\alpha\in R(A)$，有 V 中的元素 β，使 $\alpha=A\beta$. 由直和的定义知，存在 $\beta_1\in V$ 和 $\gamma\in N(A)$，使 $\beta=A\beta_1+\gamma$. 于是有 $\alpha=A\beta=A^2\beta_1$，即 $\alpha=A^2\beta_1\in R(A^2)$，故 $R(A)\subseteq R(A^2)$. 显然 $R(A^2)\subseteq R(A)$，这样 $R(A^2)=R(A)$，即 $\operatorname{rank}(\boldsymbol{A}^2)=\operatorname{rank}(\boldsymbol{A})$.

注 由该题结论可知，若存在正整数 $m>1$，使 $\boldsymbol{A}^m=\boldsymbol{A}$，则有 $V=R(A)\oplus N(A)$.

例 4 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 分别是 $n\times s$ 和 $s\times p$ 矩阵. n 维行向量 $\boldsymbol{X}$ 满足 $\boldsymbol{XAB}=\mathbf{0}$. 令 $V=\{\boldsymbol{Y}\mid\boldsymbol{Y}=\boldsymbol{XA},\boldsymbol{XAB}=\mathbf{0}\}$. 求证 V 是一个线性空间，且 $\dim V=\operatorname{rank}(\boldsymbol{A})-\operatorname{rank}(\boldsymbol{AB})$.

证 不难直接验证 V 构成一个线性空间. 设 $W=\{\boldsymbol{X}\mid\boldsymbol{XAB}=\mathbf{0}\}$，则 W 是方程组 $\boldsymbol{XAB}=\mathbf{0}$ 的解空间，于是 $\dim W=n-\operatorname{rank}(\boldsymbol{AB})$. 令 $m=n-\operatorname{rank}(\boldsymbol{AB})$.

令 $W_0=\{\boldsymbol{X}\mid\boldsymbol{XAB}=\mathbf{0},\boldsymbol{XA}=\mathbf{0}\}$，则 $W_0=\{\boldsymbol{X}\mid\boldsymbol{XA}=\mathbf{0}\}$. 于是 W_0 是方程组 $\boldsymbol{XA}=\mathbf{0}$ 的解空间. 这样 $\dim W_0=n-\operatorname{rank}(\boldsymbol{A})=r$.

显然 $W_0\subseteq W$. 取 W_0 的一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$，扩充为 W 的一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\alpha}_{r+1},\cdots,\boldsymbol{\alpha}_m$. $V=\{\boldsymbol{Y}\mid\boldsymbol{Y}=\boldsymbol{XA},\boldsymbol{XAB}=\mathbf{0}\}$，易见

$$V=L(\boldsymbol{\alpha}_1\boldsymbol{A},\boldsymbol{\alpha}_2\boldsymbol{A},\cdots,\boldsymbol{\alpha}_r\boldsymbol{A},\boldsymbol{\alpha}_{r+1}\boldsymbol{A},\cdots,\boldsymbol{\alpha}_m\boldsymbol{A}).$$

又 $\boldsymbol{\alpha}_i\boldsymbol{A}=\mathbf{0},1\leqslant i\leqslant r$，于是

$$V=L(\boldsymbol{\alpha}_{r+1}\boldsymbol{A},\cdots,\boldsymbol{\alpha}_m\boldsymbol{A}).$$

下面证明 $\boldsymbol{\alpha}_{r+1}\boldsymbol{A},\cdots,\boldsymbol{\alpha}_m\boldsymbol{A}$ 线性无关. 设 $k_{r+1}\boldsymbol{\alpha}_{r+1}\boldsymbol{A}+k_{r+2}\boldsymbol{\alpha}_{r+2}\boldsymbol{A}+\cdots+k_m\boldsymbol{\alpha}_m\boldsymbol{A}=\mathbf{0}$，则

$$(k_{r+1}\boldsymbol{\alpha}_{r+1}+k_{r+2}\boldsymbol{\alpha}_{r+2}+\cdots+k_m\boldsymbol{\alpha}_m)\boldsymbol{A}=\mathbf{0}.$$

于是 $k_{r+1}\boldsymbol{\alpha}_{r+1}+k_{r+2}\boldsymbol{\alpha}_{r+2}+\cdots+k_m\boldsymbol{\alpha}_m\in W_0$，因此

$$k_{r+1}\boldsymbol{\alpha}_{r+1}+k_{r+2}\boldsymbol{\alpha}_{r+2}+\cdots+k_m\boldsymbol{\alpha}_m=l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_r\boldsymbol{\alpha}_r.$$

又$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性无关，因此有 $k_j=0,l_i=0,r+1\leqslant j\leqslant m,1\leqslant i\leqslant r$. 因此

$$\dim V=m-r,$$

即

$$\dim V=n-\operatorname{rank}(\boldsymbol{AB})-(n-\operatorname{rank}(\boldsymbol{A}))=\operatorname{rank}(\boldsymbol{A})-\operatorname{rank}(\boldsymbol{AB}).$$

例 5 证明在 n 维欧氏空间 V 中两两夹钝角(即夹角大于 $\pi/2$)的向量不能多于 $n+1$ 个.

证 对 n 作数学归纳法，$n=1$ 时结论显然成立，设对 $n-1$ 维欧氏空间结论成立.

当 $\dim V=n$ 时，若 V 内有 $n+2$ 个向量 $\alpha_1,\cdots,\alpha_{n+1},\alpha_{n+2}$ 两两夹钝角，则有

$$(\alpha_i,\alpha_j)<0,1\leqslant i<j\leqslant n+2.$$

令 $M=L(\alpha_{n+2})$，则 $V=M\oplus M^{\perp}$，对于 $i=1,2,\cdots,n+1$，设 $\alpha_i=k_i\alpha_{n+2}+\beta_i(\beta_i\in M^{\perp})$. 因 $(\alpha_i,\alpha_{n+2})=k_i(\alpha_{n+2},\alpha_{n+2})<0$，故 $k_i<0$. 又当 $i\neq j$ 时，$(\alpha_i,\alpha_j)=k_ik_j(\alpha_{n+2},\alpha_{n+2})+(\beta_i,\beta_j)<0$.

而 $k_ik_j(\alpha_{n+2},\alpha_{n+2})>0$，故 $(\beta_i,\beta_j)<0$. 这样 $M^{\perp}$ 是 $n-1$ 维欧氏空间，$\beta_1,\beta_2\cdots,\beta_{n+1}$ 是 $M^{\perp}$ 内 $n+1$ 个两两夹钝角的向量，与归纳假设矛盾.

例 6 $\boldsymbol{A}$ 是 n 阶实矩阵，证明存在正交矩阵 $\boldsymbol{T}$，使 $\boldsymbol{T}^{-1}\boldsymbol{AT}$ 是三角矩阵的充要条件为 $\boldsymbol{A}$ 的特征方程的根全是实数.

证 若 $\boldsymbol{T}^{-1}\boldsymbol{AT}$ 为三角矩阵，其中 $\boldsymbol{T}$ 为正交矩阵，设它对角线上的元素为 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 于是

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=|\boldsymbol{T}^{-1}||(\lambda\boldsymbol{E}-\boldsymbol{A})||\boldsymbol{T}|=|\lambda\boldsymbol{E}-\boldsymbol{T}^{-1}\boldsymbol{AT}|=\prod_{i=1}^{n}(\lambda-\lambda_i),$$

即 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的特征值. 由于 $\boldsymbol{A}$ 和 $\boldsymbol{T}$ 都是实矩阵，则 $\boldsymbol{T}^{-1}\boldsymbol{AT}$ 也是实矩阵，从而 $\boldsymbol{A}$ 的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 全为实数.

反之，我们对 n 用数学归纳法证明存在正交矩阵 $\boldsymbol{T}$，使 $\boldsymbol{T}^{-1}\boldsymbol{AT}$ 为三角矩阵. 当 $n=1$ 时，显然结论成立. 假设对 $n-1$ 结论成立，我们证明对 n 结论也成立.

设$\boldsymbol{\alpha}_1$是 $\boldsymbol{A}$ 属于 λ_1 的特征向量，不妨设$\boldsymbol{\alpha}_1$为单位向量. 设 $\boldsymbol{A}$ 的全部特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 将 α_1 扩充为 n 维欧氏空间的一组基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$. 设 $\boldsymbol{A}$ 为线性变换 σ 在某一标准正交基下的矩阵. 设 σ 在$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$下的矩阵为

$$\boldsymbol{H}=\begin{pmatrix}\lambda_1 & \boldsymbol{\beta}\\ \boldsymbol{0} & \boldsymbol{D}\end{pmatrix},$$

必存在正交矩阵 $\boldsymbol{M}$，使 $\boldsymbol{M}^{-1}\boldsymbol{AM}=\boldsymbol{H}$. $\boldsymbol{A}$ 与 $\boldsymbol{H}$ 有相同的特征值，从而 $\boldsymbol{D}$ 有特征值 $\lambda_2,\cdots,\lambda_n$. 由归纳法假定，存在正交矩阵 $\boldsymbol{G}$，使 $\boldsymbol{G}^{-1}\boldsymbol{DG}$ 为三角矩阵，且对角线上的元素为 $\lambda_2,\cdots,\lambda_n$. 不妨设 $\boldsymbol{G}^{-1}\boldsymbol{DG}$ 为上三角矩阵.

令 $\boldsymbol{Q}=\begin{pmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{G}\end{pmatrix}$，则 $\boldsymbol{Q}^{-1}=\boldsymbol{Q}^{\mathrm{T}}=\begin{pmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{G}^{\mathrm{T}}\end{pmatrix}$. 令 $\boldsymbol{T}=\boldsymbol{MQ}$，则 $\boldsymbol{T}$ 是正交矩阵，且

$$T^{\mathrm{T}}AT = Q^{\mathrm{T}}M^{\mathrm{T}}AMQ = Q^{\mathrm{T}}HQ = \begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & G^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\lambda_1 & \beta\\ \mathbf{0} & D\end{pmatrix}\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & G\end{pmatrix} = \begin{bmatrix}\lambda_1 & \beta G\\ \mathbf{0} & G^{\mathrm{T}}DG\end{bmatrix}.$$

于是存在正交矩阵 T，使 $T^{-1}AT=\begin{bmatrix}\lambda_1 & \beta G\\ \mathbf{0} & G^{\mathrm{T}}DG\end{bmatrix}$ 为三角矩阵.

例 7　设 $A\in \mathbf{C}^{n\times n}$ 是非奇异矩阵，那么 A^{-1} 相似于 A^* 当且仅当存在非奇异矩阵 $B\in \mathbf{C}^{n\times n}$，使得 $A=B^{-1}B^*$.

证　对于某个非奇异矩阵 $B\in \mathbf{C}^{n\times n}$，如果有 $A=B^{-1}B^*$，那么，$A^{-1}=(B^*)^{-1}B$，并且 $B^*A^{-1}(B^*)^{-1}=B(B^*)^{-1}=(B^{-1}B^*)^*=A^*$，这样，$A^{-1}$ 可经相似矩阵 B^* 相似于 A^*.

反过来，如果 A^{-1} 相似于 A^*，那么存在非奇异矩阵 $S\in \mathbf{C}^{n\times n}$，使得 $SA^{-1}S^{-1}=A^*$. 对 $\theta\in\mathbf{R}$，令 $S_\theta\equiv \mathrm{e}^{\mathrm{i}\theta}S$，且注意到 $S_\theta A^{-1}S_\theta^{-1}=\mathrm{e}^{\mathrm{i}\theta}SA^{-1}(\mathrm{e}^{-\mathrm{i}\theta}S^{-1})=SA^{-1}S^{-1}=A^*$. 另一方面，$S_\theta=A^*S_\theta A$，且 $S_\theta^*=A^*S_\theta^*A$. 相加这两个恒等式便得到 $H_\theta=A^*H_\theta A$，其中 $H_\theta\equiv S_\theta+S_\theta^*$ 是 Hermite 矩阵. 如果 H_θ 是奇异矩阵，那么存在某个非零 $x\in\mathbf{C}^n$，使得 $\mathbf{0}=H_\theta x=S_\theta x+S_\theta^* x$，因而，$-x=S_\theta^{-1}S_\theta^* x=\mathrm{e}^{-2\mathrm{i}\theta}S^{-1}S^*x$，且 $S^{-1}S^*x=-\mathrm{e}^{2\mathrm{i}\theta}x$. 选取值 $\theta=\theta_0\in[0,2\pi]$，使 $-\mathrm{e}^{2\mathrm{i}\theta_0}$ 不是 $S^{-1}S^*$ 的特征值，所得到的 Hermite 矩阵 $H\equiv H_{\theta_0}$ 是非奇异矩阵，且有性质 $H=A^*HA$.

现在，选取任一复数 α，使 $|\alpha|=1$，且 α 不是 A^* 的特征值. 令 $B\equiv\beta(\alpha E-A^*)H$，其中复参数 $\beta\neq 0$ 是有待选定的，不难看出 B 是非奇异矩阵. 为了使 $A=B^{-1}B^*$ 或 $BA=B^*$，经计算 $B^*=H(\overline{\beta\alpha}E-\overline{\beta}A)$，而 $BA=\beta(\alpha E-A^*)HA=\beta(\alpha HA-A^*HA)=\beta(\alpha HA-H)=H(\alpha\beta A-\beta E)$. 如果能选定一个非零 β 使 $\beta=-\overline{\beta\alpha}$，就完成了证明，而当 $\alpha=\mathrm{e}^{\mathrm{i}\varphi}$ 时，只要取 $\beta=\mathrm{e}^{\mathrm{i}(\pi-\varphi)}/2$ 就行.

例 8　任何 n 阶矩阵 A 的高次幂 $A^s(s\geqslant n)$ 必可表示为 A 的最高不超过 $n-1$ 次的多项式，当 A 是可逆矩阵时，A^{-1} 亦然.

证　不妨设 A 的特征多项式为

$$\Delta_A(\lambda)=|\lambda E-A|=\lambda^n-a_1\lambda^{n-1}+\cdots+(-1)^{n-1}a_{n-1}\lambda+(-1)^n a_n.$$

根据凯莱-哈密顿定理可得

$$\Delta_A(A)=A^n-a_1A^{n-1}+\cdots+(-1)^{n-1}a_{n-1}A+(-1)^n a_n E=\mathbf{0}.$$

当 $s\geqslant n$ 时，用 $\Delta_A(\lambda)$ 除 λ^s 可得

$$\lambda^s=q(\lambda)\cdot\Delta_A(\lambda)+r(\lambda),$$

其中 $r(\lambda)$ 的次数小于等于 $n-1$，从而 $A^s=q(A)\cdot\Delta_A(A)+r(A)=r(A)$.

当 A 可逆时，$a_n=|A|\neq 0$，此时从 $\Delta_A(A)=\mathbf{0}$ 可得

$$\frac{(-1)^{n-1}}{a_n}(A^{n-1}-a_1A^{n-2}+\cdots+(-1)^{n-1}a_{n-1}E)A=E,$$

由此得到

$$A^{-1}=\frac{(-1)^{n-1}}{a_n}(A^{n-1}-a_1A^{n-2}+\cdots+(-1)^{n-1}a_{n-1}E).$$

例 9　设 A,B 是两个 n 阶实正交矩阵，证明 $|AB|=1$ 当且仅当 $n-\mathrm{rank}(A+B)$ 为偶数.

证　因 $|A|,|B|\in\{\pm 1\}$，故 $|AB|=|A^{-1}B|$. 又因为

$$\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B})=\operatorname{rank}[\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{A}^{-1}\boldsymbol{B})]=\operatorname{rank}(\boldsymbol{E}+\boldsymbol{A}^{-1}\boldsymbol{B}).$$

令 $\boldsymbol{C}=\boldsymbol{A}^{-1}\boldsymbol{B}$,则原题可化为：证明$|\boldsymbol{C}|=1$ 当且仅当 $n-\operatorname{rank}(\boldsymbol{E}+\boldsymbol{C})$为偶数.

显然 $\boldsymbol{C}$ 是实正交矩阵,因此存在酉矩阵 $\boldsymbol{U}$,使得

$$\bar{\boldsymbol{U}}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{U}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{C}$ 的特征值,$|\lambda_1|=|\lambda_2|=\cdots=|\lambda_n|$.

由于 $\boldsymbol{C}$ 的虚特征值成对出现,故$|\boldsymbol{C}|=\lambda_1\lambda_2\cdots\lambda_n=(-1)^r$,$r$ 是$\lambda_1,\lambda_2,\cdots,\lambda_n$ 中等于-1 的 λ_i 的个数,而

$$\operatorname{rank}(\boldsymbol{E}+\boldsymbol{C})=\operatorname{rank}\begin{pmatrix}1+\lambda_1 & & & \\ & 1+\lambda_2 & & \\ & & \ddots & \\ & & & 1+\lambda_n\end{pmatrix}=n-r,$$

故$|\boldsymbol{C}|=1$,当且仅当 $n-\operatorname{rank}(\boldsymbol{E}+\boldsymbol{C})$为偶数.

例 10 (1) 设 $\boldsymbol{A}$ 为正定 Hermite 矩阵,$\boldsymbol{B}$ 为反 Hermite 矩阵. 试证明：$\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 的特征值实部为零.

(2) 设 $\boldsymbol{A}$ 是 $n(n>1)$阶正定矩阵. $\boldsymbol{\alpha}$ 是非零列向量,且$\boldsymbol{\alpha}\in\mathbf{R}^n$. 令 $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}$,求 $\boldsymbol{B}$ 的最大特征值以及 $\boldsymbol{B}$ 的属于这个特征值的特征子空间的维数和一个基.

证 (1) 首先,注意到 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 有着相同的非零特征值,因此我们只需证明它们中的任意一个成立即可. 不妨设 λ 为 $\boldsymbol{AB}$ 的任意一个特征值,于是$|\lambda\boldsymbol{E}-\boldsymbol{AB}|=0$. 由于 $\boldsymbol{A}$ 为正定 Hermite 矩阵,所以存在可逆矩阵 $\boldsymbol{Q}$,使得 $\boldsymbol{A}=\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q}$,将此式代入$|\lambda\boldsymbol{E}-\boldsymbol{AB}|=0$ 可得

$$\begin{aligned}0&=|\lambda\boldsymbol{E}-\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q}\boldsymbol{B}|\\&=|\lambda\boldsymbol{Q}^{\mathrm{H}}(\boldsymbol{Q}^{\mathrm{H}})^{-1}-\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q}\boldsymbol{B}|\\&=|\boldsymbol{Q}^{\mathrm{H}}|\,|\lambda\boldsymbol{E}-\boldsymbol{Q}\boldsymbol{B}\boldsymbol{Q}^{\mathrm{H}}|\,|(\boldsymbol{Q}^{\mathrm{H}})^{-1}|.\end{aligned}$$

从而有$|\lambda\boldsymbol{E}-\boldsymbol{Q}\boldsymbol{B}\boldsymbol{Q}^{\mathrm{H}}|=0$,这表明 λ 也是 $\boldsymbol{Q}\boldsymbol{B}\boldsymbol{Q}^{\mathrm{H}}$ 的特征值. 注意到 $\boldsymbol{Q}\boldsymbol{B}\boldsymbol{Q}^{\mathrm{H}}$ 是一个反 Hermite 矩阵,而反 Hermite 矩阵的特征值为零或纯虚数,所以 λ 为零或纯虚数.

(2) 由于 $\boldsymbol{A}$ 正定,$\boldsymbol{\alpha}\neq\boldsymbol{0}$,可得 $\lambda=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}>0$,$\boldsymbol{A}\boldsymbol{\alpha}\neq\boldsymbol{0}$,$\operatorname{rank}(\boldsymbol{B})=\operatorname{rank}(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha})=1$,由 $\operatorname{rank}(\boldsymbol{B})=1$,知方程组 $\boldsymbol{BX}=\boldsymbol{0}$ 解空间 W_0 的维数为 $n-1$,W_0 同时也是 $\boldsymbol{B}$ 的属于特征值 0 的特征子空间.

由 $\lambda>0$,$\boldsymbol{A}\boldsymbol{\alpha}\neq\boldsymbol{0}$ 和 $\boldsymbol{B}\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{A}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha})\boldsymbol{A}\boldsymbol{\alpha}=\lambda\boldsymbol{A}\boldsymbol{\alpha}$,知 λ 是 $\boldsymbol{B}$ 的特征值,$\boldsymbol{A}\boldsymbol{\alpha}$ 是 $\boldsymbol{B}$ 的属于特征值λ 的特征向量. 设 $\boldsymbol{B}$ 的属于这个特征值的特征子空间为 W_λ,由 $\lambda\neq0$,$W_\lambda\cap W_0=\{0\}$,所以

$$\dim W_\lambda+\dim W_0=\dim(W_\lambda+W_0)\leqslant n.$$

即 $\dim W_\lambda\leqslant1$,而 $\boldsymbol{A}\boldsymbol{\alpha}\neq\boldsymbol{0}$,$\boldsymbol{A}\boldsymbol{\alpha}\in W_\lambda$,$\dim W_\lambda=1$,$W_\lambda$ 的一组基为 $\boldsymbol{A}\boldsymbol{\alpha}$,且

$$\dim W_\lambda = 1 \Rightarrow \dim W_\lambda + \dim W_0 = n,$$

因此，$\lambda=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}>0$ 是 $\boldsymbol{B}$ 的唯一非零特征值，也是 $\boldsymbol{B}$ 的最大特征值.

例 11 设 $\boldsymbol{C}=\begin{pmatrix}\boldsymbol{0} & \boldsymbol{E}_{n-1}\\ 1 & \boldsymbol{0}\end{pmatrix}$，证明 $\boldsymbol{C}$ 是可以对角化的，并由此证明：

(1) 存在可逆矩阵 $\boldsymbol{P}$，使对任意循环矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1}\\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2}\\ \vdots & \vdots & \vdots & & \vdots\\ a_1 & a_2 & a_3 & \cdots & a_0\end{pmatrix},$$

满足 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角矩阵；

(2) 任意 n 阶矩阵 $\boldsymbol{M}$ 可以对角化当且仅当 $\boldsymbol{M}$ 与某个循环矩阵相似.

证 显然 $\boldsymbol{C}$ 的特征多项式 $\Delta_C(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{C}|=\lambda^n-1$ 没有重根，于是其最小多项式 $m_C(\lambda)=\Delta_C(\lambda)=\lambda^n-1$ 也没有重根，因此存在可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{P}=\begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_n\end{pmatrix},$$

其中 $w_1,w_2,\cdots,w_n$ 是 1 的全部 n 次单位根.

(1) 注意到 $\boldsymbol{A}=a_0\boldsymbol{E}+a_1\boldsymbol{C}+a_2\boldsymbol{C}^2+\cdots+a_{n-1}\boldsymbol{C}^{n-1}$，记 $f(\lambda)=a_0+a_1\lambda+\cdots+a_{n-1}\lambda^{n-1}$，则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{-1}f(\boldsymbol{C})\boldsymbol{P}=f(\boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{P})=\begin{pmatrix} f(w_1) & & & \\ & f(w_2) & & \\ & & \ddots & \\ & & & f(w_n)\end{pmatrix}$$

是对角矩阵.

(2) 当 $\boldsymbol{M}$ 相似于某个循环矩阵时，由(1)即知 $\boldsymbol{M}$ 是可以对角化的，反过来，若存在可逆矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{Q}^{-1}\boldsymbol{M}\boldsymbol{Q}=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

则由拉格朗日插值公式，存在次数小于等于 $n-1$ 的多项式 $g(\lambda)=b_0+b_1\lambda+\cdots+b_{n-1}\lambda^{n-1}$，使得

$$g(w_i)=\lambda_i, 1\leqslant i\leqslant n.$$

于是

$$\boldsymbol{P}^{-1}g(\boldsymbol{C})\boldsymbol{P}=\begin{pmatrix} g(w_1) & & & \\ & g(w_2) & & \\ & & \ddots & \\ & & & g(w_n) \end{pmatrix}=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

因此 $\boldsymbol{M}$ 相似于循环矩阵 $g(\boldsymbol{C})$.

例 12 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 均为 n 阶实对称矩阵，且 $\boldsymbol{B}$ 为正定矩阵，$\boldsymbol{A}-\boldsymbol{B}$ 为半正定矩阵，证明：$|\boldsymbol{A}|-|\boldsymbol{B}|\geqslant 0$.

证 由于 $\boldsymbol{B}$ 为正定的，$\boldsymbol{A}-\boldsymbol{B}$ 为半正定的，则存在可逆矩阵 $\boldsymbol{P}_1$，使 $\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}_1=\boldsymbol{E}$，且 $\boldsymbol{P}_1^{\mathrm{T}}(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{P}_1$ 也是对称的，故存在正交矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{P}_1^{\mathrm{T}}(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{P}_1\boldsymbol{Q}=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},\quad \lambda_i\geqslant 0.$$

取 $\boldsymbol{P}=\boldsymbol{P}_1\boldsymbol{Q}$，则 $\boldsymbol{P}$ 可逆，且

$$\begin{aligned} |\boldsymbol{P}^{\mathrm{T}}|(|\boldsymbol{B}|+|\boldsymbol{A}-\boldsymbol{B}|)|\boldsymbol{P}| &= |\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}|+|\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{P}| \\ &= 1+\lambda_1\lambda_2\cdots\lambda_n. \end{aligned}$$

从而

$$(|\boldsymbol{B}|+|\boldsymbol{A}-\boldsymbol{B}|)=(1+\lambda_1\lambda_2\cdots\lambda_n)\frac{1}{|\boldsymbol{P}|^2},$$

$$\begin{aligned} |\boldsymbol{P}^{\mathrm{T}}||\boldsymbol{A}||\boldsymbol{P}| &= |\boldsymbol{P}^{\mathrm{T}}||\boldsymbol{B}+\boldsymbol{A}-\boldsymbol{B}||\boldsymbol{P}| \\ &= |\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}+\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{P}| \\ &= (\lambda_1+1)(\lambda_2+1)\cdots(\lambda_n+1) \\ &\geqslant 1+\lambda_1\lambda_2\cdots\lambda_n. \end{aligned}$$

因此 $|\boldsymbol{A}|\geqslant(1+\lambda_1\lambda_2\cdots\lambda_n)\dfrac{1}{|\boldsymbol{P}|^2}$，故 $(|\boldsymbol{B}|+|\boldsymbol{A}-\boldsymbol{B}|)\leqslant|\boldsymbol{A}|$，从而 $|\boldsymbol{A}-\boldsymbol{B}|\leqslant|\boldsymbol{A}|-|\boldsymbol{B}|$，又因为 $\boldsymbol{A}-\boldsymbol{B}$ 为半正定的，所以 $|\boldsymbol{A}-\boldsymbol{B}|\geqslant 0$，故 $|\boldsymbol{A}|-|\boldsymbol{B}|\geqslant 0$.

例 13 求方程 $x^2+y^2=z^2$ 的整数解.

分析 将不定方程写成矩阵行列式的形式，由行列式为零知矩阵方程有非零解. 再由此解出不定方程的整数解.

解 首先，将该方程写成行列式形式

$$\det\begin{pmatrix} x & z-y \\ z+y & x \end{pmatrix}=0.$$

由行列式为零，知以上述矩阵为系数的矩阵方程有非零解. 故有不全为零的数 a 和 b，使得

$$\begin{pmatrix} x & z-y \\ z+y & x \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=\boldsymbol{0},$$

即

$$\begin{pmatrix} a & -b & b \\ b & a & a \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}.$$

解得

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda\begin{pmatrix} -2ab \\ b^2-a^2 \\ a^2+b^2 \end{pmatrix}, \quad \lambda \in \mathbf{R}.$$

于是,不定方程的整数解为

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{c}{\gcd(2ab, b^2-a^2, a^2+b^2)}\begin{pmatrix} -2ab \\ b^2-a^2 \\ a^2+b^2 \end{pmatrix},$$

其中 a,b 和 c 为任意整数,且 a 和 b 不全为零.

注 此题的更进一步形式即为 1637 年法国数学家费马(Fermat)提出的费马大定理:当自然数 $n>2$ 时,不定方程 $x^n+y^n=z^n$ 无正整数解.

四、习题解答

1. 设 $a_1, a_2, \cdots, a_n$ 是互不相同的 n 个数,且

$$f_i = (x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n), \quad i=1,2,\cdots,n,$$

(1) 证明:$f_i(i=1,2,\cdots,n)$是 $P[x]_n$ 的一组基底;

(2) 取 $a_1, a_2, \cdots, a_n$ 是全体 n 次单位根,求由基底 $1, x, \cdots, x^{n-1}$ 到 $f_1, f_2, \cdots, f_n$ 的过渡矩阵.

证 (1) 由题意,只需证 $f_1(x), f_2(x), \cdots, f_n(x)$线性无关即可. 现假设

$$k_1f_1(x)+k_2f_2(x)+\cdots+k_nf_n(x)=0.$$

分别以 $x=a_i$ 代入上式,得到 $k_if_i(a_i)=0$. 因为 $f_i(a_i)\neq 0$,故 $k_i=0, i=1,2,\cdots,n$. 由此可得 $f_1(x), f_2(x), \cdots, f_n(x)$线性无关. 又因为 $\dim P[x]_n=n$,可知 $f_1(x), f_2(x), \cdots, f_n(x)$ 为 $P[x]_n$ 的基.

(2) 取全部 n 次单位根为 $a_1=\varepsilon_1, a_2=\varepsilon_2, \cdots, a_n=\varepsilon_n$,则

$$x^n-1=(x-\varepsilon_1)(x-\varepsilon_2)\cdots(x-\varepsilon_n)=(x-\varepsilon_i)f_i(x), \quad i=1,2,\cdots,n.$$

于是可得

$$f_i(x)=\frac{x^n-1}{x-\varepsilon_i}=x^{n-1}+\varepsilon_ix^{n-2}+\varepsilon_i^2x^{n-3}+\cdots+\varepsilon_i^{n-1},$$

故所求过渡矩阵为

$$\begin{pmatrix} \varepsilon_1^{n-1} & \varepsilon_2^{n-1} & \cdots & \varepsilon_n^{n-1} \\ \varepsilon_1^{n-2} & \varepsilon_2^{n-2} & \cdots & \varepsilon_n^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

2. 设 V_1,V_2 是线性空间 V 的两个非平凡子空间,证明:在 V 中存在 α,使 $\alpha\notin V_1,\alpha\notin V_2$ 同时成立.

证 法 1 因为 V_1,V_2 都是非平凡子空间,故存在向量 α,β,使 $\alpha\notin V_1,\beta\notin V_2$. 如果有 $\alpha\notin V_2$,则 α 即为所求向量. 假设 $\alpha\in V_2$,则对数域 F 中任意数 k,有 $k\alpha+\beta\notin V_2$,若不然,则由于 $\alpha\in V_2$,故 $k\alpha\in V_2$,于是 $\beta=(k\alpha+\beta)-k\alpha\in V_2$,这与 $\beta\notin V_2$ 矛盾. 现取 $k_1\neq k_2$,则 $k_1\alpha+\beta$, $k_2\alpha+\beta$ 不可能都属于 V_1,若不然,则 $(k_1\alpha+\beta)-(k_2\alpha+\beta)=(k_1-k_2)\alpha\in V_1$,于是 $\alpha=\dfrac{1}{k_1-k_2}(k_1-k_2)\alpha\in V_1$,这与 $\alpha\notin V_1$ 矛盾,故 $k_1\alpha+\beta,k_2\alpha+\beta$ 中至少有一个不属于 V_1,同时它也不属于 V_2,即为所求向量.

法 2 因为 V_1 是非平凡子空间,故存在 $\alpha\notin V_1$,如果 $\alpha\notin V_2$,则结论成立. 若 $\alpha\in V_2$,则由于 V_2 也是非平凡子空间,故存在向量 $\beta\notin V_2$,若 $\beta\notin V_1$,则结论成立,若 $\beta\in V_1$,则有 $\alpha\notin V_1,\beta\in V_1;\alpha\in V_2,\beta\notin V_2$. 于是可推得 $\gamma=\alpha+\beta\notin V_1$,并且 $\gamma\notin V_2$.

3. 设 $f(x_1,x_2,\cdots,x_n)$ 是一个秩为 n 的二次型,证明:有 $\mathbf{R}^n$ 的一个 $\dfrac{1}{2}(n-|s|)$ 维子空间 V_1 存在(s 为符号差数),使对任意的 $(x_1,x_2,\cdots,x_n)\in V_1$,都有 $f(x_1,x_2,\cdots,x_n)=0$.

证 设 f 的正惯性指数为 p,负惯性指数为 q,$p+q=n$. 则存在可逆矩阵 $\boldsymbol{C}$,$\boldsymbol{Y}=\boldsymbol{CX}$,使得

$$f=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2,$$

因

$$\frac{1}{2}(n-|s|)=\frac{1}{2}(n-|p-q|)=\begin{cases} q, & \text{当 } p>q \text{ 时}, \\ p, & \text{当 } p\leqslant q \text{ 时}. \end{cases}$$

故下面仅就 $p\leqslant q$ 时给出证明(当 $p>q$ 时同理可证). 将 $\boldsymbol{Y}=\boldsymbol{CX}$ 展开可得方程组

$$\begin{cases} c_{1,1}x_1+c_{1,2}x_2+\cdots+c_{1,n}x_n=y_1, \\ \qquad\vdots \\ c_{p,1}x_1+c_{p,2}x_2+\cdots+c_{p,n}x_n=y_p, \\ c_{p+1,1}x_1+c_{p+1,2}x_2+\cdots+c_{p+1,n}x_n=y_{p+1}, \\ \qquad\vdots \\ c_{n,1}x_1+c_{n,2}x_2+\cdots+c_{n,n}x_n=y_n. \end{cases}$$

取向量

$$\boldsymbol{\varepsilon}_1=(1,0,\cdots,0,1,0,\cdots,0),$$
$$\boldsymbol{\varepsilon}_2=(0,1,\cdots,0,0,1,\cdots,0),$$
$$\vdots$$
$$\boldsymbol{\varepsilon}_p=(0,0,\cdots1,0,0,\cdots,1,\cdots,0).$$

显然$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_p$线性无关，代入$\boldsymbol{Y}=\boldsymbol{CX}$方程组中可得解$\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p$线性无关，且满足

$$\boldsymbol{Ca}_1=\boldsymbol{\varepsilon}_1,\boldsymbol{Ca}_2=\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{Ca}_p=\boldsymbol{\varepsilon}_p.$$

下面证p维子空间$L(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p)$即为所求空间$V_1$.

设$\boldsymbol{x}_0\in V_1$，且满足关系$\boldsymbol{x}_0=k_1\boldsymbol{a}_1+k_2\boldsymbol{a}_2+\cdots+k_p\boldsymbol{a}_p$，则

$$\boldsymbol{y}_0=\boldsymbol{Cx}_0=k_1\boldsymbol{Ca}_1+k_2\boldsymbol{Ca}_2+\cdots+k_p\boldsymbol{Ca}_p=k_1\boldsymbol{\varepsilon}_1+k_2\boldsymbol{\varepsilon}_2+\cdots+k_p\boldsymbol{\varepsilon}_p,$$

代入二次型f中可得

$$\begin{aligned}f&=\boldsymbol{Y}_0^{\mathrm{T}}\begin{pmatrix}\boldsymbol{E}_p&\boldsymbol{0}\\\boldsymbol{0}&-\boldsymbol{E}_q\end{pmatrix}\boldsymbol{Y}_0\\&=(k_1\boldsymbol{\varepsilon}_1+k_2\boldsymbol{\varepsilon}_2+\cdots+k_p\boldsymbol{\varepsilon}_p)^{\mathrm{T}}\begin{pmatrix}\boldsymbol{E}_p&\boldsymbol{0}\\\boldsymbol{0}&-\boldsymbol{E}_q\end{pmatrix}(k_1\boldsymbol{\varepsilon}_1+k_2\boldsymbol{\varepsilon}_2+\cdots+k_p\boldsymbol{\varepsilon}_p)\\&=(k_1^2+k_2^2+\cdots+k_p^2-k_1^2-k_2^2-\cdots-k_p^2-0-\cdots-0)\\&=0.\end{aligned}$$

4. 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是线性空间$V$的一组基，$T$是$V$上的线性变换. 证明：$T$可逆的充分必要条件是$T\varepsilon_1,T\varepsilon_2,\cdots,T\varepsilon_n$线性无关.

证 先证必要性. 若T可逆，则存在逆变换T^{-1}使得$TT^{-1}=T^{-1}T=E$. 且T^{-1}也是V的线性变换. 如果$T\varepsilon_1,T\varepsilon_2,\cdots,T\varepsilon_n$线性相关，则$T^{-1}(T\varepsilon_1),T^{-1}(T\varepsilon_2),\cdots,T^{-1}(T\varepsilon_n)$，即$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$也线性相关，这与假设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是基相矛盾，故$T\varepsilon_1,T\varepsilon_2,\cdots,T\varepsilon_n$线性无关.

下面证充分性. 若$T\varepsilon_1,T\varepsilon_2,\cdots,T\varepsilon_n$线性无关，则由于$V$是$n$维线性空间，所以它也是$V$的一组基，从而对$V$中任一向量$\alpha_1$，存在$k_1,k_2,\cdots,k_n$，使得

$$\begin{aligned}\alpha_1&=k_1(T\varepsilon_1)+k_2(T\varepsilon_2)+\cdots+k_n(T\varepsilon_n)\\&=T(k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n),\end{aligned}$$

即存在$\alpha=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n$，使$T\alpha=\alpha_1$，因此$T$为$V$的满射变换. 如果又有$\beta=l_1\varepsilon_1+l_2\varepsilon_2+\cdots+l_n\varepsilon_n$，满足$T\beta=\alpha_1$，即

$$\begin{aligned}T\beta&=l_1T\varepsilon_1+l_2T\varepsilon_2+\cdots+l_nT\varepsilon_n\\&=k_1T\varepsilon_1+k_2T\varepsilon_2+\cdots+k_nT\varepsilon_n,\end{aligned}$$

那么，因$T\varepsilon_1,T\varepsilon_2,\cdots,T\varepsilon_n$是基，所以$l_i=k_i(i=1,2,\cdots,n)$，即$\alpha=\beta$. 因此，$T$又是单射变换，从而$T$为可逆变换.

5. 设T是线性空间V上的线性变换，如果$T^{k-1}(\xi)\neq 0$，但$T^k(\xi)=0$. 求证$\xi,T(\xi),\cdots,T^{k-1}(\xi)(k>0)$线性无关.

证 (1) 当$k=1$时，$k-1=0$，此时只有一个向量ξ，显然线性无关.

(2) 当$k>1$时，设有$t_0,t_2,\cdots,t_{k-1}$使得

$$t_0\xi+t_1T(\xi)+t_2T^2(\xi)+\cdots+t_{k-1}T^{k-1}(\xi)=0,\qquad (*)$$

对上式两边同时作用 T^{k-1} 得

$$t_0T^{k-1}(\xi)+t_1T^k(\xi)+t_2TT^k(\xi)+\cdots+t_{k-1}T^{k-2}T^k(\xi)=0.$$

由 $T^{k-1}(\xi)\neq 0,T^k(\xi)=0$，得 $t_0=0$，代入(*)式得

$$t_1T(\xi)+t_2T^2(\xi)+\cdots+t_{k-1}T^{k-1}(\xi)=0.$$

再对上式两边作用 T^{k-2} 可得 $t_1=0$，这样一直继续下去，最后可得 $t_i=0,i=0,1,\cdots,k-1$. 故 $\xi,T(\xi),\cdots,T^{k-1}(\xi)$ 线性无关.

6. (1) 设 λ_1,λ_2 是线性变换 T 的两个不同的特征值，$\varepsilon_1,\varepsilon_2$ 是分别属于 λ_1,λ_2 的特征向量，证明：$\varepsilon_1+\varepsilon_2$ 不是 T 的特征向量.

(2) 证明：如果线性空间 V 中每一个非零向量都是线性变换 T 的特征向量，则 T 是数乘变换.

证 (1) 由题意可知 $T\varepsilon_1=\lambda_1\varepsilon_1,T\varepsilon_2=\lambda_2\varepsilon_2$，因 T 是线性变换，则有 $T(\varepsilon_1+\varepsilon_2)=\lambda_1\varepsilon_1+\lambda_2\varepsilon_2$. 假设 $\varepsilon_1+\varepsilon_2$ 是 T 的特征向量，与之对应的特征值为 λ，则 $T(\varepsilon_1+\varepsilon_2)=\lambda(\varepsilon_1+\varepsilon_2)$，故有

$$T(\varepsilon_1+\varepsilon_2)=\lambda(\varepsilon_1+\varepsilon_2)=\lambda_1\varepsilon_1+\lambda_2\varepsilon_2.$$

即 $(\lambda-\lambda_1)\varepsilon_1+(\lambda-\lambda_2)\varepsilon_2=0$，因 λ_1,λ_2 是 T 的两个不同的特征值，故 $\varepsilon_1,\varepsilon_2$ 线性无关，由此可得 $\lambda=\lambda_1=\lambda_2$，矛盾，假设不成立. 故 $\varepsilon_1+\varepsilon_2$ 不是 T 的特征向量.

(2) 取 V 的一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，由题意可设 $T\varepsilon_i=\lambda_i\varepsilon_i,i=0,1,\cdots,n$. 对任意 $i=0,1,\cdots,n$，若 $\lambda_1\neq\lambda_2$，则由(1)得，$\varepsilon_1+\varepsilon_2$ 不是 T 的特征向量，与题设不符，故

$$\lambda_i=\lambda_1,\quad i=1,2,\cdots,n.$$

因此，$T\varepsilon_i=\lambda_1\varepsilon_i,i=0,1,\cdots,n$. 对 V 中任意向量 $\alpha=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n$，有

$$\begin{aligned}T\alpha&=k_1T\varepsilon_1+k_2T\varepsilon_2\cdots+k_nT\varepsilon_n\\&=k_1\lambda_1\varepsilon_1+k_2\lambda_1\varepsilon_2+\cdots+k_n\lambda_1\varepsilon_n=\lambda_1\alpha.\end{aligned}$$

所以 $T=\lambda_1E$ 是数乘变换.

7. 设 V 是复数域上的 n 维线性空间，T_1,T_2 是 V 上的线性变换，且 $T_1T_2=T_2T_1$，证明：

(1) 如果 λ_0 是 T_1 的特征值，则 V_{λ_0} 是 T_2 的不变子空间；

(2) T_1,T_2 至少有一个公共的特征向量.

证 (1) 由题意，对任意的 $\alpha\in V_{\lambda_0}$，$T_1(T_2\alpha)=T_2(T_1\alpha)=T_2\lambda_0\alpha=\lambda_0T_2\alpha$，因此 $T_2\alpha\in V_{\lambda_0}$，故 V_{λ_0} 也是 T_2 的不变子空间.

(2) 由于 T_2 是复数域上的 n 维线性变换，故 T_2 必有特征值，不妨令 λ_2 是 T_2 的一个特征值，由(1)知 V_{λ_2} 是 T_1 的不变子空间，设 $T_1|V_{\lambda_2}$ 是 T_1 在 V_{λ_2} 上的诱导线性变换，于是 $T_1|V_{\lambda_2}$ 有特征值 μ 及相应的特征向量 $\beta\in V_{\lambda_2}$，使 $T_1\beta=T_1|V_{\lambda_2}(\beta)=\mu\beta$，又因为 $\beta\in V_{\lambda_2}$，故 $T_2\beta=\lambda_2\beta$，所以 β 是 T_1 与 T_2 的公共特征向量.

8. 设 V 是复线性空间，而线性变换 T 在基底 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是一 Jordan 块，证明：

(1) V 中包含 ε_n 的不变子空间只有 V 本身；

(2) V 中任一不变子空间都包含 ε_1；

(3) V 不能分解成两个非平凡的不变子空间的直和.

证 (1) 设 T 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}\lambda & & & & \\ 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda\end{pmatrix},$$

亦即有

$$\begin{aligned}T\varepsilon_1&=\lambda\varepsilon_1+\varepsilon_2,\\ T\varepsilon_2&=\lambda\varepsilon_2+\varepsilon_3,\\ &\vdots\\ T\varepsilon_{n-1}&=\lambda\varepsilon_{n-1}+\varepsilon_n,\\ T\varepsilon_n&=\lambda\varepsilon_n.\end{aligned}\tag{**}$$

设 W 是 V 的任意一个包含 ε_1 的不变子空间，则 $\lambda\varepsilon_1,T\varepsilon_1\in W$，从而由(**)式中第一式知

$$\varepsilon_2=T\varepsilon_1-\lambda\varepsilon_1\in W.$$

再由第二式得

$$\varepsilon_3=T\varepsilon_2-\lambda\varepsilon_2\in W.$$

如此继续下去，可推得 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 都属于 W，从而 $W=V$.

(2) 设 W 是任意一个非零不变子空间，故有 $a\in W$，且 $a\neq 0$. 设

$$a=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n,$$

且 k_i 为第一个不等于零的系数，即

$$a=k_i\varepsilon_i+\cdots+k_n\varepsilon_n\in W.$$

由于 W 是对 T 不变的子空间，故 $Ta\in W$，于是

$$(T-\lambda I)a=Ta-\lambda a\in W.$$

其中 I 表示恒等变换. 但由(**)式得

$$(T-\lambda I)\varepsilon_i=\varepsilon_{i+1},\quad (T-\lambda I)\varepsilon_n=0,$$

故有

$$\begin{aligned}(T-\lambda I)a&=k_i(T-\lambda I)\varepsilon_i+k_{i+1}(T-\lambda I)\varepsilon_{i+1}+\cdots+k_n(T-\lambda I)\varepsilon_n\\ &=k_i\varepsilon_{i+1}+k_{i+1}\varepsilon_{i+2}+\cdots+k_{n-1}\varepsilon_n\\ &=a_1\in W.\end{aligned}$$

同样由 $a_1\in W$，有

$$(T-\lambda I)a_1=Ta_1-\lambda a_1\in W,$$

即

$$(T-\lambda I)^2 a=(T-\lambda I)a_1=k_i\varepsilon_{i+2}+k_{i+1}\varepsilon_{i+3}+\cdots+k_{n-2}\varepsilon_n$$
$$=a_2\in W.$$

继续这个过程，一般地有

$$(T-\lambda I)^l a=k_i\varepsilon_{i+l}+k_{i+1}\varepsilon_{i+l+1}+\cdots+k_{n-l}\varepsilon_n\in W,\quad l=1,2,\cdots,n-i.$$

特别当 $l=n-i$ 时，有

$$(T-\lambda I)^{n-i}a=k_i\varepsilon_n\in W.$$

但 $k_i\neq 0$，所以 $\varepsilon_n\in W$，即 W 包含 ε_n.

(3) 若不然，设 $V=V_1\oplus V_2$，其中 V_1,V_2 是 V 的两个非平凡不变子空间，则由(2)知

$$\varepsilon_n\in V_1,\varepsilon_n\in V_2,$$

这与 $V_1\cap V_2=\{0\}$ 矛盾.

9. 设 T 是线性空间 V 上的线性变换，证明：T 对应的矩阵的行列式为零的充要条件是 T 以零作为一个特征值.

证 设线性变换 T 对应的矩阵为 $\boldsymbol{A}$，则 T 的特征值之积为 $|\boldsymbol{A}|$，令 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 T 的所有特征值，则 $\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$，于是

$$|\boldsymbol{A}|=0\Leftrightarrow\lambda_1\lambda_2\cdots\lambda_n=0\Leftrightarrow \text{至少有一个 } \lambda_i=0.$$

即 T 以零作为一个特征值.

10. 如果 $T_1,T_2,\cdots,T_s$ 是线性空间 V 的 s 个两两不同的线性变换，则 V 中必存在向量 α，使 $T_1\alpha,T_2\alpha,\cdots,T_s\alpha$ 也两两不同.

证 令 $V_{ij}=\{x\mid x\in V,T_ix=T_jx\}$，其中 $i<j,i,j=1,2,\cdots,s$，易证 V_{ij} 为 V 的子空间. 因为 $T_1,T_2,\cdots,T_s$ 两两不相同，故对 $T_i,T_j(i<j)$，总存在向量 β，使 $T_i\beta\neq T_j\beta$，即 $\beta\notin V_{ij}$，从而 $V_{ij}\neq V,i,j=1,2,\cdots,s$. 设$\{V_{ij}\}$中有 m 个是零空间，其余的都是非平凡子空间，故存在共同向量 α 不属于每个非平凡子空间，又因为 α 显然不能是零向量，因此 α 当然也不属于零空间，故α 不属于每个 V_{ij}，即 $T_i\alpha\neq T_j\alpha$，亦即 $T_1\alpha,T_2\alpha,\cdots,T_s\alpha$ 两两不相同.

11. 证明：反对称实矩阵的特征值是零或纯虚数.

证 不妨设此实反对称矩阵为 $\boldsymbol{A}$，其属于特征值 λ 的特征向量为 $\boldsymbol{x}$，即 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$. 两端左乘 $\boldsymbol{x}^{\mathrm{H}}$，可得 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{Ax}=\lambda\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}$. 两端再取共轭转置，并利用 $\boldsymbol{A}$ 为实反对称矩阵，可得 $-\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}=\bar{\lambda}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}$. 从而有 $(\lambda-\bar{\lambda})\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=0$. 因为 $\boldsymbol{x}\neq\boldsymbol{0}$，所以 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}\neq 0$，于是有 $\lambda-\bar{\lambda}=0$，即 λ 为零或纯虚数.

12. 设 η 是欧氏空间中一单位向量，定义

$$T(\alpha)=\alpha-2(\eta,\alpha)\eta.$$

证明：(1) T 是正交变换，这样的正交变换称为镜面反射；

(2) T 是第二类的(即 T 对应的矩阵的行列式为 -1)；

(3) 如果 n 维欧氏空间中，正交变换 T 以 1 作为一个特征值，且属于特征值 1 的特征子空间 V_1 的维数是 $n-1$，则 T 是镜面反射.

证 (1) 对欧氏空间中任意元素 α,β 和实数 k_1,k_2,有

$$\begin{aligned}T(k_1\alpha+k_2\beta)&=k_1\alpha+k_2\beta-2(\eta,k_1\alpha+k_2\beta)\eta\\&=k_1\alpha+k_2\beta-2k_1(\eta,\alpha)\eta-2k_2(\eta,\beta)\eta\\&=k_1T\alpha+k_2T\beta,\end{aligned}$$

所以 T 是线性变换.

又因为

$$\begin{aligned}(T\alpha,T\beta)&=[\alpha-2(\eta,\alpha)\eta,\beta-2(\eta,\beta)\eta]\\&=(\alpha,\beta)-2(\eta,\alpha)(\eta,\beta)-2(\eta,\alpha)(\eta,\beta)+4(\eta,\alpha)(\eta,\beta)(\eta,\eta),\end{aligned}$$

注意到$(\eta,\eta)=1$,所以$(T\alpha,T\beta)=(\alpha,\beta)$,故 T 是正交变换.

(2) 由于 η 是单位向量,将它扩充成欧氏空间的一组标准正交基 $\eta,\varepsilon_2,\cdots,\varepsilon_n$,则

$$\begin{cases}T\eta=\eta-2(\eta,\eta)\eta=-\eta,\\T\varepsilon_i=\varepsilon_i-2(\eta,\varepsilon_i)\eta=\varepsilon_i,\quad i=2,3,\cdots,n.\end{cases}$$

即

$$T(\eta,\varepsilon_2,\cdots,\varepsilon_n)=(\eta,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}-1&&&\\&1&&\\&&\ddots&\\&&&1\end{pmatrix}.$$

可见 T 在基 $\eta,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵的行列式等于-1,所以 T 是第二类的.

(3) 既然 T 的特征值有 n 个,由已知,T 有 $n-1$ 个特征值为 1,另一个也须为实数,不妨设为 λ_0,则存在一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,使得

$$T(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}\lambda_0&&&\\&1&&\\&&\ddots&\\&&&1\end{pmatrix}.$$

因为 T 是正交变换,所以$(\varepsilon_1,\varepsilon_1)=(T\varepsilon_1,T\varepsilon_1)=\lambda_0^2(\varepsilon_1,\varepsilon_1)$. 从而 $\lambda_0^2=1$. 但 V_1 是 $n-1$ 维的,所以 $\lambda_0=-1$,于是

$$T\varepsilon_1=-\varepsilon_1,T\varepsilon_i=\varepsilon_i,\quad i=2,3,\cdots,n.$$

因为属于实对称矩阵的不同特征值的特征向量必正交,所以

$$(\varepsilon_1,\varepsilon_i)=0,\quad i=2,3,\cdots,n.$$

现令 $\eta=\dfrac{1}{|\varepsilon_1|}\varepsilon_1$,则 η 是单位向量,且与 $\varepsilon_2,\cdots,\varepsilon_n$ 正交,则 $\eta,\varepsilon_2,\cdots,\varepsilon_n$ 为欧氏空间 V 的一组基. 又因为

$$T\eta=T\left(\frac{1}{|\varepsilon_1|}\varepsilon_1\right)=\frac{1}{|\varepsilon_1|}T\varepsilon_1=\frac{1}{|\varepsilon_1|}(-\varepsilon_1)=-\eta,$$

任取 $\alpha=k_1\eta+k_2\varepsilon_2+\cdots+k_n\varepsilon_n\in V$，有

$$(\alpha,\eta)=(k_1\eta+k_2\varepsilon_2+\cdots+k_n\varepsilon_n,\eta)=k_1,$$

故

$$\begin{aligned}T\alpha&=k_1T\eta+k_2T\varepsilon_2+\cdots+k_nT\varepsilon_n=-k_1\eta+k_2\varepsilon_2+\cdots+k_n\varepsilon_n\\&=k_1\eta+k_2\varepsilon_2+\cdots+k_n\varepsilon_n-2k_1\eta\\&=\alpha-2(\alpha,\eta)\eta.\end{aligned}$$

可见 T 是镜面反射.

13. 设 $\boldsymbol{A}$ 是 n 阶实对称矩阵. 证明：$\boldsymbol{A}$ 正定的充要条件是 $\boldsymbol{A}$ 的特征值全大于零.

证 设二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 经过正交变换 $\boldsymbol{X}=\boldsymbol{T}\boldsymbol{Y}$，可使得

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2,$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值. 由于 $\boldsymbol{A}$ 为正定的充分必要条件是 $\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2$ 正定，而后者为正定的充分必要条件是 $\lambda_i>0(i=1,2,\cdots,n)$，得证.

14. 设 $\boldsymbol{A}$ 是 n 阶实矩阵. 证明：存在正交矩阵 $\boldsymbol{Q}$ 使 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ 为三角矩阵的充要条件是 $\boldsymbol{A}$ 的特征值全为实数.

证 为确定起见，这里三角矩阵不妨设为上三角矩阵.

先证必要性. 不妨设

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\&c_{22}&\cdots&c_{2n}\\&&\ddots&\vdots\\&&&c_{nn}\end{pmatrix},$$

其中 $\boldsymbol{Q},\boldsymbol{A}$ 均为实矩阵，从而 c_{ij} 都是实数. 又因为相似矩阵有相同的特征多项式，所以

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=|\lambda\boldsymbol{E}-\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}|=\begin{vmatrix}\lambda-c_{11}&-c_{12}&\cdots&-c_{1n}\\&\lambda-c_{22}&\cdots&-c_{2n}\\&&\ddots&\vdots\\&&&\lambda-c_{nn}\end{vmatrix}$$

$$=(\lambda-c_{11})(\lambda-c_{22})\cdots(\lambda-c_{nn}).$$

从而 $\boldsymbol{A}$ 的 n 个特征值 $c_{11},c_{22},\cdots,c_{nn}$ 均为实数.

再证充分性. 设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 为 $\boldsymbol{A}$ 的所有不同的实特征值，则 $\boldsymbol{A}$ 与某一 Jordan 形矩阵 $\boldsymbol{J}$ 相似，即存在可逆实矩阵 $\boldsymbol{P}_0$，使

$$\boldsymbol{P}_0^{-1}\boldsymbol{A}\boldsymbol{P}_0=\boldsymbol{J},$$

其中 $\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1&&&\\&\boldsymbol{J}_2&&\\&&\ddots&\\&&&\boldsymbol{J}_s\end{pmatrix}$，且

$$J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}, \quad i = 1, 2, \cdots, s.$$

由于 λ_i 都是实数，所以 $\boldsymbol{J}$ 为上三角实矩阵.

另一方面，矩阵 $\boldsymbol{P}_0$ 可以分解为

$$\boldsymbol{P}_0 = \boldsymbol{Q}_0 \boldsymbol{S}_0.$$

其中 $\boldsymbol{Q}_0$ 是正交矩阵，$\boldsymbol{S}_0$ 为上三角矩阵，于是

$$\boldsymbol{P}_0^{-1} \boldsymbol{A} \boldsymbol{P}_0 = \boldsymbol{S}_0^{-1} \boldsymbol{Q}_0^{-1} \boldsymbol{A} \boldsymbol{Q}_0 \boldsymbol{S}_0 = \boldsymbol{J},$$

即

$$\boldsymbol{Q}_0^{-1} \boldsymbol{A} \boldsymbol{Q}_0 = \boldsymbol{S}_0 \boldsymbol{J} \boldsymbol{S}_0^{-1}.$$

由于 $\boldsymbol{S}_0, \boldsymbol{J}, \boldsymbol{S}_0^{-1}$ 都是上三角矩阵，因而它们的积也为上三角矩阵，即充分性得证.

15. 设 $\boldsymbol{A}, \boldsymbol{B}$ 都是实对称矩阵. 证明：存在正交矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{B}$ 的充要条件是 $\boldsymbol{A}, \boldsymbol{B}$ 的特征值全相等.

证 必要性是显然的，因为相似矩阵有相同的特征值.

现证充分性. 设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的特征值，则它们也是 $\boldsymbol{B}$ 的特征值. 于是存在正交矩阵 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$，使得

$$\boldsymbol{X}^{-1} \boldsymbol{A} \boldsymbol{X} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = \boldsymbol{Y}^{-1} \boldsymbol{B} \boldsymbol{Y}.$$

所以 $\boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X}\boldsymbol{Y}^{-1} = \boldsymbol{B}$，令 $\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{Y}^{-1}$，则 $\boldsymbol{Q}$ 也是正交矩阵，从而 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{B}$.

16. 证明：如果 T 是正交变换，则 T 的不变子空间的正交补也是 T 的不变子空间.

证 设 W 是 T 的任意一个不变子空间，下面证 $W^{\perp}$ 也是 T 的不变子空间.

取 $\xi_1, \xi_2, \cdots, \xi_m$ 是 W 的一组标准正交基，将其扩充成 n 维空间的一组标准正交基 $\xi_1, \xi_2, \cdots, \xi_m, \xi_{m+1}, \xi_{m+2}, \cdots, \xi_n$，则

$$W = L(\xi_1, \xi_2, \cdots, \xi_m), \quad W^{\perp} = L(\xi_{m+1}, \xi_{m+2}, \cdots, \xi_n).$$

因为 T 是正交变换，所以 $T\xi_1, T\xi_2, \cdots, T\xi_n$ 也是一组标准正交基，又由于 W 是 T 的任意一个不变子空间，$T\xi_1, T\xi_2, \cdots, T\xi_m \in W$，且为 W 的一组标准正交基，于是

$$T\xi_{m+1}, T\xi_{m+2}, \cdots, T\xi_n \in W^{\perp}.$$

任取 $\alpha = k_{m+1}\xi_{m+1} + \cdots + k_n\xi_n \in W^{\perp}$，那么

$$T\alpha = k_{m+1} T\xi_{m+1} + k_{m+2} T\xi_{m+2} + \cdots + k_n T\xi_n \in W^{\perp}.$$

故 $W^{\perp}$ 是 T 的不变子空间.

17. 设 $\boldsymbol{A}$ 是 n 阶实对称矩阵，且 $\boldsymbol{A}^2=\boldsymbol{E}$. 证明：存在正交矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{E}_{n-r}\end{pmatrix}.$$

证 法 1 因为 $\boldsymbol{A}$ 是 n 阶实对称矩阵，所以存在 n 阶矩阵 $\boldsymbol{T}$，使得

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

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的 n 个特征值(重根按重数列出). 于是

$$\begin{aligned}\boldsymbol{T}^{-1}\boldsymbol{A}^2\boldsymbol{T}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}&=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\cdot\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\\&=\mathrm{diag}(\lambda_1^2,\lambda_2^2,\cdots,\lambda_n^2).\end{aligned}$$

又因为 $\boldsymbol{A}^2=\boldsymbol{E}$，所以

$$\boldsymbol{E}=\boldsymbol{T}^{-1}\boldsymbol{E}\boldsymbol{T}=\boldsymbol{T}^{-1}\boldsymbol{A}^2\boldsymbol{T}=\mathrm{diag}(\lambda_1^2,\lambda_2^2,\cdots,\lambda_n^2).$$

因此有 $\lambda_i=\pm1(i=1,2,\cdots,n)$. 不妨设 $\lambda_i=1$ 的重数为 r，则 $\lambda_i=-1$ 的重数为 $n-r$. 只要通过置换将 $\lambda_i=1$ 集中排列在前面，则有正交矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{E}_{n-r}\end{pmatrix}.$$

法 2 因为 $\boldsymbol{A}$ 是 n 阶实对称矩阵，且 $\boldsymbol{A}^2=\boldsymbol{E}$，若令 $g(x)=x^2-1$，则 $g(x)$ 为 $\boldsymbol{A}$ 的零多项式，且无重根，故 $\boldsymbol{A}$ 相似于对角矩阵. 设 λ 为 $\boldsymbol{A}$ 的任一特征值，则 $\lambda=\pm1$. 不妨设 $\lambda_i=-1$ 的重数为 $n-r$. 只要将 $\lambda_i=1$ 集中排列在前面，则有正交矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{E}_{n-r}\end{pmatrix}.$$

18. 设 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 是一实二次型，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的特征值，且 $\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$. 证明：对 $\forall\boldsymbol{X}\in\mathbf{R}^n$，有

$$\lambda_1\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\leqslant\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\leqslant\lambda_n\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}.$$

证 由题意知 $\boldsymbol{A}$ 是实对称矩阵，故存在正交矩阵 $\boldsymbol{Q}$，使得

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

其中 $\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$ 为 $\boldsymbol{A}$ 的 n 个特征值. 于是 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}-\lambda_1\boldsymbol{E}$ 的特征值非负，即矩阵 $\boldsymbol{A}-\lambda_1\boldsymbol{E}$ 半正定. 于是对任意的实向量 $\boldsymbol{X}$，有 $\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{A}-\lambda_1\boldsymbol{E})\boldsymbol{X}\geqslant0$. 故 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\geqslant\lambda_1\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$. 同理可证 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\leqslant\lambda_n\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$. 所以

$$\lambda_1\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\leqslant\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\leqslant\lambda_n\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}.$$

19. 设二次型 $f(x_1,x_2,\cdots,x_n)$ 对应的矩阵为 $\boldsymbol{A}$，λ 是 $\boldsymbol{A}$ 的特征值. 证明：存在 $\mathbf{R}^n$ 中的非零向量 $(\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n)$，使 $f(\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n)=\lambda(\tilde{x}_1^2+\tilde{x}_2^2+\cdots+\tilde{x}_n^2)$.

证 设 λ 是矩阵 $\boldsymbol{A}$ 的特征值，则存在属于 λ 的非零特征向量 $\boldsymbol{\varepsilon}=(\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n)$，使得 $\boldsymbol{A}\boldsymbol{\varepsilon}=\lambda\boldsymbol{\varepsilon}$，左乘 $\boldsymbol{\varepsilon}^{\mathrm{T}}$，得

$$\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}^{\mathrm{T}}\lambda\boldsymbol{\varepsilon}=\lambda\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\varepsilon},$$

即

$$f(\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n)=\lambda(\tilde{x}_1^2+\tilde{x}_2^2+\cdots+\tilde{x}_n^2).$$

20. 设 α,β 是欧氏空间中两个不同的单位向量. 证明：存在一个镜面反射 T,使 $T(\alpha)=\beta$.

证 记 n 维欧氏空间为 V,当 η 为欧氏空间 V 的单位向量时,由

$$T(\gamma)=\gamma-2(\eta,\gamma)\eta(\gamma\in V)$$

所确定的正交变换 T 是一个镜面反射,代入单位向量 α,有 $T(\alpha)=\alpha-2(\eta,\alpha)\eta$,若记 $\beta=\alpha-2(\eta,\alpha)\eta$,则 $\alpha-\beta=2(\eta,\alpha)\eta$,因为 α,β 是欧氏空间中两个不同的单位向量,所以 $(\eta,\alpha)\neq0$,故可解得 $\eta=\dfrac{\alpha-\beta}{2(\eta,\alpha)}$,又有

$$\begin{aligned}(\eta,\alpha)&=\left(\frac{\alpha-\beta}{2(\eta,\alpha)},\alpha\right)=\frac{1}{2(\eta,\alpha)}(\alpha-\beta,\alpha)\\&=\frac{1}{2(\eta,\alpha)}[(\alpha,\alpha)-(\alpha,\beta)]\\&=\frac{1}{2(\eta,\alpha)}[1-(\alpha,\beta)],\end{aligned}$$

即 $(\eta,\alpha)^2=\dfrac{1}{2}[1-(\alpha,\beta)]$. 于是只要取 $\eta=\dfrac{\alpha-\beta}{\sqrt{2[1-(\alpha,\beta)]}}$,就有 $(\eta,\eta)=1$,即 η 为欧氏空间 V 中的单位向量,从而 T 是一个镜面反射,且 $T(\alpha)=\alpha-2(\eta,\alpha)\eta=\beta$.

21. 证明：n 维欧氏空间中任一正交变换都可以表示成一系列镜面反射的乘积.

证 设 A 是 n 维欧氏空间 V 上的任一正交变换,取 V 的一组标准正交基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$,则 $\eta_1=A(\varepsilon_1),\eta_2=A(\varepsilon_2),\cdots,\eta_n=A(\varepsilon_n)$ 也是 V 的一组标准正交基,此时,若 $\eta_1=\varepsilon_1,\eta_2=\varepsilon_2,\cdots,\eta_n=\varepsilon_n$,则 A 是一个恒等变换,只要作镜面反射

$$A_1(\gamma)=\gamma-2(\varepsilon_1,\gamma)\varepsilon_1,\quad\forall\gamma\in V,$$

则有 $A_1(\varepsilon_1)=-\varepsilon_1,A_1(\varepsilon_j)=\varepsilon_j(j=2,3,\cdots,n)$,且 $A=A_1A_1$,结论成立.

若 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 与 $\eta_1,\eta_2,\cdots,\eta_n$ 不全相同,不妨设 $\eta_1\neq\varepsilon_1$,则 η_1,ε_1 为两个不同的单位向量,由上面第 20 题知,存在镜面反射 A_1,使 $A_1(\varepsilon_1)=\eta_1$. 令 $A_1(\varepsilon_j)=\xi_j(j=2,3,\cdots,n)$,若 $\xi_j=\eta_j(j=2,3,\cdots,n)$,则 $A=A_1$,结论成立. 否则可设 $\xi_2\neq\eta_2$,再作镜面反射 $A_2(\gamma)=\gamma-2(\eta,\gamma)\eta(\forall\gamma\in V)$,其中 $\eta=\dfrac{\xi_2-\eta_2}{\sqrt{2[1-(\xi_2-\eta_2)]}}$,则 $A_2(\xi)=\eta_2$,因为 $(\eta_1,\eta_2)=0,(\xi_2,\eta_1)=0$,所以

$$(\eta,\eta_1)=\left(\frac{\xi_2-\eta_2}{\sqrt{2[1-(\xi_2-\eta_2)]}},\eta_1\right)=0,$$

且 $A_2(\eta_1)=\eta_1-2(\eta,\eta_1)\eta=\eta_1,A_2(\xi_i)=\xi_i(i=3,\cdots,n)$. 如此继续下去,设

$$\varepsilon_1,\varepsilon_2\cdots,\varepsilon_n\xrightarrow{A_1}\eta_1,\quad\xi_2,\cdots,\xi_n\xrightarrow{A_2}\eta_1,\eta_2,$$

$$\xi_2,\cdots,\xi_n\xrightarrow{A_3}\cdots\xrightarrow{A_s}\eta_1,\eta_2,\cdots,\eta_n(s\leqslant n),$$

则 $A=A_sA_{s-1}\cdots A_2A_1$，其中 $A_j(j=1,2,\cdots,s)$都是镜面反射，证毕.

22. 设 $\boldsymbol{A},\boldsymbol{B}$ 是两个 n 阶实对称矩阵，且 $\boldsymbol{B}$ 是正交矩阵. 证明：存在 n 阶实可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$ 与 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{BP}$ 同时为对角矩阵.

证 因为 $\boldsymbol{B}$ 是正定矩阵，所以存在一个 n 阶实可逆矩阵 $\boldsymbol{C}$，使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{BC}=\boldsymbol{E}$，其中 $\boldsymbol{E}$ 为 n 阶单位矩阵. 又因为 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 还是 n 阶实对称矩阵，所以也存在一个 n 阶实正交矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC})\boldsymbol{Q}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 的特征值，于是只要令 $\boldsymbol{P}=\boldsymbol{CQ}$，则 $\boldsymbol{P}$ 为可逆矩阵，此时有

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}=\boldsymbol{Q}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC})\boldsymbol{Q}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),$$

且

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{BP}=\boldsymbol{Q}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{BC})\boldsymbol{Q}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{E}.$$

23. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，$\boldsymbol{x},\boldsymbol{y}\in\mathbf{C}^n$，$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}$，则$(\boldsymbol{Ax},\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{A}^{\mathrm{H}}\boldsymbol{y})$.

证 既然对 $\forall\boldsymbol{x},\boldsymbol{y}\in\mathbf{C}^n$，有$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}$，则

$$(\boldsymbol{Ax},\boldsymbol{y})=\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{y}=(\boldsymbol{x},\boldsymbol{A}^{\mathrm{H}}\boldsymbol{y}).$$

24. 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，则

(1) $N(\boldsymbol{A})=N(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$，$N(\boldsymbol{A}^{\mathrm{H}})=N(\boldsymbol{AA}^{\mathrm{H}})$；

(2) $R(\boldsymbol{A})=R(\boldsymbol{AA}^{\mathrm{H}})$，$R(\boldsymbol{A}^{\mathrm{H}})=R(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$；

(3) $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\mathrm{rank}(\boldsymbol{AA}^{\mathrm{H}})$.

证 (1) 设 $\boldsymbol{x}\in N(\boldsymbol{A})$，则 $\boldsymbol{Ax}=\boldsymbol{0}$，两边同时左乘 $\boldsymbol{A}^{\mathrm{H}}$，可得 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{0}=\boldsymbol{0}$，故 $\boldsymbol{x}\in N(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$.

反之，若 $\boldsymbol{x}\in N(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$，假设 $\boldsymbol{x}\notin N(\boldsymbol{A})$，则 $\boldsymbol{Ax}\neq\boldsymbol{0}$. 由 $\boldsymbol{x}\in N(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$，得 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Ax}=\boldsymbol{0}$，即存在 $\boldsymbol{y}=\boldsymbol{Ax}\neq\boldsymbol{0}$，使得 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{y}=\boldsymbol{0}$，由此可得 $\boldsymbol{y}\in R(\boldsymbol{A})$，且 $\boldsymbol{y}\in N(\boldsymbol{A}^{\mathrm{H}})$，又因为 $R(\boldsymbol{A})\perp N(\boldsymbol{A}^{\mathrm{H}})$，故 $(\boldsymbol{y},\boldsymbol{y})=0$，与 $\boldsymbol{y}\neq\boldsymbol{0}$ 矛盾，所以假设不成立.

综上可得 $N(\boldsymbol{A})=N(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$. 同理可证 $N(\boldsymbol{A}^{\mathrm{H}})=N(\boldsymbol{AA}^{\mathrm{H}})$.

(2) 若 $\boldsymbol{y}\in R(\boldsymbol{A})$，假设 $\boldsymbol{y}\notin R(\boldsymbol{AA}^{\mathrm{H}})$，故不存在 $\boldsymbol{z}$，使得 $\boldsymbol{y}=\boldsymbol{AA}^{\mathrm{H}}\boldsymbol{z}$，即不存在 $\boldsymbol{x}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{z}$，使得 $\boldsymbol{y}=\boldsymbol{Ax}$，与 $\boldsymbol{y}\in R(\boldsymbol{A})$矛盾，故假设不成立，所以 $\boldsymbol{y}\in R(\boldsymbol{AA}^{\mathrm{H}})$.

反之，若 $\boldsymbol{y}\in R(\boldsymbol{AA}^{\mathrm{H}})$，则存在 $\boldsymbol{z}$，使得 $\boldsymbol{y}=\boldsymbol{AA}^{\mathrm{H}}\boldsymbol{z}$，即存在 $\boldsymbol{x}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{z}$，使得 $\boldsymbol{y}=\boldsymbol{Ax}$，故 $\boldsymbol{y}\in R(\boldsymbol{A})$. 综上可证 $R(\boldsymbol{A})=R(\boldsymbol{AA}^{\mathrm{H}})$，同理可证 $R(\boldsymbol{A}^{\mathrm{H}})=R(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$.

(3) 因为 $\mathrm{rank}(\boldsymbol{A})=\dim R(\boldsymbol{A})$，故由(2)的结论显然可得 $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\mathrm{rank}(\boldsymbol{AA}^{\mathrm{H}})$.

25. 求证：方程 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}$ 对于任意的 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{b}\in\mathbf{C}^m$ 一定有解.

证 对任意 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$ 和任意 $\boldsymbol{b}\in\mathbf{C}^m$，显然有 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}\in R(\boldsymbol{A}^{\mathrm{H}})$. 由于 $R(\boldsymbol{A}^{\mathrm{H}})=R(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$(见第 24 题(2))，从而 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}\in R(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$，即方程 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}$ 一定有解.

26. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$是正交投影，则 $\boldsymbol{A}$ 的特征值非 0 即 1.

证 设 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量为 $\boldsymbol{x}$，即 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$，则有 $\boldsymbol{A}^2\boldsymbol{x}=\lambda^2\boldsymbol{x}$. 因为 $\boldsymbol{A}$ 是正交投影，所以 $\boldsymbol{A}^2=\boldsymbol{A}$，进而有 $\lambda^2-\lambda=0$，即 $\lambda=0$ 或 1.

27. 若 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}^2\in\mathbf{C}^{n\times n}$，$\mathrm{rank}(\boldsymbol{A})=r$，则存在酉阵 $\boldsymbol{U}\in\mathbf{C}^{n\times n}$，使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}.$$

证　既然 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}^2\in\mathbf{C}^{n\times n}$，类似于第 26 题的证明，可知矩阵 $\boldsymbol{A}$ 的特征值非 0 即 1. 又 $\mathrm{rank}(\boldsymbol{A})=r$，则 $\boldsymbol{A}$ 的值为 1 的特征值个数为 r. 又 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$，可知存在酉阵 $\boldsymbol{U}\in\mathbf{C}^{n\times n}$，使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}.$$

第二章 向量与矩阵的范数

在研究 P^n(P 为数域)上的向量或者 $P^{m\times n}$ 上的矩阵时,经常说向量或矩阵的"大"和"小",这是以什么为依据的呢?这里的"大"和"小"又是指的是什么呢?在什么情况下,可以说两个向量或矩阵"很接近"或"相差很远"呢?

要回答这些问题,可以用到的一个主要方法就是本章研究的内容——向量范数与矩阵范数.范数可以看作是我们通常所说的空间的长度概念的推广.在数学上,这种推广是不可缺少的;而且在分析和评价有关数值计算的各种算法中,它也是必要的.因此,研究范数的各种性质及其应用是相当重要的.

本章将介绍范数的基本概念、主要结论、典型例题和课后习题等内容.

一、基本概念

1. 向量范数

定义 1 设映射 $\|\cdot\|:\mathbf{C}^n\to\mathbf{R}$ 满足

(1) 正定条件 $\|\boldsymbol{x}\|\geqslant 0$,当且仅当 $\boldsymbol{x}=\mathbf{0}$ 时 $\|\boldsymbol{x}\|=0$;

(2) 齐次条件 $\|\lambda\boldsymbol{x}\|=|\lambda|\,\|\boldsymbol{x}\|$,对所有 $\lambda\in\mathbf{C},\boldsymbol{x}\in\mathbf{C}^n$ 成立;

(3) 三角不等式 $\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|,\ \forall\boldsymbol{x},\boldsymbol{y}\in\mathbf{C}^n$.

则称映射 $\|\cdot\|$ 为 $\mathbf{C}^n$ 上的向量 $\boldsymbol{x}$ 的范数,定义了范数的空间 $\mathbf{C}^n$ 又叫做一个线性赋范空间.

向量范数是用来刻画向量大小的一种度量,它是以线性空间的元素为自变量,且满足非负性、齐次性和三角不等式的一类实值函数.下面给出几种常见的 $\mathbf{C}^n$ 上的向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的范数.

(1) 1-范数 $\|\boldsymbol{x}\|_1=\sum\limits_{i=1}^{n}|x_i|$;

(2) 2-范数 $\|\boldsymbol{x}\|_2=\left(\sum\limits_{i=1}^{n}|x_i|^2\right)^{1/2}$;

(3) ∞-范数 $\|\boldsymbol{x}\|_\infty=\max\limits_{1\leqslant i\leqslant n}|x_i|$;

(4) p-范数 $\|\boldsymbol{x}\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p},1\leqslant p<\infty$；

(5) 椭圆范数 $\|\boldsymbol{x}\|=\sqrt{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}}$，$\boldsymbol{A}$ 为 Hermite 正定矩阵.

其中 p-范数也称为 Hölder 范数. 通过简单计算不难看出，当 $p=1$ 时，$\|\boldsymbol{x}\|_p$ 即为 $\|\boldsymbol{x}\|_1$；当 $p=2$ 时，$\|\boldsymbol{x}\|_p$ 即为 $\|\boldsymbol{x}\|_2$；当 $p\to\infty$ 时，$\|\boldsymbol{x}\|_p\to\|\boldsymbol{x}\|_\infty$. 这说明，前三种范数为 Hölder 范数的特殊情形，并且 $\mathbf{C}^n$ 上的向量范数不唯一.

定义 2 设 P 为数域，在 $V^n(P)$ 上定义了 $\|\cdot\|_a$，$\|\cdot\|_b$ 两种向量范数，若存在常数 $C_1>0$，$C_2>0$，使得

$$\|\boldsymbol{x}\|_a\leqslant C_1\|\boldsymbol{x}\|_b,\ \|\boldsymbol{x}\|_b\leqslant C_2\|\boldsymbol{x}\|_a,\quad \forall \boldsymbol{x}\in V^n(P),$$

则称 $V^n(P)$ 上的两个范数 $\|\cdot\|_a$ 与 $\|\cdot\|_b$ 等价.

2. 矩阵范数

定义 3 设 $\boldsymbol{A}\in P^{m\times n}$，若映射 $\|\cdot\|: P^{m\times n}\to\mathbf{R}$ 满足

(1) 正定性 $\|\boldsymbol{A}\|>0,\forall \boldsymbol{0}\neq\boldsymbol{A}\in P^{m\times n}$；

(2) 齐次性 $\|\lambda\boldsymbol{A}\|=|\lambda|\,\|\boldsymbol{A}\|,\forall\lambda\in\mathbf{R},\forall\boldsymbol{A}\in P^{m\times n}$；

(3) 三角不等式 $\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|,\forall\boldsymbol{A},\boldsymbol{B}\in P^{m\times n}$.

则称映射 $\|\cdot\|$ 为 $P^{m\times n}$ 上的矩阵范数.

几种 $P^{m\times n}$ 上常见的矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 的矩阵范数规纳如下：

(1) m_1-范数 $\|\boldsymbol{A}\|_{m_1}=\sum_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|$；

(2) m_2-范数(F-范数) $\|\boldsymbol{A}\|_{m_2}=\left(\sum_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|^2\right)^{1/2}$；

(3) m_∞-范数 $\|\boldsymbol{A}\|_{m_1}=\max\limits_{i,j}\{|a_{ij}|\},1\leqslant i\leqslant m,1\leqslant j\leqslant n$；

(4) G-范数 $\|\boldsymbol{A}\|_{\mathrm{G}}=\sqrt{mn}\max\limits_{i,j}\{|a_{ij}|\},1\leqslant i\leqslant m,1\leqslant j\leqslant n$.

其中范数 $\|\cdot\|$ 的下标 m 表示矩阵(matrix)的意思. 很显然，矩阵 m_1-范数，m_2-范数和 m_∞-范数分别是向量 1-范数，2-范数和 ∞-范数的推广.

定义 4 设 $\|\cdot\|_a: P^{m\times l}\to\mathbf{R}$，$\|\cdot\|_b: P^{l\times n}\to\mathbf{R}$ 和 $\|\cdot\|_c: P^{m\times n}\to\mathbf{R}$ 是矩阵范数. 如果

$$\|\boldsymbol{AB}\|_c\leqslant\|\boldsymbol{A}\|_a\|\boldsymbol{B}\|_b,\quad\forall\boldsymbol{A}\in P^{m\times l},\quad\forall\boldsymbol{B}\in P^{l\times n},$$

则称 $\|\cdot\|_a$，$\|\cdot\|_b$ 和 $\|\cdot\|_c$ 是相容的. 特别地，如果

$$\|\boldsymbol{AB}\|\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{B}\|,$$

则称 $\|\cdot\|$ 是自相容的矩阵范数，或简称 $\|\cdot\|$ 是相容的矩阵范数.

3. 算子范数

定义 5 设 $\|\cdot\|_a$ 是 P^n 上的向量范数，$\|\cdot\|_m$ 是 $P^{n\times n}$ 上的矩阵范数，且

$$\|Ax\|_a \leqslant \|A\|_m \|x\|_a, \quad A \in P^{n\times n}, \quad x \in P^n,$$

则称 $\|\cdot\|_m$ 为与向量范数 $\|\cdot\|_a$ 相容的矩阵范数.

定义 6 设 $\|\cdot\|_a$ 是 P^n 上的向量范数，$\|\cdot\|_m$ 是 $P^{n\times n}$ 上的矩阵范数，则

$$\|A\|_a = \max_{x\neq 0} \frac{\|Ax\|_a}{\|x\|_a} \left(= \max_{\|u\|=1} \|Au\|_a\right)$$

是与向量范数 $\|x\|_a$ 相容的矩阵范数. 称此矩阵范数为从属于向量范数 $\|x\|_a$ 的算子范数.

定义 7 设 $A\in \mathbf{C}^{n\times n}$，则 $r(A)=\max\limits_i|\lambda_i|$ 称为方阵 A 的谱半径，其中 $\lambda_i(i=1,2,\cdots,n)$ 为矩阵 A 的特征值.

算子范数的计算归结为求函数的约束极值，因此求解算子范数比较困难. 下面给出几个特殊情况下的算子范数：

(1) 从属于向量范数 $\|x\|_1=\sum\limits_{i=1}^n |x_i|$ 的算子范数（又称为列和范数）$\|A\|_1=\max\limits_j\sum\limits_{i=1}^n |a_{ij}|$；

(2) 从属于向量范数 $\|x\|_\infty=\max\limits_i|x_i|$ 的算子范数（又称为行和范数）$\|A\|_\infty=\max\limits_i\sum\limits_{j=1}^n |a_{ij}|$；

(3) 从属于向量范数 $\|x\|_2$ 的算子范数（又称为谱范数）$\|A\|_2=\sqrt{r(A^{\mathrm{H}}A)}$，其中 $r(A)$ 为矩阵 $A\in\mathbf{C}^{n\times n}$ 的谱半径.

定义 8 设 $\|\cdot\|_a$，$\|\cdot\|_b$ 均是 P^n 上的向量范数，$\|\cdot\|_{a,b}$ 是 $P^{n\times n}$ 上的矩阵范数，且

$$\|Ax\|_a \leqslant \|A\|_{a,b} \|x\|_b, \quad \forall A \in P^{n\times n}, \quad \forall x \in P^n,$$

则称 $\|\cdot\|_{a,b}$ 是与向量范数 $\|\cdot\|_a$，$\|\cdot\|_b$ 相容的矩阵范数.

定义 9 设 $\|\cdot\|_a$，$\|\cdot\|_b$ 均是 P^n 上的向量范数，且 $A\in P^{n\times n}$，则

$$\|A\|_{a,b} = \max_{x\neq 0} \frac{\|Ax\|_b}{\|x\|_a} \left(= \max_{\|u\|=1} \|Au\|_a\right)$$

是与向量范数 $\|\cdot\|_a$，$\|\cdot\|_b$ 相容的矩阵范数，叫做 $P^{n\times n}$ 上的广义算子范数（或从属于 $\|\cdot\|_a$，$\|\cdot\|_b$ 的算子范数）.

4. 酉不变范数

以下用 U_n 表示 $n\times n$ 酉阵的集合.

定义 10 设 $A\in\mathbf{C}^{m\times n}$，若映射 $\|\cdot\|:\mathbf{C}^{m\times n}\to\mathbf{R}$ 满足

(1) $A\neq 0\Rightarrow \|A\|>0$；

(2) $\|\lambda A\|=|\lambda|\,\|A\|$，$\forall\lambda\in\mathbf{C}$；

(3) $\|A+B\|\leqslant\|A\|+\|B\|$，$\forall A,B\in\mathbf{C}^{m\times n}$；

(4) $\|UAV\|=\|A\|$，$\forall U\in U_m$，$\forall V\in U_n$；

(5) $\| \boldsymbol{A} \| = \| \boldsymbol{A} \|_2$, $\text{rank}(\boldsymbol{A})=1$.

则称映射 $\| \cdot \|$ 为 $\mathbf{C}^{m\times n}$ 上的酉不变范数.

注 上面定义 10 中(5)的等价条件为

(5)′ 若 $\boldsymbol{A}=\boldsymbol{x}\boldsymbol{y}^{\mathrm{H}}$, $\boldsymbol{x}\in\mathbf{C}^m$, $\boldsymbol{y}\in\mathbf{C}^n$, $\boldsymbol{x}\neq\mathbf{0}$, $\boldsymbol{y}\neq\mathbf{0}$, 则 $\| \boldsymbol{A} \| = \| \boldsymbol{x} \|_2 \| \boldsymbol{y} \|_2$.

5. 矩阵的测度

定义 11 设 $\boldsymbol{A}\in P^{n\times n}$, $\| \cdot \|$ 是给定的算子范数,若

$$\mu(\boldsymbol{A}) = \lim_{h\to 0^+} \frac{\| \boldsymbol{E}_n + h\boldsymbol{A} \| - 1}{h}$$

存在,则称 $\mu(\boldsymbol{A})$ 为对应于 $\| \cdot \|$ 的 $\boldsymbol{A}$ 的矩阵测度.

算子范数 $\| \cdot \|_1$, $\| \cdot \|_2$ 与 $\| \cdot \|_\infty$ 对应的矩阵测度分别为

$$\mu_1(\boldsymbol{A}) = \max_j \Big\{ \mathrm{Re}\, a_{jj} + \sum_{\substack{i=1\\ i\neq j}}^n | a_{ij} | \Big\},$$

$$\mu_2(\boldsymbol{A}) = \max_i \lambda_i \left(\frac{\boldsymbol{A}^{\mathrm{H}} + \boldsymbol{A}}{2} \right),$$

$$\mu_\infty(\boldsymbol{A}) = \max_i \Big\{ \mathrm{Re}\, a_{ii} + \sum_{\substack{j=1\\ j\neq i}}^n | a_{ij} | \Big\}.$$

6. 范数的应用

定义 12 设 $\boldsymbol{A}$ 是可逆矩阵,称

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

是矩阵 $\boldsymbol{A}$ 相对矩阵范数 $\| \cdot \|_p$ 的条件数.

二、主要结论

1. 范数的等价性

(1) $V^n(P)$上的任意两个向量范数均等价. 即对于 $V^n(P)$中任意两个向量范数 $\| \boldsymbol{x} \|_a$ 与 $\| \boldsymbol{x} \|_b$,存在与 $\boldsymbol{x}$ 无关的正数 $M>m>0$,使得

$$m \| \boldsymbol{x} \|_a \leqslant \| \boldsymbol{x} \|_b \leqslant M \| \boldsymbol{x} \|_a, \quad \forall x \in V^n(P).$$

(2) $P^{m\times n}$ 上的任意两个矩阵范数均等价. 即对于 $P^{m\times n}$ 中任意两个矩阵范数 $\| \boldsymbol{A} \|_a$ 与 $\| \boldsymbol{A} \|_b$,存在与 $\boldsymbol{A}$ 无关的正数 $M>m>0$,使得

$$m \| \boldsymbol{A} \|_a \leqslant \| \boldsymbol{A} \|_b \leqslant M \| \boldsymbol{A} \|_a, \quad \forall \boldsymbol{A} \in P^{m\times n}.$$

2. 矩阵范数与向量范数的相容性

(1) 设 $\boldsymbol{x}\in P^n$, $\boldsymbol{A}\in P^{m\times n}$,则

① $\|A\|_{m_1}$ 和 $\|A\|_1$ 都与 $\|x\|_1$ 相容；

② $\|A\|_{m_2}$ 和 $\|A\|_2$ 都与 $\|x\|_2$ 相容；

③ $\|A\|_G$ 和 $\|A\|_\infty$ 都与 $\|x\|_\infty$ 相容；

④ $\|A\|_{m_\infty}$ 是不相容的矩阵范数.

(2) 对于任意给定的矩阵范数，一定存在与之相容的向量范数.

(3) 对于任意给定的向量范数，一定存在与之相容的矩阵范数.

(4) 一种矩阵范数可以与多种向量范数相容.

(5) 多种矩阵范数可以与一种向量范数相容.

(6) 并非任意的矩阵范数与任意的向量范数都相容.

3. 矩阵 Frobenius 范数与谱范数

在矩阵范数中，Frobenius 范数 $\|A\|_{m_2}$ 和谱范数 $\|A\|_2$ 用途最多，它具有下述性质.

设 $A=(\alpha_1,\alpha_2,\cdots,\alpha_n)\in P^{n\times n}$，$U,V\in P^{n\times n}$ 为任意的酉矩阵，$x,y\in P^n$，则有

(1) $\|A\|_F^2=\|A\|_{m_2}^2=\sum\limits_{i=1}^{n}\|\alpha_i\|_2^2$，其中 $\|\alpha_i\|_2^2=\alpha_i^H\alpha_i$ 是 P^n 中的向量范数；

(2) $\|A\|_{m_2}^2=\mathrm{tr}(A^HA)=\sum\limits_{i=1}^{n}\lambda_i(A^HA)$；

(3) $\|A\|_{m_2}=\|U^HAV\|_{m_2}=\|UAV^H\|_{m_2}=\|UA\|_{m_2}=\|AV\|_{m_2}=\|UAV\|_{m_2}$；

(4) $\|A\|_2=\|A^H\|_2=\|A^T\|_2=\|\bar{A}\|_2$；

(5) $\|A^HA\|_2=\|AA^H\|_2=\|A\|_2^2$；

(6) $\|UA\|_2=\|AV\|_2=\|UAV\|_2=\|A\|_2$；

(7) $\|A\|_2=\max\limits_{\|x\|_2=\|y\|_2=1}|y^HAx|$；

(8) $\|A\|_2^2\leqslant\|A\|_1\|A\|_\infty$；

(9) 设 $A\in C_r^{m\times n}$ 的正奇异值为 $\sigma_1,\sigma_2,\cdots,\sigma_r$，则

$$\|A\|_2=\max_i\sigma_i,\quad \|A\|_{m_2}=\Big(\sum_{i=1}^{r}\sigma_i^2\Big)^{1/2}.$$

4. 酉不变范数

(1) Frobenius 范数 $\|A\|_{m_2}$ 和谱范数 $\|A\|_2$ 是酉不变范数.

(2) $\|\cdot\|_1$ 与 $\|\cdot\|_\infty$ 不是酉不变范数.

(3) 设 $A,B\in C^{m\times n}$，它们的奇异值分别为 $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_n\geqslant 0$ 和 $\tau_1\geqslant\tau_2\geqslant\cdots\geqslant\tau_n\geqslant 0$. 如果 $\sigma_i\leqslant\tau_i(i=1,2,\cdots,n)$，则对于 $C^{m\times n}$ 上的任一酉不变范数 $\|\cdot\|$，必有 $\|A\|\leqslant\|B\|$.

(4) 设 $\|\cdot\|$ 为 $C^{m\times n}$ 上的酉不变范数，则有

$$\|AB\|\leqslant\|A\|_2\|B\|,\quad \forall A\in C^{m\times m},B\in C^{m\times n},$$

$$\|AB\| \leqslant \|A\| \|B\|_2, \quad \forall A \in \mathbf{C}^{m\times n}, B \in \mathbf{C}^{n\times n},$$

$$\|Ax\|_2 \leqslant \|A\| \|x\|_2, \quad \forall A \in \mathbf{C}^{m\times n}, x \in \mathbf{C}^n,$$

$$\|A\|_2 \leqslant \|A\| \leqslant r\|A\|_2, \quad \forall A \in \mathbf{C}^{m\times n}, r = \min\{m, n\}.$$

(5) 设 $\|\cdot\|$ 为 $\mathbf{C}^{n\times n}$ 上的酉不变范数，则它是相容的矩阵范数，即

$$\|AB\| \leqslant \|A\| \|B\|, \quad \forall A, B \in \mathbf{C}^{n\times n}.$$

5. 方阵的范数

设 $A \in \mathbf{C}^{n\times n}$，$r(A)$ 为矩阵 $A \in \mathbf{C}^{n\times n}$ 的谱半径.

(1) 若有矩阵范数 $\|\cdot\|$，使得 $\|A\| < 1$，则 $E - A$ 为可逆矩阵，且有

$$\|(E-A)^{-1}\| \leqslant \frac{\|E\|}{1-\|A\|}, \quad \|E-(E-A)^{-1}\| \leqslant \frac{\|A\|}{1-\|A\|};$$

(2) 若 $A^{\mathrm{H}}A = AA^{\mathrm{H}}$，则 $r(A) = \|A\|_2$；

(3) 对任意给定的相容矩阵范数 $\|\cdot\|$，有 $r(A) \leqslant \|A\|$；

(4) 对任意给定的矩阵 A 和正数 ε，存在矩阵范数 $\|\cdot\|$，使得 $\|A\|_M \leqslant r(A) + \varepsilon$.

6. 矩阵测度的性质

设 $A \in P^{n\times n}$，$\|\cdot\|$ 为算子范数，则

(1) $\mu(E_n) = 1, \mu(-E_n) = -1, \mu(\mathbf{0}) = 0$；

(2) $\mu(aA) = a\mu(A), 0 \leqslant a \in P$；

(3) $\mu(A + aE_n) = \mu(A) + a, a \in P$；

(4) $\mu(A+B) \leqslant \mu(A) + \mu(B), \forall A, B \in P^{n\times n}$；

(5) $-\|A\| \leqslant -\mu(-A) \leqslant \mu(A) \leqslant \|A\|$；

(6) $-\mu(-A) \leqslant \mathrm{Re}(\lambda_i(A)) \leqslant \mu(A) \leqslant \|A\|$；

(7) $\|Ax\| \geqslant \max\{-\mu(-A), -\mu(A)\}\|x\|, \forall x \in P^n$.

7. 矩阵逆及线性方程组的摄动

(1) 设 A 可逆，δA 为摄动矩阵，且 $\|A^{-1}\delta A\|_a < 1$，则

① $A + \delta A$ 为可逆矩阵；

② $(A+\delta A)^{-1} = (E+F)A^{-1}$，其中 $\|F\|_a \leqslant \dfrac{\|A^{-1}\delta A\|_a}{1-\|A^{-1}\delta A\|_a}$；

③ $\dfrac{\|A^{-1}-(A+\delta A)^{-1}\|_a}{\|A^{-1}\|_a} \leqslant \dfrac{\|A^{-1}\delta A\|_a}{1-\|A^{-1}\delta A\|_a}$.

(2) 设 $K(A) = \|A\|_a \|A^{-1}\|_a$ 是条件数，若 $\|A^{-1}\|_a \|\delta A\|_a < 1$，则

$$\|F\|_a \leqslant \frac{K(A)\dfrac{\|\delta A\|_a}{\|A\|_a}}{1-K(A)\dfrac{\|\delta A\|_a}{\|A\|_a}},$$

$$\frac{\| \boldsymbol{A}^{-1} - (\boldsymbol{A} + \delta\boldsymbol{A})^{-1} \|_a}{\| \boldsymbol{A}^{-1} \|_a} \leqslant \frac{K(\boldsymbol{A}) \dfrac{\| \delta\boldsymbol{A} \|_a}{\| \boldsymbol{A} \|_a}}{1 - K(\boldsymbol{A}) \dfrac{\| \delta\boldsymbol{A} \|_a}{\| \boldsymbol{A} \|_a}}.$$

(3) 若 $\boldsymbol{A}$ 是酉矩阵，且 $\| \delta\boldsymbol{A} \|_2 < 1$，则有

$$\| \boldsymbol{F} \|_a \leqslant \frac{\| \delta\boldsymbol{A} \|_a}{1 - \| \delta\boldsymbol{A} \|_a}, \quad \frac{\| \boldsymbol{A}^{-1} - (\boldsymbol{A} + \delta\boldsymbol{A})^{-1} \|_a}{\| \boldsymbol{A}^{-1} \|_a} \leqslant \frac{\| \delta\boldsymbol{A} \|_a}{1 - \| \delta\boldsymbol{A} \|_a}.$$

(4) 在方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 中，$\boldsymbol{A}$ 固定且可逆，令 $\boldsymbol{b} \neq \boldsymbol{0}$ 且有小摄动 $\delta\boldsymbol{b}$，则解方程组

$$\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b},$$

得

$$\frac{\| \delta\boldsymbol{x} \|_a}{\| \boldsymbol{x} \|_a} \leqslant K(\boldsymbol{A}) \frac{\| \delta\boldsymbol{b} \|_a}{\| \boldsymbol{b} \|_a}.$$

(5) 设 $\boldsymbol{A}$ 可逆，$\boldsymbol{0} \neq \boldsymbol{b} \in \mathbf{C}^n$，$\| \boldsymbol{A}^{-1} \|_a \| \delta\boldsymbol{A} \|_a < 1$，方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的解是 $\boldsymbol{x}$，则方程组 $(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$ 有唯一解 $\boldsymbol{x} + \delta\boldsymbol{x}$，并且满足不等式

$$\frac{\| \delta\boldsymbol{x} \|_a}{\| \boldsymbol{x} \|_a} \leqslant \frac{K(\boldsymbol{A})}{r(\boldsymbol{A})} \left(\frac{\| \delta\boldsymbol{A} \|_a}{\| \boldsymbol{A} \|_a} + \frac{\| \delta\boldsymbol{b} \|_a}{\| \boldsymbol{b} \|_a} \right),$$

其中 $r(\boldsymbol{A}) = 1 - K(\boldsymbol{A}) \dfrac{\| \delta\boldsymbol{A} \|_a}{\| \boldsymbol{A} \|_a} > 0$.

三、典型例题

例 1 设 $\boldsymbol{A} = (a_{ij})_{m \times n}$，定义实数 $\| \boldsymbol{A} \|_G = \sqrt{mn} \max\limits_{i,j} |a_{ij}|$，$1 \leqslant i \leqslant m$，$1 \leqslant j \leqslant n$. 证明：$\| \boldsymbol{A} \|_G$ 是 $\boldsymbol{C}^{m \times n}$ 上的矩阵范数，且与向量的 2-范数相容.

证 首先证明 $\| \boldsymbol{A} \|_G$ 是矩阵范数. (1) $\boldsymbol{A} = \boldsymbol{0}$ 时，$a_{ij} = 0$，$\max\limits_{i,j} |a_{ij}| = 0$，从而 $\| \boldsymbol{A} \|_G = 0$；当 $\boldsymbol{A} \neq \boldsymbol{0}$ 时，存在 $a_{i_0 j_0} \neq 0$，$\| \boldsymbol{A} \|_G = \sqrt{mn} \max\limits_{i,j} |a_{ij}| \geqslant \sqrt{mn} |a_{i_0 j_0}| > 0$. 齐次性和三角不等式容易验证成立.

下面证明相容性. 论证乘法相容性. 设 $\boldsymbol{B} = (b_{ij})_{n \times p}$，则有

$$\begin{aligned}
\| \boldsymbol{A}\boldsymbol{B} \|_G &= \sqrt{mp} \max_{i,j} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right| \leqslant \sqrt{mp} \max_{i,j} \left(\sum_{k=1}^{n} |a_{ik}| |b_{kj}| \right) \\
&\leqslant \sqrt{mp} \cdot n \cdot \max_{i,k} |a_{ik}| \max_{k,j} |b_{kj}| \\
&= \sqrt{mn} \max_{i,j} |a_{ij}| \sqrt{np} \max_{i,j} |b_{ij}| = \| \boldsymbol{A} \|_G \| \boldsymbol{B} \|_G.
\end{aligned}$$

因此，$\| \boldsymbol{A} \|_G$ 是矩阵范数.

设 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，则有

$$\| \boldsymbol{A}\boldsymbol{x} \|_2^2 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right|^2 \leqslant \sum_{i=1}^{m} \left(\sum_{j=1}^{n} |a_{ij}| |x_j| \right)^2$$

$$\leqslant \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2\right)\left(\sum_{j=1}^{n} |x_j|^2\right)$$
$$\leqslant \|\boldsymbol{A}\|_G^2 \|\boldsymbol{x}\|_2^2,$$

即 $\|\boldsymbol{Ax}\|_2 \leqslant \|\boldsymbol{A}\|_G \|\boldsymbol{x}\|_2$.

例 2 设有可逆矩阵 $\boldsymbol{S}\in\mathbf{R}^{n\times n}$, $\boldsymbol{x}\in\mathbf{R}^n$, 且 $\|\boldsymbol{x}\|_S = \|\boldsymbol{Sx}\|_2$ 是 $\mathbf{R}^n$ 上的向量范数.

(1) 若 $\|\boldsymbol{A}\|_S$ 表示 $\mathbf{R}^{n\times n}$ 上从属于向量范数 $\|\boldsymbol{x}\|_S$ 的算子范数,试导出 $\|\boldsymbol{A}\|_S$ 与矩阵的 2-范数之间的关系.

(2) 给定非零列向量 $\boldsymbol{y}\in\mathbf{R}^n$, 证明 $\|\boldsymbol{x}\| = \|\boldsymbol{xy}^{\mathrm{T}}\|_S$ 是 $\mathbf{R}^n$ 上的向量范数.

证 (1) 根据算子范数的定义,有

$$\|\boldsymbol{A}\|_S = \max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{Ax}\|_S}{\|\boldsymbol{x}\|_S} = \max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{SAx}\|_2}{\|\boldsymbol{Sx}\|_2}.$$

令 $\boldsymbol{y}=\boldsymbol{Sx}$, 因为 $\boldsymbol{S}$ 可逆, 所以当 $\boldsymbol{x}\neq\boldsymbol{0}$ 时, $\boldsymbol{y}\neq\boldsymbol{0}$, 于是有

$$\|\boldsymbol{A}\|_S = \max_{\boldsymbol{y}\neq\boldsymbol{0}} \frac{\|\boldsymbol{SAS}^{-1}\boldsymbol{y}\|_2}{\|\boldsymbol{y}\|_2} = \|\boldsymbol{SAS}^{-1}\|_2.$$

(2) 非负性 $\boldsymbol{x}=\boldsymbol{0}$ 时, $\|\boldsymbol{x}\| = \|\boldsymbol{xy}^{\mathrm{T}}\|_S = \|\boldsymbol{0}\|_S = 0$; $\boldsymbol{x}\neq\boldsymbol{0}$ 时, $\boldsymbol{xy}^{\mathrm{T}}\neq\boldsymbol{0}$, $\|\boldsymbol{x}\| = \|\boldsymbol{xy}^{\mathrm{T}}\|_S > 0$.

齐次性 对任意 $c\in\mathbf{R}$, 有 $\|c\boldsymbol{x}\| = \|(c\boldsymbol{x})\boldsymbol{y}^{\mathrm{T}}\|_S = |c|\,\|\boldsymbol{xy}^{\mathrm{T}}\|_S = |c|\,\|\boldsymbol{x}\|$.

三角不等式 对任意 $\boldsymbol{x}_1, \boldsymbol{x}_2\in\mathbf{R}^n$, 有

$$\begin{aligned}\|\boldsymbol{x}_1+\boldsymbol{x}_2\| &= \|(\boldsymbol{x}_1+\boldsymbol{x}_2)\boldsymbol{y}^{\mathrm{T}}\|_S = \|\boldsymbol{x}_1\boldsymbol{y}^{\mathrm{T}}+\boldsymbol{x}_2\boldsymbol{y}^{\mathrm{T}}\|_S \\ &\leqslant \|\boldsymbol{x}_1\boldsymbol{y}^{\mathrm{T}}\|_S + \|\boldsymbol{x}_2\boldsymbol{y}^{\mathrm{T}}\|_S \\ &= \|\boldsymbol{x}_1\| + \|\boldsymbol{x}_2\|.\end{aligned}$$

例 3 设 $\|\cdot\|$ 是 $\mathbf{C}^{m\times n}$ 上的矩阵算子范数, 若 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$ 满足 $\|\boldsymbol{A}\|<1$, 证明 $\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\| \leqslant \dfrac{1}{1-\|\boldsymbol{A}\|}$.

证 因为 $r(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|<1$, 故 $(\boldsymbol{E}-\boldsymbol{A})^{-1}$ 存在. 又因为

$$(\boldsymbol{E}-\boldsymbol{A})^{-1}(\boldsymbol{E}-\boldsymbol{A}) = \boldsymbol{E}, (\boldsymbol{E}-\boldsymbol{A})^{-1} = \boldsymbol{E}+(\boldsymbol{E}-\boldsymbol{A})^{-1}\boldsymbol{A},$$

所以

$$\begin{aligned}&\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\| \leqslant \|\boldsymbol{E}\| + \|(\boldsymbol{E}-\boldsymbol{A})^{-1}\boldsymbol{A}\| \cdot \|\boldsymbol{A}\| \\ &= 1+\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\boldsymbol{A}\| \cdot \|\boldsymbol{A}\|,\end{aligned}$$

即 $\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\| \leqslant \dfrac{1}{1-\|\boldsymbol{A}\|}$.

例 4 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 可逆, $\boldsymbol{B}\in\mathbf{C}^{n\times n}$, 若对某种矩阵范数 $\|\cdot\|$, 有 $\|\boldsymbol{B}\| < \|\boldsymbol{A}^{-1}\|^{-1}$, 则 $\boldsymbol{A}+\boldsymbol{B}$ 可逆.

证 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{A}^{-1}\boldsymbol{B})$. 因为 $\|\boldsymbol{B}\| < \|\boldsymbol{A}^{-1}\|^{-1}$, 所以 $r(\boldsymbol{A}^{-1}\boldsymbol{B}) < \|\boldsymbol{A}^{-1}\boldsymbol{B}\| \leqslant \|\boldsymbol{A}^{-1}\|\,\|\boldsymbol{B}\| < 1$, 即 $(\boldsymbol{E}+\boldsymbol{A}^{-1}\boldsymbol{B})$ 的特征值非零. 又因为 $\boldsymbol{A}$ 可逆, 所以 $\boldsymbol{A}+\boldsymbol{B}$ 可逆.

例 5 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$, 则 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$ 当且仅当 $r(\boldsymbol{A})<1$.

证 若 $\boldsymbol{A}^k\to\boldsymbol{0}$,设 λ 为 $\boldsymbol{A}$ 的任意特征值,$\boldsymbol{x}\neq\boldsymbol{0}$ 是满足 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ 的特征向量,仅当 $|\lambda|<1$ 时才有 $\boldsymbol{A}^k\boldsymbol{x}=\lambda^k\boldsymbol{x}\to\boldsymbol{0}$. 由 λ 的任意性,有 $r(\boldsymbol{A})<1$. 反过来,如果 $r(\boldsymbol{A})<1$,则存在某个矩阵范数 $\|\cdot\|$,使得 $\|\boldsymbol{A}\|<1$. 又由于当 $k\to\infty$ 时,有 $\|\boldsymbol{A}^k\|\leqslant\|\boldsymbol{A}\|^k\to 0$,因此 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$.

例 6 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,如果存在矩阵范数 $\|\cdot\|$,使得 $\|\boldsymbol{E}-\boldsymbol{A}\|<1$,则 $\boldsymbol{A}$ 是可逆矩阵,且

$$\boldsymbol{A}^{-1}=\sum_{k=0}^{\infty}(\boldsymbol{E}-\boldsymbol{A})^k.$$

证 设 λ 为矩阵 $\boldsymbol{A}$ 的任意特征值. 如果 $\|\boldsymbol{E}-\boldsymbol{A}\|<1$,也就是 $\boldsymbol{E}-\boldsymbol{A}$ 矩阵的特征值满足 $-1<1-\lambda<1$,即 $0<\lambda<2$. 由于 λ 的任意性,则 $\boldsymbol{A}$ 是可逆矩阵.

又因为级数 $\sum\limits_{k=0}^{\infty}z^k$ 的收敛半径为 1,所以

$$\sum_{k=0}^{\infty}(\boldsymbol{E}-\boldsymbol{A})^k$$

收敛于某个矩阵 $\boldsymbol{C}$. 又因为

$$\begin{aligned}\boldsymbol{A}\sum_{k=0}^{N}(\boldsymbol{E}-\boldsymbol{A})^k&=[\boldsymbol{E}-(\boldsymbol{E}-\boldsymbol{A})]\sum_{k=0}^{N}(\boldsymbol{E}-\boldsymbol{A})^k\\&=\boldsymbol{E}-(\boldsymbol{E}-\boldsymbol{A})^{N+1},\end{aligned}$$

当 $N\to\infty$ 时,由上面例 5 的证明知

$$\boldsymbol{A}\sum_{k=0}^{N}(\boldsymbol{E}-\boldsymbol{A})^k=\boldsymbol{E}-(\boldsymbol{E}-\boldsymbol{A})^{N+1}\to\boldsymbol{E},$$

所以 $\boldsymbol{C}=\boldsymbol{A}^{-1}$.

四、习题解答

1. 设 $a_1,a_2,\cdots,a_n$ 均为正数,$\boldsymbol{x}\in\mathbf{C}^n$,且 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$. 证明函数

$$f(\boldsymbol{x})=\left(\sum_{i=1}^{n}a_i\mid x_i\mid^2\right)^{1/2}$$

在 $\mathbf{C}^n$ 上定义了一个向量范数.

证 (1) 正定性对 $\forall\boldsymbol{x}\neq\boldsymbol{0}$,有 $f(\boldsymbol{x})>0$,当 $\boldsymbol{x}=\boldsymbol{0}$ 时,$f(\boldsymbol{x})=0$.

(2) 齐次性 $f(\lambda\boldsymbol{x})=\left(\sum\limits_{i=1}^{n}a_i\mid\lambda x_i\mid^2\right)^{1/2}=|\lambda|\cdot\left(\sum\limits_{i=1}^{n}a_i\mid x_i\mid^2\right)^{1/2}=|\lambda|\cdot f(\boldsymbol{x})$.

(3) 三角不等式

$$\begin{aligned}f(\boldsymbol{x}+\boldsymbol{y})^2&=\left(\sum_{i=1}^{n}a_i\mid x_i+y_i\mid^2\right)\\&=\left(\sum_{i=1}^{n}a_i(\mid x_i\mid^2+\mid y_i\mid^2+\bar{x}_iy_i+x_i\bar{y}_i)\right)\end{aligned}$$

$$\leqslant f(\boldsymbol{x})^2+f(\boldsymbol{y})^2+\sum_{i=1}^{n}(2a_i\mid x_i\mid\cdot\mid y_i\mid)$$

$$\leqslant f(\boldsymbol{x})^2+f(\boldsymbol{y})^2+\sum_{i=1}^{n}(2a_i\mid x_i\mid\cdot\mid y_i\mid)$$

$$\leqslant f(\boldsymbol{x})^2+f(\boldsymbol{y})^2+\sum_{i=1}^{n}(2\mid\sqrt{a_i}x_i\mid\cdot\mid\sqrt{a_i}y_i\mid)$$

$$\leqslant f(\boldsymbol{x})^2+f(\boldsymbol{y})^2+2\Big(\sum_{i=1}^{n}a_i\mid x_i\mid^2\Big)^{1/2}\cdot\Big(\sum_{i=1}^{n}a_i\mid y_i\mid^2\Big)^{1/2}$$

$$=f(\boldsymbol{x})^2+f(\boldsymbol{y})^2+2f(\boldsymbol{x})\cdot f(\boldsymbol{y})$$

$$=(f(\boldsymbol{x})+f(\boldsymbol{y}))^2.$$

所以函数 $f(\boldsymbol{x})$是一个向量范数.

2. 证明：对于 $\mathbf{R}^1$ 中的任何向量范数 $\|x\|$，一定有 $\|x\|=\lambda|x|$，其中 $\lambda>0$.

证 对任意向量范数 $\|x\|$，根据向量范数的定义和性质，且 $x\in\mathbf{R}^1$，有

$$\|x\|=\|x\cdot 1\|=|x|\cdot\|1\|=\lambda\mid x\mid,\quad 其中\ \lambda=\|1\|>0.$$

3. 设 $\|\boldsymbol{x}\|$ 是 P^n 中的向量范数，$\boldsymbol{A}\in P^{n\times n}$，则 $\|\boldsymbol{Ax}\|$ 也是 P^n 中的向量范数的充要条件为 $\boldsymbol{A}$ 是可逆矩阵.

证 必要性. 如果矩阵 $\boldsymbol{A}$ 不可逆，则存在 $\boldsymbol{x}\neq\boldsymbol{0}$，使得 $\boldsymbol{Ax}=\boldsymbol{0}$，即 $\|\boldsymbol{Ax}\|=\boldsymbol{0}$，这与向量范数的正定性矛盾，所以矩阵 $\boldsymbol{A}$ 可逆.

充分性. 若矩阵 $\boldsymbol{A}$ 可逆，则对任意的 $\boldsymbol{x}\neq\boldsymbol{0}$，$\boldsymbol{Ax}\neq\boldsymbol{0}$，所以 $\|\boldsymbol{Ax}\|>0$，正定性满足；又 $\|\lambda\boldsymbol{Ax}\|=|\lambda|\cdot\|\boldsymbol{Ax}\|$，齐次性满足；再由 $\|\boldsymbol{A}(\boldsymbol{x}+\boldsymbol{y})\|=\|\boldsymbol{Ax}+\boldsymbol{Ay}\|\leqslant\|\boldsymbol{Ax}\|+\|\boldsymbol{Ay}\|$，三角不等式也满足. 故 $\|\boldsymbol{Ax}\|$ 是向量范数.

4. 证明：

(1) $\|\boldsymbol{A}\|_{m_2}=[\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})]^{1/2}$；

(2) $\|\boldsymbol{A}\|_{m_2}$ 与 $\|\boldsymbol{x}\|_2$ 是相容的；

(3) $\|\boldsymbol{A}\|_a$ 与 $\|\boldsymbol{x}\|_1$，$\|\boldsymbol{x}\|_2$ 均相容；

(4) $\|\boldsymbol{AB}\|_{m_2}\leqslant\min\{\|\boldsymbol{A}\|_2\|\boldsymbol{B}\|_{m_2},\|\boldsymbol{A}\|_{m_2}\|\boldsymbol{B}\|_2\}$.

证 设 $\boldsymbol{A}=(a_{ij})\in P^{n\times n}$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in P^n$.

(1) 令 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$. 根据定义有 $\|\boldsymbol{A}\|_{m_2}^2=\sum_{i=1}^{n}\sum_{j=1}^{n}\mid a_{ij}\mid^2$，$\|\boldsymbol{\alpha}_j\|_2^2=\sum_{i=1}^{n}\mid a_{ij}\mid^2$，$j=1,2,\cdots,n$，所以有 $\sum_{j=1}^{n}\|\boldsymbol{\alpha}_j\|^2=\|\boldsymbol{A}\|_{m_2}^2$，同时有 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{H}}\\ \boldsymbol{\alpha}_2^{\mathrm{H}}\\ \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{H}}\end{pmatrix}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=$

$$\begin{pmatrix} \boldsymbol{\alpha}_1^{\mathrm{H}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_1^{\mathrm{H}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{H}}\boldsymbol{\alpha}_n \\ \boldsymbol{\alpha}_2^{\mathrm{H}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2^{\mathrm{H}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_2^{\mathrm{H}}\boldsymbol{\alpha}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\alpha}_n^{\mathrm{H}}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_n^{\mathrm{H}}\boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n^{\mathrm{H}}\boldsymbol{\alpha}_n \end{pmatrix}$$，所以有 $\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \sum_{j=1}^{n} \boldsymbol{\alpha}_j^{\mathrm{H}}\boldsymbol{\alpha}_j = \|\boldsymbol{A}\|_{m_2}^2$.

(2) 令 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_1\\ \boldsymbol{A}_2\\ \vdots\\ \boldsymbol{A}_n\end{pmatrix}$，其中 $\boldsymbol{A}_i$ 为 $\boldsymbol{A}$ 的行向量，则

$$\begin{aligned}\|\boldsymbol{Ax}\|_2^2 &= \sum_{i=1}^{n} |\boldsymbol{A}_i\boldsymbol{x}|^2 = \sum_{i=1}^{n} |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n|^2 \\ &\leqslant \sum_{i=1}^{n}\left(\sum_{j=1}^{n}|a_{ij}|^2 \sum_{j=1}^{n}|x_j|^2\right) \\ &= \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)\sum_{j=1}^{n}|x_j|^2 \\ &= \|\boldsymbol{A}\|_{m_2}^2 \|\boldsymbol{x}\|_2^2\end{aligned}$$

故有

$$\|\boldsymbol{Ax}\|_2 \leqslant \|\boldsymbol{A}\|_{m_2}\|\boldsymbol{x}\|_2.$$

(3) 因为

$$\begin{aligned}\|\boldsymbol{Ax}\|_1 &= \sum_{i=1}^{n}\left|\sum_{j=1}^{n}a_{ij}x_j\right| \leqslant \sum_{i=1}^{n}\sum_{i=1}^{n}|a_{ij}x_j| \\ &\leqslant \sum_{i=1}^{n}\max_{1\leqslant j\leqslant n}\{|a_{ij}|\}\sum_{j=1}^{n}|x_j| \\ &\leqslant n\cdot\max_{1\leqslant i,j\leqslant n}\{|a_{ij}|\}\sum_{j=1}^{n}|x_j| \\ &= \|\boldsymbol{A}\|_a\cdot\|\boldsymbol{x}\|_1;\end{aligned}$$

$$\begin{aligned}\|\boldsymbol{Ax}\|_2^2 &= \sum_{i=1}^{n}|a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n|^2 \\ &\leqslant \sum_{i=1}^{n}\left(\sum_{j=1}^{n}|a_{ij}|^2\sum_{j=1}^{n}|x_j|^2\right) \\ &= \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)\cdot\sum_{j=1}^{n}|x_j|^2 \\ &\leqslant n^2\max_{1\leqslant i,j\leqslant n}\{|a_{ij}|^2\}\cdot\sum_{j=1}^{n}|x_j|^2 \\ &= \|\boldsymbol{A}\|_a^2\cdot\|\boldsymbol{x}\|_2^2.\end{aligned}$$

所以 $\|A\|_a$ 与 $\|x\|_1$，$\|x\|_2$ 均相容.

(4) 令 $B=(\beta_1,\cdots,\beta_j,\cdots,\beta_n)$，因为 $\|A\beta_j\|_2\leqslant\|A\|_2\|\beta_j\|_2, j=1,2,\cdots,n$，同时有

$$\begin{aligned}\|AB\|_{m_2}^2 &= \|A(\beta_1,\beta_2,\cdots,\beta_n)\|_{m_2}^2\\ &= \|A\beta_1\|_2^2+\|A\beta_2\|_2^2+\cdots+\|A\beta_n\|_2^2\\ &\leqslant \|A\|_2^2(\|\beta_1\|_2^2+\|\beta_2\|_2^2+\cdots+\|\beta_n\|_2^2)\\ &= \|A\|_2^2\|B\|_{m_2}^2\end{aligned}$$

由上述结果有

$$\begin{aligned}\|AB\|_{m_2} = \|(AB)^{\mathrm{H}}\|_{m_2} &= \|B^{\mathrm{H}}A^{\mathrm{H}}\|_{m_2}\\ &\leqslant \|B^{\mathrm{H}}\|_2\|A^{\mathrm{H}}\|_{m_2}\\ &= \|B\|_2\|A\|_{m_2},\end{aligned}$$

所以(4)成立.

5. 若 $A\in P^{m\times r}$，且 $A^{\mathrm{H}}A=E_r$，则

$$\|A\|_2=1,\quad \|A\|_{m_2}=\sqrt{r}.$$

证 根据定义有 $\|A\|_2=\sqrt{r(A^{\mathrm{H}}A)}=\sqrt{r(E)}=1$，$\|A\|_{m_2}=\sqrt{\mathrm{tr}(A^{\mathrm{H}}A)}=\sqrt{\mathrm{tr}(E)}=\sqrt{r}$.

6. 设 x, Ax 的向量范数为 $\|\cdot\|_2$，证明：它对应的算子范数是

$$\|A\|_2=\max_{\|x\|_2=1}\|Ax\|_2=\max\{\sigma_1,\sigma_2,\cdots,\sigma_n\}.$$

证 对任意矩阵 A，存在酉矩阵 U, V，使得矩阵 A 的奇异值分解为 $A=UDV$. 其中 σ_1，$\sigma_2,\cdots,\sigma_n$ 是矩阵 A 的奇异值，$D=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_n)$. 根据定义有 $\|A\|_2=\sqrt{r(A^{\mathrm{H}}A)}=\sqrt{r((UDV)^{\mathrm{H}}UDV)}=\sqrt{r(V^{\mathrm{H}}D^2V)}=\sqrt{r(D^2)}=\max\{\sigma_1,\sigma_2,\cdots,\sigma_n\}$.

7. 若 $\|\cdot\|$ 是算子范数，则

(1) $\|E\|=1$；

(2) $\|A^{-1}\|\geqslant\|A\|^{-1}$；

(3) $\|A^{-1}\|^{-1}=\min\limits_{x\neq 0}\dfrac{\|Ax\|}{\|x\|}$.

证 根据算子范数定义，有 $\|A\|=\max\limits_{x\neq 0}\dfrac{\|Ax\|}{\|x\|}$，于是可得

(1) $\|E\|=\max\limits_{x\neq 0}\dfrac{\|Ex\|}{\|x\|}=\max\limits_{x\neq 0}\dfrac{\|x\|}{\|x\|}=1$；

(2) $1=\|E\|=\|AA^{-1}\|\leqslant\|A\|\,\|A^{-1}\|$，$\|A^{-1}\|\geqslant\|A\|^{-1}$；

(3) $\|A^{-1}\|=\max\limits_{x\neq 0}\dfrac{\|A^{-1}x\|}{\|x\|}$，令 $y=A^{-1}x$，则 $x=Ay$，得 $\|A^{-1}\|=\max\limits_{y\neq 0}\dfrac{\|y\|}{\|Ay\|}$，从而

$$\|A^{-1}\|^{-1}=\frac{1}{\max\limits_{y\neq 0}\dfrac{\|y\|}{\|Ay\|}}=\min_{y\neq 0}\frac{\|Ay\|}{\|y\|}=\min_{x\neq 0}\frac{\|Ax\|}{\|x\|}.$$

8. 设 $\|A\|_v$，$\|A\|_\mu$ 是对应于两个向量范数 $\|x\|_v$，$\|x\|_\mu=\|Bx\|_v$ 的算子范数，B 可

逆,则

$$\|A\|_\mu = \|BAB^{-1}\|_\nu.$$

证 根据定义,有 $\|A\|_\mu = \max\limits_{x\neq 0}\dfrac{\|Ax\|_\mu}{\|x\|_\mu}$,把 $\|x\|_\mu = \|Bx\|_\nu$代入,得到 $\|A\|_\mu = \max\limits_{x\neq 0}\dfrac{\|BAx\|_\nu}{\|Bx\|_\nu}$,令 $y=Bx$,则 $x=B^{-1}y$,故 $\|A\|_\mu = \max\limits_{y\neq 0}\dfrac{\|BAB^{-1}y\|_\nu}{\|y\|_\nu} = \|BAB^{-1}\|_\nu$.

9. 设 $\|x\|_a$,$\|x\|_b$是 $\mathbf{C}^n$ 上的两个向量范数,a_1,a_2 是两个正实数,证明:

(1) $\max\{\|x\|_a, \|x\|_b\} = \|x\|_c$;

(2) $a_1\|x\|_a + a_2\|x\|_b = \|x\|_d$

都是 $\mathbf{C}^n$ 上的向量范数.

证 只需要证明(1)和(2)满足范数定义中的三个条件即可.

(1) 先考察正定性. 当 $x\neq 0$ 时,$\|x\|_a>0$,$\|x\|_b>0$,则 $\|x\|_c>0$;当 $x=0$ 时,$\|x\|_a=0$,$\|x\|_b=0$,则 $\|x\|_c=0$.

齐次性显然成立.

计算三角不等式,得

$$\begin{aligned}\|x+y\|_c &= \max\{\|x+y\|_a, \|x+y\|_b\}\\ &\leqslant \max\{\|x\|_a+\|y\|_a, \|x\|_b+\|y\|_b\}\\ &\leqslant \max\{\|x\|_a, \|x\|_b\}+\max\{\|y\|_a, \|y\|_b\}\\ &= \|x\|_c+\|y\|_c.\end{aligned}$$

(2) 正定性和齐次性同(1),下面证明三角不等式.

$$\begin{aligned}\|x+y\|_d &= a_1\|x+y\|_a + a_2\|x+y\|_b\\ &\leqslant a_1(\|x\|_a+\|y\|_a)+a_2(\|x\|_b+\|y\|_b)\\ &= \|x\|_d+\|y\|_d.\end{aligned}$$

10. 证明:

$$\frac{1}{\sqrt{n}}\|A\|_F \leqslant \|A\|_2 \leqslant \|A\|_F.$$

证 因为

$$\|A\|_2^2 = r(A^HA) \leqslant \lambda_1+\lambda_2+\cdots+\lambda_n = \operatorname{tr}(A^HA) = (\|A\|_F)^2,$$

即 $\|A\|_2\leqslant\|A\|_F$,其中 λ_i 为半正定矩阵 A^HA 的特征值. 又由于

$$(\|A\|_F)^2 = \lambda_1+\lambda_2+\cdots+\lambda_n \leqslant n\cdot r(A^HA) = n\cdot\|A\|_2^2,$$

即$\dfrac{1}{\sqrt{n}}\|A\|_F\leqslant\|A\|_2$. 证毕.

11. 设 $\|A\|_a$是 $\mathbf{C}^{n\times n}$ 上的相容矩阵范数,B,C 都是 n 阶可逆矩阵,且 $\|B^{-1}\|_a$及 $\|C^{-1}\|_a$都小于或等于 1,证明对任何 $A\in\mathbf{C}^{n\times n}$,

$$\|A\|_b = \|BAC\|_a$$

定义了 $\mathbf{C}^{n\times n}$ 上的一个相容矩阵范数.

证 首先证明 $\|\boldsymbol{A}\|_b=\|\boldsymbol{BAC}\|_a$ 是一个矩阵范数.

正定性 对任意 $\boldsymbol{A}\neq\boldsymbol{0}$,则 $\boldsymbol{BAC}\neq\boldsymbol{0}$,即 $\|\boldsymbol{BAC}\|_a>0$,且 $\|\boldsymbol{BAC}\|_a=0$ 当且仅当 $\boldsymbol{A}=\boldsymbol{0}$.

齐次性 $\|\lambda\boldsymbol{A}\|_b=\|\boldsymbol{B}(\lambda\boldsymbol{A})\boldsymbol{C}\|_a=\lambda\|\boldsymbol{BAC}\|_a=\lambda\|\boldsymbol{A}\|_b$.

三角不等式 $\|\boldsymbol{A}_1+\boldsymbol{A}_2\|_b=\|\boldsymbol{B}(\boldsymbol{A}_1+\boldsymbol{A}_2)\boldsymbol{C}\|_a\leqslant\|\boldsymbol{BA}_1\boldsymbol{C}\|_a+\|\boldsymbol{BA}_2\boldsymbol{C}\|_a=\|\boldsymbol{A}_1\|_b+\|\boldsymbol{A}_2\|_b$.

下面证明相容性.

$$\begin{aligned}\|\boldsymbol{A}_1\boldsymbol{A}_2\|_b=\|\boldsymbol{B}(\boldsymbol{A}_1\boldsymbol{A}_2)\boldsymbol{C}\|_a&=\|(\boldsymbol{BA}_1\boldsymbol{C})\boldsymbol{C}^{-1}\boldsymbol{B}^{-1}(\boldsymbol{BA}_2\boldsymbol{C})\|_a\\&\leqslant\|\boldsymbol{BA}_1\boldsymbol{C}\|_a\|\boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\|_a\|\boldsymbol{BA}_2\boldsymbol{C}\|_a\\&\leqslant\|\boldsymbol{BA}_1\boldsymbol{C}\|_a\|\boldsymbol{C}^{-1}\|_a\|\boldsymbol{B}^{-1}\|_a\|\boldsymbol{BA}_2\boldsymbol{C}\|_a\\&\leqslant\|\boldsymbol{BA}_1\boldsymbol{C}\|_a\|\boldsymbol{BA}_2\boldsymbol{C}\|_a\\&=\|\boldsymbol{A}_1\|_b\|\boldsymbol{A}_2\|_b.\end{aligned}$$

证毕.

12. 设 $\|\boldsymbol{A}\|_a$ 是 $\mathbf{C}^{n\times n}$ 上的矩阵范数,$\boldsymbol{D}$ 是 n 阶可逆矩阵,证明对任何 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,表达式

$$\|\boldsymbol{A}\|_b=\|\boldsymbol{D}^{-1}\boldsymbol{AD}\|_a$$

是 $\mathbf{C}^{n\times n}$ 上的一个矩阵范数.

证 证法同 11 题的前半部分.

第三章 矩阵的分解

矩阵分解就是通过线性变换，将某个给定矩阵分解为两个或三个结构比较简单或性质比较熟悉的矩阵的乘积(个别情况下分解为两个矩阵的和). 这一方法在线性方程组求解、模型参数估计等许多典型问题中经常使用，所以矩阵分解对矩阵理论及近代计算数学以及工程技术问题的发展起了重要作用.

譬如，矩阵的奇异值分解在图像压缩中有着很重要的应用. 现假定一幅图像有 $n\times n$ 个像素，如果将这 n^2 个数据一起传送，往往显得数据量太大. 因此，能否利用一种技术，只传送比较少的数据，在接收端还能够利用这些数据重构原图像呢?

假设 n 阶方阵 $\boldsymbol{A}$ 表示要传送的原 $n\times n$ 个像素，由矩阵 $\boldsymbol{A}$ 的奇异值分解可得 $\boldsymbol{A}=\boldsymbol{U\Sigma V}^{\mathrm{T}}$，其中，奇异值按照由大到小的顺序排列. 如果从中选择 k 个大奇异值以及与这些奇异值对应的左右奇异向量逼近原图像，就可以使用 $k(2n+1)$ 个数值代替原来的 $n\times n$ 个图像数据. 这 $k(2n+1)$ 个选出的新数据是矩阵 $\boldsymbol{A}$ 的前 k 个奇异值，$n\times n$ 左奇异向量矩阵 $\boldsymbol{U}$ 的前 k 列元素和 $n\times n$ 右奇异向量矩阵 $\boldsymbol{V}$ 的前 k 列元素.

比率 $\rho=\dfrac{n^2}{k(2n+1)}$ 称为图像的压缩比. 显然，选出的大奇异值的个数 k 应该满足 $k(2n+1)<n^2$，即 $k<\dfrac{n^2}{2n+1}$. 因此，在传送图像的过程中，无需传送 $n\times n$ 个原始数据，只需要传送 $k(2n+1)$ 个相关奇异值和奇异向量的数据即可. 在接收端，在接收到奇异值 $\sigma_1,\sigma_2,\cdots,\sigma_k$ 以及左奇异向量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_k$ 和右奇异向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k$ 后，即可通过截尾的奇异值分解公式 $\hat{\boldsymbol{A}}=\sum\limits_{i=1}^{k}\sigma_i\boldsymbol{u}_i\boldsymbol{v}_i^{\mathrm{T}}$ 重构出原图像. 一个容易理解的事实是，若 k 值偏小，即压缩比 ρ 偏大，则重构的图像的质量有可能不令人满意. 反之，过大的 k 值又会导致压缩比过小，降低图像压缩和传送的效率. 所以，需要针对各种不同种类的图像，选择合适的压缩比，以兼顾图像传送效率和重构质量.

一、基本概念

1. 三角矩阵的定义

(1) 如果 $a_{ii}(i=1,2,\cdots,n)$ 均为正实数，$a_{ij}\in\mathbf{C}(a_{ij}\in\mathbf{R})(i<j,i=1,2,\cdots,n-1;j=i+1,i+2,\cdots,n)$，则上三角矩阵 $\boldsymbol{R}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$ 称为正线上三角复(实)矩阵. 特别当 $a_{ii}=1(i=1,2,\cdots,n)$ 时，$\boldsymbol{R}$ 称为单位上三角复(实)矩阵.

(2) 如果 $a_{ii}(i=1,2,\cdots,n)$ 均为正实数，$a_{ij}\in\mathbf{C}(a_{ij}\in\mathbf{R})(i>j,j=1,2,\cdots,n-1;i=j+1,j+2,\cdots,n)$，则下三角矩阵 $\boldsymbol{L}=\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ 称为正线下三角复(实)矩阵. 特别当 $a_{ii}=1(i=1,2,\cdots,n)$ 时，$\boldsymbol{L}$ 称为单位下三角复(实)矩阵.

2. 单纯矩阵　若矩阵 $\boldsymbol{A}$ 的每个特征值的代数重复度与几何重复度相等，则称 $\boldsymbol{A}$ 为单纯矩阵.

3. Hermite 矩阵　设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，若 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$，则称 $\boldsymbol{A}$ 是 Hermite 矩阵；若 $\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$，则称 $\boldsymbol{A}$ 是反 Hermite 矩阵. 当 $\boldsymbol{A}$ 为实对称矩阵时，$\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，所以实对称矩阵是 Hermite 矩阵的特殊情形.

4. 正规矩阵　若 n 阶复矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$，则称 $\boldsymbol{A}$ 为正规矩阵. 当 $\boldsymbol{A}$ 为 n 阶实矩阵且满足 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 时，则称 $\boldsymbol{A}$ 为实正规矩阵.

显然，对角矩阵、酉矩阵、Hermite 矩阵、反 Hermite 矩阵都是正规矩阵；正交矩阵、实对称矩阵和实反对称矩阵都是实正规矩阵. 但正规矩阵并不一定是 Hermite 矩阵.

5. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 是 Hermite 矩阵，对任意非零向量 $\boldsymbol{X}\in\mathbf{C}^{n}$，均有

$$f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{X}>0\quad(\geqslant 0),$$

则称二次型 $f(\boldsymbol{X})$ 是正定(半正定)二次型，此时称系数矩阵 $\boldsymbol{A}$ 正定(半正定).

6. 广义正定矩阵　设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$，若对于任何非零向量 $\boldsymbol{X}\in\mathbf{R}^{n}$，均有

$$\boldsymbol{X}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{X}>0,$$

则称矩阵 $\boldsymbol{A}$ 为广义正定矩阵.

7. 奇异值　设 $\boldsymbol{A}\in\mathbf{C}_r^{n\times n}$，$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值为

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r\geqslant\lambda_{r+1}=\cdots=\lambda_n=0,$$

则称 $\sigma_i=\sqrt{\lambda_i}(i=1,2,\cdots,r)$ 为矩阵 $\boldsymbol{A}$ 的正奇异值.

8. 酉等价　设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{m\times n}$，如果存在酉矩阵 $\boldsymbol{U}\in\mathbf{C}^{m\times m}$ 和 $\boldsymbol{V}\in\mathbf{C}^{n\times n}$，使得 $\boldsymbol{A}=\boldsymbol{UBC}$，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉等价.

二、主要结论

1. n 阶非奇异方阵的三角分解的存在性

定理 1　设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$，则 $\boldsymbol{A}$ 可唯一地分解为

$$\boldsymbol{A}=\boldsymbol{U}_1\boldsymbol{R},$$

其中 $\boldsymbol{U}_1$ 是酉矩阵，$\boldsymbol{R}$ 是正线上三角复矩阵；或者，$\boldsymbol{A}$ 可唯一地分解为

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}_2,$$

其中 $\boldsymbol{L}$ 是正线下三角复矩阵，$\boldsymbol{U}_2$ 是酉矩阵.

若矩阵 $\boldsymbol{A}\in\mathbf{R}_n^{n\times n}$，则可立即得到下面推论.

推论 1　设 $\boldsymbol{A}\in\mathbf{R}_n^{n\times n}$，则 $\boldsymbol{A}$ 可唯一地分解为

$$\boldsymbol{A}=\boldsymbol{Q}_1\boldsymbol{R},$$

其中 $\boldsymbol{Q}_1$ 是正交矩阵，$\boldsymbol{R}$ 是正线上三角实矩阵；或者，$\boldsymbol{A}$ 可唯一地分解为

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{Q}_2,$$

其中 $\boldsymbol{L}$ 是正线上三角实矩阵，$\boldsymbol{Q}_2$ 是正交矩阵.

推论 2　设 $\boldsymbol{A}$ 是实对称正定矩阵，则存在唯一正线上三角实矩阵 $\boldsymbol{R}$，使得

$$\boldsymbol{A}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}.$$

推论 3　设 $\boldsymbol{A}$ 是正定 Hermite 矩阵，则存在唯一正线上三角复矩阵 $\boldsymbol{R}$，使得

$$\boldsymbol{A}=\boldsymbol{R}^{\mathrm{H}}\boldsymbol{R}.$$

定理 2　设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$，用 $\boldsymbol{L}$ 表示下三角复矩阵，$\boldsymbol{L}^*$ 为单位下三角复矩阵，$\boldsymbol{R}$ 为上三角复矩阵，$\boldsymbol{R}^*$ 为单位上三角复矩阵，$\boldsymbol{D}$ 为对角矩阵，则下列命题等价：

(1) $\boldsymbol{A}$ 的各顺序主子式满足 $\Delta_k=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}\neq 0(k=1,2,\cdots,n)$；

(2) $\boldsymbol{A}$ 可唯一地分解为 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{R}^*$，并且 $\boldsymbol{L}$ 的主对角线上的元素不为零；

(3) $\boldsymbol{A}$ 可唯一地分解为 $\boldsymbol{A}=\boldsymbol{L}^*\boldsymbol{D}\boldsymbol{R}^*$，并且 $\boldsymbol{D}$ 的主对角线上的元素不为零；

(4) $\boldsymbol{A}$ 可唯一地分解为 $\boldsymbol{A}=\boldsymbol{L}^*\boldsymbol{R}$，并且 $\boldsymbol{R}$ 的主对角线上的元素不为零.

2. 任意矩阵的三角分解

定理 3　当 $\boldsymbol{A}$ 是行满秩或列满秩矩阵时，可得如下结论：

(1) 若 $\boldsymbol{A}\in\mathbf{C}_m^{m\times n}$，则存在 m 阶正线下三角复矩阵 $\boldsymbol{L}$ 和 n 阶酉矩阵 $\boldsymbol{U}$，使得

$$A = (L\ 0)U.$$

（2）若 $A\in C_n^{m\times n}$，则存在 m 阶酉矩阵 U 和 n 阶正线下三角复矩阵 R，使得

$$A = U\begin{pmatrix} R \\ 0 \end{pmatrix}.$$

为叙述方便，我们用 $U_m^{m\times n}$ 表示 m 个两两正交的单位行向量组成的矩阵的集合，用 $U_n^{m\times n}$ 表示 n 个两两正交的单位列向量组成的矩阵的集合，于是有下面定理.

定理 4　（1）若 $A\in C_m^{m\times n}$ 则 A 可唯一地分解为

$$A = LU,$$

其中 L 是 m 阶正线下三角复矩阵，$U\in U_m^{m\times n}$；

（2）若 $A\in C_m^{m\times n}$，则 A 可唯一地分解为

$$A = UR,$$

其中 $U\in U_m^{m\times n}$，R 是 n 阶正线上三角复矩阵.

若 A 既不是行满秩也不是列满秩矩阵，则可得下面定理.

定理 5　设 $A\in C_r^{m\times n}(r<\min\{m,n\})$，则存在酉矩阵 $U\in C^{m\times m}$，$V\in C^{n\times n}$ 及 r 阶正线下三角矩阵 L，使得

$$A = U\begin{pmatrix} L & 0 \\ 0 & 0 \end{pmatrix}V,$$

或者，存在酉矩阵 $U\in C^{m\times m}$，$V\in C^{n\times n}$ 及 r 阶正线上三角矩阵 R，使得

$$A = U\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}V.$$

3. 谱分解定理

矩阵的特征值在振动问题和稳定性问题中有着非常重要的作用. 相似矩阵有相同的特征值，于是人们总希望能找到其中的对角矩阵或 Jordan 标准形矩阵，即用最简单的形式来表示已知矩阵，这就是下面的矩阵谱分解.

定理 6　矩阵 A 的任意特征值的几何重复度不大于它的代数重复度.

定理 7　A 是单纯矩阵的充要条件是 A 与对角矩阵相似.

定理 8　设 $A\in C_n^{n\times n}$ 是单纯矩阵，则设 A 可分解为一系列幂等矩阵 $A_i(i=1,2,\cdots,n)$ 的加权和，即

$$A = \sum_{i=1}^{n}\lambda_i A_i,$$

其中 $\lambda_i(i=1,2,\cdots,n)$ 是 A 的特征值.

幂等矩阵 A_i 具有如下性质：

（1）幂等性　$A_i^2 = A_i$；

（2）分离性　$A_iA_j = 0(j\neq i)$；

(3) 可加性 $\sum_{i=1}^{n}\boldsymbol{A}_i=\boldsymbol{E}_n$.

定理 9 设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$，它有 k 个相异特征值 $\lambda_i(i=1,2,\cdots,k)$，则 $\boldsymbol{A}$ 是单纯矩阵的充要条件是存在 k 个矩阵 $\boldsymbol{A}_i(i=1,2,\cdots,k)$，满足(1)$\boldsymbol{A}_i\boldsymbol{A}_j=\begin{cases}\boldsymbol{A}_i, & j=i,\\ \boldsymbol{0}, & j\neq i;\end{cases}$ (2) $\sum_{i=1}^{n}\boldsymbol{A}_i=\boldsymbol{E}_n$ (3) $\boldsymbol{A}=\sum_{i=1}^{n}\lambda_i\boldsymbol{A}_i$.

定理 10 n 阶复矩阵 $\boldsymbol{A}$ 是正规矩阵的充要条件是 $\boldsymbol{A}$ 与对角矩阵酉相似，即存在 n 阶酉矩阵 $\boldsymbol{U}$，使得 $\boldsymbol{A}=\boldsymbol{U}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\boldsymbol{U}^{\mathrm{H}}$，其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值.

定理 11 设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$，它有 k 个相异特征值 $\lambda_i(i=1,2,\cdots,k)$，则 $\boldsymbol{A}$ 是正规矩阵的充要条件是存在 k 个矩阵 $\boldsymbol{A}_i(i=1,2,\cdots,k)$满足

(1) $\boldsymbol{A}_i\boldsymbol{A}_j=\begin{cases}\boldsymbol{A}_i, & j=i,\\ \boldsymbol{0}, & j\neq i;\end{cases}$

(2) $\sum_{i=1}^{n}\boldsymbol{A}_i=\boldsymbol{E}_n$;

(3) $\boldsymbol{A}=\sum_{i=1}^{n}\lambda_i\boldsymbol{A}_i$;

(4) $\boldsymbol{A}_i^{\mathrm{H}}=\boldsymbol{A}_i(i=1,2,\cdots,k)$.

当矩阵 $\boldsymbol{A}$ 不能与对角矩阵相似时，可得下面定理.

定理 12 若 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$，则 $\boldsymbol{A}$ 可分解为 $\boldsymbol{A}=\sum_{i=1}^{k}(\lambda_i\boldsymbol{A}_i+\boldsymbol{B}_i)$，其中 n 阶矩阵 $\boldsymbol{A}_i,\boldsymbol{B}_i(i=1,2,\cdots,k)$满足

(1) $\boldsymbol{A}_i\boldsymbol{A}_j=\begin{cases}\boldsymbol{A}_i, & j=i,\\ \boldsymbol{0}, & j\neq i;\end{cases}$

(2) $\boldsymbol{B}_i\boldsymbol{B}_j=\boldsymbol{0}(j\neq i)$;

(3) $\sum_{i=1}^{n}\boldsymbol{A}_i=\boldsymbol{E}_n$.

4. Hermite 矩阵的分解及性质

定理 13 设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$ 是 Hermite 矩阵，则

(1) $(\boldsymbol{A\alpha},\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{A\beta})$，$\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbf{C}^n$;

(2) $\boldsymbol{A}$ 的特征值均为实数;

(3) 属于 $\boldsymbol{A}$ 的不同特征值的特征向量正交.

定理 14 设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$ 是反 Hermite 矩阵，则 $\boldsymbol{A}$ 的特征值均为纯虚数.

定理 15 设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$ 是 Hermite 矩阵，则下列命题等价：

(1) $\boldsymbol{A}$ 是正定矩阵；

(2) $\boldsymbol{A}$ 的特征值全为实数；

(3) $\boldsymbol{A}$ 与 $\boldsymbol{E}$ 合同；

(4) $\boldsymbol{A}$ 的顺序主子式全为正.

定理 16 设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$ 是正定 Hermite 矩阵，则下列命题等价

(1) $\boldsymbol{A}$ 的主对角线上的元素均大于零；

(2) 存在正定 Hermite 矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}=\boldsymbol{B}^2$；

(3) $\boldsymbol{A}$ 的任意 k 行和对应的 k 列组成的主子阵式正定的，即

$$\boldsymbol{A}_{i_1i_2\cdots i_k}=\begin{pmatrix} a_{i_1i_1} & a_{i_1i_2} & \cdots & a_{i_1i_k} \\ a_{i_2i_1} & a_{i_2i_2} & \cdots & a_{i_2i_k} \\ \vdots & \vdots & & \vdots \\ a_{i_ki_1} & a_{i_ki_2} & \cdots & a_{i_ki_k} \end{pmatrix},\quad 1\leqslant i_1<i_2<\cdots<i_k\leqslant n$$

是正定矩阵；

(4) 设 $\boldsymbol{A}$ 的对角线上的元素为 $a_{ii}(i=1,2,\cdots,n)$，则有 $\det\boldsymbol{A}\leqslant\prod\limits_{i=1}^{n}a_{ii}$，等号成立当且仅当 $\boldsymbol{A}$ 是对角矩阵.

定理 17 设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$ 是正定 Hermite 矩阵，则下列命题等价：

(1) $\boldsymbol{A}$ 是半正定矩阵；

(2) $\boldsymbol{A}$ 的特征值非负；

(3) $\boldsymbol{A}$ 与 $\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}$ 合同，其中 $r=\operatorname{rank}(\boldsymbol{A})$；

(4) $\boldsymbol{A}$ 的顺序主子式均非负，即 $\det(\boldsymbol{A}_{i_1i_2\cdots i_k})\geqslant 0(1\leqslant i_1<i_2<\cdots<i_k\leqslant n,k=1,2,\cdots,n)$.

定理 18 设 $\boldsymbol{A}\in\mathbf{C}_n^{n\times n}$ 是正定 Hermite 矩阵，则 $\boldsymbol{A}$ 可分解为

$$\boldsymbol{A}=(\tilde{\boldsymbol{L}}\boldsymbol{D}^{\frac{1}{2}})(\tilde{\boldsymbol{L}}\boldsymbol{D}^{\frac{1}{2}})^{\mathrm{H}}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{H}},$$

其中 $\boldsymbol{L}=\tilde{\boldsymbol{L}}\boldsymbol{D}^{\frac{1}{2}}$，$\tilde{\boldsymbol{L}}$ 是单位下三角矩阵，对角矩阵 $\boldsymbol{D}^{\frac{1}{2}}$ 可表示为

$$\boldsymbol{D}^{\frac{1}{2}}=\operatorname{diag}\left(\sqrt{\Delta_1},\sqrt{\frac{\Delta_2}{\Delta_1}},\cdots,\sqrt{\frac{\Delta_n}{\Delta_{n-1}}}\right),$$

这里 $\Delta_k(k=1,2,\cdots,n)$ 是 $\boldsymbol{A}$ 的 k 阶顺序主子式.

把长方形矩阵 $\boldsymbol{A}$ 分解为两个与 $\boldsymbol{A}$ 同秩的因子的乘积，进而讨论不同分解之间的关系，这一方法在处理广义逆矩阵的问题中起着非常重要的作用.

定理 19（最大秩分解定理） 设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$，则存在矩阵 $\boldsymbol{B}\in\mathbf{C}_r^{m\times r}$，$\boldsymbol{D}\in\mathbf{C}_r^{r\times n}$，使得 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{D}$.

定理 20 设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$，且 $\boldsymbol{A}=\boldsymbol{B}_1\boldsymbol{D}_1=\boldsymbol{B}_2\boldsymbol{D}_2$ 均为 $\boldsymbol{A}$ 的最大秩分解，则

(1) 存在 r 阶可逆矩阵 $\boldsymbol{Q}$,使得 $\boldsymbol{B}_1=\boldsymbol{B}_2\boldsymbol{Q},\boldsymbol{D}_1=\boldsymbol{Q}^{-1}\boldsymbol{D}_2$;

(2) $\boldsymbol{D}_1^{\mathrm{H}}(\boldsymbol{D}_1\boldsymbol{D}_1^{\mathrm{H}})(\boldsymbol{B}_1^{\mathrm{H}}\boldsymbol{B}_1)^{-1}\boldsymbol{B}_1^{\mathrm{H}}=\boldsymbol{D}_2^{\mathrm{H}}(\boldsymbol{D}_2\boldsymbol{D}_2^{\mathrm{H}})(\boldsymbol{B}_2^{\mathrm{H}}\boldsymbol{B}_2)^{-1}\boldsymbol{B}_2^{\mathrm{H}}$.

5. 矩阵的奇异值分解

矩阵的奇异值分解是现代矩阵理论研究的一项非常重要的内容,在线性动态系统的识别、实验数据的处理、古典控制理论中的频率测算等方面都有极为广泛的应用.

定理 21 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$,则

(1) $\operatorname{rank}(\boldsymbol{A})=\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\operatorname{rank}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})$;

(2) $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A},\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的特征值均为非负实数;

(3) $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的非零特征值相同.

定理 22 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉等价,则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的正奇异值.

注 矩阵 $\boldsymbol{A}$ 的奇异值还有下列结果,有兴趣的读者不妨加以证明.

定理 23 矩阵 $\boldsymbol{A}^{\mathrm{H}}$ 与矩阵 $\boldsymbol{A}$ 有相同的非零奇异值;若矩阵 $\boldsymbol{A}$ 为可逆矩阵,则 $\boldsymbol{A}^{-1}$ 的奇异值与 $\boldsymbol{A}$ 的奇异值互为倒数;若 $\boldsymbol{A}$ 为正定矩阵,则 $\boldsymbol{A}$ 的奇异值与 $\boldsymbol{A}$ 的特征值相同.

三、典型例题

例 1 用 Schmidt 正交化方法求矩阵 $\boldsymbol{A}$ 的分解,其中

$$\boldsymbol{A}=\begin{pmatrix}0&1&1\\1&1&0\\1&0&1\end{pmatrix}.$$

解 令 $\boldsymbol{\alpha}_1=(0,1,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,0,1)^{\mathrm{T}}$,由 Schmidt 正交化方法得

$$\boldsymbol{\beta}_1=(0,1,1)^{\mathrm{T}}=\boldsymbol{\alpha}_1,\quad |\boldsymbol{\beta}_1|=\sqrt{2},\quad \boldsymbol{\xi}_1=\frac{1}{\sqrt{2}}(0,1,1)^{\mathrm{T}},$$

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-(\boldsymbol{\alpha}_2,\boldsymbol{\xi}_1)\boldsymbol{\xi}_1=\left(1,\frac{1}{2},-\frac{1}{2}\right)^{\mathrm{T}},\quad (\boldsymbol{\alpha}_2,\boldsymbol{\xi}_1)=\frac{1}{\sqrt{2}},\quad |\boldsymbol{\beta}_2|=\frac{\sqrt{6}}{2},$$

$$\boldsymbol{\xi}_2=\frac{1}{\sqrt{6}}(2,1,-1)^{\mathrm{T}},\quad (\boldsymbol{\xi}_1,\boldsymbol{\alpha}_3)=\frac{1}{\sqrt{2}},\quad (\boldsymbol{\xi}_2,\boldsymbol{\alpha}_3)=\frac{1}{\sqrt{6}},$$

$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-(\boldsymbol{\alpha}_3,\boldsymbol{\xi}_1)\boldsymbol{\xi}_1-(\boldsymbol{\alpha}_3,\boldsymbol{\xi}_2)\boldsymbol{\xi}_2=\left(\frac{2}{3},-\frac{2}{3},\frac{2}{3}\right)^{\mathrm{T}},$$

$$|\boldsymbol{\beta}_3|=\frac{2}{\sqrt{3}},\quad \boldsymbol{\xi}_3=\frac{1}{\sqrt{3}}(1,-1,1)^{\mathrm{T}}.$$

于是

$$Q=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3)=\begin{bmatrix}0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{bmatrix},\quad R=\begin{bmatrix}\sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ 0 & \frac{\sqrt{6}}{2} & \frac{1}{\sqrt{6}}\\ 0 & 0 & \frac{2}{\sqrt{3}}\end{bmatrix},$$

可得

$$A=QR.$$

例 2 设 $A=\begin{pmatrix}2 & 0 & 1\\ 1 & 2 & 0\end{pmatrix}$，求 A 的奇异值分解.

解 (1) 计算 $A^{\mathrm{T}}A$ 及 A 的奇异值 σ_i，由

$$A^{\mathrm{T}}A=\begin{bmatrix}5 & 2 & 0\\ 2 & 4 & 0\\ 2 & 0 & 1\end{bmatrix},$$

$|\lambda E-A^{\mathrm{T}}A|=\lambda(\lambda-3)(\lambda-7)$，故 $A^{\mathrm{T}}A$ 的特征值为 $7,3,0$，A 的正奇异值为 $\sigma_1=\sqrt{7}$，$\sigma_2=3$.

(2) 求 $A^{\mathrm{T}}A$ 的 3 个标准正交特征向量. 对于 $\lambda_1=7$，求得

$$\boldsymbol{\alpha}_1=(3,2,1)^{\mathrm{T}},\quad \boldsymbol{\varepsilon}_1=\frac{1}{\sqrt{14}}(3,2,1)^{\mathrm{T}},$$

由 $\lambda_1=3$，求得

$$\boldsymbol{\alpha}_2=(1,-2,1)^{\mathrm{T}},\quad \varepsilon_2=\frac{1}{\sqrt{6}}(1,-2,1)^{\mathrm{T}},$$

由 $\lambda_1=0$，求得

$$\boldsymbol{\alpha}_3=(2,-1,-4)^{\mathrm{T}},\quad \boldsymbol{\varepsilon}_3=\frac{1}{\sqrt{21}}(2,-1,-4)^{\mathrm{T}},$$

于是可得酉矩阵，

$$V_1=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2).$$

计算得到

$$U_1=AV_1\Sigma^{-1}=\begin{pmatrix}2 & 0 & 1\\ 1 & 2 & 0\end{pmatrix}\begin{bmatrix}\frac{3}{\sqrt{14}} & \frac{1}{\sqrt{6}}\\ \frac{2}{\sqrt{14}} & -\frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{14}} & \frac{1}{\sqrt{6}}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{7}} & 0\\ 0 & \frac{1}{\sqrt{3}}\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}.$$

令 $U=U_1$，$V=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3)$，则有

$$A=U\begin{pmatrix}\sqrt{7} & & \\ & \sqrt{3} & \\ & & 0\end{pmatrix}V^{\mathrm{H}}.$$

例 3 求矩阵 $A=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}$ 的奇异值分解.

解 $B=A^{\mathrm{T}}A=\begin{pmatrix}1&0&1\\0&1&1\\1&1&2\end{pmatrix}$ 的特征值对应的特征向量依次为

$$\boldsymbol{\xi}_1=\begin{pmatrix}1\\1\\2\end{pmatrix},\quad \boldsymbol{\xi}_2=\begin{pmatrix}1\\-1\\0\end{pmatrix},\quad \boldsymbol{\xi}_3=\begin{pmatrix}1\\1\\-1\end{pmatrix}.$$

于是可得

$$\operatorname{rank}A=2,\quad \boldsymbol{\Sigma}=\begin{pmatrix}\sqrt{3}&0\\0&1\end{pmatrix},$$

且有正交矩阵

$$V=\begin{pmatrix}\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{6}}&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}\\ \frac{2}{\sqrt{6}}&0&-\frac{1}{\sqrt{3}}\end{pmatrix}.$$

计算得到

$$U_1=AV_1\boldsymbol{\Sigma}^{-1}=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\ 0&0\end{pmatrix},$$

$$U_2=\begin{bmatrix}0\\0\\1\end{bmatrix},\quad U=(U_1U_2)=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\ \frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\ 0&0&1\end{pmatrix}.$$

则 A 的奇异值分解为

$$A = U\begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}V^{\mathrm{T}}.$$

总结 A 的奇异值分解的步骤如下：

(1) 计算 $A^{\mathrm{H}}A$，求 $A^{\mathrm{H}}A$ 的 n 个特征值 λ_i 及 A 的正奇异值 $\sigma_i=\sqrt{\lambda_i}$，$i=1,2,\cdots,r$.

(2) 求 $A^{\mathrm{H}}A$ 的 n 个特征向量，并用正交化方法化为标准正交向量 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$，它们构成酉矩阵

$$V = (\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_r,\boldsymbol{\varepsilon}_{r+1},\cdots,\boldsymbol{\varepsilon}_n).$$

(3) 取 $V_1=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_r)$，$\boldsymbol{\Sigma}_1=\mathrm{diag}(\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2,\cdots,\boldsymbol{\sigma}_n)$，并计算

$$U_1 = AV_1\boldsymbol{\Sigma}_1^{-1}(U_1\text{ 的 }r\text{ 个列正交}).$$

(4) 在 $\mathbf{C}^m$ 中，取与 U_1 的列向量正交的 $m-r$ 个标准正交向量组成 U_2，使得 $(U_1U_2)=U$ 为酉矩阵，从而求得 U，则

$$A = U\begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}V^{\mathrm{T}}.$$

例 4 设 A 是 n 阶可逆实矩阵，则 A 可表示成一个正交矩阵 Q 与正定矩阵 S 的乘积，即 $A=QS$.

证 构造矩阵 $B=A^{\mathrm{T}}A$，因为 A 是可逆的，则 B 为实对称矩阵，由本章习题 5 的证明过程可知，存在一个 n 阶方阵 S，使得

$$B = S^2,$$

其中 S 为对称正定矩阵. 令 Q 为 n 阶正交矩阵，即 $Q^{\mathrm{T}}Q=E$，则有

$$B = S^2 = SIS = SQ^{\mathrm{T}}QS = (SQ^{\mathrm{T}})(QS) = (QS)^{\mathrm{T}}(QS),$$

故 $A=QS$，其中 Q 为 n 阶正交矩阵，S 为 n 阶对称正定矩阵.

例 5 若 A 和 B 是任意两个 $m\times n$ 矩阵，则 $A^{\mathrm{H}}A=B^{\mathrm{H}}B$ 当且仅当存在一个 $m\times n$ 酉矩阵 Q，使得

$$QA = B.$$

证 充分性. 若存在酉矩阵 Q，使得 $QA=B$，则 $B^{\mathrm{H}}B=A^{\mathrm{H}}Q^{\mathrm{H}}QA=A^{\mathrm{H}}A$.

必要性. 令 A 和 B 的奇异值分解分别为

$$A = U_A\boldsymbol{\Sigma}_AV_A^{\mathrm{H}},\quad B = U_B\boldsymbol{\Sigma}_BV_B^{\mathrm{H}},$$

上式中的 U_A 和 U_B 均为 $m\times m$ 酉矩阵，V_A 和 V_B 均为 $n\times n$ 酉矩阵，而 $m\times n$ 矩阵 $\boldsymbol{\Sigma}_A$ 和 $\boldsymbol{\Sigma}_B$ 分别包含了矩阵 A 和 B 的非负奇异值. 由于

$$A^{\mathrm{H}}A = V_A\boldsymbol{\Sigma}_A^{\mathrm{H}}\boldsymbol{\Sigma}_AV_A^{\mathrm{H}},\quad B^{\mathrm{H}}B = V_B\boldsymbol{\Sigma}_B^{\mathrm{H}}\boldsymbol{\Sigma}_BV_B^{\mathrm{H}}.$$

若 $A^{\mathrm{H}}A=B^{\mathrm{H}}B$，则有 $V_A=V_B$ 和 $\boldsymbol{\Sigma}_A=\boldsymbol{\Sigma}_B$. 定义矩阵 $Q=U_BU_A^{\mathrm{H}}$，易知

$$QA = U_BU_A^{\mathrm{H}}A = U_BU_A^{\mathrm{H}}U_A\boldsymbol{\Sigma}_AV_A^{\mathrm{H}} = U_B\boldsymbol{\Sigma}_BV_B^{\mathrm{H}} = B.$$

例 6 设 $A=(a_{ij})_{n\times n}$，$\lambda_i(i=1,2,\cdots,n)$ 为 A 的特征值，若 $\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2=\sum_{i=1}^{n}|\lambda_i|^2$，

证明 $\boldsymbol{A}$ 是正规矩阵.

证 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}\lambda_1 & t_{12} & \cdots & t_{1n}\\ & \lambda_2 & \cdots & t_{2n}\\ & & \ddots & \vdots\\ & & & \lambda_n\end{bmatrix}$, $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}=\begin{bmatrix}\bar{\lambda}_1 & & & \\ \bar{t}_{12} & \bar{\lambda}_2 & & \\ \vdots & \vdots & \ddots & \\ \bar{t}_{1n} & \bar{t}_{2n} & \cdots & \bar{\lambda}_n\end{bmatrix}$,

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}\lambda_1\bar{\lambda}_1 & * & \cdots & * \\ * & |\lambda_2|^2+|t_{12}|^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ * & \cdots & * & |\lambda_n|^2+\sum_{i=1}^{n}|t_{12}|\end{bmatrix},$$

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 &= \operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\operatorname{tr}(\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U})\\ &= \sum_{i=1}^{n}|\lambda_i|^2+\sum_{i=1}^{n}\sum_{j=1}^{n}|t_{ij}|^2(i>j\text{ 时},t_{ij}=0).\end{aligned}$$

由题设 $\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2=\sum_{i=1}^{n}|\lambda_i|^2$,可得 $\sum_{i=1}^{n}\sum_{j=1}^{n}|t_{ij}|^2=0(i>j$ 时,$t_{ij}=0)$,从而有

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

所以 $\boldsymbol{A}$ 可以与对角矩阵相似,因而 $\boldsymbol{A}$ 是正规矩阵.

例 7 设 $\boldsymbol{A}$ 是正规矩阵,证明:

(1) $\boldsymbol{A}$ 的特征向量也是 $\boldsymbol{A}^{\mathrm{H}}$ 的特征向量;

(2) $\forall \boldsymbol{X}\in\mathbf{C}^m$,$\boldsymbol{AX}$ 与 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{X}$ 的长度相等.

证 (1) $\boldsymbol{A}$ 为正规矩阵,则有酉矩阵,使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},\quad \boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}=\begin{bmatrix}\bar{\lambda}_1 & & & \\ & \bar{\lambda}_2 & & \\ & & \ddots & \\ & & & \bar{\lambda}_n\end{bmatrix},$$

其中 $\boldsymbol{U}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为 $\boldsymbol{A}$ 的特征向量,由上两式可知

$$\boldsymbol{A}\boldsymbol{\alpha}_i=\lambda_i\boldsymbol{\alpha}_i\quad \boldsymbol{A}^{\mathrm{H}}\boldsymbol{\alpha}_i=\bar{\lambda}_i\boldsymbol{\alpha}_i,$$

故 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{H}}$ 有相同的特征向量.

(2) 由 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$,得

$$|\boldsymbol{A}^{\mathrm{H}}\boldsymbol{X}|^2=(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{X})^{\mathrm{H}}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{X})=\boldsymbol{X}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{X}=\boldsymbol{X}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{X}=(\boldsymbol{AX})^{\mathrm{H}}(\boldsymbol{AX})=|\boldsymbol{AX}|^2.$$

所以

$$|A^{\mathrm{H}}X|=|AX|.$$

例 8 设 A 为正规矩阵，则 A 的谱分解式有

$$A=\lambda_1U_1U_1^{\mathrm{H}}+\lambda_2U_2U_2^{\mathrm{H}}+\cdots+\lambda_nU_nU_n^{\mathrm{H}},$$

其中 $U_1,U_2,\cdots,U_n$ 是 A 的 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 对应的标准正交特征向量.

证 A 是正规矩阵，则存在酉矩阵 U，使得

$$U^{\mathrm{H}}AU=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},\quad A=U\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}U^{\mathrm{H}}.$$

把 U 列分块为 $U=(U_1\quad U_2\quad\cdots\quad U_n)$，则

$$A=(U_1\quad U_2\quad\cdots\quad U_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}\begin{pmatrix}U_1^{\mathrm{H}}\\ U_2^{\mathrm{H}}\\ \vdots\\ U_n^{\mathrm{H}}\end{pmatrix}$$

$$=\lambda_1U_1U_1^{\mathrm{H}}+\lambda_2U_2U_2^{\mathrm{H}}+\cdots+\lambda_nU_nU_n^{\mathrm{H}}.$$

例 9 设 Hermite 矩阵 A 非负定，证明：存在一个三角矩阵 T，使得 $A=TT^{\mathrm{H}}$.

证 由于 Hermite 矩阵为正规矩阵，由定理 10 知

$$A=U\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)U^{\mathrm{H}}.$$

令 $P=\mathrm{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n})U^{\mathrm{H}}$，则有 $A=P^{\mathrm{H}}P$. 由定理 1 知

$$P=UR,$$

其中 U 是酉矩阵，R 是上三角复矩阵，所以

$$A=(UR)^{\mathrm{H}}UR=R^{\mathrm{H}}R.$$

令 $T=R^{\mathrm{H}}$，则可得 $A=TT^{\mathrm{H}}$.

例 10 试证：如果 A 为 n 阶正规矩阵，且 $Ax=\lambda x, Ay=\mu y$，其中 $\lambda\neq\mu$，则 x 与 y 正交.

证 A 为 n 阶正规矩阵 $\Rightarrow A=U^{\mathrm{H}}\Lambda U\Rightarrow$

$$Ax=\lambda x\Rightarrow U^{\mathrm{H}}\Lambda Ux=\lambda x\Rightarrow \Lambda Ux=\lambda Ux\xrightarrow{x'=Ux}\Lambda x'=\lambda x'.$$

设 $x'=(x_1',x_2',\cdots,x_n')^{\mathrm{T}}\Rightarrow\lambda_i\neq\lambda$ 时，$x_i'=0$，

$$Ay=\lambda y\Rightarrow U^{\mathrm{H}}\Lambda Uy=\mu y\Rightarrow \Lambda Uy=\mu Uy\Rightarrow\xrightarrow{y'=Uy}\Lambda y'=\lambda y'.$$

设 $y'=(y_1',y_2',\cdots,y_n')^{\mathrm{T}}\Rightarrow\mu_i\neq\mu$ 时，$y_i'=0$，

$$\lambda\neq\mu\Rightarrow(x',y')=0.$$

故

$$0=(x',y')=(Ux)^{\mathrm{H}}Uy=x^{\mathrm{H}}U^{\mathrm{H}}Uy=x^{\mathrm{H}}y=(x,y).$$

四、习题解答

1. 求矩阵

$$A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$$

的谱分解.

解 (1) 求特征值 $|\lambda E - A| = (\lambda_1 - 3)(\lambda_2 + 1) = 0$,所以特征值为 $\lambda_1 = 3, \lambda_2 = -1$.

(2) 求特征向量 $\lambda_1 = 3$ 对应的特征向量为 $p_1 = (1,2)^T$;$\lambda_2 = -1$ 对应的特征向量为 $p_2 = (1,-2)^T$.

(3) 谱分解 令 $P = (p_1, p_2) = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}$,则 $P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & -\frac{1}{4} \end{pmatrix} = \begin{pmatrix} \omega_1^T \\ \omega_2^T \end{pmatrix}$. 令 $A_1 = p_1 \omega_1^T = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ 1 & \frac{1}{2} \end{pmatrix}$,$A_2 = p_2 \omega_2^T = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} \\ -1 & \frac{1}{2} \end{pmatrix}$,则谱分解式为 $A = 3A_1 - A_2$.

2. 求单纯矩阵

$$A = \begin{pmatrix} -29 & 6 & 18 \\ -20 & 5 & 12 \\ -40 & 8 & 25 \end{pmatrix}$$

的谱分解式.

解 提示:先求出矩阵 A 的特征值 $\lambda_1, \lambda_2, \lambda_3$ 及其相应特征向量,并对特征向量进行单位正交化,再根据特征向量构造矩阵 P,使得 $A = P\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)P^{-1}$,最后根据 P 和 P^{-1} 的行和列可写出 A 的谱分解式.

3. 设 $\lambda_i (i = 1, 2, \cdots, n)$ 是正规矩阵 $A \in \mathbf{C}^{n\times n}$ 的特征值,证明:$|\lambda_i|^2 (i = 1, 2, \cdots, n)$ 是 $A^H A$ 与 AA^H 的特征值.

证 根据题设 A 酉相似与对角矩阵,即

$$A = U\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)U^H,$$

其中 U 为酉矩阵,则

$$\begin{aligned} A^H A &= (U\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)U^H)^H (U\mathrm{diag}(\lambda_1, \lambda_2 \cdots, \lambda_n)U^H) \\ &= U\mathrm{diag}(|\lambda_1|^2, |\lambda_2|^2, \cdots, |\lambda_n|^2)U^H, \end{aligned}$$

即 $A^H A$ 的特征值为 $|\lambda_i|^2 (i = 1, 2, \cdots, n)$,同理可证 $|\lambda_i|^2 (i = 1, 2, \cdots, n)$ 也是 AA^H 的特征值.

4. 设 $\boldsymbol{A}$ 是 $n\times n$ 的实对称矩阵，并且 $\boldsymbol{A}^2=\boldsymbol{0}$，问能用几种方法证明 $\boldsymbol{A}=\boldsymbol{0}$.

证 法 1 设 λ 是矩阵 $\boldsymbol{A}$ 的一个特征值，$\boldsymbol{x}$ 是对应于 λ 的一个非零特征向量，即 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，$\boldsymbol{A}^2\boldsymbol{x}=\lambda^2\boldsymbol{x}=\boldsymbol{0}$，则 $\lambda^2=0$，即 $\lambda=0$，所以矩阵 $\boldsymbol{A}$ 的特征值全为零，又 $\boldsymbol{A}$ 酉相似于对角矩阵 $\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，所以 $\boldsymbol{A}=\boldsymbol{0}$.

法 2 设 $\boldsymbol{A}\neq\boldsymbol{0}$，则 $\boldsymbol{A}^2=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\neq\boldsymbol{0}$，与题设矛盾，所以结论成立.

5. 试证：对于每一个实对称矩阵 $\boldsymbol{A}$，都存在一个 n 阶方阵 $\boldsymbol{S}$，使 $\boldsymbol{A}=\boldsymbol{S}^3$.

证 矩阵 $\boldsymbol{A}$ 是一个对称矩阵，则 $\boldsymbol{A}$ 酉相似于一个对角矩阵，即

$$\boldsymbol{A}=\boldsymbol{U}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\boldsymbol{U}^{\mathrm{H}}.$$

令 $\boldsymbol{D}=\mathrm{diag}(\lambda_1^{\frac{1}{3}},\lambda_2^{\frac{1}{3}},\cdots,\lambda_n^{\frac{1}{3}})$，则 $\boldsymbol{D}^3=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$. 又由 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{D}^3\boldsymbol{U}^{\mathrm{H}}=(\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{\mathrm{H}})(\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{\mathrm{H}})(\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{\mathrm{H}})$，令 $\boldsymbol{S}=\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{\mathrm{H}}$，则 $\boldsymbol{A}=\boldsymbol{S}^3$.

6. 证明：一个正规矩阵若是三角矩阵，则它一定是对角矩阵.

证 由 $\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$ 为正规矩阵，必有 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 成立，又

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix} \sum\limits_{j=1}^{n}a_{1j}^2 & * & \cdots & * \\ * & \sum\limits_{j=2}^{n}a_{2j}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ * & \cdots & * & a_{nn}^2 \end{pmatrix},$$

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix} a_{11}^2 & * & \cdots & * \\ * & \sum\limits_{i=1}^{2}a_{i2}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ * & \cdots & * & \sum\limits_{i=1}^{n}a_{in}^2 \end{pmatrix},$$

推得 $a_{ij}=0$，$\forall i\neq j$，$i,j\in\mathbf{N}$.

7. 证明：正规矩阵 $\boldsymbol{A}$ 是幂零阵（$\boldsymbol{A}^2=\boldsymbol{0}$）的充要条件是 $\boldsymbol{A}=\boldsymbol{0}$.

证 充分性. 若 $\boldsymbol{A}=\boldsymbol{0}$，则结论显然.

必要性. 若 $\boldsymbol{A}^2=\boldsymbol{0}$，由题设矩阵 $\boldsymbol{A}$ 是正规矩阵，则 $\boldsymbol{A}$ 酉相似于一个对角矩阵，即

$$\boldsymbol{A}=\boldsymbol{U}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\boldsymbol{U}^{\mathrm{H}},$$

$$\boldsymbol{A}^2=\boldsymbol{U}\mathrm{diag}(\lambda_1^2,\lambda_2^2,\cdots,\lambda_n^2)\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{0},$$

即

$$\operatorname{diag}(\lambda_1^2,\lambda_2^2,\cdots,\lambda_n^2)=\mathbf{0}.$$

所以

$$\lambda_1=\lambda_2=\cdots=\lambda_n=0,$$

即 $\boldsymbol{A}=\mathbf{0}$. 结论成立.

8. 求矩阵 $\boldsymbol{A}=\begin{pmatrix}3&-2&-4\\-2&6&-2\\-4&-2&3\end{pmatrix}$ 的谱分解式,并给出 $\boldsymbol{A}^n$ 的表达式.

解 先求矩阵 $\boldsymbol{A}$ 的特征值. 由 $\det(\lambda\boldsymbol{E}-\boldsymbol{A})=(\lambda+2)(\lambda-7)^2$,得矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1=-2,\lambda_{2,3}=7$,对应的特征向量分别为

$$\boldsymbol{u}_1=\left(\frac{2}{3},-\frac{1}{3},-\frac{2}{3}\right)^{\mathrm{T}},$$

$$\boldsymbol{u}_2=(-0.5758,0.7982,0.1767)^{\mathrm{T}},$$

$$\boldsymbol{u}_3=(0.4732,0.5017,-0.7241)^{\mathrm{T}},$$

令 $\boldsymbol{U}=(\boldsymbol{u}_1\quad\boldsymbol{u}_2\quad\boldsymbol{u}_3)\boldsymbol{U}^{\mathrm{H}}=\begin{pmatrix}\boldsymbol{v}_1^{\mathrm{T}}\\\boldsymbol{v}_2^{\mathrm{T}}\\\boldsymbol{v}_3^{\mathrm{T}}\end{pmatrix}$,$\boldsymbol{A}_1=\boldsymbol{u}_1\boldsymbol{v}_1^{\mathrm{T}}$,$\boldsymbol{A}_2=\boldsymbol{u}_2\boldsymbol{v}_2^{\mathrm{T}}$,$\boldsymbol{A}_3=\boldsymbol{u}_3\boldsymbol{v}_3^{\mathrm{T}}$,则 $\boldsymbol{A}$ 的谱分解为 $\boldsymbol{A}=-2\boldsymbol{A}_1+7\boldsymbol{A}_2+7\boldsymbol{A}_3$. 所以 $\boldsymbol{A}^n=(-2)^n\boldsymbol{A}_1+7^n\boldsymbol{A}_2+7^n\boldsymbol{A}_3$.

9. 证明:如果一个实对称矩阵 $\boldsymbol{A}$ 的主对角元都大于零,则 $\boldsymbol{A}$ 至少有一个正的特征值.

证 设矩阵 $\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{n\times n}$,由于矩阵 $\boldsymbol{A}$ 是对称矩阵,则其特征值 $\lambda_i(i=1,2,\cdots,n)$都是实数,根据矩阵特征值与矩阵迹的关系,可得

$$0<a_{11}+a_{22}+\cdots+a_{nn}=\lambda_1+\lambda_2+\cdots+\lambda_n,$$

所以矩阵 $\boldsymbol{A}$ 至少有一个正的特征值.

10. 求下列矩阵的最大秩分解式.

(1) $\boldsymbol{A}=\begin{pmatrix}1&-1&1&1\\-1&1&-1&-1\\1&1&-1&-1\\1&-1&1&1\end{pmatrix}$, (2) $\boldsymbol{A}=\begin{pmatrix}1&2&3&6\\2&4&6&12\\1&2&3&6\\2&4&6&12\end{pmatrix}$.

解 (1) 对矩阵 $\boldsymbol{A}$ 实行行初等变换,得

$$\boldsymbol{A}\to\begin{pmatrix}1&0&0&0\\0&1&-1&-1\\0&0&0&0\\0&0&0&0\end{pmatrix}.$$

取

$$B=\begin{pmatrix}1&-1\\-1&1\\1&1\\1&-1\end{pmatrix},\quad D=\begin{pmatrix}1&0&0&0\\0&1&-1&-1\end{pmatrix},$$

则 $A=BD$ 就是矩阵 A 的最大秩分解.

(2) 同理对矩阵 A 进行行初等变换,可得

$$A\to\begin{pmatrix}1&2&3&6\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}.$$

取

$$B=\begin{pmatrix}1\\2\\1\\2\end{pmatrix},\quad D=(1\quad 2\quad 3\quad 6),$$

则 $A=BD$ 就是矩阵 A 的最大秩分解.

11. 设矩阵为

$$A=\begin{pmatrix}-1&\mathrm{i}&0\\-\mathrm{i}&0&-\mathrm{i}\\0&\mathrm{i}&-1\end{pmatrix},\quad B=\begin{pmatrix}0&\mathrm{i}&1\\-\mathrm{i}&0&0\\1&0&0\end{pmatrix}.$$

试问:A 与 B 是正规矩阵吗?若是,通过酉变换把它们化成相似对角矩阵.

解 由于

$$A^{\mathrm{H}}A=AA^{\mathrm{H}}=\begin{pmatrix}2&-\mathrm{i}&1\\\mathrm{i}&2&\mathrm{i}\\1&-\mathrm{i}&-1\end{pmatrix},$$

所以矩阵 A 是正规矩阵.

矩阵 A 的特征值为 $-2,-1,1$,其对应的特征向量构成的矩阵为

$$U=\begin{bmatrix}-0.5774&-0.7071&0.4082\\-0.5774\mathrm{i}&0&-0.8165\mathrm{i}\\-0.5774&0.7071&0.4082\end{bmatrix},$$

则酉变换为

$$U^{\mathrm{H}}AU=\mathrm{diag}(-2,-1,1).$$

12. 设矩阵 A 的最大秩分解为 $A=BC$,证明:

$$Ax=0\Leftrightarrow Cx=0$$

证 充分性显然.

必要性用反证法. 如果存在向量 $\boldsymbol{x}$，使得 $\boldsymbol{Ax}=\boldsymbol{0}$，但 $\boldsymbol{Cx}\neq\boldsymbol{0}$，令 $\boldsymbol{y}=\boldsymbol{Cx}$，则 $\boldsymbol{y}\neq\boldsymbol{0}$. 由于 $\boldsymbol{A}=\boldsymbol{BC}$ 是矩阵 $\boldsymbol{A}$ 的最大秩分解，则矩阵 $\boldsymbol{B}$ 的列向量是线性无关的，如果 $\boldsymbol{y}\neq\boldsymbol{0}$，则 $\boldsymbol{By}\neq\boldsymbol{0}$，从而 $\boldsymbol{BCx}\neq\boldsymbol{0}$，与题设矛盾，所以 $\boldsymbol{Cx}=\boldsymbol{0}$.

13. 若 $\boldsymbol{A}=(a_{ij})$ 是 n 阶正定矩阵，则有

$$\det(\boldsymbol{A})\leqslant a_{11}a_{22}\cdots a_{nn},$$

当且仅当 $\boldsymbol{A}$ 为对角矩阵时等式才成立(这就是 Hadamard 不等式).

证 记 $d_i=a_{ii}^{-\frac{1}{2}}$，$i=1,2,\cdots,n$，并设 $\boldsymbol{D}=\mathrm{diag}(d_1,d_2,\cdots,d_n)$. 于是证明 $\det(\boldsymbol{A})\leqslant a_{11}a_{22}\cdots a_{nn}$ 等价于证明 $\det(\boldsymbol{DAD})\leqslant 1$. 由 $\boldsymbol{DAD}$ 的构造知，其对角元素都为 1. 令 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{DAD}$ 的全部特征值，则有

$$\det(\boldsymbol{DAD})=\prod_{i=1}^{n}\lambda_i\leqslant\left(\frac{1}{n}\sum_{i=1}^{n}\lambda_i\right)^n=\left(\frac{1}{n}\mathrm{tr}(\boldsymbol{DAD})\right)^n=1.$$

上式利用的算术-几何均值不等式的等号成立当且仅当所有的 $\lambda_i=1$. 因为 $\boldsymbol{A}$ 是正定矩阵，可以对角化，所以这种情况出现当且仅当 $\boldsymbol{A}$ 为对角矩阵.

14. 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$ 均为正定的 Hermite 矩阵，则 $\boldsymbol{AB}$ 为正定的 Hermite 矩阵的充要条件是 $\boldsymbol{AB}=\boldsymbol{BA}$.

证 必要性. 设 $\boldsymbol{AB}$ 为正定的 Hermite 矩阵，根据定义有 $(\boldsymbol{AB})^{\mathrm{H}}=\boldsymbol{AB}$，即 $\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{AB}$，同时有 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$，$\boldsymbol{B}^{\mathrm{H}}=\boldsymbol{B}$，所以 $\boldsymbol{AB}=\boldsymbol{BA}$.

充分性. 设 $\boldsymbol{AB}=\boldsymbol{BA}$，则 $(\boldsymbol{AB})^{\mathrm{H}}=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{BA}=\boldsymbol{AB}$，于是矩阵 $\boldsymbol{AB}$ 是 Hermite 矩阵. 由于矩阵 $\boldsymbol{A}$ 是正定 Hermite 矩阵，故存在一个正定的 Hermite 矩阵 $\boldsymbol{S}$，使得 $\boldsymbol{A}=\boldsymbol{S}^2$，则有 $\boldsymbol{AB}=\boldsymbol{S}^2\boldsymbol{B}$. 对矩阵 $\boldsymbol{AB}$ 施行相似变换得 $\boldsymbol{S}^{-1}(\boldsymbol{AB})\boldsymbol{S}=\boldsymbol{SBS}=\boldsymbol{S}^{\mathrm{H}}\boldsymbol{BS}$，则矩阵 $\boldsymbol{AB}$ 与矩阵 $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{BS}$ 有相同的特征值，且 $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{BS}$ 是 Hermite 矩阵.

对 $\forall\boldsymbol{x}\neq\boldsymbol{0}$，可得 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{S}^{\mathrm{H}}\boldsymbol{BSx}=(\boldsymbol{Sx})^{\mathrm{H}}\boldsymbol{B}(\boldsymbol{Sx})>0$，即 $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{BS}$ 是正定的 Hermite 矩阵，所以其所有的特征值为正，从而矩阵 $\boldsymbol{AB}$ 所有的特征值为正，即矩阵 $\boldsymbol{AB}$ 为正定的 Hermite 矩阵.

第四章 特征值的估计与摄动

在工程技术中，涉及某个振动系统的长期稳定性微分方程的求解问题时，往往需要证明一个矩阵的所有特征值都位于左半平面内；在统计学或数值分析中，往往需要证明矩阵是正定的；在线性代数方程组迭代求解的收敛性分析中，需要估计一个矩阵的谱半径是否小于 1，从而判定迭代法是否收敛以及收敛的速度快慢等，这些都需要矩阵特征值的估计.

从实际问题得到的矩阵 $\boldsymbol{A}$ 的元素往往带有一定的误差，由误差得到的摄动矩阵有时候对问题的影响是微乎其微的，但有时候却导致矩阵是病态的，结果严重失真，从而需要对矩阵特征值做摄动分析.

本章针对以上两个问题展开讨论.

一、基本概念

定义 1 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$，称

$$S_i=\{z\in\mathbf{C}: |z-a_{ii}|\leqslant R_i\}$$

为矩阵 $\boldsymbol{A}$ 在复平面上的第 i 个 Gerschgorin 圆（盖尔圆），其中

$$R_i=R_i(\boldsymbol{A})=\sum_{j\neq i}|a_{ij}|, i=1,2,\cdots,n$$

称为 S_i 的半径.

定义 2 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$，若

$$|a_{ii}|\geqslant R_i(\boldsymbol{A}),\quad i=1,2,\cdots,n,$$

则称 $\boldsymbol{A}$ 为行对角占优矩阵；若 $\boldsymbol{A}^{\mathrm{T}}$ 为行对角占优矩阵，则称 $\boldsymbol{A}$ 为列对角占优矩阵；若 $\boldsymbol{A}$ 为行列对角占优矩阵. 若式中均为严格不等式，则称 $\boldsymbol{A}$ 为相应的严格对角占优矩阵.

定义 3 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$ 为 Hermite 矩阵，$x\in\boldsymbol{C}^n$，称

$$R(\boldsymbol{x})=\frac{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}},\quad \boldsymbol{x}\neq\boldsymbol{0}$$

为 $\boldsymbol{A}$ 的 Rayleigh 商.

Hermite 矩阵有很多好的性质，在实际应用中经常遇到.

二、主要结论

定理 1 (Schur)设 $A=(a_{ij})\in \mathbf{C}^{n\times n}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$,则

$$\sum_{i=1}^{n}|\lambda_i|^2\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2=\|A\|_F^2,$$

且等号成立当且仅当 A 为正规矩阵.

Schur 不等式在理论上极为重要，很多学者在此基础之上给出了更多的新结果.

设 $A=(a_{ij})\in \mathbf{C}^{n\times n}$，记

$$B=(b_{ij})=(A+A^{\mathrm{H}})/2,\quad C=(c_{ij})=(A-A^{\mathrm{H}})/2,$$

显然 B,C 分别为 Hermite 和反 Hermite 矩阵.

以下若无特别说明，总假设 A,B,C 的特征值集合分别为 $\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$，$\{\mu_1,\mu_2,\cdots,\mu_n\}$，$\{\mathrm{i}\gamma_1,\mathrm{i}\gamma_2,\cdots,\mathrm{i}\gamma_n\}$，且满足

$$|\lambda_1|\geqslant|\lambda_2|\geqslant\cdots\geqslant|\lambda_n|,$$
$$\mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_n,$$
$$\gamma_1\geqslant\gamma_2\geqslant\cdots\geqslant\gamma_n.$$

定理 2 (Hirsch)设 $A=(a_{ij})\in \mathbf{C}^{n\times n}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则

$$|\lambda_i|\leqslant n\max_{i,j}|a_{ij}|,$$
$$|\mathrm{Re}\,(\lambda_i)|\leqslant n\max_{i,j}|b_{ij}|,$$
$$|\mathrm{Im}\,\lambda_i|\leqslant n\max_{i,j}|c_{ij}|.$$

定理 3 (Bendixson)设 A 为 n 阶实矩阵，则 A 的任一特征值 λ_i 满足

$$|\mathrm{Im}\,\lambda_i|\leqslant\sqrt{\frac{n(n-1)}{2}}\max_{i,j}|c_{ij}|.$$

定理 4 设 $A=(a_{ij})\in \mathbf{C}^{n\times n}$，$B$，$C$，$\lambda_i$，$\mu_i$，$\gamma_i$ 定义同上，则有

$$\mu_n\leqslant\mathrm{Re}\,(\lambda_i)\leqslant\mu_1,$$
$$\gamma_n\leqslant\mathrm{Im}\,(\lambda_i)\leqslant\gamma_1.$$

定理 5 (Browne)设 $A=(a_{ij})\in \mathbf{C}^{n\times n}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，奇异值为 $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_n$，则有

$$\sigma_n\leqslant|\lambda_i|\leqslant\sigma_1,\quad i=1,2,\cdots n.$$

定理 6 (Hadamard 不等式)设 $A=(a_{ij})\in \mathbf{C}^{n\times n}$，则有

$$\prod_{i=1}^{n}|\lambda_i(A)|=|\det A|\leqslant\Big[\prod_{j=1}^{n}\sum_{i=1}^{n}|a_{ij}|^2\Big]^{\frac{1}{2}},$$

且等号成立的充要条件或是 A 的某一列全为零元，或 $(a^i,a^j)=0(i\neq j)$，这里 a^i 表示 A 的第 i 列.

下面给出利用矩阵的元素来确定 $\boldsymbol{A}$ 的特征值分布区域的几个定理,其中包括著名的 Gerschgorin 圆盘定理.

定理 7 (Gerschgorin 圆盘定理 1)设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$,则 $\boldsymbol{A}$ 的任一特征值满足

$$\lambda_i \in S = \bigcup_{j=1}^{n} S_j, \quad i = 1,2,\cdots,n.$$

定理 7 只说明了矩阵 $\boldsymbol{A}$ 的特征值均在其余全部盖尔圆的并集中,而没有明确在哪个连通部分中有几个特征值,下面的定理改进了这一点.

定理 8 (Gerschgorin 圆盘定理 2)设 n 阶方阵 $\boldsymbol{A}$ 的 n 个盖尔圆盘中有 k 个圆盘的并形成一个连通区域,且余下的 $n-k$ 个圆盘都不相交,则在这个区域中恰好有 $\boldsymbol{A}$ 的 k 个特征值.

注 定理 7 中的区域 $\boldsymbol{S}$ 常称为 $\boldsymbol{A}$ 的行 Gerschgorin 区域,因为 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 有相同的特征值,所以将定理 7 和定理 8 应用于 $\boldsymbol{A}^{\mathrm{T}}$,则得到 $\boldsymbol{A}$ 的列的 Gerschgorin 圆盘定理. 记为

$$C_j(\boldsymbol{A}) = \sum_{\substack{i=1\\ i\neq j}}^{n} |a_{ij}|.$$

由定理 7 和定理 8,可以得到几个重要推论.

推论 1 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$,则 $\boldsymbol{A}$ 的所有特征值均包含在 $\boldsymbol{A}$ 的列盖尔圆区域

$$G = G_1 \cup G_2 \cup \cdots \cup G_n$$

中,其中 G_j 为 $\boldsymbol{A}$ 的列盖尔圆盘,即

$$G_j = \{z \in C\colon |z - a_{jj}| \leqslant C_j(\boldsymbol{A})\},$$

此外,若这些圆盘中的 k 个圆盘之并形成一个连通区域,且与所有其余 $n-k$ 个圆盘不相交,则在这个区域内恰好有 $\boldsymbol{A}$ 的 k 个特征值.

推论 2 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$,则 $\boldsymbol{A}$ 的全部特征值均落在平面区域

$$T = \left(\bigcup_{i=1}^{n} S_i\right) \cap \left(\bigcup_{j=1}^{n} G_j\right)$$

之中.

注 有两个或两个以上的盖尔圆构成的连通部分,可能在一个盖尔圆中有两个或两个以上的特征值,而在另外的一个或几个盖尔圆中没有特征值.

推论 3 设 n 阶矩阵 $\boldsymbol{A}$ 的 n 个圆盘两两互不相交,则 $\boldsymbol{A}$ 相似于对角矩阵.

推论 4 设 n 阶实矩阵 $\boldsymbol{A}$ 的 n 个圆盘两两互不相交,则 $\boldsymbol{A}$ 的特征值全为实数.

定理 9 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$,$p_1,p_2,\cdots,p_n$ 均大于零,则 $\boldsymbol{A}$ 的任一特征值

$$\lambda_i \in \left(\bigcup_{i=1}^{n} Q_i\right) \cap \left(\bigcup_{j=1}^{n} P_j\right), \quad i = 1,2,\cdots,n.$$

其中

$$r_i = \frac{1}{p_i}\sum_{\substack{j=1\\j\neq i}}^{n} |a_{ij}|p_j,\quad Q_i = \{z\in \mathbf{C}: |z-a_{ii}|\leqslant r_i\},\quad i=1,2,\cdots,n.$$

$$t_j = p_j\sum_{\substack{i=1\\i\neq j}}^{n} \frac{|a_{ij}|}{p_j},\quad P_j = \{z\in \mathbf{C}: |z-a_{ii}|\leqslant t_j\},\quad j=1,2,\cdots,n.$$

定理 10 设 $\mathbf{A}=(a_{ij})\in\mathbf{C}^{n\times n}$为行(或列)严格对角占优的，则

(1) $\mathbf{A}$ 为可逆矩阵，且 $\lambda_i\in\bigcup_{i=1}^{n} S_i$，这里 $S_i=\{z\in\mathbf{C}:|z-a_{ii}|<|a_{ii}|\}$；

(2) 若 $\mathbf{A}$ 的所有主对角元均为正数，则 $\mathbf{A}$ 的所有特征值都有正实部；

(3) 若 $\mathbf{A}$ 为 Hermite 阵，且 $\mathbf{A}$ 的所有主对角元都是正数，则 $\mathbf{A}$ 的所有特征值均为正数.

定理 11 设 $\mathbf{A}=(a_{ij})\in\mathbf{C}^{n\times n}$，则谱半径 $r(\mathbf{A})$满足

(1) $r(\mathbf{A})\leqslant\min\left\{\max\limits_{i}\sum\limits_{j=1}^{n}|a_{ij}|,\max\limits_{j}\sum\limits_{i=1}^{n}|a_{ij}|\right\}$；

(2) $r(\mathbf{A})\leqslant\min\limits_{p_1,p_2,\cdots,p_n\geqslant 0}\max\limits_{1\leqslant i\leqslant n}\frac{1}{p_i}\sum\limits_{j=1}^{n}p_j|a_{ij}|,r(\mathbf{A})\leqslant\min\limits_{p_1,p_2,\cdots,p_n\geqslant 0}\max\limits_{1\leqslant i\leqslant n}\frac{1}{p_j}\sum\limits_{i=1}^{n}p_i|a_{ij}|$.

定理 12 (Ostrowski)设 $\mathbf{A}=(a_{ij})\in\mathbf{C}^{n\times n}$，$\alpha\in[0,1]$为给定的数，则 $\mathbf{A}$ 的所有特征值位于 n 个圆盘的并集

$$\bigcup_{i=1}^{n}\{z\in\mathbf{C}: |z-a_{ii}|\leqslant R_i^{\alpha}C_i^{1-\alpha}\}$$

中，其中

$$R_i=\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|,\quad C_i=\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ji}|.$$

定理 13 (Brauer)设 $\mathbf{A}=(a_{ij})\in\mathbf{C}^{n\times n}$，则 $\mathbf{A}$ 的所有特征值位于$\frac{n(n-1)}{2}$个 Cassini 卵形的并集

$$\bigcup_{i=1}^{n}\{z\in\mathbf{C}: |z-a_{ii}||z-a_{jj}|\leqslant R_iR_j\}$$

之中.

定理 14 设不可约矩阵 $\mathbf{A}=(a_{ij})\in\mathbf{C}^{n\times n}$的一个特征值 λ 在其 n 个盖尔圆$\{|z-a_{ii}|\leqslant R_i\}(i=1,2,\cdots,n)$的并集的边界上，则所有的 n 个圆周

$$|z-a_{ii}|=R_i,\quad i=1,2,\cdots,n$$

都通过点 λ.

定理 15 若 $\mathbf{A}=(a_{ij})\in\mathbf{C}^{n\times n}$不可约，且存在 i_0 使得 $\sum\limits_{j=1}^{n}|a_{i_0 j}|<\|\mathbf{A}\|_\infty$，则有 $r(\mathbf{A})<$

$\| \boldsymbol{A} \|_{\infty}$.

等距的差分格式方程求解调和方程的第一类边值问题及有限元法求解某些结构问题时所产生的刚度矩阵都归结为 Hermite 矩阵特征值问题的研究. 下面给出 Hermite 矩阵特征值的一些特性.

定理 16 (Rayleigh-Ritz) 设 $\boldsymbol{A}=(a_{ij})\in \mathbf{C}^{n\times n}$ 为 Hermite 矩阵，则

$$\lambda_n \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} \leqslant \boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x} \leqslant \lambda_1 \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}, \quad \forall \boldsymbol{x} \in \mathbf{C}^n,$$

$$\lambda_{\max} = \lambda_1 = \max_{x\neq 0} R(\boldsymbol{x}) = \max_{x^{\mathrm{H}}x=1} \boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x},$$

$$\lambda_{\min} = \lambda_n = \max_{x\neq 0} R(\boldsymbol{x}) = \max_{x^{\mathrm{H}}x=1} \boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}.$$

注 1. 任意 Hermite 矩阵 $\boldsymbol{A}$ 均可以对角化，且与对角矩阵酉相似，即

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^{\mathrm{H}},$$

其中 $\boldsymbol{U}$ 为酉矩阵，$\boldsymbol{\Sigma}$ 为对角线元素为实数的对角矩阵. 即 Hermite 矩阵的特征值是实数，且特征向量相互正交.

2. Hermite 矩阵也是一类重要的正规矩阵. 正规矩阵还包括斜 Hermite 矩阵和酉矩阵等几类特殊矩阵，下面给出特殊正规矩阵与一般正规矩阵的特征值的情况比较.

Hermite 矩阵$\Leftrightarrow$特征值均为实数；斜 Hermite 矩阵$\Leftrightarrow$特征值均在虚轴上；

酉矩阵$\Leftrightarrow$特征值的复数模均等于 1；其他正规矩阵$\Leftrightarrow$特征值无特殊要求.

定理 17 (Courant-Fischer) 设 $\boldsymbol{A}=(a_{ij})\in \mathbf{C}^{n\times n}$ 为 Hermite 矩阵，其特征值为 $\lambda_n \leqslant \lambda_{n-1} \leqslant \cdots \leqslant \lambda_1$，$k$ 为给定的整数，$1\leqslant k\leqslant n$，则

$$\min_{\omega_1,\omega_2,\cdots,\omega_{n-k}\in \mathbf{C}^n} \ \max_{\substack{x\neq 0,\, x\in \mathbf{C}^n \\ x\perp \omega_1,\omega_2,\cdots,\omega_{n-k}}} R(\boldsymbol{x}) = \lambda_k,$$

$$\max_{\omega_1,\omega_2,\cdots,\omega_{k-1}\in \mathbf{C}^n} \ \min_{\substack{x\neq 0,\, x\in \mathbf{C}^n \\ x\perp \omega_1,\omega_2,\cdots,\omega_{k-1}}} R(\boldsymbol{x}) = \lambda_k.$$

Courant-Fischer 定理有许多重要的应用，下面的 Weyl 定理，它讨论了 $\boldsymbol{A}+\boldsymbol{B}$ 的特征值与 $\boldsymbol{A}$ 的特征值的比较问题，这里假设 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$.

定理 18 (Weyl) 设 $\boldsymbol{A},\boldsymbol{B}\in \mathbf{C}^{n\times n}$ 为 Hermite 矩阵，则 $\forall k=1,2,\cdots,n$，有

$$\lambda_k(\boldsymbol{A}) + \lambda_n(\boldsymbol{B}) \leqslant \lambda_k(\boldsymbol{A}+\boldsymbol{B}) \leqslant \lambda_k(\boldsymbol{A}) + \lambda_1(\boldsymbol{B}).$$

定理 19 设 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\in \mathbf{C}^{n\times n}$，$\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，$\boldsymbol{\delta}\in \mathbf{C}^{n\times n}$，$\boldsymbol{A}+\boldsymbol{\delta}$ 的特征值为 $\mu_1,\mu_2,\cdots,\mu_n$，则对任一 μ_j，存在 λ_i，使得

$$|\lambda_i - \mu_j| \leqslant \| \boldsymbol{P}^{-1}\boldsymbol{\delta}\boldsymbol{P} \|_{\infty},$$

此外，若是一个重数 m 的特征值，且圆盘 $S_i=\{z\in \mathbf{C}: |z-\lambda_i| \leqslant \| \boldsymbol{P}^{-1}\boldsymbol{\delta}\boldsymbol{P} \|_{\infty}\}$ 和圆盘 $S_k=\{z\in \mathbf{C}: |z-\lambda_k| \leqslant \| \boldsymbol{P}^{-1}\boldsymbol{\delta}\boldsymbol{P} \|_{\infty}\}$ $(\lambda_k\neq \lambda_i)$ 不相交，则 S_i 正好包含着 $\boldsymbol{A}+\boldsymbol{\delta}$ 的 m 个特征值.

三、例题解答

例 1 证明矩阵 $A=\begin{pmatrix} 2 & \frac{2}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & 4 & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \cdots & 2n \end{pmatrix}$能够相似于对角矩阵，且 A 的特征值都是实数.

证 因为 A 的 n 个盖尔圆为 $|z-2|\leqslant 1$，$|z-2k|\leqslant\frac{n-1}{n}$，$k=2,\cdots,n$. 它们都是孤立的，所以 A 有 n 个不同的特征值，从而相似于对角矩阵. 又因为 $G_k(k=1,2,\cdots,n)$关于实轴对称，且 A 为实矩阵，所以由定理 8 推论 4 知，G_k 中的特征值都是实数.

例 2 用 Gerschgorin 圆盘定理证明：矩阵

$$A=\begin{pmatrix} 2 & \frac{1}{2} & \frac{1}{2^2} & \frac{1}{2^3} & \frac{1}{2^4} \\ \frac{2}{3} & 4 & \frac{2}{3^2} & \frac{2}{3^3} & \frac{2}{3^4} \\ \frac{3}{4} & \frac{3}{4^2} & 6 & \frac{3}{4^3} & \frac{3}{4^4} \\ \frac{4}{5} & \frac{4}{5^2} & \frac{4}{5^3} & 8 & \frac{4}{5^4} \\ \frac{5}{6} & \frac{5}{6^2} & \frac{5}{6^3} & \frac{5}{6^4} & 10 \end{pmatrix}$$

能够相似于对角矩阵且 A 的特征值都是正实数.

证 A 的 5 个盖尔圆盘为

$$G_i=\left\{z \,\middle|\, |z-2i|\leqslant 1-\frac{1}{(i+1)^4}\right\},\quad i=1,2,3,4,5$$

它们都是孤立的，从而矩阵有 5 个互异特征值，所以矩阵能够相似于对角矩阵，再由 G_i 关于实轴对称且都在 y 坐标轴右边，以及实矩阵的复数特征值成对共扼出现的性质知，G_i 中的特征值必为正实数，所以 A 的特征值都是正实数.

例 3 试隔离矩阵 $A=\begin{pmatrix} 20 & 5 & 0.8 \\ 4 & 10 & 1 \\ 1 & 2 & 10\mathrm{i} \end{pmatrix}$的特征值.

解 A 的 3 个盖尔圆为

$G_1:|z-20|\leqslant 5.8$，$G_2:|z-10|\leqslant 5$，$G_3:|z-10\mathrm{i}|\leqslant 3$，$G_1$ 与 G_2 相交；而 G_3 孤立，其中恰好有 $\boldsymbol{A}$ 的一个特征值，记作 λ_3. 如选取

$$\boldsymbol{D}=\operatorname{diag}(1,1,2),$$

则

$$\boldsymbol{B}=\boldsymbol{DAD}^{-1}=\begin{pmatrix}20 & 5 & 0.4\\ 4 & 10 & 0.5\\ 2 & 4 & 10\mathrm{i}\end{pmatrix}$$

的 3 个盖尔圆为

$$G_1':|z-20|\leqslant 5.4,\quad G_2':|z-10|\leqslant 4.5,\quad G_3':|z-10\mathrm{i}|\leqslant 6,$$

易知，这是 3 个孤立的盖尔圆，每个盖尔圆中恰好有 $\boldsymbol{B}$(也是 $\boldsymbol{A}$)一个特征值，注意到，G_3'中的特征值就是 G_3 中的特征值 λ_3，所以 $\boldsymbol{A}$ 的 3 个特征值分别位于 G_1'，G_2'和 G_3'之中.

注　对于矩阵 $\boldsymbol{A}$，选取适当正数 $p_1,p_2,\cdots,p_n$，可以获得只含 $\boldsymbol{A}$ 的一个特征值的孤立盖尔圆. 选取 $p_1,p_2,\cdots,p_n$ 的一般方法是：观察 $\boldsymbol{A}$ 的 n 个盖尔圆，使得第 i 个盖尔圆 G_i'的半径大(或小一些)，取 $p_i>1$(或 $p_i<1$)，而取 $p_1=\cdots=p_{i-1}=p_{i+1}=\cdots=p_n=1$. 此时，$\boldsymbol{B}=\boldsymbol{DAD}^{-1}$的第 i 个盖尔圆 G_i'的半径比 G_i 的半径大(或小)，而其余盖尔圆半径相对变小(或变大). 但是，这种隔离矩阵特征值的方法还不能用于任意的具有互异特征值的矩阵，比如主对角线上有相同元素的矩阵.

例 4　若 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$不可约，且存在 i_0 使得 $\sum\limits_{j=1}^{n}|a_{i_0j}|<\|\boldsymbol{A}\|_\infty$，则有 $r(\boldsymbol{A})<\|\boldsymbol{A}\|_\infty$.

证　因为 $r(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|_\infty$，所以只需证明 $r(\boldsymbol{A})\neq\|\boldsymbol{A}\|_\infty$即可. 对于任意 $z\in\bigcup\limits_{i=1}^{n}G_i$，存在 i_1 使 $|z-a_{i_1i_1}|\leqslant R_{i_1}$，由三角不等式可得

$$|z|\leqslant R_{i_1}+|a_{i_1i_1}|=\sum_{j=1}^{n}|a_{i_1j}|\leqslant\|\boldsymbol{A}\|_\infty,$$

即 $z\in\{z\,|\,|z|\leqslant\|\boldsymbol{A}\|_\infty\}=\overline{S}(0;\|\boldsymbol{A}\|_\infty)$，因此并集 $\bigcup\limits_{i=1}^{n}G_i\subset\overline{S}(0;\|\boldsymbol{A}\|_\infty)$.

假设在定理条件下，$r(\boldsymbol{A})=\|\boldsymbol{A}\|_\infty$，则 $\boldsymbol{A}$ 至少有一个特征值 λ_0 满足

$$|\lambda_0|=r(\boldsymbol{A})=\|\boldsymbol{A}\|_\infty,$$

所以 λ_0 在 $\overline{S}(0;\|\boldsymbol{A}\|_\infty)$的边界上. 从而 λ_0 亦在 $\boldsymbol{A}$ 的 n 个盖尔圆并集的边界上. 由定理 14 知

$$|\lambda_0-a_{ii}|=R_i,\quad i=1,2,\cdots,n.$$

特别地有

$$|\lambda_0-a_{i_0i_0}|=R_{i_0},$$

于是

$$|\lambda_0| \leqslant R_{i_0} + |a_{i_0 i_0}| = \sum_{j=1}^{n} |a_{i_0 j}| < \|\boldsymbol{A}\|_\infty,$$

这与假设矛盾，故应有 $r(\boldsymbol{A}) \neq \|\boldsymbol{A}\|_\infty$.

四、习题解答

1. 证明：实对称矩阵 $\boldsymbol{A}$ 的所有特征值在区间$[a,b]$上的充要条件是对任何 $\lambda_0<a$，$\boldsymbol{A}-\lambda_0\boldsymbol{E}$ 是正定矩阵；而对任何 $\lambda_0<a$，$\boldsymbol{A}-\lambda_0\boldsymbol{E}$ 是负定矩阵.

证 因为 $\boldsymbol{A}$ 为实对称矩阵，所以存在正交矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{A} = \boldsymbol{Q}^{\mathrm{T}}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\boldsymbol{Q},$$

其中特征值 $\lambda_i \in [a,b]$.

$$\boldsymbol{A} - \lambda_0\boldsymbol{E} = \boldsymbol{Q}^{\mathrm{T}}\mathrm{diag}\{\lambda_1-\lambda_0,\lambda_2-\lambda_0,\cdots,\lambda_n-\lambda_0\}\boldsymbol{Q},$$

所以对于 $\forall\lambda_0<a,\lambda_i-\lambda_0>0$，可知 $\boldsymbol{A}$ 为正定矩阵；$\forall\lambda_0>b,\lambda_i-\lambda_0<0$，可知 $\boldsymbol{A}$ 为负定矩阵.

2. 设 $\boldsymbol{A},\boldsymbol{B}$ 都是实对称矩阵，$\boldsymbol{A}$ 的一切特征值在区间$[a,b]$上，$\boldsymbol{B}$ 的一切特征值在区间$[c,d]$上. 证明：$\boldsymbol{A}+\boldsymbol{B}$ 的特征值必在区间$[a+c,b+d]$上.

证 设 $\boldsymbol{A},\boldsymbol{B}$ 的特征值分别为

$$b \geqslant \lambda_1(\boldsymbol{A}) \geqslant \lambda_2(\boldsymbol{A}) \geqslant \cdots \geqslant \lambda_n(\boldsymbol{A}) \geqslant a, \quad d \geqslant \lambda_1(\boldsymbol{B}) \geqslant \lambda_2(\boldsymbol{B}) \geqslant \cdots \geqslant \lambda_n(\boldsymbol{B}) \geqslant c,$$

又因为 $\boldsymbol{A},\boldsymbol{B}$ 为实对称矩阵，所以 $\boldsymbol{A},\boldsymbol{B}$ 为 Hermite 矩阵，由定理 18 知，$\boldsymbol{A}+\boldsymbol{B}$ 的特征值为 $\lambda_k(\boldsymbol{A}+\boldsymbol{B})$，$\forall k=1,2,\cdots,n$. 于是有

$$\lambda_k(\boldsymbol{A}) + \lambda_n(\boldsymbol{B}) \leqslant \lambda_k(\boldsymbol{A}+\boldsymbol{B}) \leqslant \lambda_k(\boldsymbol{A}) + \lambda_1(\boldsymbol{B}), \quad k=1,2,\cdots,n.$$

即

$$\begin{aligned} a+c &\leqslant \lambda_k(\boldsymbol{A}) + c \leqslant \lambda_k(\boldsymbol{A}) + \lambda_n(\boldsymbol{B}) \leqslant \lambda_k(\boldsymbol{A}+\boldsymbol{B}) \\ &\leqslant \lambda_k(\boldsymbol{A}) + \lambda_1(\boldsymbol{B}) \leqslant \lambda_k(\boldsymbol{A}) + d \leqslant b+d, \quad k=1,2,\cdots,n. \end{aligned}$$

3. 设 $\boldsymbol{P}$ 是酉矩阵，$\boldsymbol{A}=\mathrm{diag}(a_1,a_2,\cdots,a_n)$，证明 $\boldsymbol{PA}$ 的特征值 μ 满足不等式

$$m \leqslant |\mu| \leqslant M,$$

其中 $m=\min\limits_i|a_i|$，$M=\max\limits_i|a_i|$.

证 因为 $\boldsymbol{P}$ 是酉矩阵，所以 $\boldsymbol{P}^{\mathrm{H}}\boldsymbol{P}=\boldsymbol{E}$，又因为 $(\boldsymbol{PA})^{\mathrm{H}}(\boldsymbol{PA})=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{PA}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$，所以由 Browne 定理知，$\boldsymbol{PA}$ 的特征值 μ 满足不等式

$$\min_i\{\sqrt{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}\} = \min_i\{\sqrt{(\boldsymbol{PA})^{\mathrm{H}}\boldsymbol{PA}}\} \leqslant |\mu| \leqslant \max_i\{\sqrt{(\boldsymbol{PA})^{\mathrm{H}}\boldsymbol{PA}}\} = \max_i\{\sqrt{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}\},$$

而

$$\min_i\{\sqrt{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}\} = \min_i|a_i| = m, \quad \max_i\{\sqrt{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}\} = \max_i|a_i| = M,$$

所以

$$m \leqslant |\mu| \leqslant M.$$

4. 用圆盘定理证明 $\boldsymbol{A}=\begin{pmatrix} 9 & 1 & -2 & 1 \\ 0 & 8 & 1 & 1 \\ -1 & 0 & 4 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ 至少有两个实特征值.

证 $\boldsymbol{A}$ 的 4 个盖尔圆为

$$G_1=\{z \mid |z-9| \leqslant 4\},\quad G_2=\{z \mid |z-8| \leqslant 2\},$$
$$G_3=\{z \mid |z-4| \leqslant 1\},\quad G_4=\{z \mid |z-1| \leqslant 1\},$$

它们构成的两个连通区域部分为 $S_1=G_1 \cup G_2 \cup G_3$，$S_2=G_4$，易知 S_1 与 S_2 都关于实轴对称，因为实矩阵的复特征值必成对共轭出现，所以 S_2 中含有 $\boldsymbol{A}$ 的一个特征值，而 S_1 中至少含有 $\boldsymbol{A}$ 的一个实特征值，因此 $\boldsymbol{A}$ 中至少有两个实特征值.

5. 估计矩阵

$$\boldsymbol{A}=\begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{2} & \mathrm{i} & 0 \\ 0 & -\frac{\mathrm{i}}{2} & 5 & \frac{\mathrm{i}}{2} \\ -1 & 0 & 0 & 5\mathrm{i} \end{pmatrix}$$

的特征值的分布范围.

解 $\boldsymbol{A}$ 的四个盖尔圆为

$$S_1: |z-1| \leqslant \frac{1}{2}+\frac{1}{2}+0=1;$$
$$S_2: \left|z-\frac{3}{2}\right| \leqslant \frac{1}{2}+|\mathrm{i}|=\frac{3}{2};$$
$$S_3: |z-5| \leqslant \left|-\frac{\mathrm{i}}{2}\right|+\left|\frac{\mathrm{i}}{2}\right|=1;$$
$$S_4: |z-5\mathrm{i}| \leqslant 1,$$

于是，由定理 7 知 $\boldsymbol{A}$ 的特征值均落在 $\bigcup\limits_{i=1}^{4} S_i$ 之中.

6. 用圆盘定理估计矩阵

$$\boldsymbol{A}=\begin{pmatrix} 7 & -16 & 8 \\ -16 & 7 & -8 \\ 8 & -8 & -5 \end{pmatrix}$$

的特征值和 $\boldsymbol{A}$ 的谱半径，然后选取一组正数 p_1, p_2, p_3 对 $\boldsymbol{A}$ 的特征值作更细的估计.

解 $\boldsymbol{A}$ 的 3 个特征值在它的两个盖尔圆 $|z-7| \leqslant 24$，$|z+5| \leqslant 16$ 的并集中，且 $r(\boldsymbol{A}) \leqslant 31$. 因为矩阵 $\boldsymbol{A}$ 有相同的主对角元素，所以，无法通过选取正数 p_1, p_2, p_3 给出更

精细的估计.

7. 证明 $\boldsymbol{A}=\begin{pmatrix}1/4 & 1/4 & 1/4 & 1/4\\ 1/5 & 2/5 & 1/5 & 1/5\\ 1/6 & 1/6 & 3/6 & 1/6\\ 1/7 & 1/7 & 1/7 & 3/7\end{pmatrix}$的谱半径满足 $r(\boldsymbol{A})<1$.

证 因为 $\boldsymbol{A}$ 不可约，且 $\sum_{j=1}^{4}|a_{4j}|=\frac{6}{7}<1=\|\boldsymbol{A}\|_\infty$，所以由定理 15 知 $r(\boldsymbol{A})<1$.

8. 证明 $\boldsymbol{A}=\begin{pmatrix}1/4 & 1/4 & 1/4 & 1/4\\ 1/5 & 2/5 & 1/5 & 1/5\\ 1/6 & 1/6 & 3/6 & 1/6\\ 1/7 & 1/7 & 1/7 & 4/7\end{pmatrix}$的谱半径为 $r(\boldsymbol{A})=1$.

证 因为 $r(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|_\infty=1$，且 $\det(1\boldsymbol{E}-\boldsymbol{A})=0$，所以 $r(\boldsymbol{A})=1$.

9. 举例说明：

(1) 在由两个盖尔圆构成的连通部分中，可以在每一个盖尔圆中恰有一个特征值.

(2) 不一定每个盖尔圆中必有一个特征值.

解 (1) 如 $\boldsymbol{A}=\begin{pmatrix}1 & 2\\ 2 & -1\end{pmatrix}$，故 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=\lambda^2-5=0$，$\lambda_{1,2}=\pm\sqrt{5}$.

(2) 如 $\boldsymbol{A}=\begin{pmatrix}1 & -0.8\\ 0.5 & 0\end{pmatrix}$，故 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=\lambda^2-\lambda+0.4=0$，$\lambda_{1,2}=\frac{1}{2}(1\pm\mathrm{i}\sqrt{0.6})$.

10. 应用 Ostrowski 定理(或推论)，证明

$$\boldsymbol{A}=\begin{pmatrix}6 & 5 & 1 & 2\\ 1 & 7 & 0 & 2\\ 0 & 4 & 7 & 5\\ 2 & 0 & 1 & 5\end{pmatrix}$$

的谱半径 $r(\boldsymbol{A})$小于 13.

证 因为

$$P_1T_1=14\times 9=126,\quad P_2T_2=10\times 16=160,$$
$$P_3T_3=16\times 9=144,\quad P_4T_4=8\times 14=112.$$

所以由定理 2 的推论 2 可得 $r(\boldsymbol{A})\leqslant\max_i(P_iT_i)^{\frac{1}{2}}=\sqrt{160}<13$.

11. 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$，满足 $|a_{ij}|>\sum_{j\neq i}|a_{ij}|\ (i=1,2,\cdots,n)$，则

(1) $\boldsymbol{A}$ 可逆；

(2) $|\det\boldsymbol{A}|\geqslant\prod_{i=1}^{n}(|a_{ii}|-\sum_{j\neq i}|a_{ij}|)$.

证 (1) 因为 $\boldsymbol{A}$ 为严格对角占优矩阵，由定理 4 知，$\boldsymbol{A}$ 可逆.

(2) 设 $d_i=|a_{ii}|-\sum\limits_{j\neq i}|a_{ij}|>0$，令

$$B=\begin{pmatrix}\frac{a_{11}}{d_1} & \frac{a_{12}}{d_1} & \cdots & \frac{a_{1n}}{d_1}\\ \frac{a_{21}}{d_2} & \frac{a_{22}}{d_2} & \cdots & \frac{a_{2n}}{d_2}\\ \vdots & \vdots & & \vdots\\ \frac{a_{n1}}{d_n} & \frac{a_{n2}}{d_n} & \cdots & \frac{a_{nn}}{d_n}\end{pmatrix},$$

则有 $|\det B|=\dfrac{|\det A|}{d_1d_2\cdots d_n}$. 所以 B 盖尔圆盘 $S_i=\left(\dfrac{a_{ii}}{d_i},\dfrac{\sum\limits_{j=1,j\neq i}^{n}a_{ij}}{d_i}\right)$，又 $\dfrac{|a_{ii}|-\sum\limits_{j\neq i}|a_{ij}|}{d_i}=1$，$\dfrac{a_{ii}}{d_i}>1$ 并且 S_i 与单位圆相外切，故矩阵 B 的特征值的模均大于 1，所以 $|\det B|\geqslant 1$，即 $\dfrac{|\det A|}{d_1d_2\cdots d_n}\geqslant 1$. 所以 $|\det A|\geqslant\prod\limits_{i=1}^{n}(|a_{ii}|-\sum\limits_{j\neq i}|a_{ij}|)$ 成立.

12. 若 $A=(a_{ij})\in C^{n\times n}$ 奇异，则存在某个 i_0，使 $|a_{i_0i_0}|\leqslant\sum\limits_{j\neq i_0}|a_{i_0j}|$.

证 (反证法)若对任意的 i，均有 $|a_{ii}|>\sum\limits_{j\neq i}|a_{ij}|$，则 A 为严格对角占优，即 A 非奇异，这与 $A=(a_{ij})\in C^{n\times n}$ 奇异矛盾.

13. 设 $A=(a_{ij})\in C^{n\times n}$ 可逆，λ 为特征值，则 $\|A^{-1}\|_2^{-1}\leqslant|\lambda|\leqslant\|A\|_2$.

证 设 A 的特征值为 λ，对应的特征向量为 $x\neq 0$，则 $Ax=\lambda x$. 因为 $\|\lambda x\|_2=\lambda\|x\|_2$，$\|Ax\|_2\leqslant\|A\|_2\|x\|_2$，所以 $\lambda\leqslant\|A\|_2$. 又因为 $A^{-1}x=\dfrac{x}{\lambda}$，同理可得 $\|A^{-1}\|_2^{-1}\leqslant|\lambda|$. 所以 $\|A^{-1}\|_2^{-1}\leqslant|\lambda|\leqslant\|A\|_2$ 成立.

14. 设 $R_A(x)$ 是 $A=A^H\in C^{n\times n}$ 的 Rayleigh 商，证明：(1) $R_A(\lambda x)=R_A(x)$，$\forall 0\neq\lambda\in C$，$0\neq x\in C^n$；(2) 存在 $0\neq x_i\in C^n(i=1,2,\cdots,n)$，使 $R_A(x_i)=\lambda_i(A)$.

证 (1) $\forall 0\neq\lambda\in C$，$0\neq x\in C^n$，有

$$R_A(\lambda x)=\frac{(\lambda x)^HA(\lambda x)}{(\lambda x)^H(\lambda x)}=\frac{|\lambda|^2x^HAx}{|\lambda|^2x^Hx}=\frac{x^HAx}{x^Hx}=R_A(x);$$

(2) 令 x_i 是矩阵 A 的特征值 $\lambda_i(A)$ 所对应的特征向量，即 $Ax_i=\lambda_i(A)x_i$，$x_i\neq 0$，则有

$$R_A(x_i)=\frac{x_i^HAx_i}{x_i^Hx_i}=\frac{\lambda_i(A)x_i^Hx_i}{x_i^Hx_i}=\lambda_i(A).$$

15. 设实对称阵 A 和 B 的特征值分别是

$$\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n \quad 和 \quad u_1\leqslant u_2\leqslant\cdots\leqslant u_n,$$

若对单位向量 $\boldsymbol{x}$,恒有 $|\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{B}-\boldsymbol{A})\boldsymbol{x}|\leqslant\varepsilon(\varepsilon>0)$，则 $|u_k-\lambda_k|\leqslant\varepsilon(k=1,2,\cdots,n)$.

证 设 $\boldsymbol{A}$ 的属于 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的标准正交向量系为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$,$\boldsymbol{B}$ 的属于 $u_1,u_2,\cdots,u_n$ 的标准正交向量系为 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$.

设 $V_0^k(\boldsymbol{x})=L(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k)$,$V_0^k(\boldsymbol{y})=L(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_k)$,可得

$$\begin{aligned}\lambda_k&=\min_{V^k}\{\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\mid\boldsymbol{x}\in V^k,\|\boldsymbol{x}\|_2=1\}\}\\&\leqslant\min_{V^k}\{\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\mid\boldsymbol{x}\in V_0^k(\boldsymbol{y}),\|\boldsymbol{x}\|_2=1\}\}\\&\leqslant\min_{V^k}\{\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}+\varepsilon\mid\boldsymbol{x}\in V_0^k(\boldsymbol{y}),\|\boldsymbol{x}\|_2=1\}\}\\&=\min_{V^k}\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}\mid\boldsymbol{x}\in V_0^k(\boldsymbol{y}),\|\boldsymbol{x}\|_2=1\}+u_k+\varepsilon,\quad k=1,2,\cdots,n.\end{aligned}$$

同理可得 $u_k\leqslant\lambda_k+\varepsilon(k=1,2,\cdots,n)$，因此 $|u_k-\lambda_k|\leqslant\varepsilon(k=1,2,\cdots,n)$.

16.(Weyl 定理)设实对称矩阵 $\boldsymbol{A}$，$\boldsymbol{A}+\boldsymbol{Q}$ 和 $\boldsymbol{Q}$ 的特征值分别是 $\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$，$u_1\leqslant u_2\leqslant\cdots\leqslant u_n$,$\gamma_1\leqslant\gamma_2\leqslant\cdots\leqslant\gamma_n$，则

$$\lambda_k+\gamma_1\leqslant u_k\leqslant\lambda_k+\gamma_n,\quad k=1,2,\cdots n.$$

证 设 $\boldsymbol{A}$ 的属于 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的标准正交向量系为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$,$\boldsymbol{A}+\boldsymbol{Q}$ 的属于 $u_1,u_2,\cdots,u_n$ 的标准正交向量系为 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$,

设 $V_0^k(\boldsymbol{x})=L(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k)$,$V_0^k(\boldsymbol{y})=L(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_k)$,可得

$$\begin{aligned}u_k&=\min_{V^k}\max\{\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{Q})\boldsymbol{x}\mid\boldsymbol{x}\in V_k,\|\boldsymbol{x}\|_2=1\}\\&\leqslant\min_{V^k}[\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\mid\boldsymbol{x}\in V^k,\|\boldsymbol{x}\|_2=1\}\\&\qquad+\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}\mid\boldsymbol{x}\in V^k,\|\boldsymbol{x}\|_2=1\}]\\&\leqslant\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\mid\boldsymbol{x}\in V_0^k(\boldsymbol{x}),\|\boldsymbol{x}\|_2=1\}+\gamma_n=\lambda_k+\gamma_n,\quad k=1,2,\cdots,n\end{aligned}$$

$$\begin{aligned}\lambda_k&=\min_{V^k}\max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\mid\boldsymbol{x}\in V^k,\|\boldsymbol{x}\|_2=1\}\\&=\min_{V^k}\max\{\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{Q})\boldsymbol{x}+\boldsymbol{x}^{\mathrm{T}}(-\boldsymbol{Q})\boldsymbol{x}\mid\boldsymbol{x}\in V^k,\|\boldsymbol{x}\|_2=1\}\\&=\min_{V^k}\max\{\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{Q})\boldsymbol{x}\mid\boldsymbol{x}\in V^k,\|\boldsymbol{x}\|_2=1\}\\&\qquad+\max\{\boldsymbol{x}^{\mathrm{T}}(-\boldsymbol{Q})\boldsymbol{x}\mid\boldsymbol{x}\in V^k,\|\boldsymbol{x}\|_2=1\}\\&\leqslant\max\{\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{Q})\boldsymbol{x}\mid\boldsymbol{x}\in V_0^k(\boldsymbol{y}),\|\boldsymbol{x}\|_2=1\}+(-\gamma_1)=u_k-\gamma_1,\quad k=1,2,\cdots,n,\end{aligned}$$

所以

$$\lambda_k+\gamma_1\leqslant u_k\leqslant\lambda_k+\gamma_n,\quad k=1,2,\cdots,n.$$

第五章 矩阵分析

在研究数值方法以及线性系统的控制理论等方面的问题时，有必要对矩阵进行类似于数学分析中的极限、级数、微分和积分等运算，这就是矩阵分析的内容. 矩阵序列的极限运算是矩阵分析的基础，矩阵级数在通过建立矩阵函数来表示系统微分方程的解时常常用到. 事实上，在自动控制理论及一些实际问题中，经常遇到求矩阵的特殊导数的问题；在线性控制系统中，常常涉及求解线性微分方程组的问题. 这些情况下，利用矩阵函数表示线性微分方程组的解，形式比较简单，从而简化了复杂的系统状态方程的求解问题.

一、基本概念

1. 矩阵序列

(1) 矩阵序列　称 $\boldsymbol{A}^{(k)}=(a_{ij}^{(k)})_{m\times n}(k=1,2,\cdots)$为一个 $m\times n$ 矩阵序列，记为$\{\boldsymbol{A}^{(k)}\}$.

(2) 收敛性　对矩阵序列$\{\boldsymbol{A}^{(k)}\}$，$\boldsymbol{A}^{(k)}=(a_{ij}^{(k)})_{m\times n}$，若

$$\lim_{k\to\infty}a_{ij}^{(k)}=a_{ij},\qquad i=1,2,\cdots,m;\quad j=1,2,\cdots,n,$$

则称矩阵序列$\{\boldsymbol{A}^{(k)}\}$收敛于矩阵 $\boldsymbol{A}=(a_{ij})$，记为$\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}$. 特别地，若$\lim\limits_{k\to\infty}\boldsymbol{A}^{k}=\boldsymbol{A}$，则称 $\boldsymbol{A}$ 为收敛矩阵.

2. 矩阵级数

(1) 矩阵级数　设$\{\boldsymbol{A}^{(k)}\}$是 $\mathbf{C}^{m\times n}$ 中的矩阵序列，称无穷和

$$\boldsymbol{A}^{(1)}+\boldsymbol{A}^{(2)}+\cdots+\boldsymbol{A}^{(k)}+\cdots$$

为矩阵级数，记为 $\sum\limits_{k=1}^{\infty}\boldsymbol{A}^{(k)}$.

(2) 收敛性　记 $\boldsymbol{S}^{(N)}=\sum\limits_{k=1}^{N}\boldsymbol{A}^{(k)}$ 为矩阵级数$\sum\limits_{k=1}^{\infty}\boldsymbol{A}^{(k)}$ 的部分和，若$\lim\limits_{N\to\infty}\boldsymbol{S}^{(N)}=\boldsymbol{S}$，则称矩阵级数$\sum\limits_{k=1}^{\infty}\boldsymbol{A}^{(k)}$ 收敛且和为 $\boldsymbol{S}$，记为 $\boldsymbol{S}=\sum\limits_{k=1}^{\infty}\boldsymbol{A}^{(k)}$.

(3) 绝对收敛　若 mn 个数项级数 $\sum\limits_{k=1}^{\infty} a_{ij}^{(k)}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 都是绝对收敛的，则称矩阵级数 $\sum\limits_{k=1}^{\infty}\boldsymbol{A}^{(k)}$ 是绝对收敛的.

(4) 幂级数　称 $\sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$ 为矩阵 $\boldsymbol{A}$ 的幂级数，而 $\sum\limits_{k=0}^{\infty}\boldsymbol{A}^k$ 为矩阵 $\boldsymbol{A}$ 的 Neumann 级数.

3. 矩阵函数

由幂级数定义的矩阵函数　设幂级数 $\sum\limits_{k=0}^{\infty} a_k z^k$ 的收敛半径为 r，和函数为 $f(z)$，即 $\sum\limits_{k=0}^{\infty} a_k z^k = f(z)$，$|z|<r$. 若矩阵 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 满足 $r(A)<r$，则称收敛矩阵幂级数 $\sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$ 的和为矩阵函数，记为 $f(\boldsymbol{A})$，即 $f(\boldsymbol{A})=\sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$.

4. 矩阵的微分和积分

(1) 函数矩阵的微分和积分　称 $\boldsymbol{A}(t)=(a_{ij}(t))_{m\times n}$ 为函数矩阵，若每个关于 t 的函数 $a_{ij}(t)$ 在 $[a,b]$ 上可导(可积)，则称 $\boldsymbol{A}(t)$ 在 $[a,b]$ 上可导(可积)，且规定 $\boldsymbol{A}(t)$ 的导数为

$$\boldsymbol{A}'(t)=(a'_{ij}(t))_{m\times n}\quad 或\quad \frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}=\left(\frac{\mathrm{d}a_{ij}(t)}{\mathrm{d}t}\right)_{m\times n}.$$

规定 $\boldsymbol{A}(t)$ 在 $[a,b]$ 上的积分为

$$\int_a^b \boldsymbol{A}(t)\mathrm{d}t=\left(\int_a^b a_{ij}(t)\mathrm{d}t\right)_{m\times n}.$$

(2) 数量函数对矩阵变量的导数　设 $f(\boldsymbol{X})$ 是以矩阵 $\boldsymbol{X}=(x_{ij})_{m\times n}$ 为自变量的 mn 元函数，且 $\dfrac{\partial f}{\partial x_{ij}}(i=1,2,\cdots,m;\ j=1,2,\cdots,n)$ 均存在，则数量函数 $f(\boldsymbol{X})$ 对矩阵 $\boldsymbol{X}$ 的导数为

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{X}}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{m\times n}=\begin{pmatrix}\dfrac{\partial f}{\partial x_{11}} & \cdots & \dfrac{\partial f}{\partial x_{1n}}\\ \vdots & & \vdots\\ \dfrac{\partial f}{\partial x_{m1}} & \cdots & \dfrac{\partial f}{\partial x_{mn}}\end{pmatrix}.$$

(3) 矩阵值函数对矩阵变量的导数　设 $\boldsymbol{X}=(x_{ij})_{m\times n}$，由 mn 元函数 $f_{ij}(\boldsymbol{X})(i=1,2,\cdots,r;j=1,2,\cdots,s)$ 定义的矩阵值函数 $\boldsymbol{F}(\boldsymbol{X})=(f_{ij}(\boldsymbol{X}))_{r\times s}$ 对矩阵 $\boldsymbol{X}$ 的导数为

$$\frac{\mathrm{d}\boldsymbol{F}}{\mathrm{d}\boldsymbol{X}}=\begin{pmatrix}\dfrac{\partial \boldsymbol{F}}{\partial x_{11}} & \cdots & \dfrac{\partial \boldsymbol{F}}{\partial x_{1n}}\\ \vdots & & \vdots\\ \dfrac{\partial \boldsymbol{F}}{\partial x_{m1}} & \cdots & \dfrac{\partial \boldsymbol{F}}{\partial x_{mn}}\end{pmatrix},$$

其中$\dfrac{\partial \boldsymbol{F}}{\partial x_{ij}}=\begin{pmatrix} \dfrac{\partial f_{11}}{\partial x_{ij}} & \cdots & \dfrac{\partial f_{1s}}{\partial x_{ij}} \\ \vdots & & \vdots \\ \dfrac{\partial f_{r1}}{\partial x_{ij}} & \cdots & \dfrac{\partial f_{rs}}{\partial x_{ij}} \end{pmatrix}$, $i=1,2,\cdots,m$; $j=1,2,\cdots,n$.

二、主要结论

1. 矩阵序列的收敛性

(1) 设$\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}$, $\lim\limits_{k\to\infty}\boldsymbol{B}^{(k)}=\boldsymbol{B}$，其中 $\boldsymbol{A}^{(k)}$, $\boldsymbol{B}^{(k)}$ $(k=1,2,\cdots)$, $\boldsymbol{A}$, $\boldsymbol{B}$ 为适当阶数的矩阵，$a,b\in\mathbf{C}$，则

① $\lim\limits_{k\to\infty}(a\boldsymbol{A}^{(k)}+b\boldsymbol{B}^{(k)})=a\boldsymbol{A}+b\boldsymbol{B}$;

② $\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)}=\boldsymbol{A}\boldsymbol{B}$;

③ 当 $\boldsymbol{A}^{(k)}$ $(k=1,2,\cdots)$与 $\boldsymbol{A}$ 均可逆时，有$\lim\limits_{k\to\infty}(\boldsymbol{A}^{(k)})^{-1}=\boldsymbol{A}^{-1}$.

(2) 矩阵序列$\{\boldsymbol{A}^{(k)}\}$收敛的充要条件是，对任何一种范数 $\|\cdot\|$，都有$\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}-\boldsymbol{A}\|=0$.

(3) $\boldsymbol{A}\in\mathbf{C}^{n\times n}$为收敛矩阵的充要条件是 $\boldsymbol{A}$ 的谱半径满足 $r(\boldsymbol{A})<1$.

2. 矩阵级数的收敛性

(1) 设 $\boldsymbol{A}^{(k)}=(a_{ij}^{(k)})_{m\times n}$，则矩阵级数 $\sum\limits_{k=1}^{\infty}\boldsymbol{A}^{(k)}$ 收敛的充要条件是 mn 个数项级数 $\sum\limits_{k=1}^{\infty}a_{ij}^{(k)}$ $(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 都收敛.

(2) 矩阵级数$\sum\limits_{k=1}^{\infty}\boldsymbol{A}^{(k)}$ 绝对收敛的充要条件是级数$\sum\limits_{k=1}^{\infty}\|\boldsymbol{A}^{(k)}\|$ 在任一矩阵范数下收敛.

(3) 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，则 Neumann 级数$\sum\limits_{k=0}^{\infty}\boldsymbol{A}^{k}$ 收敛的充要条件是 $\boldsymbol{A}$ 为收敛矩阵，且其和为 $(\boldsymbol{E}-\boldsymbol{A})^{-1}$.

(4) 设幂级数 $f(z)=\sum\limits_{k=0}^{\infty}c_k z^k$ 的收敛半径为 r，则当 $r(\boldsymbol{A})<r$ 时，矩阵幂级数$\sum\limits_{k=0}^{\infty}c_k\boldsymbol{A}^k$ 绝对收敛，当 $r(\boldsymbol{A})>r$ 时，矩阵幂级数$\sum\limits_{k=0}^{\infty}c_k\boldsymbol{A}^k$ 发散.

3. 矩阵函数值的计算

(1) 利用相似对角化(仅用于 $\boldsymbol{A}$ 可对角化的情形)

① 求出可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$;

② 计算 $f(\boldsymbol{A})=\boldsymbol{P}\mathrm{diag}(f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n))\boldsymbol{P}^{-1}$.

(2) Jordan 标准形法

① 求出可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\boldsymbol{J}_1,\boldsymbol{J}_2,\cdots,\boldsymbol{J}_s)$；其中 $\boldsymbol{J}_i$ 为 m_i 阶的 Jordan 块.

② 对 $i=1,2,\cdots,s$，计算 $f^{(k)}(\lambda_i)(k=0,1,\cdots,m_i-1)$，并构造 m_i 阶矩阵

$$f(\boldsymbol{J}_i)=\begin{pmatrix} f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & \frac{1}{1!}f'(\lambda_i) \\ & & & f(\lambda_i) \end{pmatrix},$$

③ 计算 $f(\boldsymbol{A})=\boldsymbol{P}\mathrm{diag}(f(\boldsymbol{J}_1),f(\boldsymbol{J}_2),\cdots,f(\boldsymbol{J}_s))\boldsymbol{P}^{-1}$.

(3) 数项级数求和法(递推公式计算法)

见本章例 9.

4. 矩阵函数的性质

(1) $\mathrm{e}^{\mathrm{i}\boldsymbol{A}}=\cos\boldsymbol{A}+\mathrm{i}\sin\boldsymbol{A},\cos\boldsymbol{A}=\frac{1}{2}(\mathrm{e}^{\mathrm{i}\boldsymbol{A}}+\mathrm{e}^{-\mathrm{i}\boldsymbol{A}}),\sin A=\frac{1}{2\mathrm{i}}(\mathrm{e}^{\mathrm{i}\boldsymbol{A}}-\mathrm{e}^{-\mathrm{i}\boldsymbol{A}})$.

(2) 若 $\boldsymbol{AB}=\boldsymbol{BA}$，则 $\mathrm{e}^{\boldsymbol{A}}\mathrm{e}^{\boldsymbol{B}}=\mathrm{e}^{\boldsymbol{B}}\mathrm{e}^{\boldsymbol{A}}=\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$(由此易得$(\mathrm{e}^{\boldsymbol{A}})^{-1}=\mathrm{e}^{-\boldsymbol{A}},(\mathrm{e}^{\boldsymbol{A}})^m=\mathrm{e}^{m\boldsymbol{A}}$).

(3) 若 $\boldsymbol{AB}=\boldsymbol{BA}$，则

$$\cos(\boldsymbol{A}+\boldsymbol{B})=\cos\boldsymbol{A}\cos\boldsymbol{B}-\sin\boldsymbol{A}\sin\boldsymbol{B},\quad \cos 2\boldsymbol{A}=\cos^2\boldsymbol{A}-\sin^2\boldsymbol{A},$$
$$\sin(\boldsymbol{A}+\boldsymbol{B})=\sin\boldsymbol{A}\cos\boldsymbol{B}+\cos\boldsymbol{A}\sin\boldsymbol{B},\quad \sin 2\boldsymbol{A}=2\sin\boldsymbol{A}\cos\boldsymbol{A}.$$

5. 矩阵微分和积分的性质

(1) 设函数矩阵 $\boldsymbol{A}(t),\boldsymbol{B}(t)$ 均可微，则

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}(t)+\boldsymbol{B}(t))=\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}+\frac{\mathrm{d}\boldsymbol{B}(t)}{\mathrm{d}t},$$
$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}(t)\boldsymbol{B}(t))=\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}\cdot\boldsymbol{B}(t)+\boldsymbol{A}(t)\cdot\frac{\mathrm{d}\boldsymbol{B}(t)}{\mathrm{d}t},$$
$$\frac{\mathrm{d}}{\mathrm{d}t}(a(t)\boldsymbol{A}(t))=\frac{\mathrm{d}a(t)}{\mathrm{d}t}\cdot\boldsymbol{A}(t)+a(t)\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}.$$

(2) 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，则

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{t\boldsymbol{A}}=\boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}}=\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{A},$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\cos(t\boldsymbol{A})=-\boldsymbol{A}(\sin(t\boldsymbol{A}))=-(\sin(t\boldsymbol{A}))\boldsymbol{A},$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\sin(t\boldsymbol{A})=\boldsymbol{A}(\cos(t\boldsymbol{A}))=(\cos(t\boldsymbol{A}))\boldsymbol{A}.$$

(3) 设函数矩阵 $\boldsymbol{A}(t)$,$\boldsymbol{B}(t)$在$[a,b]$上可积，则

$$\int_a^b(\boldsymbol{A}(t)+\boldsymbol{B}(t))\mathrm{d}t=\int_a^b\boldsymbol{A}(t)\mathrm{d}t+\int_a^b\boldsymbol{B}(t)\mathrm{d}t,\quad \int_a^b\lambda\boldsymbol{A}(t)\mathrm{d}t=\lambda\int_a^b\boldsymbol{A}(t)\mathrm{d}t,\lambda \text{ 为常数.}$$

$$\int_a^b\boldsymbol{A}(t)\boldsymbol{B}\mathrm{d}t=\left(\int_a^b\boldsymbol{A}(t)\mathrm{d}t\right)\boldsymbol{B},\quad \boldsymbol{B} \text{ 为常数矩阵.}$$

$$\int_a^b\boldsymbol{A}\boldsymbol{B}(t)\mathrm{d}t=\boldsymbol{A}\left(\int_a^b\boldsymbol{B}(t)\mathrm{d}t\right),\quad \boldsymbol{A} \text{ 为常数矩阵.}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_0^\tau\boldsymbol{A}(\tau)\mathrm{d}\tau\right)=\boldsymbol{A}(t),\quad \int_a^b\boldsymbol{A}'(t)\mathrm{d}t=\boldsymbol{A}(b)-\boldsymbol{A}(a).$$

6. 微分方程组的解

设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,$\boldsymbol{c}\in\mathbf{C}^n$,$\boldsymbol{x}(t)=(x_1(t),x_2(t),\cdots,x_n(t))^{\mathrm{T}}$,$\boldsymbol{f}(t)=(f_1(t),f_2(t),\cdots,f_n(t))^{\mathrm{T}}$，则一阶线性常系数微分方程组的初值问题

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{f}(t),\\ \boldsymbol{x}(t_0)=\boldsymbol{c}\end{cases}$$

的解为

$$\boldsymbol{x}(t)=\mathrm{e}^{\boldsymbol{A}(t-t_0)}\boldsymbol{c}+\mathrm{e}^{\boldsymbol{A}t}\int_{t_0}^t\mathrm{e}^{-\boldsymbol{A}\tau}\boldsymbol{f}(\tau)\mathrm{d}\tau.$$

三、典型例题

例 1 证明：若对任一矩阵范数 $\|\cdot\|$，方阵 $\boldsymbol{A}$ 满足 $\|\boldsymbol{A}\|<1$，则 $\boldsymbol{A}$ 为收敛矩阵.

证 因为 $\|\boldsymbol{A}^k\|<\|\boldsymbol{A}\|^k$，所以当 $\|\boldsymbol{A}\|<1$ 时，有 $\lim\limits_{k\to\infty}\|\boldsymbol{A}^k\|=0$，也即 $\lim\limits_{k\to\infty}\|\boldsymbol{A}^k-\boldsymbol{0}\|=0$，于是 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$，故 $\boldsymbol{A}$ 为收敛矩阵.

例 2 (1) 已知 $\boldsymbol{A}=\begin{pmatrix}5&-2&-4\\-2&2&2\\-4&2&5\end{pmatrix}$，求 $\boldsymbol{A}^{-1}$. (2) 设 $\boldsymbol{A}=\begin{pmatrix}3&1&-1\\1&2&-1\\2&1&0\end{pmatrix}$，求 $\mathrm{e}^{\boldsymbol{A}t}$.

解 (1) 由 $\det(\lambda\boldsymbol{E}-\boldsymbol{A})=(\lambda-1)^2(\lambda-10)$，求得 $\lambda_1=1$ 的线性无关特征向量为 $(-1,-2,0)^{\mathrm{T}}$ 和 $(-1,0,-1)^{\mathrm{T}}$，而 $\lambda_2=10$ 的特征向量为 $(2,-1,-2)^{\mathrm{T}}$.

令 $\boldsymbol{P}=\begin{pmatrix}-1&-1&2\\-2&0&-1\\0&-1&-2\end{pmatrix}$，则 $\boldsymbol{P}^{-1}=\begin{pmatrix}-1/9&-4/9&1/9\\-4/9&2/9&-5/9\\2/9&-1/9&-2/9\end{pmatrix}$，且 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(1,1,10)$.

于是

$$f(\boldsymbol{A})=\boldsymbol{A}^{-1}=\boldsymbol{P}\begin{pmatrix}1&&\\&1&\\&&1/10\end{pmatrix}\boldsymbol{P}^{-1}=\begin{pmatrix}3/5&1/5&2/5\\1/5&9/10&-1/5\\2/5&-1/5&3/5\end{pmatrix}.$$

(2) $\det(\lambda \boldsymbol{E}-\boldsymbol{A})=(\lambda-1)(\lambda-2)^2$，求得 $\lambda_1=1$ 的特征向量为 $(0,1,1)^{\mathrm{T}}$，而 $\lambda_2=2$ 的线性无关特征向量只有一个 $(1,0,1)^{\mathrm{T}}$. 因此需要再求出它的一个广义特征向量 $(1,1,1)^{\mathrm{T}}$. 取 $\boldsymbol{P}=\begin{pmatrix}0&1&1\\1&0&1\\1&1&1\end{pmatrix}$，则 $\boldsymbol{P}^{-1}=\begin{pmatrix}-1&0&1\\0&-1&1\\1&1&-1\end{pmatrix}$，$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=\begin{pmatrix}1&0&0\\0&2&1\\0&0&2\end{pmatrix}=\begin{pmatrix}\boldsymbol{J}_1&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{J}_2\end{pmatrix}$，所以

$$\begin{aligned}
\mathrm{e}^{\boldsymbol{A}t}&=\boldsymbol{P}\mathrm{diag}(f(\boldsymbol{J}_1),f(\boldsymbol{J}_2))\boldsymbol{P}^{-1}\\
&=\begin{pmatrix}0&1&1\\1&0&1\\1&1&1\end{pmatrix}\begin{pmatrix}\mathrm{e}^t&0&0\\0&\mathrm{e}^{2t}&t\mathrm{e}^{2t}\\0&0&\mathrm{e}^{2t}\end{pmatrix}\begin{pmatrix}-1&0&1\\0&-1&1\\1&1&-1\end{pmatrix}\\
&=\begin{pmatrix}(t+1)\mathrm{e}^{2t}&t\mathrm{e}^{2t}&-t\mathrm{e}^{2t}\\-\mathrm{e}^t+\mathrm{e}^{2t}&\mathrm{e}^{2t}&\mathrm{e}^t+\mathrm{e}^{2t}\\-\mathrm{e}^t+(t+1)\mathrm{e}^{2t}&t\mathrm{e}^{2t}&\mathrm{e}^t-t\mathrm{e}^{2t}\end{pmatrix}.
\end{aligned}$$

例 3 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，证明：

(1) $\dfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{t\boldsymbol{A}}=\boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}}=\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{A}$，

(2) $\dfrac{\mathrm{d}}{\mathrm{d}t}\cos(t\boldsymbol{A})=-\boldsymbol{A}(\sin(t\boldsymbol{A}))=-(\sin(t\boldsymbol{A}))\boldsymbol{A}$，

(3) $\dfrac{\mathrm{d}}{\mathrm{d}t}\sin(t\boldsymbol{A})=\boldsymbol{A}(\cos(t\boldsymbol{A}))=(\cos(t\boldsymbol{A}))\boldsymbol{A}$.

证 (1) 首先有 $\mathrm{e}^{t\boldsymbol{A}}=\sum\limits_{k=0}^{\infty}\dfrac{1}{k!}t^k\boldsymbol{A}^k$. 因等式右边 t 的幂级数对 t 一致收敛，逐项微分得

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{t\boldsymbol{A}})=\sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}\boldsymbol{A}^k.$$

于是

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{t\boldsymbol{A}}=\sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}\boldsymbol{A}^k=\boldsymbol{A}\sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}\boldsymbol{A}^{k-1}=\boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}}.$$

同理

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{t\boldsymbol{A}}=\sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}\boldsymbol{A}^k=\left(\sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}\boldsymbol{A}^{k-1}\right)\boldsymbol{A}=\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{A}.$$

(2)
$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(\cos(t\boldsymbol{A}))&=\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}(t\boldsymbol{A})^{2k}=\sum_{k=1}^{\infty}\frac{(-1)^k\cdot 2k\cdot t^{2k-1}}{(2k)!}\boldsymbol{A}^{2k}\\
&=-\boldsymbol{A}\left(\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(2k-1)!}\cdot(t\boldsymbol{A})^{2k-1}\right)=-\boldsymbol{A}(\sin(t\boldsymbol{A})).
\end{aligned}$$

同理可证明(3).

例 4 设 $\boldsymbol{C},\boldsymbol{D}\in\mathbf{R}^{n\times n}$，$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{C} & \boldsymbol{0}\\ \boldsymbol{D} & \boldsymbol{E}_n\end{pmatrix}$，且 $r(\boldsymbol{C})<1$，求 $\lim\limits_{k\to\infty}\boldsymbol{A}^k$，并计算

$$\lim_{k\to\infty}\begin{pmatrix}0.1 & 0.3 & 0 & 0\\ 0 & 0.6 & 0 & 0\\ 1 & 0 & 1 & 0\\ 1 & 2 & 0 & 1\end{pmatrix}^k.$$

解 因为

$$\boldsymbol{A}^2=\boldsymbol{A}\boldsymbol{A}=\begin{pmatrix}\boldsymbol{C} & \boldsymbol{0}\\ \boldsymbol{D} & \boldsymbol{E}_n\end{pmatrix}\begin{pmatrix}\boldsymbol{C} & \boldsymbol{0}\\ \boldsymbol{D} & \boldsymbol{E}_n\end{pmatrix}=\begin{bmatrix}\boldsymbol{C}^2 & \boldsymbol{0}\\ \boldsymbol{DC}+\boldsymbol{D} & \boldsymbol{E}_n\end{bmatrix},$$

$$\boldsymbol{A}^3=\boldsymbol{A}^2\boldsymbol{A}=\begin{bmatrix}\boldsymbol{C}^2 & \boldsymbol{0}\\ \boldsymbol{DC}+\boldsymbol{D} & \boldsymbol{E}_n\end{bmatrix}\begin{pmatrix}\boldsymbol{C} & \boldsymbol{0}\\ \boldsymbol{D} & \boldsymbol{E}_n\end{pmatrix}=\begin{bmatrix}\boldsymbol{C}^3 & \boldsymbol{0}\\ \boldsymbol{D}(\boldsymbol{C}^2+\boldsymbol{C}+\boldsymbol{E}_n) & \boldsymbol{E}_n\end{bmatrix},$$

递推得

$$\boldsymbol{A}^k=\begin{bmatrix}\boldsymbol{C}^k & \boldsymbol{0}\\ \boldsymbol{D}(\boldsymbol{C}^{k-1}+\cdots+\boldsymbol{C}+\boldsymbol{E}_n) & \boldsymbol{E}_n\end{bmatrix},$$

由已知条件 $r(\boldsymbol{C})<1$，可知 $\boldsymbol{C}$ 为收敛矩阵，即 $\lim\limits_{k\to\infty}\boldsymbol{C}^k=\boldsymbol{0}$，且 $\lim\limits_{k\to\infty}(\boldsymbol{C}^{k-1}+\cdots+\boldsymbol{C}+\boldsymbol{E}_n)=(\boldsymbol{E}_n-\boldsymbol{C})^{-1}$. 所以

$$\lim_{k\to\infty}\boldsymbol{A}^k=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{D}(\boldsymbol{E}_n-\boldsymbol{C})^{-1} & \boldsymbol{E}_n\end{bmatrix}.$$

当 $\boldsymbol{A}$ 为所给矩阵时，可见 $\boldsymbol{C}=\begin{pmatrix}0.1 & 0.3\\ 0 & 0.6\end{pmatrix}$，$\boldsymbol{D}=\begin{pmatrix}1 & 0\\ 1 & 2\end{pmatrix}$. 容易判断 $r(\boldsymbol{C})=0.6<1$，且

$$\boldsymbol{D}(\boldsymbol{E}_n-\boldsymbol{C})^{-1}=\begin{pmatrix}1 & 0\\ 1 & 2\end{pmatrix}\begin{pmatrix}0.9 & -0.3\\ 0 & 0.4\end{pmatrix}^{-1}=\begin{pmatrix}1 & 0\\ 1 & 2\end{pmatrix}\begin{pmatrix}10/9 & 5/6\\ 0 & 5/2\end{pmatrix}=\begin{pmatrix}10/9 & 5/6\\ 10/9 & 35/6\end{pmatrix}.$$

于是有

$$\lim_{k\to\infty}\boldsymbol{A}^k=\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 10/9 & 5/6 & 1 & 0\\ 10/9 & 35/6 & 0 & 1\end{bmatrix}.$$

例 5 已知 $\boldsymbol{A}=\begin{pmatrix}0.5 & 1\\ 1 & -1\end{pmatrix}$，证明矩阵幂级数 $\sum\limits_{k=1}^{\infty}\dfrac{k}{2^{k+2}}\boldsymbol{A}^k$ 收敛，并求其和.

解 由 $\det(\lambda\boldsymbol{E}-\boldsymbol{A})=\begin{vmatrix}\lambda-0.5 & -1\\ -1 & \lambda+1\end{vmatrix}=(\lambda+1.5)(\lambda-1)=0$，得矩阵 $\boldsymbol{A}$ 的两个特征值 $\lambda_1=-1.5$，$\lambda_2=1$，故 $r(\boldsymbol{A})=1.5$. 而级数 $\sum\limits_{k=1}^{\infty}\dfrac{k}{2^{k+2}}x^k$ 的收敛半径为 $r=\lim\limits_{k\to\infty}\dfrac{k}{2^{k+2}}\cdot\dfrac{2^{k+3}}{k+1}=2$，$r(\boldsymbol{A})=1.5<2=r$，从而矩阵幂级数 $\sum\limits_{k=1}^{\infty}\dfrac{k}{2^{k+2}}\boldsymbol{A}^k$ 收敛.

又 $\sum\limits_{k=1}^{\infty}\ \dfrac{k}{2^{k+2}}x^k = \dfrac{x}{4}\sum\limits_{k=1}^{\infty}\ \dfrac{k}{2^k}x^{k-1} = \dfrac{x}{4}\left(\sum\limits_{k=1}^{\infty}\dfrac{x^k}{2^k}\right)' = \dfrac{x}{4}\left(\dfrac{x}{2-x}\right)' = \dfrac{x}{4}\cdot\dfrac{2}{(2-x)^2} =$
$\dfrac{1}{2}x\,(2-x)^{-2}$，于是

$$\begin{aligned}\sum_{k=1}^{\infty}\ \frac{k}{2^{k+2}}\boldsymbol{A}^k &= \frac{1}{2}\boldsymbol{A}(2\boldsymbol{E}-\boldsymbol{A})^{-2} = \frac{1}{2}\begin{pmatrix}0.5 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}1.5 & -1\\ -1 & 3\end{pmatrix}^{-2}\\ &= \begin{pmatrix}1/4 & 1/2\\ 1/2 & -1/2\end{pmatrix}\begin{pmatrix}40/49 & 18/49\\ 18/49 & 13/49\end{pmatrix} = \begin{pmatrix}19/49 & 11/49\\ 11/49 & 5/98\end{pmatrix}.\end{aligned}$$

例 6 (1) 若 $\boldsymbol{A}^2=\boldsymbol{A}$，求 $\sin\boldsymbol{A}$.

(2) 若 $\boldsymbol{A}^2=\boldsymbol{E}$，求 $\sin\boldsymbol{A}$.

解 (1) 若 $\boldsymbol{A}^2=\boldsymbol{A}$，设 $\operatorname{rank}(\boldsymbol{A})=r$，则 $\boldsymbol{A}$ 的特征值为 1 或 0，于是必存在可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{A}$ 可对角化为 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}=\boldsymbol{J}$，故

$$\begin{aligned}\sin A &= \boldsymbol{P}\sin\boldsymbol{J}\cdot\boldsymbol{P}^{-1} = \boldsymbol{P}\begin{pmatrix}\sin 1 &&&&&\\ & \ddots &&&&\\ && \sin 1 &&&\\ &&& 0 &&\\ &&&& \ddots &\\ &&&&& 0\end{pmatrix}\boldsymbol{P}^{-1}\\ &= \sin 1\cdot\boldsymbol{P}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{P}^{-1} = \sin 1\cdot\boldsymbol{A}\end{aligned}$$

(2) 当 $\boldsymbol{A}^2=\boldsymbol{E}$ 时，$\boldsymbol{A}$ 的特征值为 1 或 -1，设 $\boldsymbol{A}$ 有 r 个特征值为 1，$n-r$ 个特征值为 -1，则存在可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{E}_{n-r}\end{pmatrix}\boldsymbol{P}^{-1}$，于是

$$\begin{aligned}\sin\boldsymbol{A} &= \boldsymbol{P}\begin{pmatrix}\sin 1 &&&&&\\ & \ddots &&&&\\ && \sin 1 &&&\\ &&& -\sin 1 &&\\ &&&& \ddots &\\ &&&&& -\sin 1\end{pmatrix}\boldsymbol{P}^{-1}\\ &= \sin 1\cdot P\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{E}_{n-r}\end{pmatrix}\boldsymbol{P}^{-1} = \sin 1\cdot A.\end{aligned}$$

例 7 设 $\boldsymbol{A},\boldsymbol{B}\in\boldsymbol{C}^{n\times n}$，当 $\boldsymbol{AB}=\boldsymbol{BA}$ 时，计算可知 $\mathrm{e}^{\boldsymbol{A}}\mathrm{e}^{\boldsymbol{B}}=\mathrm{e}^{\boldsymbol{B}}\mathrm{e}^{\boldsymbol{A}}=\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$.

(1) 试举例说明，一般情况下 $\mathrm{e}^{\boldsymbol{A}}\mathrm{e}^{\boldsymbol{B}}=\mathrm{e}^{\boldsymbol{B}}\mathrm{e}^{\boldsymbol{A}}=\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$ 可能不成立；

(2) 证明 $|\mathrm{e}^{\boldsymbol{A}}|\cdot|\mathrm{e}^{\boldsymbol{B}}|=|\mathrm{e}^{\boldsymbol{B}}|\cdot|\mathrm{e}^{\boldsymbol{A}}|=|\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}|$ 总是成立.

解　(1) 取 $\boldsymbol{A}=\begin{pmatrix}0&1\\0&0\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$，则 $\boldsymbol{AB}=\begin{pmatrix}0&1\\0&0\end{pmatrix}\neq\boldsymbol{BA}=\begin{pmatrix}0&0\\0&0\end{pmatrix}$.

容易看出 $\boldsymbol{A},\boldsymbol{B}$ 已经是 Jordan 标准形. 于是 $\mathrm{e}^{\boldsymbol{A}}=\begin{pmatrix}1&1\\0&1\end{pmatrix}$，$\mathrm{e}^{\boldsymbol{B}}=\begin{pmatrix}1&0\\0&\mathrm{e}\end{pmatrix}$，而 $\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}0&1\\0&1\end{pmatrix}$，令 $\boldsymbol{P}=\begin{pmatrix}-1&1\\0&1\end{pmatrix}$，则 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{PJP}^{-1}=\begin{pmatrix}-1&1\\0&1\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix}\begin{pmatrix}-1&1\\0&1\end{pmatrix}$，$\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}=\begin{pmatrix}-1&1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\0&\mathrm{e}\end{pmatrix}\begin{pmatrix}-1&1\\0&1\end{pmatrix}=\begin{pmatrix}1&\mathrm{e}-1\\0&\mathrm{e}\end{pmatrix}$. 显然，$\mathrm{e}^{\boldsymbol{A}}\mathrm{e}^{\boldsymbol{B}}=\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\0&\mathrm{e}\end{pmatrix}=\begin{pmatrix}1&\mathrm{e}\\0&\mathrm{e}\end{pmatrix}\neq\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$，$\mathrm{e}^{\boldsymbol{B}}\mathrm{e}^{\boldsymbol{A}}=\begin{pmatrix}1&0\\0&\mathrm{e}\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}=\begin{pmatrix}1&1\\0&\mathrm{e}\end{pmatrix}\neq\mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$.

(2) 设 $\boldsymbol{A}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_s$，相应重数分别为 $m_1,m_2,\cdots,m_s$. 再设 $\boldsymbol{A}$ 的 Jordan 标准形为 $\boldsymbol{J}$，则有可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{J}=\begin{bmatrix}\boldsymbol{J}_1&&\\&\ddots&\\&&\boldsymbol{J}_s\end{bmatrix},$$

其中 $\boldsymbol{J}_i=\begin{bmatrix}\lambda_i&1&&\\&\lambda_i&\ddots&\\&&\ddots&1\\&&&\lambda_i\end{bmatrix}_{m_i\times m_i}$，$i=1,2,\cdots,s$.

于是

$$\mathrm{e}^{\boldsymbol{A}}=\boldsymbol{P}\mathrm{e}^{\boldsymbol{J}}\boldsymbol{P}^{-1}=\boldsymbol{P}\begin{bmatrix}\mathrm{e}^{\boldsymbol{J}_1}&&\\&\ddots&\\&&\mathrm{e}^{\boldsymbol{J}_s}\end{bmatrix}\boldsymbol{P}^{-1},$$

其中 $\mathrm{e}^{\boldsymbol{J}_i}=\begin{bmatrix}\mathrm{e}^{\lambda_i}&\mathrm{e}^{\lambda_i}&\frac{1}{2!}\mathrm{e}^{\lambda_i}&\cdots&\frac{1}{(m_i-1)!}\mathrm{e}^{\lambda_i}\\&\mathrm{e}^{\lambda_i}&\mathrm{e}^{\lambda_i}&\ddots&\vdots\\&&\mathrm{e}^{\lambda_i}&\ddots&\frac{1}{2!}\mathrm{e}^{\lambda_i}\\&&&\ddots&\mathrm{e}^{\lambda_i}\\&&&&\mathrm{e}^{\lambda_i}\end{bmatrix}_{m_i\times m_i}$，$i=1,2,\cdots,s$.

因此有

$$|\mathrm{e}^{\boldsymbol{A}}|=|\boldsymbol{P}||\mathrm{e}^{\boldsymbol{J}}||\boldsymbol{P}^{-1}|=|\mathrm{e}^{\boldsymbol{J}}|=\begin{vmatrix}\mathrm{e}^{\boldsymbol{J}_1}&&&\\&\mathrm{e}^{\boldsymbol{J}_2}&&\\&&\ddots&\\&&&\mathrm{e}^{\boldsymbol{J}_s}\end{vmatrix}=|\mathrm{e}^{\boldsymbol{J}_1}||\mathrm{e}^{\boldsymbol{J}_2}|\cdots|\mathrm{e}^{\boldsymbol{J}_s}|$$

$$=\mathrm{e}^{m_1\lambda_1}\mathrm{e}^{m_2\lambda_2}\cdots\mathrm{e}^{m_s\lambda_s}=\mathrm{e}^{m_1\lambda_1+m_2\lambda_2+\cdots+m_s\lambda_n}=\mathrm{e}^{\mathrm{tr}\boldsymbol{A}}.$$

利用上式结论即得

$$|e^{\boldsymbol{A}}|\cdot|e^{\boldsymbol{B}}|=|e^{\boldsymbol{B}}|\cdot|e^{\boldsymbol{A}}|=e^{\mathrm{tr}\boldsymbol{A}}\cdot e^{\mathrm{tr}\boldsymbol{B}}=e^{\mathrm{tr}\boldsymbol{A}+\mathrm{tr}\boldsymbol{B}}=e^{\mathrm{tr}(\boldsymbol{A}+\boldsymbol{B})}=|e^{\boldsymbol{A}+\boldsymbol{B}}|.$$

例 8 设 $\boldsymbol{X}$ 为 $n\times m$ 矩阵，$\boldsymbol{A},\boldsymbol{B}$ 分别为 $n\times n$ 和 $m\times n$ 的常数矩阵，证明：

(1) $\dfrac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}(\mathrm{tr}(\boldsymbol{B}\boldsymbol{X}))=\dfrac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}(\mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}))=\boldsymbol{B}^{\mathrm{T}}$；

(2) $\dfrac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}(\mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}))=(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})\boldsymbol{X}$.

证 (1) 设 $\boldsymbol{B}=(b_{ij})_{m\times n}$，$\boldsymbol{X}=(x_{ij})_{n\times m}$，则 $\boldsymbol{B}\boldsymbol{X}=\left(\sum\limits_{k=1}^{n}b_{ik}x_{kj}\right)_{m\times m}$，于是

$$\mathrm{tr}(\boldsymbol{B}\boldsymbol{X})=\sum_{k=1}^{n}b_{1k}x_{k1}+\cdots+\sum_{k=1}^{n}b_{jk}x_{kj}+\cdots+\sum_{k=1}^{n}b_{mk}x_{km},$$

$$\frac{\partial\mathrm{tr}(\boldsymbol{B}\boldsymbol{X})}{\partial x_{ij}}=b_{ji},\quad i=1,2,\cdots,n;\ j=1,2,\cdots m.$$

所以

$$\frac{\mathrm{d}}{\mathrm{d}X}(\mathrm{tr}(\boldsymbol{B}\boldsymbol{X}))=\begin{pmatrix}b_{11}&b_{21}&\cdots&b_{m1}\\b_{12}&b_{22}&\cdots&b_{m2}\\\vdots&\vdots&&\vdots\\b_{1n}&b_{2n}&\cdots&b_{mn}\end{pmatrix}=\boldsymbol{B}^{\mathrm{T}}.$$

又因为 $\mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}})=\mathrm{tr}((\boldsymbol{B}\boldsymbol{X})^{\mathrm{T}})=\mathrm{tr}(\boldsymbol{B}\boldsymbol{X})$，故有

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}(\mathrm{tr}(\boldsymbol{B}\boldsymbol{X}))=\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{X}}(\mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}))=\boldsymbol{B}^{\mathrm{T}}.$$

(2) 设 $\boldsymbol{A}=(a_{ij})_{n\times n},\boldsymbol{X}=(x_{ij})_{n\times m}$，$f=\mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X})$，则有

$$\boldsymbol{X}^{\mathrm{T}}=\begin{pmatrix}x_{11}&x_{21}&\cdots&x_{n1}\\x_{12}&x_{22}&\cdots&x_{n2}\\\vdots&\vdots&&\vdots\\x_{1m}&x_{2m}&\cdots&x_{nm}\end{pmatrix},\quad \boldsymbol{A}\boldsymbol{X}=\begin{pmatrix}\sum\limits_{k=1}^{n}a_{1k}x_{k1}&\sum\limits_{k=1}^{n}a_{1k}x_{k2}&\cdots&\sum\limits_{k=1}^{n}a_{1k}x_{km}\\\sum\limits_{k=1}^{n}a_{2k}x_{k1}&\sum\limits_{k=1}^{n}a_{2k}x_{k2}&\cdots&\sum\limits_{k=1}^{n}a_{2k}x_{km}\\\vdots&\vdots&&\vdots\\\sum\limits_{k=1}^{n}a_{nk}x_{k1}&\sum\limits_{k=1}^{n}a_{nk}x_{k2}&\cdots&\sum\limits_{k=1}^{n}a_{nk}x_{km}\end{pmatrix}.$$

$$f=\sum_{l=1}^{n}x_{l1}\sum_{k=1}^{n}a_{lk}x_{k1}+\cdots+\sum_{l=1}^{n}x_{lj}\sum_{k=1}^{n}a_{lk}x_{kj}+\cdots+\sum_{l=1}^{n}x_{lm}\sum_{k=1}^{n}a_{lk}x_{km}.$$

$$\frac{\partial f}{\partial x_{ij}}=\frac{\partial}{\partial x_{ij}}\left(\sum_{l=1}^{n}x_{lj}\sum_{k=1}^{n}a_{lk}x_{kj}\right)=\sum_{l=1}^{n}\left[\frac{\partial x_{lj}}{\partial x_{ij}}\sum_{k=1}^{n}a_{lk}x_{kj}+x_{lj}\cdot\frac{\partial}{\partial x_{ij}}\left(\sum_{k=1}^{n}a_{lk}x_{kj}\right)\right]$$

$$=\sum_{k=1}^{n}a_{ik}x_{kj}+\sum_{l=1}^{n}a_{li}x_{lj}.$$

所以

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{X}}=\left(\frac{\mathrm{d}f}{\mathrm{d}x_{ij}}\right)_{n\times m}=\boldsymbol{AX}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}=(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})\boldsymbol{X}.$$

例 9　已知 $\boldsymbol{A}$ 为一个四阶矩阵，其特征值为 $0,0,\pi,-\pi$，求 $\sin\boldsymbol{A},\cos\boldsymbol{A},\mathrm{e}^{\boldsymbol{A}}$.

解　根据 $\boldsymbol{A}$ 的特征值可知 $\boldsymbol{A}$ 的特征方程为

$$\varphi(\lambda)=\det(\lambda\boldsymbol{E}-\boldsymbol{A})=\lambda^2(\lambda-\pi)(\lambda+\pi)=\lambda^4-\pi^2\lambda^2=0.$$

由 Hamilton-Cayley 定理可知 $\varphi(\boldsymbol{A})=\boldsymbol{0}$，也即 $\boldsymbol{A}^4=\pi^2\boldsymbol{A}^2$. 于是有

$$\begin{aligned}\sin\boldsymbol{A}&=\boldsymbol{A}-\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{5!}\boldsymbol{A}^5-\frac{1}{7!}\boldsymbol{A}^7+\frac{1}{9!}\boldsymbol{A}^9-\cdots\\&=\boldsymbol{A}-\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{5!}\pi^2\boldsymbol{A}^3-\frac{1}{7!}\pi^4\boldsymbol{A}^3+\frac{1}{9!}\pi^6\boldsymbol{A}^3-\cdots\\&=\boldsymbol{A}+\left(-\frac{1}{3!}+\frac{1}{5!}\pi^2-\frac{1}{7!}\pi^4+\frac{1}{9!}\pi^6-\cdots\right)\boldsymbol{A}^3\\&=\boldsymbol{A}+\frac{\sin\pi-\pi}{\pi^3}\boldsymbol{A}^3=\boldsymbol{A}-\pi^{-2}\boldsymbol{A}^3,\end{aligned}$$

$$\begin{aligned}\cos A&=\boldsymbol{E}-\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{4!}\boldsymbol{A}^4-\frac{1}{6!}\boldsymbol{A}^6+\frac{1}{8!}\boldsymbol{A}^8-\cdots\\&=\boldsymbol{E}-\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{4!}\pi^2\boldsymbol{A}^2-\frac{1}{6!}\pi^4\boldsymbol{A}^2+\frac{1}{8!}\pi^6\boldsymbol{A}^2-\cdots\\&=\boldsymbol{E}+\left(-\frac{1}{2!}+\frac{1}{4!}\pi^2-\frac{1}{6!}\pi^4+\frac{1}{8!}\pi^6-\cdots\right)\boldsymbol{A}^2\\&=\boldsymbol{A}+\frac{\cos\pi-1}{\pi^2}\boldsymbol{A}^2=\boldsymbol{E}-2\pi^{-2}\boldsymbol{A}^2,\end{aligned}$$

$$\begin{aligned}\mathrm{e}^{\boldsymbol{A}}&=\boldsymbol{E}+\boldsymbol{A}+\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{4!}\boldsymbol{A}^4+\cdots\\&=\boldsymbol{E}+\boldsymbol{A}+\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{4!}\pi^2\boldsymbol{A}^2+\frac{1}{5!}\pi^2\boldsymbol{A}^3+\frac{1}{6!}\pi^4\boldsymbol{A}^2+\frac{1}{7!}\pi^4\boldsymbol{A}^3+\cdots\\&=\boldsymbol{E}+\boldsymbol{A}+\left(\frac{1}{2!}+\frac{1}{4!}\pi^2+\frac{1}{6!}\pi^4+\cdots\right)\boldsymbol{A}^2+\left(\frac{1}{3!}+\frac{1}{5!}\pi^2+\frac{1}{7!}\pi^4+\cdots\right)\boldsymbol{A}^3\\&=\boldsymbol{E}+\boldsymbol{A}+\frac{1}{\pi^2}\left(\frac{\mathrm{e}^{\pi}+\mathrm{e}^{-\pi}}{2}-1\right)\boldsymbol{A}^2+\frac{1}{\pi^3}\left(\mathrm{e}^{\pi}-\pi-\frac{\mathrm{e}^{\pi}+\mathrm{e}^{-\pi}}{2}\right)\boldsymbol{A}^3.\end{aligned}$$

例 10　设 $\boldsymbol{A}=\begin{pmatrix}-2&1&0\\-4&2&0\\1&0&1\end{pmatrix}$，$\boldsymbol{f}(t)=(1,2,\mathrm{e}^t-1)^{\mathrm{T}}$. 求线性微分方程组初值问题

$$\begin{cases}\dfrac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=\boldsymbol{Ax}(t)+\boldsymbol{f}(t),\\\boldsymbol{x}(0)=(1,1,-1)^{\mathrm{T}}\end{cases}$$

的解.

解 由 $\det(\lambda \boldsymbol{E}-\boldsymbol{A})=\lambda^2(\lambda-1)$，求得特征值 $\lambda_1=0,\lambda_2=1$. 进一步求得 $\lambda_1=0$ 的一个特征向量 $\boldsymbol{x}_1=(-2,-4,2)^{\mathrm{T}}$ 和一个广义特征向量 $\boldsymbol{x}_2=(1,0,1)^{\mathrm{T}}$，以及 $\lambda_2=1$ 的特征向量 $\boldsymbol{x}_3=(0,0,-1)^{\mathrm{T}}$. 取 $\boldsymbol{P}=\begin{pmatrix}-2&1&0\\-4&0&0\\2&1&-1\end{pmatrix}$，则 $\boldsymbol{P}^{-1}=\begin{pmatrix}0&-1/4&0\\1&-1/2&0\\1&-1&-1\end{pmatrix}$，$\boldsymbol{A}$ 的 Jordan 标准形为 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}0&1&0\\0&0&0\\0&0&1\end{pmatrix}$. 计算得到

$$\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{x}(0)=\boldsymbol{P}\begin{pmatrix}1&t&0\\0&1&0\\0&0&\mathrm{e}^t\end{pmatrix}\boldsymbol{P}^{-1}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\begin{pmatrix}1-2t&t&0\\-4t&1+2t&0\\1+2t-\mathrm{e}^t&-1-t+\mathrm{e}^t&\mathrm{e}^t\end{pmatrix}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\begin{pmatrix}1-t\\1-2t\\t-\mathrm{e}^t\end{pmatrix},$$

$$\int_0^t \mathrm{e}^{\boldsymbol{A}(t-\tau)}\boldsymbol{f}(\tau)\mathrm{d}\tau=\int_0^t\boldsymbol{P}\begin{pmatrix}1&t-\tau&0\\0&1&0\\0&0&\mathrm{e}^{t-\tau}\end{pmatrix}\boldsymbol{P}^{-1}\begin{pmatrix}1\\2\\\mathrm{e}^\tau-1\end{pmatrix}\mathrm{d}\tau=\int_0^t\begin{pmatrix}1\\2\\\mathrm{e}^t-1\end{pmatrix}\mathrm{d}\tau=\begin{pmatrix}t\\2t\\t\mathrm{e}^t-t\end{pmatrix}.$$

于是初值问题的解为

$$\boldsymbol{x}(t)=\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{x}(0)+\mathrm{e}^{\boldsymbol{A}t}\int_0^t\mathrm{e}^{-\boldsymbol{A}\tau}\boldsymbol{f}(\tau)\mathrm{d}\tau=\mathrm{e}^{\boldsymbol{A}t}x(0)+\int_0^t\mathrm{e}^{\boldsymbol{A}(t-\tau)}\boldsymbol{f}(\tau)\mathrm{d}\tau$$

$$=\begin{pmatrix}1-t\\1-2t\\t-\mathrm{e}^t\end{pmatrix}+\begin{pmatrix}t\\2t\\t\mathrm{e}^t-t\end{pmatrix}=\begin{pmatrix}1\\1\\(t-1)\mathrm{e}^t\end{pmatrix}.$$

四、习题解答

1. 设 $A=\begin{pmatrix}0&c&c\\c&0&c\\c&c&0\end{pmatrix}$. 讨论 c 取何值时 $\boldsymbol{A}$ 为收敛矩阵.

解 由于 $|\lambda\boldsymbol{E}_3-\boldsymbol{A}|=\begin{vmatrix}\lambda&-c&-c\\-c&\lambda&-c\\-c&-c&\lambda\end{vmatrix}=(\lambda+c)^2(\lambda-2c)$，所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=2c$，$\lambda_2=\lambda_3=-c$，于是 $r(\boldsymbol{A})=2|c|$，而矩阵 $\boldsymbol{A}$ 收敛的充要条件是 $r(\boldsymbol{A})<1$，即 $-\dfrac{1}{2}<c<\dfrac{1}{2}$.

2. 若 $\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}$，证明 $\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}\|=\|\boldsymbol{A}\|$，其中 $\boldsymbol{A}^{(k)},\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\|\cdot\|$ 为 $\mathbf{C}^{m\times n}$ 中的任一矩阵范数. 并问该命题的逆命题是否成立，为什么?

证 由于 $\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}\Leftrightarrow\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}-\boldsymbol{A}\|=0$，再利用矩阵范数的三角不等式推知

$$\big|\,\|\boldsymbol{A}^{(k)}\|-\|\boldsymbol{A}\|\,\big|\leqslant\|\boldsymbol{A}^{(k)}-\boldsymbol{A}\|,$$

所以有$\lim\limits_{k\to\infty}|\,\|\boldsymbol{A}^{(k)}\|-\|\boldsymbol{A}\|\,|=0$，即$\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}\|=\|\boldsymbol{A}\|$.

该命题的逆命题不成立，例如取 $\boldsymbol{A}^{(k)}=\begin{pmatrix}(-1)^k & \frac{1}{k}\\ 1 & 0\end{pmatrix}$，$\boldsymbol{A}=\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}$，并取矩阵范数为 Frobenius 范数，则有$\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}\|=\lim\limits_{k\to\infty}\sqrt{1+1+\frac{1}{k^2}}=\sqrt{2}=\|\boldsymbol{A}\|$，但$\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}$不存在，所以$\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}\neq\boldsymbol{A}$.

3. 设 $\boldsymbol{A}^{(k)}\in\mathbf{C}^{m\times n}$，$\boldsymbol{B}^{(k)}\in\mathbf{C}^{n\times l}$，$\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}$，$\lim\limits_{k\to\infty}\boldsymbol{B}^{(k)}=\boldsymbol{B}$，证明$\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)}=\boldsymbol{AB}$.

证 $\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)}=\boldsymbol{AB}\Leftrightarrow\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)}-\boldsymbol{AB}\|=0$，利用矩阵范数的性质有

$$\begin{aligned}\|\boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)}-\boldsymbol{AB}\|&=\|\boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)}-\boldsymbol{AB}^{(k)}+\boldsymbol{AB}^{(k)}-\boldsymbol{AB}\|\\&\leqslant\|(\boldsymbol{A}^{(k)}-\boldsymbol{A})\boldsymbol{B}^{(k)}\|+\|\boldsymbol{A}(\boldsymbol{B}^{(k)}-\boldsymbol{B})\|\\&\leqslant\|\boldsymbol{B}^{(k)}\|\,\|(\boldsymbol{A}^{(k)}-\boldsymbol{A})\|+\|\boldsymbol{A}\|\,\|\boldsymbol{B}^{(k)}-\boldsymbol{B}\|.\end{aligned}$$

由已知条件$\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}$，$\lim\limits_{k\to\infty}\boldsymbol{B}^{(k)}=\boldsymbol{B}$及第 2 题结论知$\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}-\boldsymbol{A}\|=0$，$\lim\limits_{k\to\infty}\|\boldsymbol{B}^{(k)}-\boldsymbol{B}\|=0$，$\lim\limits_{k\to\infty}\|\boldsymbol{B}^{(k)}\|=\|\boldsymbol{B}\|$. 由此可见，上面不等式的右边趋于 0，所以$\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)}-\boldsymbol{AB}\|=0$.

4. 设 $\boldsymbol{A}^{(k)}\in\mathbf{C}^{n\times n}$，$\lim\limits_{k\to\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}$，$(\boldsymbol{A}^{(k)})^{-1}$和 $\boldsymbol{A}^{-1}$都存在，证明$\lim\limits_{k\to\infty}(\boldsymbol{A}^{(k)})^{-1}=\boldsymbol{A}^{-1}$.

证 记 $\mathrm{adj}(\boldsymbol{A})$为矩阵 $\boldsymbol{A}$ 的伴随矩阵，A_{ij} 为 $\boldsymbol{A}$ 中元素 a_{ij} 的代数余子式，则$(\boldsymbol{A}^{(k)})^{-1}=\dfrac{\mathrm{adj}(\boldsymbol{A}^{(k)})}{\det(\boldsymbol{A}^{(k)})}$，其中

$$\mathrm{adj}(\boldsymbol{A}^{(k)})=\begin{pmatrix}A_{11}^{(k)} & A_{21}^{(k)} & \cdots & A_{n1}^{(k)}\\ A_{12}^{(k)} & A_{22}^{(k)} & \cdots & A_{n2}^{(k)}\\ \vdots & \vdots & & \vdots\\ A_{1n}^{(k)} & A_{2n}^{(k)} & \cdots & A_{nn}^{(k)}\end{pmatrix}.$$

易知 $A_{ij}^{(k)}$ 是 $\boldsymbol{A}^{(k)}$ 中元素的 $n-1$ 次多项式，由多项式函数的连续性知$\lim\limits_{k\to\infty}A_{ij}^{(k)}=A_{ij}$，故$\lim\limits_{k\to\infty}\mathrm{adj}(\boldsymbol{A}^{(k)})=\mathrm{adj}(\boldsymbol{A})$. 同理 $\det(\boldsymbol{A}^{(k)})$是 $\boldsymbol{A}^{(k)}$ 中元素的 n 次多项式，所以$\lim\limits_{k\to\infty}\det(\boldsymbol{A}^{(k)})=\det(\boldsymbol{A})\neq0$，于是$\lim\limits_{k\to\infty}(\boldsymbol{A}^{(k)})^{-1}=\lim\limits_{k\to\infty}\dfrac{\mathrm{adj}(\boldsymbol{A}^{(k)})}{\det(\boldsymbol{A}^{(k)})}=\dfrac{\mathrm{adj}(\boldsymbol{A})}{\det(\boldsymbol{A})}=\boldsymbol{A}^{-1}$.

5. 设矩阵级数 $\sum\limits_{k=0}^{\infty}\boldsymbol{A}^{(k)}$ 收敛(绝对收敛)，证明$\sum\limits_{k=0}^{\infty}\boldsymbol{P}\boldsymbol{A}^{(k)}\boldsymbol{Q}$也收敛(绝对收敛)，且

$$\sum_{k=0}^{\infty}\boldsymbol{P}\boldsymbol{A}^{(k)}\boldsymbol{Q}=\boldsymbol{P}\Big(\sum_{k=0}^{\infty}\boldsymbol{A}^{(k)}\Big)\boldsymbol{Q},$$

其中 $\boldsymbol{A}^{(k)}\in\mathbf{C}^{m\times n}$，$\boldsymbol{P}\in\mathbf{C}^{s\times m}$，$\boldsymbol{Q}\in\mathbf{C}^{n\times l}$.

证 记 $\boldsymbol{S}^{(N)}=\sum\limits_{k=0}^{N}\boldsymbol{P}\boldsymbol{A}^{(k)}\boldsymbol{Q}=\boldsymbol{P}\Big(\sum\limits_{k=0}^{N}\boldsymbol{A}^{(k)}\Big)\boldsymbol{Q}$，于是

$$\sum_{k=0}^{\infty}\boldsymbol{P}\boldsymbol{A}^{(k)}\boldsymbol{Q}=\lim_{N\to\infty}\boldsymbol{S}^{(N)}=\boldsymbol{P}\Big(\lim_{N\to\infty}\sum_{k=0}^{N}\boldsymbol{A}^{(k)}\Big)\boldsymbol{Q}=\boldsymbol{P}\Big(\sum_{k=0}^{\infty}\boldsymbol{A}^{(k)}\Big)\boldsymbol{Q},$$

可见若 $\sum_{k=0}^{\infty}\boldsymbol{A}^{(k)}$ 收敛，则 $\sum_{k=0}^{\infty}\boldsymbol{P}\boldsymbol{A}^{(k)}\boldsymbol{Q}$ 也收敛. 如果 $\sum_{k=0}^{\infty}\boldsymbol{A}^{(k)}$ 绝对收敛，则 $\sum_{k=0}^{\infty}\|\boldsymbol{A}^{(k)}\|$ 收敛. 又由于 $\|\boldsymbol{P}\boldsymbol{A}^{(k)}\boldsymbol{Q}\|\leqslant\|\boldsymbol{P}\|\,\|\boldsymbol{A}^{(k)}\|\,\|\boldsymbol{Q}\|\leqslant c\|\boldsymbol{A}^{(k)}\|$，其中 c 是与 k 无关的正常数，由比较判别法知 $\sum_{k=0}^{\infty}\|\boldsymbol{P}\boldsymbol{A}^{(k)}\boldsymbol{Q}\|$ 收敛，故 $\sum_{k=0}^{\infty}\boldsymbol{P}\boldsymbol{A}^{(k)}\boldsymbol{Q}$ 也绝对收敛.

6. 讨论下列幂级数的敛散性：

(1) $\sum_{k=1}^{\infty}\frac{1}{k^2}\begin{pmatrix}1 & 7\\ -1 & -3\end{pmatrix}^k$；　　(2) $\sum_{k=0}^{\infty}\frac{k}{6^k}\begin{pmatrix}1 & -8\\ -2 & 1\end{pmatrix}^k$.

解 (1) 设 $\boldsymbol{A}=\begin{pmatrix}1 & 7\\ -1 & -3\end{pmatrix}$，可求得 $\boldsymbol{A}$ 的特征值为 $\lambda_1=\lambda_2=-2$，所以 $r(\boldsymbol{A})=2$. 幂级数 $\sum_{k=1}^{\infty}\frac{1}{k^2}x^k$ 的收敛半径为 $r=\lim_{k\to\infty}\left|\frac{a_k}{a_{k+1}}\right|=\lim_{k\to\infty}\frac{(k+1)^2}{k^2}=1$.

由 $r(\boldsymbol{A})=2>r$ 知,矩阵幂级数 $\sum_{k=1}^{\infty}\frac{1}{k^2}\boldsymbol{A}^k$ 发散.

(2) 设 $\boldsymbol{B}=\begin{pmatrix}1 & -8\\ -2 & 1\end{pmatrix}$，可求得 $\boldsymbol{B}$ 的特征值为 $\lambda_1=-3$，$\lambda_2=5$，所以 $r(\boldsymbol{B})=5$. 又因为幂级数 $\sum_{k=0}^{\infty}\frac{k}{6^k}x^k$ 的收敛半径为 $r=\lim_{k\to\infty}\left|\frac{a_k}{a_{k+1}}\right|=\lim_{k\to\infty}\frac{k}{6^k}\,\frac{6^{k+1}}{k+1}=6$，$r(\boldsymbol{B})<r$，所以矩阵幂级数 $\sum_{k=0}^{\infty}\frac{k}{6^k}\boldsymbol{B}^k$ 绝对收敛.

7. 计算 $\sum_{k=0}^{\infty}\begin{pmatrix}0.1 & 0.7\\ 0.3 & 0.6\end{pmatrix}^k$.

解 设 $A=\begin{pmatrix}0.1 & 0.7\\ 0.3 & 0.6\end{pmatrix}$，由于 $\|\boldsymbol{A}\|_\infty=0.9<1$，故矩阵幂级数 $\sum_{k=0}^{\infty}\boldsymbol{A}^k$ 收敛，且其和为 $(\boldsymbol{E}-\boldsymbol{A})^{-1}=\frac{2}{3}\begin{pmatrix}4 & 7\\ 3 & 9\end{pmatrix}$.

8. 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$，$\boldsymbol{AB}=\boldsymbol{BA}$，证明：

$\sin(\boldsymbol{A}+\boldsymbol{B})=\sin\boldsymbol{A}\cos\boldsymbol{B}+\cos\boldsymbol{A}\sin\boldsymbol{B}$，　$\cos(\boldsymbol{A}+\boldsymbol{B})=\cos\boldsymbol{A}\cos\boldsymbol{B}-\sin\boldsymbol{A}\sin\boldsymbol{B}$.

证 由 $\boldsymbol{AB}=\boldsymbol{BA}$，有 $e^{\boldsymbol{A}}e^{\boldsymbol{B}}=e^{\boldsymbol{A}+\boldsymbol{B}}$，

$$\begin{aligned}\sin(\boldsymbol{A}+\boldsymbol{B})&=\frac{1}{2i}(e^{i(\boldsymbol{A}+\boldsymbol{B})}-e^{-i(\boldsymbol{A}+\boldsymbol{B})})=\frac{1}{2i}(e^{i\boldsymbol{A}}e^{i\boldsymbol{B}}-e^{-i\boldsymbol{A}}e^{-i\boldsymbol{B}})\\&=\frac{1}{2i}[(\cos\boldsymbol{A}+i\sin\boldsymbol{A})(\cos\boldsymbol{B}+i\sin\boldsymbol{B})\\&\quad-(\cos\boldsymbol{A}-i\sin\boldsymbol{A})(\cos\boldsymbol{B}-i\sin\boldsymbol{B})]\\&=\sin\boldsymbol{A}\cos\boldsymbol{B}+\cos\boldsymbol{A}\sin\boldsymbol{B}.\end{aligned}$$

同理可证

$$\cos(\boldsymbol{A}+\boldsymbol{B})=\cos\boldsymbol{A}\cos\boldsymbol{B}-\sin\boldsymbol{A}\sin\boldsymbol{B}.$$

9. 设 $\boldsymbol{A}=\begin{pmatrix}2&1&0\\0&0&1\\0&1&0\end{pmatrix}$，求 $e^{\boldsymbol{A}t}$，$\sin\boldsymbol{A}t$.

解 由 $|\lambda\boldsymbol{E}_3-\boldsymbol{A}|=(\lambda+1)(\lambda-1)(\lambda-2)=0$，求得 $\boldsymbol{A}$ 的特征值为 $\lambda_1=-1$，$\lambda_2=1$，$\lambda_3=2$，于是存在可逆矩阵

$$\boldsymbol{P}=\begin{pmatrix}1&-1&1\\-3&1&0\\3&1&0\end{pmatrix},\quad \boldsymbol{P}^{-1}=\frac{1}{6}\begin{pmatrix}0&-1&1\\0&3&3\\6&4&2\end{pmatrix},$$

使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}-1&&\\&1&\\&&2\end{bmatrix}.$$

再根据矩阵函数值的计算公式得

$$\begin{aligned}e^{\boldsymbol{A}t}&=\boldsymbol{P}\mathrm{diag}(e^{-t},\ e^{t},\ e^{2t})\boldsymbol{P}^{-1}\\&=\frac{1}{6}\begin{pmatrix}6e^{2t}&4e^{2t}-3e^{t}-e^{-t}&2e^{2t}-3e^{t}+e^{-t}\\0&3e^{t}+3e^{-t}&3e^{t}-3e^{-t}\\0&3e^{t}-3e^{-t}&3e^{t}+3e^{-t}\end{pmatrix},\\ \sin\boldsymbol{A}t&=\boldsymbol{P}\mathrm{diag}(\sin(-t),\ \sin t,\ \sin 2t)\boldsymbol{P}^{-1}\\&=\frac{1}{6}\begin{pmatrix}\sin 2t&4\sin 2t-2\sin t&2\sin 2t-4\sin t\\0&0&6\sin t\\0&6\sin t&0\end{pmatrix}.\end{aligned}$$

10. 设 $\boldsymbol{A}=\begin{pmatrix}-1&-2&6\\-1&0&3\\-1&-1&4\end{pmatrix}$，求 $e^{\boldsymbol{A}t}$，$\cos\boldsymbol{A}t$.

解 由

$$|\lambda\boldsymbol{E}_3-\boldsymbol{A}|=\begin{vmatrix}\lambda+1&2&-6\\1&\lambda&-3\\1&1&\lambda-4\end{vmatrix}=(\lambda-1)^3=0$$

得出 $\boldsymbol{A}$ 的特征值 $\lambda=1$，解齐次线性方程组 $(\boldsymbol{A}-\boldsymbol{E}_3)\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{pmatrix}-2&-2&6\\-1&-1&3\\-1&-1&3\end{pmatrix}\boldsymbol{x}=\boldsymbol{0},$$

得出 $\lambda=1$ 的两个线性无关的特征向量 $\boldsymbol{\alpha}_1=(-1,1,0)^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=(2,1,1)^{\mathrm{T}}$. 又对 $\boldsymbol{\alpha}_2$，因非齐次方

程组$(\boldsymbol{A}-\boldsymbol{E}_3)\boldsymbol{\beta}=\boldsymbol{\alpha}_2$相容，故可求得其解$\boldsymbol{\beta}=(-1,0,0)^{\mathrm{T}}$. 由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}$构造可逆矩阵$\boldsymbol{P}=\begin{pmatrix}-1&2&-1\\1&1&0\\0&1&0\end{pmatrix}$，$\boldsymbol{P}^{-1}=\begin{pmatrix}0&1&-1\\0&0&1\\-1&-1&3\end{pmatrix}$，使$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}$为$\boldsymbol{A}$的 Jordan 标准形. 于是

$$\mathrm{e}^{\boldsymbol{A}t}=\boldsymbol{P}\begin{pmatrix}\mathrm{e}^t&0&0\\0&\mathrm{e}^t&t\mathrm{e}^t\\0&0&\mathrm{e}^t\end{pmatrix}\boldsymbol{P}^{-1}=\begin{pmatrix}-1&2&-1\\1&1&0\\0&1&0\end{pmatrix}\begin{pmatrix}\mathrm{e}^t&0&0\\0&\mathrm{e}^t&t\mathrm{e}^t\\0&0&\mathrm{e}^t\end{pmatrix}\begin{pmatrix}0&1&-1\\0&0&1\\-1&-1&3\end{pmatrix}$$
$$=\mathrm{e}^t\begin{pmatrix}1-2t&-2t&6t\\-t&1-t&3t\\-t&-t&1+3t\end{pmatrix},$$

$$\cos\boldsymbol{A}t=\boldsymbol{P}\begin{pmatrix}\cos t&0&0\\0&\cos t&-t\sin t\\0&0&\cos t\end{pmatrix}\boldsymbol{P}^{-1}$$
$$=\begin{pmatrix}2t\sin t+\cos t&2t\sin t&-6t\sin t\\t\sin t&\cos t+t\sin t&-3t\sin t\\t\sin t&t\sin t&\cos t-3t\sin t\end{pmatrix}.$$

11. 设$\boldsymbol{A}=\begin{pmatrix}1&0&0&0\\1&1&0&0\\0&1&1&0\\0&0&1&1\end{pmatrix}$，求$\ln\boldsymbol{A}$.

解 法一 事实上，可证明$f(\boldsymbol{A}^{\mathrm{T}})=[f(\boldsymbol{A})]^{\mathrm{T}}$成立. 本题中$\boldsymbol{A}^{\mathrm{T}}$为一 Jordan 标准形矩阵，由$f(\boldsymbol{A})=\ln\boldsymbol{A}$,可知$f(1)=0,f'(1)=1,f''(1)=-1,f'''(1)=2$. 所以

$$\ln\boldsymbol{A}=[\ln\boldsymbol{A}^{\mathrm{T}}]^{\mathrm{T}}=\begin{pmatrix}f(1)&f'(1)&\frac{f''(1)}{2!}&\frac{f'''(1)}{3!}\\&f(1)&f'(1)&\frac{f''(1)}{2!}\\&&f(1)&f'(1)\\&&&f(1)\end{pmatrix}^{\mathrm{T}}$$
$$=\begin{pmatrix}0&1&-\frac{1}{2}&\frac{1}{3}\\&0&1&-\frac{1}{2}\\&&0&1\\&&&0\end{pmatrix}^{\mathrm{T}}=\begin{pmatrix}0&&&\\1&0&&\\-\frac{1}{2}&1&0&\\\frac{1}{3}&-\frac{1}{2}&1&0\end{pmatrix}.$$

法二 对 $\boldsymbol{A}$ 求得 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}1&1&&\\&1&1&\\&&1&1\\&&&1\end{pmatrix}=\boldsymbol{J}$，进一步得到

$$\ln\boldsymbol{A}=\boldsymbol{P}\ln\boldsymbol{J}\cdot\boldsymbol{P}^{-1}=\begin{pmatrix}0&0&0&0\\1&0&0&0\\-\frac{1}{2}&1&0&0\\\frac{1}{3}&-\frac{1}{2}&1&0\end{pmatrix}.$$

12. 设 $\boldsymbol{A}(t)$ 和 $\boldsymbol{A}^{-1}(t)$ 均为 n 阶可微矩阵，证明：

$$\frac{\mathrm{d}\boldsymbol{A}^{-1}(t)}{\mathrm{d}t}=-\boldsymbol{A}^{-1}(t)\left(\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}\right)\boldsymbol{A}^{-1}(t).$$

证 对 $\boldsymbol{A}(t)\boldsymbol{A}^{-1}(t)=\boldsymbol{E}$ 两端关于 t 求导数，可得

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}\cdot\boldsymbol{A}^{-1}(t)+\boldsymbol{A}(t)\cdot\frac{\mathrm{d}\boldsymbol{A}^{-1}(t)}{\mathrm{d}t}=\boldsymbol{0}.$$

两边左乘 $\boldsymbol{A}^{-1}(t)$ 并移项，即得

$$\frac{\mathrm{d}\boldsymbol{A}^{-1}(t)}{\mathrm{d}t}=-\boldsymbol{A}^{-1}(t)\left(\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}\right)\boldsymbol{A}^{-1}(t).$$

13. 设 $f(\boldsymbol{X})=\mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}),\boldsymbol{X}\in\mathbf{R}^{m\times n}$，求 $\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{X}}$.

解 这是数量函数对矩阵变量的导数. 设 $\boldsymbol{X}=(x_{ij})_{m\times n}$，则

$$f(\boldsymbol{X})=\|\boldsymbol{X}\|_F^2=\sum_{s=1}^{m}\sum_{t=1}^{n}x_{st}^2=\mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}).$$

又因为 $\frac{\partial f}{\partial x_{ij}}=2x_{ij}(i=1,2,\cdots,m;\ j=1,2,\cdots,n)$，所以

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{X}}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{m\times n}=(2x_{ij})_{m\times n}=2\boldsymbol{X}.$$

14. 设 $\boldsymbol{A}\in\mathbf{R}^{m\times n},\boldsymbol{x}\in\mathbf{R}^n,\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{Ax}$，求 $\frac{\mathrm{d}\boldsymbol{F}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}},\frac{\mathrm{d}\boldsymbol{F}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}$.

解 设 $\boldsymbol{A}=(a_{ij})_{m\times n},\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，由于

$$\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{Ax}=\left(\sum_{k=1}^{n}a_{1k}x_k,\cdots,\sum_{k=1}^{n}a_{mk}x_k\right)^{\mathrm{T}},$$

所以 $\frac{\partial\boldsymbol{F}}{\partial x_i}=(a_{1i},\cdots,a_{mi})^{\mathrm{T}},\frac{\mathrm{d}\boldsymbol{F}}{\mathrm{d}\boldsymbol{x}}=\left(\frac{\partial\boldsymbol{F}}{\partial x_1},\cdots,\frac{\partial\boldsymbol{F}}{\partial x_n}\right)^{\mathrm{T}}=(a_{11},\cdots,a_{m1},\cdots,a_{1n},\cdots,a_{mn})^{\mathrm{T}}$.

而

$$\frac{\mathrm{d}\boldsymbol{F}}{\mathrm{d}\boldsymbol{x}^{\mathrm{T}}}=\left(\frac{\partial\boldsymbol{F}}{\partial x_1},\cdots,\frac{\partial\boldsymbol{F}}{\partial x_n}\right)=\begin{bmatrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{m1}&\cdots&a_{mn}\end{bmatrix}=\boldsymbol{A}.$$

15. 设 $\boldsymbol{X}\in\mathbf{R}^{n\times n}$,$\det(\boldsymbol{X})\neq 0$，$f(\boldsymbol{X})=\det(\boldsymbol{X})$. 证明$\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{X}}=(\det(\boldsymbol{X}))(\boldsymbol{X}^{-1})^{\mathrm{T}}$.

证 设 $\boldsymbol{X}=(x_{ij})_{n\times n}$，记 x_{ij} 的代数余子式为 X_{ij}，$\boldsymbol{X}$ 的伴随矩阵为 $\mathrm{adj}(\boldsymbol{X})$. 将 $\det(\boldsymbol{X})$ 按第 i 行展开，得

$$f(\boldsymbol{X})=\det(\boldsymbol{X})=x_{i1}X_{i1}+\cdots+x_{ij}X_{ij}+\cdots+x_{in}X_{in},$$

所以$\dfrac{\partial f}{\partial x_{ij}}=X_{ij}(i,j=1,2,\cdots,n)$，从而有

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{X}}=(X_{ij})_{n\times n}=(\mathrm{adj}(\boldsymbol{X}))^{\mathrm{T}}=((\det(\boldsymbol{X}))\boldsymbol{X}^{-1})^{\mathrm{T}}=(\det(\boldsymbol{X}))(\boldsymbol{X}^{-1})^{\mathrm{T}}.$$

第六章 广义逆矩阵

广义逆矩阵是通常逆矩阵的推广，它的研究起源和一大应用就是对相容线性方程组的一般解、最小范数解，不相容方程组的最小二乘解、最佳最小二乘解（最佳逼近解）做出了统一的理论解释，解决了一般线性方程组的求解问题. 除此之外，广义逆矩阵在数理统计、系统理论、最优化理论、现代控制理论，网络理论等许多领域也有广泛的应用，是矩阵理论的一个重要分支.

一、基本概念

1. Penrose 的四个矩阵方程

(1) $\boldsymbol{AGA}=\boldsymbol{A}$； (2) $\boldsymbol{GAG}=\boldsymbol{G}$； (3) $(\boldsymbol{GA})^{\mathrm{H}}=\boldsymbol{GA}$； (4) $(\boldsymbol{AG})^{\mathrm{H}}=\boldsymbol{AG}$.

2. 矩阵的单边逆

设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，若存在矩阵 $\boldsymbol{G}$，满足 $\boldsymbol{GA}=\boldsymbol{E}_n$（$\boldsymbol{AG}=\boldsymbol{E}_m$），则称 $\boldsymbol{G}$ 为 $\boldsymbol{A}$ 的左（右）逆矩阵记为 $\boldsymbol{G}=\boldsymbol{A}_{\mathrm{L}}^{-1}$（$\boldsymbol{G}=\boldsymbol{A}_{\mathrm{R}}^{-1}$），并称 $\boldsymbol{A}$ 左（右）可逆.

3. 几种常用的广义逆矩阵

(1) 矩阵 $\boldsymbol{A}$ 的广义逆矩阵 $\boldsymbol{A}^{-}$ 满足 Penrose 方程(1)的矩阵，其全体记为 $\boldsymbol{A}\{1\}$；

(2) 矩阵 $\boldsymbol{A}$ 的自反广义逆矩阵 $\boldsymbol{A}_{\mathrm{r}}^{-}$ 满足 Penrose 方程(1)和(2)的矩阵，其全体记为 $\boldsymbol{A}\{1,2\}$；

(3) 矩阵 $\boldsymbol{A}$ 的$\{1,3\}$逆满足 Penrose 方程(1)和(3)的矩阵，其全体记为 $\boldsymbol{A}\{1,3\}$；

(4) 矩阵 $\boldsymbol{A}$ 的$\{1,4\}$逆满足 Penrose 方程(1)和(4)的矩阵，其全体记为 $\boldsymbol{A}\{1,4\}$；

(5) 矩阵 $\boldsymbol{A}$ 的 M-P 广义逆矩阵 $\boldsymbol{A}^{+}$ 满足 Penrose 方程(1)～(4)的矩阵.

4. 线性方程组的解

(1) 最小范数解　设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{b}\in\mathbf{C}^{m}$，当方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解时，称其解集中范数最小的解为最小范数解.

(2) 最小二乘解　设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{b}\in\mathbf{C}^m$，当方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 无解时，若存在 $\boldsymbol{x}_0\in\mathbf{C}^n$，使得 $\|\boldsymbol{Ax}_0-\boldsymbol{b}\|_2=\min\limits_{\boldsymbol{x}\in\mathbf{C}^n}\|\boldsymbol{Ax}-\boldsymbol{b}\|_2$，则称 $\boldsymbol{x}_0$ 为 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解.

特别地，把 $\boldsymbol{Ax}=\boldsymbol{b}$ 的所有最小二乘解中范数最小的解称为最佳最小二乘解(最佳逼近解).

二、主要结论

1. 矩阵的单边逆的存在条件和计算公式

(1) 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，则 $\boldsymbol{A}$ 左可逆的充要条件是 $\boldsymbol{A}$ 列满秩(或 $N(\boldsymbol{A})=\{0\}$)，$\boldsymbol{A}$ 右可逆的充要条件是 $\boldsymbol{A}$ 行满秩(或 $R(\boldsymbol{A})=\mathbf{C}^m$).

(2) 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$ 是左可逆矩阵，则 $\boldsymbol{G}=(\boldsymbol{A}_1^{-1}-\boldsymbol{BA}_2\boldsymbol{A}_1^{-1}\quad \boldsymbol{B})\boldsymbol{P}$ 为 $\boldsymbol{A}$ 的一个左逆矩阵，其中 $\boldsymbol{B}\in\mathbf{C}^{n\times(m-n)}$ 为任意矩阵，$\boldsymbol{P}$ 为使 $\boldsymbol{PA}=\begin{pmatrix}\boldsymbol{A}_1\\ \boldsymbol{A}_2\end{pmatrix}$($\boldsymbol{A}_1$ 为 n 阶可逆方阵)的行初等变换矩阵.

(3) 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$ 是右可逆矩阵，则 $\boldsymbol{G}=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{A}_1^{-1}-\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\boldsymbol{D}\\ \boldsymbol{D}\end{pmatrix}$ 为 $\boldsymbol{A}$ 的一个右逆矩阵. 其中 $\boldsymbol{D}\in\mathbf{C}^{(n-m)\times m}$ 为任意矩阵，$\boldsymbol{Q}$ 为使 $\boldsymbol{AQ}=(\boldsymbol{A}_1\quad \boldsymbol{A}_2)$ ($\boldsymbol{A}_1$ 为 m 阶可逆阵)的列初等变换矩阵.

2. 广义逆矩阵的计算公式

(1) 利用 $\boldsymbol{A}$ 的最大秩分解求 $\boldsymbol{A}^-$ 和 $\boldsymbol{A}_r^-$　设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$ 的最大秩分解为 $\boldsymbol{A}=\boldsymbol{BD}$，则 $\boldsymbol{A}^-=\boldsymbol{D}_{\mathrm{R}}^{-1}\boldsymbol{B}_{\mathrm{L}}^{-1}$ 是 $\boldsymbol{A}$ 的一个广义逆矩阵，同时还是 $\boldsymbol{A}$ 的自反广义逆矩阵.

(2) 利用矩阵的行交换和列交换法求 $\boldsymbol{A}^-$ 和 $\boldsymbol{A}_r^-$　设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$，通过行交换(对应初等变换矩阵 $\boldsymbol{P}$)和列交换(对应初等变换矩阵 $\boldsymbol{Q}$)将矩阵 $\boldsymbol{A}$ 化为 $\boldsymbol{A}_1$，即

$$\boldsymbol{PAQ}=\boldsymbol{A}_1=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix},$$

其中 $\boldsymbol{A}_{11}$ 是 r 阶可逆矩阵，则 $\boldsymbol{A}^-=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{A}_{11}^{-1} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{P}$ 为 $\boldsymbol{A}$ 的一个广义逆矩阵，同时还是 $\boldsymbol{A}$ 的自反广义逆矩阵.

(3) 利用矩阵的初等变换化为标准形求 $\boldsymbol{A}^-$　设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$，则必存在可逆矩阵 $\boldsymbol{P}$，$\boldsymbol{Q}$，使得

① $\boldsymbol{PAQ}=\boldsymbol{B}_1=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{B}_{12}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}$，此时 $\boldsymbol{A}^-=\boldsymbol{QB}_1^-\boldsymbol{P}=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{G}_{12}\\ \boldsymbol{G}_{21} & \boldsymbol{G}_{22}\end{pmatrix}\boldsymbol{P}$，其中 $\boldsymbol{B}_{12}\boldsymbol{G}_{21}=\boldsymbol{0}$，$\boldsymbol{G}_{12}$，$\boldsymbol{G}_{22}$ 是相应维数的任意矩阵.

② $\boldsymbol{PAQ}=\boldsymbol{B}_2=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}$，此时 $\boldsymbol{A}^-=\boldsymbol{QB}_2^-\boldsymbol{P}=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{G}_{12}\\ \boldsymbol{G}_{21} & \boldsymbol{G}_{22}\end{pmatrix}\boldsymbol{P}$，其中 $\boldsymbol{G}_{12}$，$\boldsymbol{G}_{21}$，$\boldsymbol{G}_{22}$ 是相应

维数的任意矩阵.

(4) 用最大秩分解法求 $\boldsymbol{A}^{+}$ 设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$ 的最大秩分解为 $\boldsymbol{A}=\boldsymbol{BD}$，则

$$\boldsymbol{A}^{+}=\boldsymbol{D}^{+}\boldsymbol{B}^{+}=\boldsymbol{D}^{\mathrm{H}}(\boldsymbol{DD}^{\mathrm{H}})^{-1}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{H}}.$$

(5) 用奇异值分解法求 $\boldsymbol{A}^{+}$

① 设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$ 的奇异值分解为 $\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{D}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\boldsymbol{V}$，则 $\boldsymbol{A}^{+}=\boldsymbol{V}^{\mathrm{H}}\begin{pmatrix}\boldsymbol{D}_r^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}$.

② 设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$，$\lambda_i$ 和 $\boldsymbol{\alpha}_i(i=1,2,\cdots,r)$ 分别是 $\boldsymbol{AA}^{\mathrm{H}}$ 的 r 个非零特征值和相应的单位正交特征向量，记 $\Delta_r=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_r)$，$\boldsymbol{U}_1=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)$，则 $\boldsymbol{A}^{+}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}_1\Delta_r^{-1}\boldsymbol{U}_1^{\mathrm{H}}$.

(6) 用谱分解法求 $\boldsymbol{A}^{+}$ 设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$，$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\sum\limits_{i=1}^{r}\lambda_i\boldsymbol{A}_i$ 为谱分解式，$\lambda_i(i=1,2,\cdots,r)$ 为 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的 r 个相异特征值，则 $\boldsymbol{A}^{+}=\sum\limits_{i=1}^{r}\lambda_i^{-}\dfrac{P_i(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})}{P_i(\lambda_i)}\boldsymbol{A}^{\mathrm{H}}$. 其中 $\lambda_i^{-}=\begin{cases}\lambda_i^{-1}, & \lambda_i\neq 0,\\ 0, & \lambda_i=0,\end{cases}$ $P_i(\lambda)=\prod\limits_{\substack{j=1\\ j\neq i}}^{r}(\lambda-\lambda_j)$.

(7) 用极限算法求 $\boldsymbol{A}^{+}$ 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，则

$$\boldsymbol{A}^{+}=\lim_{\delta\to 0}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}+\delta^2\boldsymbol{E}_n)^{-1}\boldsymbol{A}^{\mathrm{H}}=\lim_{\delta\to 0}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}}+\delta^2\boldsymbol{E}_m)^{-1}.$$

(8) 用级数展开法求 $\boldsymbol{A}^{+}$ 设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$，$\sigma_i(i=1,2,\cdots,r)$ 是 $\boldsymbol{A}$ 的 n 个奇异值，又令 $c=\max\limits_{1\leqslant i\leqslant n}\sigma_i^2$，取 $0<a<\dfrac{c}{2}$，则 $\boldsymbol{A}^{+}=a\sum\limits_{k=0}^{\infty}(\boldsymbol{E}-a\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^k\boldsymbol{A}^{\mathrm{H}}$.

3. 广义逆矩阵的性质

(1) 广义逆矩阵 $\boldsymbol{A}^{-}$ 的性质

① 任何矩阵 $\boldsymbol{A}$ 都有广义逆矩阵 $\boldsymbol{A}^{-}$.

② $\boldsymbol{A}$ 的广义逆矩阵 $\boldsymbol{A}^{-}$ 唯一的充要条件是 $\boldsymbol{A}$ 为可逆方阵.

③ $(\boldsymbol{A}^{\mathrm{T}})^{-}=(\boldsymbol{A}^{-})^{\mathrm{T}}$，$(\boldsymbol{A}^{\mathrm{H}})^{-}=(\boldsymbol{A}^{-})^{\mathrm{H}}$.

④ $(\lambda\boldsymbol{A})^{-}=\lambda^{-}\boldsymbol{A}^{-}$.

⑤ $R(\boldsymbol{AA}^{-})=R(\boldsymbol{A})$，$N(\boldsymbol{A}^{-}\boldsymbol{A})=N(\boldsymbol{A})$.

⑥ $\boldsymbol{A}^{-}\in\boldsymbol{A}\{1,2\}\Leftrightarrow\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^{-})$.

(2) 自反广义逆矩阵 $\boldsymbol{A}_{\mathrm{r}}^{-}$ 的性质

① 任何矩阵 $\boldsymbol{A}$ 都有广义逆矩阵 $\boldsymbol{A}_{\mathrm{r}}^{-}$.

② 若 $\boldsymbol{X},\boldsymbol{Y}$ 都是 $\boldsymbol{A}$ 的广义逆矩阵，则 $\boldsymbol{Z}=\boldsymbol{XAY}$ 为 $\boldsymbol{A}$ 的自反广义逆矩阵.

③ 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，则 $\boldsymbol{X}=(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{-}\boldsymbol{A}^{\mathrm{H}}$，$\boldsymbol{Y}=\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-}$ 都是 $\boldsymbol{A}$ 的自反广义逆矩阵.

④ $\boldsymbol{AA}_{\mathrm{r}}^{-}$ 和 $\boldsymbol{A}_{\mathrm{r}}^{-}\boldsymbol{A}$ 都是幂等矩阵.

⑤ 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，则存在 $\boldsymbol{A}_{\mathrm{r}}^{-}$，使 $R(\boldsymbol{A})\oplus N(\boldsymbol{A}_{\mathrm{r}}^{-})=\mathbf{C}^m$，$N(\boldsymbol{A})\oplus R(\boldsymbol{A}_{\mathrm{r}}^{-})=\mathbf{C}^n$.

⑥ $R(AA_r^-)=R(A),N(AA_r^-)=N(A_r^-)$.

(3) M-P 广义逆 A^+ 的性质

① $(A^+)^+=A,(A^T)^+=(A^+)^T,(A^H)^+=(A^+)^H$.

② $R(A^+)=R(A^H),N(A^+)=N(A^H)$.

③ $AA^+=P_{R(A)},A^+A=P_{R(A^+)}$.

④ $R(A)=R(A^H)\Leftrightarrow AA^+=A^+A$.

⑤ $(A^HA)^+=A^+(A^H)^+,(AA^H)^+=(A^H)^+A^+$.

⑥ $(A^HA)^+=A^+(AA^H)^+A=A^H(AA^H)^+(A^H)^+$.

⑦ $AA^+=(AA^H)(AA^H)^+=(AA^H)^+(AA^H),A^+A=(A^HA)(A^HA)^+=(A^HA)^+(A^HA)$.

⑧ $(AB)^+=B^+A^+\Leftrightarrow R(A^HAB)\subset R(B),R(BB^HA^H)\subset R(A^H)$.

4. 广义逆矩阵与方程的求解

(1) 设 $A\in C^{m\times n},B\in C^{p\times q},D\in C^{m\times q}$，则矩阵方程 $AXB=D$ 有解的充要条件是存在 A^-，B^-，使得 $AA^-DB^-B=D$. 此时其通解为

$$X=A^-DB^-+Y-A^-AYBB^-,\quad \forall Y\in C^{n\times p}.$$

特别地，$AX=D$ 有解的充要条件是 $AA^-D=D$，通解为 $X=A^-D+Y-A^-AY$，$\forall Y\in C^{n\times p}$. $XB=D$ 有解的充要条件是 $DB^-B=D$，通解为 $X=DB^-+Y-YBB^-$，$\forall Y\in C^{p\times m}$.

(2) 设 $A_1\in C^{m\times n},D_1\in C^{m\times l},A_2\in C^{l\times p},D_2\in C^{n\times p}$，则 $\begin{cases}A_1X=D_1\\XA_2=D_2\end{cases}$，有公共解的充要条件是 $A_1A_1^-D_1=D_1,D_2A_2^-A_2=D_2,A_1D_2=D_1A_2$，通解为

$$X=X_0+(E_n-A_1^-A_1)Y(E_l-A_2A_2^-),\quad \forall Y\in C^{n\times l}.$$

其中 X_0 为原矩阵方程组的一个解.

(3) 设 $A\in C^{m\times n},b\in C^m$，则

① $Ax=b$ 有解的充要条件是 $AA^-b=b$，通解为 $x=A^-b+(E_n-A^-A)y(\forall y\in C^n)$. 最小范数解为 $Db(D\in A\{1,3\})$，且唯一.

② 若 $Ax=b$ 不相容，则 $Gb(G\in A\{1,4\})$ 是 $Ax=b$ 的一个最小二乘解. 且最小二乘解通解的形式为 $x=Gb+(E_n-A^-A)u(\forall u\in C^n,G\in A\{1,4\})$. 而 $x_0=A^+b$ 是 $Ax=b$ 的唯一最佳最小二乘解，即最佳逼近解.

5. 广义逆矩阵集合的表示

以下总假设 $A\in C^{m\times n}$.

(1) $A\{1\}=\{G|G=A^-+U-A^-AUAA^-,\ \forall U\in C^{n\times m}\}$，其中 A^- 为 A 的一个$\{1\}$逆.

(2) $A\{1,2\}=\{G|\mathrm{rank}(G)=\mathrm{rank}(A),\ G\in A\{1\}\}$.

(3) $A\{1,3\}=\{G|G=A^{(1,3)}+(E-A^{(1,3)}A)U,\forall U\in C^{n\times m}\}$，其中 $A^{(1,3)}$ 为 A 的一个$\{1,3\}$逆.

(4) $\boldsymbol{A}\{1,4\}=\{\boldsymbol{G}|\boldsymbol{G}=\boldsymbol{A}^{(1,4)}+\boldsymbol{U}(\boldsymbol{E}-\boldsymbol{A}\boldsymbol{A}^{(1,4)}),\forall \boldsymbol{U}\in\mathbf{C}^{n\times m}\}$，其中 $\boldsymbol{A}^{(1,4)}$ 为 $\boldsymbol{A}$ 的一个$\{1,4\}$逆.

三、典型例题

例 1　设 $\boldsymbol{A}\in\mathbf{R}^{m\times m}$,$\boldsymbol{C}\in\mathbf{R}^{n\times n}$ 都是可逆矩阵，证明：若矩阵 $\boldsymbol{B}\in\mathbf{R}^{m\times n}$ 是左(右)可逆的，则矩阵 $\boldsymbol{ABC}$ 也是左(右)可逆的.

证　事实上，若矩阵 $\boldsymbol{B}$ 是左可逆的，设 $\boldsymbol{B}_{\mathrm{L}}^{-1}$ 为它的一个左逆矩阵，因为 $\boldsymbol{A},\boldsymbol{C}$ 都是可逆矩阵，由$(\boldsymbol{C}^{-1}\boldsymbol{B}_{\mathrm{L}}^{-1}\boldsymbol{A}^{-1})\boldsymbol{ABC}=\boldsymbol{C}^{-1}\boldsymbol{B}_{\mathrm{L}}^{-1}\boldsymbol{A}^{-1}\boldsymbol{ABC}=\boldsymbol{C}^{-1}\boldsymbol{B}_{\mathrm{L}}^{-1}\boldsymbol{BC}=\boldsymbol{C}^{-1}\boldsymbol{C}=\boldsymbol{E}_n$ 易知，$\boldsymbol{C}^{-1}\boldsymbol{B}_{\mathrm{L}}^{-1}\boldsymbol{A}^{-1}$就是 $\boldsymbol{ABC}$ 的一个左逆矩阵.

同理可证,若矩阵 $\boldsymbol{B}$ 是右可逆的，则 $\boldsymbol{C}^{-1}\boldsymbol{B}_{\mathrm{R}}^{-1}\boldsymbol{A}^{-1}$为 $\boldsymbol{ABC}$ 的一个右逆矩阵，其中 $\boldsymbol{B}_{\mathrm{R}}^{-1}$ 为 $\boldsymbol{B}$ 的一个右逆矩阵.

例 2　设 $\boldsymbol{A}=\begin{pmatrix}1&2&1\\0&1&-2\end{pmatrix}$，求 $\boldsymbol{A}^{-}$.

解　将 $\boldsymbol{A}$ 通过初等行变换和列变换化为标准形,得

$$\begin{pmatrix}\boldsymbol{A}&\boldsymbol{E}_2\\\boldsymbol{E}_3&\boldsymbol{0}\end{pmatrix}=\begin{pmatrix}1&2&1&1&0\\0&1&-2&0&1\\1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0&1&-2\\0&1&0&0&1\\1&0&-5&0&0\\0&1&2&0&0\\0&0&1&0&0\end{pmatrix}.$$

故有

$$\boldsymbol{PAQ}=\begin{pmatrix}1&-2\\0&1\end{pmatrix}\begin{pmatrix}1&2&1\\0&1&-2\end{pmatrix}\begin{pmatrix}1&0&-5\\0&1&2\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\end{pmatrix}=\boldsymbol{B},$$

易知 $\boldsymbol{B}$ 的减号广义逆为 $\boldsymbol{B}^{-}=\begin{pmatrix}1&0\\0&1\\\mu&\nu\end{pmatrix}$(其中 μ,ν 为任意实数)，由此得 $\boldsymbol{A}$ 的减号广义逆

$$\boldsymbol{A}^{-}=\boldsymbol{QB}^{-}\boldsymbol{P}=\begin{pmatrix}1&0&-5\\0&1&2\\0&0&1\end{pmatrix}\begin{pmatrix}1&0\\0&1\\\mu&\nu\end{pmatrix}\begin{pmatrix}1&-2\\0&1\end{pmatrix}=\begin{pmatrix}1-5\mu&10\mu-5\nu-2\\2\mu&-4\mu+2\nu+1\\\mu&-2\mu+\nu\end{pmatrix},$$

其中 μ,ν 为任意实数.

若取 μ,ν 为 0，则得到一个减号广义逆

$$\boldsymbol{A}^{-}=\boldsymbol{Q}\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}\boldsymbol{P}=\begin{pmatrix}1&0&-5\\0&1&2\\0&0&1\end{pmatrix}\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}\begin{pmatrix}1&-2\\0&1\end{pmatrix}=\begin{pmatrix}1&-2\\0&1\\0&0\end{pmatrix}.$$

例 3 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$,rank$(\boldsymbol{A})=r$，若有 m 阶可逆矩阵 $\boldsymbol{P}$ 和 n 阶置换矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{S}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix},\quad \boldsymbol{S}\in\mathbf{C}^{(m-r)\times(n-r)}.$$

试证：对任给 $\boldsymbol{L}\in\mathbf{C}^{(n-r)\times(m-r)}$，矩阵 $\boldsymbol{G}=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{L}\end{pmatrix}\boldsymbol{P}$ 是 $\boldsymbol{A}$ 的一个广义逆，若 $\boldsymbol{L}=\boldsymbol{0}$，则相应的 $\boldsymbol{G}$ 是 $\boldsymbol{A}$ 的一个自反广义逆.

证 因为 $\boldsymbol{A}=\boldsymbol{P}^{-1}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{S}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{Q}^{-1}$，所以

$$\begin{aligned}\boldsymbol{AGA}&=\boldsymbol{P}^{-1}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{S}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{Q}^{-1}\boldsymbol{Q}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{L}\end{pmatrix}\boldsymbol{P}\boldsymbol{P}^{-1}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{S}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{Q}^{-1}\\&=\boldsymbol{P}^{-1}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{S}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{L}\end{pmatrix}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{S}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{Q}^{-1}=\boldsymbol{P}^{-1}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{S}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{Q}^{-1}=\boldsymbol{A},\end{aligned}$$

故 $\boldsymbol{G}$ 是 $\boldsymbol{A}$ 的广义逆. 当 $\boldsymbol{L}=\boldsymbol{0}$ 时，有

$$\boldsymbol{GAG}=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{P}\boldsymbol{P}^{-1}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{S}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{Q}^{-1}\boldsymbol{Q}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{P}=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{P}=\boldsymbol{G},$$

从而 $\boldsymbol{G}$ 是 $\boldsymbol{A}$ 的自反广义逆.

例 4 设 $\boldsymbol{A}$ 是幂等阵且是 Hermite 阵(即正交投影阵)，证明 $\boldsymbol{A}^{+}=\boldsymbol{A}$.

证 设 $\boldsymbol{A}$ 的秩为 r. 因 $\boldsymbol{A}$ 是幂等阵，$\boldsymbol{A}^2=\boldsymbol{A}$，则 $\boldsymbol{A}$ 的特征值为 0 或 1，又 $\boldsymbol{A}$ 是 Hermite 阵，$\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$，则 $\boldsymbol{A}$ 可酉相似于对角矩阵，即存在酉矩阵 $\boldsymbol{U}$，使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{AU}=\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix},\quad \boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}.$$

利用 $\boldsymbol{U}$ 为酉矩阵，可得

$$\boldsymbol{A}^{+}=\left(\boldsymbol{U}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}\right)^{+}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}^{+}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{E}_r & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{A}.$$

例 5 设 $\boldsymbol{A}=\begin{bmatrix}2 & 1 & 0\\ -4 & -1 & 2\\ 3 & 2 & 1\end{bmatrix}$，求 $\boldsymbol{A}^{+}$.

解 对 $\boldsymbol{A}$ 实施初等行变换，得

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

于是 $\boldsymbol{A}$ 的最大秩分解为

$$\boldsymbol{A}=\begin{bmatrix}2 & 1\\ -4 & -1\\ 3 & 2\end{bmatrix}\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 2\end{pmatrix}=\boldsymbol{BC}.$$

而

$$CC^{\mathrm{T}}=\begin{pmatrix}1&0&-1\\0&1&2\end{pmatrix}\begin{pmatrix}1&0\\0&1\\-1&2\end{pmatrix}=\begin{pmatrix}2&-2\\-2&5\end{pmatrix},$$

$$B^{\mathrm{T}}B=\begin{pmatrix}2&-4&3\\1&-1&2\end{pmatrix}\begin{pmatrix}2&1\\-4&-1\\3&2\end{pmatrix}=\begin{pmatrix}29&12\\12&6\end{pmatrix},$$

易得

$$(CC^{\mathrm{T}})^{-1}=\begin{pmatrix}5/6&1/3\\1/3&1/3\end{pmatrix},(B^{\mathrm{T}}B)^{-1}=\begin{pmatrix}1/5&-2/5\\-2/5&29/30\end{pmatrix}.$$

于是

$$A^{+}=C^{\mathrm{T}}(CC^{\mathrm{T}})^{-1}(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}=\begin{pmatrix}1/18&-11/90&7/90\\1/18&7/90&8/45\\1/18&5/18&5/18\end{pmatrix}.$$

例 6 设 $A\in\mathbf{C}^{m\times n}$，$P,Q$ 分别是 m 阶和 n 阶可逆矩阵.

(1) 证明 $Q^{-1}A^{-}P^{-1}\in(PAQ)\{1\}$；

(2) 举例说明 $(PAQ)^{+}=Q^{-1}A^{+}P^{-1}$ 不真.

证 (1) 由 $(PAQ)(Q^{-1}A^{-}P^{-1})(PAQ)=PAA^{-}AQ=PAQ$，知 $Q^{-1}A^{-}P^{-1}\in(PAQ)\{1\}$；

(2) 如取 $A=\begin{pmatrix}1\\2\end{pmatrix}$，$P=\begin{pmatrix}1&3\\-1&2\end{pmatrix}$，$Q=(2)$，则有 $A^{+}=\begin{pmatrix}\frac{1}{5}&\frac{2}{5}\end{pmatrix}$，$PAQ=\begin{pmatrix}14\\6\end{pmatrix}$，$P^{-1}=\begin{pmatrix}2/5&-3/5\\1/5&1/5\end{pmatrix}$，$(PAQ)^{+}=\begin{pmatrix}\frac{7}{116}&\frac{3}{116}\end{pmatrix}$，$Q^{-1}A^{+}P^{-1}=\begin{pmatrix}\frac{2}{25}&-\frac{1}{50}\end{pmatrix}$，可见 $(PAQ)^{+}\neq Q^{-1}A^{+}P^{-1}$.

例 7 求方程组 $\begin{cases}x_1+2x_2+3x_3=1,\\x_1+x_3=0,\\2x_1+2x_3=0,\\2x_1+4x_2+6x_3=2\end{cases}$ 的最小范数解和通解.

解 令

$$A=\begin{pmatrix}1&2&3\\1&0&1\\2&0&2\\2&4&6\end{pmatrix},\quad x=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\quad b=\begin{pmatrix}1\\0\\0\\2\end{pmatrix}.$$

则所给方程组可写为 $Ax=b$. 由于 $\mathrm{rank}(A)=\mathrm{rank}(A\ \vdots\ b)=2$，故方程组相容. 可证明 $G=A^{\mathrm{T}}(AA^{\mathrm{T}})^{-}\in A\{1,3\}$（利用 A 的满秩分解来验证 M-P 方程(1)和(3)). 计算可得

$$AA^{\mathrm{T}}=\begin{pmatrix}1&2&3\\1&0&1\\2&0&2\\2&4&6\end{pmatrix}\begin{pmatrix}1&1&2&2\\2&0&0&4\\3&1&2&6\end{pmatrix}=\begin{pmatrix}14&4&8&28\\4&2&4&8\\8&4&8&16\\28&8&16&56\end{pmatrix}.$$

先求得 AA^{T} 的一个减号广义逆

$$(AA^{\mathrm{T}})^{-}=\begin{pmatrix}1/6&-1/3&0&0\\-1/3&7/6&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}.$$

从而得到 A 的一个$\{1,3\}$广义逆

$$G=A^{\mathrm{T}}(AA^{\mathrm{T}})^{-}=\begin{pmatrix}1&1&2&2\\2&0&0&4\\3&1&2&6\end{pmatrix}\begin{pmatrix}1/6&-1/3&0&0\\-1/3&7/6&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

$$=\begin{pmatrix}-1/6&5/6&0&0\\1/3&-2/3&0&0\\1/6&1/6&0&0\end{pmatrix}.$$

故方程组的最小范数解为

$$x=Gb=\begin{pmatrix}-1/6&5/6&0&0\\1/3&-2/3&0&0\\1/6&1/6&0&0\end{pmatrix}\begin{pmatrix}1\\0\\0\\2\end{pmatrix}=\begin{pmatrix}-1/6\\1/3\\1/6\end{pmatrix}.$$

而方程组的通解为

$$x=Gb+(E-GA)u$$

$$=\begin{pmatrix}-1/6\\1/3\\1/6\end{pmatrix}+\left(\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\begin{pmatrix}-1/6&5/6&0&0\\1/3&-2/3&0&0\\1/6&1/6&0&0\end{pmatrix}\begin{pmatrix}1&2&3\\1&0&1\\2&0&2\\2&4&6\end{pmatrix}\right)\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}$$

$$=\frac{1}{3}\begin{pmatrix}-\frac{1}{2}+u_1+u_2-u_3\\1+u_1+u_2-u_3\\\frac{1}{2}-u_1-u_2+u_3\end{pmatrix}.$$

例 8 已知 $A=\begin{pmatrix}-1&2&1\\-1&2&1\\0&3&2\end{pmatrix}$，$b=\begin{pmatrix}-1\\1\\1\end{pmatrix}$，求方程组 $Ax=b$ 的最佳逼近解.

解 因为 $\mathrm{rank}(\boldsymbol{A})=2\neq 3=\mathrm{rank}(\boldsymbol{A}\,\vdots\,\boldsymbol{b})$，所以方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 是不相容的. 由矩阵的初等行变换有 $\boldsymbol{A}=\begin{pmatrix}-1&2&1\\-1&2&1\\0&3&2\end{pmatrix}\to\begin{pmatrix}1&0&1/3\\0&1&2/3\\0&0&0\end{pmatrix}$，取 $\boldsymbol{B}=\begin{pmatrix}-1&2\\-1&2\\0&3\end{pmatrix}$，$\boldsymbol{C}=\begin{pmatrix}1&0&1/3\\0&1&2/3\end{pmatrix}$，则 $\boldsymbol{A}=\boldsymbol{BC}$ 为 $\boldsymbol{A}$ 的最大秩分解. 计算

$$\boldsymbol{B}^{+}=(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}=\begin{pmatrix}-\frac{1}{2}&-\frac{1}{2}&\frac{2}{3}\\0&0&\frac{1}{3}\end{pmatrix},\quad \boldsymbol{C}^{+}=\boldsymbol{C}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}})^{-1}=\begin{pmatrix}13/14&-1/7\\-1/7&5/7\\3/14&3/7\end{pmatrix}.$$

则 $\boldsymbol{A}^{+}=\boldsymbol{C}^{+}\boldsymbol{B}^{+}=\begin{pmatrix}-13/28&-13/28&4/7\\1/14&1/14&1/7\\-3/28&-3/28&2/7\end{pmatrix}$. 所以方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最佳逼近解为

$$\boldsymbol{x}=\boldsymbol{A}^{+}\boldsymbol{b}=\begin{pmatrix}4/7\\1/7\\2/7\end{pmatrix}.$$

四、习题解答

1. 设矩阵 $\boldsymbol{A}$ 为

$$\boldsymbol{A}=\begin{pmatrix}2&3&1&-1\\5&8&0&1\\1&2&-2&3\end{pmatrix},$$

求广义逆矩阵 $\boldsymbol{A}^{-}$，$\boldsymbol{A}_{r}^{-}$.

解 用矩阵初等变换来求广义逆 $\boldsymbol{A}^{-}$.

$$\begin{pmatrix}2&3&1&-1&1&0&0\\5&8&0&1&0&1&0\\1&2&-2&3&0&0&1\end{pmatrix}\to\begin{pmatrix}1&2&-2&3&0&0&1\\2&3&1&-1&1&0&0\\0&0&0&0&-2&1&-1\end{pmatrix},$$

$$\to\begin{pmatrix}1&2&-2&3&0&0&1\\0&-1&5&-7&1&0&-2\\0&0&0&0&-2&1&-1\end{pmatrix}\to\begin{pmatrix}1&0&8&-11&2&0&-3\\0&1&-5&7&-1&0&2\\0&0&0&0&-2&1&-1\end{pmatrix},$$

$$\boldsymbol{PAQ}=\begin{pmatrix}2&0&-3\\-1&0&2\\-2&1&-1\end{pmatrix}\begin{pmatrix}2&3&1&1\\5&8&0&1\\1&2&-2&3\end{pmatrix}\boldsymbol{E}_4=\begin{pmatrix}1&0&8&-11\\0&1&-5&7\\0&0&0&0\end{pmatrix}=\boldsymbol{B}_1.$$

取 $B_1^- = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$，则 $A^- = QB_1^- P = E_4 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & -3 \\ -1 & 0 & 2 \\ -2 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -3 \\ -1 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

再用最大秩分解来求 A_r^-. 先用初等行变换化 A 为行标准形矩阵 $\widetilde{A}$，得

$$A = \begin{pmatrix} 2 & 3 & 1 & -1 \\ 5 & 8 & 0 & 1 \\ 1 & 2 & -2 & 3 \end{pmatrix} \to \widetilde{A} = \begin{pmatrix} 1 & 0 & 8 & -11 \\ 0 & 1 & -5 & 7 \\ 0 & 0 & 7 & 0 \end{pmatrix},$$

则 $A = BD = \begin{pmatrix} 2 & 3 \\ 5 & 8 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 8 & -11 \\ 0 & 1 & -5 & 7 \end{pmatrix}$为 A 的一个最大秩分解.

然后用初等行变换求 B 的单边逆. 由 $\begin{pmatrix} 2 & 3 & 1 & 0 & 0 \\ 5 & 8 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 & 0 & -3 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & -2 & 1 & -1 \end{pmatrix}$，得

$B_L^{-1} = \begin{pmatrix} 2 & 0 & -3 \\ -1 & 0 & 2 \end{pmatrix}$. 容易看出 $D_R^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$，于是

$$A_r^- = D_R^{-1} B_L^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & -3 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -3 \\ -1 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

2. 设 $A \in C^{n\times n}$，证明：总有广义逆矩阵 A^- 存在.

证 若 $A = 0_{m\times n}$，此时任给 $X \in C^{n\times m}$，都有 $0X0 = 0$，故 $X = A^-$.

若 $A \neq 0_{m\times n}$，设 $\operatorname{rank}(A) = r > 0$，则存在 m 阶可逆矩阵 P 与 n 阶可逆矩阵 Q，使得

$$A = P\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q.$$

设 $G = Q^{-1}\begin{pmatrix} E_r & X \\ Y & Z \end{pmatrix} P^{-1}$，其中 $X \in C^{r\times(m-r)}$，$Y \in C^{(n-r)\times r}$，$Z \in C^{(n-r)\times(m-r)}$ 为任意矩阵. 则

$$\begin{aligned} AGA &= P\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} QQ^{-1} \begin{pmatrix} E_r & X \\ Y & Z \end{pmatrix} P^{-1} P \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q \\ &= P\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_r & X \\ Y & Z \end{pmatrix} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q \\ &= P\begin{pmatrix} E_r & X \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q = P\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q = A, \end{aligned}$$

故 $G=A^-$.

3. 设 $A\in C^{m\times n}$，证明 $(A^-)^T\in A^T\{1\}$.

证 因为 $A^T\{1\}=\{G\mid A^TGA^T=A^T,\ \forall G\in C^{m\times n}\}$，而 $A^T(A^-)^TA^T=(AA^-A)^T$，故 $(A^-)^T\in A^T\{1\}$.

4. 设 $P\in C^{m\times m}$, $Q\in C^{n\times n}$ 均为可逆矩阵，且有 $B=PAQ$，证明：$Q^{-1}A^-P^{-1}\in B\{1\}$.

证 因为 $B(Q^{-1}A^-P^{-1})B=PAQQ^{-1}A^-P^{-1}PAQ=PAA^-AQ=PAQ=B$，所以 $Q^{-1}A^-P^{-1}\in B\{1\}$.

5. 证明：$0_{m\times n}$ 的自反广义逆矩阵仅为 $0_{n\times m}$.

证 $\forall G\in C^{n\times m}$，有 $0_{m\times n}G0_{m\times n}=0_{m\times n}$，可见 G 为 $0_{m\times n}$ 的广义逆矩阵. 要使 G 是 $0_{m\times n}$ 的自反广义逆矩阵，还需 $G0_{m\times n}G=G$ 成立，但 $G0_{m\times n}G=0_{n\times m}$，所以 $G=0_{n\times m}$.

6. 设矩阵 $A=\begin{pmatrix}1&0&2\\2&1&5\\0&1&-1\\1&3&-1\end{pmatrix}$，求 M-P 广义逆矩阵 A^+.

解 容易验证 rank$(A)=3$，A 为列满秩矩阵，所以 $A^+=(A^HA)^{-1}A^H$. 又

$$A^HA=\begin{pmatrix}1&2&0&1\\0&1&1&3\\2&5&-1&-1\end{pmatrix}\begin{pmatrix}1&0&2\\2&1&5\\0&1&-1\\1&3&-1\end{pmatrix}=\begin{pmatrix}6&5&11\\5&11&1\\11&1&31\end{pmatrix},$$

所以

$$A^+=(A^HA)^{-1}A^H=\begin{pmatrix}6&5&11\\5&11&1\\11&1&31\end{pmatrix}^{-1}\begin{pmatrix}1&2&0&1\\0&1&1&3\\2&5&-1&-1\end{pmatrix}$$

$$=\begin{pmatrix}\frac{85}{11}&-\frac{36}{11}&-\frac{29}{11}\\-\frac{36}{11}&\frac{65}{44}&\frac{49}{44}\\-\frac{29}{11}&\frac{49}{44}&\frac{41}{44}\end{pmatrix}\begin{pmatrix}1&2&0&1\\0&1&1&3\\2&5&-1&-1\end{pmatrix}$$

$$=\begin{pmatrix}\frac{27}{11}&-1&-\frac{7}{11}&\frac{6}{11}\\-\frac{23}{22}&\frac{1}{2}&\frac{4}{11}&\frac{1}{22}\\-\frac{17}{22}&\frac{1}{2}&\frac{2}{11}&-\frac{5}{22}\end{pmatrix}.$$

7. 设 $A\in C^{m\times n}$，$D,G\in A\{1\}$，试证明：$GAD\in A\{1,2\}$.

证 由 $D,G\in A\{1\}$，得 $ADA=A,AGA=A$，故有

$$A(GAD)A=(AGA)DA=ADA=A,$$
$$(GAD)A(GAD)=G(ADA)GAD=G(AGA)D=GAD.$$

所以 $GAD\in A\{1,2\}$.

8. 设 $A^2=A=A^H$，试证明：$A=A^+$.

证 由条件 $A^2=A=A^H$，可得

$$AAA=A^2A=A^2=A,\quad (AA)^H=A^HA^H=AA.$$

由以上两式易见矩阵 A 与它本身满足 M-P 广义逆定义的四个方程，所以 $A=A^+$.

9. 设 $A=A^H$，证明：

$(A^2)^+=(A^+)^2,\quad AA^+=A^+A,\quad A^+A^2=A^2A^+,\quad A^2(A^2)^+=(A^2)^+A^2=AA^+$.

证 (1) 利用已知 $A=A^H$ 和 M-P 广义逆 A^+ 的性质⑤，有

$$(A^2)^+=(AA^H)^+=(A^H)^+A^+=A^+A^+=(A^+)^2.$$

(2) 由 $A=A^H$ 和 A^+ 的性质，有 $AA^+=(AA^+)^H=(A^+)^HA^H=(A^H)^+A^H=A^+A$.

(3) 利用(2)的结论 $AA^+=A^+A$，有 $A^+A^2=A^+AA=AA^+A=AAA^+=A^2A^+$.

(4) 由(1)和(2)中结论，得 $A^2(A^2)^+=A^2(A^+)^2=AAA^+A^+=AA^+AA^+=AA^+$，同理可得 $(A^2)^+A^2=AA^+$.

10. 若 A 的最大秩分解为 $A=BC$，证明：$A^+=C^+B^+$.

证 $\forall A\in C_r^{m\times n}$，由已知 $A=BC$ 为 A 的最大秩分解，可得 $A^+=C^H(CC^H)^{-1}(B^HB)^{-1}B^H$. 由于 B 是列满秩矩阵，则 $B=BE_r$ 为 B 的最大秩分解，于是 $B^+=(B^HB)^{-1}B^H$，同理，由 C 是行满秩矩阵，得 $C^+=C^H(CC^H)^{-1}$，所以 $A^+=C^+B^+$.

11. 证明：$(A^+)^+=A$.

证 易知 A 与 A^+ 满足 M-P 广义逆定义的四个方程，故有 $(A^+)^+=A$.

12. 试证明：

$$(A^HA)^+=A^+(A^H)^+,\quad (AA^H)^+=(A^H)^+A^+,$$
$$(A^HA)^+=A^+(AA^H)^+A=A^H(AA^H)^+(A^H)^+,$$
$$AA^+=(AA^H)(AA^H)^+=(AA^H)^+(AA^H).$$

证 先证 $(A^HA)^+=A^+(A^H)^+$，$(AA^H)^+=(A^H)^+A^+$.

设 $A=BD$ 是 A 的最大秩分解，则 $A^+=D^H(DD^H)^{-1}(B^HB)^{-1}B^H$，且不难验证 $A^HA=D^H(B^HBD)$ 是 A^HA 的最大秩分解，故

$$\begin{aligned}(A^HA)^+&=(D^HB^HBD)^+\\&=(B^HBD)^H(B^HBDD^HB^HB)^{-1}(DD^H)^{-1}D\\&=D^HB^HB(B^HB)^{-1}(DD^H)^{-1}(B^HB)^{-1}(DD^H)^{-1}D\\&=[D^H(DD^H)^{-1}(B^HB)^{-1}B^H][B(B^HB)^{-1}(DD^H)^{-1}D]\\&=A^+(A^H)^+\end{aligned}$$

同理可证$(\boldsymbol{AA}^{\mathrm{H}})^{+}=(\boldsymbol{A}^{\mathrm{H}})^{+}\boldsymbol{A}^{+}$.

再证$(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{+}=\boldsymbol{A}^{+}(\boldsymbol{AA}^{\mathrm{H}})^{+}\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{+}(\boldsymbol{A}^{\mathrm{H}})^{+}$.

可验证,$\boldsymbol{AA}^{\mathrm{H}}=\boldsymbol{B}(\boldsymbol{DD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}})$是$\boldsymbol{AA}^{\mathrm{H}}$的最大秩分解,于是有

$$\begin{aligned}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{+}&=(\boldsymbol{BD})^{\mathrm{H}}(\boldsymbol{DD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}})^{\mathrm{H}}(\boldsymbol{DD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{BDD}^{\mathrm{H}})^{-1}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{H}}\\&=\boldsymbol{D}^{\mathrm{H}}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})(\boldsymbol{DD}^{\mathrm{H}})(\boldsymbol{DD}^{\mathrm{H}})^{-1}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}(\boldsymbol{DD}^{\mathrm{H}})^{-1}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{H}}\\&=\boldsymbol{D}^{\mathrm{H}}(\boldsymbol{DD}^{\mathrm{H}})^{-1}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{H}}\\&=\boldsymbol{A}^{+}\end{aligned}$$

$$\begin{aligned}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{+}&=\boldsymbol{A}^{+}(\boldsymbol{A}^{\mathrm{H}})^{+}=\boldsymbol{A}^{+}(\boldsymbol{A}^{+})^{\mathrm{H}}=\boldsymbol{A}^{+}[\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{+}]^{\mathrm{H}}\\&=\boldsymbol{A}^{+}[(\boldsymbol{AA}^{\mathrm{H}})^{+}]^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}^{+}(\boldsymbol{AA}^{\mathrm{H}})^{+}\boldsymbol{A},\end{aligned}$$

同理可证$(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{+}=\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{+}(\boldsymbol{A}^{\mathrm{H}})^{+}$.

最后证$\boldsymbol{AA}^{+}=(\boldsymbol{AA}^{\mathrm{H}})(\boldsymbol{AA}^{\mathrm{H}})^{+}=(\boldsymbol{AA}^{\mathrm{H}})^{+}(\boldsymbol{AA}^{\mathrm{H}})$.

$$\begin{aligned}(\boldsymbol{AA}^{\mathrm{H}})^{+}\boldsymbol{AA}^{\mathrm{H}}&=[\boldsymbol{B}(\boldsymbol{DD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}})]^{+}\boldsymbol{BDD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}\\&=(\boldsymbol{DD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}})^{\mathrm{H}}(\boldsymbol{DD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{BDD}^{\mathrm{H}})^{-1}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{BDD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}\\&=\boldsymbol{B}(\boldsymbol{DD}^{\mathrm{H}})(\boldsymbol{DD}^{\mathrm{H}})^{-1}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}(\boldsymbol{DD}^{\mathrm{H}})\boldsymbol{DD}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}\\&=\boldsymbol{BD}[\boldsymbol{D}^{\mathrm{H}}(\boldsymbol{DD}^{\mathrm{H}})^{-1}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{H}}]\\&=\boldsymbol{AA}^{+},\end{aligned}$$

故有$\boldsymbol{AA}^{+}=(\boldsymbol{AA}^{\mathrm{H}})^{+}\boldsymbol{AA}^{\mathrm{H}}$,同理可证$\boldsymbol{AA}^{+}=\boldsymbol{AA}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{+}$.

13. 设$\boldsymbol{U}\in\mathbf{C}^{m\times m}$与$\boldsymbol{V}\in\mathbf{C}^{n\times n}$均是酉矩阵,证明:$(\boldsymbol{UAV}^{\mathrm{H}})^{+}=\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}$.

证 利用$\boldsymbol{U},\boldsymbol{V}$为酉矩阵和$\boldsymbol{A}^{+}$的运算性质,容易验证$\boldsymbol{UAV}^{\mathrm{H}}$和$\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}$满足M-P广义逆定义的四个方程:

(1) $\boldsymbol{UAV}^{\mathrm{H}}\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{UAV}^{\mathrm{H}}=\boldsymbol{UAA}^{+}\boldsymbol{AV}^{\mathrm{H}}=\boldsymbol{UAV}^{\mathrm{H}}$;

(2) $\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{UAV}^{\mathrm{H}}\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{VA}^{+}\boldsymbol{AA}^{+}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}$;

(3) $(\boldsymbol{UAV}^{\mathrm{H}}\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}})^{\mathrm{H}}=(\boldsymbol{UAA}^{+}\boldsymbol{U}^{\mathrm{H}})^{\mathrm{H}}=\boldsymbol{U}(\boldsymbol{AA}^{+})^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{UAA}^{+}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{UAV}^{\mathrm{H}}\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}$;

(4) 类似(3)可得$(\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{UAV}^{\mathrm{H}})^{\mathrm{H}}=\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{UAV}^{\mathrm{H}}$.

所以$(\boldsymbol{UAV}^{\mathrm{H}})^{+}=\boldsymbol{VA}^{+}\boldsymbol{U}^{\mathrm{H}}$.

14. 若$\boldsymbol{A}$是正规矩阵,证明:(1) $\boldsymbol{A}^{+}\boldsymbol{A}=\boldsymbol{AA}^{+}$;(2) $(\boldsymbol{A}^{n})^{+}=(\boldsymbol{A}^{+})^{n}$.

证 设$\boldsymbol{A}$为m阶正规矩阵,则存在酉矩阵$\boldsymbol{U}$,使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{AU}=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_m\end{pmatrix},\quad \boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_m\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}.$$

由13题结论知

$$\boldsymbol{A}^{+}=\boldsymbol{U}\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_m\end{pmatrix}^{+}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}\begin{pmatrix}\lambda_1^{+}&&&\\&\lambda_2^{+}&&\\&&\ddots&\\&&&\lambda_m^{+}\end{pmatrix}\boldsymbol{U}^{\mathrm{H}},$$

其中 $\lambda_i^+=\begin{cases}1/\lambda_i, & \lambda_i\neq 0,\\ 0, & \lambda_i=0,\end{cases} i=1,2,\cdots,m.$

$$(1)\ \boldsymbol{A}^+\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\lambda_1^+ & & & \\ & \lambda_2^+ & & \\ & & \ddots & \\ & & & \lambda_m^+\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}$$

$$=\boldsymbol{U}\begin{pmatrix}\lambda_1^+\lambda_1 & & & \\ & \lambda_2^+\lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m^+\lambda_m\end{pmatrix}\boldsymbol{U}^{\mathrm{H}},$$

$$\boldsymbol{A}\boldsymbol{A}^+=\boldsymbol{U}\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\begin{pmatrix}\lambda_1^+ & & & \\ & \lambda_2^+ & & \\ & & \ddots & \\ & & & \lambda_m^+\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}$$

$$=\boldsymbol{U}\begin{pmatrix}\lambda_1^+\lambda_1 & & & \\ & \lambda_2^+\lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m^+\lambda_m\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}.$$

所以 $\boldsymbol{A}^+\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^+$.

$$(2)\ (\boldsymbol{A}^n)^+=\left(\boldsymbol{U}\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}\cdots\boldsymbol{U}\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}\right)^+$$

$$=\left(\boldsymbol{U}\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m\end{pmatrix}^n\boldsymbol{U}^{\mathrm{H}}\right)^+=\left(\boldsymbol{U}\begin{pmatrix}\lambda_1^n & & & \\ & \lambda_2^n & & \\ & & \ddots & \\ & & & \lambda_m^n\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}\right)^+$$

$$=\boldsymbol{U}\begin{pmatrix}(\lambda_1^+)^n & & & \\ & (\lambda_2^+)^n & & \\ & & \ddots & \\ & & & (\lambda_m^+)^n\end{pmatrix}\boldsymbol{U}^{\mathrm{H}}$$

$$=U\begin{pmatrix}\lambda_1^+ & & & \\ & \lambda_2^+ & & \\ & & \ddots & \\ & & & \lambda_m^+\end{pmatrix}U^HU\begin{pmatrix}\lambda_1^+ & & & \\ & \lambda_2^+ & & \\ & & \ddots & \\ & & & \lambda_m^+\end{pmatrix}U^H\cdots U\begin{pmatrix}\lambda_1^+ & & & \\ & \lambda_2^+ & & \\ & & \ddots & \\ & & & \lambda_m^+\end{pmatrix}U^H$$

$$=(A^+)^n.$$

15. 若 $ABA=A,(BA)^H=BA,AGA=A,(AG)^H=AG$，则 $BAG=A^+$.

证 由已知条件验证 BAG 满足 M-P 广义逆定义的四个方程：

(1) $A(BAG)A=(ABA)GA=AGA=A$;

(2) $(BAG)A(BAG)=B(AGA)BAG=BABAG=BAG$;

(3) $(ABAG)^H=(AG)^H=AG=ABAG$;

(4) $(BAGA)^H=(BA)^H=BA=BAGA$.

所以 $BAG=A^+$.

16. 试证明：$(A\otimes B)^+=A^+\otimes B^+$.

证 根据 Kronecker 乘积的性质有

$$(A\otimes B)(A^+\otimes B^+)(A\otimes B)=(AA^+A)\otimes(BB^+B)=A\otimes B;$$

$$(A^+\otimes B^+)(A\otimes B)(A^+\otimes B^+)=(A^+AA^+)\otimes(B^+BB^+)=A^+\otimes B^+;$$

$$[(A\otimes B)(A^+\otimes B^+)]^H=[(AA^+)\otimes(BB^+)]^H=(AA^+)^H\otimes(BB^+)^H$$
$$=(AA^+)\otimes(BB^+)=(A\otimes B)(A^+\otimes B^+);$$

同理有

$$[(A^+\otimes B^+)(A\otimes B)]^H=(A^+\otimes B^+)(A\otimes B).$$

所以 $(A\otimes B)^+=A^+\otimes B^+$.

17. 已知下列矩阵 A，试用各种方法求 A^+.

(1) $A=\begin{pmatrix}1 & 2\\ 0 & 0\\ 2 & 4\end{pmatrix}$;　(2) $A=\begin{pmatrix}1 & 2 & -1\\ 0 & -1 & 2\end{pmatrix}$;　(3) $A=\begin{pmatrix}\mathrm{i} & 0\\ 1 & \mathrm{i}\\ 0 & 1\end{pmatrix}$;

(4) $A=\begin{pmatrix}1 & 2 & 0\\ 0 & 0 & 2\\ 2 & 4 & 0\end{pmatrix}$;　(5) $A=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 1 & 0 & 0\\ 2 & 1 & -1\end{pmatrix}$.

解 (1) 奇异值分解法

$$AA^H=\begin{pmatrix}1 & 2\\ 0 & 0\\ 2 & 4\end{pmatrix}\begin{pmatrix}1 & 0 & 2\\ 2 & 0 & 4\end{pmatrix}=\begin{pmatrix}5 & 0 & 10\\ 0 & 0 & 0\\ 10 & 0 & 20\end{pmatrix},$$

由 $|\lambda E_3-AA^H|=\begin{vmatrix}\lambda-5 & 0 & -10\\ 0 & \lambda & 0\\ -10 & 0 & \lambda-20\end{vmatrix}=\lambda^2(\lambda-25)=0$，得 AA^H 的特征值 $\lambda_1=25,\lambda_2=0$. 而对

应于 $\lambda_1=25$ 的单位特征向量为 $\boldsymbol{\alpha}_1=\left(\frac{1}{\sqrt{5}},0,\frac{2}{\sqrt{5}}\right)^{\mathrm{T}}$，故 $U_1=(\boldsymbol{\alpha}_1)$.

$$\boldsymbol{A}^{+}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}_1\Delta_r^{-1}\boldsymbol{U}_1^{\mathrm{H}}=\begin{pmatrix}1&0&2\\2&0&4\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{5}}\\0\\\frac{2}{\sqrt{5}}\end{pmatrix}\times\frac{1}{25}\times\left(\frac{1}{\sqrt{5}},0,\frac{2}{\sqrt{5}}\right)=\begin{pmatrix}\frac{1}{25}&0&\frac{2}{25}\\\frac{2}{25}&0&\frac{4}{25}\end{pmatrix}.$$

(2) 最大秩分解法显然 $\boldsymbol{A}$ 是行满秩矩阵，则

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\begin{pmatrix}1&2&-1\\0&-1&2\end{pmatrix}\begin{pmatrix}1&0\\2&-1\\-1&2\end{pmatrix}=\begin{pmatrix}6&-4\\-4&5\end{pmatrix},\quad 易得(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})^{-1}=\frac{1}{14}\begin{pmatrix}5&4\\4&6\end{pmatrix}.$$

所以

$$\boldsymbol{A}^{+}=\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})^{-1}=\begin{pmatrix}1&0\\2&-1\\-1&2\end{pmatrix}\cdot\frac{1}{14}\begin{pmatrix}5&4\\4&6\end{pmatrix}=\frac{1}{14}\begin{pmatrix}5&4\\6&2\\3&8\end{pmatrix}.$$

(3) 极限算法

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\begin{pmatrix}-\mathrm{i}&1&0\\0&-\mathrm{i}&1\end{pmatrix}\begin{pmatrix}\mathrm{i}&0\\1&\mathrm{i}\\0&1\end{pmatrix}=\begin{pmatrix}2&\mathrm{i}\\-\mathrm{i}&2\end{pmatrix},\quad \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}+\delta^2\boldsymbol{E}_2=\begin{pmatrix}2+\delta^2&\mathrm{i}\\-\mathrm{i}&2+\delta^2\end{pmatrix}.$$

容易算得

$$(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}+\delta^2\boldsymbol{E}_2)^{-1}=\frac{1}{(\delta^2+1)(\delta^2+3)}\begin{pmatrix}\delta^2+2&-\mathrm{i}\\\mathrm{i}&\delta^2+2\end{pmatrix},$$

于是

$$\begin{aligned}\boldsymbol{A}^{+}&=\lim_{\delta\to0}(\boldsymbol{A}^{\mathrm{H}}A+\delta^2\boldsymbol{E}_2)^{-1}\boldsymbol{A}^{\mathrm{H}}\\&=\lim_{\delta\to0}\frac{1}{(\delta^2+1)(\delta^2+3)}\begin{pmatrix}\delta^2+2&-\mathrm{i}\\\mathrm{i}&\delta^2+2\end{pmatrix}\begin{pmatrix}-\mathrm{i}&1&0\\0&-\mathrm{i}&1\end{pmatrix}\\&=\lim_{\delta\to0}\frac{1}{(\delta^2+1)(\delta^2+3)}\begin{pmatrix}-\mathrm{i}(\delta^2+2)&\delta^2+1&-\mathrm{i}\\1&-\mathrm{i}(\delta^2+1)&\delta^2+2\end{pmatrix}\\&=\begin{pmatrix}-\frac{2\mathrm{i}}{3}&\frac{1}{3}&-\frac{\mathrm{i}}{3}\\\frac{1}{3}&-\frac{\mathrm{i}}{3}&\frac{2}{3}\end{pmatrix}.\end{aligned}$$

(4) 谱分解法

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\begin{pmatrix}1&0&2\\2&0&4\\0&2&0\end{pmatrix}\begin{pmatrix}1&2&0\\0&0&2\\2&4&0\end{pmatrix}=\begin{pmatrix}5&10&0\\10&20&0\\0&0&4\end{pmatrix},$$

$$|\lambda \boldsymbol{E}_3 - \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}| = \begin{vmatrix} \lambda-5 & -10 & 0 \\ -10 & \lambda-20 & 0 \\ 0 & 0 & \lambda-4 \end{vmatrix} = \lambda(\lambda-4)(\lambda-25)=0,$$

故 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值为 $\lambda_1=4,\lambda_2=25,\lambda_3=0$. 于是 $\lambda_1^-=\frac{1}{4},\lambda_2^-=\frac{1}{25},\lambda_3^-=0$.

设

$$P_1(\lambda)=(\lambda-\lambda_2)(\lambda-\lambda_3)=\lambda(\lambda-25),$$
$$P_2(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_3)=\lambda(\lambda-4),$$
$$P_3(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)=(\lambda-4)(\lambda-25).$$

可求得

$$P_1(\lambda_1)=4(4-25)=-84, P_2(\lambda_2)=25(25-4)=525,$$
$$P_1(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}-25\boldsymbol{E}_3)\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A},\quad P_2(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}-4\boldsymbol{E}_3)\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}.$$

于是

$$\boldsymbol{A}^+=\left[\frac{1}{4}\times\frac{(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}-25\boldsymbol{E}_3)\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}{-84}+\frac{1}{25}\times\frac{(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}-4\boldsymbol{E}_3)\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}}{525}\right]\boldsymbol{A}^{\mathrm{H}}$$

$$=\begin{pmatrix} \frac{1}{25} & 0 & \frac{2}{25} \\ \frac{2}{25} & 0 & \frac{4}{25} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

(5) 最大秩分解法

容易得到 $\boldsymbol{A}$ 的一个满秩分解表达式

$$\boldsymbol{A}=\boldsymbol{BC}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix},$$

$$\boldsymbol{CC}^{\mathrm{T}}=\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix},\quad \boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=\begin{pmatrix} 6 & 2 \\ 2 & 2 \end{pmatrix},$$

$$(\boldsymbol{CC}^{\mathrm{T}})^{-1}=\frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},\quad (\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}=\frac{1}{4}\begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}.$$

于是

$$\boldsymbol{A}^+=\boldsymbol{C}^{\mathrm{T}}(\boldsymbol{CC}^{\mathrm{T}})^{-1}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}=\frac{1}{8}\begin{pmatrix} 2 & -2 & 2 & 2 \\ -1 & 3 & -1 & 1 \\ 1 & -3 & 1 & -1 \end{pmatrix}.$$

18. 证明：方程组 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}$ 是相容的，其中 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{b}\in\mathbf{C}^{m}$.

证 方程 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}$ 相容的充要条件是 $\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\mid\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b})=\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$，显然 $\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\mid\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b})\geqslant\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$，同时有

$$\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\mid\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b})=\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{A}\mid\boldsymbol{b}))\leqslant\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}})=\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}),$$

所以 $\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\mid\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b})=\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$，方程 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}$ 相容.

19. 已知

$$\boldsymbol{A}=\begin{pmatrix}1&2\\0&0\\2&4\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}1\\0\\2\end{pmatrix},$$

求方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的通解及最小范数解.

解 显然 $\operatorname{rank}(\boldsymbol{A}\mid\boldsymbol{b})=\operatorname{rank}(\boldsymbol{A})=2$，方程组是相容的. 容易得到 $\boldsymbol{A}$ 的一个最大秩分解为 $\boldsymbol{A}=(1\quad 0\quad 2)^{\mathrm{T}}(1\quad 2)=\boldsymbol{B}\boldsymbol{D}$，而

$$\boldsymbol{B}^{+}=(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}=\left[(1\quad 0\quad 2)\begin{pmatrix}1\\0\\2\end{pmatrix}\right]^{-1}(1\quad 0\quad 2)=\begin{pmatrix}\frac{1}{5}&0&\frac{2}{5}\end{pmatrix},$$

$$\boldsymbol{D}^{+}=\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{D}^{\mathrm{T}})^{-1}=\begin{pmatrix}1\\2\end{pmatrix}\left((1\quad 2)\begin{pmatrix}1\\2\end{pmatrix}\right)^{-1}=\begin{pmatrix}\frac{1}{5}\\\frac{2}{5}\end{pmatrix}.$$

所以

$$\boldsymbol{A}^{+}=\boldsymbol{D}^{+}\boldsymbol{B}^{+}=\begin{pmatrix}\frac{1}{5}\\\frac{2}{5}\end{pmatrix}\begin{pmatrix}\frac{1}{5}&0&\frac{2}{5}\end{pmatrix}=\begin{pmatrix}\frac{1}{25}&0&\frac{2}{25}\\\frac{2}{25}&0&\frac{4}{25}\end{pmatrix}.$$

方程组的通解为

$$\begin{aligned}\boldsymbol{x}&=\boldsymbol{A}^{+}\boldsymbol{b}+(\boldsymbol{E}_2-\boldsymbol{A}^{+}\boldsymbol{A})\begin{pmatrix}u_1\\u_2\end{pmatrix}\\&=\begin{pmatrix}\frac{1}{25}&0&\frac{2}{25}\\\frac{2}{25}&0&\frac{4}{25}\end{pmatrix}\begin{pmatrix}1\\0\\2\end{pmatrix}+\left[\begin{pmatrix}1&0\\0&1\end{pmatrix}-\begin{pmatrix}\frac{1}{25}&0&\frac{2}{25}\\\frac{2}{25}&0&\frac{4}{25}\end{pmatrix}\begin{pmatrix}1&2\\0&0\\2&4\end{pmatrix}\right]\begin{pmatrix}u_1\\u_2\end{pmatrix}\\&=\begin{pmatrix}\frac{1}{5}\\\frac{2}{5}\end{pmatrix}+\begin{pmatrix}\frac{4}{5}&-\frac{2}{5}\\-\frac{2}{5}&\frac{1}{5}\end{pmatrix}\begin{pmatrix}u_1\\u_2\end{pmatrix}=\begin{pmatrix}\frac{1}{5}\\\frac{2}{5}\end{pmatrix}+\begin{pmatrix}\frac{4}{5}\\-\frac{2}{5}\end{pmatrix}u_1+\begin{pmatrix}-\frac{2}{5}\\\frac{1}{5}\end{pmatrix}u_2,\quad u_1,u_2\in\boldsymbol{R}.\end{aligned}$$

最小范数解为 $\boldsymbol{A}^{+}\boldsymbol{b}=\begin{pmatrix}\frac{1}{5}&\frac{2}{5}\end{pmatrix}^{\mathrm{T}}$.

20. 验证下列方程组是不相容的，并用 $\boldsymbol{A}^+$ 求它的最佳逼近解.

(1) $\begin{pmatrix}0&0&2\\1&1&0\\0&0&1\\1&1&1\end{pmatrix}\boldsymbol{x}=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$； (2) $\begin{pmatrix}0&2\mathrm{i}&\mathrm{i}&0&4+2\mathrm{i}&1\\0&0&0&-3&-6&-2-3\mathrm{i}\\0&2&1&1&4-6\mathrm{i}&1\end{pmatrix}\boldsymbol{x}=\begin{pmatrix}-\mathrm{i}\\1\\1\end{pmatrix}$.

解 (1) 显然 $3=\operatorname{rank}(\boldsymbol{A}\mid\boldsymbol{b})\neq\operatorname{rank}(\boldsymbol{A})=2$，故方程组不相容. 将矩阵 $\boldsymbol{A}$ 通过初等行变换变为 $\widetilde{\boldsymbol{A}}$，得

$$\boldsymbol{A}=\begin{pmatrix}0&0&2\\1&1&0\\0&0&1\\1&1&1\end{pmatrix}\rightarrow\widetilde{\boldsymbol{A}}=\begin{pmatrix}1&1&0\\0&0&1\\0&0&0\\0&0&0\end{pmatrix}.$$

故可得 $\boldsymbol{A}$ 的最大秩分解 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C}=\begin{pmatrix}0&2\\1&0\\0&1\\1&1\end{pmatrix}\begin{pmatrix}1&1&0\\0&0&1\end{pmatrix}$. 又

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=\begin{pmatrix}2&1\\1&6\end{pmatrix},\quad \boldsymbol{C}\boldsymbol{C}^{\mathrm{T}}=\begin{pmatrix}2&0\\0&1\end{pmatrix},$$

$$(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}=\frac{1}{11}\begin{pmatrix}6&-1\\-1&2\end{pmatrix},\quad (\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}})^{-1}=\begin{pmatrix}\frac{1}{2}&0\\0&1\end{pmatrix}.$$

所以

$$\begin{aligned}\boldsymbol{A}^+&=\boldsymbol{C}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}})^{-1}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\\&=\frac{1}{11}\begin{pmatrix}1&0\\1&0\\0&1\end{pmatrix}\begin{pmatrix}\frac{1}{2}&0\\0&1\end{pmatrix}\begin{pmatrix}6&-1\\-1&2\end{pmatrix}\begin{pmatrix}0&1&0&1\\2&0&1&1\end{pmatrix}\\&=\frac{1}{22}\begin{pmatrix}-2&6&-1&5\\-2&6&-1&5\\8&-2&4&2\end{pmatrix}.\end{aligned}$$

最佳逼近解为

$$\boldsymbol{x}_0=\boldsymbol{A}^+\boldsymbol{b}=\frac{1}{22}\begin{pmatrix}-2&6&-1&5\\-2&6&-1&5\\8&-2&4&2\end{pmatrix}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}=\frac{1}{11}\begin{pmatrix}4\\4\\6\end{pmatrix}.$$

21. 已知

$$\boldsymbol{A}=\begin{pmatrix}1&0&2\\2&0&4\end{pmatrix}^{\mathrm{T}},\quad \boldsymbol{b}=(0,1,0)^{\mathrm{T}},$$

求方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解和最佳逼近解.

解 因 $2=\text{rank}(\boldsymbol{A} \mid \boldsymbol{b})\neq\text{rank}(\boldsymbol{A})=1$，故方程组不相容. 易得 $\boldsymbol{A}$ 的一个最大秩分解为

$$\boldsymbol{A}=\boldsymbol{BC}=\begin{pmatrix}1\\0\\2\end{pmatrix}(1\quad 2),$$

所以

$$\boldsymbol{A}^{+}=\boldsymbol{C}^{\mathrm{T}}(\boldsymbol{CC}^{\mathrm{T}})^{-1}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}=\begin{pmatrix}1\\2\end{pmatrix}\cdot\frac{1}{5}\cdot\frac{1}{5}(1\quad 0\quad 2)=\frac{1}{25}\begin{pmatrix}1&0&2\\2&0&4\end{pmatrix},$$

最佳逼近解为

$$\boldsymbol{x}_0=\boldsymbol{A}^{+}\boldsymbol{b}=\frac{1}{25}\begin{pmatrix}1&0&2\\2&0&4\end{pmatrix}\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}.$$

方程组的最小二乘解为

$$\begin{aligned}\boldsymbol{x}&=\boldsymbol{A}^{+}\boldsymbol{b}+(\boldsymbol{E}_2-\boldsymbol{A}^{+}\boldsymbol{A})\begin{pmatrix}u_1\\u_2\end{pmatrix}\\&=\begin{pmatrix}0\\0\end{pmatrix}+\left(\begin{pmatrix}1&0\\0&1\end{pmatrix}-\begin{pmatrix}\frac{1}{25}&0&\frac{2}{25}\\\frac{2}{25}&0&\frac{4}{25}\end{pmatrix}\begin{pmatrix}1&2\\0&0\\2&4\end{pmatrix}\right)\begin{pmatrix}u_1\\u_2\end{pmatrix}\\&=\begin{pmatrix}\frac{4}{5}&-\frac{2}{5}\\-\frac{2}{5}&\frac{1}{5}\end{pmatrix}\begin{pmatrix}u_1\\u_2\end{pmatrix}=\begin{pmatrix}\frac{4}{5}\\-\frac{2}{5}\end{pmatrix}u_1+\begin{pmatrix}-\frac{2}{5}\\\frac{1}{5}\end{pmatrix}u_2,\quad u_1,u_2\in\mathbf{R}.\end{aligned}$$

22. 设 $\boldsymbol{A}$ 是对称矩阵，$\boldsymbol{M}=\boldsymbol{A}^{+}$，证明：$\boldsymbol{M}^2=(\boldsymbol{A}^2)^{+}$.

证 略，见第 9 题.

23. 设 $\boldsymbol{A}\in\mathbf{C}^{m\times m}$ 和 $\boldsymbol{B}\in\mathbf{C}^{n\times n}$ 均可逆，证明：

(1) 若 $\boldsymbol{D}\in\mathbf{C}^{m\times n}$ 是左可逆的，则 $\boldsymbol{ADB}$ 是左可逆的；

(2) 若 $\boldsymbol{D}\in\mathbf{C}^{m\times n}$ 是右可逆的，则 $\boldsymbol{ADB}$ 是右可逆的.

证 (1) 若 $\boldsymbol{D}$ 左可逆，则存在 $\boldsymbol{D}_{\mathrm{L}}^{-1}\in\mathbf{C}^{n\times m}$，使得 $\boldsymbol{D}_{\mathrm{L}}^{-1}\boldsymbol{D}=\boldsymbol{E}_n$，由于 $\boldsymbol{A},\boldsymbol{B}$ 均可逆，故$(\boldsymbol{B}^{-1}\boldsymbol{D}_{\mathrm{L}}^{-1}\boldsymbol{A}^{-1})\boldsymbol{ADB}=\boldsymbol{E}_n$. 所以 $\boldsymbol{ADB}$ 是左可逆的，且$(\boldsymbol{ADB})_{\mathrm{L}}^{-1}=\boldsymbol{B}^{-1}\boldsymbol{D}_{\mathrm{L}}^{-1}\boldsymbol{A}^{-1}$.

(2) 类似(1)可证得$(\boldsymbol{ADB})_{\mathrm{R}}^{-1}=\boldsymbol{B}^{-1}\boldsymbol{D}_{\mathrm{R}}^{-1}\boldsymbol{A}^{-1}$.

24. 已知 $\boldsymbol{A}_1=\begin{pmatrix}1&0&0\\0&0&0\end{pmatrix}$ 和 $\boldsymbol{A}_2=\begin{pmatrix}1&0\\0&0\\0&0\end{pmatrix}$，求 $\boldsymbol{A}_1^{+}$ 和 $\boldsymbol{A}_2^{+}$.

解 (1) $\boldsymbol{A}_1$ 的最大秩分解为 $\boldsymbol{A}_1=\begin{pmatrix}1\\0\end{pmatrix}(1\quad 0\quad 0)=\boldsymbol{BC}$，所以

$$A_1^+ = C^T(CC^T)^{-1}(B^TB)^{-1}B^T = \begin{pmatrix}1\\0\\0\end{pmatrix}\cdot 1\cdot 1\cdot(1\quad 0) = \begin{pmatrix}1&0\\0&0\\0&0\end{pmatrix}.$$

(2) 因为 $A_2 = A_1^T$，所以 $A_2^+ = (A_1^T)^+ = (A_1^+)^T = \begin{pmatrix}1&0&0\\0&0&0\end{pmatrix}$.

25. 已知一组数据：(−3,9),(−2,6),(0,2),(1,1),求数据拟合的最佳二次抛物线，并计算误差.

解 本题实际上是要求参数 β_i，使函数 $y=\beta_0+\beta_1x+\beta_2x^2$ 最佳拟合数据点(−3,9)，(−2,6)，(0,2),(1,1). 也即求方程组

$$A\beta = \begin{pmatrix}1&-3&9\\1&-2&4\\1&0&0\\1&1&1\end{pmatrix}\begin{pmatrix}\beta_0\\\beta_1\\\beta_2\end{pmatrix} = \begin{pmatrix}9\\6\\2\\1\end{pmatrix} = b$$

的最佳逼近解.

因系数矩阵 A 是列满秩的，求得 $(A^TA)^{-1} = \frac{1}{90}\begin{pmatrix}54&-21&-15\\-21&49&20\\-15&20&10\end{pmatrix}$，故最佳逼近解为

$$\begin{aligned}\beta &= A^+ b = (A^TA)^{-1}A^Tb\\ &= \frac{1}{90}\begin{pmatrix}54&-21&-15\\-21&49&20\\-15&20&10\end{pmatrix}\begin{pmatrix}1&1&1&1\\-3&-2&0&1\\9&4&0&1\end{pmatrix}\begin{pmatrix}9\\6\\2\\1\end{pmatrix}\\ &= \begin{pmatrix}2\\-4/3\\1/3\end{pmatrix}.\end{aligned}$$

于是数据拟合的最佳二次抛物线为 $y=2-\frac{4}{3}x+\frac{1}{3}x^2$. 误差为

$$\|A\beta-b\|_2 = \left\|\begin{pmatrix}1&-3&9\\1&-2&4\\1&0&0\\1&1&1\end{pmatrix}\begin{pmatrix}2\\-4/3\\1/3\end{pmatrix}-\begin{pmatrix}9\\6\\2\\1\end{pmatrix}\right\|_2 = \left\|\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\right\|_2 = 0.$$

第七章 非负矩阵

非负矩阵理论是元素非负的实矩阵的理论，它起源于由 Perron 发现，后来由 Frobenius 发展的关于非负矩阵谱半径的一个优美结果，其基本特征已成为矩阵理论的经典内容之一. 它广泛应用于数值分析、运筹学、数理经济学、概率论、弹性系统微震动理论、系统稳定性分析、组合数学和控制论等诸多学科中. 作为矩阵理论研究的一个分支，这类矩阵具有一般矩阵所不具有的一系列重要而优美的性质（如 Perron-Frobenius 定理）. 作为解决问题的工具和手段，它与其他学科的联系日益紧密，已成为矩阵理论研究中最活跃的领域之一.

一、基本概念

1. 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{m\times n}$，$\boldsymbol{B}=(b_{ij})\in\mathbf{C}^{m\times n}$，记 $\boldsymbol{B}\geqslant 0$，如果所有 $b_{ij}\geqslant 0$；记 $\boldsymbol{B}>0$，如果所有 $b_{ij}>0$；记 $\boldsymbol{A}\geqslant\boldsymbol{B}$，如果所有 $a_{ij}\geqslant b_{ij}$；记 $\boldsymbol{A}>\boldsymbol{B}$，如果所有 $a_{ij}>b_{ij}$；记 $|\boldsymbol{A}|=(|a_{ij}|)$. 若 $\boldsymbol{A}\geqslant 0$，则称 $\boldsymbol{A}$ 为非负矩阵；若 $\boldsymbol{A}>0$，则称 $\boldsymbol{A}$ 为正矩阵.

2. 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$，若 $\boldsymbol{A}$ 为一个零阶矩阵或者对 $n\geqslant 2$ 有置换矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{C}\\ \boldsymbol{0} & \boldsymbol{D}\end{pmatrix}$，其中 $\boldsymbol{B}$ 和 $\boldsymbol{D}$ 为阶数大于等于 1 的方阵，则称 $\boldsymbol{A}$ 为可约的. 若 $\boldsymbol{A}$ 不是可约的，就称 $\boldsymbol{A}$ 为不可约的.

3. 若对有向图 D 的任何一对有序结点 (i,j)，$i\neq j$，都有一条由有向弧组成的有向路径 $\overset{\frown}{il_1}$，$\overset{\frown}{l_1l_2}$，$\cdots$，$\overset{\frown}{l_{r-1}j}$ 连接结点 i 与 j，则称有向图 D 为强连通的.

4. 若 n 阶非负矩阵 $\boldsymbol{A}$ 的各行元素之和均为 1，则称 $\boldsymbol{A}$ 为（行）随机矩阵；若非负矩阵 $\boldsymbol{A}$ 的各列元素之和均为 1，则称 $\boldsymbol{A}$ 为（列）随机矩阵；若 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 均为行随机矩阵，则称 $\boldsymbol{A}$ 为双随机矩阵.

5. 非负矩阵 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$，若 $\boldsymbol{A}$ 是不可约的且只有一个最大模特征值，则称 $\boldsymbol{A}$ 为素矩阵.

6. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 是非负矩阵. 如果存在正整数 m，使得 $\boldsymbol{A}^m$ 为正矩阵，则称 $\boldsymbol{A}$ 是素矩阵或本原矩阵或幂正矩阵.

7. 给定素矩阵 $\boldsymbol{A}$，使得 $\boldsymbol{A}^m>0$ 的最小正整数 m 称为 $\boldsymbol{A}$ 的素数指数或本原指数，记为 $\gamma(\boldsymbol{A})$.

二、主要结论

1. 设 $\boldsymbol{A},\boldsymbol{B}\in \mathbf{C}^{n\times n}$，若 $|\boldsymbol{A}|\leqslant \boldsymbol{B}$，则 $r(\boldsymbol{A})\leqslant r(|\boldsymbol{A}|)\leqslant r(\boldsymbol{B})$.

2. 设 $\boldsymbol{A},\boldsymbol{B}\in \mathbf{C}^{n\times n}$，若 $0\leqslant \boldsymbol{A}\leqslant \boldsymbol{B}$，则 $r(\boldsymbol{A})\leqslant r(\boldsymbol{B})$.

3. 设 $\boldsymbol{A}\in \mathbf{C}^{n\times n}$，$\boldsymbol{A}\geqslant 0$，

(1) 若 $\boldsymbol{A}_1$ 为 $\boldsymbol{A}$ 的任一主子阵，则 $r(\boldsymbol{A}_1)\leqslant r(\boldsymbol{A})$.

(2) 若 $\boldsymbol{A}$ 的所有行和都相等，则 $r(\boldsymbol{A})=\|\boldsymbol{A}\|_\infty$；若 $\boldsymbol{A}$ 的所有列和都相等，则 $r(\boldsymbol{A})=\|\boldsymbol{A}\|_1$.

(3) $\min\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{n}a_{ij}\leqslant r(\boldsymbol{A})\leqslant \max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{n}a_{ij}$，$\min\limits_{1\leqslant j\leqslant n}\sum\limits_{i=1}^{n}a_{ij}\leqslant r(\boldsymbol{A})\leqslant \max\limits_{1\leqslant j\leqslant n}\sum\limits_{i=1}^{n}a_{ij}$.

(4) 若对所有 $i=1,2,\cdots,n$，有 $\sum\limits_{j=1}^{n}a_{ij}>0$，则 $r(\boldsymbol{A})>0$. 特别地，若 $\boldsymbol{A}>0$，则 $r(\boldsymbol{A})>0$.

(5) 对任意正向量 $\boldsymbol{x}=(x_1,x_2,\cdots x_n)^{\mathrm{T}}\in C^n$，有

$$\min_{1\leqslant i\leqslant n}\frac{1}{x_i}\sum_{j=1}^{n}a_{ij}x_j\leqslant r(\boldsymbol{A})\leqslant \max_{1\leqslant i\leqslant n}\frac{1}{x_i}\sum_{j=1}^{n}a_{ij}x_j,$$

$$\min_{1\leqslant j\leqslant n}x_j\sum_{i=1}^{n}\frac{a_{ij}}{x_i}\leqslant r(\boldsymbol{A})\leqslant \max_{1\leqslant j\leqslant n}x_j\sum_{i=1}^{n}\frac{a_{ij}}{x_i}.$$

(6) 若 $\boldsymbol{A}$ 有正特征向量 $\boldsymbol{x}$，则相应的特征值为 $r(\boldsymbol{A})$，即如果 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，$\boldsymbol{x}>0$，$\boldsymbol{A}\geqslant 0$，则 $\lambda=r(\boldsymbol{A})$.

(7) 若 $\boldsymbol{A}$ 有正特征向量 $\boldsymbol{x}$，则有

$$r(\boldsymbol{A})=\max_{x>0}\min_{1\leqslant i\leqslant n}\frac{1}{x_i}\sum_{j=1}^{n}a_{ij}x_j=\min_{x>0}\max_{1\leqslant i\leqslant n}\frac{1}{x_i}\sum_{j=1}^{n}a_{ij}x_j.$$

(8) $(\boldsymbol{E}-\boldsymbol{A})^{-1}$ 存在且 $(\boldsymbol{E}-\boldsymbol{A})^{-1}\geqslant 0$ 的充要条件是 $r(\boldsymbol{A})<1$.

4. 设 $\boldsymbol{A}\in \mathbf{C}^{n\times n}$，$\boldsymbol{x}\in \mathbf{R}^n$，$\boldsymbol{A}\geqslant 0$，$\boldsymbol{x}>0$，若存在 $\alpha>0$，$\beta>0$，使得 $\alpha\boldsymbol{x}\leqslant \boldsymbol{A}\boldsymbol{x}\leqslant \beta\boldsymbol{x}$，则 $\alpha\leqslant r(\boldsymbol{A})\leqslant\beta$；

若 $\alpha\boldsymbol{x}<\boldsymbol{A}\boldsymbol{x}(\alpha\boldsymbol{x}\leqslant\boldsymbol{A}\boldsymbol{x})$，则 $\alpha<r(\boldsymbol{A})(\alpha\leqslant r(\boldsymbol{A}))$；若 $\boldsymbol{A}\boldsymbol{x}<\beta\boldsymbol{x}(\boldsymbol{A}\boldsymbol{x}\leqslant\beta\boldsymbol{x})$，则 $r(\boldsymbol{A})<\beta(r(\boldsymbol{A})\leqslant\beta)$.

5. 设 $\boldsymbol{A}\in \mathbf{R}^{n\times n}$，$\boldsymbol{A}>0$.

(1) 若 $\boldsymbol{A}$ 的某特征值 λ 满足 $|\lambda|=r(\boldsymbol{A})$，相应的特征向量为 $\mathbf{x}\neq\mathbf{0}$，则 $\boldsymbol{A}|\boldsymbol{x}|=r(\boldsymbol{A})|\boldsymbol{x}|$，$|\boldsymbol{x}|>0$.

(2) $r(\boldsymbol{A})>0$ 且 $r(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的一个特征值，相应的特征向量为正向量.

(3) 若 λ 为 $\boldsymbol{A}$ 的特征值且 $\lambda\neq r(\boldsymbol{A})$，则 $\lambda<r(\boldsymbol{A})$.

(4) 若 $\boldsymbol{y},\boldsymbol{z}$ 为满足 $\boldsymbol{A}\boldsymbol{y}=r(\boldsymbol{A})\boldsymbol{y}$，$\boldsymbol{A}\boldsymbol{z}=r(\boldsymbol{A})\boldsymbol{z}$ 的非零向量，则存在某个常数 $a\in\mathbf{C}$，使得 $\boldsymbol{y}=a\boldsymbol{z}$.

(5) 存在唯一的向量 $\boldsymbol{x}$，使得 $\boldsymbol{A}\boldsymbol{x}=r(\boldsymbol{A})\boldsymbol{x}$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)>0$ 及 $\sum\limits_{i=1}^{n}x_i=1$.

(6) $\lim\limits_{m\to\infty}\boldsymbol{B}^m=\lim\limits_{m\to\infty}(r(\boldsymbol{A})^{-1}\boldsymbol{A})^m=L=\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}$，其中，向量 $\boldsymbol{x},\boldsymbol{y}$ 满足

$$\boldsymbol{A}\boldsymbol{x}=r(\boldsymbol{A})\boldsymbol{x},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}=r(\boldsymbol{A})\boldsymbol{y},\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=1.$$

(7) $r(\boldsymbol{A})$为 $\boldsymbol{A}$ 的单重特征值.

(8) 矩阵 $\boldsymbol{A}$ 的谱半径 $r(\boldsymbol{A})>0$；相应于 $\boldsymbol{A}$ 的特征向量为正向量；对于 $\boldsymbol{A}$ 的任一异于 $r(\boldsymbol{A})$的特征值 λ，均满足 $|\lambda|<r(\boldsymbol{A})$.

(9) 若 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x},\boldsymbol{x}\neq 0$，以及 $|\lambda|=r(\boldsymbol{A})$，则对某个 $\theta\in\mathbf{R}$，有 $\mathrm{e}^{-i\theta}\boldsymbol{x}=|\boldsymbol{x}|>0$ 或 $\boldsymbol{x}=\mathrm{e}^{i\theta}|\boldsymbol{x}|>0$.

(10) 若 $\rho<R$，则 $\boldsymbol{A}$ 的最大特征值 $r=r(\boldsymbol{A})$满足

$$\rho+\eta\left(\frac{1}{\sqrt{\delta}}-1\right)\leqslant r\leqslant R-\eta(1-\sqrt{\delta}),$$

其中 $R=\max\limits_i\sum\limits_{j=1}^n a_{ij},\rho=\min\sum\limits_{j=1}^n a_{ij},\eta=\min\limits_{i,j}a_{ij},\delta=\max\limits_{r_i<r_j}(r_i/r_j)$，$r_i$ 为 $\boldsymbol{A}$ 的第 i 行行和.

(11) 若 r 为 $\boldsymbol{A}$ 的最大特征值，则

$$\rho+\eta\left(\frac{1}{\sigma}-1\right)\leqslant r\leqslant R-\eta(1-\sigma),$$

其中 $\sigma=\sqrt{(\rho-\eta)/(R-\eta)},R=\max\limits_i\sum\limits_{j=1}^n a_{ij},\rho=\min\sum\limits_{j=1}^n a_{ij},\eta=\min\limits_{i,j}a_{ij},\delta=\max\limits_{r_i<r_j}(r_i/r_j)$，$r_i$ 为 $\boldsymbol{A}$ 的第 i 行行和.

(12) 若 $\boldsymbol{A},r,R,\rho,\eta$ 和 r_i 的定义同上，则

$$\rho+\eta(h-1)\leqslant r\leqslant R-\eta\left(1-\frac{1}{g}\right),$$

其中

$$g=\frac{R-2\eta+\sqrt{R^2-4\eta(R-\rho)}}{2(\rho-\eta)},$$

$$h=\frac{-\rho+2\eta+\sqrt{\rho^2+4\eta(R-\rho)}}{2\eta}.$$

6. 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n},B=(b_{ij})_{n\times n}\in\mathbf{R}^{n\times n}$满足 $\boldsymbol{B}\geqslant|\boldsymbol{A}|$，则 $\boldsymbol{A}$ 的每个特征值均位于区域 $\bigcup\limits_{i=1}^n\{z\in\mathbf{C}:|z-a_{ii}|\leqslant r(\boldsymbol{B})-b_{ii}\}$ 之中.

7. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n},\boldsymbol{A}\geqslant 0$，则

(1) $r(\boldsymbol{A})$为 $\boldsymbol{A}$ 的特征值，且存在非负向量 $\boldsymbol{x}\geqslant 0,\boldsymbol{x}\neq\boldsymbol{0}$，使得 $\boldsymbol{A}\boldsymbol{x}=r(\boldsymbol{A})\boldsymbol{x}$.

(2) $r(\boldsymbol{A})=\max\limits_{\substack{\boldsymbol{x}\geqslant 0\\ \boldsymbol{x}\neq\boldsymbol{0}}}\min\limits_{\substack{1\leqslant i\leqslant n\\ x_i\neq 0}}\dfrac{1}{x_i}\sum\limits_{j=1}^n a_{ij}x_j$.

8. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n},\boldsymbol{A}\geqslant 0,\boldsymbol{x}\in\mathbf{C}^n,\boldsymbol{x}\geqslant 0,\boldsymbol{x}\neq\boldsymbol{0}$，若有 $\alpha\in\mathbf{R}$，使得 $\boldsymbol{A}\boldsymbol{x}\geqslant\alpha\boldsymbol{x}$，则 $r(\boldsymbol{A})\geqslant\alpha$.

9. 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n},n\geqslant 2$，则

(1) $\boldsymbol{A}$ 为不可约矩阵的充要条件是对集合 N^* 的任意分割 N_1 与 $N_2(N_1\cap N_2=\varphi,N_1\cup N_2=N^*)$，$\boldsymbol{A}$ 中有元素 $a_{ij}\neq 0,i\in N_1,j\in N_2$.

(2) $\boldsymbol{A}$ 为不可约矩阵的充要条件是 $\boldsymbol{A}$ 的有向图 $D(\boldsymbol{A})$ 是强连通的.

10. $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$ 为不可约矩阵的充要条件是 $(\boldsymbol{E}+|\boldsymbol{A}|)^{n-1}>0$.

11. 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$, $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值(计重特征值), 则 $\lambda_1+1,\lambda_2+1,\cdots,\lambda_n+1$ 是 $\boldsymbol{E}+\boldsymbol{A}$ 的特征值, 且 $r(\boldsymbol{E}+\boldsymbol{A})\leqslant 1+r(\boldsymbol{A})$, 若还有 $\boldsymbol{A}\geqslant 0$, 则 $r(\boldsymbol{E}+\boldsymbol{A})=1+r(\boldsymbol{A})$.

12. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$, $\boldsymbol{A}\geqslant 0$ 且对某个 $k\geqslant 1$, 有 $\boldsymbol{A}^k>0$, 则 $r(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的代数单重特征值.

13. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 为非负不可约矩阵, 则

(1) $r(\boldsymbol{A})>0$;

(2) $r(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的特征值;

(3) 存在正向量 $\boldsymbol{x}$, 使得 $\boldsymbol{A}\boldsymbol{x}=r(\boldsymbol{A})\boldsymbol{x}$;

(4) $r(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的代数(因而为几何)单重特征值.

14. 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$, 若 $\boldsymbol{A}$ 是非负不可约矩阵, $\boldsymbol{A}\geqslant|\boldsymbol{B}|$, 则 $r(\boldsymbol{A})\geqslant r(\boldsymbol{B})$, 若 $r(\boldsymbol{A})=r(\boldsymbol{B})$, 又 $\lambda=\mathrm{e}^{\mathrm{i}\varphi}r(\boldsymbol{B})(0\leqslant\varphi\leqslant 2\pi)$ 是 $\boldsymbol{B}$ 的特征值, 则存在 $\theta_1,\theta_2,\cdots,\theta_n\in\mathbf{R}$, 使得 $\boldsymbol{B}=\mathrm{e}^{\mathrm{i}\varphi}\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^{-1}$, 其中 $\boldsymbol{D}=\mathrm{diag}(\mathrm{e}^{\mathrm{i}\theta_1},\mathrm{e}^{\mathrm{i}\theta_2},\cdots,\mathrm{e}^{\mathrm{i}\theta_n})$.

15. 设 $\boldsymbol{A}$ 为不可约非负矩阵, $\boldsymbol{B}\geqslant\boldsymbol{A}$, $\boldsymbol{B}\neq\boldsymbol{A}$, 则 $r(\boldsymbol{B})>r(\boldsymbol{A})$.

16. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 为素矩阵, 则

(1) $\lim\limits_{m\to\infty}(r(\boldsymbol{A})^{-1}\boldsymbol{A})^m=L>0$, 其中 $\boldsymbol{L}=\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}$, $\boldsymbol{A}\boldsymbol{x}=r(\boldsymbol{A})\boldsymbol{x}$, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}=r(\boldsymbol{A})\boldsymbol{y}$, $\boldsymbol{x}>0$, $\boldsymbol{y}>0$, $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=1$.

(2) $\boldsymbol{A}^k$ 为不可约矩阵, $k=1,2,\cdots,n$.

(3) 对正整数 $k\leqslant(n-1)n^n$, 有 $\boldsymbol{A}^k>0$.

17. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$, P_i 和 P_j 为有向图 $D(\boldsymbol{A})$ 的给定结点, 则在 P_i 和 P_j 之间存在 $D(\boldsymbol{A})$ 中的一条长为 m 的有向道路的充要条件是 $(|\boldsymbol{A}|^m)_{ij}\neq 0$.

18. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$, 则 $|\boldsymbol{A}|^m>0$ 的充要条件是从 $D(\boldsymbol{A})$ 中的每个结点 P_i 到其余任一结点 P_j 都存在一条长恰为 m 的有向道路.

19. 设 $\boldsymbol{A}$ 为 n 阶非负矩阵, 则

(1) $\boldsymbol{A}$ 为素矩阵的充要条件是 $\boldsymbol{A}^m>0$ 对某个 $m\geqslant 1$ 成立.

(2) $\boldsymbol{A}$ 为素矩阵的充要条件是 $\boldsymbol{A}^{n^2-2n+2}>0$.

20. 设 $\boldsymbol{A}$ 为 n 阶素矩阵, $D(\boldsymbol{A})$ 中最简单有向回路长为 s, 则 $\gamma(\boldsymbol{A})\leqslant n+s(n-2)$.

21. (1) 设 $\boldsymbol{A}$ 为随机矩阵, 则 $\boldsymbol{A}$ 的最大特征值为 1;

(2) 设 $\boldsymbol{A}$ 为非负矩阵, 则 $\boldsymbol{A}$ 为随机矩阵的充要条件是 $\boldsymbol{e}$ 为 $\boldsymbol{A}$ 的相应于特征值 1 的特征向量, 其中 $\boldsymbol{e}=(1,1,\cdots,1)^{\mathrm{T}}$.

22. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 为非负矩阵, $r(\boldsymbol{A})>0$, $\boldsymbol{z}>0$, 满足 $\boldsymbol{A}\boldsymbol{z}=r(\boldsymbol{A})\boldsymbol{z}$, 则 $r(\boldsymbol{A})^{-1}\boldsymbol{A}$ 相似于一个随机矩阵.

23. 设 $\boldsymbol{A}=(a_{ij})$ 为随机矩阵, $\omega=\min\limits_i a_{ii}$, 则对 $\boldsymbol{A}$ 的任一特征值 λ_i, 都有 $|\lambda_i-\omega|\leqslant 1-\omega$.

24. n 阶矩阵 $\boldsymbol{A}$ 为双随机矩阵的充要条件是对某个 $N<\infty$，存在置换矩阵 $\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_N$ 和正纯量 $a_1,a_2,\cdots,a_N\in\mathbf{R}$，使得 $a_1+a_2+\cdots+a_N=1$，且 $\boldsymbol{A}=a_1\boldsymbol{P}_1+a_2\boldsymbol{P}_2+\cdots+a_N\boldsymbol{P}_N$.

25. 设 $\boldsymbol{A}$ 是素矩阵，则

(1) $\boldsymbol{A}^{\mathrm{T}}$ 是素矩阵；

(2) 对任意的正整数 k，$\boldsymbol{A}^k$ 是素矩阵；

(3) 对任意的非负矩阵 $\boldsymbol{B}$，$\boldsymbol{A}+\boldsymbol{B}$ 是素矩阵.

26. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 为素矩阵，则

(1) $r(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的具有正特征向量的特征值；

(2) $\boldsymbol{A}$ 的特征值 $r(\boldsymbol{A})$ 为一个单根；

(3) $\boldsymbol{A}$ 的其他特征值的模小于 $r(\boldsymbol{A})$；

(4) $\gamma(\boldsymbol{A})\leqslant(n-1)^2+1$.

27. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 是素矩阵，且对某个正整数 h 和矩阵 $\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^h$，至少有 $d(>0)$ 个正对角元，则有 $\gamma(\boldsymbol{A})\leqslant n-d+(n-1)h$.

28. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 是素矩阵，并且 a_{ij} 与 a_{ji} 同时为正或同时为零，则 $\gamma(\boldsymbol{A})\leqslant 2(n-1)$.

三、典型例题

1. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$，$\|\boldsymbol{A}\|$ 是 $\mathbf{R}^{n\times n}$ 上的任意一种矩阵范数，则 $r(\boldsymbol{A})=\lim\limits_{k\to\infty}\|\boldsymbol{A}^k\|^{\frac{1}{k}}$.

证 对任意的正整数 k，有 $r(\boldsymbol{A}^k)=[r(\boldsymbol{A})]^k$，而 $r(\boldsymbol{A}^k)\leqslant\|\boldsymbol{A}^k\|$，于是 $[r(\boldsymbol{A})]^k\leqslant\|\boldsymbol{A}^k\|$，从而 $r(\boldsymbol{A})\leqslant\|\boldsymbol{A}^k\|^{\frac{1}{k}}$，取极限得 $r(\boldsymbol{A})\leqslant\lim\limits_{k\to\infty}\|\boldsymbol{A}^k\|^{\frac{1}{k}}$.

另一方面，对任意的 $\varepsilon>0$，令 $\boldsymbol{B}=\dfrac{1}{r(\boldsymbol{A})+\varepsilon}\boldsymbol{A}$，则 $r(\boldsymbol{B})<1$. 因而 $\lim\limits_{k\to\infty}\boldsymbol{B}^k=\boldsymbol{0}$. 于是存在正整数 N，当 $k>N$ 时，有 $\|\boldsymbol{B}^k\|<1$. 即 $\dfrac{1}{[r(\boldsymbol{A})+\varepsilon]^k}\|\boldsymbol{A}^k\|<1$，于是 $\|\boldsymbol{A}^k\|^{\frac{1}{k}}<r(\boldsymbol{A})+\varepsilon$，取极限，由 ε 的任意性，有 $\lim\limits_{k\to\infty}\|\boldsymbol{A}^k\|^{\frac{1}{k}}\leqslant r(\boldsymbol{A})$，故 $\lim\limits_{k\to\infty}\|\boldsymbol{A}^k\|^{\frac{1}{k}}=r(\boldsymbol{A})$.

2. 设 $\boldsymbol{A}=(a_{ij})_{n\times n}\geqslant 0$ 为不可约矩阵. 则 $\boldsymbol{A}$ 不能有两个线性无关的非负特征向量.

证 设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}>0$ 为 $\boldsymbol{A}$ 的对应于 $r(\boldsymbol{A})$ 的特征向量. 此外，$\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}\geqslant 0$ 为 $\boldsymbol{A}$ 的对应于特征值 a 的特征向量，且 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 线性无关. 因 $r(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的单重特征值，故 $a\neq r(\boldsymbol{A})$. 进一步，记 $\boldsymbol{z}=(z_1,z_2,\cdots,z_n)^{\mathrm{T}}$ 为 $\boldsymbol{A}^{\mathrm{T}}$ 的对应于 $r(\boldsymbol{A}^{\mathrm{T}})=r(\boldsymbol{A})$ 的正特征向量. 于是有

$$r(\boldsymbol{A})(\boldsymbol{y},\boldsymbol{z})=(\boldsymbol{y},r(\boldsymbol{A})\boldsymbol{z})=(\boldsymbol{y},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z})=\boldsymbol{z}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{y}=(\boldsymbol{A}\boldsymbol{y},\boldsymbol{z})=(a\boldsymbol{y},\boldsymbol{z})=a(\boldsymbol{y},\boldsymbol{z})$$

即 $(r(\boldsymbol{A})-a)(\boldsymbol{y},\boldsymbol{z})=0$，而 $a\neq r(\boldsymbol{A})$，故必有 $(\boldsymbol{y},\boldsymbol{z})=0$，此与 $\boldsymbol{y}\geqslant 0,\boldsymbol{z}>0,\boldsymbol{y}\neq 0$ 时 $(\boldsymbol{y},\boldsymbol{z})>0$ 矛盾！故 $\boldsymbol{y}$ 不存在.

3. 设 $\boldsymbol{A}$ 是素矩阵，则对任意的正整数 m，$\boldsymbol{A}^m$ 是素矩阵.

证　因为 $r(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的单重特征值，而且是 $\boldsymbol{A}$ 的模为 $r(\boldsymbol{A})$ 的唯一特征值，所以 $(r(\boldsymbol{A}))^m$ 是 $\boldsymbol{A}^m$ 的单特征值，而且是 $\boldsymbol{A}^m$ 的模为 $(r(\boldsymbol{A}))^m$ 的唯一特征值. 因此，对 $m\geqslant 1$，由于有 $\boldsymbol{A}^m\geqslant 0$，要证明 $\boldsymbol{A}^m$ 是素矩阵，只需证明 $\boldsymbol{A}^m$ 是不可约的.

用反证法. 假定对某个正整数 v，$\boldsymbol{A}^v$ 是可约的，不妨设 $\boldsymbol{A}^v=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{C}\\ \boldsymbol{0} & \boldsymbol{D}\end{pmatrix}$，其中 $\boldsymbol{B},\boldsymbol{D}$ 为子方阵. 设 $\boldsymbol{A}\boldsymbol{x}=r(\boldsymbol{A})\boldsymbol{x}$，$\boldsymbol{x}>0$，于是有 $\boldsymbol{A}^v\boldsymbol{x}=(r(\boldsymbol{A}))^v\boldsymbol{x}$，对应于 $\boldsymbol{B},\boldsymbol{D}$，将 $\boldsymbol{x}$ 写成 $\boldsymbol{x}=\begin{pmatrix}x^{(1)}\\ x^{(2)}\end{pmatrix}$，从而有 $d^v x^{(2)}=(r(\boldsymbol{A}))^v x^{(2)}$，这表明 $(r(\boldsymbol{A}))^v$ 是 $\boldsymbol{D}$ 的特征值.

另一方面，再考虑 $\boldsymbol{A}^{\mathrm{T}}$，它也是非负的和不可约的，那么必存在向量 $\boldsymbol{y}>0$，使得 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}=r(\boldsymbol{A})\boldsymbol{y}$. 类似地，可推出 $(r(\boldsymbol{A}))^v$ 是 $\boldsymbol{B}^{\mathrm{T}}$ 的特征值，从而也是 $\boldsymbol{B}$ 的特征值.

这样，$(r(\boldsymbol{A}))^v$ 是 $\boldsymbol{D}$ 的特征值又是 $\boldsymbol{B}$ 的特征值，则 $r(\boldsymbol{A})$ 成了 $\boldsymbol{A}$ 的多重特征值，矛盾！因此，对所有的 $m\geqslant 1$，$\boldsymbol{A}^m$ 是不可约的. 从而 $\boldsymbol{A}^m$ 是素矩阵.

4. 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{n\times n}$，$\boldsymbol{A}\geqslant 0$，$\boldsymbol{A}$ 不可约，而且 $a_{ii}>0$，$i=1,2,\cdots,n$，证明 $\boldsymbol{A}^{n-1}>0$.

证　可选取非负不可约方阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}\geqslant\gamma(\boldsymbol{E}+\boldsymbol{B})$，$\gamma>0$. 于是有 $\boldsymbol{A}^{n-1}\geqslant\gamma^{n-1}(\boldsymbol{E}+\boldsymbol{B})^{n-1}>0$.

5. 非负不可约矩阵 $\boldsymbol{A}$ 的“模等于 $r(\boldsymbol{A})$ 的特征值”唯一吗？

解　非负不可约矩阵 $\boldsymbol{A}$ 的谱半径 $r(\boldsymbol{A})$ 是它的单重特征值，但模等于 $r(\boldsymbol{A})$ 的特征值不唯一.

例如，对非负不可约矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & 2 & 0\\ 2 & 1 & 3\\ 0 & 2 & 1\end{pmatrix}$，其谱半径 $r(\boldsymbol{A})=1+\sqrt{10}$ 就是它的一个正的单特征值，而属于 $r(\boldsymbol{A})$ 的正特征向量是 $(2,\sqrt{10},2)^{\mathrm{T}}$. 并且“模等于 $r(\boldsymbol{A})$ 的特征值”$1+\sqrt{10}$ 也只有一个.

但对于非负不可约矩阵 $\boldsymbol{A}=\begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & 1\\ 1 & 0 & 0 & 0 & 0\end{pmatrix}$，$r(\boldsymbol{A})=1$，$\boldsymbol{A}$ 的 n 个特征值是 $\lambda_j=\mathrm{e}^{\mathrm{i}\frac{2j\pi}{n}}$，$j=0,1,\cdots,n-1$，从而 $\boldsymbol{A}$ 的特征值的模都等于 $r(\boldsymbol{A})$. 综上，对于一个非负素矩阵 $\boldsymbol{A}$，除特征值 $r(\boldsymbol{A})$ 外，其余特征值的模都小于 $r(\boldsymbol{A})$，但对非负不可约矩阵，该结论不再成立.

6. 设 $\boldsymbol{A}$ 为不可约随机矩阵，则 $\lim\limits_{m\to\infty}\boldsymbol{A}^m$ 存在的充分必要条件是 $\boldsymbol{A}$ 为素矩阵.

证　由于 $\boldsymbol{A}$ 的谱半径 $r(\boldsymbol{A})$ 等于 1，且 $\boldsymbol{A}$ 的任一模等于 1 的特征值所对应的 Jordan 块都是一阶的，所以结论成立.

7. 在某区人口流动问题的研究中,出现随机矩阵 $\boldsymbol{P}=\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$ 的幂 $\boldsymbol{P}^m$ 的极限问题,求 $\lim\limits_{m\to\infty}\boldsymbol{P}^m$.

解 由于 $\boldsymbol{E}+\boldsymbol{P}>0$,故 $\boldsymbol{P}$ 不可约. 又因为 $\boldsymbol{P}^2=\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}>0$,所以 $\boldsymbol{P}$ 是素矩阵,故可知 $\lim\limits_{m\to\infty}\boldsymbol{P}^m$. 又由归纳法可证明 $\boldsymbol{P}^m=\begin{pmatrix} p_m & p_{m+1} & p_{m+1} \\ p_{m+1} & p_m & p_{m+1} \\ p_{m+1} & p_{m+1} & p_m \end{pmatrix}$,其中 $p_m=\frac{1}{3}\left(1+\frac{(-1)^m}{2^{m-1}}\right)$,因此,$\lim\limits_{m\to\infty}\boldsymbol{P}^m=\frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

四、习题解答

1. 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$,$\boldsymbol{x}\in\mathbf{C}^n$,证明:

(1) $|\boldsymbol{Ax}|\leqslant|\boldsymbol{A}|\,|\boldsymbol{x}|$;

(2) $|\boldsymbol{AB}|\leqslant|\boldsymbol{A}|\,|\boldsymbol{B}|$;

(3) 若 $0\leqslant\boldsymbol{A}\leqslant\boldsymbol{B}$,则 $0\leqslant\boldsymbol{A}^m\leqslant\boldsymbol{B}^m$($m$ 为正整数).

证 (用归纳法)

当 $m=1$ 时,结论成立.

设结论对 $m-1$ 成立,即 $\boldsymbol{A}^{m-1}\leqslant\boldsymbol{B}^{m-1}$,那么 $\boldsymbol{B}^{m-1}-\boldsymbol{A}^{m-1}\geqslant0$,又 $\boldsymbol{B}\geqslant\boldsymbol{A}\geqslant0$,故

$$\begin{aligned}\boldsymbol{B}^m-\boldsymbol{A}^m&=\boldsymbol{B}^m-\boldsymbol{AB}^{m-1}+\boldsymbol{AB}^{m-1}-\boldsymbol{A}^m\\&=\boldsymbol{B}(\boldsymbol{B}^{m-1}-\boldsymbol{A}^{m-1})+(\boldsymbol{B}-\boldsymbol{A})\boldsymbol{A}^{m-1}\geqslant0,\end{aligned}$$

从而 $\boldsymbol{B}^m\geqslant\boldsymbol{A}^m\geqslant0$.

2. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,$\boldsymbol{B}=|\boldsymbol{A}|$,则 $\|\boldsymbol{A}\|_{m_2}=\left(\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}}=\|\boldsymbol{B}\|_{m_2}$,但 $\|\boldsymbol{A}\|_2=\|\boldsymbol{B}\|_2$ 未必成立.

证　设 $\boldsymbol{A}=\begin{pmatrix}1 & -2 & 3\\ 7 & 8 & -1\\ 6 & -9 & 2\end{pmatrix}$，$\boldsymbol{B}=|\boldsymbol{A}|=\begin{pmatrix}1 & 2 & 3\\ 7 & 8 & 1\\ 6 & 9 & 2\end{pmatrix}$，则 $\|\boldsymbol{A}\|_2=\sqrt{r(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})}\approx 12.4998$，$\|\boldsymbol{B}\|_2=\sqrt{4(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})}\approx 15.5077$，即 $\|\boldsymbol{A}\|_2\neq\|\boldsymbol{B}\|_2$.

对于 $\|\cdot\|_{m_2}$，显然有

$$\|\boldsymbol{A}\|_{m_2}=\Big(\sum_{j=1}^{n}\sum_{i=1}^{n}|a_{ij}|^2\Big)^{\frac{1}{2}}=\|\boldsymbol{B}\|_{m_2}.$$

3．构造一个相似于 $\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$ 且不含零元的矩阵，它的谱半径是什么？它是非负矩阵吗？

解　令 $\boldsymbol{P}=\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$，$\boldsymbol{A}=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}-1 & 1\\ -1 & 1\end{pmatrix}$，则 $\boldsymbol{P}^{-1}=\begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix}$，那么 $\boldsymbol{PAP}^{-1}=\boldsymbol{B}$，即 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相似，且 $\boldsymbol{B}$ 不是非负矩阵，其谱半径为 0.

4．设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$，$0\leqslant\boldsymbol{A}<\boldsymbol{B}$，则 $r(\boldsymbol{A})<r(\boldsymbol{B})$.

证　由 $0\leqslant\boldsymbol{A}<\boldsymbol{B}$，可知 $\boldsymbol{B}-\boldsymbol{A}>0$，且由 $\boldsymbol{A}\geqslant 0$ 知 $\boldsymbol{Ax}=r(\boldsymbol{A})\boldsymbol{x}(\boldsymbol{x}\geqslant 0,\boldsymbol{x}\neq\boldsymbol{0})$，则 $0<(\boldsymbol{B}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{Bx}-\boldsymbol{Ax}=\boldsymbol{Bx}-r(\boldsymbol{A})\boldsymbol{x}$，即 $\boldsymbol{Bx}>r(\boldsymbol{A})\boldsymbol{x}$，从而有 $(\boldsymbol{Bx})_i>r(\boldsymbol{A})x_i(x_i>0)$，$\dfrac{(\boldsymbol{Bx})_i}{x_i}>r(\boldsymbol{A})$，$r(\boldsymbol{B})\geqslant\min\limits_{i}\dfrac{(\boldsymbol{B}x)_i}{x_i}>r(\boldsymbol{A})$，即 $r(\boldsymbol{B})>r(\boldsymbol{A})$.

5．设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，$\boldsymbol{A}\geqslant 0$，证明：$r(\boldsymbol{A})\leqslant\max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{n}a_{ij}$.

证　令 $\beta=\max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{n}a_{ij}$，构造一个矩阵 $\boldsymbol{B}$，使得 $0\leqslant\boldsymbol{A}\leqslant\boldsymbol{B}$ 且 $\beta=\sum\limits_{j=1}^{n}b_{ij}(i=1,2,\cdots,n)$. 例如，若 $\beta=0$，令 $\boldsymbol{B}=\boldsymbol{0}$ 即可，若 $\beta>0$，令 $b_{ij}=\beta a_{ij}\Big(\sum\limits_{j=1}^{n}a_{ij}\Big)^{-1}$，则有 $r(\boldsymbol{B})=\beta$，且 $r(\boldsymbol{A})\leqslant r(\boldsymbol{B})$.

6．证明不可约矩阵不可能有零行或零列.

证　设其第 k 行为零行，令 $\boldsymbol{P}=\begin{pmatrix}1 & & & & \\ & \ddots & & & \\ & & 0 & \cdots & 1\\ & & \vdots & \ddots & \vdots\\ & & 1 & \cdots & 0\end{pmatrix}$，则有 $\boldsymbol{PAP}^{\mathrm{T}}=\begin{pmatrix}\widetilde{\boldsymbol{A}} & \widetilde{\boldsymbol{b}}\\ \boldsymbol{0} & \widetilde{\boldsymbol{c}}\end{pmatrix}$，从而为可约矩阵.

7．设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，$\boldsymbol{x}\in\mathbf{R}^n$，$\boldsymbol{A}\geqslant 0$，$\boldsymbol{x}>0$，且有 $\beta>0$ 使得 $\boldsymbol{Ax}\leqslant\beta\boldsymbol{x}(\boldsymbol{Ax}<\beta\boldsymbol{x})$，证明：$r(\boldsymbol{A})\leqslant\beta(r(\boldsymbol{A})<\beta)$.

证　若 $\boldsymbol{Ax}\leqslant\beta\boldsymbol{x}$，则 $\sum\limits_{j-1}^{n}a_{ij}x_j\leqslant\beta x_i$，从而有 $\dfrac{1}{x_i}\sum\limits_{j=1}^{n}a_{ij}x_j\leqslant\beta$，$\max\limits_{1\leqslant i\leqslant n}\dfrac{1}{x_i}\sum\limits_{j=1}^{n}a_{ij}x_j\leqslant\beta$，从而有

$r(\boldsymbol{A}) \leqslant \max\limits_{1\leqslant i\leqslant n} \dfrac{1}{x_i}\sum\limits_{j=1}^{n} a_{ij}x_j \leqslant \beta$，若 $\boldsymbol{Ax} < \beta\boldsymbol{x}$，则存在 β'，使得 $\beta' < \beta$，且 $\boldsymbol{Ax} \leqslant \beta'\boldsymbol{x}$，于是有 $r(\boldsymbol{A})\leqslant\beta'<\beta$.

8. 若 $\boldsymbol{A}\geqslant 0$，且对某个 k 有 $\boldsymbol{A}^k>0$，证明 $r(\boldsymbol{A})>0$.

证 设 $\boldsymbol{A}$ 的全部特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则 $\boldsymbol{A}^k$ 的全部特征值为 $\lambda_1^k,\lambda_2^k,\cdots,\lambda_n^k$，从而 $r(\boldsymbol{A}^k)=\max\limits_{i}|\lambda_i^k|=(\max\limits_{i}|\lambda_i|)^k=(r(\boldsymbol{A}))^k$. 由 $\boldsymbol{A}^k>0$，可知 $r(\boldsymbol{A}^k)>0$，从而 $r(\boldsymbol{A})=\sqrt[k]{r(\boldsymbol{A}^k)}>0$.

9. 给出一个二阶矩阵 $\boldsymbol{A}$，使得 $\boldsymbol{A}\geqslant 0$，且 $\boldsymbol{A}$ 不是正的，但 $\boldsymbol{A}^2>0$.

解 令 $\boldsymbol{A}=\begin{pmatrix}1&1\\1&0\end{pmatrix}$，则 $\boldsymbol{A}\geqslant 0$，$\boldsymbol{A}$ 不是正的，且

$$\boldsymbol{A}^2=\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}1&1\\1&0\end{pmatrix}=\begin{pmatrix}2&1\\1&1\end{pmatrix}>0.$$

这说明，正矩阵是素矩阵，但反过来不成立. 例如，若 $\boldsymbol{B}=\begin{bmatrix}0&1&1\\1&0&0\\1&1&1\end{bmatrix}$，则 $\boldsymbol{B}$ 不是正矩阵，但 $\boldsymbol{B}^4>0$，所以 $\boldsymbol{B}$ 是素矩阵.

10. 设非负矩阵 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$，若 $\boldsymbol{A}$ 有正特征向量 $\boldsymbol{x}$，则对所有 $m=1,2,\cdots$ 和 $i=1,2,\cdots,n$，有 $\sum\limits_{j=1}^{n}a_{ij}^{(m)} \leqslant \left(\dfrac{\max\limits_{1\leqslant k\leqslant m}x_k}{\min\limits_{1\leqslant k\leqslant m}x_k}\right)r(\boldsymbol{A})^m$，$\left(\dfrac{\min\limits_{1\leqslant k\leqslant m}x_k}{\max\limits_{1\leqslant k\leqslant m}x_k}\right)r(\boldsymbol{A})^m \leqslant \sum\limits_{j=1}^{n}a_{ij}^{(m)}$，其中 $\boldsymbol{A}^m=(a_{ij}^{(m)})$. 特别地，若 $r(\boldsymbol{A})>0$，则对 $m=1,2,\cdots$，都有 $r(\boldsymbol{A})^{-1}\boldsymbol{A}^m$ 的各元一致有界.

证 若 $\boldsymbol{Ax}=r(\boldsymbol{A})\boldsymbol{x}$，则 $\boldsymbol{A}^m\boldsymbol{x}=r(\boldsymbol{A})^m\boldsymbol{x}$. 若 $\boldsymbol{A}\geqslant 0$，则 $\boldsymbol{A}^m\geqslant 0$，且有 $(r(\boldsymbol{A}))^m\max\limits_{1\leqslant k\leqslant n}x_k \geqslant$

$r(\boldsymbol{A}^m)x_i=(\boldsymbol{A}^m\boldsymbol{x})_i=\sum\limits_{j=1}^{n}a_{ij}^{(m)}x_j \geqslant \min\limits_{1\leqslant k\leqslant n}x_k\cdot\sum\limits_{j=1}^{n}a_{ij}^{(m)}, i=1,2,\cdots,n.$

由于 $\boldsymbol{x}>0$，两边同除以 $\min\limits_{1\leqslant k\leqslant n}x_k$，即得结论.

类似地，有

$$\begin{aligned}(r(\boldsymbol{A}))^m\min_{1\leqslant k\leqslant n}x_k &\leqslant r(\boldsymbol{A}^m)x_i=(\boldsymbol{A}^m\boldsymbol{x})_i=\sum_{j=1}^{n}a_{ij}^{(m)}x_j\\ &\leqslant \max_{1\leqslant k\leqslant n}x_k\cdot\sum_{j=1}^{n}a_{ij}^{(m)},\ i=1,2,\cdots,n.\end{aligned}$$

再由 $\boldsymbol{x}>0$，两边同除以 $\max\limits_{1\leqslant k\leqslant n}x_k$，即得结论.

11. 若 $\boldsymbol{A}$ 为 n 阶正矩阵，证明 $r(\boldsymbol{A})=\max\limits_{x>0}\min\limits_{1\leqslant i\leqslant n}\dfrac{1}{x_i}\sum\limits_{j=1}^{n}a_{ij}x_j=\min\limits_{x>0}\max\limits_{1\leqslant i\leqslant n}\dfrac{1}{x_i}\sum\limits_{j=1}^{n}a_{ij}x_j$.

证 $\forall x>0$，令 $\boldsymbol{D}=\mathrm{diag}(x_1,x_2,\cdots,x_n)$，$\boldsymbol{B}=\boldsymbol{D}^{-1}\boldsymbol{AD}$，则显然 $\boldsymbol{B}>0$，且 $\min\limits_{1\leqslant i\leqslant n}\left(\dfrac{\sum\limits_{j=1}^{n}a_{ij}x_j}{x_i}\right)<$

$$r(\boldsymbol{A})<\max_{1\leqslant i\leqslant n}\left(\frac{\sum\limits_{j=1}^{n}a_{ij}x_j}{x_i}\right)，或\ r(\boldsymbol{A})=\frac{\sum\limits_{j=1}^{n}a_{ij}x_j}{x_i}(i=1,2,\cdots,n)，而且有\sup_{x>0}\left\{\min_{1\leqslant i\leqslant n}\left(\frac{\sum\limits_{j=1}^{n}a_{ij}x_j}{x_i}\right)\right\}\leqslant$$

$$r(\boldsymbol{A})\leqslant\inf_{x>0}\left\{\max_{1\leqslant i\leqslant n}\left(\frac{\sum\limits_{j=1}^{n}a_{ij}x_j}{x_i}\right)\right\}，$$特别地，选择 $\boldsymbol{x}>0$ 为 $r(\boldsymbol{A})$ 对应的特征向量，则等式成立.

12. 设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 为正矩阵，证明存在唯一向量 $\boldsymbol{x}$，使得 $\boldsymbol{A}\boldsymbol{x}=r(\boldsymbol{A})\boldsymbol{x}$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)>0$ 及 $\sum\limits_{i=1}^{n}x_i=1$.

证　由 $\boldsymbol{A}>0$ 可知，存在 $\boldsymbol{y}>0$，$\boldsymbol{A}\boldsymbol{y}=r(\boldsymbol{A})\boldsymbol{y}$，令 $\alpha=\sum\limits_{i=1}^{n}y_i$，$\boldsymbol{x}=\dfrac{1}{\alpha}\boldsymbol{y}$，则有

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x},\quad 且\quad \sum_{i=1}^{n}x_i=1.$$

13. 设 $\boldsymbol{A}$ 为 n 阶正矩阵，若存在某个 $\boldsymbol{x}\in\mathbf{C}^n$，$\boldsymbol{x}\geqslant0$，$\boldsymbol{x}\neq0$，$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，试证 $\boldsymbol{x}$ 为 Perron 向量的倍数且 $\lambda=r(\boldsymbol{A})$.

证　由 $\boldsymbol{A}>0$，$\boldsymbol{x}\geqslant0$，$\boldsymbol{x}\neq0$，可知 $\boldsymbol{A}\boldsymbol{x}>0$，而 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，故 $\boldsymbol{x}>0$. 设 $\boldsymbol{y}\in\mathbf{C}^n$，$\boldsymbol{y}>0$，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}=r(\boldsymbol{A})\boldsymbol{y}$，则有

$$\lambda(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{A}\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y})=(\boldsymbol{x},r(\boldsymbol{A})\boldsymbol{y})=r(\boldsymbol{A})(\boldsymbol{x},\boldsymbol{y}),$$

即

$$(\lambda-r(\boldsymbol{A}))(\boldsymbol{x},\boldsymbol{y})=0,$$

但 $(\boldsymbol{x},\boldsymbol{y})>0$，故 $\lambda=r(\boldsymbol{A})$. 由 $V_{r(\boldsymbol{A})}=\{\boldsymbol{x}\mid\boldsymbol{A}\boldsymbol{x}=r(\boldsymbol{A})\boldsymbol{x}\}$ 为一维子空间可知，$\boldsymbol{x}$ 为 Perron 向量的倍数.

14. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，$\boldsymbol{x}\in\mathbf{R}^n$，$\boldsymbol{A}\geqslant0$，$\boldsymbol{x}\geqslant0$，$\beta\geqslant0$，若 $\boldsymbol{A}\boldsymbol{x}<\beta\boldsymbol{x}$ ($\boldsymbol{A}\boldsymbol{x}\leqslant\beta\boldsymbol{x}$)，证明

$$r(\boldsymbol{A})<\beta\ (r(\boldsymbol{A})\leqslant\beta)$$

不一定成立.

证　令 $\boldsymbol{A}=\begin{pmatrix}1&0\\0&2\end{pmatrix}$，$\boldsymbol{x}=\begin{pmatrix}1\\0\end{pmatrix}$，则 $\boldsymbol{A}\boldsymbol{x}=\begin{pmatrix}1&0\\0&2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}$，故 $1\cdot\boldsymbol{x}\leqslant\boldsymbol{A}\boldsymbol{x}=1\cdot\boldsymbol{x}\leqslant1\cdot\boldsymbol{x}$，但 $r(\boldsymbol{A})=2>1$.

15. 求出 $\boldsymbol{A}=\begin{pmatrix}1&2&0&0&2\\0&0&4&0&0\\0&0&0&2&0\\0&8&0&0&0\\4&0&0&0&0\end{pmatrix}$ 的谱半径.

解　由矩阵 $\boldsymbol{A}$ 的特征多项式为 $(\lambda-4)^2(\lambda-3)(\lambda^2+4\lambda+7)$，易得矩阵 $\boldsymbol{A}$ 的谱半径为 $r(\boldsymbol{A})=4$.

16. 证明矩阵 $\boldsymbol{A}=\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}$是素矩阵，求它的特征值，给出 $r(\boldsymbol{A})$的上界，讨论 $r(\boldsymbol{A})$的精确值是什么？

解 $\boldsymbol{A}^2=\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}=\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}>0$，故 $\boldsymbol{A}$ 是素矩阵. $\boldsymbol{A}$ 的特征值为$\dfrac{1\pm\sqrt{5}}{2}$，$r(\boldsymbol{A})=\dfrac{1+\sqrt{5}}{2}$，$r(\boldsymbol{A})\leqslant 2$.

17. 证明：若 $\boldsymbol{A}\geqslant 0$ 以及 $\boldsymbol{A}^k>0$，则 $\boldsymbol{A}^m>0$，$\forall m\geqslant k$.

证 若存在 $m(m\geqslant k)$，使得 $\boldsymbol{A}^m\geqslant 0$，且某个元素为 0，则由 $\boldsymbol{A}^m=\boldsymbol{A}^{m-k}\boldsymbol{A}^k\geqslant 0$，以及 $\boldsymbol{A}^k>0$，可知 $\boldsymbol{A}^{m-k}$的某行元素全为 0.

若 $m-k<k$，则由此可知 $\boldsymbol{A}^k$ 的对应行元素皆为 0，这与 $\boldsymbol{A}^k>0$ 矛盾！

若 $m-k\geqslant k$，令 $m=k\cdot l+r$（r 为 m 除以 k 的余数），则 $\boldsymbol{A}^m=\boldsymbol{A}^r(\boldsymbol{A}^k)^l$，从而可知 $\boldsymbol{A}^r$ 的某行元素全为 0，由此得到 $\boldsymbol{A}^k$ 的某行元素全为 0，这与 $\boldsymbol{A}^k>0$ 矛盾！故 $\boldsymbol{A}^m>0$，$\forall m\geqslant k$.

附录

复习题一

一、判断题(40 分)(对者打√，错者打×)

1. 设 $\boldsymbol{x}\in\mathbf{C}^n$，$\boldsymbol{U}$ 为 n 阶酉矩阵，则 $\|\boldsymbol{Ux}\|_2=\|\boldsymbol{x}\|_2$. (　　)

2. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$,则 $\|\boldsymbol{A}\|_{m_2}^2\geqslant\sum\limits_{i=1}^{n}|\lambda_i|^2$. (　　)

3. 如果 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbf{C}^n$，则 $\|\boldsymbol{x}\|=|x_1|^2$ 为向量范数. (　　)

4. 设 $\boldsymbol{x}\in\mathbf{C}^n$，$\|\boldsymbol{x}\|_\infty\leqslant\|\boldsymbol{x}\|_1\leqslant n\|\boldsymbol{x}\|_\infty$. (　　)

5. 设 $\boldsymbol{A}$ 为 n 阶酉矩阵，则 $\boldsymbol{AA}^+=\boldsymbol{A}^+\boldsymbol{A}=\boldsymbol{E}$. (　　)

6. 若 $\boldsymbol{A}\in\mathbf{C}_r^{m\times r}$，则 $\boldsymbol{A}_{\mathrm{L}}^{-1}=(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{A}^{\mathrm{H}}$. (　　)

7. 若 $\|\cdot\|$ 为算子范数，则 $\|\boldsymbol{A}\|^{-1}\geqslant\|\boldsymbol{A}^{-1}\|$. (　　)

8. $\begin{pmatrix}1 & \mathrm{i}\\ \mathrm{i} & 1\end{pmatrix}$ 和 $\begin{pmatrix}\mathrm{i} & \mathrm{i}\\ \mathrm{i} & 1\end{pmatrix}$ 都是复对称矩阵，故均为正规矩阵. (　　)

9. 设 $r(\boldsymbol{A})$为矩阵 $\boldsymbol{A}$ 的谱半径，则 $r(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|_{m_\infty}$. (　　)

10. 设 $\|\cdot\|_m$为自相容矩阵范数,$\boldsymbol{a}$ 为任一适当维数的非零向量,则 $\|\boldsymbol{x}\|=\|\boldsymbol{xa}^{\mathrm{H}}\|_m$ 是与 $\|\cdot\|_m$相容的向量范数,其中 $\boldsymbol{x}$ 和 $\boldsymbol{a}$ 为相同维数的列向量。 (　　)

二、设 $\boldsymbol{A}$ 是幂等矩阵($\boldsymbol{A}^2=\boldsymbol{A}$)，但 $\boldsymbol{A}\neq\boldsymbol{E}$,证明 $\boldsymbol{A}$ 不是严格对角占优矩阵.(10 分)

三、设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$,$\lambda$ 是 $\boldsymbol{B}=(|a_{ij}|)$的特征值，且存在向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ $(x_i>0,i=1,2,\cdots,n)$使得$\boldsymbol{Bx}=\lambda\boldsymbol{x}$，记 $\boldsymbol{D}=\mathrm{diag}(x_1,x_2,\cdots,x_n)$. 证明 $\boldsymbol{D}^{-1}\boldsymbol{BD}$ 的每个 Gerschgorin 圆都经过 λ.(10 分)

四、证明如果 $\boldsymbol{A}$ 为正规矩阵，且对所有 $\boldsymbol{x}\in\mathbf{C}^n$，都有 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{Ax}\leqslant 0$，那么 $\boldsymbol{A}$ 的所有特征值均非正. 如果有 $\mathrm{tr}(\boldsymbol{A})=0$，那么 $\boldsymbol{A}=\boldsymbol{0}$.(10 分)

五、求矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & 0 & 1\\ -1 & 1 & -1\\ 1 & -1 & 1\end{pmatrix}$的最大秩分解，并求 $\boldsymbol{A}^+$.(10 分)

六、求矩阵 $\boldsymbol{A}=\begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$ 的谱分解，并计算 $\boldsymbol{A}^{10}$.（10 分）

七、(1) 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$, $\boldsymbol{B}=\begin{pmatrix} \boldsymbol{0} & \boldsymbol{A} \\ \boldsymbol{A}^{H} & \boldsymbol{0} \end{pmatrix}$，证明 $\|\boldsymbol{B}\|_2=\|\boldsymbol{A}\|_2$.（5 分）

(2) 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$，$\boldsymbol{A}$ 是非奇异矩阵，$\boldsymbol{B}$ 是奇异矩阵，如果 $\|\cdot\|$ 是任意自相容的矩阵范数，证明 $\|\boldsymbol{A}-\boldsymbol{B}\|\geqslant 1/\|\boldsymbol{A}^{-1}\|$.（5 分）

复习题二

一、判断题(40 分)(对者打√，错者打×)

1. 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$ 的奇异值分别为 $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_n>0$，$\sigma_1'\geqslant\sigma_2'\geqslant\cdots\geqslant\sigma_n'>0$，如果 $\sigma_i>\sigma_i'$ $(i=1,2,\cdots,n)$，则 $\|\boldsymbol{A}^{+}\|_2>\|\boldsymbol{B}^{+}\|_2$.（　　）

2. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 为正规矩阵，则 $\boldsymbol{A}$ 的谱半径为 $r(\boldsymbol{A})=\|\boldsymbol{A}\|_2$.（　　）

3. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 可逆，$\boldsymbol{B}\in\mathbf{C}^{n\times n}$，若对算子范数 $\|\cdot\|$，有 $\|\boldsymbol{A}^{-1}\|\cdot\|\boldsymbol{B}\|<1$，则 $\boldsymbol{A}+\boldsymbol{B}$ 可逆.（　　）

4. 设 $\boldsymbol{A}=\begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$ 为一非零实矩阵，则 $-(a_1^2+a_2^2+a_3^2)^{-1}\boldsymbol{A}$ 为 $\boldsymbol{A}$ 的一个广义逆矩阵.（　　）

5. 设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，$\boldsymbol{P}$ 为 m 阶酉矩阵，则 $\boldsymbol{PA}$ 与 $\boldsymbol{A}$ 有相同的奇异值.（　　）

6. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，且 $\boldsymbol{A}$ 的所有列和都相等，则 $r(\boldsymbol{A})=\|\boldsymbol{A}\|_\infty$.（　　）

7. 如果 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbf{C}^n$，则 $\|\boldsymbol{x}\|=\min\limits_{1\leqslant i\leqslant n}|x_i|$ 是向量范数.（　　）

8. $\boldsymbol{A}=\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 6 & 2 \\ 0 & 1 & 1 & 8 \end{pmatrix}$ 至少有 2 个实特征值.（　　）

9. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，则矩阵范数 $\|\boldsymbol{A}\|_{m_\infty}$ 与向量的 1-范数相容.（　　）

10. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 是不可逆矩阵，则对任意一个自相容矩阵范数 $\|\cdot\|$，有 $\|\boldsymbol{E}-\boldsymbol{A}\|\geqslant 1$，其中 $\boldsymbol{E}$ 为单位矩阵.（　　）

二、设 $\boldsymbol{A}$ 是 Hermite 矩阵($\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$)，且 $\boldsymbol{A}$ 的特征值满足 $\lambda_1=\lambda_2=\cdots=\lambda_n$，证明矩阵 $\boldsymbol{A}$ 的 Rayleigh 商恒等于 λ_1.（5 分）

三、已知 $\mathbf{C}^{n\times n}$ 中的两种矩阵算子范数 $\|\cdot\|_a$ 与 $\|\cdot\|_b$，对于任意矩阵 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，验证

$$\|A\| = \|A\|_a + \|A\|_b$$

是 $\mathbf{C}^{n\times n}$ 中的相容矩阵范数.(10 分)

四、用 Gerschgorin 圆盘定理证明矩阵

$$A=\begin{pmatrix} 2 & \frac{1}{n} & \vdots & \frac{1}{n} & \frac{1}{n} \\ \frac{1}{2n} & 4 & \cdots & \frac{1}{n} & \frac{1}{n} \\ \vdots & \vdots & & \vdots & \frac{1}{n} \\ \frac{1}{(n-1)n} & \frac{1}{(n-1)n} & \cdots & 2(n-1) & \frac{1}{n} \\ \frac{1}{n^2} & \frac{1}{n^2} & \cdots & \frac{1}{n^2} & 2n \end{pmatrix} \in \mathbf{C}^{n\times n}$$

的特征值为两两不相等的正实数.(10 分)

五、证明:

(1) $\begin{pmatrix} A \\ 0 \end{pmatrix}^{+} = (A^{+} \quad 0)$,其中 $A\in \mathbf{C}^{m\times n}$;(5 分)

(2) 设 λ 是矩阵 $A\in \mathbf{C}^{m\times m}$ 的任意特征值,则 $|\lambda| \leqslant \sqrt[m]{\|A^m\|}$,其中矩阵范数 $\|\cdot\|$ 是算子范数.(5 分)

六、设矩阵 $A\in \mathbf{C}^{m\times n}$ 的非零奇异值为 $\sigma_1,\sigma_2,\cdots,\sigma_r(r>0)$,证明:

$$\|A\|_F = \left(\sum_{i=1}^{r}\sigma_i^2\right)^{\frac{1}{2}}.\text{(10 分)}$$

七、已知

$$A=\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \end{pmatrix},\quad b=\begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix},$$

(1) 求矩阵 A 的最大秩分解;

(2) 求 A^{+};

(3) 用广义逆矩阵方法判断方程组 $Ax=b$ 是否有解.(15 分)

复习题三

一、判断题(40 分)(对者打√,错者打×)

1. A 为 n 阶实对称矩阵,对 $\mathbf{R}^n$ 中的列向量 x,定义 $\|x\| = \sqrt{x^{\mathrm{T}}Ax}$,则 $\|x\|$ 为向量 x 的范数. (　　)

2. 设 $\boldsymbol{A}$ 为 n 阶 Hermite 矩阵，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 是矩阵 $\boldsymbol{A}$ 的特征值，则 $\|\boldsymbol{A}\|_{m_2}^2=\sum\limits_{i=1}^n\lambda_i^2$.

( )

3. 如果 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，且 $\boldsymbol{A}\neq\boldsymbol{0}$，$(\boldsymbol{AA}^-)^{\mathrm{H}}=\boldsymbol{AA}^-$，则 $\|\boldsymbol{AA}^-\|_2=n$. ( )

4. 设 $\|\cdot\|_a$ 为从属于向量范数 $\|\boldsymbol{x}\|_a$ 的算子范数，$\boldsymbol{H}=\boldsymbol{E}-2\boldsymbol{uu}^{\mathrm{H}}$，其中 $\boldsymbol{E}$ 为 n 阶单位矩阵，$\boldsymbol{u}\in\mathbf{C}^n$，且 $\|\boldsymbol{u}\|_2=1$，则 $\|\boldsymbol{H}\|_a=n$. ( )

5. 设 $\boldsymbol{A}=\begin{pmatrix}1/5 & 1/5 & 1/5 & 1/5\\ 1/6 & 2/6 & 1/6 & 1/6\\ 1/7 & 1/7 & 3/7 & 1/7\\ 1/8 & 1/8 & 1/8 & 4/8\end{pmatrix}$，则矩阵 $\boldsymbol{A}$ 的谱半径满足 $r(A)<1$. ( )

6. 若 $\boldsymbol{A}\in\mathbf{C}^{m\times m}(m>1)$ 严格对角占优，则 $\boldsymbol{A}$ 的谱半径满足 $r(\boldsymbol{A})<\|2\boldsymbol{A}\|_{m_\infty}$. ( )

7. 若设 $x\in\mathbf{R}^n$，则 $\|\boldsymbol{x}\|_2\leqslant\|\boldsymbol{x}\|_1\leqslant\sqrt{n}\|\boldsymbol{x}\|_2$. ( )

8. 设 $\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0 & 1\\ 1 & 1 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1\end{pmatrix}$，则 $\|\boldsymbol{A}^+\|_{m_1}=1$. ( )

9. 设 $\boldsymbol{G}$ 为矩阵 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}(r<n)$ 的广义逆 $\boldsymbol{A}$ ，$\boldsymbol{A}=\boldsymbol{BD}$ 为 $\boldsymbol{A}$ 的最大秩分解，则 $\mathrm{rank}(\boldsymbol{DGB})=n$. ( )

10. 设 $\boldsymbol{A}=\begin{pmatrix}0.9 & 0.01 & 0.12\\ 0.01 & 0.8 & 0.13\\ 0.01 & 0.02 & 0.4\end{pmatrix}$，则 $\boldsymbol{A}$ 的特征值均为实数. ( )

二、设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\|\boldsymbol{A}\|=\sqrt{mn}\max\limits_{i,j}|a_{ij}|$，证明：

(1) $\|\boldsymbol{A}\|$ 为矩阵范数；

(2) $\|\boldsymbol{A}\|$ 与向量 2-范数相容.（10 分）

三、设 $\boldsymbol{A}=\begin{pmatrix}1/4 & 1/4 & 1/4 & 1/4\\ 1/5 & 2/5 & 1/5 & 1/5\\ 1/6 & 1/6 & 3/6 & 1/6\\ 1/7 & 1/7 & 1/7 & 4/7\end{pmatrix}$，利用盖尔圆盘定理计算矩阵 $\boldsymbol{A}$ 的谱半径 $r(\boldsymbol{A})$.（10 分）

四、试证：如果 $\boldsymbol{A}$ 为 n 阶正规矩阵，且 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ 以及 $\boldsymbol{Ay}=\mu\boldsymbol{y}$，其中 $\lambda\neq\mu$，那么 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交.（10 分）

五、设 $D\in\mathbf{C}^{m\times n}$ 为列满秩矩阵，$\boldsymbol{D}^+$ 为 M-P 广义逆，$\boldsymbol{A}\in\mathbf{C}^{n\times n}$，证明：$\|\boldsymbol{A}\|=\|\boldsymbol{DAD}^+\|_2$ 为 $\mathbf{C}^{n\times n}$ 上的矩阵范数.（10 分）

六、设有线性方程组$\begin{cases}x_1+2x_2+2x_2+x_4=1,\\ x_1+x_2+x_3+x_4=1,\\ 2x_1+x_2+x_3+2x_4=1,\end{cases}$

(1) 求方程组的系数矩阵 $\boldsymbol{A}$ 的最大秩分解；

(2) 计算 $\boldsymbol{A}^+$；

(3) 求线性方程组的最佳逼近解. (10 分)

七、(1) 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}(n>1)$为严格对角占优矩阵，$\boldsymbol{D}=\mathrm{diag}(a_{11},a_{22},\cdots,a_{nn})$，其中 $a_{ii}(i=1,2,\cdots,n)$为 $\boldsymbol{A}$ 的对角元素，$\boldsymbol{E}$ 为 n 阶单位矩阵，则存在一个矩阵范数 $\|\cdot\|$，使得 $r(\boldsymbol{E}-\boldsymbol{D}^{-1}\boldsymbol{A})<1$. (5 分)

(2) 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，$\varepsilon$ 为任意给定的正数，$r(\boldsymbol{A})$为矩阵的谱半径. 证明：至少存在一个矩阵范数 $\|\boldsymbol{A}\|$，使得 $\|\boldsymbol{A}\|\leqslant r(\boldsymbol{A})+\varepsilon$. (5 分)

复习题四

一、判断题(40 分)(对者打√，错者打×)

1. 设矩阵 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，且 $\boldsymbol{A}\neq\boldsymbol{0}$，则 $\|\boldsymbol{A}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{+}\boldsymbol{A}^{\mathrm{H}}\|_2=1$. (　　)

2. 设 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$ 的奇异值为 $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_n$，则 $\|\boldsymbol{A}\|_2^2=\sum\limits_{i=1}^{n}\sigma_i^2$. (　　)

3. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，且有某种算子范数 $\|\cdot\|$，使得 $\|\boldsymbol{A}\|<1$，则 $\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\|>\dfrac{1}{1-\|\boldsymbol{A}\|}$，其中 $\boldsymbol{E}$ 为 n 阶单位矩阵. (　　)

4. 设 $\boldsymbol{A}=\boldsymbol{E}-2\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}}$，$\boldsymbol{u}\in\mathbf{C}^n$ 且 $\|\boldsymbol{u}\|_2=1$，则 $\|\boldsymbol{A}\|_{m_2}=\sqrt{n}$. (　　)

5. 设 $\boldsymbol{A}=\begin{pmatrix}1&2&3&4\\2&4&6&8\\1&1&0&1\end{pmatrix}$，则 $\boldsymbol{A}$ 的 M-P 广义逆 $\boldsymbol{A}^+$ 的秩满足 $\mathrm{rank}(\boldsymbol{A}^+)=1$. (　　)

6. 若 $\boldsymbol{A}$ 为列满秩矩阵，则$(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{H}}$ 既是 $\boldsymbol{A}$ 的左逆又是 $\boldsymbol{A}$ 的 M-P 广义逆 $\boldsymbol{A}^+$. (　　)

7. 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为线性空间 V^n 的一组基，$x=x_1\varepsilon_1+x_2\varepsilon_2+\cdots x_n\varepsilon_n\in V^n$，则 $\|x\|=k_1|x_1|^2+k_2|x_2|^2+\cdots k_n|x_n|^2(k_i\geqslant0)$是 V^n 上向量 x 的范数. (　　)

8. 设 $\boldsymbol{A}=\begin{pmatrix}30&3&2\\2&10&2\\8&1&0\end{pmatrix}$，则 $\boldsymbol{A}$ 有 3 个实特征值. (　　)

9. 设 $\boldsymbol{G}$ 为矩阵 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}(r<n)$的广义逆 $\boldsymbol{A}^-$，$\boldsymbol{A}=\boldsymbol{B}\boldsymbol{D}$ 为 $\boldsymbol{A}$ 的最大秩分解，则 $\|\boldsymbol{D}\boldsymbol{G}\boldsymbol{B}\|_2=r$. (　　)

10. 设 $A=(a_{ij})\in C^{n\times n}(n>1)$ 为严格对角占优矩阵，$D=\mathrm{diag}(a_{11},a_{22},\cdots,a_{nn})$，$B=E-D^{-1}A$，则 B 的谱半径满足 $r(B)\geqslant 1$.　（　　）

二、设矩阵 U 是酉矩阵，$A=\mathrm{diag}(a_1,a_2,\cdots,a_n)$，证明：$UA$ 的任意特征值 λ 都满足不等式 $\min\limits_i\{|a_i|\}\leqslant|\lambda|\leqslant\max\limits_i\{|a_i|\}$.（10 分）

三、设 $\|\cdot\|_a$ 是 $C^{n\times n}$ 上相容的矩阵范数，矩阵 B,C 都是 n 阶可逆矩阵，且 $\|B^{-1}\|_a$ 及 $\|C^{-1}\|_a$ 都小于或等于 1，证明：对任意矩阵 $A\in C^{n\times n}$，表达式

$$\|A\|_b=\|BAC\|_a$$

定义了 $C^{n\times n}$ 上的一个相容的矩阵范数.（10 分）

四、已知矩阵 $A=\begin{pmatrix}0&1&-1\\-1&0&1\\1&-1&0\\0&1&-1\end{pmatrix}$，$b=\begin{pmatrix}1\\-1\\0\\1\end{pmatrix}$.

(1) 求矩阵 A 的最大秩分解；

(2) 求 A^+；

(3) 用广义逆矩阵方法判断方程组 $Ax=b$ 是否有解；

(4) 求方程组 $Ax=b$ 的最小范数解或最佳逼近解(要求指出所求的是哪种解).（10 分）

五、用 Gerschgorin 圆盘定理证明：矩阵

$$A=\begin{pmatrix}2&\frac{1}{2}&\frac{1}{2^2}&\frac{1}{2^3}&\frac{1}{2^4}\\ \frac{2}{3}&4&\frac{2}{3^2}&\frac{2}{3^3}&\frac{2}{3^4}\\ \frac{3}{4}&\frac{3}{4^2}&6&\frac{3}{4^3}&\frac{3}{4^4}\\ \frac{4}{5}&\frac{4}{5^2}&\frac{4}{5^3}&8&\frac{4}{5^4}\\ \frac{5}{6}&\frac{5}{6^2}&\frac{5}{6^3}&\frac{5}{6^4}&10\end{pmatrix}$$

能够相似于对角矩阵且 A 的特征值都是正实数.（10 分）

六、设矩阵 $A,B\in C^{m\times n}$，$\sigma_1(M)$ 表示矩阵 $M\in C^{m\times n}$ 的最大奇异值，证明：

(1) $\sigma_1(AB)\leqslant\sigma_1(A)\cdot\sigma_1(B)$；

(2) $\sigma_1(A+B)\leqslant\sigma_1(A)+\sigma_1(B)$.（10 分）

七、(1) 设 A 是可逆矩阵，λ 是 A 的一个特征值，对于任意的算子范数 $\|\cdot\|$，证明 $|\lambda|\geqslant\dfrac{1}{\|A^{-1}\|}$.（5 分）

(2) 设矩阵 $A=(a_{ij})\in \mathbf{R}^{n\times n}$，$R_i=\sum_{j=1}^{n}|a_{ij}|$，证明：$\operatorname{rank}(A)\geqslant \sum_{i=1}^{n}\frac{|a_{ii}|}{R_i}$，这里 $\operatorname{rank}(A)$ 表示矩阵 A 的秩，约定在和式中 $\frac{0}{0}=0$.（5 分）

复习题五

一、判断题(40 分)(对者打√，错者打×)

1. 设 $A=UDV$，其中 U,V 为酉矩阵，则 $A^{+}=V^{H}D^{+}U^{H}$.（　　）
2. 设 $A\in \mathbf{C}^{n\times n}$是不可逆矩阵，则对任意的矩阵范数 $\|\cdot\|$，有 $\|E-A\|\geqslant 1$.（　　）
3. 如果 $A\in \mathbf{C}^{n\times n}$，则 $\|x\|=\|Ax\|_2(x\in \mathbf{R}^n)$是 $\mathbf{R}^n$ 中的向量范数.（　　）
4. 设 $A=\begin{pmatrix}\frac{\pi}{2} & 0 & 0 & 0\\ 0 & -\frac{\pi}{2} & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{pmatrix}$，则 $\sin A=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{pmatrix}$.（　　）
5. 设 $A\in \mathbf{C}_n^{m\times n}$，则 $A^{+}A$ 的特征值全是 1.（　　）
6. 任何矩阵都有自反广义逆.（　　）
7. 方阵 A 的任意盖尔圆不一定都包含 A 的特征值.（　　）
8. 任意一个实对称矩阵 A 的主对角元素大于零，则 A 至少有一个正特征值.（　　）
9. 任意方阵的谱半径 $r(A)$小于等于任何矩阵范数 $\|A\|$.（　　）
10. 奇异矩阵存在 QR 分解.（　　）

二、设 A 是 n 阶实对称矩阵，且 A 的特征值分别为 $1,\sqrt{2},\cdots,\sqrt{n}$，试求 $\|A\|_F$，$r(A)$，$\|A\|_2$，$\operatorname{cond}(A)_2$，$\|(A^{-1})^m\|_2$.（10 分）

三、设 n 阶方阵 $A=(a_{ij})_{n\times n}$，且 $\sum_{j=1}^{n}|a_{ij}|<1(i=1,2,\cdots,n)$，证明：$A$ 的每一个特征值 λ 满足 $|\lambda|<1$.（10 分）

四、已知矩阵 $A=\begin{pmatrix}1 & 0 & 1 & 1\\ 2 & 1 & 2 & 1\\ 2 & 0 & 2 & 2\\ 4 & 2 & 4 & 2\end{pmatrix}$，$b=\begin{pmatrix}-1\\ 5\\ -2\\ 10\end{pmatrix}$.

(1) 求矩阵 A 的最大秩分解；

(2) 求 A^{+}；

(3) 用广义逆矩阵方法判断方程组 $Ax=b$ 是否有解；

(4) 求方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小范数解或最佳逼近解(要求指出所求的是哪种解). (15 分).

五、设 $A=\begin{pmatrix} \frac{1}{\sqrt{2}} & a & 0 \\ 0 & 0 & 1 \\ b & c & 0 \end{pmatrix}$，问 a，b，c 为何值时，$\boldsymbol{A}$ 为正交矩阵？(10 分)

六、设 $\boldsymbol{A}\in\mathbf{R}^{n\times m}$，$\boldsymbol{B}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$，试证：

(1) $\boldsymbol{B}$ 是半正定矩阵；

(2) 当 $\boldsymbol{A}$ 的列向量组线性无关时，$\boldsymbol{B}$ 为正定矩阵. (10 分)

七、证明：$(\boldsymbol{A}^{+})^{+}=\boldsymbol{A}$. (5 分)

复习题六

一、判断题(40 分)(对者打√，错者打×)

1. 设$(\boldsymbol{x},\boldsymbol{y})$是 $V^n(\mathbf{R})$上的内积且 $\|\boldsymbol{x}\|=\sqrt{(\boldsymbol{x},\boldsymbol{x})}$，则 $\|\boldsymbol{x}-\boldsymbol{y}\|^2+\|\boldsymbol{x}+\boldsymbol{y}\|^2=2(\|\boldsymbol{x}\|^2+\|\boldsymbol{y}\|^2)$. (　　)

2. 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，则 $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})$. (　　)

3. 如果 $\boldsymbol{x}\in\mathbf{C}^n$，$\|\boldsymbol{x}\|_a$，$\|\boldsymbol{x}\|_b$是 $\mathbf{C}^n$ 上的两个向量范数，则 $\|\boldsymbol{x}\|=\|\boldsymbol{x}\|_a+k\|\boldsymbol{x}\|_b$，$k\in\mathbf{R}$ 是向量范数. (　　)

4. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，必有 $\|\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\|_2=\|\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\|_2=\|\boldsymbol{A}\|_2^2$. (　　)

5. 设 $\boldsymbol{A}$ 为 n 阶矩阵，则 $\boldsymbol{A}\boldsymbol{A}^{+}=\boldsymbol{A}^{+}\boldsymbol{A}$. (　　)

6. 若 $\boldsymbol{A}\in\mathbf{C}_m^{m\times r}$，则 $\boldsymbol{A}_{\mathrm{L}}^{-1}=(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{H}}$. (　　)

7. 若 $\|\cdot\|$ 为算子范数，则 $\|\boldsymbol{E}\|=1$，$\boldsymbol{E}$ 为单位矩阵. (　　)

8. 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，且 $\boldsymbol{A}$ 的所有列和都相等，则 $r(\boldsymbol{A})=\|\boldsymbol{A}\|_\infty$. (　　)

9. 一个正规矩阵若是三角矩阵，则它一定是对角矩阵. (　　)

10. 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$，且 $\boldsymbol{A}<\boldsymbol{B}$，则 $r(\boldsymbol{A})<r(\boldsymbol{B})$. (　　)

二、设 $\boldsymbol{A},\boldsymbol{B}$ 是两个 n 阶正交矩阵，且 $\det(\boldsymbol{AB})=-1$，试证明：

(1) $\det(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B})=\det(\boldsymbol{AB})=\det(\boldsymbol{AB}^{\mathrm{T}})=-1$；(5 分)

(2) $\det(\boldsymbol{A}+\boldsymbol{B})=0$. (5 分)

三、用 Gerschgorin 圆定理证明 $\boldsymbol{A}=\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 0 & 6 & 2 \\ 0 & 1 & 1 & 8 \end{pmatrix}$ 至少有 2 个实特征值. (10 分)

四、已知矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & -1 & 0 & -1 & 1\\ 0 & 0 & 2 & -2 & 2\\ -1 & 1 & -1 & 2 & -1\end{pmatrix}$，$b=\begin{pmatrix}2\\ -2\\ -1\end{pmatrix}$，

(1) 求矩阵 $\boldsymbol{A}$ 的最大秩分解；

(2) 求 $\boldsymbol{A}^{+}$；

(3) 用广义逆矩阵方法判断方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 是否有解；

(4) 求方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的最小范数解或最佳逼近解(要求指出所求的是哪种解).（15 分）

五、设列向量 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbf{R}^n$，证明：$\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|_2=\|\boldsymbol{\alpha}\|_2+\|\boldsymbol{\beta}\|_2$ 的充分必要条件是 $\boldsymbol{\alpha}=k\boldsymbol{\beta}$，$0<k\in\mathbf{R}$，且 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}\geqslant 0$.（10 分）

六、证明：$\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{M}=\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{N}$ 的充分必要条件是 $\boldsymbol{A}\boldsymbol{M}=\boldsymbol{A}\boldsymbol{N}$.（5 分）

七、证明：(1) 设 $\boldsymbol{A}=\begin{pmatrix}0 & c & c\\ c & 0 & c\\ c & c & 0\end{pmatrix}$. 讨论 c 为何值时 $\boldsymbol{A}$ 为收敛矩阵；(5 分)

(2) 设 $\boldsymbol{A}$ 是 n 阶可逆矩阵，$\boldsymbol{B}$ 是 n 阶矩阵，若对某种矩阵范数 $\|\cdot\|$，有 $\|\boldsymbol{B}\|<\dfrac{1}{\|\boldsymbol{A}^{-1}\|}$，证明 $\boldsymbol{A}+\boldsymbol{B}$ 可逆.（5 分）

答　　案

复习题一

一、判断题

1. (√)　$\|\boldsymbol{U}\boldsymbol{x}\|_2^2=(\boldsymbol{U}\boldsymbol{x},\ \boldsymbol{U}\boldsymbol{x})=\boldsymbol{x}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=\|\boldsymbol{x}\|_2^2$.

2. (√)　$\boldsymbol{A}\in\mathbf{C}^{n\times n}\Rightarrow\boldsymbol{A}=\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^{\mathrm{H}}\Rightarrow\|\boldsymbol{A}\|_{m_2}^2=\|\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^{\mathrm{H}}\|_{m_2}^2=\|\boldsymbol{R}\|_{m_2}^2\geqslant\|\boldsymbol{R}^2\|_{m_2}=\sum\limits_{i=1}^{n}|\lambda_i|^2$.

3. (×)　例如 $\boldsymbol{x}=(0,1,0,\cdots,0)\neq\boldsymbol{0}$，但 $\|\boldsymbol{x}\|=0$.

4. (√)　$\|\boldsymbol{x}\|_\infty=\max\limits_i|x_i|\leqslant\sum\limits_{i=1}^{n}|x_i|=\|\boldsymbol{x}\|_1\leqslant n\max\limits_i|x_i|=n\|\boldsymbol{x}\|_\infty$.

5. (√)　因为 $\boldsymbol{A}^{+}=\boldsymbol{A}^{\mathrm{H}}$，故结论成立.

6. (×)　$\boldsymbol{A}_{\mathrm{L}}^{-1}=(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{H}}$，故结论不成立.

7. (×)　$1=\|\boldsymbol{A}\boldsymbol{A}^{-1}\|\leqslant\|\boldsymbol{A}\|\ \|\boldsymbol{A}^{-1}\|$，故结论不成立.

8. (×)　$\begin{pmatrix}1 & \mathrm{i}\\ \mathrm{i} & 1\end{pmatrix}$为正规矩阵而$\begin{pmatrix}\mathrm{i} & \mathrm{i}\\ \mathrm{i} & 1\end{pmatrix}$非正规，因为$\begin{pmatrix}\mathrm{i} & \mathrm{i}\\ \mathrm{i} & 1\end{pmatrix}\begin{pmatrix}-\mathrm{i} & -\mathrm{i}\\ -\mathrm{i} & 1\end{pmatrix}\neq\begin{pmatrix}-\mathrm{i} & -\mathrm{i}\\ -\mathrm{i} & 1\end{pmatrix}\begin{pmatrix}\mathrm{i} & \mathrm{i}\\ \mathrm{i} & 1\end{pmatrix}$.

9. (×)　$\boldsymbol{A}=\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}$，则 $\|\boldsymbol{A}\|_{m_\infty}=1$，而 $\rho(\boldsymbol{A})=1.618$.

10. (√) $\|Ax\| = \|Axa^{H}\|_m \leqslant \|A\|_m \|xa^{H}\|_m = \|A\|_m \|x\|$.

二、A 是严格对角占优矩阵$\Rightarrow A$ 可逆$\Rightarrow A=E$，矛盾.

三、

$$D^{-1}BD = \begin{pmatrix} \frac{1}{x_1} & & & \\ & \frac{1}{x_2} & & \\ & & \ddots & \\ & & & \frac{1}{x_n} \end{pmatrix} \begin{pmatrix} |a_{11}| & |a_{12}| & \cdots & |a_{1n}| \\ |a_{21}| & |a_{22}| & \cdots & |a_{2n}| \\ \vdots & \vdots & & \vdots \\ |a_{n1}| & |a_{n2}| & \cdots & |a_{nn}| \end{pmatrix} \begin{pmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{pmatrix}$$

$$= \begin{pmatrix} \frac{|a_{11}|}{x_1} & \frac{|a_{12}|}{x_1} & \cdots & \frac{|a_{1n}|}{x_1} \\ \frac{|a_{21}|}{x_2} & \frac{|a_{22}|}{x_2} & \cdots & \frac{|a_{2n}|}{x_2} \\ \vdots & \vdots & & \vdots \\ \frac{|a_{n1}|}{x_n} & \frac{|a_{n2}|}{x_n} & \cdots & \frac{|a_{nn}|}{x_n} \end{pmatrix} \begin{pmatrix} x_1 & & & \\ & x_2 & & \\ & & \ddots & \\ & & & x_n \end{pmatrix}$$

$$= \begin{pmatrix} |a_{11}| & \frac{|a_{12}|x_2}{x_1} & \cdots & \frac{|a_{1n}|x_n}{x_1} \\ \frac{|a_{21}|x_1}{x_2} & |a_{22}| & \cdots & \frac{|a_{2n}|x_n}{x_2} \\ \vdots & \vdots & & \vdots \\ \frac{|a_{n1}|x_1}{x_n} & \frac{|a_{n2}|x_2}{x_n} & \cdots & |a_{nn}| \end{pmatrix}$$

$\Rightarrow D^{-1}BD$ 的每个 Gerschgorin 圆为 $S_i = \{z \in \mathbf{C}:\ |z-|a_{ii}|| \leqslant R_i\}$，$R_i = \frac{1}{x_i}\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|x_j$，

又 $Bx=\lambda x \Rightarrow |\lambda-|a_{ii}|| = \frac{1}{x_i}\sum_{\substack{j=1\\ j\neq i}}^{n}(|a_{ij}|x_j) = R_i, i=1,2,\cdots,n$，所以结论成立.

四、A 为正规矩阵$\Rightarrow A=U\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)U^{H} \Rightarrow \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)=U^{H}AU \Rightarrow \lambda_i = u_i^{H}Au_i \leqslant 0$；又因为 $\mathrm{tr}(A)=\lambda_1+\lambda_2+\cdots+\lambda_n=0 \Rightarrow \lambda_i=0 \Rightarrow A=\mathbf{0}$.

五、$A=BD=\begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$，$B^{H}B=\begin{pmatrix} 3 & -2 \\ -2 & 2 \end{pmatrix}$，$DD^{H}=\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$，

$$(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B})^{-1}=\begin{pmatrix}1 & 1\\ 1 & 1.5\end{pmatrix},\ (\boldsymbol{D}\boldsymbol{D}^{\mathrm{H}})^{-1}=\begin{pmatrix}0.5 & 0\\ 0 & 1\end{pmatrix},\ \boldsymbol{A}^{+}=\begin{pmatrix}0.5 & 0.5 & -0.5\\ 0.5 & 1 & -1\\ 0.5 & 0.5 & -0.5\end{pmatrix}.$$

六、$|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda+2)(\lambda-1)^2$，$\boldsymbol{P}=\begin{pmatrix}-1 & -2 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\end{pmatrix}$，$\boldsymbol{P}^{-1}=\begin{pmatrix}1 & 2 & 0\\ -1 & -1 & 0\\ -1 & -2 & 1\end{pmatrix}$，

$$\boldsymbol{A}_1=\begin{pmatrix}-1\\ 1\\ 1\end{pmatrix}(1\quad 2\quad 0)=\begin{pmatrix}-1 & -2 & 0\\ 1 & 2 & 0\\ 1 & 2 & 0\end{pmatrix},$$

$$\boldsymbol{A}_2=\begin{pmatrix}-2 & 0\\ 1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}-1 & -1 & 0\\ -1 & -2 & 1\end{pmatrix}=\begin{pmatrix}2 & 2 & 0\\ -1 & -1 & 0\\ -1 & -2 & 1\end{pmatrix},$$

$$\boldsymbol{A}=-2\boldsymbol{A}_1+\boldsymbol{A}_2,\ \boldsymbol{A}^{10}=\begin{bmatrix}-1022 & -2046 & 0\\ 1023 & 2047 & 0\\ 1023 & 2046 & 1\end{bmatrix}.$$

七、(1) $\boldsymbol{B}=\begin{pmatrix}\boldsymbol{0} & \boldsymbol{A}\\ \boldsymbol{A}^{\mathrm{H}} & \boldsymbol{0}\end{pmatrix}\Rightarrow\boldsymbol{B}^{\mathrm{H}}=\begin{pmatrix}\boldsymbol{0} & \boldsymbol{A}\\ \boldsymbol{A}^{\mathrm{H}} & \boldsymbol{0}\end{pmatrix}\Rightarrow\boldsymbol{B}\boldsymbol{B}^{\mathrm{H}}=\begin{pmatrix}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\end{pmatrix}\Rightarrow$

$r(\boldsymbol{B}\boldsymbol{B}^{\mathrm{H}})=r(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})\Rightarrow\|\boldsymbol{B}\|_2=\|\boldsymbol{A}\|_2$.

(2) $\boldsymbol{B}=\boldsymbol{A}-(\boldsymbol{A}-\boldsymbol{B})=\boldsymbol{A}[\boldsymbol{E}-\boldsymbol{A}^{-1}(\boldsymbol{A}-\boldsymbol{B})]\Rightarrow\boldsymbol{E}-\boldsymbol{A}^{-1}(\boldsymbol{A}-\boldsymbol{B})$奇异$\Rightarrow$

$1\leqslant r[\boldsymbol{A}^{-1}(\boldsymbol{A}-\boldsymbol{B})]\leqslant\|\boldsymbol{A}^{-1}\|\cdot\|(\boldsymbol{A}-\boldsymbol{B})\|$.

复习题二

一、判断题

1. (×). 2. (√). 3. (√). 4. (√). 5. (√). 6. (×). 7. (×).
8. (√). 9. (×). 10. (√).

二、因为 $\boldsymbol{A}$ 是 Hermite 矩阵，由 Rayleigh-Ritz 定理知，

$$\lambda_{\max}=\max_{\boldsymbol{x}\neq\boldsymbol{0}}R(\boldsymbol{x}),\ \lambda_{\min}=\min_{\boldsymbol{x}\neq\boldsymbol{0}}R(\boldsymbol{x}),$$

又由 $\boldsymbol{A}$ 的特征值 $\lambda_1=\lambda_2=\cdots=\lambda_n$，所以 $\boldsymbol{A}$ 的 Rayleigh 商恒等于 λ_1.

三、非负性. $\boldsymbol{A}=\boldsymbol{0}$ 时，$\|\boldsymbol{A}\|_a=0$，$\|\boldsymbol{A}\|_b=0$，从而 $\|\boldsymbol{A}\|=0$；$\boldsymbol{A}\neq\boldsymbol{0}$ 时，$\|\boldsymbol{A}\|_a>0$，$\|\boldsymbol{A}\|_b>0$，从而 $\|\boldsymbol{A}\|>0$.

齐次性和三角不等式容易验证成立. 下面证明相容性.

设 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{n\times n}$，则有

$$\|AB\| = \|AB\|_a + \|AB\|_b \leqslant \|A\|_a\|B\|_a + \|A\|_b\|B\|_b$$
$$\leqslant (\|A\|_a + \|A\|_b)(\|B\|_a + \|B\|_b) = \|A\|\|B\|.$$

因此，$\|A\|$ 是矩阵范数.

四、A 的 n 个 Gerschgorin 圆为

$$G_1: |z-2| \leqslant \frac{n-1}{n} < 1; \quad G_i: |z-2i| \leqslant \frac{n-1}{n} < 1, \quad i=2,3,\cdots,n,$$

它们都是孤立的，没有公共部分，所以 A 有 n 个不同的特征值. 因为关于实轴对称，且 A 是实矩阵，故 $G_i(i=1,2,\cdots,n)$ 中只能有实特征值.

五、(1) 由定义直接验证即可；

(2) 设 A 的属于 λ 的特征向量为 x，即 $Ax=\lambda x$，进一步有 $A^m x=\lambda^m x$. 取向量范数 $\|x\|$ 和题中所说的矩阵范数相容,则有 $|\lambda|^m\|x\| = \|\lambda^m x\| = \|A^m x\| \leqslant \|A^m\|\|x\|$，即 $|\lambda|^m \leqslant \|A^m\|$，也就是 $|\lambda| \leqslant \sqrt[m]{\|A^m\|}$.

六、由 A 的奇异值分解可知，存在 m 阶酉矩阵 U 及 n 阶酉矩阵 V，使得 $A=U\begin{pmatrix}D & 0\\ 0 & 0\end{pmatrix}V$，其中 $D=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_r)$，则有

$$\|A\|_F^2 = \left\|U\begin{pmatrix}D^2 & 0\\ 0 & 0\end{pmatrix}U^H\right\| = \left\|\begin{pmatrix}D^2 & 0\\ 0 & 0\end{pmatrix}\right\| = \sum_{i=1}^r \sigma_i^2,$$

即 $\|A\|_F = \left(\sum_{i=1}^r \sigma_i^2\right)^{\frac{1}{2}}$.

七、(1) $A=\begin{pmatrix}1 & 0\\ 0 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 1 & 0 & 1\\ 0 & 1 & 1 & 0\end{pmatrix}=BD$；

(2) $B^+=(B^TB)^{-1}B^T=\dfrac{1}{3}\begin{pmatrix}2 & -1 & 1\\ -1 & 2 & 1\end{pmatrix}$， $D^+=D^T(DD^T)^{-1}=\dfrac{1}{5}\begin{pmatrix}2 & -1\\ 1 & 2\\ -1 & 3\\ 2 & -1\end{pmatrix}$，

$$A^+=D^+B^+=\frac{1}{15}\begin{pmatrix}5 & -4 & 1\\ 0 & 3 & 3\\ -5 & 7 & 2\\ 5 & -4 & 1\end{pmatrix};$$

(3) $AA^+b=A(1,1,0,1)^T=(3,1,4)^T=b$，故 $Ax=b$ 有解.

复习题三

一、判断题

1. (×) 因为非负性不成立，故结论错误.

2. (√) A 为 n 阶 Hermite 矩阵 $\Rightarrow \|A\|_{m_2}^2 = \|U\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)U^H\|_{m_2}^2 =$

$\| \mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n) \|_{m_2}^2 = \sum_{i=1}^{n}\lambda_i^2$.

3. (×) $\boldsymbol{AA}^-$ 为幂等矩阵$\Rightarrow \boldsymbol{AA}^-$ 的特征值为 0 或 1. 又 $\boldsymbol{A}\neq\boldsymbol{0}\Rightarrow \mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{AA}^-)\geqslant 1$ $\Rightarrow \boldsymbol{AA}^-\neq\boldsymbol{0}\Rightarrow 1$ 是 $\boldsymbol{AA}^-$ 的特征值$\Rightarrow \|\boldsymbol{AA}^-\|_2=\sqrt{\max\ (\lambda_i((\boldsymbol{AA}^-)^{\mathrm{H}}\boldsymbol{AA}^-))}=\max \lambda_i(\boldsymbol{AA}^-)=1$.

4. (×) $\boldsymbol{H}=\boldsymbol{E}-2\boldsymbol{uu}^{\mathrm{H}}\Rightarrow \boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}=\boldsymbol{E}\Rightarrow \boldsymbol{H}$ 为 n 阶酉矩阵$\Rightarrow \|\boldsymbol{H}\|_a=1$.

5. (√) 因为 $\|\boldsymbol{A}\|_\infty<1$，故结论成立.

6. (√) $\boldsymbol{A}\in\mathbf{C}^{m\times m}(m>1)$严格对角占优$\Rightarrow \lambda\in\bigcup_{i=1}^{m}\{z\in\mathbf{C}^m: |z-a_{ii}|\leqslant C_i<|a_{ii}|\}\Rightarrow$ $r(\boldsymbol{A})<2\max_{1\leqslant i\leqslant m}|a_{ii}|=2\|\boldsymbol{A}\|_{m_\infty}$，故结论成立.

7. (√) 因为 $\|\boldsymbol{x}\|_2^2=|x_1|^2+|x_2|^2+\cdots+|x_2|^2\leqslant \|\boldsymbol{x}\|_1^2$，$\|\boldsymbol{x}\|_1=\sum_{i=1}^{n}|x_1|=$ $\sum_{i=1}^{n}|x_1|\cdot 1\leqslant\sqrt{n}\left(\sum_{i=1}^{n}|x_1|^2\right)^{1/2}=\sqrt{n}\|\boldsymbol{x}\|_2$

8. (×) $\boldsymbol{A}^+=\dfrac{1}{8}\begin{pmatrix}3 & 3 & -1 & -1\\ -1 & 3 & 3 & -1\\ -1 & -1 & 3 & 3\\ 3 & -1 & -1 & 3\end{pmatrix}\Rightarrow \|\boldsymbol{A}^+\|_{m_1}=4$，故结论不成立.

9. (×) $\boldsymbol{G}$ 为矩阵 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$ 的广义逆，$\boldsymbol{A}=\boldsymbol{BD}$ 为 $\boldsymbol{A}$ 的最大秩分解$\Rightarrow \boldsymbol{BDGBD}=\boldsymbol{BD}\Rightarrow$ $\boldsymbol{DGB}=\boldsymbol{E}_r\Rightarrow \mathrm{rank}(\boldsymbol{DGB})=r$.

10. (√) 取 $\boldsymbol{D}=\mathrm{diag}(1,1,0.1)$，则 $\boldsymbol{D}^{-1}\boldsymbol{AD}$ 的 3 个盖尔圆盘为 $|z-0.9|\leqslant 0.022$，$|z-0.8|\leqslant 0.023$，$|z-0.4|\leqslant 0.3$，3 个盖尔圆盘彼此孤立，故 $\boldsymbol{A}$ 的特征值均为实数.

二、(1) 非负性. $\boldsymbol{A}=\boldsymbol{0}$ 时，$a_{ij}=0$，$\max_{i,j}|a_{ij}|=0$，从而 $\|\boldsymbol{A}\|=0$；$\boldsymbol{A}\neq\boldsymbol{0}$ 时，存在 $a_{i_0j_0}\neq 0$，$\|\boldsymbol{A}\|=\sqrt{mn}\max_{i,j}|a_{ij}|\geqslant\sqrt{mn}|a_{i_0j_0}|>0$.

容易验证齐次性和三角不等式成立.

乘法相容性. 设 $\boldsymbol{B}\in\mathbf{C}^{n\times p}$，则有

$$
\begin{aligned}
\|\boldsymbol{AB}\| &= \sqrt{mp}\max_{i,j}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right|\leqslant\sqrt{mp}\max_{i,j}\left(\sum_{k=1}^{n}|a_{ik}||b_{kj}|\right)\\
&\leqslant\sqrt{mp}\cdot n\cdot\max_{i,j}|a_{ij}|\cdot\max_{i,j}|b_{ij}|=\|\boldsymbol{A}\|\cdot\|\boldsymbol{B}\|
\end{aligned}
$$

因此，$\|\boldsymbol{A}\|$ 是矩阵范数.

设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，则有

$$
\begin{aligned}
\|\boldsymbol{Ax}\|_2^2 &= \sum_{i=1}^{m}\left|\sum_{j=1}^{n}a_{ij}x_j\right|^2\leqslant\sum_{i=1}^{m}\left(\sum_{j=1}^{n}|a_{ij}|\cdot|x_j|\right)^2\\
&\leqslant\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)\left(\sum_{j=1}^{n}|x_j|^2\right)\leqslant\|\boldsymbol{A}\|^2\cdot\|\boldsymbol{x}\|_2^2
\end{aligned}
$$

三、设 $r(\boldsymbol{A})$ 为矩阵 $\boldsymbol{A}$ 的谱半径，由盖尔圆盘定理知

$$r(\boldsymbol{A}) \leqslant \min\left\{\max_i \sum_{j=1}^n |a_{ij}|, \max_j \sum_{i=1}^n |a_{ij}|\right\}.$$

显然矩阵 $\boldsymbol{A}$ 的所有行和均为 1，所以 $r(\boldsymbol{A}) \leqslant \|\boldsymbol{A}\|_\infty$，并且 1 为矩阵 $\boldsymbol{A}$ 的特征值，也就是说 $r(\boldsymbol{A}) = \|\boldsymbol{A}\|_\infty = 1$.

四. $\boldsymbol{A}$ 为 n 阶正规矩阵 $\Rightarrow \boldsymbol{A} = \boldsymbol{U}^{\mathrm{H}}\boldsymbol{\Lambda}\boldsymbol{U} \Rightarrow \boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \Rightarrow \boldsymbol{U}^{\mathrm{H}}\boldsymbol{\Lambda}\boldsymbol{U}\boldsymbol{x} = \lambda\boldsymbol{x} \Rightarrow \boldsymbol{\Lambda}\boldsymbol{U}\boldsymbol{x} = \lambda\boldsymbol{U}\boldsymbol{x}$，$\boldsymbol{x}' = \boldsymbol{U}\boldsymbol{x}$ $\quad \boldsymbol{\Lambda}\boldsymbol{x}' = \lambda\boldsymbol{x}'$，设 $\boldsymbol{x}' = (x_1', x_2', \cdots, x_n')^{\mathrm{T}} \Rightarrow \lambda_i \neq \lambda$ 时，$x_i' = 0$，$\boldsymbol{A}\boldsymbol{y} = \lambda\boldsymbol{y} \Rightarrow \boldsymbol{U}^{\mathrm{H}}\boldsymbol{\Lambda}\boldsymbol{U}\boldsymbol{y} = \mu\boldsymbol{y} \Rightarrow \boldsymbol{\Lambda}\boldsymbol{U}\boldsymbol{y} = \mu\boldsymbol{U}\boldsymbol{y} \Rightarrow \boldsymbol{y}' = \boldsymbol{U}\boldsymbol{y}$ $\quad \boldsymbol{\Lambda}\boldsymbol{y}' = \lambda\boldsymbol{y}'$，设 $\boldsymbol{y}' = (y_1', y_2', \cdots, y_n')^{\mathrm{T}} \Rightarrow \mu_i \neq \mu$ 时，$y_i' = 0$，$\lambda \neq \mu \Rightarrow (\boldsymbol{x}', \boldsymbol{y}') = 0$，$0 = (\boldsymbol{x}', \boldsymbol{y}') = (\boldsymbol{U}\boldsymbol{x})^{\mathrm{H}}\boldsymbol{U}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{y} = (\boldsymbol{x}, \boldsymbol{y})$.

五、非负性. $\boldsymbol{A} = \boldsymbol{0}$ 时，$\|\boldsymbol{A}\| = \|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^+\|_2 = \|\boldsymbol{0}\|_2 = 0$；$\boldsymbol{A} \neq \boldsymbol{0}$ 时，由于 $\boldsymbol{D} \in \mathbf{C}^{m\times n}$ 为列满秩矩阵，所以 $\|\boldsymbol{A}\| = \|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^+\|_2 > 0$.

齐次性和三角不等式很容易验证成立.

相容性. 设 $\boldsymbol{A}, \boldsymbol{B} \in \mathbf{C}^{n\times n}$，则有

$$\begin{aligned} \|\boldsymbol{A}\boldsymbol{B}\| &= \|\boldsymbol{D}\boldsymbol{A}\boldsymbol{B}\boldsymbol{D}^+\|_2 = \|(\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^+)(\boldsymbol{D}\boldsymbol{B}\boldsymbol{D}^+)\|_2 \\ &\leqslant \|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^+\|_2 \|\boldsymbol{D}\boldsymbol{B}\boldsymbol{D}^+\|_2 = \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\|. \end{aligned}$$

因此，$\|\boldsymbol{A}\| = \|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^+\|_2$ 为 $\mathbf{C}^{n\times n}$ 上的矩阵范数.

六、(1) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$，

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} = \boldsymbol{B}\boldsymbol{D}.$$

(2) $\boldsymbol{B}^+ = (\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}} = \begin{pmatrix} 6 & 5 \\ 5 & 6 \end{pmatrix}^{-1} \boldsymbol{B}^{\mathrm{T}} = \dfrac{1}{11}\begin{pmatrix} -4 & 1 & 7 \\ 7 & 1 & -4 \end{pmatrix}$，

$$\boldsymbol{D}^+ = \boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{D}^{\mathrm{T}})^{-1} = \boldsymbol{D}^{\mathrm{T}}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\boldsymbol{A}^+ = \boldsymbol{D}^+\boldsymbol{B}^+ = \frac{1}{22}\begin{pmatrix} -4 & 1 & 7 \\ 7 & 1 & -4 \\ 7 & 1 & -4 \\ -4 & 1 & 7 \end{pmatrix}.$$

(3) $\boldsymbol{x}_0 = \boldsymbol{A}^+\boldsymbol{b} = \dfrac{2}{11}(1,1,1,1)^{\mathrm{T}}$.

七、(1) 容易算得矩阵 $\boldsymbol{E}-\boldsymbol{D}^{-1}\boldsymbol{A}$ 的对角元素为 0，由于 $\boldsymbol{A}$ 为严格对角占优矩阵，则 $r(\boldsymbol{E}-\boldsymbol{D}^{-1}\boldsymbol{A})\leqslant\|\boldsymbol{E}-\boldsymbol{D}^{-1}\boldsymbol{A}\|_\infty<1$，得证.

(2) 由 Schur 三角化定理知，存在酉矩阵 $\boldsymbol{U}$ 和上三角矩阵$\boldsymbol{\Delta}$,使得 $\boldsymbol{A}=\boldsymbol{U}^*\boldsymbol{\Delta U}$，令 $\boldsymbol{D}_t=\mathrm{diag}(t,t^2,\cdots,t^n)$，则

$$\boldsymbol{D}_t\boldsymbol{\Delta D}_t^{-1}=\begin{pmatrix}\lambda_1 & t^{-1}d_{12} & t^{-2}d_{13} & \cdots & t^{-n+1}d_{1n}\\ 0 & \lambda_2 & t^{-1}d_{23} & \cdots & t^{-n+2}d_{2n}\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & t^{-1}d_{n-1,n}\\ 0 & 0 & 0 & \cdots & \lambda_n\end{pmatrix},$$

因此，对足够大的 $t>0$,$\boldsymbol{D}_t\boldsymbol{\Delta D}_t^{-1}$ 的所有非对角元的绝对值之和小于 ε. 特别是，对足够大的 t,有 $\|\boldsymbol{D}_t\boldsymbol{\Delta D}_t^{-1}\|_1\leqslant r(\boldsymbol{A})+\varepsilon$，这样，如果定义矩阵范数为

$$\|\boldsymbol{B}\| = \|\boldsymbol{D}_t\boldsymbol{U}^*\boldsymbol{BUD}_t^{-1}\|_1= \|(\boldsymbol{UD}_t^{-1})\boldsymbol{B}(\boldsymbol{UD}_t^{-1})\|_1,$$

其中 $\boldsymbol{B}\in\mathbf{C}^{n\times n}$为任意矩阵,再选择足够大的 t,则可以构造出合适的矩阵 $\boldsymbol{A}$,满足 $\|\boldsymbol{A}\|\leqslant r(\boldsymbol{A})+\varepsilon$.

复习题四

一、判断题

1. (√) $\boldsymbol{B}=\boldsymbol{A}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^+\boldsymbol{A}^{\mathrm{H}}\Rightarrow\boldsymbol{B}^{\mathrm{H}}=\boldsymbol{B}\Rightarrow\|\boldsymbol{B}\|_2=\rho(\boldsymbol{B})$;$\boldsymbol{B}^2=\boldsymbol{B}\Rightarrow\boldsymbol{B}$ 的特征值为 0 或 1，则 $\boldsymbol{A}\neq\boldsymbol{0}\Rightarrow\rho(\boldsymbol{B})=1$.

2. (×).

3. (×) $\boldsymbol{E}=(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}-\boldsymbol{A})^{-1}=(\boldsymbol{E}-\boldsymbol{A})^{-1}-\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})^{-1}\Rightarrow(\boldsymbol{E}-\boldsymbol{A})^{-1}=\boldsymbol{E}+\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})^{-1}\Rightarrow \|(\boldsymbol{E}-\boldsymbol{A})^{-1}\|=\|\boldsymbol{E}+\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})^{-1}\|\leqslant\|\boldsymbol{E}\|+\|\boldsymbol{A}\|\ \|(\boldsymbol{E}-\boldsymbol{A})^{-1}\|\Rightarrow\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\|\leqslant \dfrac{\|\boldsymbol{E}\|}{1-\|\boldsymbol{A}\|}=\dfrac{1}{1-\|\boldsymbol{A}\|}$.

4. (√) 因为 $\boldsymbol{A}^{\mathrm{H}}=(\boldsymbol{E}-2\boldsymbol{uu}^{\mathrm{H}})^{\mathrm{H}}=\boldsymbol{E}-(2\boldsymbol{uu}^{\mathrm{H}})^{\mathrm{H}}=\boldsymbol{E}-2\boldsymbol{uu}^{\mathrm{H}}=\boldsymbol{A}$，$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=(\boldsymbol{E}-2\boldsymbol{uu}^{\mathrm{H}})(\boldsymbol{E}-2\boldsymbol{uu}^{\mathrm{H}})=\boldsymbol{E}-2\boldsymbol{uu}^{\mathrm{H}}-2\boldsymbol{uu}^{\mathrm{H}}+4\boldsymbol{uu}^{\mathrm{H}}\boldsymbol{uu}^{\mathrm{H}}=\boldsymbol{E}$,

所以 $\|\boldsymbol{A}\|_{m_2}=\sqrt{\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})}=\sqrt{n}$.

5. (×) 6. (√). 7. (×). 8. (√). 9. (×). 10. (×)

二、因为矩阵 $\boldsymbol{U}$ 是酉矩阵，所有$|\lambda|\leqslant\|\boldsymbol{UA}\|_2=\|\boldsymbol{A}\|_2=\max\limits_i\{|a_i|\}$；令关于特征值 λ 的特征向量为 $\boldsymbol{x}$，即$\boldsymbol{UAx}=\lambda\boldsymbol{x}$，则

$$|\lambda|^2\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=(\boldsymbol{UAx},\boldsymbol{UAx})=(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{Ax},\boldsymbol{x})\geqslant\min_i\{|a_i|^2\}(\boldsymbol{x},\boldsymbol{x})=\min_i\{|a_i|^2\}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x},$$

即$\min\limits_i\{|a_i|\}\leqslant|\lambda|$.

三、首先证明 $\|\boldsymbol{A}\|_b=\|\boldsymbol{BAC}\|_a$是一个矩阵范数.

正定性　对任意 $A\neq 0$，则 $BAC\neq 0$，即 $\| BAC\|_a>0$，且 $\| BAC\|_a=0$ 当且仅当 $A=0$.

齐次性　$\|\lambda A\|_b=\|B(\lambda A)C\|_a=|\lambda|\,\|BAC\|_a=|\lambda|\,\|A\|_b$.

三角不等式　$\|A_1+A_2\|_b=\|B(A_1+A_2)C\|_{a}\leqslant\|BA_1C\|_a+\|BA_2C\|_a$

$$=\|A_1\|_b+\|A_2\|_b.$$

下面再证明相容性.

$$\begin{aligned}\|A_1A_2\|_b&=\|B(A_1A_2)C\|_a=\|(BA_1C)C^{-1}B^{-1}(BA_2C)\|_a\\&\leqslant\|BA_1C\|_a\,\|C^{-1}B^{-1}\|_a\,\|BA_2C\|_a\\&\leqslant\|BA_1C\|_a\,\|C^{-1}\|_a\,\|B^{-1}\|_a\,\|BA_2C\|_a\\&\leqslant\|BA_1C\|_a\,\|BA_2C\|_a\\&=\|A_1\|_b\,\|A_2\|_b.\end{aligned}$$

四、(1) $A=\begin{pmatrix}0&1\\-1&0\\1&-1\\0&1\end{pmatrix}\begin{pmatrix}1&0&-1\\0&1&-1\end{pmatrix}=BD$；

(2) $B^+=(B^{\mathrm T}B)^{-1}B^{\mathrm T}=\dfrac{1}{5}\begin{pmatrix}1&-3&2&1\\2&-1&-1&2\end{pmatrix}$，

$D^+=D^{\mathrm T}(DD^{\mathrm T})^{-1}=\dfrac{1}{3}\begin{pmatrix}2&-1\\-1&2\\-1&-1\end{pmatrix}$，

$A^+=D^+B^+=\dfrac{1}{15}\begin{pmatrix}0&-5&5&0\\3&1&-4&3\\-3&4&-1&-3\end{pmatrix}$，

(3) $AA^+b=b=\begin{pmatrix}1\\-1\\0\\1\end{pmatrix}$，方程组 $Ax=b$ 有解；

(4) 最小范数解 $x_0=A^+b=\dfrac{1}{3}\begin{pmatrix}1\\1\\-2\end{pmatrix}$.

五、A 的 5 个盖尔圆盘为 $G_i=\left\{z\,\|\,z-2i|\leqslant 1-\dfrac{1}{(i+1)^4}\right\}(i=1,2,3,4,5)$，它们都是孤立的，从而矩阵有 5 个互异特征值，所以矩阵能够相似于对角矩阵. 再由 G_i 关于 x 轴对称且都在 y 坐标轴右边，以及实矩阵的复数特征值成对共扼出现的性质知，G_i 中的特征值必为正实数，所以 A 的特征值都是正实数.

六、(1) $\sigma_1(\boldsymbol{AB})=\|\boldsymbol{AB}\|_2\leqslant\|\boldsymbol{A}\|_2\cdot\|\boldsymbol{B}\|_2\leqslant\sigma_1(\boldsymbol{A})\cdot\sigma_1(\boldsymbol{B})$.

(2) $\sigma_1(\boldsymbol{A}+\boldsymbol{B})=\|\boldsymbol{A}+\boldsymbol{B}\|_2\leqslant\|\boldsymbol{A}\|_2+\|\boldsymbol{B}\|_2\leqslant\sigma_1(\boldsymbol{A})+\sigma_1(\boldsymbol{B})$.

七、(1) 设 $\boldsymbol{A}$ 关于 λ 的特征向量为 $\boldsymbol{x}$，即 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$. 因为 $\boldsymbol{A}$ 可逆，则 $\frac{1}{\lambda}\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{x}$，进一步有 $\left|\frac{1}{\lambda}\right|\|\boldsymbol{x}\|=\|\boldsymbol{A}^{-1}\boldsymbol{x}\|\leqslant\|\boldsymbol{A}^{-1}\|\ \|\boldsymbol{x}\|$，即 $\frac{1}{|\lambda|}\leqslant\|\boldsymbol{A}^{-1}\|$，因此，$|\lambda|\geqslant\frac{1}{\|\boldsymbol{A}^{-1}\|}$.

(2) 由于用一非零数乘以矩阵的一行的所有元素不改变矩阵的秩，因而可以假设矩阵 $\boldsymbol{A}$ 的主对角元素 $a_{ii}\geqslant 0$，且所有的 $R_i=1$ 或者为 0，则矩阵 $\boldsymbol{A}$ 的所有特征值都在单位圆内，且可以证明 $\operatorname{rank}(\boldsymbol{A})\geqslant\sum_{i=1}^{n}a_{ii}\left(\left(\sum_{i=1}^{n}a_{ii}\right)=\operatorname{tr}(\boldsymbol{A})=\sum_{i=1}^{n}\lambda_i\leqslant\sum_{i=1}^{n}|\lambda_i|\leqslant\boldsymbol{A}\right.$的非零特征值的个数$\leqslant\operatorname{rank}(\boldsymbol{A})$).

复习题五

一、判断题

1. (√). 2. (×). 3. (×). 4. (√). 5. (√). 6. (√). 7. (√).
8. (√). 9. (×). 10. (×).

二、因为 $\boldsymbol{A}$ 是 n 阶实对称矩阵，所以是正规矩阵，进一步由 Schur 定理得 $\sum_{i=1}^{n}|\lambda_i|^2=\|\boldsymbol{A}\|_F^2$，即有 $\|\boldsymbol{A}\|_F=\sqrt{\sum_{i=1}^{n}|\lambda_i|^2}=\sqrt{\frac{n(n+1)}{2}}$；$r(\boldsymbol{A})=\sqrt{n}$； $\|\boldsymbol{A}\|_2=\sqrt{r(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})}=\sqrt{n}$；$\operatorname{cond}(\boldsymbol{A})_2=\frac{\sqrt{n}}{1}=\sqrt{n}$；$\|(\boldsymbol{A}^{-1})^m\|_2=1$.

三、由 $\boldsymbol{A}$ 的 n 个盖尔圆盘为 $G_i=\{z\,\|\,z-a_{ii}|\leqslant\sum_{j\neq i}|a_{ij}|\}(i=1,2,\cdots,n)$，故有

$$|\lambda|\leqslant\sum_{j=1}^{n}|a_{ij}|<1,i=1,2,\cdots,n.$$

四、(1) $\boldsymbol{A}=\begin{pmatrix}1&0\\2&1\\2&0\\4&2\end{pmatrix}\begin{pmatrix}1&0&1&1\\0&1&0&-1\end{pmatrix}=\boldsymbol{BD}$；

(2) $\boldsymbol{B}^+=(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}=\frac{1}{5}\begin{pmatrix}1&0&2&0\\-2&1&-4&2\end{pmatrix}$，$\boldsymbol{D}^+=\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{DD}^{\mathrm{T}})^{-1}=\frac{1}{5}\begin{pmatrix}2&1\\1&3\\2&1\\1&-2\end{pmatrix}$，

$$\boldsymbol{A}^+=\boldsymbol{D}^+\boldsymbol{B}^+=\frac{1}{25}\begin{pmatrix}0&1&0&2\\-5&3&-10&6\\0&1&0&2\\5&-2&10&-4\end{pmatrix};$$

(3) $\boldsymbol{A}^+\boldsymbol{b}=(1\quad 4\quad 1\quad 3)^{\mathrm{T}}$，$\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{b}=\boldsymbol{b}$,故 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有解；

(4) 极小范数解 $\boldsymbol{x}_0=\boldsymbol{A}^+\boldsymbol{b}=(1\quad 4\quad 1\quad 3)^{\mathrm{T}}$.

五、若 $\boldsymbol{A}$ 为正交矩阵，则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$，所以有$\frac{1}{2}+a^2=1$，$\frac{b}{\sqrt{2}}+ac=0$，$b^2+c^2=1$，解得 $a=b=-c=\frac{1}{\sqrt{2}}$，或$a=-b=c=\frac{1}{\sqrt{2}}$，或 $a=b=c=-\frac{1}{\sqrt{2}}$，或$-a=b=c=\frac{1}{\sqrt{2}}$.

六、(1) 令 $\boldsymbol{x}\in\mathbf{R}^n$ 的列向量，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x})\geqslant 0$，故 $\boldsymbol{B}$ 是半正定矩阵；

(2) $\boldsymbol{x}=\boldsymbol{0}$ 时，$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}=(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x})=0$；当 $\boldsymbol{x}\neq\boldsymbol{0}$ 时，由 $\boldsymbol{A}$ 的列向量组线性无关，得 $\boldsymbol{A}\boldsymbol{x}\neq\boldsymbol{0}$，也就是 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}=(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x})>0$，因此，$\boldsymbol{B}$ 为正定矩阵.

七、直接写出矩阵 $\boldsymbol{A}$ 的 M-P 广义逆的 4 个条件，再根据矩阵广义逆的唯一性，即可得证.

复习题六

一、判断题

1. (√).　2. (√).　3. (×).　4. (√).　5. (×).　6. (×).　7. (√).
8. (×).　9. (√).　10. (×)

二、(1) 因为 $\det(\boldsymbol{AB})=\det(\boldsymbol{A})\det(\boldsymbol{B})=-1$，所以

$$\begin{aligned}\det(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B})&=\det(\boldsymbol{A}^{\mathrm{T}})\det(\boldsymbol{B})=\det(\boldsymbol{A})\det(\boldsymbol{B})=\det(\boldsymbol{A})\det(\boldsymbol{B}^{\mathrm{T}})\\&=\det(\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}})=-1;\end{aligned}$$

(2) 因 $\boldsymbol{A},\boldsymbol{B}$ 是两个 n 阶正交矩阵，所以

$$\begin{aligned}\det(\boldsymbol{A}+\boldsymbol{B})&=\det(\boldsymbol{A})\det(\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}})\det(\boldsymbol{B})=\det(\boldsymbol{AB})\det(\boldsymbol{A}+\boldsymbol{B})\\&=-\det(\boldsymbol{A}+\boldsymbol{B}),\end{aligned}$$

即 $\det(\boldsymbol{A}+\boldsymbol{B})=0$.

三、设按行得到的 4 个 Gerschgorin 圆分别为 G_1,G_2,G_3,G_4，容易算得 G_1 是孤立的，有一个实特征值，连通部分 $G_2\cup G_3\cup G_4$ 至少有一个实特征值，因此，命题得证.

四、(1) $\boldsymbol{A}=\begin{pmatrix}1&0\\0&2\\-1&-1\end{pmatrix}\begin{pmatrix}1&-1&0&-1&1\\0&0&1&-1&1\end{pmatrix}=\boldsymbol{BD}$；

(2) $\boldsymbol{B}^+=(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}=\frac{1}{9}\begin{pmatrix}5&-2&-4\\-1&4&-1\end{pmatrix}$，

$$\boldsymbol{D}^+=\boldsymbol{D}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{D}^{\mathrm{T}})^{-1}=\frac{1}{8}\begin{pmatrix}3&-2\\-3&2\\-2&4\\-1&-2\\1&2\end{pmatrix},$$

$$A^+=D^+B^+=\frac{1}{72}\begin{pmatrix}17&-14&-10\\-17&14&10\\-14&20&4\\-3&-6&6\\3&6&-6\end{pmatrix};$$

(3) $A^+b=(1,-1,-1,0,0)^T$，$AA^+b=b$,故 $Ax=b$ 有解；

(4) 极小范数解 $x_0=A^+b=(1,-1,-1,0,0)^T$.

五、必要性. $\|\alpha+\beta\|_2^2=(\|\alpha\|_2+\|\beta\|_2)^2\Rightarrow\alpha^T\beta=\|\alpha\|_2\|\beta\|_2\geqslant 0$；当$\beta\neq 0$ 时，取 $k=\frac{\|\alpha\|_2}{\|\beta\|_2}$，则有 $\|\alpha-k\beta\|_2^2=0$，$\alpha-k\beta=0$，即$\alpha=k\beta$；$\beta=0$ 显然成立.

充分性. $\alpha=0$ 显然成立；当$\alpha\neq 0$ 时，由$\alpha=k\beta$，$\alpha^T\beta\geqslant 0$，且有

$$\|\alpha+\beta\|_2=(1+k)\|\alpha\|_2=\|\alpha\|_2+\|k\alpha\|_2=\|\alpha\|_2+\|\beta\|_2.$$

六、充分性. 显然成立；

必要性. 设 $A=BD$ 是 A 的最大秩分解，则有 $A^+=D^H(DD^H)^{-1}(B^HB)^{-1}B^H$，由 $A^+AM=A^+AN$，可得

$$D^H(DD^H)^{-1}(B^HB)^{-1}B^HAM=D^H(DD^H)^{-1}(B^HB)^{-1}B^HAN$$

$\Rightarrow(B^HB)^{-1}B^HAM=(B^HB)^{-1}B^HAN\Rightarrow B^HAM=B^HAN\Rightarrow B^H(AM-AN)=0$，因为 B 是列满秩矩阵，也就是说 B^H 行满秩，所以 $AM-AN=0$，即 $AM=AN$.

七、(1) 若 A 为收敛矩阵，则 $r(A)<1$，计算得 $r(A)=2|c|$，所以当$|c|<\frac{1}{2}$时，A 为收敛矩阵.

(2) $A+B=A(E+A^{-1}B)$，由于 $\|A^{-1}B\|\leqslant\|A^{-1}\|\ \|B\|<1$，所以 $\|-A^{-1}B\|<1$. 令 λ 为 $A^{-1}B$ 任意一个特征值，则$|\lambda|<1$，也就是说 $0<1+\lambda$ 为 $E+A^{-1}B$ 的特征值，即$E+A^{-1}B$ 可逆，从而 $A+B$ 可逆.