矩阵论入门

武同锁　喻厚义　编著

科学出版社

矩阵论入门

武同锁　喻厚义　编著

科学出版社

北京

内 容 简 介

本书共 5 章，第 1 章是简要的预备知识，包括线性代数(矩阵消元法、置换矩阵、Schmidt 正交化、镜面反射、分块矩阵的乘法)，以及一元多项式的互素与整除；第 2 章是矩阵的各种分解式，也是对大学阶段线性代数的复习与提升，包括正规矩阵与酉相似、矩阵分解式、Moore-Penrose 广义逆以及 Hermite 半正定矩阵的唯一幂表达定理；第 3 章是较为完整的线性变换理论，也是本书的理论核心，包括若干关于线性变换与矩阵的一一对应定理、根子空间分解定理以及 Jordan 标准形的简要现代处理、线性空间与线性映射(矩阵)的张量积与外幂；第 4 章是矩阵分析，包括向量范数及其诱导的矩阵范数、矩阵函数概要、特征值的估计 (几个圆盘定理)、非负方阵与正方阵以及三个相关的核心定理、随机矩阵. 第 4 章与第 2 章一起构成工科矩阵理论的核心内容，技巧性强且具有重要的应用背景. 第 5 章收集了有关矩阵理论应用的一些关键词，方便读者搜索应用. 本书配备部分具有一定难度的题目(标记*)，这些题目也是矩阵理论的重要内容；基于这一考量，对部分较难的题目给出了提示或解答.

本书内容的编排，遵循由浅入深原则，特别强调逻辑一致性；在重视技巧性的同时，适度强调一定的思想性.

本书可供学习过线性代数(高等代数)、高等数学 (数学分析) 的理、工、管理等专业本科生和研究生使用，也可供具有一定数学基础的数学爱好者使用.

图书在版编目(CIP)数据

矩阵论入门/武同锁，喻厚义编著. —北京：科学出版社，2020.5

ISBN 978-7-03-064936-2

I. ①矩… II. ①武… ②喻… III. ①矩阵论-研究 IV. ①O151.21

中国版本图书馆 CIP 数据核字 (2020) 第 068153 号

责任编辑：王丽平 李香叶／责任校对：邹慧卿
责任印制：吴兆东／封面设计：陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

北京虎彩文化传播有限公司印刷
科学出版社发行 各地新华书店经销
*
2020 年 5 月第 一 版 开本：720 × 1000 B5
2020 年 5 月第一次印刷 印张：10 1/2
字数：200 000

定价：78.00 元

(如有印装质量问题，我社负责调换)

前　言

矩阵理论是理工科的一门重要基础课, 一向以高技巧性而闻名, 在其他学科理论、工程技术、网络与人工智能、运筹与最优化、科学计算等诸多领域具有广泛的重要应用. 国内外已经有了大量的优秀专著或教材, 那么为什么又编撰一本同类的教材?

为了回答这个问题, 我们先从对思维与技巧 (ideas and tricks) 的关系理解谈起. 众所周知, 矩阵理论的技巧性非常强, 所以不介绍和重视技巧是不现实的. 但是若在数学的任一分支的教材中过分强调技巧, 恐怕不利于学科的发展和传播, 因为这不利于引起读者 (包括学生) 的兴趣. 我们认为应该适度加强在思维创新性上的训练和提高. 基于这个考量, 第 3 章 (线性变换部分) 在内容广度和深度上比一般教材做得更深入一些; 在技巧性高强的矩阵理论本身的编排和介绍上, 也适度地在思维创新性上有所加强, 比如讲清楚引进新概念的自然性、合理性和重要性. 再比如在 4.2 节, 通过自然的实例强调引进向量范数概念的迫切性与合理性; 在向量范数的基础上自然引进由其诱导出的矩阵范数; 并在内容选取上始终以诱导出的矩阵范数为核心, 而不是强调抽象的矩阵范数 (虽然后者更具有一般性). 结束范数的内容后, 后续内容选取以及安排, 围绕谱空间特别是谱半径的概念展开; 大部分选材都是以谱半径概念为纲, 这是整个矩阵分析最核心的部分. 在本书编写过程中, 我们在这方面做了一定的尝试和努力, 希望能够引起读者的浓厚兴趣, 以利于他们更进一步的自主钻研. 实际上, 矩阵理论框架十分宏大, 很多主题本身也正处在发展进程中, 要在 48 学时的短暂教学过程中做到面面俱到, 任何教师都面临艰难的选择. 我们自然要在课本 (和课堂教学) 中, 通过选取最核心、最基本的概念, 介绍最为典型的思想方法、技术手段与技巧, 将读者尽快地带进该领域; 若能在学习的过程中引起他们的浓厚兴趣, 能引导他们主动自主钻研和查询资料, 就是本书 (以及课程教学过程) 的最大成功. 本书在这方面做了一些尝试和努力, 这也是出版这个短篇教材的初衷.

本书的编排次序有自己的特色, 且与已有教材有较大的不同. 在内容处理上, 也有若干创新. 例如, 将矩阵分解式与 Moore-Penrose 广义逆提前到第 2 章. 对最小多项式的处理, 传统上大多使用 Jordan 标准形理论. 我们认为, 过多使用超级大定理对于训练学生的逻辑和思维能力不利. 使用一元多项式中互素的理论和分块矩阵基本技巧, 给出了最小多项式的初等处理. 对于 Jordan 标准形理论, 采用了最新的研究成果, 采取直接又初等的处理方式. 引导读者在引进矩阵乘法概念的开始, 就

循序渐进地掌握分块矩阵的乘法. 分块矩阵在矩阵理论中已经被普遍应用, 为了帮助读者更好地掌握, 在预备知识部分, 加入“分块矩阵的乘法”部分, 希望有益于读者. 对于第 4 章矩阵分析, 则重点突出范数及其应用 (正矩阵与非负矩阵). 这些内容很好地体现了分析和代数的思想与技巧的融合.

本书第 1 章和第 2 章作者曾经多次在上海交通大学暑假小学期课程“数学的思维与技巧”(共 2 学分, 与数学分析老师分享课时) 讲授过, 内容比较成熟和稳定, 在暑假小学期课程上教授的效果不错; 第 3 章是作者为致远学院 (非数学专业) 本科生讲授 80 学时线性代数的标准内容之一, 也有了十几年的经历和沉淀; 对第 4 章内容的选取, 作者作了较为广泛的调研和梳理, 这章在上海交通大学研究生荣誉课程矩阵理论上讲过两遍.

本书来自讲稿, 不追求内容的全面, 只强调逻辑性与自洽性, 也没有配备过多习题. 由于篇幅限制, 矩阵理论的应用部分收录较少, 读者当然可自行查找和阅读.

作者认为在研究生阶段, 习题的重要性与本科阶段相比有所减弱, 建议读者尝试阅读有关文献 [2,9,10,11], 并逐步培养追踪或主动查找大量专题文献的能力及习惯, 在这个过程中能很好地使用互联网搜索工具.

本书得到了上海交通大学数学科学学院、上海交通大学研究生院、西南大学数学与统计学院、科学出版社的大力支持和帮助, 作者在此一并表示衷心感谢. 本书的完成还得到了上海市自然科学基金 (项目编号: 19ZR1424100) 的资助.

欢迎读者提出宝贵的建议, 或指出书中的疏漏与不妥, 有关问题可发至电子邮件, 或至作者主页查询相关事宜. 武同锁的邮箱: tswu@sjtu.du.cn; 个人主页: http://math.sjtu.edu.cn/faculty/wuts/; 地址: 上海交通大学数学科学学院. 喻厚义的邮箱: yuhouyi@swu.edu.cn; 个人主页: http://math.swu.edu.cn/s/math/index2jiangshitea42sub1.html; 地址: 西南大学数学与统计学院.

作　者

2020 年 3 月 14 日

符 号 说 明

$=:$ 或 $\triangleq$: 定义或记号 (记为)

$\binom{n}{i}$ 或 C_n^i: 从给定的 n 个元素中任取 i 个元素的取法 (组合数)

$\deg f(x)$: 多项式 $f(x)$ 的次数

$p(A)$: 由多项式 $p(x)$ 和矩阵 A 确定的矩阵多项式

A_{ij}: a_{ij} 在 A 中的代数余子式

$A(i,j)$: 矩阵 A 的 (i,j) 位置的元素 a_{ij}

$A(i)$: 矩阵 A 的第 i 列 (列向量)

$(i)A$: A 的第 i 行 (行向量)

$\det(A)$: 方阵 A 的行列式

$|c|$: 复数 $c=a+bi$ 的模长 $\sqrt{a^2+b^2}$

$|A|$: $(|a_{ij}|)_{n\times n}$ (不是行列式)

$A^\dagger$: 矩阵 A 的 Moore-Penrose 广义逆

A^D: 方阵 A 的 Drazin 广义逆

A^{T}: 矩阵 A 的转置矩阵

$\overline{A}$: 矩阵 $A=(a_{ij})_{n\times n}$ 的共轭矩阵 $\overline{A}=:(\overline{a_{ij}})_{n\times n}$

$A^\star \overset{\text{or}}{=} A^{\mathrm{H}}$: A 的共轭转置矩阵 $\overline{A}^{\mathrm{T}}$

A^*: A 的伴随矩阵, 其第一行为 $A_{11},A_{21},\cdots,A_{n1}$

$\sigma^\star$: 内积空间 V 上线性变换 σ 的共轭线性变换, 满足

$$[\sigma(\alpha),\beta]=[\alpha,\sigma^\star(\beta)],\quad \forall\alpha,\beta\in V$$

$\mathrm{diag}\{a_1,a_2,\cdots,a_n\}$: 对角矩阵

ε_i : $(i,1)$ 位置的元素是 1, 其余位置的元素为 0 的列向量

E : 单位矩阵 $\mathrm{diag}\{1,1,\cdots,1\}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$

e : 列向量 $(1,1,\cdots,1)^{\mathrm{T}}$

J : 任意位置的元素都是 1 的 n 阶方阵, 即 ee^{T}

$\mathrm{diag}\{A_1,A_2,\cdots,A_n\}$: 分块对角矩阵

$\mathrm{GL}_n(\mathbb{F})$: 数域 $\mathbb{F}$ 上的所有 n 阶可逆矩阵全体

$M_n(\mathbb{F})$ 或 $\mathbb{F}^{n\times n}$: 数域 $\mathbb{F}$ 上的所有 n 阶矩阵全体

$O_n(\mathbb{F})$: 数域 $\mathbb{F}$ 上的所有 n 阶正交矩阵全体

$U_n(\mathbb{F})$: 数域 $\mathbb{F}$ 上的所有 n 阶酉阵全体

$\mathbb{F}^{n\times 1}$ 或 $\mathbb{F}^n$: 数域 $\mathbb{F}$ 上的 n 维列向量全体

$\alpha_1,\cdots,\alpha_u \stackrel{l}{\Longrightarrow} \beta_1,\cdots,\beta_v$: 每个 β_i 都可由 $\alpha_1,\cdots,\alpha_u$ 线性表出

V_A^λ 或 V_λ^A: 方阵 A 的属于特征值 λ 的特征子空间,

$$V_\lambda^A =: \{\alpha \mid A\alpha = \lambda\alpha\}(\text{简记为 } V_\lambda)$$

$\dim({}_{\mathbb{F}}V)$: 数域 $\mathbb{F}$ 上的线性空间 V 的维数

$r(A)$: 矩阵 A 的秩

$\mathrm{tr}\,(A)$: 方阵 A 的迹 $\sum\limits_{i=1}^{n} a_{ii} = \sum\limits_{i=1}^{n} \lambda_i$

$\rho(B)$: 方阵 B 的谱半径, 即 $\max\limits_{1\leqslant i\leqslant n} |\lambda_i|$ (特征根的模长最大值)

$A \xrightarrow{r} \cdots \xrightarrow{r} B$: 对 A 进行一系列的初等行变换, 化为矩阵 B

$A \xrightarrow[c]{} \cdots \xrightarrow[c]{} B$: 对 A 进行一系列的初等列变换, 化为矩阵 B

$\mathrm{im}(\sigma)$ 或 $\sigma(V)$: 线性空间 V 上线性变换 σ 的像子空间

$\ker(\sigma)$: 线性空间 V 上线性变换 σ 的核子空间

$$\ker(\sigma) =: \{\alpha \in V \mid \sigma(\alpha) = 0_V\}$$

$r(\sigma)$: 线性变换 σ 的秩, 即 $\dim_{\mathbb{F}} \mathrm{im}(\sigma)$

$\mathrm{Ndeg}(\sigma)$: 线性变换 σ 的零化度 (null degree), 即 $\dim_{\mathbb{F}} \ker(\sigma)$

$r\{\alpha_1,\alpha_2,\cdots,\alpha_t\}$: 向量组 $\alpha_1,\alpha_2,\cdots,\alpha_t$ 的秩

$\beta\perp W$: 内积空间中的向量 β 与子空间 W 中所有向量正交

$[\alpha,\beta]$: 通常指 $\alpha^\star\beta$

$U\oplus V$: 线性空间 U 与 V 的直和

$\mathrm{End}_{\mathbb{F}}(V)$: V 上线性变换全体

$\|\alpha\|$: 在第 4 章表示向量 α 的向量范数; 在其余章表示 α 的长度 $\sqrt{[\alpha,\alpha]} =: \sqrt{\alpha^\star\alpha}$

$\|A\|$: 矩阵 A 的范数

$\Gamma(A)$: 方阵 A 的有向图 G, 其中 $V(G) = \{P_i \mid 1\leqslant i\leqslant n\}$, 而

$$\{P_i, P_j\} \in E(G) \Longleftrightarrow a_{ij} \neq 0 \ (1\leqslant j\leqslant n)$$

目录

第1章 预备知识

1.1 线性代数

1. 矩阵消元法 (又叫 Gauss 消元法) 与三种初等矩阵

用 E_{ij} 表示 (i,j) 位置的元素是 1 而其余位置的元素均为 0 的方阵.

第 I 型初等矩阵 $P(k(i))$, 是将非零数 k 乘以 n 阶单位矩阵 E 的第 i 行而得到的方阵. 注意 $P(k(i))^{-1} = P\left(\frac{1}{k}(i)\right)$.

第 II 型初等矩阵 P_{ij}, 是交换 E 的 i,j 两行的位置得到的; 第 II 型初等矩阵又叫**初等置换矩阵**. 注意 $P_{ij}^2 = E$, 即 $P_{ij}^{-1} = P_{ij}$.

第 III 型初等矩阵形为 $P((i)+l(j)) =: E + l\cdot E_{ij}$ $(i \neq j)$. 注意 $(E+lE_{ij})^{-1} = E + (-l)\cdot E_{ij}$.

此外, 对一个矩阵 $B_{m\times n}$ **左乘**一个 m 阶初等矩阵, 相当于对 B 作一次相应的**初等行变换**; 对 B 作一次初等列变换, 结果等于用相应的 n 阶初等矩阵右乘于 B. 例如, 对于一个 3×4 的矩阵 A, 左乘以初等矩阵 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$, 相当于将 A 的第1行的元素各乘以 -3 加到了第 3 行的对应位置元素; 而右乘以矩阵 $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, 相当于将 A 的第3列的元素各乘以 -3 加到了第 1 列的对应位置元素.

矩阵 (行) 消元法, 指的是对一个矩阵进行多次的初等行变换, 化为 (行) 阶梯形或者标准阶梯形. 矩阵消元法是进行矩阵运算的灵魂, 这种方法在我国古代数学巨著《九章算术》中又被称为中国消元法 (或又叫孙子消元法), 而在西方教材中则通常被叫作 Gauss 消元法.

2. 正交矩阵、酉阵与置换矩阵

对于方阵 A_n, 记

$$\overline{A} =: (\overline{a_{ij}}), \quad A^{\star} \overset{\text{or}}{=} A^{\mathrm{H}} =: \overline{A}^{\mathrm{T}},$$

其中 $\overline{a}_{ij}$ 是复数 $a_{ij} =: c+di$ 的共轭数 $c-di$. 如果方阵 A 满足

$$AA^\star = E = A^\star A, \tag{1}$$

即 $A^{-1}=A^\star$, 则称 A 是一个**酉阵**. 实的酉阵叫作**正交矩阵**.

如果 $A_n=(\alpha_1,\cdots,\alpha_n)$, 则 A 是一个酉阵当且仅当

$$[\alpha_i,\alpha_j] =: \alpha_i^\star\alpha_j = \delta_{ij}, \quad \forall 1\leqslant i,j\leqslant n, \tag{2}$$

即当 $i=j$ 时, 有 $[\alpha_i,\alpha_j]=1$; 当 $i\neq j$ 时, 有 $[\alpha_i,\alpha_j]=0$.

一个**置换矩阵**是这样的一个方阵, 其每一行和每一列恰好有一个位置数是 1, 其余均为零. 直接易于验证: 置换矩阵都是正交矩阵, 即置换矩阵 P 满足 $P^{\mathrm{T}}P=E$. 当然, 这一点也可由下面的分解式得到.

命题 1.1.1 任一置换矩阵都可分解为有限个初等置换矩阵 (即第 II 型初等矩阵) 的乘积; 反之亦然.

证明 只验证第一个结论. 对于阶数 n 用数学归纳法.

当 $n=1$ 时, 结论当然成立.

现在假设置换矩阵 P 是 n 阶的. 如果 P 的第一列中的元素 1 出现在 $(1,1)$ 位置, 则由数学归纳法完成论证; 如果它出现在 $(i,1)$ 位置, 则 $P_{1i}\cdot P$ 仍为置换矩阵, 且其 $(1,1)$ 位置元素为 1. 因此 $P_{1i}P$ 形为 $\begin{pmatrix}1 & 0\\ 0 & Q\end{pmatrix}$, 其中 Q 是一个 $n-1$ 阶的置换矩阵. 根据归纳假设, 可设 $Q=Q_1\cdots Q_u$, 其中 $Q_i(i=1,2,\cdots,u)$ 都是 $n-1$ 阶初等置换矩阵. 命 $P_i=\begin{pmatrix}1 & 0\\ 0 & Q_i\end{pmatrix}$, 则有

$$P=P(1,i)\cdot P_1\cdots P_u,$$

其中 $P_1,\cdots,P_u,P(1,i)$ 均为初等置换矩阵. 由此还可看出, 每个 n 阶置换矩阵可以写成至多 n 个初等置换矩阵的乘积. □

上述命题当然提供了置换矩阵的一种基本分解式. 注意到单位矩阵 E 是置换矩阵, 且矩阵乘法满足结合律, 因此, 根据命题 1.1.1, 马上得到如下的推论.

推论 1.1.2 所有的 n 阶置换矩阵 (共 $n!$ 个) 作成一个**乘法群**, 即还有

(1) 两个 n 阶置换矩阵的乘积也是置换矩阵;

(2) 置换矩阵的逆矩阵也是置换矩阵.

类似地, 可知所有 n 阶酉阵 (正交矩阵) 全体作成一个群, 分别叫作**酉群**(**正交群**); 数域 $\mathbb{F}$ 上的 n 阶可逆矩阵全体也作成一个群, 叫作第 n 个一般线性群, 记为 $\mathrm{GL}_n(\mathbb{F})$. 除了酉群和正交群外, $\mathrm{GL}_n(\mathbb{F})$ 还包含了很多其他群作为子群. 例如, 所

有形为

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & 1 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & 1 \end{pmatrix}$$

的 n 阶方阵全体关于矩阵乘法作成一个群. 所有 n 阶可逆上三角方阵作成一个群.

一般地, 假设 G 是带有一个二元运算 $\cdot$ 的非空集合. 如果以下三个条件得到满足, 则称 G 关于运算 $\cdot$ 作成一个群:

(a) **结合律** $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$, $\forall g_i \in G$, $i = 1, 2, 3$;

(b) **存在单位元** $\exists e \in G$ s.t. $e \cdot g = g = g \cdot e$, $\forall g \in G$;

(易证: 若存在, 则唯一, 记为 1_G.)

(c) **存在可逆元** $\forall g \in G$, 存在 $h \in G$, 使得 $g \cdot h = h \cdot g = 1_G$.

(可证: h 是唯一的, 记为 $g^{-1} =: h$.)

习题与扩展内容

习题 1 写出所有的 3 阶置换矩阵, 并把非初等置换矩阵写成初等置换矩阵之乘积.

习题 2 考虑具有如下性质的 n 阶矩阵全体 G: 每一行、每一列恰好有一个位置元素是 ± 1, 其余位置为零. 求证: G 关于乘法作成一个群, 并计算 $|G|$.

习题 3 假设 $i_1, i_2, \cdots, i_n$ 是 $1, 2, \cdots, n$ 的一个排列. 求置换矩阵 P 使得

$$P^{\mathrm{T}} \cdot \operatorname{diag}\{i_1, i_2, \cdots, i_n\} \cdot P = \operatorname{diag}\{1, 2, \cdots, n\}.$$

3. Schmidt 正交化与单位化过程

对于复空间 $\mathbb{C}^n$, 其中的内积是如下定义的: $[\alpha, \beta] = \alpha^* \beta$; 对于内积空间 $\mathbb{C}^n$ 中的给定线性无关向量组 $\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_m$, 存在一个标准程序 (Schmidt 正交化过程) 去求得与 $\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_m$ 等价的正交向量组:

$$\beta_1 = \alpha_1,$$

$$\beta_2 = \alpha_2 - \frac{[\beta_1, \alpha_2]}{[\beta_1, \beta_1]} \beta_1,$$

$$\beta_3 = \alpha_3 - \frac{[\beta_2, \alpha_3]}{[\beta_2, \beta_2]} \beta_2 - \frac{[\beta_1, \alpha_3]}{[\beta_1, \beta_1]} \beta_1,$$

$$\cdots\cdots$$

$$\beta_m = \alpha_m - \sum_{i=1}^{m-1} \frac{[\beta_i, \alpha_m]}{[\beta_i, \beta_i]} \beta_i.$$

注记　与正交化过程相伴的是**单位化**(也叫**标准化**), 即将每个 β_i 进一步单位化, 得与 β_i 同方向的单位向量 $\gamma_i = \frac{1}{||\beta_i||}\beta_i$, 这里 $||\beta_i|| = \sqrt{[\beta_i, \beta_i]}$, 所得向量组 $\gamma_1, \cdots, \gamma_m$ 是标准正交组, 即满足类似推论 1.1.2 (2) 中的 m^2 个等式.

如果从一个 $n \times m$ 列满秩矩阵 $A = (\alpha_1, \alpha_2, \cdots, \alpha_m)$ 出发, 对于其列向量施行上述的正交化、单位化过程; 以 $\gamma_1, \gamma_2, \cdots, \gamma_m$ 依次作为列向量所得矩阵记为 Q, 则有

$$Q^\star Q = E_m, \quad Q = (\alpha_1, \alpha_2, \cdots, \alpha_m)B = AB, \tag{3}$$

其中 B_m 是上三角方阵, 其对角线元素均为正实数, (i, i) 的位置元素为 $\frac{1}{||\beta_i||}$, 因此有 $A = QR$, 其中 $Q^\star Q = E_m$, R 是 m 阶上三角方阵, R 的对角线元素全为正实数 (事实上, $R(i, i) = ||\beta_i||$). 特别地, 如果 A 是可逆方阵, 则 A 有分解式 $A = QR$, 其中 Q 是一个酉阵, 而 R 是主对角线元素均为正实数的上三角方阵. 这种分解式叫作**QR-分解**.

命题 1.1.3　任一列满秩矩阵 $A_{n\times m}$ 有如下的唯一分解式 (叫作矩阵 A 的 QR-分解): $A = QR$, 其中 $Q^\star Q = E_m$, 而 R_m 是对角线元素均为正实数的上三角方阵.

关于唯一性的论证将在 2.2 节给出; 作为练习, 读者也可自己尝试给出一个严格的验证.

4. **镜面反射矩阵与** Householder **变换**

以下是由美国学者 A. S. Householder 在 1958 年发现的有趣结果.

命题 1.1.4　对于任一单位向量 $\beta \in V =: \mathbb{C}^{m\times 1}$, 存在内积空间 V 上的如下酉线性变换

$$\sigma : V \to V, \quad \alpha \mapsto \alpha - 2[\beta, \alpha]\beta = (E - 2\beta\beta^\star) \cdot \alpha.$$

换言之, 对于任一单位向量 β, 可以构造 Hermite 酉阵

$$H = E - 2\beta\beta^\star \quad (\text{s.t. } H^\star = H, H^2 = E).$$

此矩阵叫作**镜面反射矩阵**, 或者 Householder 矩阵.

证明　(1) σ 是线性变换, 即验算

$$\sigma(k_1\alpha_1 + k_2\alpha_2) = k_1\sigma(\alpha_1) + k_2\sigma(\alpha_2).$$

(2) σ 是酉变换, 即验算

$$[\sigma(\alpha_1), \sigma(\alpha_2)] = [\alpha_1, \alpha_2].$$

□

镜面反射变换的几何意义是非常明确的, 下面将在欧氏空间情形下加以解释. 首先, $[\beta, \alpha]\beta$ 是向量 α 在单位矢量 β 方向上的投影矢量. 所以 $\alpha - 2[\beta, \alpha]\beta$ 是 α 在

由 α,β 确定的平面上关于与 β 垂直的矢量的对称矢量 (方向相反, 所以才叫镜面反射. 镜面指的是以 β 作为法向量的平面). 详细见图 1.1 (三维空间情形), 其中 $\overrightarrow{AO}$ 为 α, 而 $\overrightarrow{OC}$ 的方向为 β 的方向, 是水平平面的法向量. 此外

$$\overrightarrow{OC}=2\overrightarrow{OM}=-2[\beta,\alpha]\beta.$$

根据矢量加法的三角形法则, 易见

$$2\overrightarrow{OM}=\overrightarrow{OC}=\overrightarrow{OB}+\overrightarrow{OA},\quad \text{i.e.,}\quad \overrightarrow{OB}-\alpha=2[\beta,-\alpha]\beta,$$

即 $\overrightarrow{OB}=\alpha-2[\beta,\alpha]\beta$. 注意到

$$\sigma:\alpha\mapsto\alpha-2[\beta,\alpha]\beta$$

是一个线性变换, 叫作**镜面反射变换**或者 **Householder 变换**.

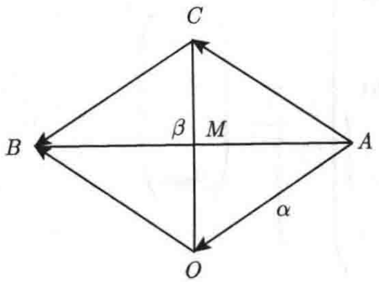

图 1.1 镜面反射变换

习题与扩展内容

习题 1 求一个镜面反射实矩阵 $B=:E-2\alpha\alpha^{\mathrm{T}}$, 使得

$$B(1,1,1,1)^{\mathrm{T}}=(1,0,0,\sqrt{3})^{\mathrm{T}}.$$

习题 2 求证: 内积空间 $\mathbb{C}^{m\times 1}$ 上的一个线性变换 σ 是关于某个单位向量 α 的镜面反射, 当且仅当 σ 在某个标准正交基下的矩阵为 $\mathrm{diag}\{-1,E_{m-1}\}$.

5. 如何理解分块矩阵的乘法?

对于初学者来说, 理解、掌握并熟练使用分块矩阵技巧进行矩阵乘法运算是一个较难的事情, 但是只要遵循由易到难的原则, 细心又有耐心, 经过一定的练习, 是不难达到熟练使用的目的的.

(1) 对于

$$C_{m\times r}=A_{m\times\boxed{n}}B_{\boxed{n}\times r},$$

假设 B 的各列依次为 $\beta_1,\beta_2,\cdots,\beta_r$, 亦即有

$$(b_{ij})_{n\times r}=B=(\beta_1,\beta_2,\cdots,\beta_r).$$

根据矩阵乘法的定义可知, AB 的第 1 列为 $AB(1)$, 这是因为

$$\begin{pmatrix} \sum_{i=1}^{n} a_{1i}b_{i1} \\ \sum_{i=1}^{n} a_{2i}b_{i1} \\ \vdots \\ \sum_{i=1}^{n} a_{mi}b_{i1} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix} = A\beta_1 = A \cdot B(1).$$

注意, 如果假设 $A = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, 则

$$A \cdot B(1) = \begin{pmatrix} \sum_{i=1}^{n} a_{1i}b_{i1} \\ \sum_{i=1}^{n} a_{2i}b_{i1} \\ \vdots \\ \sum_{i=1}^{n} a_{mi}b_{i1} \end{pmatrix} = \sum_{i=1}^{n} \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} \cdot b_{i1} = \sum_{i=1}^{n} b_{i1} \cdot A(i) = \sum_{i=1}^{n} b_{i1}\alpha_i,$$

即

$$\boxed{(\alpha_1, \alpha_2, \cdots, \alpha_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} \alpha_i \cdot b_i = \sum_{i=1}^{n} b_i \cdot \alpha_i,} \quad \forall b_i \in \mathbb{F},\ \forall \alpha_i \in \mathbb{F}^n. \tag{4}$$

同理, AB 的第 i 列为 $A\beta_i = A \cdot B(i)$, 从而有

$$\boxed{A(\beta_1, \beta_2, \cdots, \beta_r) = AB = (A\beta_1, A\beta_2, \cdots, A\beta_r).}$$

由此可知, 如果 $B = (C, D)$, 其中

$$C = (\beta_1, \beta_2, \cdots, \beta_t), \quad D = (\beta_{t+1}, \beta_{t+2}, \cdots, \beta_r),$$

则将有

$$A(C, D) = AB = (AC, AD),$$

这是因为

$$AC = A(\beta_1, \beta_2, \cdots, \beta_t) = (A\beta_1, A\beta_2, \cdots, A\beta_t), \quad AD = (A\beta_{t+1}, A\beta_{t+2}, \cdots, A\beta_r).$$

基于同样的理由, 将有

$$\boxed{A(B_1,B_2,\cdots,B_v)=(AB_1,AB_2,\cdots,AB_v).}$$

(2) 同理, 如果 $A=\begin{pmatrix}\gamma_1\\ \gamma_2\\ \vdots\\ \gamma_m\end{pmatrix}$ 或 $A=\begin{pmatrix}A_1\\ A_2\\ \vdots\\ A_u\end{pmatrix}$, 其中 $\gamma_i(i=1,2,\cdots,m)$ 是行向量, 而 $A_i(i=1,2,\cdots,u)$ 是由 A 的连续相邻行向量组成的子矩阵, 则有

$$\boxed{\begin{pmatrix}\gamma_1\\ \gamma_2\\ \vdots\\ \gamma_m\end{pmatrix}B=\begin{pmatrix}\gamma_1 B\\ \gamma_2 B\\ \vdots\\ \gamma_m B\end{pmatrix},\quad \begin{pmatrix}A_1\\ A_2\\ \vdots\\ A_u\end{pmatrix}B=\begin{pmatrix}A_1 B\\ A_2 B\\ \vdots\\ A_u B\end{pmatrix}.}$$

(3) 对于

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)=A_{m\times n},\quad \begin{pmatrix}\delta_1^{\mathrm{T}}\\ \delta_2^{\mathrm{T}}\\ \vdots\\ \delta_n^{\mathrm{T}}\end{pmatrix}=B_{n\times s},$$

我们猜测

$$\boxed{(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}\delta_1^{\mathrm{T}}\\ \delta_2^{\mathrm{T}}\\ \vdots\\ \delta_n^{\mathrm{T}}\end{pmatrix}=\sum_{i=1}^{n}\alpha_i\delta_i^{\mathrm{T}}}=\sum_{i=1}^{n}\begin{pmatrix}\cdot\\ \cdot\\ \vdots\\ \cdot\end{pmatrix}\cdot(\cdot,\cdot,\cdots,\cdot).$$

事实上, 如果假设 $B=(b_{ij})$, 则 $\delta_i^{\mathrm{T}}=(b_{i1},b_{i2},\cdots,b_{is})(i=1,2,\cdots,n)$, 从而有

$$\alpha_i\delta_i^{\mathrm{T}}=\alpha_i(b_{i1},b_{i2},\cdots,b_{is})=(\alpha_i b_{i1},\alpha_i b_{i2},\cdots,\alpha_i b_{is})=(b_{i1}\alpha_i,b_{i2}\alpha_i,\cdots,b_{is}\alpha_i),$$

所以, 根据 (4) 式得到

$$\begin{aligned}\sum_{i=1}^{n}\alpha_i\delta_i^{\mathrm{T}}&=\left(\sum_{i=1}^{n}b_{i1}\alpha_i,\ \sum_{i=1}^{n}b_{i2}\alpha_i,\ \cdots,\ \sum_{i=1}^{n}b_{is}\alpha_i\right)\\ &=(\,A\cdot B(1),A\cdot B(2),\cdots,A\cdot B(s)\,)\\ &=A(B(1),B(2),\cdots,B(s)\,)=AB=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}\delta_1^{\mathrm{T}}\\ \delta_2^{\mathrm{T}}\\ \vdots\\ \delta_n^{\mathrm{T}}\end{pmatrix}.\end{aligned}$$

现在继续假设 AB 有意义, 并假设

$$A=(A_1,A_2),\quad B=\begin{pmatrix}B_1\\B_2\end{pmatrix}$$

已经给定, 其中 A_1B_1 有意义 (从而 A_2B_2 也有意义). 此时, 可以进一步假设 $A_i=(\alpha_{i1},\alpha_{i2},\cdots,\alpha_{ir_i})$, $B_i=\begin{pmatrix}\beta_{i1}^{\mathrm{T}}\\\beta_{i2}^{\mathrm{T}}\\\vdots\\\beta_{ir_i}^{\mathrm{T}}\end{pmatrix}$, 则有

$$\boxed{(A_1,A_2)\cdot\begin{pmatrix}B_1\\B_2\end{pmatrix}=\sum_{j=1}^{r_1}\alpha_{1j}\beta_{1j}^{\mathrm{T}}+\sum_{j=1}^{r_2}\alpha_{2j}\beta_{2j}^{\mathrm{T}}=A_1B_1+A_2B_2.}$$

(4) 假设 AB 有意义, 并假设

$$A=\begin{pmatrix}A_1&A_2\\A_3&A_4\end{pmatrix},\quad B=\begin{pmatrix}B_1&B_2\\B_3&B_4\end{pmatrix}.$$

记

$$L_1=(A_1,A_2),\quad L_2=(A_3,A_4),\quad H_1=\begin{pmatrix}B_1\\B_3\end{pmatrix},\quad H_2=\begin{pmatrix}B_2\\B_4\end{pmatrix}.$$

则有

$$A=\begin{pmatrix}L_1\\L_2\end{pmatrix},\quad B=(H_1,H_2),$$

从而得到

$$AH_1=\begin{pmatrix}L_1\\L_2\end{pmatrix}H_1=\begin{pmatrix}L_1H_1\\L_2H_1\end{pmatrix}=\begin{pmatrix}A_1B_1+A_2B_3\\A_3B_1+A_4B_3\end{pmatrix}$$

以及

$$\begin{pmatrix}A_1&A_2\\A_3&A_4\end{pmatrix}\cdot\begin{pmatrix}B_1&B_2\\B_3&B_4\end{pmatrix}=A(H_1,H_2)=(AH_1,AH_2),$$

从而有

$$\boxed{\begin{pmatrix}A_1&A_2\\A_3&A_4\end{pmatrix}\begin{pmatrix}B_1&B_2\\B_3&B_4\end{pmatrix}=\begin{pmatrix}A_1B_1+A_2B_3&A_1B_2+A_2B_4\\A_3B_1+A_4B_3&A_3B_2+A_4B_4\end{pmatrix}.}$$

(5) 如果采用归纳法, 由 (4) 中最后一式不难得到更一般情形的分块矩阵的乘