计算方法丛书

矩阵扰动分析

孙继广 著

科学出版社

计算方法丛书

矩阵扰动分析

孙继广 著

科学出版社

1987

内 容 简 介

本书系统地论述了矩阵扰动分析的理论、方法、技巧和新的进展，重点在于矩阵空间的范数与度量、代数特征值问题的扰动理论、广义逆与最小二乘问题的扰动理论. 每节都附有难易程度不同的习题.

本书可供计算数学工作者、工程技术人员、高等学校有关专业的高年级学生、研究生及教师参考.

计算方法丛书

矩阵扰动分析

孙继广 著

责任编辑 林 鹏 张鸿林

科学出版社出版

北京朝阳门内大街137号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1987年2月第一版 开本：850×1168 1/32

1987年2月第一次印刷 印张：12 1/8

印数：0001—3,750 字数：312,000

统一书号：13031·3417

本社书号：5033·13—1

定价：3.40元

前　　言

所谓"矩阵扰动分析"就是研究矩阵元素的变化对于矩阵问题的解的影响．这个课题的研究，不仅对于矩阵论和算子理论，而且对于矩阵计算，都有重要意义．

与矩阵计算有关的扰动理论及其主要结果是在最近二、三十年里得到的．近十年来，随着各种科学计算问题的深入与扩大以及计算机的发展，矩阵扰动理论又有不少新的进展；同时，还存在许多问题有待于进一步解决．

这本书是从计算数学的角度，利用代数方法，系统地、比较全面地阐述矩阵扰动理论的主要结果．据作者所知，国内外还没有一部这样的专门著作．

本书除了介绍必要的预备知识之外，重点在以下三个方面：1.矩阵空间上的范数与度量，2.代数特征值问题（包括普通的和广义的特征值问题）的扰动理论，3.广义逆与最小二乘问题的扰动理论．本书在内容上，力求向读者展示矩阵扰动理论中最基本、最重要的知识和方法技巧，以及若干新的进展，并且对所述结果给出完整的和严格的数学论证．书中每一节都安排几道难易程度不同的习题，书末附有比较详细的参考文献目录．

作者曾以本书的部分章节作为教材，在中国科学院研究生院讲授"矩阵扰动分析"课程．作者希望，已经学习过线性代数、数学分析和复变函数论基础教程的读者，在阅读本书时不致于感到困难．

作者感谢冯康教授以及中国科学院计算中心科学计算部的同志们的鼓励与支持，感谢复旦大学蒋尔雄同志对书稿仔细认真的审阅．

请读者对本书提出批评指正．

孙继广

1984 年国庆节

符 号 表

$\mathbf{C}^{m\times n}$	所有$m\times n$复元素矩阵的全体
$\mathbf{R}^{m\times n}$	所有$m\times n$实元素矩阵的全体
$\mathbf{C}_r^{m\times n}$	$\mathbf{C}^{m\times n}$ 中所有秩为r的矩阵的全体
$\mathbf{C}^n$	所有复n维列向量的全体(即 $\mathbf{C}^{n\times 1}$)
$\mathbf{R}^n$	所有实n维列向量的全体(即 $\mathbf{R}^{n\times 1}$)
$\mathbf{C}$	所有复数的全体(即$\mathbf{C}^1$)
$\mathbf{R}$	所有实数的全体(即$\mathbf{R}^1$)
$\mathscr{U}_n$	所有$n\times n$酉阵的全体
$\bar{A}$	矩阵A的共轭
A^T	矩阵A的转置
A^H	矩阵A的共轭转置(即 $\bar{A}^T$)
A^{-1}	矩阵A的逆
$A^\dagger$	矩阵A的广义逆(即A的 Moore-Penrose 广义逆)
$A>0$	指矩阵A是正定的 Hermite 阵
$A\geqslant 0$	指矩阵A是半正定的 Hermite 阵
I	单位矩阵
$I^{(n)}$	$n\times n$单位矩阵
0	零矩阵
$R(A)$	由矩阵A的所有列向量所张成的子空间
$N(A)$	矩阵A的零空间
G_l^n	$\mathbf{C}^n$ 内所有l维列空间的全体
P_A	到 $R(A)$ 上的正交投影算子
$\det A$	矩阵A的行列式
$\mathrm{rank}(A)$	矩阵A的秩
$\mathrm{tr}(A)$	矩阵A的迹

$\rho(A)$	矩阵 A 的谱半径
$\lambda(A)$	矩阵 A 的所有特征值的全体
$\sigma(A)$	矩阵 A 的所有奇异值的全体
$\sigma_{\max}(A)$	矩阵 A 的最大奇异值
$\sigma_{\min}(A)$	矩阵 A 的最小奇异值
$\|\boldsymbol{x}\|_2$	向量 $\boldsymbol{x}$ 的 Euclid 长度
$\|A\|_2$	矩阵 A 的谱范数
$\|A\|_F$	矩阵 A 的 Frobenius 范数
$\in$	元素属于
$\subseteq$	集合含于
$\cup$	集合的并
$\cap$	集合的交
$\varnothing$	空集
$\Leftrightarrow$	等价
$\Rightarrow$	蕴涵

目　　录

第一章 预备知识

§1 特征值与特征向量

本节列举矩阵代数的几条熟知的定义和结论.

定义 1.1. 设 $A\in\mathbf{C}^{n\times n}$. 如果存在$\lambda\in\mathbf{C}$和非零向量$\boldsymbol{x}\in\mathbf{C}^n$, 使得 $A\boldsymbol{x}=\lambda\boldsymbol{x}$, 则$\lambda$叫做$A$的特征值, $\boldsymbol{x}$叫做A的属于特征值λ的特征向量.

A的所有特征值的全体,叫做A的谱(spectrum), 记作$\lambda(A)$.

据定义 1.1, $\lambda\in\lambda(A)$ 的必要与充分条件是

$$\det(\lambda I-A)=0.$$

$p(\lambda)\equiv\det(\lambda I-A)$ 叫做A的特征多项式. 如果A有r个不同的特征值 $\lambda_1,\cdots,\lambda_r$, 其重数分别为$n(\lambda_1),\cdots,n(\lambda_r)$, 则

$$p(\lambda)=\prod_{i=1}^{r}(\lambda-\lambda_i)^{n(\lambda_i)},\quad \lambda_i\neq\lambda_j\ (i\neq j),$$

$$\sum_{i=1}^{r}n(\lambda_i)=n,$$

其中 $n(\lambda_i)$ 叫做 λ_i的代数重数.如果

$$\operatorname{rank}(\lambda_i I-A)=n-m(\lambda_i),$$

则 $m(\lambda_i)$ 叫做 λ_i的几何重数, 它表示A的属于λ_i的线性无关特征向量的个数. 显然有

$$1\leqslant m(\lambda_i)\leqslant n(\lambda_i)\leqslant n.$$

定理 1.1. 设 $A\in\mathbf{C}^{n\times n}$ 有r个不同的特征值 $\lambda_1,\cdots,\lambda_r$, 其代数重数分别为 $n(\lambda_1),\cdots,n(\lambda_r)$, 则必存在非奇异阵 $P\in\mathbf{C}^{n\times n}$, 使得

$$P^{-1}AP=J\equiv\operatorname{diag}(J_1(\lambda_1),\cdots,J_r(\lambda_r)),\tag{1.1}$$

其中

$$J_i(\lambda_i)=\mathrm{diag}(J_i^{(1)}(\lambda_i),\cdots,J_i^{(k_i)}(\lambda_i))\in\mathbf{C}^{n(\lambda_i)\times n(\lambda_i)},\ 1\leqslant i\leqslant r, \tag{1.2}$$

$$J_i^{(k)}(\lambda_i)=\begin{pmatrix}\lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1\\ & & & \lambda_i\end{pmatrix}\in\mathbf{C}^{n_k(\lambda_i)\times n_k(\lambda_i)},1\leqslant k\leqslant k_i, \tag{1.3}$$

$$\sum_{k=1}^{k_i}n_k(\lambda_i)=n(\lambda_i),\ 1\leqslant i\leqslant r;$$

并且除了 $J_i^{(k)}(\lambda_i)$ ($1\leqslant k\leqslant k_i$, $1\leqslant i\leqslant r$) 的编排次序可以改变外，J 是唯一确定的.

(1.3) 式所示的每个矩阵 $J_i^{(k)}(\lambda_i)$ ($1\leqslant k\leqslant k_i$, $1\leqslant i\leqslant r$) 叫做 Jordan 块;矩阵 J 叫做 A 的 Jordan 标准形.

定理 1.1 是矩阵论的基本结果之一，它的证明可以在普通的线性代数教程中找到.

定义 1.2. 设 $A\in\mathbf{C}^{n\times n}$. 如果存在非奇异阵 $Q\in\mathbf{C}^{n\times n}$,使得

$$Q^{-1}AQ=\mathrm{diag}(\lambda_1,\cdots,\lambda_n),$$

则 A 叫做可对角化阵(又叫单构阵,或者可正规化阵).

易知下面 5 种叙述互相等价:

1) A 是可对角化阵,

2) $\mathbf{C}^n$ 存在由 A 的特征向量构成的一组基底,

3) A 的初等因子都是线性的,

4) A 的 Jordan 标准形中的 Jordan 块都是 1 阶的,

5) $m(\lambda_i)=n(\lambda_i)$, $\forall\lambda_i\in\lambda(A)$.

定义 1.3. 设 $A\in\mathbf{C}^{n\times n}$. 如果 $AA^H=A^HA$, 则称 A 为正规矩阵 (以下简称之为正规阵); 如果 $A^H=A$, 则称 A 为 Hermite 阵;如果 $A^T=\bar{A}=A$,则称 A 为实对称阵;如果 $A^HA=I$, 则称 A 为酉阵;如果 $A^TA=I$ 并且 $\bar{A}=A$, 则称 A 为实正交阵.

下面的定理 1.2 是熟知的(§ 2 习题 4 提出了一种证法).

定理 1.2 (Schur 定理). 设 $A\in\mathbf{C}^{n\times n}$, 则存在酉阵 U, 使得

$$U^HAU=T, \tag{1.4}$$

其中T是上三角阵;而且适当选取U,可使T的对角线元素按任一指定顺序排列.

(1.4)式右端的T,叫做A的 Schur 上三角形式. 可记

$$T = \Lambda + M, \quad \Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n),$$

其中M为严格上三角阵(即对角线元素为零的上三角阵).

从定理 1.2 可得

推论 1.1. 下列结论成立:

1) A是正规阵$\Longleftrightarrow$存在酉阵U,使得

$$U^H A U = \mathrm{diag}(\lambda_1, \cdots, \lambda_n),$$

即$\mathbf{C}^n$存在由A的特征向量构成的标准正交基.

2) A是 Hermite 阵$\Longleftrightarrow A$是正规阵, 并且$\lambda(A) \subseteq \mathbf{R}$.

3) A是酉阵$\Longleftrightarrow A$是正规阵,并且

$$\lambda(A) \subseteq \mathscr{S} \equiv \{\lambda \in \mathbf{C}: |\lambda| = 1\}.$$

同理可得实对称阵和实正交阵的相类似的结论.

定义 1.4. 设 $A \in \mathbf{C}^{n\times n}$ 是 Hermite 阵. 如果

$$\boldsymbol{x}^H A \boldsymbol{x} > 0, \ \forall \boldsymbol{x} \in \mathbf{C}^n, \ \boldsymbol{x} \neq 0$$

(或 $\boldsymbol{x}^H A \boldsymbol{x} \geqslant 0, \ \forall \boldsymbol{x} \in \mathbf{C}^n$), 则$A$叫做正定(或半正定)阵.

以下用 $A > 0$ ($A \geqslant 0$) 表示A是正定(半正定)阵.

设 $A \in \mathbf{C}^{n\times n}$ 为 Hermite 阵. 容易验证下面的5种叙述是互相等价的:

1) A是正定(半正定)阵,

2) A的每个特征值$>0(\geq 0)$,

3) A的每个主子阵都是正定(半正定)阵,

4) A的每个主子式(即主子阵的行列式)$>0(\geq 0)$,

5) 对任一$n \times n$非奇异阵 Q, $Q^H A Q$ 是正定(半正定)阵.

习题

1. 设A是正规阵,同时A是上三角阵. 证明A是对角阵.

2. 设A为可对角化阵. 试证: A的特征值皆为实数的必要与充分条件是存在正定阵H,使得 HA 为 Hermite 阵.

3. 设 $A\in\mathbf{C}^{n\times n}$，$A$ 的所有特征值互不相等。则 A 的特征向量构成 $\mathbf{C}^n$ 的一组基底。

4. 设 $AB=BA$，其中 B 为幂零阵（即存在自然数 m，使得 $B^m=0$）。则 $\det(A+B)=\det A$。

5. 设 $A=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}$，其中 A_{ij} 均为方阵，并且满足 $A_{11}A_{21}=A_{21}A_{11}$。则

$$\det A=\det(A_{11}A_{22}-A_{21}A_{12}).$$

6. 设 $\det\begin{pmatrix}A & \boldsymbol{a}\\ \boldsymbol{a}^H & \alpha\end{pmatrix}=0$。则

$$\det\begin{pmatrix}A & \boldsymbol{a}\\ \boldsymbol{a}^H & \beta\end{pmatrix}=(\beta-\alpha)\det A.$$

7. 设 A 与 B 均为可对角化阵。试证下列三个条件互相等价:

i) $AB=BA$,

ii) 存在非奇异阵 X，使 $X^{-1}AX$ 与 $X^{-1}BX$ 同时为对角阵,

iii) 存在一个可对角化阵 C 及一对多项式 $p(\lambda)$ 与 $q(\lambda)$，使得 $A=p(C)$ 和 $B=q(C)$.

8. 至少给出两种不同的方法，计算行列式 $\det(I-\sigma\boldsymbol{u}\boldsymbol{v}^H)$，其中 $\boldsymbol{u},\boldsymbol{v}\in\mathbf{C}^n$，$\sigma\in\mathbf{C}$.

9. 证明 AB 与 BA 的特征多项式除了一个 λ 的幂之外是相等的.

10. 设 H_1 为正定阵，H_2 是与 H_1 同阶的 Hermite 阵. 试证: H_1+H_2 为正定阵的必要与充分条件是 $H_1^{-1}H_2$ 的特征值均大于 -1.

§2 初 等 矩 阵

2.1 初等矩阵的一般形式

定义 2.1. 形如

$$E(\boldsymbol{u},\boldsymbol{v};\sigma)\equiv I-\sigma\boldsymbol{u}\boldsymbol{v}^H$$

的矩阵叫做初等矩阵，其中 $\boldsymbol{u},\boldsymbol{v}\in \mathbf{C}^n$，$\sigma\in \mathbf{C}$.

初等矩阵是数值代数的基本工具之一.

显然 $E(\boldsymbol{u},\boldsymbol{v};0)=I$. 所以下面仅讨论 $\sigma\neq 0$ 的情形.

初等矩阵的性质:

1）特征向量. 若 $\boldsymbol{u}\notin \boldsymbol{v}^{\perp}$（$\boldsymbol{v}^{\perp}$ 表示与 $\boldsymbol{v}$ 正交的 $n-1$ 维子空间），则 $E(\boldsymbol{u},\boldsymbol{v};\sigma)$ 有 n 个线性无关的特征向量，该组特征向量由 $\boldsymbol{u}$ 及 $\boldsymbol{v}^{\perp}$ 中任取一组基底构成；若 $\boldsymbol{u}\in \boldsymbol{v}^{\perp}$，则$E(\boldsymbol{u},\boldsymbol{v};\sigma)$仅有 $n-1$ 个线性无关的特征向量，该组特征向量由 $\boldsymbol{v}^{\perp}$ 中任取一组基底构成.

证明:

首先在 $\boldsymbol{v}^{\perp}$ 中任取一组基底 $\{\boldsymbol{u}_i\}_{i=1}^{n-1}$，有

$$E(\boldsymbol{u},\boldsymbol{v};\sigma)\ \boldsymbol{u}_i=\boldsymbol{u}_i,\quad i=1,\cdots,n-1.$$

因此 $\boldsymbol{u}_i(i=1,\cdots,n-1)$是 $E(\boldsymbol{u},\boldsymbol{v};\sigma)$ 的属于特征值 1 的特征向量.

若 $\boldsymbol{u}\notin \boldsymbol{v}^{\perp}$，则有

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

即 $\boldsymbol{u}$ 也是 $E(\boldsymbol{u},\boldsymbol{v};\sigma)$ 的特征向量，相应的特征值是 $1-\sigma\boldsymbol{v}^H\boldsymbol{u}$.

若 $\boldsymbol{u}\in \boldsymbol{v}^{\perp}$，则这时 $E(\boldsymbol{u},\boldsymbol{v};\sigma)$ 的任一特征向量 $\boldsymbol{x}$ 必有 $\boldsymbol{x}\in \boldsymbol{v}^{\perp}$. 事实上，由

$$E(\boldsymbol{u},\boldsymbol{v};\sigma)\boldsymbol{x}=q\boldsymbol{x}$$

可得

$$(1-q)\boldsymbol{x}=\sigma\boldsymbol{v}^H\boldsymbol{x}\boldsymbol{u}.$$

如果 $q=1$，则有 $\boldsymbol{v}^H\boldsymbol{x}=0$，即 $\boldsymbol{x}\in \boldsymbol{v}^{\perp}$；如果 $q\neq 1$，假定 $\boldsymbol{x}\notin \boldsymbol{v}^{\perp}$，则 $\boldsymbol{x}$ 应与 $\boldsymbol{u}$ 共线，但这与已知 $\boldsymbol{u}\in \boldsymbol{v}^{\perp}$ 矛盾，所以这时亦有 $\boldsymbol{x}\in \boldsymbol{v}^{\perp}$.

2）特征值. 在 $\boldsymbol{v}^{\perp}$ 内取一标准正交基 $\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$，构造酉阵

$$U=\left(\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|_2},\ \boldsymbol{u}_2,\cdots,\boldsymbol{u}_n\right).$$

由