



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

## 《计算方法丛书》编委会

主 副 编	主 编 委	冯 康				
		石钟慈	李岳生			
		王仁宏	王汝权	孙继广	何旭初	吴文达
		李庆扬	李德元	林 群	周毓麟	席少霖
		徐利治	郭本瑜	袁兆鼎	黄鸿慈	蒋尔雄
		雷晋干	滕振寰			

## 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

• 1 •

## 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 $\mathbf{R}^1$ )
$\mathcal{U}_n$	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
$A \geq 0$	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
$\text{rank}(A)$	矩阵 $A$ 的秩
$\text{tr}(A)$	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
- $\|\mathbf{x}\|_2$  向量  $\mathbf{x}$  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
- $\emptyset$  空集
- $\Leftrightarrow$  等价
- $\Rightarrow$  蕴涵

# 目 录

第一章 预备知识.....	1
§ 1 特征值与特征向量 .....	1
习题 .....	3
§ 2 初等矩阵 .....	4
2.1 初等矩阵的一般形式 .....	4
2.2 初等下三角阵 .....	7
2.3 初等 Hermite 阵 .....	8
习题 .....	10
§ 3 矩阵分解 .....	11
习题 .....	21
§ 4 值域 .....	21
习题 .....	25
§ 5 Kronecker 乘积 .....	25
5.1 基本概念 .....	25
5.2 应用举例:线性矩阵方程 .....	26
习题 .....	28
§ 6 广义逆 .....	28
6.1 基本概念 .....	28
6.2 基本性质 .....	31
习题 .....	34
§ 7 投影 .....	34
7.1 幂等阵与投影 .....	35
7.2 正交投影 .....	38
7.3 $AA^\dagger$ 与 $A^\dagger A$ 的几何意义 .....	40
7.4 应用举例:线性最小二乘问题 .....	42
习题 .....	43
§ 8 行列式 .....	44



8.1	Binet-Cauchy 公式	44
8.2	Hadamard 不等式	46
	习题	50
	第一章说明	51
第二章	范数与度量	52
§ 1	$\mathbb{C}^n$ 上的范数	52
	习题	57
§ 2	$\mathbb{C}^{m \times n}$ 上的范数	58
2.1	基本概念	58
2.2	算子范数	61
	习题	69
§ 3	$\mathbb{C}^{m \times n}$ 上的酉不变范数	70
3.1	定义	70
3.2	von Neumann 不等式	72
3.3	SG 函数	75
3.4	酉不变范数的性质	83
	习题	86
§ 4	$\mathbb{G}_7$ 上的度量	88
4.1	基本概念	88
4.2	关于 $\ \sin \theta(Z, W)\ _2$	90
4.3	关于 $\ \sin \theta(Z, W)\ $	95
4.4	其它的度量	97
	习题	103
	第二章说明	104
第三章	特征值问题扰动分析	105
§ 1	特征值问题的稳定性	105
1.1	特征值的连续性	105
1.2	扰动性质的数学描述	110
	习题	113
§ 2	Gerschgorin 理论	113
2.1	Gerschgorin 定理	113
2.2	应用举例	116

习题	119
§ 3 Hermite 阵的特征值	120
3.1 极小极大定理	120
3.2 极小极大定理的一般形式	124
3.3 Hermite 扰动	131
3.4 关于奇异值的扰动	134
习题	136
§ 4 正规阵与可正规化阵的特征值	138
4.1 正规阵与可正规化阵	138
4.2 Hoffman-Wielandt 定理	139
4.3 Bauer-Fike 定理	145
4.4 Hermite 阵的任意扰动	146
习题	151
§ 5 一般方阵的特征值	152
5.1 推广的 Bauer-Fike 定理	152
5.2 Henrici 定理	154
5.3 正规性偏离度的估计	159
5.4 Henrici 定理(续)	164
5.5 举例	171
习题	173
§ 6 条件数	173
6.1 特征值问题病态程度的数据标准	173
6.2 几种条件数之间的关系	178
习题	184
§ 7 特征空间的扰动界限	185
7.1 Rayleigh 商和剩余	185
7.2 Davis-Kahan $\sin \theta$ 定理	192
7.3 与近似特征空间有关的其它估计	198
习题	204
§ 8 不变子空间的扰动界限	205
8.1 不变子空间	205
8.2 一个非线性方程及其解的估计	210

8.3 剩余与矩阵分离度 .....	215
8.4 扰动定理 .....	222
习题 .....	224
第三章说明 .....	225
第四章 广义特征值问题扰动分析 .....	227
§ 1 基本概念 .....	227
1.1 正则对与奇异对 .....	228
1.2 特征值与特征向量 .....	230
1.3 广义特征值问题的稳定性 .....	235
1.4 几类重要的正则对 .....	243
习题 .....	246
§ 2 Gerschgorin 理论 .....	247
2.1 Gerschgorin 型定理 .....	247
2.2 应用举例 .....	251
习题 .....	255
§ 3 定型对的特征值 .....	255
3.1 Crawford 数 $c(A, B)$ 的性质 .....	255
3.2 $\mathbb{D}(n)$ 上的一种投影度量 .....	257
3.3 Weyl-Лидский 型定理 .....	260
3.4 关于广义奇异值的扰动 .....	266
习题 .....	271
§ 4 正规对、可对角化对与一般正则对的特征值 .....	272
4.1 Hoffman-Wielandt 型定理 .....	273
4.2 Bauer-Fike 型定理 .....	278
4.3 Henrici 型定理 .....	281
4.4 $d_2(Z, W)$ 与 $d_F(Z, W)$ 的估计 .....	285
习题 .....	287
§ 5 特征空间的扰动界限 .....	288
5.1 特征空间 .....	288
5.2 $\sin \theta$ 第一定理 .....	290
5.3 $\sin \theta$ 第二定理 .....	300
习题 .....	303

§ 6 广义不变子空间的扰动界限 .....	304
6.1 广义不变子空间 .....	304
6.2 算子 $T(P, Q)$ 和函数 $\text{dif}$ .....	306
6.3 逼近定理与扰动定理 .....	313
习题 .....	316
第四章说明 .....	317
第五章 广义逆与最小二乘问题扰动分析 .....	319
§ 1 矩阵逆与线性方程组解的扰动 .....	319
1.1 矩阵逆的扰动界限 .....	321
1.2 线性方程组解的扰动界限 .....	323
习题 .....	325
§ 2 广义逆扰动分析 .....	325
2.1 关于一对投影 .....	325
2.2 锐角扰动 .....	333
2.3 广义逆的扰动界限 .....	336
习题 .....	348
§ 3 投影的扰动 .....	348
3.1 关于投影的连续性 .....	348
3.2 投影的扰动界限 .....	350
习题 .....	355
§ 4 线性最小二乘问题扰动分析 .....	355
习题 .....	361
第五章说明 .....	361
参考文献 .....	363

# 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  $A \in \mathbb{C}^{n \times n}$ . 如果存在  $\lambda \in \mathbb{C}$  和非零向量  $\mathbf{x} \in \mathbb{C}^n$ , 使得  $A\mathbf{x} = \lambda\mathbf{x}$ , 则  $\lambda$  叫做  $A$  的特征值,  $\mathbf{x}$  叫做  $A$  的属于特征值  $\lambda$  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  $\lambda_1, \dots, \lambda_r$ , 其重数分别为  $n(\lambda_1), \dots, n(\lambda_r)$ , 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  $\lambda_i$  的代数重数. 如果

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

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

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

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

$$P^{-1}AP = J \equiv \text{diag}(J_1(\lambda_1), \dots, J_r(\lambda_r)), \quad (1.1)$$

其中

$$J_i(\lambda_i) = \text{diag}(J_i^{(1)}(\lambda_i), \dots, J_i^{(k_i)}(\lambda_i)) \in \mathbb{C}^{n(\lambda_i) \times n(\lambda_i)}, \quad 1 \leq i \leq r, \quad (1.2)$$

$$J_i^{(k)}(\lambda_i) = \begin{pmatrix} \lambda_i & & & 1 & & \\ & \ddots & & \vdots & & \\ & & \ddots & \ddots & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{n_k(\lambda_i) \times n_k(\lambda_i)}, \quad 1 \leq k \leq k_i, \quad (1.3)$$
$$\sum_{k=1}^{k_i} n_k(\lambda_i) = n(\lambda_i), \quad 1 \leq i \leq r;$$

并且除了  $J_i^{(k)}(\lambda_i)$  ( $1 \leq k \leq k_i$ ,  $1 \leq i \leq r$ ) 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  $J_i^{(k)}(\lambda_i)$  ( $1 \leq k \leq k_i$ ,  $1 \leq i \leq r$ ) 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  $A \in \mathbb{C}^{n \times n}$ . 如果存在非奇异阵  $Q \in \mathbb{C}^{n \times n}$ , 使得

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

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)  $\mathbb{C}^n$  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  $A \in \mathbb{C}^{n \times n}$ . 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  $A \in \mathbb{C}^{n \times n}$ , 则存在酉阵  $U$ , 使得

$$U^H A U = T, \quad (1.4)$$

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

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

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  $\Leftrightarrow$  存在酉阵  $U$ , 使得

$$U^H A U = \text{diag}(\lambda_1, \dots, \lambda_n),$$

即  $\mathbb{C}^n$  存在由  $A$  的特征向量构成的标准正交基.

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

3)  $A$  是酉阵  $\Leftrightarrow A$  是正规阵, 并且

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

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  $A \in \mathbb{C}^{n \times n}$  是 Hermite 阵. 如果

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

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

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

设  $A \in \mathbb{C}^{n \times n}$  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  $A \in \mathbb{C}^{n \times n}$ ,  $A$  的所有特征值互不相等. 则  $A$  的特征向量构成  $\mathbb{C}^n$  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  $\det \begin{pmatrix} A & \alpha \\ \alpha^H & \alpha \end{pmatrix} = 0$ . 则

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

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  $\det(I - \sigma uv^H)$ , 其中  $u, v \in \mathbb{C}^n$ ,  $\sigma \in \mathbb{C}$ .

9. 证明  $AB$  与  $BA$  的特征多项式除了一个  $\lambda$  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

$$E(u, v; \sigma) \equiv I - \sigma uv^H$$



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

初等矩阵是数值代数的基本工具之一.

显然  $E(u, v; 0) = I$ . 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  $u \notin v^\perp$  ( $v^\perp$  表示与  $v$  正交的  $n-1$  维子空间), 则  $E(u, v; \sigma)$  有  $n$  个线性无关的特征向量, 该组特征向量由  $u$  及  $v^\perp$  中任取一组基底构成; 若  $u \in v^\perp$ , 则  $E(u, v; \sigma)$  仅有  $n-1$  个线性无关的特征向量, 该组特征向量由  $v^\perp$  中任取一组基底构成.

证明:

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

$$E(u, v; \sigma) u_i = u_i, \quad i = 1, \dots, n-1.$$

因此  $u_i (i = 1, \dots, n-1)$  是  $E(u, v; \sigma)$  的属于特征值 1 的特征向量.

若  $u \notin v^\perp$ , 则有

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

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

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

事实上, 由

$$E(u, v; \sigma)x = qx$$

可得

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

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

2) 特征值. 在  $v^\perp$  内取一标准正交基  $u_2, \dots, u_n$ , 构造酉阵

$$U = \left( \frac{v}{\|v\|_2}, u_2, \dots, u_n \right).$$

由