矩阵迭代分析

R. S. 瓦格 著

上海科学技术出版社

矩陣迭代分析

〔美〕R. S. 瓦格 著
蔣尔雄 游光永 張玉德 譯

上海科学技术出版社

內 容 提 要

本书以非負矩陣的理論为基础，对近年来随着数字計算机的飞跃发展而建立起来的現代循环迭代法作了綜合性的論述。全书共分九章。第一章介紹向量模、矩陣模、方向图、对角占优陣等基本概念，第二章討論非負矩陣的理論，第三、四、五章分析逐次超松弛迭代法的几种变形，第六章介紹导出椭圓型方程的差分逼近的几种观点，第七章討論交替方向隐式法，第八章研究抛物型偏微分方程的矩陣方法，最后一章处理最佳迭代参数的估計問題。另有两篇附录，列出了数值求解的整个过程。每节之后，附有一定数量的练习，其中大部分可視为正文內容的推广，对讀者有一定帮助。

本书可供数学系計算数学专业作为教学参考书，也可供从事实际計算工作的工程技术人員参考。

MATRIX ITERATIVE ANALYSIS

R. S. Varga

Prentice-Hall, Inc., 1962

矩 陣 迭 代 分 析

蔣尔雄 游兆永 張玉德 譯

上海科学技术出版社出版（上海瑞金二路 450 号）

上海市书刊出版业营业許可証出 093 号

上海市印刷五厂印刷　　新华书店上海发行所发行

开本 850×1156 1/32　印張 9 18/32　排版字数 254,000

1966 年 7 月第 1 版　1966 年 7 月第 1 次印刷

印数 1—1,800

統一书号 13119·691　定价（科六）1.40 元

目　录

第一章　矩陣性质和概念 …… 1

1.1　引言 …… 1

1.2　一个簡单的例子 …… 3

1.3　模和譜半徑 …… 7

1.4　矩陣的譜半徑的界和方向图 …… 15

1.5　对角占优陣 …… 21

书目与討論 …… 23

第二章　非負矩陣 …… 25

2.1　非負矩陣的譜半徑 …… 25

2.2　循环陣和本原陣 …… 33

2.3　可約矩陣 …… 43

2.4　非負矩陣和方向图 …… 45

书目与討論 …… 50

第三章　基本迭代法和比較定理 …… 52

3.1　点 Jacobi、Gauss-Seidel 和逐次超松弛迭代法 …… 52

3.2　平均收斂率 …… 57

3.3　Stein-Rosenberg 定理 …… 63

3.4　Ostrowski-Reich 定理 …… 69

3.5　Stieltjes 陣和 M-陣 …… 75

3.6　矩陣的正規分裂 …… 81

书目与討論 …… 88

第四章　逐次超松弛迭代法 …… 91

4.1　p-循环陣 …… 91

4.2　p-循环陣的逐次超松弛迭代法 …… 99

4.3　最佳松弛因子在理論上的确定 …… 102

4.4　矩陣的 p-循环理論的推广 …… 108

4.5　漸近收斂率 …… 119

书目与討論 …… 121

第五章 半迭代法……123
5.1 半迭代法和 Chebyshev 多項式……123
5.2 半迭代法和逐次超松弛迭代法的关系……131
5.3 平均收斂率的比較:弱循环情形……138
5.4 循环归化及有关的迭代法……143
书目与討論……147
第六章 椭圓型差分方程的导出与求解……150
6.1 一个簡单的两点边値問題……150
6.2 一般二阶常微分方程……161
6.3 高維情况差分逼近的导出……169
6.4 因子分解技巧和块迭代法……181
6.5 模型問題的漸近收斂率……188
书目与討論……192
第七章 交替方向隐式迭代法……195
7.1 Peaceman-Rachford 迭代法……195
7.2 可交換情形……204
7.3 不可交换情形……216
7.4 Peaceman-Rachford 迭代法的变形……223
书目与討論……230
第八章 抛物型偏微分方程的矩陣方法……233
8.1 半离散逼近……233
8.2 本性正矩陣……239
8.3 $\exp(-tS)$的矩陣逼近……244
8.4 与解椭圓型差分方程的各种迭代法的联系……252
书目与討論……261
第九章 加速参数的估計……264
9.1 非負矩陣理論的应用……264
9.2 等周不等式的应用……272
书目与討論……277
附录……278
参考文献……284
符号表……298
索引……299

第一章

矩陣性质和概念

1.1 引　　言

从本书的名称——矩陣迭代分析——来看，我們应在这里討論一切属于迭代性质的矩陣数値方法. 但是事实上，这样龐大的目标已改为更实用的目标：我們只是較詳細地考虑数值分析的一个較小的分支，即关于用迭代法求解从偏微分方程的离散逼近所引出的矩陣方程的有效解法. 这些矩陣方程的矩陣通常具有稀疏的特点，即它的很大一部分元素等于零. 此外，这些矩陣的非零元素都以一定的規律出現，因此，相对于数字計算机来說，即使非常高阶的矩陣，也能有效地存儲. 各种循环迭代法都理想地适用于这种矩陣方程，这是因为每一步所要求的数字計算机存儲量或算术运算量都相对地少. 作为在数字計算机上曾成功地用循环迭代法解出的問題大小的一个例子，Westinghouse 电力公司的 Bettis 原子能实驗室在 1960 年經常使用一个二維程序，它以处理 20000 阶的 Laplace 型矩陣方程作为特殊情形❶.

用迭代法求解大型綫性方程組的想法当然不是新的，至少可追溯到 Gauss (1823). 以后，Southwell (1946) 和他的学生系統地考虑了实用物理学和工程問題的数値解法，对迭代法的应用起了眞正的推动作用. Southwell 所倡議的松弛迭代法是非循环迭代法，多年来，在用紙笔或台式計算机进行必要的算术运算中，被成功地使用着，又，这个方法在以人的观察去指导計算的全部过程时

---

❶ 这个称为“PDQ-4”的程序是专为带有 32000 个单詞的磁心存儲器的 Philco-2000 計算机而編制的. Bettis 的更大的努力是用三維程序“TNT-1”去处理 108000 阶的偶合的矩陣方程.

特别有效. 对于大型数字計算机来說，人的观察一般是难于有效地加入到計算机程序中去的. 因此，数学家們就着手寻找使基本循环迭代法或古典迭代法加速收斂的途徑. 上述迭代法和非循环迭代法恰好相反，它是在求解矩陣方程的过程中一經最初規定以后不再变更的迭代法. 我們在这里将只限于考虑循环迭代法（簡称为迭代法）；非循环迭代法的理論和应用在别的地方已有十分丰富的論述❶，且它們一般在大型数字計算机上并不使用.

数值分析中有关循环迭代法这一領域內的大部分近代活动，它的基础是 Frankel (1950)，Geiringer (1949)，Reich (1949)，Stein-Rosenberg (1948)，以及 Young (1950) 等一系列論文，所有这些論文都是在数字計算机显示出巨大的威力时出現的. 由于这些論文对这一領域內的流行研究的潮流所产生的巨大冲击，所以我們认为把上面提及的作者在 1948 年左右开始的工作称之为現代矩陣迭代分析是合适的. 从这个观点出发，我們的第一个目标是要描述現代矩陣迭代分析从它开始到現在的一些基本結果.

这里假定讀者已具备了矩陣和綫代数理論的基本知識，即 Birkhoff 与 MacLane (1953)，Faddeeva (1959)，及 Bellman (1960) 等书中叙述到的內容. 于是，举例来說，假定讀者已經知道什么是复矩陣的 Jordan 法式.

除了能够独立地閱讀的若干孤立部分外，我們的第二个目标是使这里的內容能适当地自給自足和完备. 象我們将要看到的，这里的矩陣迭代分析的发展主要依賴于 Perron (1907) 和 Frobenius (1908～1912) 关于具有非負元素的矩陣的早期研究；于是，我們的第一个目标不仅要描述这一領域內的基本結果，同时也要利用非負矩陣的 Perron-Frobenius 理論作为解釋这些結果的基础. 按照本书內容能自給自足的这个目标，对于 Perron-Frobenius 理論，虽然 Gantmacher (1959) 的书中有专章討論，但我們仍在第二章中作了专門討論.

我們的第三个目标是为这些理論对偏微分方程数值解法的实

---

❶ 在本章末尾的书目与討論中給出有参考文献.

际应用感到极大兴趣的讀者提供充分的数值細节. 为了这个目的,在附录A和B中收集的都是說明性例子,它們指出了从問題的列出、矩陣方程的推导、各种迭代法的应用、到数字計算机輸出的数值結果的最后檢驗等整个过程. 对眞正数值应用感兴趣的讀者会极力主張把附录中例子的細节全部作出. 在每一章,我們也为讀者安排了一些练习;这些练习不仅用来檢查掌握該章內容的熟练程度,而且在很多情形下还指出了沒有包含在本书中的有意义的理論結果和扩充. 带星号的练习对部分讀者可能要求更多的努力.

本书的內容是这样安排的,即从自共軛椭圓型偏微分方程到矩陣方程的一般推导(第六章)要等建立了大部分理論之后才討論. 有些讀者可能认为故意要他在接触任何应用之前先負担一大批"非本质"(从数值的观点来看)的定理和引理. 为了緩和这些負担以及說明給出这些理論的动机,在下一节将通过从 Dirichlet 問題的数值解法所引出的一个特殊的簡单例子,来指出非負矩陣怎样自然地出現. 最后,第一章的其余部分将討論矩陣数值分析的某些基本概念和結果.

由于篇幅关系,对若干重要的有关問題只能作簡要的叙述. 舍入誤差的影响的分析,以及当网格步长趋于零时綫性方程組的离散解向有关偏微分方程的連續解收斂的問題,一般都要求使用十分不同于矩陣迭代分析中所用的数学工具. 对于这些問題,我們在这一章的书目与討論中列出了重要的参考文献.

1.2 一个簡单的例子

現在考慮单位正方形上的 Dirichlet 問題的数值解法,即寻找定义于閉单位正方形并在正方形內部滿足 Laplace 方程

$$(1.1)\qquad \frac{\partial^2 u(x,\,y)}{\partial x^2}+\frac{\partial^2 u(x,\,y)}{\partial y^2}=u_{xx}(x,\,y)+u_{yy}(x,\,y)=0,\qquad 0<x,\,y<1$$

的函数 $u(x,\,y)$ 的近似. 設 Γ 表示正方形的边界,則 $u(x,\,y)$ 除了

应满足微分方程(1.1)之外，还应满足 Dirichlet 边界条件

(1.2) $$u(x, y)=g(x, y),\quad (x, y)\in\Gamma,$$

其中 $g(x, y)$ 是定义于 Γ 上的某一特定函数. 在这单位正方形上，用边长 $h=1/3$ 作出均匀的正方形网格，如图 1 所示，用适当的足标把由水平和垂直綫段相交而得的内部和边界上的交点(网格节点)編号. 我們只找寻函数 $u(x, y)$ 在单位正方形的内部节点上的近似值，来代替找寻对所有 $0<x, y<1$ 都适合(1.1)以及边界条件(1.2)的函数 $u(x, y)$. 求 $u(x, y)$ 的这样的近似虽然有許多方法 (第六章)，但一个簡单的方法是把 $u(x, y)$ 分别对两个变量展为 Taylor 級数. 假定 $u(x, y)$ 是足够次可微的，则

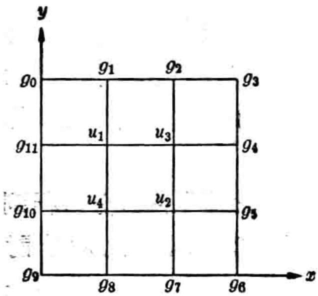

图 1

(1.3) $$u(x_0\pm h, y_0)=u(x_0, y_0)\pm hu_x(x_0, y_0)+\frac{h^2}{2}u_{xx}(x_0, y_0)\pm\frac{h^3}{3!}u_{xxx}(x_0, y_0)+\frac{h^4}{4!}u_{xxxx}(x_0, y_0)\pm\cdots,$$

(1.3′) $$u(x_0, y_0\pm h)=u(x_0, y_0)\pm hu_y(x_0, y_0)+\frac{h^2}{2}u_{yy}(x_0, y_0)\pm\frac{h^3}{3!}u_{yyy}(x_0, y_0)+\frac{h^4}{4!}u_{yyyy}(x_0, y_0)\pm\cdots,$$

其中点 (x_0, y_0) 及其四个邻点 $(x_0\pm h, y_0)$，$(x_0, y_0\pm h)$ 都是閉单位正方形上的点. 于是得

(1.4) $$\begin{aligned}&\frac{1}{h^2}\{u(x_0+h, y_0)+u(x_0-h, y_0)+u(x_0, y_0+h)\\&\quad+u(x_0, y_0-h)-4u(x_0, y_0)\}\\&=\{u_{xx}(x_0, y_0)+u_{yy}(x_0, y_0)\}\\&\quad+\frac{h^2}{12}\{u_{xxxx}(x_0, y_0)+u_{yyyy}(x_0, y_0)\}+\cdots.\end{aligned}$$

从(1.1)知(1.4)的右方第一项为零，又如略去有系数 h^2 和更高阶

的項,则近似地有

$$(1.4') \qquad u(x_0, y_0) \doteq \frac{1}{4}\{u(x_0+h, y_0)+u(x_0-h, y_0) + u(x_0, y_0+h)+u(x_0, y_0-h)\}.$$

設

$$u_1 \equiv u\left(\frac{1}{3}, \frac{2}{3}\right), \quad u_2 \equiv u\left(\frac{2}{3}, \frac{1}{3}\right),$$

$$u_3 \equiv u\left(\frac{2}{3}, \frac{2}{3}\right) \quad 及 \quad u_4 \equiv u\left(\frac{1}{3}, \frac{1}{3}\right),$$

又类似地, g_9 为給定函数 $g(x, y)$ 在原点 $x=y=0$ 的值,等等,現在用(1.4′)来分别定义 $u_i (1\leqslant i\leqslant 4)$ 的近似值 w_i:

$$(1.5) \qquad \begin{aligned} w_1 &= \frac{1}{4}(w_3+w_4+g_1+g_{11}), \\ w_2 &= \frac{1}{4}(w_3+w_4+g_5+g_7), \\ w_3 &= \frac{1}{4}(w_1+w_2+g_2+g_4), \\ w_4 &= \frac{1}{4}(w_1+w_2+g_8+g_{10}), \end{aligned}$$

这便是四个未知数 w_i 的四个綫性方程,每一个方程把未知函数 $u(x, y)$ 在一个网格内点上的近似值表示为 $u(x, y)$ 在相邻节点上的近似值的平均值. 用矩陣記号,(1.5)能写成

$$(1.5') \qquad A\mathbf{w}=\mathbf{k},$$

其中

$$(1.6) \qquad A=\begin{bmatrix} 1 & 0 & -1/4 & -1/4 \\ 0 & 1 & -1/4 & -1/4 \\ -1/4 & -1/4 & 1 & 0 \\ -1/4 & -1/4 & 0 & 1 \end{bmatrix},$$

$$\mathbf{w}=\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} \quad 及 \quad \mathbf{k}=\frac{1}{4}\begin{bmatrix} g_1+g_{11} \\ g_5+g_7 \\ g_2+g_4 \\ g_8+g_{10} \end{bmatrix}.$$

此处 $\mathbf{k}$ 是一个向量,它的分量能从已知边界值 g_i 計算出来. 显然,

矩陣 A 能写成 $I-B$，其中

(1.7) $$B=\frac{1}{4}\begin{bmatrix}0&0&1&1\\0&0&1&1\\1&1&0&0\\1&1&0&0\end{bmatrix}.$$

显然，A 和 B 都是实对称矩陣，又矩陣 B 的元素都是非負实数. 矩陣 B 的特征多項式为

(1.8) $$\phi(\mu)=\det(\mu I-B)=\mu^2(\mu^2-1/4),$$

所以 B 的特征值为 $\mu_1=-1/2$，$\mu_2=0=\mu_3$ 及 $\mu_4=1/2$，于是

$$\max_{1\leqslant i\leqslant 4}|\mu_i|=1/2.$$

因 A 的特征值 ν_i 是 $1-\mu_i$ 的形式，故 A 的特征值显然为正实数，从而推知 A 为一实对称正定矩陣. 因 A 非奇，它的逆 A^{-1} 是唯一确定的，并可明显地表为

(1.9) $$A^{-1}=\frac{1}{6}\begin{bmatrix}7&1&2&2\\1&7&2&2\\2&2&7&1\\2&2&1&7\end{bmatrix},$$

于是 A^{-1} 的元素都是正实数. 我們将在后面看到(在第六章)，象矩陣 B 的特征值的模小于 1 以及矩陣 A 只有正实元素等这些簡单的性质，对于从自共軛二阶椭圓型偏微分方程导出的矩陣方程，一般都成立.

因为能把矩陣方程(1.5′)等价地写为

(1.10) $$\mathbf{w}=B\mathbf{w}+\mathbf{k},$$

所以对这个簡单的問題，可引出第一个(循环)迭代法，称为点 *Jacobi* 或点全步法❶. 設 $\mathbf{w}^{(0)}$ 为(1.5′)的唯一(因 A 非奇)解向量 $\mathbf{w}$ 的任意实或复的近似向量，則从

(1.11) $$\mathbf{w}^{(m+1)}=B\mathbf{w}^{(m)}+\mathbf{k},\quad m\geqslant 0$$

逐次定义一迭代向量 $\mathbf{w}^{(m)}$ 的序列. 要問的第一个問題是关于(1.11)的收斂性，即是否每一 $\lim\limits_{m\to\infty} w_j^{(m)}$ 存在，又假定这些极限存在

❶ 这个迭代法还有其他名称，見 3.1 节.

时，是否每一极限等于方程的解的第 j 分量 w_j? 为了回答这些問題，令

$$\boldsymbol{\varepsilon}^{(m)} \equiv \mathbf{w}^{(m)} - \mathbf{w}, \quad m \geqslant 0,$$

其中 $\boldsymbol{\varepsilon}^{(m)}$ 是对应于迭代向量 $\mathbf{w}^{(m)}$ 的誤差向量. 从 (1.11) 减去 (1.10)，得

$$\boldsymbol{\varepsilon}^{(m+1)} = B\boldsymbol{\varepsilon}^{(m)},$$

由此用归納法得

(1.12) $$\boldsymbol{\varepsilon}^{(m)} = B^m\boldsymbol{\varepsilon}^{(0)}, \quad m \geqslant 0.$$

对第 j 分量，显然，当且仅当 $\lim\limits_{m\to\infty} w_j^{(m)}$ 存在时 $\lim\limits_{m\to\infty} \varepsilon_j^{(m)}$ 存在，又如这些极限都存在，則当且仅当 $\lim\limits_{m\to\infty} \varepsilon_j^{(m)} = 0$ 时 $\lim\limits_{m\to\infty} w_j^{(m)} = w_j$. 所以，按 (1.12)，如果希望誤差向量的每一分量在极限情形下消失，就要找寻一些条件，这些条件对所有向量 $\boldsymbol{\varepsilon}^{(0)}$，保証

(1.13) $$\lim_{m\to\infty} B^m\boldsymbol{\varepsilon}^{(0)} = 0.$$

但找寻保証 (1.13) 成立的条件等价于决定什么时候

(1.14) $$\lim_{m\to\infty} B^m = O$$

成立，其中 O 是 $n\times n$ 零矩陣. 这将于下节中討論.

1.3 模和譜半徑

向量模，矩陣模[1]，以及矩陣的譜半徑的概念在迭代的数值分析中起着重要的作用. 象用长度来比較两个向量一样，用某些度量或模来比較两个矩陣也是方便的. 我們将看到，在某些确切的意义下，这将是决定两个迭代法中哪一个能較快地收斂的基础.

首先，令 $V_n(C)$ 为复数域 C 上的列向量 $\mathbf{x}$ 的 n 維向量空間，这里向量 $\mathbf{x}$，它的轉置 $\mathbf{x}^T$ 以及它的共軛轉置 $\mathbf{x}^*$ 分别記为

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{x}^T = [x_1 \ x_2 \ \cdots \ x_n], \quad \mathbf{x}^* = [\bar{x}_1 \ \bar{x}_2 \ \cdots \ \bar{x}_n],$$

[1] 向量模和矩陣模也称为向量范数和矩陣范数. ——譯者注

其中 $x_1, x_2, \cdots, x_n$ 为复数，又 $\bar{x}_i$ 是 x_i 的共軛复数.

定义 1.1　令 $\mathbf{x}$ 为 $V_n(C)$ 的(列)向量．则

$$\|\mathbf{x}\| \equiv (\mathbf{x}^*\mathbf{x})^{1/2} = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \tag{1.15}$$

称为 $\mathbf{x}$ 的欧氏模(或长).

用这个定义，下列結果是熟知的.

定理 1.1　設 $\mathbf{x}$ 和 $\mathbf{y}$ 为 $V_n(C)$ 的向量，則

$$\begin{aligned} &\|\mathbf{x}\| > 0,\ \text{除非}\ \mathbf{x} = \mathbf{0}; \\ &\text{設}\ \alpha\ \text{为純量，則}\ \|\alpha\mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\|; \\ &\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|. \end{aligned} \tag{1.16}$$

設有 $V_n(C)$ 的无穷向量序列 $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots$，如果

$$\lim_{m\to\infty} x_j^{(m)} = x_j, \quad \text{对所有}\ 1 \leqslant j \leqslant n\ \text{成立},$$

其中 $x_j^{(m)}$ 和 x_j 分别为向量 $\mathbf{x}^{(m)}$ 和 $\mathbf{x}$ 的第 j 分量，則称这向量序列收斂于 $V_n(C)$ 的向量 $\mathbf{x}$．类似地，$V_n(C)$ 的无穷向量級数 $\sum_{m=0}^{\infty} \mathbf{y}^{(m)}$ 收斂于 $V_n(C)$ 的向量 $\mathbf{y}$ 是指

$$\lim_{N\to\infty} \sum_{m=0}^{N} y_j^{(m)} = y_j, \quad \text{对所有}\ 1 \leqslant j \leqslant n\ \text{成立}.$$

用欧氏模，則从定义 1.1 得

$$\|\mathbf{x}^{(m)} - \mathbf{x}\| \to 0, \quad m \to \infty$$

的充要条件是向量序列 $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \cdots$ 收斂于向量 $\mathbf{x}$；又类似地，

$$\left\| \sum_{m=0}^{N} \mathbf{y}^{(m)} - \mathbf{y} \right\| \to 0, \quad N \to \infty$$

的充要条件是无穷級数 $\sum_{m=0}^{\infty} \mathbf{y}^{(m)}$ 收斂于向量 $\mathbf{y}$.

在以后討論中将反复用到的另一基本定义为

定义 1.2　令 $A = (a_{i,j})$ 为 $n \times n$ 复矩陣，具有特征值 λ_i, $1 \leqslant i \leqslant n$．則称

$$\rho(A) \equiv \max_{1 \leqslant i \leqslant n} |\lambda_i| \tag{1.17}$$

为矩陣 A 的譜半徑.

在几何上，如果 A 的所有特征值 λ_i 都画在复 z-平面上，則

$\rho(A)$ 是中心在原点并包含矩陣 A 的所有特征值的最小圓❶ $|z|\leqslant R$ 的半徑.

現在,对每一个有复数元素的 $n\times n$ 矩陣 A,賦予一个象向量模 $\|\mathbf{x}\|$ 一样、具有类似于长度性质(1.16)的非負实数.

定义 1.3 設 $A=(a_{i,j})$ 为 $n\times n$ 复矩陣,則称

$$\|A\|=\sup_{\mathbf{x}\neq\mathbf{0}}\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \tag{1.18}$$

为矩陣 A 的譜模.

象从向量 $\mathbf{x}$ 的欧氏模所得到的性质一样,矩陣的譜模的基本性质如下:

定理 1.2 設 A 和 B 是两个 $n\times n$ 矩陣,則

$$\begin{aligned}&\|A\|>0,\quad 除非\ A\equiv O\ (零矩陣);\\&若\ \alpha\ 为純量,則\ \|\alpha A\|=|\alpha|\cdot\|A\|;\\&\|A+B\|\leqslant\|A\|+\|B\|;\\&\|A\cdot B\|\leqslant\|A\|\cdot\|B\|.\end{aligned} \tag{1.19}$$

此外,对所有向量 $\mathbf{x}$,成立着

$$\|A\mathbf{x}\|\leqslant\|A\|\cdot\|\mathbf{x}\|, \tag{1.20}$$

又在 $V_n(C)$ 中存在一非零向量 $\mathbf{y}$,使

$$\|A\mathbf{y}\|=\|A\|\cdot\|\mathbf{y}\|. \tag{1.21}$$

【証】 (1.19)和(1.20)的結果可直接从定理 1.1 和定义 1.3 推得. 为了建立(1.21),首先看到当 $\mathbf{x}$ 换为 $\alpha\mathbf{x}$ (α 为純量)时,$\|A\mathbf{x}\|/\|\mathbf{x}\|$ 不改变. 因此可写

$$\|A\|=\sup_{\|\mathbf{x}\|=1}\|A\mathbf{x}\|.$$

但适合 $\|\mathbf{x}\|=1$ 的所有向量 $\mathbf{x}$ 的集合是紧致的,又 $A\mathbf{x}$ 是定义在这集合上的連續函数,故存在一向量 $\mathbf{y}$,$\|\mathbf{y}\|=1$,使

$$\|A\|=\sup_{\|\mathbf{x}\|=1}\|A\mathbf{x}\|=\|A\mathbf{y}\|,$$

这便完成了証明.

---

❶ 精确地說,$|z-a|\leqslant R$ 的所有点的集合称为圓(或圓盘),而它的由 $|z-a|=R$ 定义的子集称为圓周.

联系定义 1.2 和 1.3 的譜半徑和譜模，我們有

推論 对任意 $n\times n$ 复矩陣 A，都有

$$\|A\|\geqslant\rho(A). \tag{1.22}$$

【証】 設 λ 为 A 的任一特征值，$\mathbf{x}$ 为对应于特征值 λ 的特征向量，則 $A\mathbf{x}=\lambda\mathbf{x}$. 于是从定理 1.1 和 1.2 得

$$|\lambda|\cdot\|\mathbf{x}\|=\|\lambda\mathbf{x}\|=\|A\mathbf{x}\|\leqslant\|A\|\cdot\|\mathbf{x}\|,$$

由此可得 $\|A\|\geqslant|\lambda|$ 对 A 的所有特征值都成立，这便証明了 (1.22).

用 Faddeeva (1959) 和 Householder (1958) 的术語，(1.20) 是說矩陣的譜模相容于向量的欧氏模，而 (1.21) 是說矩陣的譜模从屬于向量的欧氏模. 适合定理 1.1 和 1.2 的性质的向量模和矩陣模还可用很多其他方法来定义. 虽然在本节练习中給出了向量模和矩陣模的其他定义，但这里仅集中討論向量的欧氏模和矩陣的譜模，因为这两种模在以后的討論中一般来說已經够用.

矩陣的譜模也能用譜半徑来表示. 若以

$$A=\begin{bmatrix}a_{1,1} & a_{1,2} & \cdots & a_{1,n}\\ a_{2,1} & a_{2,2} & \cdots & a_{2,n}\\ \vdots & & & \vdots\\ a_{n,1} & a_{n,2} & \cdots & a_{n,n}\end{bmatrix},\qquad A^T=\begin{bmatrix}a_{1,1} & a_{2,1} & \cdots & a_{n,1}\\ a_{1,2} & a_{2,2} & \cdots & a_{n,2}\\ \vdots & & & \vdots\\ a_{1,n} & a_{2,n} & \cdots & a_{n,n}\end{bmatrix},$$

$$A^*=\begin{bmatrix}\bar{a}_{1,1} & \bar{a}_{2,1} & \cdots & \bar{a}_{n,1}\\ \bar{a}_{1,2} & \bar{a}_{2,2} & \cdots & \bar{a}_{n,2}\\ \vdots & & & \vdots\\ \bar{a}_{1,n} & \bar{a}_{2,n} & \cdots & \bar{a}_{n,n}\end{bmatrix}$$

分别表示具有复数元素 $a_{i,j}$ 的 $n\times n$ 矩陣 A，它的轉置 A^T 以及它的共軛轉置 A^*，則矩陣的乘积 A^*A 也是 $n\times n$ 矩陣.

定理 1.3 設 $A=(a_{i,j})$ 为 $n\times n$ 复矩陣，則

$$\|A\|=[\rho(A^*A)]^{1/2}. \tag{1.23}$$

【証】 矩陣 A^*A 是非負定 Hermite 陣，即对任一向量 $\mathbf{x}$，有

$$(A^*A)^*=A^*A \quad 以及 \quad \mathbf{x}^*A^*A\mathbf{x}=\|A\mathbf{x}\|^2\geqslant 0.$$

因 A^*A 为 Hermite 陣，令 $\{\boldsymbol{\alpha}_i\}_{i=1}^n$ 为 A^*A 的正交特征向量組，即 $A^*A\boldsymbol{\alpha}_i=\nu_i\boldsymbol{\alpha}_i$，其中 $0\leqslant\nu_1\leqslant\nu_2\leqslant\cdots\leqslant\nu_n$，对所有 $i\neq j$ 有 $\boldsymbol{\alpha}_i^*\boldsymbol{\alpha}_j=0$,

对所有 $1\leqslant i\leqslant n$ 有 $\boldsymbol{\alpha}_i^*\boldsymbol{\alpha}_i=1$. 設

$$\mathbf{x}=\sum_{i=1}^{n} c_i\boldsymbol{\alpha}_i$$

是任一非零向量，則由直接計算得

$$\left(\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}\right)^2=\frac{\mathbf{x}^*A^*A\mathbf{x}}{\mathbf{x}^*\mathbf{x}}=\frac{\sum_{i=1}^{n}|c_i|^2\nu_i}{\sum_{j=1}^{n}|c_j|^2},$$

所以

$$0\leqslant\nu_1\leqslant\left(\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}\right)^2\leqslant\nu_n.$$

此外，选 $\mathbf{x}=\boldsymbol{\alpha}_n$ 便使上式右方的等号成立. 于是由定义 1.3 得

$$\|A\|^2=\sup_{\mathbf{x}\neq 0}\left(\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}\right)^2=\nu_n=\rho(A^*A),$$

这便完成了証明.

推論 設 A 为 $n\times n$ Hermite 陣，則

$$\|A\|=\rho(A). \tag{1.24}$$

此外，設 $g_m(x)$ 为 x 的任一 m 次实多項式，則

$$\|g_m(A)\|=\rho(g_m(A)). \tag{1.24′}$$

【証】 設 A 为 Hermite 陣，則由定义有 $A^*=A$. 于是

$$\|A\|^2=\rho(A^*A)=\rho(A^2)=\rho^2(A),$$

便証明了 $\|A\|=\rho(A)$. 若 $g_m(x)$ 是变量 x 的实多項式，則容易驗証矩陣 $g_m(A)$ 也是 Hermite 陣，于是卽得一般情形的結果，这便証明了 (1.24′).

特别是，我們現在知道 Hermite 陣的譜模和譜半徑必相等. 这对一般矩陣并不成立. 为了指出在 (1.22) 中等号不能普遍成立，只要考虑下面的簡单矩陣就够了：

$$A=\begin{pmatrix}\alpha & 1\\ 0 & \alpha\end{pmatrix},$$

其中 α 为任一复数. 因 A 为三角形矩陣并且对角元素等于 α, 故它的特征値等于 α, 从而 $\rho(A)=|\alpha|$. 另一方面，由直接計算可得

$$A^*A=\begin{bmatrix}|\alpha|^2 & \bar{\alpha}\\ \alpha & |\alpha|^2+1\end{bmatrix},$$

它的特征值是 $|\alpha|^2+\frac{1}{2}\pm\sqrt{|\alpha|^2+\frac{1}{4}}$，所以

$$\|A\|=\left\{|\alpha|^2+\frac{1}{2}+\sqrt{|\alpha|^2+\frac{1}{4}}\right\}^{1/2}>|\alpha|=\rho(A).$$

仿照向量序列和級数的收斂性定义，可以类似地定义矩陣序列和級数的收斂性．設 $A^{(0)}=(a_{i,j}^{(0)})$，$A^{(1)}=(a_{i,j}^{(1)})$，$\cdots$ 为 $n\times n$ 复矩陣的无穷序列，如果

$$\lim_{m\to\infty}a_{i,j}^{(m)}=a_{i,j},\quad \text{对所有 } 1\leqslant i,\ j\leqslant n \text{ 成立},$$

則称这序列收斂于 $n\times n$ 复矩陣 $A=(a_{i,j})$．类似地，$n\times n$ 复矩陣 $B^{(m)}=(b_{i,j}^{(m)})$ 的无穷級数 $\sum\limits_{m=0}^{\infty}B^{(m)}$ 收斂于矩陣 $B=(b_{i,j})$ 是指

$$\lim_{N\to\infty}\sum_{m=0}^{N}b_{i,j}^{(m)}=b_{i,j},\ \text{对所有 } 1\leqslant i,\ j\leqslant n \text{ 成立}.$$

用譜模，从定理 1.2 推得

$$\|A^{(m)}-A\|\to 0,\quad m\to\infty$$

的充要条件是矩陣序列 $A^{(0)}$, $A^{(1)}$, $\cdots$ 收斂于矩陣 A；又类似地，

$$\left\|\sum_{m=0}^{N}B^{(m)}-B\right\|\to 0,\quad N\to\infty$$

的充要条件是无穷級数 $\sum\limits_{m=0}^{\infty}B^{(m)}$ 收斂于矩陣 B.

現在回到上节所提出的問題．設 A 为 $n\times n$ 复矩陣，什么时候矩陣 A 的乘幂序列 A, A^2, A^3, $\cdots$ 收斂于零矩陣 O？由于这种形式的序列是我們以后討論的基础，現在作出

定义 1.4　令 A 为 $n\times n$ 复矩陣．設矩陣的序列 A, A^2, A^3, $\cdots$ 收斂于零矩陣 O，則称矩陣 A 为收斂（于零），否則称为发散．

定理 1.4　設 A 为 $n\times n$ 复矩陣，則 A 收斂的充要条件是 $\rho(A)<1$.

【証】 对于給定的矩陣 A，存在非奇 $n\times n$ 矩陣 S，化矩陣 A