数值线代数讲义

孙 瓛 著

南开大学出版社

数值线代数讲义

孙 瓛 编

南开大学出版社

1987年

内 容 提 要

本书在回顾和补充必要的线代数基本知识之后，介绍线代数数值计算的常用计算方法。包括线代数方程组的各类求解方法：直接解法，迭代解法和变分解法；代数特征值问题的各种实用的计算方法：乘幂法、反乘幂法、子空间迭代法，Jacobi方法、Givens-Householder方法、QR方法。并对这些计算方法做出必要的数值分析。

本书是根据原教育部颁发的大纲编写的教材。可供综合大学或工科高等院校使用。编者力图使本书方便于实际部门的读者和自学者阅读或参考。阅读本书的基础是数学分析和线性代数的基础知识。

数值线代数讲义

孙 瓛 编

南开大学出版社出版

（天津八里台南开大学校内）

新华书店天津发行所发行

天津市印刷晒图厂印刷

1987年3月第1版　1987年3月第1次印刷

开本：$850\times1168\ 1/32$　印张：14.75

字数：368千　印数：1-2，600

统一书号：13301·24　定价：3.00元

前　言

数值线代数是研究线代数问题的数值计算方法及其有关理论（诸如各种方法的稳定性、收敛性及误差估计等）的一门学科。它研究的最基本的问题有两个：一是非奇异线代数方程组的求解问题（包括矩阵求逆和行列式求值问题）。也就是说，对于线代数方程组$Ax=b$，在已知系数阵A（非奇异）和右端向量b的前提下，求未知向量x，使其满足方程组$Ax=b$。一般地，我们只讨论A、b的元素均为实数的情形。二是代数特征问题。即已知方阵A，求A的全部或部分特征值，以及在需要时求其相应的全部或部分特征向量。由于矩阵的特征值λ和它相应的特征向量x满足齐次方程组（$\lambda I-A$）$x=0$，所以代数特征问题也称为齐次问题。而线代数方程组的求解问题，相应地称为非齐次问题。

随着计算机科学的迅速发展，到目前，数值计算几乎在所有科技领域中，已经有并且越来越有更加广泛的应用。而许多数值计算问题，特别是来源于工程、物理问题的微分方程的数值解问题及来自许多最优化问题的有效的数值计算方法，最终都将归结为数值线代数计算问题。因此，数值线代数在现代计算数学中就占有特殊重要的基础的地位。

本书的主要内容是介绍非奇异线代数方程组的求解问题（第二、三、四章）和标准特征问题（第五章）的常用计算方法及对这些方法的简单的理论分析。对广义特征问题在最后一章中亦做简单的介绍。做为必要的数学准备，我们在开始的一章中回顾并

补充一些为本书内容直接需要的线代数基本知识。在每章之后附有练习，同时将练习的答案或提示附于本书的最后。为了便于读者做局部的查阅和进一步的深入学习，本书列出全书的符号索引和补充图书书目。

本书的目的是介绍最常用的一些数值线代数计算方法，同时对这些方法做简单而基本的理论分析，以指出这些方法的适用范围和使用效率。

本书可作为综合性大学或工科院校《数值线代数》课程的基本教材。也可以作为学习数值线代数方法的读者的自学材料。对于具备一些代数基本知识的读者，读懂本书并无太大的困难。对于只希望了解和使用有关方法的读者，完全可以略去理论分析的过程，有选择地阅读所需要的有关部分内容。作者力求使本书对从事实际工作的计算工作者和希望了解或使用其中某些计算方法的工程技术人员及科学工作者有用。

具体地说，本书的内容编排为三大部分。第一部分是线代数基本知识的回顾与补充（第一章）。主要包括线性变换与矩阵、矩阵特征值及特征向量的基本定义和结论；在计算方法的使用中不可少的各种变换矩阵——初等变换阵、置换阵、消元矩阵及平面旋转阵、镜像变换阵这两种重要的酉矩阵的基本性质和作用；向量、矩阵的极限及范数，它是对各种计算方法的功效进行理论分析的必不可少的工具；矩阵特征值的估计，特别是Hermite阵特征值做为Rayleigh商的极值的重要结论。它在代数特征问题特别是对称阵的代数特征问题的各种计算方法的理论分析中起重要作用；本章的最后部分对矩阵的奇异值分解给出基本的结论。第二部分是解线代数方程组的常用计算方法及简单的理论分析。主要包括：解线代数方程组的直接方法——Gauss（高斯）消元法、LU分解〔直接三角分解的Doolittle（杜利特尔）算法、Crout（克劳特）算法、Cholesky（乔勒斯基）分解法（平方

根法)〕、最小二乘法及简单的误差分析。对于大型稀疏线代数方程组的解法及稀疏技术做了简单的介绍（第二章）。解线代数方程组的迭代方法——Jacobi（雅可比）迭代法、Gauss—Seidel（高斯—赛德尔）迭代法、松弛迭代法及其收敛性讨论（第三章）。解线代数方程组的变分方法——最速下降法、共轭斜量法、最小二乘法及收敛性讨论（第四章）。第三部分是矩阵特征问题的常用计算方法及简单的理论分析。主要包括：求矩阵按模最大的特征值及相应特征向量的乘幂法、求矩阵按模最小的特征值及相应特征向量或求矩阵的与其给定值最接近的特征值及相应特征向量的反幂法（逆迭代）、求对称矩阵按模大的若干（少数）个特征值及相应特征向量的子空间迭代法、求对称矩阵全部特征值及相应的特征向量的Jacobi（雅可比）方法、解对称矩阵的全部或部分特征值及相应特征向量的Givens—Householder（吉文斯—豪斯赫尔德）方法、求对称矩阵部分或全部特征值及相应特征向量的带位移QR算法与求非对称阵相应问题的双重步QR算法。最后对广义特征问题及奇异值计算问题做简单的介绍（第五章）。

阅读本书的必要基础是熟悉一般高等院校的线代数教程的基本内容。本书中未列入或未证明而用到的线代数的全部结果，几乎可在任何一本线代数教科书（例如，北京大学编的《高等代数》）中找到。

常用符号表

$\mathbf{R}^n$（$\mathbf{C}^n$）	n维实（复）向量空间
$\mathbf{R}^{m\times n}$（$\mathbf{C}^{m\times n}$）	$m\times n$阶实(复）矩阵的全体构成的线性空间
$A=(a_{ij})_{m\times n}$	以a_{ij}为其第i行j列元素的$m\times n$矩阵
$\mathbf{x}=(x_1, \cdots, x_n)^T$	以x_i为其第i个分量的n维列向量
A^T	A的转置矩阵
A^*	A的共轭转置阵
$\|A\|$	以A的(i, j)元素的模$\|a_{ij}\|$为（i, j）元的（与A同阶的）矩阵
det（A）	方阵A的行列式
rank（A）	矩阵A的秩
trace（A）	方阵A的迹
$\rho(A)$	方阵A的谱半径
e_i	第i个自然基向量
$\\|A\\|$	矩阵A的范数
$\\|\mathbf{x}\\|$	向量x的范数
$\{\mathbf{x}^{(k)}\}$	向量列$\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(k)}, \cdots$

$\{\mathbf{x}^{(k)}\}_{k=1}^n$ 或$\{\mathbf{x}^{(1)}, \cdots \mathbf{x}^{(n)}\}$　由$\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(n)}$构成的向量组

$(\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(n)})$	依次以$\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(n)}$为其列向量的矩阵
$\equiv$	全等于或定义为
$\mathbf{N}(A)$	A的零空间
$\mathbf{R}(A)$	A的像空间

$R^{\perp}(A)$	像空间的正交补空间
cond(A)	矩阵A关于解线性方程组的条件数
diag(λ_1, …, λ_n)	依次以λ_1, …, λ_n为对角元的对角形矩阵
ϕ	空集
$\longleftrightarrow$	充分必要条件或一一对应

目　　录

第一部分　线代数基本知识的回顾与补充

第一章　线代数基本知识的回顾与补充

§1　引　　言

本节介绍本章常用的记号和定义，并引入矩阵分块运算的结果。

用大写英文字母A、B、X、Y等表示矩阵。相应的小写字母附以下标a_{ij}、b_{ij}、x_{ij}、y_{ij}依次表示矩阵A、B、X、Y的第i行和第j列交叉位置上的元素。今后简称为（i，j）元素。如果A为m行n列矩阵，就记$A=(a_{ij})_{m\times n}$或$A_{m\times n}$。$\mathbf{C}^{m\times n}$（$\mathbf{R}^{m\times n}$）表示复（实）数域上$m\times n$矩阵的全体所构成的集合，它是线性空间。特别地，当$m=n$时$\mathbf{C}^{n\times n}$（$\mathbf{R}^{n\times n}$）是复（实）数域上n阶方阵全体的集合。$A\in\mathbf{C}^{n\times n}$（$A\in\mathbf{R}^{n\times n}$）表示$A=(a_{ij})_{n\times n}$是一个元素$a_{ij}$（$i, j=1, \cdots, n$）均为复数（实数）的n阶方阵。

记$\mathbf{C}^n\equiv\mathbf{C}^{n\times 1}$（$\mathbf{R}^n\equiv\mathbf{R}^{n\times 1}$）为复（实）数域上n维列向量的全体构成的集合——复（实）n维向量空间。以小写字母a、b、x、y及带括号的上标的$a^{(k)}$、$b^{(k)}$等表示列向量。在需要表明它的维数时可记$x\in\mathbf{C}^n$（$\mathbf{R}^n$），用以表明x是n维复（实）的列向量，以相应带下标的小写字母a_i、b_i …依次表示向量a、b、…的第i个分量。用$\|x\|$表示向量的长度，以区别其第i个分量为

$|x_i|$的列向量$|x|$。由n个向量$x^{(1)}, \cdots, x^{(n)}$构成的向量组常记为$\{x^{(k)}\}_{k=1}^{n}$或$\{x^{(1)}, \cdots, x^{(n)}\}$。而记以这n个向量为列构成的矩阵为$X=(x^{(1)}, x^{(2)}, \cdots, x^{(n)})$。除向量矩阵的元素外，一般以小写希腊字母$\xi$、$\eta$、$\lambda$、$\theta$等表示数。

以I_n表示$\mathbf{C}^{n\times n}$（$\mathbf{R}^{n\times n}$）中的单位阵，其元素为

$$\delta_{ij}=\begin{cases}0, & 如果i\neq j\\ 1, & 如果i=j\end{cases} \quad i, j=1, 2, \cdots, n。$$

当不强调单位阵的阶数或对单位阵阶数的认识不会发生混淆时，将记它为I，而省略下标n。e_i表示I的第i个列向量，即除第i个分量为1，其余分量均为零的向量。$\{e_i\}_{i=1}^{n}$构成$\mathbf{C}^n$（$\mathbf{R}^n$）的一组基，称为$\mathbf{C}^n$（$\mathbf{R}^n$）的自然基底组。e_i就称为自然基底组的第i个基底向量。

注意，只对自然基底向量才用带下标的英文小写字母e_i表示，以与惯用的记号一致。

对于数零、零向量和零矩阵——元素全为零的矩阵，我们统记为0，只在必要表明其维数或阶数时才附加下标说明。例如，$0_{m\times n}$表示$m\times n$阶零矩阵。

A^T、A^*分别表示矩阵A的转置矩阵和共轭转置矩阵。即：如果$A=(a_{ij})_{m\times n}$则A^T和A^*均为$n\times m$阵，其(i, j)元素分别为a_{ji}和$\overline{a_{ji}}$，其中$\overline{a_{ji}}$是a_{ji}的共轭复数。特别地，n维列向量$\mathbf{x}$的转置和共轭转置向量分别记为x^T和x^*。它们是n维行向量。

今后我们用$|A|$表示以$|a_{ij}|$为(i, j)元素的矩阵。为区别起见，我们记矩阵A的行列式为det(A)。rank(A)表示矩阵A的秩。trace(A)表示方阵A的所有对角元素之和，称为A的迹。

对于矩阵的乘积运算，如下结果是熟知的。设$A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{n\times l}$，则

——$C\equiv A\cdot B=(c_{ij})_{m\times l}$，其中

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj},\quad i=1,\cdots,m;$$

$$j=1,\cdots,l;$$

——$(A\cdot B)^T=B^T\cdot A^T$；　$(A\cdot B)^*=B^*\cdot A^*$。

——当$m=n=l$时，如果A、B均可逆，则$A\cdot B$可逆，且$(A\cdot B)^{-1}=B^{-1}\cdot A^{-1}$。

——当$m=n$时，如果A可逆，则A^T也可逆且

$$(A^T)^{-1}=(A^{-1})^T。$$

将矩阵$A=(a_{ij})_{m\times n}$分块写为$A=(A_{pq})_{s\times t}$，如果记子阵A_{pq}的阶数为$m_p\times n_q$（$p=1,\cdots,s$；$q=1,\cdots,t$），显然，$\sum_{p=1}^{s}m_p=m$，$\sum_{q=1}^{t}n_q=n$。应该注意的是，对于方阵A，如果不加说明，则其分块表示式中总假定其对角块A_{pp}是方的。对分块表示的矩阵（今后称分块矩阵）的乘法，有形同一般矩阵一样的计算法：

设$A=(A_{pk})_{s\times t}$，$B=(B_{kq})_{t\times r}$，如果A_{pk}和B_{kq}对一切$1\leqslant p\leqslant s$，$1\leqslant k\leqslant t$，$1\leqslant q\leqslant r$是可乘的，则$AB\equiv C=(C_{pq})_{s\times r}$，其中子阵

$$C_{pq}=\sum_{k=1}^{t}A_{pk}B_{kq}\qquad p=1,\cdots,s;\ q=1,\cdots,r。$$

特别地，有矩阵和向量的分块乘法法则：

$$(Ax)_p=\sum_{q=1}^{t}A_{pq}\breve{x}_q;\ p=1,\cdots,r$$

其中$A=(A_{pq})_{r\times t}$，$x=\begin{pmatrix}\breve{x}_1\\ \vdots\\ \breve{x}_t\end{pmatrix}$且$A_{pq}$与$\breve{x}_q$对$1\leqslant q\leqslant t$可乘。

最后，我们回顾有关向量正交的如下概念：

——称向量$\mathbf{x}$与向量$\mathbf{y}$正交，如果$\mathbf{y}^*\mathbf{x}=0$。这时记$\mathbf{x}\perp\mathbf{y}$；

——称向量组$\{\mathbf{x}^{(k)}\}_{k=1}^{m}$是正交组，如果其中任意两个向量均正交。即$\mathbf{x}^{(j)*}\mathbf{x}^{(i)}=0$，$(i\neq j)$。特别地，如果$\mathbf{x}^{(j)*}\mathbf{x}^{(i)}=\delta_{ij}=\begin{cases}0, & 当i\neq j\\ 1, & 当i=j\end{cases}$则称$\{\mathbf{x}^{(k)}\}_{k=1}^{m}$为标准正交组。

——称向量$\mathbf{x}$与子空间$\mathbf{W}$正交。如果对任意$\mathbf{y}\in\mathbf{W}$，有$\mathbf{y}^*\mathbf{x}=0$。这时记$\mathbf{x}\perp\mathbf{W}$或$\mathbf{x}\in\mathbf{W}^{\perp}$其中$\mathbf{W}^{\perp}$表示$\mathbf{W}$的正交补空间。

§2 线性变换与矩阵

在一般线性代数教程中已经知道，线性变换$\mathbf{A}$作为$\mathbf{C}^n\to\mathbf{C}^n$的映射，它和矩阵$A\in\mathbf{C}^{n\times n}$之间有一个一一对应的关系，即$\{\mathbf{A}:\mathbf{C}^n\to\mathbf{C}^n\}$和$\{A:A\in\mathbf{C}^{n\times n}\}$之间存在着一一对应关系。如果给定$\mathbf{C}^n$的一组基$\mathbf{x}^{(1)}$，$\mathbf{x}^{(2)}$，…，$\mathbf{x}^{(n)}$则这个一一对应关系就表示为

$$\mathbf{A}(\mathbf{x}^{(1)},\cdots,\mathbf{x}^{(n)})=(\mathbf{x}^{(1)},\cdots,\mathbf{x}^{(n)})A$$

或者

$$\begin{cases}\mathbf{A}\mathbf{x}^{(1)}=a_{11}\mathbf{x}^{(1)}+a_{21}\mathbf{x}^{(2)}+\cdots+a_{n1}\mathbf{x}^{(n)}\\ \mathbf{A}\mathbf{x}^{(2)}=a_{12}\mathbf{x}^{(1)}+a_{22}\mathbf{x}^{(2)}+\cdots+a_{n2}\mathbf{x}^{(n)}\\ \quad\cdots\\ \mathbf{A}\mathbf{x}^{(n)}=a_{1n}\mathbf{x}^{(1)}+a_{2n}\mathbf{x}^{(2)}+\cdots+a_{nn}\mathbf{x}^{(n)}\end{cases}$$

称$A=(a_{ij})_{n\times n}$为线性变换$\mathbf{A}$在基组$\mathbf{x}^{(1)}$，…，$\mathbf{x}^{(n)}$下对应的矩阵。特别地，简称线性变换在自然基底组e_1，…e_n下对应的矩阵为线性变换对应的矩阵，而略去“在自然基底组下”的字样。

对于任一向量$\mathbf{x}$及它在线性变换$\mathbf{A}$下的“像”$\mathbf{y}$的坐标间的关系，有如下结论：

对于$\mathbf{A}x = y$，如果$x = \sum_{i=1}^{n} x_i e_i$，$y = \sum_{i=1}^{n} y_i e_i$

则

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \qquad i = 1, 2, \cdots, n$$

写成向量的形式为：

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$，其中A为$\mathbf{A}$对应的矩阵。

这也可以说，给定了矩阵A，则由

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \qquad i = 1, 2, \cdots, n$$

唯一地确定了$\mathbf{C}^n$上的一个线性变换$\mathbf{A}$：$\mathbf{A}x = y$。

通常，如果不加说明地给出向量x的坐标表示$x = (x_1, x_2, \cdots, x_n)^T$则总假定是在自然基底下讨论问题的。

最后我们指出，如果设$y = Ax$确定了线性变换$\mathbf{A}$，其中$x = (x_1, x_2, \cdots, x_n)^T$，$y = (y_1, \cdots, y_n)^T$，那么，$\mathbf{A}$在另一组基底$\{\overset{\frown}{e}^{(i)}\}_{i=1}^{n}$下的矩阵$\overset{\frown}{A}$和A有何关系。或者说$x = \sum_{i=1}^{n} \overset{\frown}{x}_i \overset{\frown}{e}^{(i)}$，$y = \sum_{i=1}^{n} \overset{\frown}{y}_i \overset{\frown}{e}^{(i)}$在基组$\{\overset{\frown}{e}^{(i)}\}_{i=1}^{n}$下的坐标$\{\overset{\frown}{x}_i\}_{i=1}^{n}$和$\{\overset{\frown}{y}_i\}_{i=1}^{n}$间应满足怎样的关系。

在线代数教程中已讨论过，如果基向量$\overset{\frown}{e}^{(i)}$，$i = 1, 2,$

…，n，在基组$\{e^{(i)}, i=1, 2, \cdots, n\}$下的坐标表示为

$$\tilde{e}^{(i)}=\sum_{j=1}^{n} t_{ji} e^{(j)} \qquad i=1, 2, \cdots, n$$

即如果由基组$\{e^{(i)}\}_{i=1}^{n}$到基组$\{\tilde{e}^{(i)}\}_{i=1}^{n}$的过渡阵为

$$T=(t_{ji})_{n\times n}$$

则显然T是非奇阵，且

$$\begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_n \end{pmatrix}=\tilde{A}\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_n \end{pmatrix}$$

其中$\tilde{A}=T^{-1}AT$。

由上述结果可见：

1）线性变换$\mathbf{A}\in(\mathbf{C}^n\to\mathbf{C}^n)$的全体的集合与矩阵$A\in\mathbf{C}^{n\times n}$的全体的集合间存在着一一对应关系。而每一个线性变换$\mathbf{A}$所对应的矩阵A的具体表示式取决于$\mathbf{C}^n$中基底向量组的选择。

2）线性变换$\mathbf{A}\in(\mathbf{C}^n\to\mathbf{C}^n)$在$\mathbf{C}^n$的不同的基组$\{x^{(i)}\}_{i=1}^{n}$和$\{y^{(i)}\}_{i=1}^{n}$下的矩阵A和B是彼此相似的。

即

$$B=T^{-1}AT$$

其中相似变换阵T是由$\{x^{(i)}\}_{i=1}^{n}$到$\{y^{(i)}\}_{i=1}^{n}$的过渡矩阵。

§3 矩阵的特征值和矩阵的约化

1.矩阵的特征值和特征向量

〈**定义1.3.1**〉 称给定数域中的数λ为n阶方阵A的特征值，非零的n维列向量x为A的属于λ的（右）特征向量。如果

$$Ax=\lambda x$$

或
$$(\lambda I-A)\ x=0。$$

由线性方程组理论知，特征值λ是A的特征多项式$\varphi_A(\lambda)\equiv\det(\lambda I-A)$的根。注意到特征多项式$\varphi_A(\lambda)$是首1的n次多项式，由代数基本定理知，它在复数域中必定有且仅有n个根。也就是说，任意n阶方阵A在复数域**C**上有且仅有n个特征值（其中重特征值按其重数计算个数），记为$\lambda_i(A)$，$i=1,2,\cdots,n$。称这n个特征值的全体为A的**谱**。显然有

$$\sum_{i=1}^{n}\lambda_i(A)=\mathrm{trace}\ (A),$$

$$\prod_{i=1}^{n}\lambda_i(A)=\det\ (A)。$$

同样由线性方程组的理论知，A的属于特征值λ的特征向量的全体连同零向量构成齐次线性方程组

$$(\lambda I-A)\ x=0$$

的解空间，称它为A的对应于特征值λ的特征子空间。对于它的维数有如下结论：

〈**定理1.3.1**〉 n阶方阵A的对应于特征值λ_0的特征子空间的维数r不超过特征值λ_0的重数k。

证明，由线性方程组理论及A的相应于λ_0的特征子空间的维数r就是齐次方程组