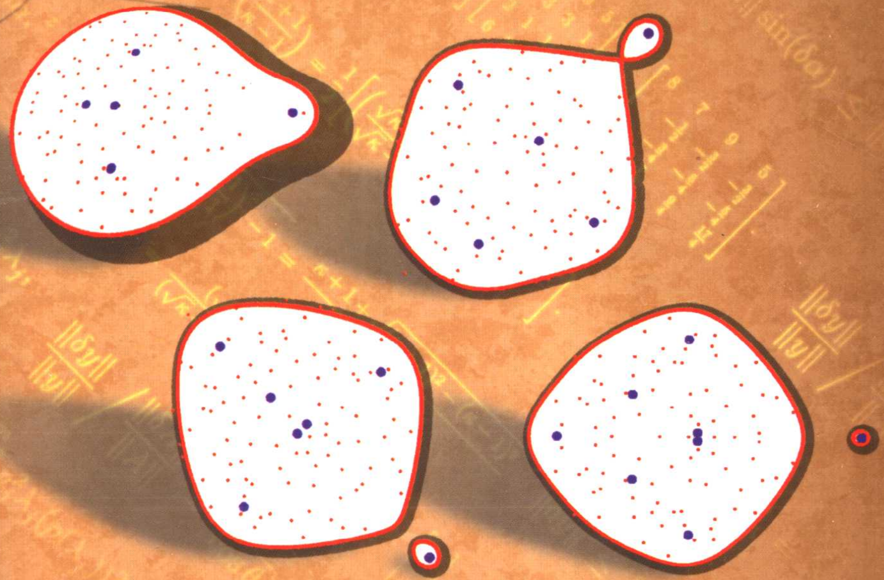
TURING 图灵数学·统计学丛书



Numerical Linear Algebra

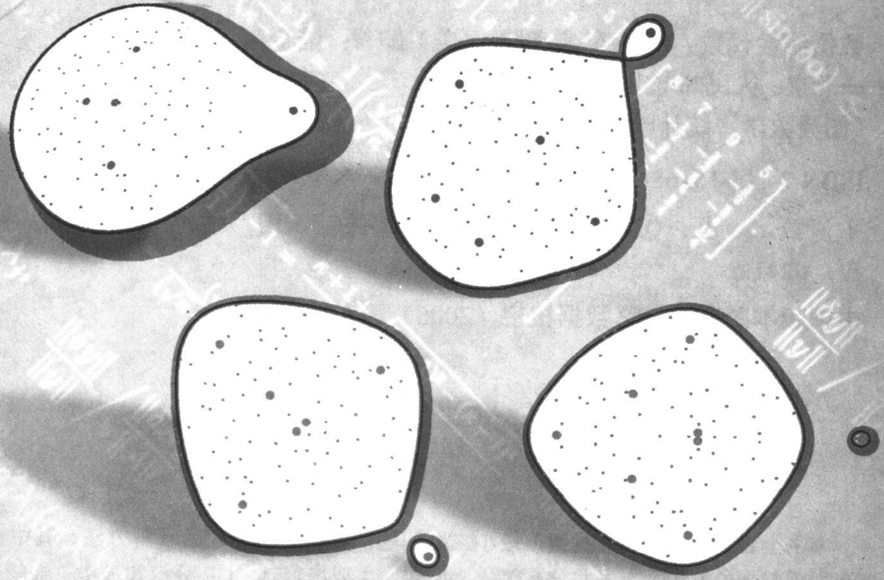
数值线性代数

Lloyd N. Trefethen
David Bau, III 著

陆金甫 关治 译

人民邮电出版社
POSTS & TELECOM PRESS

TURING 图灵数学·统计学丛书



Numerical Linear Algebra

数值线性代数

Lloyd N. Trefethen
David Bau, III 著

陆金甫 关治 译

人民邮电出版社
POSTS & TELECOM PRESS

图书在版编目（CIP）数据

数值线性代数 / （英）特雷弗腾，（美）鲍著；陆金甫，关治译.
—北京：人民邮电出版社，2006.11
（图灵数学·统计学丛书）

ISBN 7-115-15168-7

Ⅰ. 数…　Ⅱ. ①特…②鲍…③陆…④关…　Ⅲ. 线性代数计算法—教材
Ⅳ. 0241.6

中国版本图书馆 CIP 数据核字（2006）第 098728 号

内容提要

本书全面论述了线性方程组、最小二乘问题以及特征值问题的求解方法，其中包含不少新近发展起来的方法. 全书共分 6 部分，40 讲。主要内容有：QR 分解和最小二乘问题、条件与稳定性、求解线性方程组的直接方法、特征值问题及迭代方法.

本书可作为计算数学专业、科学和工程学科高年级本科生或研究生一学期的教材，也可供应用工作者参考.

图灵数学·统计学丛书

数值线性代数

◆ 著　Lloyd N.Trefethen　David Bau, III
译　陆金甫　关　治
责任编辑　武卫东

◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街 14 号
邮编　100061　电子函件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
北京铭成印刷有限公司印刷
新华书店总店北京发行所经销

◆ 开本：700 × 1000　1/16
印张：20.5
字数：425 千字　2006 年 11 月第 1 版
印数：1– 4 000 册　2006 年 11 月北京第 1 次印刷

著作权合同登记号　图字：01-2006-1278 号

ISBN 7-115-15168-7/O1 • 3

定价：39.00 元

读者服务热线：(010)88593802　印装质量热线：(010)67129223

版 权 声 明

符　　号

对正方形或矩形矩阵 $A \in \mathbb{C}^{m\times n}$，$m \geqslant n$：

QR 因子分解：$A = QR$

约化 QR 因子分解：$A = \hat{Q}\hat{R}$

SVD：$A = U\Sigma V^*$

约化 SVD：$A = \hat{U}\hat{\Sigma}V^*$

对正方形矩阵 $A \in \mathbb{C}^{m\times m}$：

LU 因子分解：$PA = LU$

楚列斯基因子分解：$A = R^*R$

特征值分解：$A = X\Lambda X^{-1}$

舒尔因子分解：$A = UTU^*$

正交投影算子：$P = \hat{Q}\hat{Q}^*$

豪斯霍尔德镜射算子：$F = I - 2\dfrac{vv^*}{v^*v}$

QR 算法：$A^k = \underline{Q}^{(k)}\underline{R}^{(k)}$，　$A^{(k)} = (\underline{Q}^{(k)})^{\mathrm{T}}A\,\underline{Q}^{(k)}$

阿诺尔迪迭代：$AQ_n = Q_{n+1}\tilde{H}_n$，　$H_n = Q_n^*AQ_n$

兰乔斯迭代：$AQ_n = Q_{n+1}\tilde{T}_n$，　$T_n = Q_n^{\mathrm{T}}AQ_n$

译 者 介 绍

陆金甫 清华大学数学科学系教授．研究方向为偏微分方程差分方法及并行计算．长期从事计算数学的教学与研究，参加和负责过多次国家自然科学基金项目．

关　治 清华大学数学科学系教授，专业方向为计算数学．长期从事基础数学与计算数学的教学与研究，编著《全国工程硕士专业学位入学资格考试考前辅导教程》及其他多部教材与专著，曾多次获清华大学优秀教学成果奖．

两译者曾合作编著了《数值分析基础》（高等教育出版社）、《偏微分方程数值方法》（清华大学出版社）等教材．

译 者 序

L. N. Trefethen 和 D. Bau, Ⅲ的《数值线性代数》是一本有特色的教材. 它适用于计算数学专业高年级本科生或研究生课程, 也是科学与工程学科研究生的一本好的参考书.

本书在不大的篇幅中, 全面论述了线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$、最小二乘问题 $\min_{x}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2$ 以及特征值问题的求解方法, 包括了基本求解方法及新近发展起来的新方法, 全书分为基础知识、QR 分解和最小二乘法问题、条件与稳定性、求解线性方程组的直接方法、特征值问题以及迭代方法 6 个部分. 特征值问题求解及迭代方法是本书的重点, 它们占了本书近一半篇幅. 全书共 40 讲, 每讲讲述 1 个或 2 个主题, 安排内容都不是太多, 便于课堂讲授.

与通常的数值线性代数教材相比较, 该教材具有显著特点. 在内容安排上, 把多处用到的 QR 分解方法放在高斯消元法前面讲述; 在迭代方法的材料选择上基本不涉及经典的方法, 而几乎全部讲述克雷洛夫子空间迭代方法; 在算法的讲述上特别注重算法的思想, 强调了各个算法之间的联系, 例如共轭梯度法与多项式逼近, 兰乔斯迭代与高斯求积方法等; 几何方面的说明在教材中起到了重要作用, 例如在阿诺尔迪的特征值方法中采用了"阿诺尔迪双纽线"来说明方法的收敛性等等.

本书正文后有一个附录《数值分析的定义》, 它是第一作者已发表过的一篇文章. 也已译出, 供参考.

本教材中有些方法细节未详细讲述, 有关内容, 读者可参考 J. W. Demmel 的 *Applied Numerical Linear Algebra* (1997)[1]和 Y. Saad 的 *Iterative Methods for Sparse Linear Systems* (2nd edition,2003) 等书.

本书从开头到第 23 讲由关治翻译, 其后部分由陆金甫翻译, 张宝琳研究员抽空看过部分译稿并提出了很好的意见.

由于译者的英文水平和学科水平都有限. 译文难免有不妥之处, 请广大读者批评指正.

译 者

2006 年 4 月

1. 该书的中译本《应用数值线性代数》也已由人民邮电出版社出版. ——编者注

前　　言

从上个世纪80年代初期开始，本书第一作者在麻省理工学院和康奈尔大学讲授数值线性代数的研究生课程. 听过这门课程的学生如今已有数以百计之多，他们都是来自工程和物理科学等相关领域的研究生. 本书就是在这门课程所讲述的内容的基础上整理而成的.

在数值线性代数领域，已经有一本百科全书式的著作，即 Golub 和 Van Loan 的 *Matrix Computations*[1]，现已有第三版. 本书绝不是要重复前人的工作，而是想以尽可能简练的方式给出基本的数学思想. 本书篇幅小，适合于一学期教学之用. 当然，我们仍希望本书的每一位读者阅读 Golub 和 Van Loan 的书，了解进一步的细节和其他相关主题，以及书中为该领域研究提供的丰富的参考文献. 最近新出了两本重要的书，由 Higham 和 Demmel 所著，我们将在本书最后的注记中提到.

数值线性代数，虽然名字略显沉闷，但是这一领域本身却是多彩且基础性的. 说它多彩，是因为它充满了极富生命力的思想，这完全不同于在数学系线性代数课程通常所强调的内容.（在每学期结束之时，学生们总是说这门课程要比他们一直想像的有更多的内涵.）说它是基础性的，是因为由于历史原因“数值”线性代数事实上就是指应用线性代数. 这里涵盖了每个数学家有效地应用向量和矩阵都需要的基本思想. 事实上，我们的课题并不仅仅局限于向量和矩阵，因为我们做的每一件事情都涉及函数和算子. 数值线性代数在真正意义上归于泛函分析，只不过它的重点总是在实际的算法思想而不是数学上的技术细节.

本书共分40讲. 我们力图使每讲围绕一个或两个中心思想，强调各主题之间的统一性并且不拘泥于细节的处理. 许多地方的处理方式并非是标准的. 这里不便一一列出（参看注记），但是我们会指出本书中一个比较特殊的地方. 我们将一反惯例，不从高斯消元法开始，那是数值线性代数中非典型的算法，极其难以分析，而对每个初学本课程的学生，它又过于熟知因而有些乏味. 代之的是，我们从 QR 因子分解开始，它更重要，而且通俗易懂，对大多数学生来说是新鲜的思想. QR 因子分解是联系数值线性代数大多数算法的纽带，包括最小二乘、特征值和奇异值问题等方法，同时也包括对这些问题和对方程组的迭代方法. 自从上个世纪70年代以来，迭代方法已经占据了科学计算的中心地位，对它们，我们安排在本书的最后一部分加以介绍.

我们希望读者能同意我们如下的见解，即如果除了微积分与微分方程之外，还有什么数学领域是数学科学的基础的话，那就是数值线性代数.

---

1. 中译本为：《矩阵计算》，袁亚湘等译，科学出版社，2001. ——译者注

致　　谢

如果没有众多人士的帮助，我们不可能完成此书．首先要感谢修读麻省理工学院课程 Math 335 和康奈尔大学课程 CS 621 的数以百计的研究生，他们在整个 10 年期间的热情和建议影响了本课程内容的选择和表达的方式．在康奈尔大学，这些学生中大约有 70 人使用了本书的初稿，他们提出了很多有益的建议．Keith Sollers 就独自指出了大量排印错误．

Trefethen 的大多数研究生在本书写作期间，从头到尾通读了这本教科书——有时是在很短时间内完成的．感谢 Jeff Baggett、Toby Driscoll、Vicki Howle、Gudbjorn Jonsson、Kim Toh 和 Divakar Viswanath 的很多建设性建议．拥有众多学生和同事，真是荣幸．

与 SIAM 的出版工作人员的合作是令人愉快的，很少能有出版单位像 SIAM 这样将灵活性和专业性集于一身．我们十分感谢 SIAM 的编辑、生产和设计人员，他们的共同努力大大增加了本书的吸引力，尤其要感谢 Beth Gallagher，她除了出色地完成了文字编辑工作之外，还自始至终在支持着我们．

康奈尔大学计算机科学系对数值线性代数及其教材出版给予的支持无与伦比．系里其他 3 位在此方面有兴趣的同事是 Tom Coleman，Charlie Van Loan 和 Steve Vavasis，我们感谢他们在使康奈尔大学成为富于吸引力的科学计算中心的过程中所做的贡献．Vavasis 阅读了本书整个的初稿并提出很多有价值的建议，Van Loan 是最初介绍 Trefethen 来到康奈尔的人．在我们的非数值方面的同事之中，感谢 Dexter Kozen，他的 *The Design and Analysis of Algorithms* 一书是本书仿效的对象，它也是一本由 40 讲组成的书．感谢 Rebekah Personius 先生，作为本系的一名行政人员，他的专业特长、勤勤恳恳和良好的工作态度给我们提供了很大的支持．

我们的另一位同事，经常到康奈尔大学访问的学者 Anne Greenbaum 为本书提供了大量建议，她是我们认识的对数值线性代数思考最深入的学者之一．

从 1995 年 9 月到 1996 年 12 月，我们的几位同事用本书的初稿讲授课程，他们提出了自己的意见，也转达了学生们的建议．其中有 Gene Golub（斯坦福大学）、Bob Lynch（普度大学）、Suely Oliveira（得克萨斯 A & M 大学）、Michael Overton（纽约大学）、Haesun Park 和 Ahmed Sameh（明尼苏达大学）、Irwin Pressmann（加拿大卡尔顿大学）、Bob Russell 和 Manfred Trummer（加拿大西蒙・费雷泽大学）以及 Hong Zhang 和 Bill Moss（美国克莱姆森大学）．在这些人中打破纪录的是 Lynch 和 Overton，他们每人都提出了长长的一系列的详细建议．虽然我们写书时已尽量一丝不苟，但他们的意见还是非常合理的，不能忽略，而现在本书中上百处改进正是因为 Lynch 或 Overton 的帮助．

要说到有什么重要的帮助使本书更加完善，我们要着重提及曼彻斯特大学的 Nick Higham 教授，他所提供的帮助（虽然他不愿意承让）是我们永远的情债，他的创造性和细致学风启迪了一批比他年幼或年长的数值分析学家（最小的是他年龄的一半，最大的是他年龄的两倍）. 虽然时间很短，Higham 却带着他特有的好意精心地阅读了本书的初稿并提出了很多页技术上的建议，其中某些建议深深地改变了本书.

几十年来，数值线性代数已经成为充满友爱和富有凝聚力的领域的一种典范. Trefethen 要特别感谢 3 位“教父级的长者”，是他们的课堂讲授首先将他吸引到这个主题，他们是 Gene Golub，Cleve Moler 和 Jim Wilkinson.

除了数值线性代数之外，还有更多事情使生命更有意义. 为此，第一作者要感谢家人 Anne、Emma(5 岁)和 Jacob(3 岁)，第二作者要感谢 Heidi Yeh.

目　录

第 I 部分
基　　础

- 第 1 讲　矩阵-向量乘法
- 第 2 讲　正交向量和矩阵
- 第 3 讲　范数
- 第 4 讲　奇异值分解
- 第 5 讲　SVD 的进一步讨论

第 1 讲 矩阵-向量乘法

我们已经知道了矩阵-向量乘法的公式，然而，第 1 讲的目的是描述一种解释这种乘积的方法，它可能是人们不太熟悉的. 如果 $\boldsymbol{b}=\boldsymbol{A}\boldsymbol{x}$，则 $\boldsymbol{b}$ 是 $\boldsymbol{A}$ 的列的一个线性组合.

1.1 熟知的定义

令 $\boldsymbol{x}$ 为一个 n 维列向量，$\boldsymbol{A}$ 为一个 $m\times n$ 矩阵（m 行 n 列）. 则矩阵-向量乘积 $\boldsymbol{b}=\boldsymbol{A}\boldsymbol{x}$是如下定义的 m 维列向量：

$$b_i=\sum_{j=1}^{n} a_{ij}x_j, \qquad i=1,\cdots,m. \tag{1.1}$$

这里，b_i 记为 b 的第 i 个元素，a_{ij}记为 A 的 i,j 元素（第 i 行第 j 列），x_j 记为 $\boldsymbol{x}$ 的第 j 个元素. 为了简化，在本书中除了几讲外，都假设这些量属于复数域$\mathbb{C}$. m 维向量空间为$\mathbb{C}^m$，$m\times n$ 矩阵的空间为$\mathbb{C}^{m\times n}$.

映射 $\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 是线性的（linear），意思是对任意的 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{C}^n$ 和任意的 $\alpha\in\mathbb{C}$，

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{x}+\boldsymbol{y}) &= \boldsymbol{A}\boldsymbol{x}+\boldsymbol{A}\boldsymbol{y},\\ \boldsymbol{A}(\alpha\boldsymbol{x}) &= \alpha\boldsymbol{A}\boldsymbol{x}.\end{aligned}$$

反之，每个由$\mathbb{C}^n$ 到$\mathbb{C}^m$ 的线性映射都对应一个 $m\times n$ 矩阵.

1.2 矩阵乘向量

令 m 维向量 $\boldsymbol{a}_j$ 记为 $\boldsymbol{A}$ 的第 j 列，那么（1.1）可重写成

$$\boldsymbol{b}=\boldsymbol{A}\boldsymbol{x}=\sum_{j=1}^{n} x_j\boldsymbol{a}_j. \tag{1.2}$$

这个方程可以如下格式显示：

$$\left[\begin{array}{c} \\ \boldsymbol{b} \\ \\ \end{array}\right]=\left[\begin{array}{c|c|c|c} & & & \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \\ & & & \end{array}\right]\begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}=x_1\left[\begin{array}{c} \\ \boldsymbol{a}_1 \\ \\ \end{array}\right]+x_2\left[\begin{array}{c} \\ \boldsymbol{a}_2 \\ \\ \end{array}\right]+\cdots+x_n\left[\begin{array}{c} \\ \boldsymbol{a}_n \\ \\ \end{array}\right].$$

在（1.2）里，$\boldsymbol{b}$ 表示为列 $\boldsymbol{a}_j$ 的线性组合. 由（1.1）到（1.2）只是记号稍作改变，然而，要正确理解数值线性代数算法，理解（1.2）中的 $\boldsymbol{A}\boldsymbol{x}$ 运算形式是非常重要的.

我们可以用下面的方法综合矩阵–向量乘积各种不同的描述. 作为数学家，我们将公式$\boldsymbol{Ax}=\boldsymbol{b}$看成$\boldsymbol{A}$作用在$\boldsymbol{x}$上产生$\boldsymbol{b}$；反之，公式（1.2）则解释为$\boldsymbol{x}$作用在$\boldsymbol{A}$上产生$\boldsymbol{b}$.

例 1.1 范德蒙德矩阵. 固定一个数列$\{x_1,x_2,\cdots,x_m\}$，如果p和q都是次数$<n$的多项式，α是一个标量，那么$p+q$和αp也是次数$<n$的多项式，而且，这些多项式在点x_i的值满足下列线性性质：

$$(p+q)(x_i) = p(x_i)+q(x_i),$$
$$(\alpha p)(x_i) = \alpha(p(x_i)).$$

因此，由次数$<n$的多项式p的系数向量到多项式样本值的向量$(p(x_1),p(x_2),\cdots,p(x_m))$的映射是线性的. 任意一个线性映射可以表示为矩阵的乘法. 这就是一个例子. 事实上，它表示为一个$m\times n$阶范德蒙德矩阵（Vandermonde matrix）

4

$$\boldsymbol{A}=\begin{bmatrix}1 & x_1 & x_1^2 & \cdots & x_1^{n-1}\\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1}\end{bmatrix}.$$

如果$\boldsymbol{c}$是p的系数列向量

$$\boldsymbol{c}=\begin{bmatrix}c_0\\ c_1\\ c_2\\ \vdots\\ c_{n-1}\end{bmatrix},\quad p(x)=c_0+c_1x+c_2x^2+\cdots+c_{n-1}x^{n-1},$$

那么，乘积$\boldsymbol{Ac}$将给出多项式样本值，即对从 1 到m的每个i，我们有

$$(\boldsymbol{Ac})_i = c_0+c_1x_i+c_2x_i^2+\cdots+c_{n-1}x_i^{n-1} = p(x_i). \tag{1.3}$$

在这个例子里，显然矩阵–向量乘积$\boldsymbol{Ac}$不必像（1.1）所示的那样，看作m个不同的数量和（每个和都给出一个关于$\boldsymbol{c}$的元素的线性组合），相反，$\boldsymbol{A}$可以看成一个以单项式的样本值为列的矩阵

$$\boldsymbol{A}=\left[\begin{array}{c|c|c|c|c}1 & x & x^2 & \cdots & x^{n-1}\end{array}\right], \tag{1.4}$$

乘积$\boldsymbol{Ac}$可以理解为（1.2）形式的单个向量的和，立刻就可以给出这些单项式的线性组合

$$Ac = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1} = p(x)\ . \qquad \square$$

本讲余下的部分将用（1.2）的观点重温线性代数的一些基本概念.

1.3 矩阵乘矩阵

对于矩阵-矩阵乘积 $B = AC$，B 的每列是 A 的列的线性组合. 为导出这个事实，我们从通常的矩阵乘积公式出发. 若 A 是 $\ell \times m$ 阶矩阵，C 是 $m \times n$ 阶矩阵，则 B 是 $\ell \times n$ 阶矩阵，其元素定义为

$$b_{ij} = \sum_{k=1}^{m} a_{ik} c_{kj}. \tag{1.5}$$

其中，b_{ij}，a_{ik}和 c_{kj}分别是 B，A 和 C 的元素. 按列来写，此乘积为

$$\left[\begin{array}{c|c|c|c} \\ b_1 & b_2 & \cdots & b_n \\ \\ \end{array}\right] = \left[\begin{array}{c|c|c|c} \\ a_1 & a_2 & \cdots & a_m \\ \\ \end{array}\right] \left[\begin{array}{c|c|c|c} c_1 & c_2 & \cdots & c_n \end{array}\right],$$

5

（1.5）成为

$$b_j = Ac_j = \sum_{k=1}^{m} c_{kj} a_k. \tag{1.6}$$

b_j 是诸列向量 a_k 与对应系数 c_{kj}的线性组合.

例 1.2 外积. 矩阵-矩阵乘积的一个简单例子是外积（outer product），这是一个 m 维列向量 u 与一个 n 维行向量 v 的乘积；结果是一个秩为 1 的 $m \times n$ 矩阵. 外积可以写成

$$\left[\begin{array}{c} \\ u \\ \\ \end{array}\right] \left[\begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array}\right] = \left[\begin{array}{c|c|c|c} \\ v_1 u & v_2 u & \cdots & v_n u \\ \\ \end{array}\right] = \begin{bmatrix} v_1 u_1 & \cdots & v_n u_1 \\ \vdots & & \vdots \\ v_1 u_m & \cdots & v_n u_m \end{bmatrix}.$$

其列是同一个向量 u 的所有倍数，类似地，其行是同一个向量 v 的所有倍数. $\square$

例 1.3 作为第 2 个实例，考虑 $B = AR$. 其中，R 是一个 $n \times n$ 阶上三角矩阵，其元素对$i \leqslant j$有 $r_{ij} = 1$，对 $i > j$ 有 $r_{ij} = 0$. 乘积可以写成

$$\left[\begin{array}{c|c|c} \\ b_1 & \cdots & b_n \\ \\ \end{array}\right] = \left[\begin{array}{c|c|c} \\ a_1 & \cdots & a_n \\ \\ \end{array}\right] \begin{bmatrix} 1 & \cdots & 1 \\ & \ddots & \vdots \\ & & 1 \end{bmatrix}.$$

公式（1.6）给出

$$b_j = Ar_j = \sum_{k=1}^{j} a_k. \tag{1.7}$$

即 $\boldsymbol{B}$ 的第 j 列是 $\boldsymbol{A}$ 的前 j 列的和，矩阵 $\boldsymbol{R}$ 是不定积分算子的离散模拟. □

1.4 值域和零空间

矩阵的值域（range）记为 range($\boldsymbol{A}$)，它是所有能表示为 $\boldsymbol{Ax}$ 形式的向量的集合. 公式（1.2）自然导出下列 range($\boldsymbol{A}$)的特征.

6 **定理 1.1** range($\boldsymbol{A}$)是由 $\boldsymbol{A}$ 的列所张成的空间.

证明 由（1.2），任意的 $\boldsymbol{Ax}$ 是 $\boldsymbol{A}$ 的列的线性组合，反之，由 $\boldsymbol{A}$ 的列所张成的空间中的任意向量 $\boldsymbol{y}$ 可以写成列的一个线性组合，$y = \sum_{j=1}^{n} x_j \boldsymbol{a}_j$. 由于由系数 x_j 构成向量 $\boldsymbol{x}$，我们有 $\boldsymbol{y} = \boldsymbol{Ax}$，因此 $\boldsymbol{y}$ 在 $\boldsymbol{A}$ 的值域中. □

由定理 1.1，矩阵 $\boldsymbol{A}$ 的值域也称为 $\boldsymbol{A}$ 的列空间（column space）.

$\boldsymbol{A} \in \mathbb{C}^{m\times n}$，其零空间（nullspace）写成 null($\boldsymbol{A}$)，它是满足 $\boldsymbol{Ax} = \boldsymbol{0}$ 的向量 $\boldsymbol{x}$ 的集合，这里的 $\boldsymbol{0}$ 是 $\mathbb{C}^m$ 中的 $\boldsymbol{0}$-向量. 每个向量 $x \in$ null($\boldsymbol{A}$)的元素给出零作为 $\boldsymbol{A}$ 的列的线性组合展开式的系数：$\boldsymbol{0} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$.

1.5 秩

矩阵的列秩（column rank）是它的列空间的维数，类似地，矩阵的行秩（row rank）是由其行所张成的空间的维数. 行秩总是等于列秩（在其他一些证明中，这是在第 4 讲和第 5 讲中讨论的奇异值分解的一个推论），所以我们简称这个数为矩阵的秩.

一个满秩（full rank）的 $m \times n$ 矩阵是具有最大可能的秩（m 和 n 的较小者）的矩阵. 这意味着一个$m \geqslant n$的满秩矩阵一定有 n 个线性无关的列，这样的矩阵也可以用这样的性质来表征：它所定义的映射是一一对应的.

定理 1.2 $m \geqslant n$ 的矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 为满秩的，当且仅当它不把两个不同的向量映射到同一个向量上.

证明 （⇒）如果 $\boldsymbol{A}$ 满秩，那么它的列线性无关，所以它们构成了 range($\boldsymbol{A}$)的一组基. 这意味着每个 $\boldsymbol{b} \in$ range($\boldsymbol{A}$)有惟一的关于 $\boldsymbol{A}$ 的列的线性展开式，因此，由（1.2），每个 $\boldsymbol{b} \in$ range($\boldsymbol{A}$)存在惟一的 $\boldsymbol{x}$ 使得 $\boldsymbol{b} = \boldsymbol{Ax}$.（⇐）反之，若 $\boldsymbol{A}$ 并非满秩，则它的列 $\boldsymbol{a}_j$ 是相关的，存在一个非平凡的线性组合使得 $\sum_{j=1}^{n} c_j \boldsymbol{a}_j = 0$. 由系数 c_j 构成的非零向量 $\boldsymbol{c}$ 满足 $\boldsymbol{Ac} = 0$，所以对任意 $\boldsymbol{x}$，有 $\boldsymbol{Ax} = \boldsymbol{A}(\boldsymbol{x} + \boldsymbol{c})$，这样 $\boldsymbol{A}$ 把不同的向量映射到同一个向量，矛盾. □