矩阵分析

R·A·合恩
C·R·约翰逊 著
杨奇 译

天津大学出版社

矩 阵 分 析

R·A·合　恩
C·R·约翰逊　著
杨　奇　译
侯自新　[illegible]

天津大学出版社

内容提要

本书从数学分析的角度论述矩阵分析的经典结果和现代结果，取材新，有一定的深度，并给出在多元微积分、复分析、微分方程、最优化、逼近理论中的许多重要应用。

主要内容有：特征值，特征向量和相似性；酉等价和正规矩阵；Hermite矩阵和对称矩阵；向量范数和矩阵范数；特征值的估计和扰动；正定矩阵；非负矩阵。

本书可作为工程、统计、经济学等专业的研究生教材或大学数学系高年级学生教材，亦是一本很有价值的自学参考书。

矩 阵 分 析

R·A·合恩 C·R·约翰逊 著

杨 奇 译

*

天津大学出版社出版

（天津大学内）

河北省永清县印刷厂印刷

新华书店天津发行所发行

*

开本：787×1092毫米1/16 印张：22 字数：549千字

1989年9月第一版 1989年9月第一次印刷

印数：1—3500

ISBN 7-5618-0141-6

O·14

定价：3.65元

译者的话

近年来，由于计算机技术的发展和普及，矩阵理论的重要性愈加显著，应用日益广泛，而矩阵理论本身也得到了迅猛的发展。然而，目前使用的教科书中所讲述的矩阵内容已不能满足实际的需要，而且处理方法陈旧，缺乏实际应用。因此，广大读者希望有一本较现代的矩阵理论教材。

1987年初，译者看到R.A.Horn和C.R.JohnSon合著的*Matrix analysis*(Cambridge大学出版社，85年第1版)的书评介绍，认为它是一种不可多得的好教科书。我立刻把书找来仔细阅读，深深地被其丰富的内容和新颖的叙述吸引住了，同时产生了将该书译成中文，介绍给国内广大读者的强烈愿望，后经教研室和数学系推荐，被天津大学出版社列入了出版计划。

关于原书的撰写意图、特色、内容概要以及使用本书的建议等可参阅序言，这里不再赘述。对于原书的某些错误（主要是排印错误）译者在力所能及的范围内作了订正，而未一一指明。为了压缩篇幅，适当对原书中的部分内容，供进一步阅读的参考文献和参考书目，以及部分习题作了删减。

在本书的翻译过程中，南开大学数学系侯自新副教授认真细致地审阅了全部译稿，并提出了很多宝贵的意见。译者自始至终还得到天津大学出版社的大力支持以及译者的同事和朋友的热情帮助，在此，一并表示衷心的感谢。由于译者水平有限，时间匆促，误译之处在所难免，恳请读者给予指正。

译者　1988年7月于天津大学数学系

序　言

线性代数和矩阵理论很久以来就已经是各数学学科的基本工具，而且就其本身的研究说来它们也是富有创造性的领域。在本书及其姐妹篇（*Topics in Matrix Analysis*）中，要给出矩阵分析的经典结果和现代结果，这些结果已被证明对应用数学是重要的。本书可作为大学生或研究生的教材，也可以作为各种不同读者的自学参考书。我们假定读者已学过一个学期的基础线性代数课程，并具备初步的分析知识。本书从讨论特征值和特征向量开始；不要求已熟悉这些概念。

除了在基础线性代数教程中会遇到有关矩阵的内容以外，实际上，任何涉及数学的领域都需要矩阵的知识，不论它是微分方程、概率统计、最优化；还是理论经济学及应用经济学，工程技术学或运筹学，这里仅举出上述几方面。然而，直到最近，很多必需的内容才零星地（或者根本没有）出现在大学生或研究生的课程中。随着应用数学的重要性不断增长，有相当多课程专门研究高等矩阵理论，这就迫切需要有一本能广泛选择各种论题的教材，它还能对有关的论题提供现代参考资料。

虽有不少深受读者喜爱的矩阵理论经典著作，但它们既不适合于一般的课堂教学，也不适合于系统的自学。某些传统的参考书给读者带来的困难是：缺乏习题、不注重应用和启发诱导；索引不完善；处理方法陈旧。较现代的书又往往趋向于作为基础教材，或趋向于作为讨论专门化论题的专著。我们的目标是撰写一本能对范围广泛的各种论题进行有效而现代处理的书。

“矩阵分析”这本书的一个意图是，它要包括由于数学分析（例如，多元微积分、复变量、微分方程、最优化和逼近理论等）的需要而产生的线性代数中的那些论题。本书的另一个意图是，矩阵分析是解决实的和复的线性代数问题的一种方法，这种方法果断地采用诸如极限、连续和幂级数这些来自分析的概念，这是考虑到这些概念有时比纯代数方法更为有效或更为自然。矩阵分析的这两个出发点就影响了对本书中所述论题的选择和处理。我们认为采用术语矩阵分析比线性代数更能准确地反映该领域的广泛内容和研究方法。

为了复习和便于查阅，本书第0章包括基础线性代数所必需内容的一个提要，另外还包括一些有用资料。第1章、第2章和第3章涉及大部分核心内容，这些内容可能包括在线性代数或矩阵理论的任何选修教程中，主要有：特征值、特征向量和相似性；酉相似、Schur三角化及其推论、正规矩阵；标准形和包括Jordan标准形在内的分解．*LU*分解，*QR*分解和友矩阵。除此以外，以下各章大体上都是独立展开的，而且以一定的深度讨论一个主要论题：

Hermite矩阵和复对称矩阵（第4章）．我们着重讨论了关于研究Hermite矩阵的特征值的变分法，其中包括优化概念的引入。

向量范数和矩阵范数（第5章）．它们对于数值线性代数算法的误差分析，对于研究矩阵幂级数和迭代过程是必不可少的。我们较详细地讨论范数的代数性质、几何性质和分析性质，并且仔细区分与矩阵范数的次乘性公理有关的和无关的关于矩阵的那些范数结果。

特征值的估计和扰动理论（第6章）．这是针对一般矩阵（不一定Hermite矩阵）来讨

论的，并且对于许多应用是很重要的。我们详细讨论了Geršgorin区域理论以及它的某些现代的改进，同时还给出了有关的图论概念。

正定矩阵（第7章）。我们相当详细地考察了正定矩阵及其包括不等式在内的各种应用讨论了极分解和奇异值分解，同时还讨论了对矩阵逼近问题的应用。

其分量是非负数和正数的矩阵（第8章）．它们起因于其中必然会出现非负量的许多应用（如概率论、经济学和工程技术等），它们的引人瞩目的理论反映了这些应用。在运用范数的基础上，我们初步讨论了关于非负矩阵、正矩阵、素矩阵和不可约矩阵的理论。

在姐妹篇中，将讨论其它一些同样值得注意的论题：值域及推广；惯性、稳定矩阵、M-矩阵和相关特殊类；矩阵方程、Kronecher 乘积和Hadamard乘积；可以把函数和矩阵联系起来的各种方法。

通过选择适合于不同读者的章节内容，本书为一个学期或两个学期的课程提供了基本教材。我们建议，为了一门特殊课程的需要，教师要对本书的章节及各节的内容事先进行仔细的选择。这大概应当包括第1章、第2章和第3章的大部分以及第4章和第5章中有关Hermite矩阵和范数的内容。

大多数章包括了一些比较专门化的或非传统的材料。例如，第2章不仅包括一个矩阵的酉三角化的Schur基本定理，而且还讨论了矩阵族的同时三角化。在关于酉等价那一节中，在介绍通常的内容之后，还讨论两个矩阵是酉等价的迹条件．第4章中关于复对称矩阵的讨论是对照推导Hermite矩阵的经典理论给出的。一个论题的基本观点出现在每一章的前几节中，而更为详细的讨论则出现在各节的末尾或以后各节中。这样处理的优点是可以依次提出各种论题，从而提高了本书作为参考书的使用价值，它还为教师提供了广泛的选择余地。

所讨论的许多结果适合于或者可经推广后适合于其它域上的矩阵，或者可以在一些更广泛的代数系统中讨论这些结果。不过，我们有意识地将定义域限制在实数域或复数域，在这种条件下就能够使用经典分析的熟悉方法以及正规的代数方法。

虽然我们一般考虑复矩阵，但大多数例子只限于实矩阵，并且不需要较深的复分析知识。熟悉复数的算术对于理解矩阵分析是必不可少的，其熟悉程度需达到附录中所规定的水平．其余几个简短的附录包含几个次要的而又必需的论题，例如，Weierstrass定理和凸性。

在本书中，给出了许多练习和习题，因为我们认为这些练习和习题对于进一步理解主要论题及其推论是必要的。练习始终是作为每节的论述部分而出现的，它们一般是很基本的，直接用于加深对概念的理解。建议读者充分选做这些练习。习题安排在每节的末尾（没有特定的顺序）；它们涉及一系列难题和典型题（有证明题，也有计算题）；可以用来扩展论题，提出特殊见解，或者为主要论题提供另外的证明。对较难的习题都给出明显的提示。有些习题的结果与其它习题中或正文本身中的结果有关。我们要特别强调，读者能尽心竭力去完成练习和钻研习题是至关重要的。

虽然本书不是一本有关应用方面的书，但为了激发学习动机，在每一章的开始都介绍几个应用来引入本章的论题。

（以下部分从略）

R.A.H.

C.R.J.

符 号 表

$\mathbf{R}$	实数域
$\mathbf{R}^n$	实 n 维列向量空间，$M_{n,1}(\mathbf{R})$
$\mathbf{C}$	复数域
$\mathbf{C}^n$	复 n 维列向量空间，$M_{n,1}(\mathbf{C})$
$\mathbf{F}$	一个域（通常指R或C）
$\mathbf{F}^n$	（域F上）分量取自F的 n 维列向量空间，$M_{n,1}(\mathbf{F})$
$M_{m,n}(\mathbf{F})$	元素取自F的$m\times n$矩阵的集合
$M_{m,n}$	$m\times n$复矩阵的集合，$M_{m,n}(\mathbf{C})$
M_n	$n\times n$复矩阵的集合，$M_{n,n}(\mathbf{C})$
A, B, C, 等等	矩阵；$A=[a_{ij}]\in M_{m,n}(\mathbf{F})$
x, y, z, 等等	列向量；$x=[x_i]\in \mathbf{F}^n$
I	$M_n(\mathbf{F})$中的单位矩阵
0	零纯量，零向量或零矩阵
$\bar{A}$	$A\in M_{m,n}(\mathbf{C})$的各元素取复共轭的矩阵
A^T	$A\in M_{m,n}(\mathbf{F})$的转置
A^*	$A\in M_{m,n}(\mathbf{C})$的Hermite伴随，$\bar{A}^T$
A^{-1}	非奇异矩阵$A\in M_n(F)$的逆
$A^{1/2}$	半正定矩阵$A\in M_n$的唯一半正定平方根
$\|A\|$	$A\in M_{m,n}$的各元素取绝对值的矩阵
A^+	$A\in M_{m,n}$的Moore-Penrose广义逆
$\mathrm{adj}A$	$A\in M_n(\mathbf{F})$的经典伴随（转置伴随）
$\mathscr{B}$	向量空间的一个基
e_i	（通常指）$\mathbf{F}^n$中第 i 个标准基向量
$[v]_{\mathscr{B}}$	向量 v 的$\mathscr{B}$坐标表示
${}_{\mathscr{B}_2}[T]_{\mathscr{B}_1}$	线性变换T的$\mathscr{B}_1$-$\mathscr{B}_2$基表示
$\binom{n}{k}$	二项式系数，$n!/[k!(n-k)!]$
$p_A(t)$	$A\in M_n(\mathbf{F})$的特征多项式
$\kappa(A)$	非奇异矩阵$A\in M_n$（关于逆相应于给定的矩阵范数）的条件数
$\det A$	$A\in M_n(\mathbf{F})$的行列式
$\oplus$	直和
$\Gamma(A)$	$A\in M_n(\mathbf{F})$的有向图
$\\|\cdot\\|^D$	向量范数$\\|\cdot\\|$的对偶范数
$f^D(\cdot)$	准范数$f(\cdot)$的对偶范数
λ	（通常指）$A\in M_n$的特征值

$\lambda_i(A)$	$A\in M_n$的特征值的集合（谱），如果A是Hermite矩阵，通常指$\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$
$n!$	阶乘，$n(n-1)(n-2)\cdots2\cdot1$
$G(A)$	$A\in M_n$的$Ger\check{s}gorin$区域
$GL(n, F)$	$M_n(F)$中的非奇异矩阵构成的群
$A\circ B$	$A, B\in M_{m,n}(F)$的Hadamard乘积
$\gamma(A)$	素矩阵$A\in M_n$的本原指标
$M(A)$	$A\in M_{m,n}$(F)的指标矩阵
$J_k(\lambda)$	具有特征值λ的k阶$Jordan$块
$\otimes$	Kronecker（张量）乘积
$q_A(t)$	$A\in M_n$(F)的极小多项式
$\lVert\cdot\rVert_1$	$\mathbf{C}^n$的l_1（和）范数；M_n上的l_1矩阵范数
$\lVert\cdot\rVert_2$	$\mathbf{C}^n$的l_2(Euclid)范数；M_n上的l_2（Frobenius)矩阵范数
$\lVert\cdot\rVert_\infty$	$\mathbf{C}^n$上的l_∞(极大）范数；M_n上的l_∞向量范数
$\lVert\cdot\rVert_p$	$\mathbf{C}^n$上的l_p范数
$\lVert\!\vert\cdot\vert\!\rVert_1$	M_n上的极大列和矩阵范数
$\lVert\!\vert\cdot\vert\!\rVert_2$	M_n上的谱矩阵范数
$\lVert\!\vert\cdot\vert\!\rVert_\infty$	M_n的极大行和矩阵范数
$r(A)$	（通常指）$A\in M_n$的数值半径
$\perp$	正交补
per$A$	$A\in M$（F)的积和式
rank$A$	$A\in M_{m,n}$(F)的秩
sgn	排列的正负号函数
$\sigma_i(A)$	$A\in M_{m,n}$的奇异值的集合；通常取$\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_{\min(m,n)}\geqslant0$
$\sigma_1(A)$	$A\in M_{m,n}$的最大奇异值，$\lVert\!\vert A\vert\!\rVert_2$
Span$S$	向量空间的子集S的张成
$\rho(A)$	$A\in M_n$的谱半径
$\sigma(A)$	$A\in M_n$的谱（特征值的集合）
$A(\alpha, \beta)$	由指标集α, β确定的$A\in M_{m,n}$(F)的子矩阵
tr$A$	$A\in M_n(F)$的迹

目　录

第0章 复习与其它

0.0 导 引

本章的目的是简要地编撰许多有用概念或结果，不加证明，其中许多概念和结果直接或间接地为涉及本书主要章节的内容奠定了基础。所编写内容的大部分可能以某种形式包括在线性代数的基础教程中，另外，还选编了一些有用的资料，这些资料往往不易在其它地方找到，或者不宜安排在以后的章节中。这样，在开始学习本书之前，这一章可作为读者的复习提要，必要时也为读者参阅有关资料提供方便。本章还规定了一些基本符号，并给出一些定义；为此，了解本章是有益的。这里，要求读者已经熟悉线性的基本概念和矩阵的基本运算，如矩阵的乘法和加法。

0.1 向量空间

在本书的论述中，虽然一般是含蓄地述及向量空间，但是，向量空间是矩阵理论的基本结构。

0.1.1 纯量域. 构成向量空间的基础是域，或者是具有乘法的纯量集．对于实际应用来说，在通常的加法和乘法运算下，基域几乎总是实数域**R**或复数域**C**（见附录A），但是，它也可能是有理数域，也可能是关于一个特定素数的整数同余类域或一些其它的域。当未指明是哪种域时，就用符号**F**表示域。为了验证一个纯量集是域，它必须在两个指定的二元运算（“加法”和“乘法，”）下封闭；两个运算必须满足结合律和交换律，且在该集合中各有一个单位元；对于所有的元素，在该集合中必须有关于加法运算的逆元素，并且对于除加法单位元（0）以外的所有元素，在该集合中有关于乘法运算的逆元素；同时，乘法运算对加法运算必须满足分配律。

0.1.2 向量空间. 域**F**上的**向量空间**是一些对象（称为向量）的集合V，它在一个二元运算（加法）下封闭，这个运算是结合的和交换的，在集合V中有一个单位元（“0”），且有加法逆元。该集合对用纯量域**F**的元素左乘向量的运算也是封闭的，且有性质：对所有的$a, b \in \mathbf{F}$，以及所有的$x, y \in V$，有$a(x+y)=ax+ay$，$(a+b)x=ax+bx$，$a(bx)=(ab)x$，以及对乘法单位元$e \in \mathbf{F}$，有$ex=x$。

对于给定的域**F**，分量取自**F**的n元组的集合$\mathbf{F}^n$（n是整数），在通常的运算（在$\mathbf{F}^n$中按分量相加）下构成**F**上的一个向量空间．特别地，$\mathbf{R}^n$和$\mathbf{C}^n$是本书的基本向量空间．具有实系数或复系数的（不超过某一指定次数的或任意次数的）多项式集合，以及区间$[a, b] \subset \mathbf{R}$上的实值或复值连续函数，或任意函数的集合也是（**R**上或**C**上）向量空间的例子．当然，在有限维空间$\mathbf{R}^n$与由$[0, 1]$上的实值连续函数组成的无限维向量空间之间，有着本质的差别．

0.1.3 子空间和张成. 向量空间V的子空间U是V的非空子集，它自身正好是同一个

纯量域上的向量空间. 例如 $\{[a, b, 0]^T: a, b \in \mathbf{R}\}$ 是$\mathbf{R}^3$的子空间. 通常，我们用某种关系来定义向量空间V的子空间，如此得到的由V中部分元素组成的集合关于V中的加法是封闭的——例如，$\mathbf{R}^3$中最后一个分量是 0 的所有元素组成的集合. 一般认为，把所得到的集合看成一个子空间比自身看成一个向量空间更有用. 在任何情况下，两个子空间的交还是子空间.

如果S是向量空间V的子集，S的张成是集合$\mathrm{Span}S \equiv \{a_1v_1 + a_2v_2 + \cdots + a_kv_k:\ a_1, a_2, \cdots a_k \in \mathbf{F},\ v_1, v_2 \cdots, v_k \in S,\ k = 1, 2 \cdots\}$. 注意，即使$S$不是子空间，$\mathrm{Span}S$总是子空间. 如果$\mathrm{Span}S = V$，就称$S$张成向量空间$V$.

0.1.4 线性相关和线性无关. 称一个向量空间中的向量组$\{x_1, x_2, \cdots x_k\}$线性相关，指的是，在纯量基域$\mathbf{F}$中存在不全为 0 的系数$a_1, a_2 \cdots a_k$，使得

$$a_1x_1 + a_2x_2 + \cdots + a_kx_k = 0$$

或等价地，某一向量x_i是其余向量的线性组合，其中系数取自$\mathbf{F}$. 例如，$\{[1, 2, 3]^T, [1, 0, -1]^T, [2, 2, 2]^T\}$是$R^3$的中线性相关组. V中的一个子集在$\mathbf{F}$上不线性相关，就说它线性无关. 例如，$\{[1, 2, 3]^T, [1, 0, -1]^T\}$是$\mathbf{R}^3$中的线性无关组. 重要的是要注意这两个概念实质上与向量组有关. 线性无关的任一非空子集线性无关；$\{0\}$是线性相关组；因而，包含 0 向量的任一集合线性相关. 可能一个向量集线性相关，而它的任一真子集是线性无关的.

0.1.5 基. 设S是向量空间V的子集，如果V的每个元素可以表示成S的诸元素（具有纯量基域中的系数）的线性组合，就称S张成V. 例如，$\{[1, 0, 0]^T [0, 1, 0]^T, [0, 0, 1]^T, [1, 0, -1]\}$在$\mathbf{R}$上张成$\mathbf{R}^3$（或在$\mathbf{C}$上张成$\mathbf{C}^3$）. 张成向量空间$V$的一个线性无关组为$V$的一个基. 基虽然不是唯一的，但是基很有用，$V$的每个元素能且只能用一种方式由基来表示，且在该基中再添加任何一个元素或从该基中去掉任一元素，上述性质不再成立. V中的无关组是V的一个基，当且仅当没有真包含它的无关组. 张成V的一个集合是V的一个基，当且仅当它没有真子集仍张成V. 每个向量空间总有一个基.

0.1.6 扩充成一个基. 向量空间V中的任何线性无关组都可以扩充为V的一个基，也就是说，给定V中的线性无关组$\{x_1, x_2 \cdots x_k\}$，存在另外的向量$x_{k+1}, \cdots, x_n, \cdots \in V$，使得$\{x_1 x_2, \cdots, x_n, \cdots\}$是$V$的一个基，把一个已知的无关组扩充为一个基，当然不是唯一的〔例如，可以把第三个分量是非零的任一向量添加到无关组$\{[1, 0, 0]^T, [0, 1, 0]^T\}$中，便得到$\mathbf{R}^3$的一个基〕. 由$[0, 1]$上实值连续函数组成的实向量空间$C[0, 1]$的例子说明，一般地，一个基未必有限；由单项式$\{1, x, x^2, x^3, \cdots\}$组成的无限集是$C[0, 1]$中的无关组.

0.1.7 维数. 如果向量空间V的某个基包含有限个元素，那么所有的基有相同的元素个数，并且称这个公共的数为向量空间的维数. 这时，就说V是有限维的，否则，就说V是无限维的. 无限维的情形（例如，$C[0, 1]$），在任意两个基的元素之间存在一个一一对应. 实向量空间$\mathbf{R}^n$有维数n. 向量空间$\mathbf{C}^n$在域$\mathbf{C}$上有维数n，但在域$\mathbf{R}$上有维数$2n$. 有时称基$\{e_1, e_2, \cdots, e_n\}$为$\mathbf{R}^n$或$\mathbf{C}^n$的标准基，其中$e_i$的第$i$个分量是1，其余分量是0.

0.1.8 同构. 如果U和V是同一个纯量域$\mathbf{F}$上的向量空间，且$f: U \to V$是可逆函数，使得对所有的$x, y \in U$和所有的$a, b \in \mathbf{F}$有$f(ax + by) = af(x) + bf(y)$，则$f$是一个同构，且称$U$和$V$同构（“结构相同”）. 同一个域上的两个有限维向量空间同构，当且仅当它们有相同的维数，于是，域$\mathbf{F}$上的任一n维向量空间同构于$\mathbf{F}^n$，因此，任一n维实向量空间

同构于R^n，而任一n维复向量空间同构于$\mathbf{C}^n$。特别是，如果V是域F上的n维向量空间，具有一个给定的基$\mathscr{B}=\{x_1,\cdots,x_n\}$，那么，因为任一元素$x\in V$可以唯一地写成$x=a_1x_1+\cdots+a_nx_n$，$a_i\in F$，$i=1,2,\cdots,n$，于是，相对于这个基，可以把$x$对应于$n$元组$[x]_{\mathscr{B}}=[a_1,\cdots a_n]^T$。对于任一个基$\mathscr{B}$，映射$x\to[x]_{\mathscr{B}}$是$V$与$\mathbf{F}^n$之间的一个同构。

0.2 矩 阵

这里所研究的对象可以用两种重要的方式来考察：一是把它看成纯量的矩形阵列，一是给每个空间指定一个基，然后把它看作两个向量空间之间的线性变换。

0.2.1 矩形列阵. 一个矩阵是由域F中若干个纯量组成的一个$m\times n$阵列.如果$m=n$，就称矩阵是方阵.F上的所有$m\times n$的矩阵集合用$M_{m,n}(\mathbf{F})$来表示，而$M_{n,n}(\mathbf{F})$简记为$M_n(\mathbf{F})$。最常见的情形是$\mathbf{F}=\mathbf{C}$（复数域），还把$M_n(\mathbf{C})$简记为M_n，$M_{m,n}(\mathbf{C})$简记为$M_{m,n}$.通常用大写字母来表示矩阵。例如，如果

$$A=\begin{bmatrix} 2 & -\frac{3}{2} & 0 \\ -1 & \pi & 4 \end{bmatrix},$$

那么$A\in M_{2,3}(R)$. 一个给定矩阵的子矩阵是位于该矩阵的一些指定的行和列的矩形阵列。例如，$[\pi,\ 4]$是上述A的子矩阵（位于第2行，第2列，第3列）.

0.2.2 线性变换. 设U和V分别是同一个纯量域F上的n维向量空间和m维向量空间；设$\mathscr{B}_U$和$\mathscr{B}_V$分别是U和V的基.我们可以分别用同构$x\to[x]_{\mathscr{B}_U}$和$y\to[y]_{\mathscr{B}_V}$把U和V中的向量表示成F上的n元组和m元组。一个线性变换是一个函数T：$U\to V$，使得对于任意纯量a_1和a_2，以及向量x_1和x_2，都有$T(a_1x_1+a_2x_2)=a_1T(x_1)+a_2T(x_2)$. 一个矩阵$A\in M_{m,n}(F)$可以用下述方式对应于一个线性变换$T$：$U\to V$：向量$y=T(x)$当且仅当$[y]_{\mathscr{B}_V}=A[x]_{\mathscr{B}_U}$。这时就称矩阵$A$表示线性变换$T$（关于基$\mathscr{B}_U$和$\mathscr{B}_V$）；表示矩阵$A$与基的选择有关。在讨论矩阵时，要意识到是在讨论关于特别选定的基下的线性交换，但借助于什么基，一般不必明言。

0.2.3 与一个已知矩阵或线性交换相关联的向量空间. 不失一般性，使F上的n维向量空间与F^n相对应，于是，就把$A\in M_{m,n}(\mathbf{F})$看作从$\mathbf{F}^n$到$\mathbf{F}^m$的线性变换（同时也看作一个阵列）.这样一个线性变换的定义域是$\mathbf{F}^n$；它的值域是$\{y\in\mathbf{F}^m:\ y=Ax$，对所有$x\in\mathbf{F}^n\}$。$A$的零空间是$\{x\in\mathbf{F}^n: Ax=0\}$. A的值域是$\mathbf{F}^m$的子空间，而A的零空间是$\mathbf{F}^n$的子空间。关于这两个的关系式是：

$$n=A\text{的零空间的维数}+A\text{的值域的维数}.$$

0.2.4 矩阵运算. 矩阵加法定义为两个同维阵列按对应元相加，并且用+（“$A+B$”）表示。它对应线性变换的加法（关于相同的基），且继承了从纯量域来的交换性和结合性。零矩阵（所有元全为0的矩阵）是矩阵加法的单位元，并且$M_{m,n}(\mathbf{F})$自身也是F上的向量空间。按通常方式定义的矩阵乘法用AB来表示，它与线性交换的复合相对应。这样，只有当$A\in M_{m,n}(\mathbf{F})$，$B\in M_{p,q}(\mathbf{F})$，且$p=n$时，它才有定义；它是结合的，一般是不交换的。例如，

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\neq\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

但是，当把矩阵限制在M_n(F)的某些有研究价值的子集时，它可以是交换的．矩阵乘法有一个单位元，即形如

$$I=\begin{bmatrix}1 & & 0\\ & 1 & \\ 0 & \ddots & 1\end{bmatrix}$$

的矩阵$I\in M$（F）．这个矩阵以及它的所有纯量倍数（称为纯量矩阵）与M_n(F)中的所有其它矩阵都可交换，并且只有纯量矩阵具有这一性质．矩阵乘法对于矩阵加法是分配的．

这里需要指出，我们总是用符号0表示以下各种术语：零纯量、零向量（所有分量都等于零纯量的向量）和零矩阵（所有的元都等于零纯量）．一般地，上下文将明确它是哪种情形，因而不会引起混淆．我们还用符号I表示任意阶数的单位矩阵．如果可能引起混淆，就指明其阶数．

0.2.5 转置与Hermite伴随． 如果$A\in[a_{ij}]\in M_{m,n}$(F)，A的转置，记作A^T，是$M_{n,m}$(F)中的一个矩阵，它的元是a_{ji}；即将原矩阵行与列调换，反之亦然．例如，

$$\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix}^T=\begin{bmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{bmatrix}.$$

显然，$(A^T)^T=A$。$A\in M_{m,n}$(C)的 Hermite伴随A^*定义为 $A^*=\bar{A}^T$，其中$\bar{A}$表示按分量取共轭．例如，

$$\begin{bmatrix}1+i & 2-i\\ -3 & -2i\end{bmatrix}^*=\begin{bmatrix}1-i & -3\\ 2+i & 2i\end{bmatrix}.$$

转置和Hermite伴随[以及将在(0.5)中讨论的矩阵的逆]都服从倒序律：$[AB]^*=B^*A^*$ 和 $[AB]^T=B^TA^T$，当然要假定乘积有定义．对于乘积的共轭，不存在倒序：$\overline{AB}=\bar{A}\bar{B}$．如果$x$，$y\in M_{n,1}=\mathbf{C}^n$，那末$y^*x$是纯量，并且它的Hermite伴随与它的复共轭相同；因此，$(y^*x)^*=(\overline{y^*x})=y^T\bar{x}$．

0.2.6 矩阵乘法的技巧． 这里，给出几个要反复用到的矩阵乘法的简单性质．

1．如果b_j表示矩阵B的第j列，那么乘积AB的第j列正好是Ab_j．

2．如果a_i表示矩阵A的第i行，那么乘积AB的第i行正好是a_iB．

解释一下，在乘积AB中，左乘以A是乘B的列，而右乘以B是乘A的行．对其中一个因子是对角矩阵的情形，在（0．9．1）中再讨论．

3．如果$A\in M_{m,n}$(F)，且$x\in\mathbf{F}^n$，那么 Ax 是（以x的坐标为系数的）A的各列的线性组合．

4．如果$A\in M_{m,n}$(F)，且$y\in\mathbf{F}^m$，那么y^TA是（以y的坐标为系数的）A的各行的线性组合．

0.3 行列式

只用一个数概括一种多变量现象，这在数学中常常很有用，其中行列式就是一例．行列式只对方阵$A\in M_n$(F)有定义，并且它可以按两种重要的，完全不同的等价方式来描述．我们把$A\in M_n$(F)的行列式记作det$A$．

0.3.1 Laplace展开. 对$A=[a_{ij}]\in M_n(F)$ 的行列式可归纳定义如下。假设行列式在$M_{n-1}(F)$上已定义，设$A_{ij}\in M_{n-1}(F)$表示从$A\in M_n(F)$中划去第i行和第j列后得到的子矩阵。于是，对所有的$i\leqslant n$，$j\leqslant n$，

$$\sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij}=\sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det A_{ij},$$

而这两个公共值就是$\det A$。等式左边是依第i行关于诸子式的Laplace展开式，而右边是依第j列的Laplace展开式[见(0.7.1)]。对于任意选择的行和列，其中任一展开式都得到A的行列式。归纳过程从1×1矩阵开始，定义它的行列式为单个元的值。于是

$$\det[a_{11}]=a_{11},$$

$$\det\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}=a_{11}a_{22}-a_{12}a_{21}$$

$$\det\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31},$$

等等。显然，如果$A\in M_n(C)$，则$\det A^T=\det A$，且$\det A^*=\det A$。

0.3.2 交错和. 受上述低维例子的启发，对于$A=[a_{ij}]\in M_n(F)$，还有

$$\det A=\sum_{\sigma}\operatorname{sgn}\sigma\prod_{i=1}^{n}a_{i\sigma(i)},$$

其中，求和取遍 1，2…，n 的所有$n!$个排列σ，排列σ的“正负号”或“正负号函数”$\operatorname{sgn}\sigma$是+1或-1，取决于由 1，2，…，n 开始到得到排列σ所需对换（或两两交换）的最小数是偶数还是奇数。于是，每个乘积

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$$

都在行列式中出现，如果σ是偶排列，则在乘积前冠以+号，如果σ是奇排列，就冠以-号。

如果系数$\operatorname{sgn}\sigma$用某些其它的函数来代替，那么所谓的广义矩阵函数就取代了$\det A$。一个例子是$\operatorname{per}A$，称为A的积和式（*permanent*），其中$\operatorname{sgn}\sigma$被恒等于1的函数所代替。

0.3.3 初等变换。 有三种简单的基本变换，人们常常可以利用这些变换把任一个矩阵化简成与该矩阵相抵的、唯一简单的形式（即标准形），以便用它来解线性方程组，计算行列式，矩阵求逆和研究矩阵的秩，等等。我们集中讨论关于行的变换，它们是如下三种变换。第一种：交换矩阵的两行。

交换第i行和第j行可以经左乘以矩阵

$$\left[\begin{array}{ccc|c|ccc|c|cc}
1&&&&&&&&&\\
&\ddots&&&&&&&&\\
&&1&&&&&&&\\
\hline
&&&0&&&&1&&\\
\hline
&&&&1&&&&&\\
&&&&&\ddots&&&&\\
&&&&&&1&&&\\
\hline
&&&1&&&&0&&\\
\hline
&&&&&&&&\ddots&\\
&&&&&&&&&1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ \text{— } i\ \text{行}\\ \\ \\ \\ \text{— } j\ \text{行}\\ \\ \\
\end{array}$$

i 列　　j 列

来实现，其中，在 i，j 位置和 j，i 位置上的两个非对角元是 1，而所有未注明的元都是 0。

第二种：用一个非零纯量乘某一行。

用一个纯量 c 乘 A 的第 i 行可以经左乘以矩阵

$$\begin{array}{cl}
\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ -\,i\text{行} \\ \\ \\ \\ \end{array} \\
\quad\quad\quad i\text{列} &
\end{array}$$

来实现，其中纯量 c 出现在 i，i 位置。

第三种：把某一行的纯量倍数加到另一行。

用 c 乘第 i 行加到第 j 行相当于把 A 左乘以矩阵

$$\begin{array}{cl}
\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & c & & \\ & & & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ -\,j\text{行} \\ \\ \end{array} , \\
\quad i\text{列} &
\end{array}$$

其中纯量 c 出现在 j，i 位置。注意，上述每个作初等变换的矩阵，正好是把相应的初等变换施于单位矩阵 I 的结果。

第一种初等变换在行列式上的作用是将行列式乘以 -1；第二种变换的作用是将它乘以纯量 c；第三种变换不改变行列式。由此可知，如果一个矩阵有一个零行，或有两行相关，或任意 k 行相关，那么它的行列式为零。一个矩阵的行列式是零，当且仅当它的诸行的一个子集线性相关。

0.3.4 行简化梯形阵. 对每个 $A\in M_{m,n}(\mathbf{F})$，在 $M_{m,n}(\mathbf{F})$ 中存在一个标准形，A 的行简化梯形（RREF），它可以经（不唯一的）一系列初等变换得到．许多矩阵有相同的RREF，每个矩阵不管经一列什么样的初等变换而得到，它只有一个RREF。RREF的定义是：

(a) 各非零行的第一个非零元是 1；

(b) 具有上述首元 1 的列的所有其它元都为零；

(c) 全由零元组成的行出现在矩阵的底部；

(d) 诸首元 1 位于从左到右的"阶梯型"之中，即下一行的首元 1 必须出现在其上一行的首元 1 的右边。

例如，

$$\begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & \pi \\ 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

是RREF．$A\in M_n$ 的行列式是非零的，当且仅当它的RREF是单位矩阵