全国工程专业学位研究生教育国家级规划教材

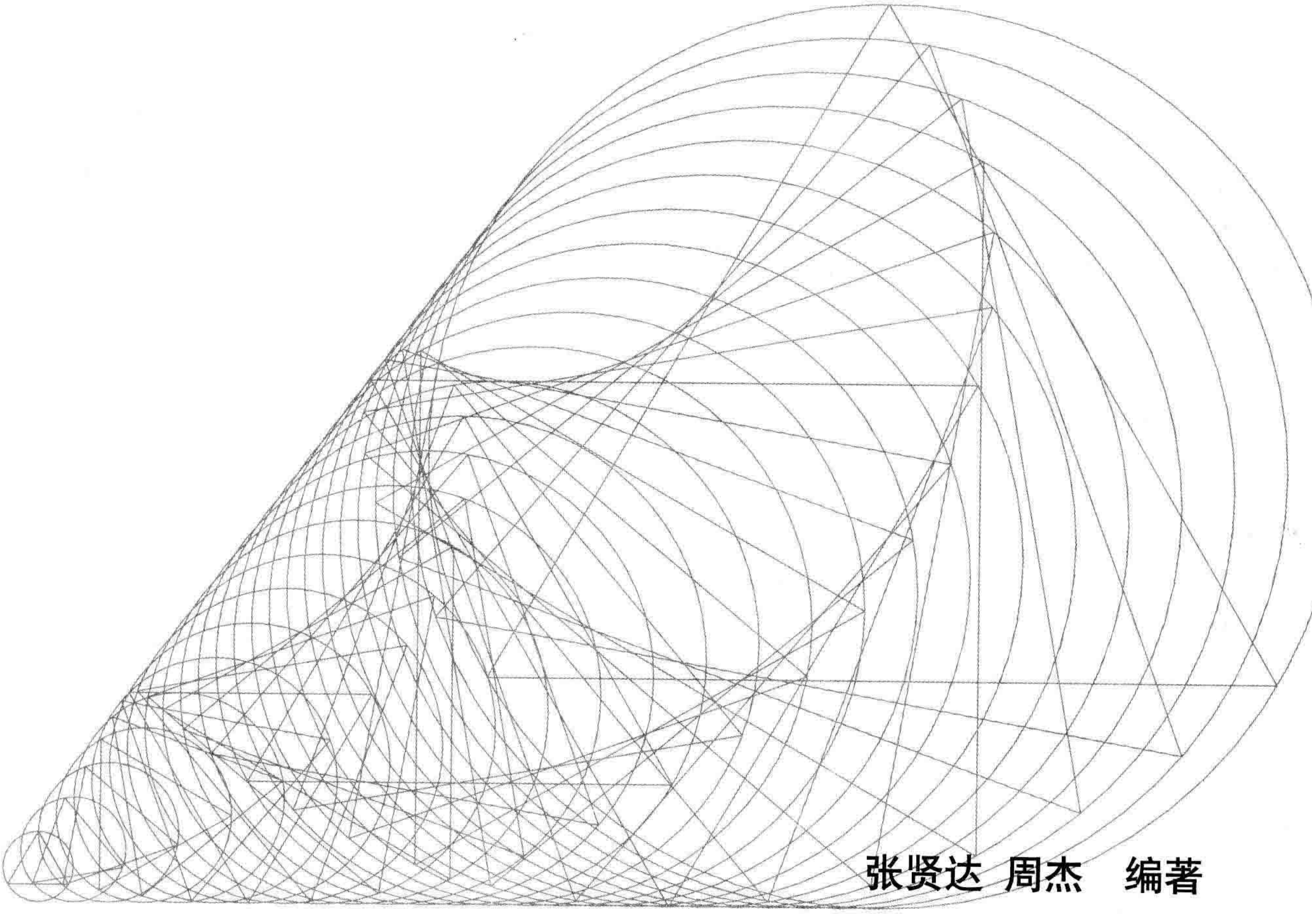
张贤达 周杰 编著

矩阵论及其工程应用

http://www.tup.com.cn

清华大学出版社

全国工程专 国家级规划教材



张贤达 周杰 编著

矩阵论及其工程应用

清华大学出版社
北京

内 容 简 介

本书主要为全国工程硕士研究生学位课程“矩阵论”的教学所编写。针对各工程领域对矩阵论相关内容的实际应用需求，确定了教材编写的基本思想是“强调问题的工程背景、注重基本概念和原理、重点介绍常用的矩阵论方法、淡化理论推导、突出应用案例”。

主要内容包括：代数与矩阵的基本概念、特殊矩阵、矩阵的相似化简与特征分析、奇异值分析、子空间分析、广义逆及矩阵方程求解、矩阵微分与梯度分析等。本书旨在主要介绍：(1)矩阵的基本理论和方法；(2)主要结果的求解思路；(3)矩阵的应用方法及有关应用案例。

本书适用于各相关工程领域的工程硕士研究生学位课程“矩阵论”作教学用书，也可作为工科各专业的大学本科生和研究生矩阵论课程的教学参考用书，还可供从事相关研究和开发工作的工程技术人员自学和参考。

图书在版编目(CIP)数据

矩阵论及其工程应用/张贤达，周杰编著. --北京：清华大学出版社，2015
全国工程专业学位研究生教育国家级规划教材
ISBN 978-7-302-41035-5

Ⅰ.①矩…　Ⅱ.①张…②周…　Ⅲ.①矩阵论-研究生-教材　Ⅳ.①O151.21

中国版本图书馆 CIP 数据核字(2015)第 169173 号

责任编辑：刘　颖
封面设计：何凤霞
责任校对：刘玉霞
责任印制：王静怡

出版发行：清华大学出版社
　　网　　址：http://www.tup.com.cn，http://www.wqbook.com
　　地　　址：北京清华大学学研大厦 A 座　　**邮　　编**：100084
　　社 总 机：010-62770175　　**邮　　购**：010-62786544
　　投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
　　质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：北京鑫海金澳胶印有限公司
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　张**：16.75　　**字　　数**：406 千字
版　　次：2015 年 9 月第 1 版　　**印　　次**：2015 年 9 月第1 次印刷
印　　数：1～3000
定　　价：38.00 元

产品编号：057904-01

前言

随着我国社会和经济发展进入新的时期，高层次工程专业人才的需求越来越大。经过认真研究与分析，全国工程专业学位研究生教育指导委员会提出了工程硕士课程教学改革设想和指导性意见，即旨在提高工程硕士研究生工程应用能力和职业能力，推动工程硕士的课程建设与教学改革，为社会培养更多高素质的应用型人才。针对工程硕士数学课程建设和教学内容改革，教指委也提出了指导性意见，希望工程硕士应具备运用数学方法和计算工具解决工程领域实际问题的能力，要求数学课程教学的改革与创新要紧紧围绕这一核心目标，注重数学在工程中的应用案例教学，加强工程硕士研究生利用数学方法和计算机工具解决实际工程问题的能力培养。

矩阵论作为工程硕士研究生的一门重要的数学课程，在很多工程领域都有着广泛的应用。根据教指委的改革思路和总要求，同时考虑到各相关工程领域课程教学的实际需求，本教材以介绍矩阵论中的基本理论和实用算法为主线，强调问题的工程背景，注重基本概念和原理，重点介绍常用的矩阵论方法和应用，淡化理论推导。这也是本教材与目前已有的其他矩阵论教材之间的最大区别。特别需要说明的是，矩阵理论和方法具有比较强的抽象性，往往使得工程硕士研究生难以理解。为了帮助学生更好地掌握相关的矩阵论方法及其应用，编者在本教材中选入了十多个经典的工程应用例子，从应用背景的介绍出发，引入所选用的矩阵论相关算法，分析了其应用的效果，以有助于读者能够站在应用的角度全面理解矩阵论相关算法的精髓与奥妙，培养工程应用意识，提高解决工程领域实际问题的能力。

教材的主要内容包括：代数与矩阵的基本概念、特殊矩阵、矩阵的相似化简、特征分析、奇异值分析、子空间分析、广义逆及矩阵方程求解、矩阵微分与梯度分析等。本书的主要目的是介绍：(1) 矩阵的基本理论和方法；(2) 主要结果的求解思路；(3) 矩阵的应用方法。建议任课教师在课程讲授中注重实际应用能力的培养，可以结合课程布置 1~2 个大作业或综合训练，以加强理论联系实践，培养学生运用矩阵论解决工程实际问题的能力。

该教材主要是针对全国工程硕士相关工程领域专业学位研究生的矩阵论课程编写的，适用于相关的工程领域包括：机械工程、材料工程、电气工程、电子与通信工程、控制工程、软件工程、建筑与土木工程、水利工程、测绘工程、地质工程、矿业工程、冶金工程、石油工程、纺织工程、轻工技术与工程、交通运输工程、船舶与海洋工程、安全工程、兵器工程、航空工程、农业工程、林业工程、环境工程、化工工程、生物医药工程、食品工程、车辆工程、工业工程、工业设计工程、生物工程、项目管理、物流工程等。同时，该教材也适于作为工科各专业的本科生和研究生的矩阵论课程教学用书或参考教材，还可供从事相关研究工作的工程技术人员参考之用。由于不同高校和不同学科的培养方案有着很大的差别，建议任课

教师根据学时安排和学科领域的需求选择相关内容讲授。我们也根据教材各章节内容在主要工程领域中的应用程度，在附录中给出了各章节的重要性分级建议和学时分配建议，供任课教师和选课学时参考。

在本书编写过程中，得到全国工程专业学位研究生教指委的领导和专家的大力支持与资助，特别是教指委副主任陈子辰教授、秘书处高彦芳主任和沈岩副主任提出了很多指导性意见；教指委数学组的专家华中科技大学齐欢教授、解放军信息工程大学韩中庚教授、重庆大学易正俊教授、武汉大学李大美教授等也都对该教材提出了很多建设性意见；各相关工程领域的专家也都从不同的工程领域实际提出了很多好的建议．在该教材的编写和编辑出版过程中，得到了清华大学出版社理工分社张秋玲社长与刘颖编辑的大力支持和帮助．在准备应用案例的过程中，清华大学自动化系研究生陈纯杰、朱海洋、雷磊、安邦、肖驰洋、马晨光等给予了很多支持和帮助。在此，编者谨以最诚挚的心情，对所有为该教材的编写出版提供帮助和支持的领导、专家和学者一并表示衷心的感谢。

鉴于编者的水平有限，教材中定有错漏和不当之处，恳请各位专家、同行和热心的读者不吝赐教．

张贤达　周杰　谨识于清华大学

2015 年 5 月

目录

第 1 章　代数与矩阵基础 1
1.1　代数与矩阵的基本概念 1
1.1.1　代数基本概念 1
1.1.2　矩阵与向量 3
1.1.3　矩阵的基本运算 4
1.2　矩阵的初等变换 6
1.2.1　初等行变换与阶梯型矩阵 7
1.2.2　初等行变换的两个应用 9
1.2.3　初等列变换 12
1.3　矩阵的性能指标 13
1.3.1　矩阵的行列式 13
1.3.2　矩阵的二次型 14
1.3.3　矩阵的特征值 14
1.3.4　矩阵的迹 15
1.3.5　矩阵的秩 16
1.4　内积与范数 18
1.4.1　向量的内积与范数 18
1.4.2　矩阵的内积与范数 22
1.5　矩阵和向量的应用案例 23
1.5.1　模式识别与机器学习中向量的相似比较 23
1.5.2　人脸识别的稀疏表示 25
本章小结 26
习题 26

第 2 章　特殊矩阵 29
2.1　置换矩阵、互换矩阵与选择矩阵 29
2.1.1　Hermitian 矩阵 29
2.1.2　置换矩阵与互换矩阵 30
2.1.3　广义置换矩阵与选择矩阵 32
2.1.4　广义置换矩阵在鸡尾酒会问题中的应用案例 33
2.2　正交矩阵与酉矩阵 34
2.3　三角矩阵 36

2.4 Vandermonde 矩阵与 Fourier 矩阵 37
2.4.1 Vandermonde 矩阵 38
2.4.2 Fourier 矩阵 40
2.5 Hadamard 矩阵 41
2.6 Toeplitz 矩阵与 Hankel 矩阵 43
2.6.1 Toeplitz 矩阵 43
2.6.2 Hankel 矩阵 44
本章小结 45
习题 45

第 3 章 矩阵的相似化简与特征分析 48
3.1 特征值分解 48
3.1.1 矩阵的特征值分解 48
3.1.2 特征值的性质 50
3.1.3 特征向量的性质 52
3.1.4 特征值分解的计算 53
3.2 矩阵与矩阵多项式的相似化简 54
3.2.1 矩阵的相似变换 54
3.2.2 矩阵的相似化简 57
3.2.3 矩阵多项式的相似化简 60
3.3 多项式矩阵及相抵化简 63
3.3.1 多项式矩阵与相抵化简的基本理论 64
3.3.2 多项式矩阵的相抵化简方法 66
3.3.3 Jordan 标准型与 Smith 标准型的相互转换 69
3.4 Cayley-Hamilton 定理及其应用 74
3.4.1 Cayley-Hamilton 定理 74
3.4.2 在矩阵函数计算中的应用 75
3.5 特征分析的应用 78
3.5.1 Pisarenko 谐波分解 78
3.5.2 主成分分析 81
3.5.3 基于特征脸的人脸识别 82
3.6 广义特征值分解 87
3.6.1 广义特征值分解及其性质 87
3.6.2 广义特征值分解算法 89
3.6.3 广义特征分析的应用 90
3.6.4 相似变换在广义特征值分解中的应用 92
本章小结 95
习题 95

第 4 章 奇异值分析 …… 100
4.1 数值稳定性与条件数 …… 100
4.2 奇异值分解 …… 102
4.2.1 奇异值分解及其解释 …… 102
4.2.2 奇异值的性质 …… 105
4.2.3 矩阵的低秩逼近 …… 107
4.2.4 奇异值分解的数值计算 …… 108
4.3 乘积奇异值分解 …… 111
4.3.1 乘积奇异值分解问题 …… 111
4.3.2 乘积奇异值分解的精确计算 …… 112
4.4 奇异值分解的工程应用案列 …… 114
4.4.1 静态系统的奇异值分解 …… 114
4.4.2 图像压缩 …… 115
4.4.3 数字水印 …… 119
4.5 广义奇异值分解 …… 123
4.5.1 广义奇异值分解的定义与性质 …… 123
4.5.2 广义奇异值分解的实际算法 …… 125
4.5.3 广义奇异值分解的应用例子 …… 128
本章小结 …… 129
习题 …… 129

第 5 章 子空间分析 …… 131
5.1 子空间的一般理论 …… 131
5.1.1 子空间的基 …… 131
5.1.2 无交连、正交与正交补 …… 133
5.1.3 子空间的正交投影与夹角 …… 135
5.2 列空间、行空间与零空间 …… 137
5.2.1 矩阵的列空间、行空间与零空间 …… 137
5.2.2 子空间基的构造：初等变换法 …… 140
5.2.3 基本空间的标准正交基构造：奇异值分解法 …… 142
5.3 信号子空间与噪声子空间 …… 144
5.4 快速子空间跟踪与分解 …… 147
5.4.1 投影逼近子空间跟踪 …… 147
5.4.2 快速子空间分解 …… 152
5.5 子空间方法的应用 …… 156
5.5.1 多重信号分类 …… 156
5.5.2 子空间白化 …… 157
5.5.3 盲信道估计的子空间方法 …… 158
本章小结 …… 164
习题 …… 164

第 6 章 广义逆与矩阵方程求解 167
6.1 广义逆矩阵 167
6.1.1 满列秩和满行秩矩阵的广义逆矩阵 167
6.1.2 Moore-Penrose 逆矩阵 168
6.2 广义逆矩阵的求取 172
6.2.1 广义逆矩阵与矩阵分解的关系 172
6.2.2 Moore-Penrose 逆矩阵的数值计算 173
6.3 最小二乘方法 175
6.3.1 普通最小二乘方法 176
6.3.2 数据最小二乘 177
6.3.3 Tikhonov 正则化方法 178
6.3.4 交替最小二乘方法 180
6.4 总体最小二乘 184
6.4.1 总体最小二乘问题 184
6.4.2 总体最小二乘解 185
6.4.3 总体最小二乘解的性能 190
6.5 约束总体最小二乘 190
6.5.1 约束总体最小二乘方法 190
6.5.2 最小二乘方法及其推广的比较 192
6.6 稀疏矩阵方程求解 193
6.6.1 L_1 范数最小化 194
6.6.2 贪婪算法 195
6.6.3 同伦算法 197
6.7 三个应用案例 198
6.7.1 恶劣天气下的图像恢复 198
6.7.2 总体最小二乘法在确定地震断层面参数中的应用 202
6.7.3 谐波频率估计 204
本章小结 209
习题 210

第 7 章 矩阵微分与梯度分析 213
7.1 Jacobian 矩阵与梯度矩阵 213
7.1.1 Jacobian 矩阵 213
7.1.2 梯度矩阵 214
7.1.3 梯度计算 215
7.2 一阶实矩阵微分与 Jacobian 矩阵辨识 217
7.2.1 一阶实矩阵微分 217
7.2.2 标量函数的 Jacobian 矩阵辨识 219
7.2.3 矩阵微分的应用举例 226

7.3 实变函数无约束优化的梯度分析 227

7.3.1 单变量函数 $f(x)$ 的平稳点与极值点 228

7.3.2 多变量函数 $f(\boldsymbol{x})$ 的平稳点与极值点 230

7.3.3 多变量函数 $f(\boldsymbol{X})$ 的平稳点与极值点 231

7.3.4 实变函数的梯度分析 233

7.4 平滑凸优化的一阶算法 235

7.4.1 凸集与凸函数 235

7.4.2 无约束凸优化的一阶算法 237

7.5 约束凸优化算法 243

7.5.1 标准约束优化问题 243

7.5.2 极小 - 极大化与极大 - 极小化方法 244

7.5.3 Nesterov 最优梯度法 248

本章小结 250

习题 250

参考文献 252

第 1 章

代数与矩阵基础

很多工程问题都可以通过数学建模转化成线性方程组，而矩阵是描述和求解线性方程组最基本和最常用的数学工具。本章将介绍矩阵的基本数学运算和重要性质。

1.1 代数与矩阵的基本概念

1.1.1 代数基本概念

群、环、域是代数的基本概念，但限于篇幅，不予赘述。这里主要介绍线性空间和线性映射 (线性变换)。

在抽象代数中，域是一种可进行加、减、乘、除四则运算的代数结构。域的概念是数域及四则运算的推广。

首先，引入几个基本符号：P 代表数域，$\mathbb{R}$ 代表实数域，$\mathbb{C}$ 表示复数域，$\mathbb{Z}$ 为整数域。

一个 m 维列向量定义为

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_m\end{bmatrix} \tag{1.1.1}$$

若其元素 x_i 取实数，即 $x_i\in\mathbb{R}$，则称其为 m 维实 (数) 向量，并记作 $\boldsymbol{x}\in\mathbb{R}^{m\times 1}$，或者简记为 $\boldsymbol{x}\in\mathbb{R}^m$。类似地，若 $x_i\in\mathbb{C}$，则称其为 m 维复向量，并记作 $\boldsymbol{x}\in\mathbb{C}^m$。

一个 m 维行向量定义为 $\boldsymbol{x}=[x_1,x_2,\cdots,x_m]$，记作 $\boldsymbol{x}\in\mathbb{R}^{1\times m}$ 或 $\boldsymbol{x}\in\mathbb{C}^{1\times m}$。为了书写的方便，常将一个 m 维列向量写成 m 维行向量的转置形式，即 $\boldsymbol{x}=[x_1,x_2,\cdots,x_m]^{\mathrm{T}}$。

一个 $m\times n$ 矩阵定义为

$$\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}=[a_{ij}]_{i=1,j=1}^{m,n} \tag{1.1.2}$$

若其元素 $a_{ij}\in\mathbb{R}$，则称其为 $m\times n$ 实矩阵，用符号表示为 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$。类似地，$\boldsymbol{A}\in\mathbb{C}^{m\times n}$ 表示 $\boldsymbol{A}$ 是一个 $m\times n$ 复矩阵。

$m \times n$ 矩阵可以利用其列向量 $\boldsymbol{a}_j = [a_{1j}, a_{2j}, \cdots, a_{mj}]^{\mathrm{T}} (j = 1, 2, \cdots, n)$ 表示为 $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n]$。

定义 1.1.1　线性空间是指在一个集合 S 上定义了加法 (且对加法是交换群) 和数乘运算，且数乘满足下列线性规则，即 $\forall \alpha, \beta \in P, \forall x, y \in S$，均有

$$\alpha(x+y) = \alpha x + \alpha y$$

$$(\alpha+\beta)x = \alpha x + \beta x$$

$$\alpha(\beta x) = (\alpha\beta)x$$

例 1.1.1　$\mathbb{R}^n$ 是线性空间，其中“加法”运算定义为

$$\boldsymbol{a} + \boldsymbol{b} = [a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n]^{\mathrm{T}}$$

数乘运算为

$$\lambda \boldsymbol{a} = [\lambda a_1, \lambda a_2, \cdots, \lambda a_n]^{\mathrm{T}}$$

其中 $\boldsymbol{a} = [a_1, a_2, \cdots, a_n]^{\mathrm{T}}$ 和 $\boldsymbol{b} = [b_1, b_2, \cdots, b_n]^{\mathrm{T}}$。

函数有定义域 (domain) 与值域 (range)。定义域是函数自变量所有可取值的集合；值域则是由定义域中一切元素所能产生的所有函数值的集合。

一个 $m \times n$ 矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 所对应的线性变换 $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{y} (\boldsymbol{y} \in \mathbb{C}^n)$ 的值域定义为 $\mathrm{Range}(\boldsymbol{A}) = \{\boldsymbol{x} | \boldsymbol{x} = \boldsymbol{A}\boldsymbol{y}, \boldsymbol{y} \in \mathbb{C}^n\}$，而零空间 (null space) 则是矩阵方程 $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}$ 的所有解向量 $\boldsymbol{v}$ 的集合，也称为 $\boldsymbol{A}$ 的核或核空间，常用数学符号表示为 $\mathrm{Null}(\boldsymbol{A}) = \{\boldsymbol{v} \in V : \boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}\}$。值域 $\mathrm{Range}(\boldsymbol{A})$ 也写作 $\mathrm{Span}(\boldsymbol{A})$，或者简记为 $R(\boldsymbol{A})$。

例 1.1.2　给定 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ ($m \times n$ 复系数矩阵)，则该矩阵的值域 $R(\boldsymbol{A}) = \{\boldsymbol{x} | \boldsymbol{x} = \boldsymbol{A}\boldsymbol{y}, \boldsymbol{y} \in \mathbb{C}^n\}$ 和零空间 $N(\boldsymbol{A}) = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$ 都是线性空间，但 $R(\boldsymbol{A}) \subset \mathbb{C}^m$，而 $N(\boldsymbol{A}) \subset \mathbb{C}^n$。

例 1.1.3　次数 $\leqslant n$ 的复系数多项式的全体 $C_n[\lambda] = \{a | a = a_n\lambda^n + \cdots + a_1\lambda + a_0,\ a_i \in \mathbb{C}\}$ 是线性空间；但次数 $= n$ 的复系数多项式的集合不是线性空间。

思考：最小的线性空间是什么？

S 和 Q 同为线性空间，且 $Q \subset S$, 则称 Q 为 S 的 线性子空间。

例 1.1.4　$\boldsymbol{A} \in \mathbb{C}^{m\times n}$，则 $R(\boldsymbol{A})$ 是 $\mathbb{C}^m$ 的子空间，而 $N(\boldsymbol{A})$ 是 $\mathbb{C}^n$ 的子空间。

S 是线性空间，元素组 $\{x_1, x_2, \cdots, x_m\} \subset S$ 是线性相关的，系指存在不全为零的系数 $a_i \in P$，使 $\sum\limits_{i=1}^{m} a_i x_i = 0$。反之，称 $\{x_1, x_2, \cdots, x_m\}$ 是线性无关的。当 $\{x_1, x_2, \cdots, x_m\}$ 线性无关时，$\sum\limits_{i=1}^{m} a_i x_i = 0 \Rightarrow a_i = 0, \quad i = 1, 2, \cdots, m$。

若在线性空间 S 中存在 n 个向量线性无关，而任何 $n+1$ 个向量均线性相关，则称 S 的维数为 n，记为 $\dim(S) = n$。

S 是域 P 上的线性空间，元素组 $\{x_1, x_2, \cdots, x_n\}$ 称为 S 的一组基，是指：

① $\{x_1, x_2, \cdots, x_n\}$ 线性无关；

② $\mathrm{Span}[x_1, x_2, \cdots, x_n] \overset{\mathrm{def}}{=} \left\{ \sum\limits_{i=1}^{n} a_i x_i, a_i \in P \right\} = S$。

注：若 $\forall x \in S, x = \sum\limits_{i=1}^{n} a_i x_i$，则坐标向量 $\boldsymbol{a} = [a_1, a_2, \cdots, a_n]^{\mathrm{T}}$ 唯一。

因为线性空间也是集合，因此可以定义交集和并集。线性空间之间还可以定义“和”运算以及“直 (接) 和”运算。线性空间的交、并、和、直和四种运算的定义如下：

$$T \cap V = \{\ x |\ x \in T\ 且\ \ x \in V\}$$
$$T \cup V = \{\ x |\ x \in T\ 或\ \ x \in V\}$$
$$T + V = \{\ x |\ x =\ y +\ z,\ y \in T,\ z \in V\}$$
$$T \oplus V\ 是指\ T + V,\ 当\ T \cap V = \{0\}$$

思考：以上哪种集合还是线性空间？

映射 $\sigma : S \to T$ 称为线性映射 (线性算子)，是指它满足

$$\sigma(\boldsymbol{a}+\boldsymbol{b}) = \sigma(\boldsymbol{a}) + \sigma(\boldsymbol{b})$$
$$\sigma(\lambda\boldsymbol{a}) = \lambda\sigma(\boldsymbol{a})$$

1.1.2 矩阵与向量

科学和工程中的很多问题都可以通过数学建模转化成一个线性方程组

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}\right\} \tag{1.1.3}$$

它使用 m 个方程描述 n 个未知量之间的线性关系。

线性方程组 (1.1.3) 的简洁表示形式为

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{1.1.4}$$

式中

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]_{i=1,j=1}^{m,n} \tag{1.1.5}$$

为 $m \times n$ 矩阵，是一个按照长方阵列排列的实数或复数集合，其中 a_{ij} 表示矩阵 $\boldsymbol{A}$ 的第 i 行、第 j 列元素，简称第 (i,j) 个元素。

在工程问题的数学建模中，矩阵 $\boldsymbol{A}$ 往往是某个物理系统的符号表示。其中，$\boldsymbol{A}$ 代表一线性系统，$\boldsymbol{x}$ 和 $\boldsymbol{b}$ 则分别表示该系统的输入激励 (不可观测) 和输出响应 (可观测)。于是，矩阵方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的求解问题便可叙述为：根据已知的线性系统参数和输出观测值，求未知的输入激励。

当 $m = n$ 时，矩阵 $\boldsymbol{A}$ 称为正方矩阵 (square matrix)；若 $m < n$，则矩阵 $\boldsymbol{A}$ 形象地称为宽矩阵 (broad matrix)；而当 $m > n$ 时，便称矩阵 $\boldsymbol{A}$ 为高矩阵 (tall matrix)。

科学和工程中遇到的向量可分为以下 3 种[59]：

(1) 物理向量 泛指既有幅值，又有方向的物理量，如速度、加速度、位移等。

(2) 几何向量　是物理向量的可视化表示，常用带方向的 (简称“有向”) 线段表示。这种有向线段称为几何向量。例如，$\boldsymbol{v}=\overrightarrow{AB}$ 表示一有向线段，其起点为 A，终点为 B。

(3) 代数向量　几何向量的代数形式表示。例如，若平面上的几何向量 $\boldsymbol{v}=\overrightarrow{AB}$ 的起点坐标为 $A=(a_1,a_2)$，终点坐标为 $B=(b_1,b_2)$，则该几何向量可以用代数形式表示为 $\boldsymbol{v}=\begin{bmatrix} b_1-a_1 \\ b_2-a_2 \end{bmatrix}$。这种用代数形式表示的几何向量称为代数向量。

根据元素取值种类的不同，代数向量又可分为以下 3 种：

(1) 常数向量　向量的元素全部为实常数或者复常数，例如 $\boldsymbol{a}=[1,5,4]^{\mathrm{T}}$ 等。

(2) 函数向量　向量的元素包含了函数值，例如 $\boldsymbol{x}=[1,x^2,\cdots,x^n]^{\mathrm{T}}$ 等。

(3) 随机向量　向量的元素为随机变量，如 $\boldsymbol{x}(n)=[x_1(n),x_2(n),\cdots,x_m(n)]^{\mathrm{T}}$，其中 $x_1(n),x_2(n),\cdots,x_m(n)$ 是 m 个随机过程或随机信号。

图 1.1.1 归纳了向量的分类。

$$
\text{向量}\begin{cases}\text{物理向量}\\ \text{代数向量}\begin{cases}\text{常数向量}\\ \text{函数向量}\\ \text{随机向量}\end{cases}\\ \text{几何向量}\end{cases}
$$

图 1.1.1　向量的分类

工程应用中遇到的往往是物理向量，几何向量是物理向量的可视化表示，而代数向量则可看作是物理向量的运算化工具。

一个 $n\times n$ 正方矩阵 $\boldsymbol{A}$ 的主对角线是指从左上角到右下角沿 $i=j,\ j=1,2,\cdots,n$ 相连接的线段。与主对角线平行的对角线称为次对角线。位于主对角线上的元素称为 $\boldsymbol{A}$ 的对角元素，它们是 $a_{ii},i=1,2,\cdots,n$。

主对角线以外元素全部为零的 $n\times n$ 矩阵称为对角矩阵，记作

$$
\boldsymbol{D}=\mathrm{diag}(d_{11},d_{22},\cdots,d_{nn}) \tag{1.1.6}
$$

若对角矩阵主对角线元素全部等于 1，则称其为单位矩阵，用符号 $\boldsymbol{I}_{n\times n}$ 示之。所有元素为零的 $m\times n$ 矩阵称为零矩阵，记为 $\boldsymbol{O}_{m\times n}$。

一个全部元素为零的向量称为零向量。为了书写的简洁，单位矩阵、零矩阵和零向量分别表示为 $\boldsymbol{I},\boldsymbol{O}$ 和 $\mathbf{0}$。

1.1.3　矩阵的基本运算

矩阵的基本运算包括矩阵的转置、共轭、共轭转置、加法、乘法和求逆。

1. 矩阵的转置

若 $\boldsymbol{A}=[a_{ij}]$ 是一个 $m\times n$ 矩阵，则 $\boldsymbol{A}$ 的转置记作 $\boldsymbol{A}^{\mathrm{T}}$，它是一个 $n\times m$ 矩阵，定义为

$$\boldsymbol{A}^{\mathrm{T}}=\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \tag{1.1.7}$$

矩阵 $\boldsymbol{A}$ 的复数共轭 $\boldsymbol{A}^*$ 仍然是一个 $m \times n$ 矩阵，定义为

$$\boldsymbol{A}^{*}=\begin{bmatrix} a_{11}^* & a_{12}^* & \cdots & a_{1n}^* \\ a_{21}^* & a_{22}^* & \cdots & a_{2n}^* \\ \vdots & \vdots & & \vdots \\ a_{m1}^* & a_{m2}^* & \cdots & a_{mn}^* \end{bmatrix} \tag{1.1.8}$$

矩阵 $\boldsymbol{A}$ 的 (复) 共轭转置 $\boldsymbol{A}^{\mathrm{H}}$ 是一个 $n \times m$ 矩阵，定义为

$$\boldsymbol{A}^{\mathrm{H}}=\begin{bmatrix} a_{11}^* & a_{21}^* & \cdots & a_{m1}^* \\ a_{12}^* & a_{22}^* & \cdots & a_{m2}^* \\ \vdots & \vdots & & \vdots \\ a_{1n}^* & a_{2n}^* & \cdots & a_{mn}^* \end{bmatrix} \tag{1.1.9}$$

共轭转置又叫 Hermitian 转置或 Hermitian 共轭。

满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$ 的正方实矩阵和 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$ 的正方复矩阵分别称为对称矩阵和 Hermitian 矩阵 (复共轭对称矩阵)。共轭转置与转置之间存在关系

$$\boldsymbol{A}^{\mathrm{H}}=(\boldsymbol{A}^*)^{\mathrm{T}}=(\boldsymbol{A}^{\mathrm{T}})^* \tag{1.1.10}$$

2. 矩阵的加法与乘法

两个 $m \times n$ 矩阵 $\boldsymbol{A}=[a_{ij}]$ 和 $\boldsymbol{B}=[b_{ij}]$ 之和记作 $\boldsymbol{A}+\boldsymbol{B}$，定义为 $[\boldsymbol{A}+\boldsymbol{B}]_{ij}=a_{ij}+b_{ij}$。

矩阵的加法服从下面的运算法则：

(1) 加法交换律 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$

(2) 加法结合律 $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$

矩阵的乘法分为以下三种：

矩阵与标量的乘法　$m\times n$ 矩阵 $\boldsymbol{A}=[a_{ij}]$ 与标量 α 的乘积的元素定义为 $[\alpha\boldsymbol{A}]_{ij}=\alpha a_{ij}$。

矩阵与向量的乘法　$m \times n$ 矩阵 $\boldsymbol{A}=[a_{ij}]$ 与 $r \times 1$ 向量 $\boldsymbol{x}=[x_1,x_2,\cdots,x_r]^{\mathrm{T}}$ 的乘积 $\boldsymbol{Ax}$ 只有当 $n=r$ 时才存在，它是一个 $m \times 1$ 向量，定义为

$$[\boldsymbol{Ax}]_i=\sum_{j=1}^{n} a_{ij}x_j, \quad i=1,2,\cdots,m$$

两个矩阵的乘法　$m\times n$ 矩阵 $\boldsymbol{A}=[a_{ij}]$ 与 $r\times s$ 矩阵 $\boldsymbol{B}=[b_{ij}]$ 的乘积 $\boldsymbol{AB}$ 只有当 $n=r$ 时才存在，它是一个 $m \times s$ 矩阵，定义为

$$[\boldsymbol{AB}]_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}, \quad i=1,2,\cdots,m;\ j=1,2,\cdots,s$$

矩阵的乘积服从下面的运算法则：

(1) 乘法结合律 若 $\boldsymbol{A}\in\mathbb{C}^{m\times n},\boldsymbol{B}\in\mathbb{C}^{n\times p},\boldsymbol{C}\in\mathbb{C}^{p\times q}$，则 $\boldsymbol{A}(\boldsymbol{BC})=(\boldsymbol{AB})\boldsymbol{C}$。

(2) 乘法左分配律 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是两个 $m\times n$ 矩阵，且 $\boldsymbol{C}$ 是一个 $n\times p$ 矩阵，则 $(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{AC}+\boldsymbol{BC}$。

(3) 乘法右分配律 若 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵，并且 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 是两个 $n\times p$ 矩阵，则 $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC}$。

(4) 若 α 是一个标量，并且 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是两个 $m\times n$ 矩阵，则 $\alpha(\boldsymbol{A}+\boldsymbol{B})=\alpha\boldsymbol{A}+\alpha\boldsymbol{B}$。

注意，矩阵的乘法一般不满足交换律，即 $\boldsymbol{AB}\neq\boldsymbol{BA}$。

3. 矩阵的求逆

矩阵与向量的乘积 $\boldsymbol{Ax}=\boldsymbol{y}$ 可视为向量 $\boldsymbol{x}$ 的线性变换。此时，矩阵 $\boldsymbol{A}$ 称为线性变换矩阵。若 $\boldsymbol{A}$ 为 $n\times n$ 矩阵且向量 $\boldsymbol{y}$ 到 $\boldsymbol{x}$ 的线性逆变换 $\boldsymbol{A}^{-1}$ 存在，则

$$\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{y} \tag{1.1.11}$$

这相当于原线性变换 $\boldsymbol{Ax}=\boldsymbol{y}$ 两边左乘 $\boldsymbol{A}^{-1}$ 得到的结果 $\boldsymbol{A}^{-1}\boldsymbol{Ax}=\boldsymbol{A}^{-1}\boldsymbol{y}$。因此，线性逆变换 $\boldsymbol{A}^{-1}$ 应该满足 $\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$ 之关系。另一方面，$\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{y}$ 也应该是可逆的，即两边左乘 $\boldsymbol{A}$ 后得到的 $\boldsymbol{Ax}=\boldsymbol{AA}^{-1}\boldsymbol{y}$ 应该与原线性变换 $\boldsymbol{Ax}=\boldsymbol{y}$ 一致，故 $\boldsymbol{A}^{-1}$ 还应该满足 $\boldsymbol{AA}^{-1}=\boldsymbol{I}$。

定义 1.1.2 令 $\boldsymbol{A}$ 是一个 $n\times n$ 矩阵。若 $n\times n$ 矩阵 $\boldsymbol{A}^{-1}$ 满足 $\boldsymbol{AA}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$，则称矩阵 $\boldsymbol{A}$ 可逆，并称 $\boldsymbol{A}^{-1}$ 是矩阵 $\boldsymbol{A}$ 的逆矩阵。

下面是共轭、转置、共轭转置和逆矩阵的性质。

(1) 矩阵的共轭、转置和共轭转置满足分配律

$$(\boldsymbol{A}+\boldsymbol{B})^*=\boldsymbol{A}^*+\boldsymbol{B}^*,\quad (\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}},\quad (\boldsymbol{A}+\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}+\boldsymbol{B}^{\mathrm{H}}$$

(2) 矩阵乘积的转置、共轭转置和逆矩阵满足关系式

$$(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}},\quad (\boldsymbol{AB})^{\mathrm{H}}=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}$$
$$(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\quad (\boldsymbol{A},\boldsymbol{B}\ \text{为可逆的正方矩阵})$$

(3) 共轭、转置和共轭转置等符号均可与求逆符号交换，即有

$$\boldsymbol{A}^{-*}=(\boldsymbol{A}^*)^{-1}=(\boldsymbol{A}^{-1})^*,\quad \boldsymbol{A}^{-\mathrm{T}}=(\boldsymbol{A}^{\mathrm{T}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}},\quad \boldsymbol{A}^{-\mathrm{H}}=(\boldsymbol{A}^{\mathrm{H}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{H}}$$

1.2 矩阵的初等变换

涉及矩阵行与行 (或者列与列) 之间的简单运算称为初等行变换或者初等列变换，二者统称初等变换。在应用中，矩阵的初等运算往往可以解决一些重要问题。例如，只使用初等行运算，就可以解决矩阵方程求解、矩阵求逆和矩阵基本空间的基向量构造等复杂问题。又如，在矩阵的相似变换中，也需要使用矩阵的初等变换。

1.2.1 初等行变换与阶梯型矩阵

定义 1.2.1 令矩阵 $\boldsymbol{A}\in\mathbb{C}^{m\times n}$ 的 m 个行向量分别为 $\boldsymbol{r}_1,\boldsymbol{r}_2,\cdots,\boldsymbol{r}_m$。下列运算称为矩阵 $\boldsymbol{A}$ 的初等行运算 (elementary row operation) 或初等行变换 (elementary row transformation):

(1) 互换矩阵的任意两行，如 $\boldsymbol{r}_p\leftrightarrow\boldsymbol{r}_q$，称为 I 型初等行变换。

(2) 一行元素同乘一个非零常数 α，如 $\alpha\boldsymbol{r}_p\to\boldsymbol{r}_p$，称为II型初等行变换。

(3) 将第 p 行元素同乘一个非零常数 β 后，加给第 q 行，即 $\beta\boldsymbol{r}_p+\boldsymbol{r}_q\to\boldsymbol{r}_q$，称为 III 型初等行变换。

假设矩阵 $\boldsymbol{A}_{m\times n}$ 经过一系列初等行运算，变换成为矩阵 $\boldsymbol{B}_{m\times n}$，则称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为行等价矩阵 (row equivalent matrix)。

一个非零行最左边的非零元素称为该行的首项元素 (leading entry)。如果首项元素等于 1，便称为首一元素 (leading 1 entry)。

从线性方程组的求解以及基本空间的基向量构造等实际应用出发，往往希望将一个矩阵经过初等行运算之后，变换为阶梯型矩阵。

定义 1.2.2 一个 $m\times n$ 矩阵称为阶梯型 (echelon form) 矩阵，若下列条件都满足:

(1) 全部由零组成的所有行都位于矩阵的底部。

(2) 每一个非零行的首项元素总是出现在上一个非零行的首项元素的右边。

(3) 首项元素下面的同列元素全部为零。

例如，下面是阶梯型矩阵的两个例子:

$$\boldsymbol{A}=\begin{bmatrix}2&*&*\\0&5&*\\0&0&3\\0&0&0\end{bmatrix},\qquad \boldsymbol{A}=\begin{bmatrix}1&*&*\\0&3&*\\0&0&0\\0&0&0\end{bmatrix}$$

式中，$*$ 表示该元素可以为任意值。

定义 1.2.3[59] 阶梯型矩阵 $\boldsymbol{A}$ 称为行简约阶梯型 (row reduced echelon form, RREF)，若 $\boldsymbol{A}$ 的每一非零行的首项元素等于 1 (即为首一元素)，并且每一个首一元素也是它所在列唯一的非零元素。

行简约阶梯型也称行阶梯标准型或 Hermite 标准型。

给定一个 $m\times n$ 矩阵 $\boldsymbol{B}$，下面的算法通过初等行变换将 $\boldsymbol{B}$ 化成行简约阶梯型矩阵。

算法 1.2.1 将 $m\times n$ 矩阵化成行简约阶梯型 [59]

步骤 1 将含有一个非零元素的列设定为最左边的第 1 列。

步骤 2 若需要，将第 1 行与其他行互换，使第 1 个非零列在第 1 行有一个非零元素。

步骤 3 如果第 1 行的首项元素为 a，则将该行的所有元素乘以 $1/a$，以使该行的首项元素等于 1，成为首一元素。

步骤 4 通过初等行变换，将其他行位于第 1 行首一元素下面的全部元素变成 0。

步骤 5 对第 $i=2,3,\cdots,m$ 行依次重复以上步骤，以使每一行的首一元素出现在上一行的首一元素的右边，并使与第 i 行首一元素同列的其他各行元素都变为 0。